



(12) 发明专利申请

(10) 申请公布号 CN 114819112 A

(43) 申请公布日 2022.07.29

(21) 申请号 202210756949.9

G06Q 10/04 (2012.01)

(22) 申请日 2022.06.30

G06Q 50/02 (2012.01)

(71) 申请人 南京北斗创新应用科技研究院有限公司

G06Q 50/16 (2012.01)

地址 211599 江苏省南京市六合区龙池街道虎跃东路8号

G06Q 50/26 (2012.01)

G06V 10/82 (2022.01)

G06V 10/771 (2022.01)

G06V 10/774 (2022.01)

(72) 发明人 杜志强 李柏延 王超 李沐春
干伟

G06V 10/80 (2022.01)

(74) 专利代理机构 南京佰腾智信知识产权代理
事务所(普通合伙) 32509

专利代理师 胡杰

(51) Int.Cl.

G06N 3/04 (2006.01)

G06N 3/08 (2006.01)

G06N 7/00 (2006.01)

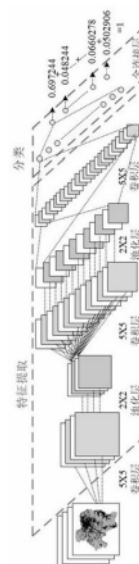
权利要求书3页 说明书9页 附图5页

(54) 发明名称

一种顾及空间分区的土地利用适宜性概率生成方法

(57) 摘要

本发明公开了一种顾及空间分区的土地利用适宜性概率生成方法,顾及土地利用变化驱动机制的空间差异性,定量描述土地利用变化建模中各驱动因素的贡献度,有助于了解各驱动因素对于土地利用变化的影响;还围绕空间分区、驱动因素数据集的构建、各驱动因素贡献度的计算以及适宜性概率的训练,解决了土地利用变化模拟中由于时空异质性考虑不充分导致模拟精度降低的问题;基于卷积神经网络,构建了自然地理和社会经济驱动因素与土地利用变化的复杂非线性关联,进而获得各土地利用的适宜性概率,有助于加深对研究区域土地利用变化时空异质性的理解,提高对邻域空间内多土地利用驱动因素的感知能力,也为长时序的土地利用变化时空模拟和预测做准备。



1. 一种顾及空间分区的土地利用适宜性概率生成方法,其特征在于:包括以下步骤:

S1、土地利用数据预处理:将土地利用数据重新分类,耕地的编码设置为1;林地的编码设置为2;灌木和草地的编码设置为3;水体的编码设置为4;建设用地的编码设置为5,建设用地包括不透水面;未利用地的编码设置为6,未利用地包括雪地、冰面、湿地以及裸地;

S2、构建多个土地利用面积信息的100m×100m分辨率土地利用数据;

S3、驱动因素数据预处理:选取12个驱动因素构建驱动因素数据集,驱动因素数据集包括自然地理驱动因素和社会经济驱动因素,自然地理驱动因素设置为3个,分别为高程、坡度以及距离水体;社会经济驱动因素设置为9个,分别为距离政府的距离、人口、GDP、高速公路、主干道、一级道路、二级道路、三级道路以及铁路;根据高速公路、主干道、一级道路、二级道路、三级道路以及铁路,构建土地利用变化社会经济驱动因素的交通可达性数据集;选取距离政府构建区域兴趣点的欧式距离数据;对驱动因素数据集进行重采样,并对结果进行归一化处理,其数学形式如下所示:

$$d' = (d - d_{\min}) / (d_{\max} - d_{\min})$$

式中, d' 表示归一化后的交通可达性, d 表示当前元胞距离上述驱动因素的欧式距离, d_{\min} 和 d_{\max} 分别表示当前元胞距离上述驱动因素的欧式距离的最小值和最大值;

S4、基于土地利用变化视角和景观格局变化视角的空间划分:根据研究区的关键时间节点,生成土地利用变化数据集,从土地利用变化和景观格局变化两个视角,对研究区域进行空间划分;

S5、选取驱动因素;

S6、基于CNN网络构建土地利用驱动机制,配合混合元胞自动机,对用地适宜性概率进行训练;

S7、计算适宜性概率,其数学形式如下所示:

$$P(x_i) = \frac{f(x_i)}{\sum_{i=1}^N f(x_i)}$$

式中, X_i 表示元胞状态,即用地类型, $P(x_i)$ 表示元胞转换为 X_i 的概率, $f(x_i)$ 表示元胞状态 X_i 的适宜性概率, $\sum_{i=1}^N f(x_i)$ 表示累积概率;

S8、按照空间分区计算各用地基于混淆矩阵的卷积网络训练结果,并输出各土地利用类型一致的多波段适宜性概率训练结果图。

2. 根据权利要求1所述的一种顾及空间分区的土地利用适宜性概率生成方法,其特征在于:所述步骤S2中,多个土地利用面积信息的100m×100m分辨率土地利用数据的构建方法包括以下步骤:

S2.1、选取100m×100m的分辨率作为土地利用变化模拟的基础;

S2.2、基于30m×30m原始分辨率的土地利用数据,构建100m×100m分辨率的渔网;

S2.3、计算100m×100m分辨率的渔网空间中各用地面积占渔网面积的比例;

S2.4、得到一个包含多个土地利用面积信息的100m×100m分辨率土地利用数据。

3. 根据权利要求1所述的一种顾及空间分区的土地利用适宜性概率生成方法,其特征在于:所述步骤S4中,从土地利用变化和景观格局变化两个视角,对研究区域进行空间划分

的方法包括以下步骤:

S4.1、从土地利用变化视角对研究区域进行空间划分:根据土地利用程度和土地利用程度变化率进行分区,分为第一梯队、第二梯队以及第三梯队;

S4.2、从景观格局变化视角对研究区域进行空间划分:选取最大斑块指数(LPI)和边缘密度(ED)作为评估指标,将空间划分为主城区和远城区。

4.根据权利要求1所述的一种顾及空间分区的土地利用适宜性概率生成方法,其特征在于:所述步骤S5中,驱动因素的选取因素包括以下步骤:

S5.1、判断驱动因素是否同时具有易获取性、全面性、时空信息一致性、空间差异性以及可量化,当同时具有时,对该驱动因素进行选取;

S5.2、构建驱动因素数据集;

S5.3、评估驱动因素对土地利用变化的重要性。

5.根据权利要求1所述的一种顾及空间分区的土地利用适宜性概率生成方法,其特征在于:所述步骤S6中,用地适宜性概率的训练方法包括以下步骤:

S6.1、卷积层:卷积网络模型输入数据首先经过卷积层,通过卷积核学习输入变量的特征表示,卷积网络模型通过多个卷积层,逐步解析图像不同维度的抽象特征,获取更深层的特征表示,卷积层的数学形式如下所示:

$$x_j^l = f\left(\sum_{i=1}^n k_j^l \times x_i^{l-1} + b_j^l\right)$$

式中, x_i^{l-1} 表示第1层卷积层的第i维输入变量, k_j^l 和 b_j^l 分别表示第1层卷积层中第j个卷积核组的权重向量和偏置项, x_j^l 表示第1层卷积层中第j个卷积核对应的输出变量, n表示第1层卷积层输入变量的总维数, f表示非线性激活函数;

S6.2、池化层:对卷积提取后的数据特征进行选择,根据信息量的不同筛选出符合要求的特征,对不重要特征进行剔除,同时降低数据维度并避免过拟合,对池化区域内的数据特征计算函数值,通过均值池化将区域中所有元素的算术平均值作为输出,通过最大值池化提取局域特征平面的最显著的特征值作为输出,以用于浅层特征的提取,池化层的数学形式如下所示:

$$x_j^l = \beta_i^l \times \text{down}(x_i^{l-1}) + b_j^l$$

式中, x_i^{l-1} 表示第1池化层的第i维输入变量, β_i^l 和 b_j^l 分别表示第1池化层中第j维变量的乘法偏置项和加法偏置项, x_j^l 表示第1池化层中第j维变量的输出变量, down表示池化函数;

S6.3、全连接层:卷积网络通过逐层的特征提取和过滤,将原始输入数据的抽象表达输入到全连接层中进行特征融合,并形成与目标空间对应的一维特征向量,全连接层的数学形式如下所示:

$$x_j^l = f\left(\sum_{i=1}^n k_{j,i}^l \times x_i^{l-1} + b_{j,i}^l\right)$$

式中, x_i^{l-1} 表示第1层全连接层的第i维神经元, $k_{j,i}^l$ 表示第1层卷积层中神经元i到神经元j的权重, $b_{j,i}^l$ 表示第1层卷积层中神经元i到神经元j的偏置项, x_j^l 表示第1层卷积层中第j个卷积核对应的输出变量, n表示全连接层输入神经元的总数, f表示非线性激活函数;

S6.4、激活层:将混合元胞的用地覆盖比作为卷积网络训练目标,全连接层采用的激活函数设置为Sigmoid激活函数,其数学形式如下所示:

$$\text{sigmoid}(x_j^l) = \frac{1}{1 + e^{-x_j^l}}$$

式中, x_j^l 表示激活层的输入。

6. 根据权利要求1所述的一种顾及空间分区的土地利用适宜性概率生成方法,其特征在于:所述步骤S3中,通过软件ArcGIS10.1对驱动因素数据集进行重采样。

7. 根据权利要求6所述的一种顾及空间分区的土地利用适宜性概率生成方法,其特征在于:所述步骤S3中,驱动因素数据集设置为原始30m30m分辨率的驱动因素数据集。

8. 根据权利要求2所述的一种顾及空间分区的土地利用适宜性概率生成方法,其特征在于:所述步骤S2.3中,各用地面积占渔网面积的比例的计算工具设置为ArcGIS10.1软件的分区统计工具。

一种顾及空间分区的土地利用适宜性概率生成方法

技术领域

[0001] 本发明涉及土地利用变化模拟技术领域,特别是涉及一种顾及空间分区的土地利用适宜性概率生成方法。

背景技术

[0002] 土地利用变化适宜性概率研究旨在揭示土地利用/覆盖变化(Land-Use/Land-Cover Change,LUCC)的内在动力机制一极及作用机理,进而预测土地利用动态演变趋势,因此受到国内外学者的广泛关注,各土地利用类型的选择概率和其适宜性概率成比例,适宜性概率越大,用地竞争力越强,用地被选作元胞下一时刻的转换状态的概率也就越大。

[0003] 土地利用变化驱动因素主要分为两大类,即自然环境因素和社会经济因素,以坡度、地形、气候以及降水等为代表的自然环境因素是长时序土地利用变化研究不可或缺的一部分;影响LUCC的社会经济因素则主要包括人口增长、经济发展、交通可达性以及土地利用管理政策。

[0004] 现有研究表明,随着城镇化水平的持续提升,以交通可达性为代表的社会经济因素,对土地利用变化施加着远超自然环境因素的影响力;具有不同时间和空间尺度的驱动因素间也存在着内部关联,从而导致了土地利用变化的时空异质性,为LUCC驱动机制分析过程和构建过程增添了复杂性和不确定性;在长时间尺度上,自然环境因素对土地利用变化的影响较为明显;而在中小尺度和短时间尺度上,由于人类活动对土地利用变化的影响日益凸显,社会经济因素逐渐占据主导地位。

[0005] 近年来,不少学者开始关注土地利用变化的时空异质性现象,尝试采用空间分区的方式对整个研究区域进行划分,进而探究不同子区域的土地利用变化模式,随着地理学计量理念的突破及智能算法的引入,土地利用变化驱动机制的研究突破了传统的实践经验为主导的定性方法,逐步走向模拟过程高效率 and 模拟结果高精度的定量化。

[0006] 智能算法的引入为土地利用变化驱动机制研究提供了更高效的手段,近年来,基于智能算法的驱动机制研究不断增多,因此,顾及研究区域驱动因素的时空依赖特征和土地利用时空演化差异,构建土地利用变化研究合理的驱动因素数据集,挖掘用地变化与自然地理和社会经济维度驱动因素的关联,是土地利用变化驱动机制研究的重中之重和关键所在。

发明内容

[0007] 为了解决以上技术问题,本发明提供一种顾及空间分区的土地利用适宜性概率生成方法,包括以下步骤:

S1、土地利用数据预处理:将土地利用数据重新分类,耕地的编码设置为1;林地的编码设置为2;灌木和草地的编码设置为3;水体的编码设置为4;建设用地的编码设置为5,建设用地包括不透水面;未利用地的编码设置为6,未利用地包括雪地、冰面、湿地以及裸地;

S2、构建多个土地利用面积信息的100m×100m分辨率土地利用数据；

S3、驱动因素数据预处理：选取12个驱动因素构建驱动因素数据集，驱动因素数据集包括自然地理驱动因素和社会经济驱动因素，自然地理驱动因素设置为3个，分别为高程、坡度以及距离水体；社会经济驱动因素设置为9个，分别为距离政府的距离、人口、GDP、高速公路、主干道、一级道路、二级道路、三级道路以及铁路；根据高速公路、主干道、一级道路、二级道路、三级道路以及铁路，构建土地利用变化社会经济驱动因素的交通可达性数据集；选取距离政府构建区域兴趣点的欧式距离数据；对驱动因素数据集进行重采样，并对结果进行归一化处理，其数学形式如下所示：

$$d' = (d - d_{\min}) / (d_{\max} - d_{\min})$$

式中， d' 表示归一化后的交通可达性， d 表示当前元胞距离上述驱动因素的欧式距离， d_{\min} 和 d_{\max} 分别表示当前元胞距离上述驱动因素的欧式距离的最小值和最大值；

S4、基于土地利用变化视角和景观格局变化视角的空间划分：根据研究区的关键时间节点，生成土地利用变化数据集，从土地利用变化和景观格局变化两个视角，对研究区域进行空间划分；

S5、选取驱动因素；

S6、基于CNN网络构建土地利用驱动机制，配合混合元胞自动机，对用地适宜性概率进行训练；

S7、计算适宜性概率，其数学形式如下所示：

$$P(x_i) = \frac{f(x_i)}{\sum_{i=1}^N f(x_i)}$$

式中， X_i 表示元胞状态，即用地类型， $P(x_i)$ 表示元胞转换为 X_i 的概率， $f(x_i)$ 表示元胞状态 X_i 的适宜性概率， $\sum_{i=1}^N f(x_i)$ 表示累积概率；

S8、按照空间分区计算各用地基于混淆矩阵的卷积网络训练结果，并输出各土地利用类型一致的多波段适宜性概率训练结果图。

[0008] 本发明进一步限定的技术方案是：

进一步的，步骤S2中，多个土地利用面积信息的100m×100m分辨率土地利用数据的构建方法包括以下步骤：

S2.1、选取100m×100m的分辨率作为土地利用变化模拟的基础；

S2.2、基于30m×30m原始分辨率的土地利用数据，构建100m×100m分辨率的渔网；

S2.3、计算100m×100m分辨率的渔网空间中各用地面积占渔网面积的比例；

S2.4、得到一个包含多个土地利用面积信息的100m×100m分辨率土地利用数据。

[0009] 前所述的一种顾及空间分区的土地利用适宜性概率生成方法，步骤S4中，从土地利用变化和景观格局变化两个视角，对研究区域进行空间划分的方法包括以下步骤：

S4.1、从土地利用变化视角对研究区域进行空间划分：根据土地利用程度和土地利用程度变化率进行分区，分为第一梯队、第二梯队以及第三梯队；

S4.2、从景观格局变化视角对研究区域进行空间划分：选取最大斑块指数(LPI)和边缘密度(ED)作为评估指标，将空间划分为主城区和远城区。

[0010] 前所述的一种顾及空间分区的土地利用适宜性概率生成方法，步骤S5中，驱动因

素的选取因素包括以下步骤：

S5.1、判断驱动因素是否同时具有易获取性、全面性、时空信息一致性、空间差异性以及可量化，当同时具有时，对该驱动因素进行选取；

S5.2、构建驱动因素数据集；

S5.3、评估驱动因素对土地利用变化的重要性。

[0011] 前所述的一种顾及空间分区的土地利用适宜性概率生成方法，步骤S6中，用地适宜性概率的训练方法包括以下步骤：

S6.1、卷积层：卷积网络模型输入数据首先经过卷积层，通过卷积核学习输入变量的特征表示，卷积网络模型通过多个卷积层，逐步解析图像不同维度的抽象特征，获取更深层的特征表示，卷积层的数学形式如下所示：

$$x_j^l = f \left(\sum_{i=1}^n k_j^l \times x_i^{l-1} + b_j^l \right)$$

式中， x_i^{l-1} 表示第1层卷积层的第i维输入变量， k_j^l 和 b_j^l 分别表示第1层卷积层中第j个卷积核组的权重向量和偏置项， x_j^l 表示第1层卷积层中第j个卷积核对应的输出变量，n表示第1层卷积层输入变量的总维数，f表示非线性激活函数；

S6.2、池化层：对卷积提取后的数据特征进行选择，根据信息量的不同筛选出符合要求的特征，对不重要特征进行剔除，同时降低数据维度并避免过拟合，对池化区域内的数据特征计算函数值，通过均值池化将区域中所有元素的算术平均值作为输出，通过最大值池化提取局域特征平面的最显著的特征值作为输出，以用于浅层特征的提取，池化层的数学形式如下所示：

$$x_j^l = \beta_j^l \times \text{down} (x_i^{l-1}) + b_j^l$$

式中， x_i^{l-1} 表示第1池化层的第i维输入变量， β_j^l 和 b_j^l 分别表示第1池化层中第j维变量的乘法偏置项和加法偏置项， x_j^l 表示第1池化层中第j维变量的输出变量，down表示池化函数；

S6.3、全连接层：卷积网络通过逐层的特征提取和过滤，将原始输入数据的抽象表达输入到全连接层中进行特征融合，并形成与目标空间对应的一维特征向量，全连接层的数学形式如下所示：

$$x_j^l = f \left(\sum_{i=1}^n k_{j,i}^l \times x_i^{l-1} + b_{j,i}^l \right)$$

式中， x_i^{l-1} 表示第1层全连接层的第i维神经元， $k_{j,i}^l$ 表示第1层卷积层中神经元i到神经元j的权重， $b_{j,i}^l$ 表示第1层卷积层中神经元i到神经元j的偏置项， x_j^l 表示第1层卷积层中第j个卷积核对应的输出变量，n表示全连接层输入神经元的总数，f表示非线性激活函数；

S6.4、激活层：将混合元胞的用地覆盖比作为卷积网络训练目标，全连接层采用的

激活函数设置为Sigmoid激活函数,其数学形式如下所示:

$$\text{sigmoid}(x_j^l) = \frac{1}{1 + e^{-x_j^l}}$$

式中, x_j^l 表示激活层的输入。

[0012] 前所述的一种顾及空间分区的土地利用适宜性概率生成方法,步骤S3中,通过软件ArcGIS10.1对驱动因素数据集进行重采样。

[0013] 前所述的一种顾及空间分区的土地利用适宜性概率生成方法,步骤S3中,驱动因素数据集设置为原始30m30m分辨率的驱动因素数据集。

[0014] 前所述的一种顾及空间分区的土地利用适宜性概率生成方法,步骤S2.3中,各用地面积占渔网面积的比例的计算工具设置为ArcGIS10.1软件的分区统计工具。

[0015] 本发明的有益效果是:

(1) 本发明中,能够顾及的土地利用变化驱动机制的空间差异性,定量描述土地利用变化建模中各驱动因素的贡献,有助于了解各驱动因素如何影响土地利用变化的;

(2) 本发明围绕空间分区、驱动因素数据集的构建、各驱动因素贡献度的计算以及适宜性概率的训练,解决了土地利用变化模拟中由于时空异质性考虑不充分导致模拟精度降低的问题;

(3) 本发明基于卷积神经网络(Convolutional Neural Network,CNN),构建了自然地理和社会经济驱动因素与土地利用变化的复杂非线性关联,进而获得了各土地利用的适宜性概率,有助于加深与研究区域土地利用变化时空异质性的理解,提高了对邻域空间内自然地理和社会经济等多土地利用驱动因素的感知能力,也为长时序的土地利用变化时空模拟和预测做准备。

附图说明

[0016] 图1为本发明中试验区域土地利用时空变化图;

图2为本发明中试验区域景观指数时空变化图;

图3为本发明中试验区域各驱动因素对耕地和林地的贡献度(SHAP值);

图4为本发明中试验区域卷积网络结构;

图5为本发明中试验区域主城区卷积网络适宜性训练混淆矩阵;

图6为本发明中试验区域远城区卷积网络适宜性训练混淆矩阵;

图7为本发明中试验各用地适宜性概率训练图。

具体实施方式

[0017] 本实施例提供一种顾及空间分区的土地利用适宜性概率生成方法,为了便于本领域普通技术人员理解和实施本发明,以近十年来土地利用变化幅度和规模较大的武汉市为例,对本发明作进一步的详细描述。

[0018] 一种顾及空间分区的土地利用适宜性概率生成方法,包括以下步骤:

S1、土地利用数据预处理:按照《地理国情普查内容与指标GDPJ01-2013》的相关规范,将土地利用数据重新分类,耕地的编码设置为1;林地的编码设置为2;灌木和草地的编码设置为3;水体的编码设置为4;建设用地的编码设置为5,建设用地包括不透水面;未利用

地的编码设置为6,未利用地包括雪地、冰面、湿地以及裸地。

[0019] S2、构建多个土地利用面积信息的100m×100m分辨率土地利用数据,包括以下步骤:

S2.1、综合考虑研究目标以及及模型运行效率、选取100m×100m的分辨率作为土地利用变化模拟的基础;

S2.2、基于30m×30m原始分辨率的土地利用数据,构建100m×100m分辨率的渔网;

S2.3、通过ArcGIS10.1软件的分区统计工具,计算100m×100m分辨率的渔网空间中各用地面积占渔网面积的比例,从而将高分辨率土地利用数据的空间分布特征聚合到低分辨率网格中,有效的防止了高分辨率数据信息的全部丢失;

S2.4、得到一个包含多个土地利用面积信息的100m×100m分辨率土地利用数据。

[0020] S3、驱动因素数据预处理:选取12个驱动因素构建驱动因素数据集,驱动因素数据集包括自然地理驱动因素和社会经济驱动因素,自然地理驱动因素设置为3个,分别为高程、坡度以及距离水体;社会经济驱动因素设置为9个,分别为距离政府的距离、人口、GDP、高速公路、主干道、一级道路、二级道路、三级道路以及铁路;根据高速公路、主干道、一级道路、二级道路、三级道路以及铁路,构建土地利用变化社会经济驱动因素的交通可达性数据集;选取距离政府构建区域兴趣点的欧式距离数据;通过软件ArcGIS10.1对原始30m30m分辨率的驱动因素数据集进行重采样,并对结果进行归一化处理,其数学形式如下所示:

$$d' = (d - d_{\min}) / (d_{\max} - d_{\min})$$

式中, d' 表示归一化后的交通可达性, d 表示当前元胞距离上述驱动因素的欧式距离, d_{\min} 和 d_{\max} 分别表示当前元胞距离上述驱动因素的欧式距离的最小值和最大值。

[0021] S4、根据研究区域土地利用时空演变特征,选取具有代表性的4个年份(2005年、2010年、2015年以及2020年)作为武汉市土地利用变化研究的关键时间节点,以6年为间隔生成2005年-2010年、2010年-2015年以及2015年-2020年的研究数据集,从土地利用变化和景观格局变化两个视角对武汉市土地利用时空演变特征进行分析,进而对研究区域进行空间划分,景观格局是对研究区域景观空间结构特征的刻画和描述,是形状不同、大小各异的景观要素的组合与分布,也是自然、社会以及人文等多类要素共同作用的结果,空间划分的方法包括以下步骤:

S4.1、从土地利用变化视角对研究区域进行空间划分:根据土地利用程度和土地利用程度变化率进行分区,如图1所示,分为第一梯队、第二梯队以及第三梯队:

第一梯队包括江汉、硚口、青山以及武昌,土地利用程度较高,均大于300,同时,在2005-2020年间,除硚口外三处区域的土地利用程度变化起伏不大,江汉、青山以及武昌的平均土地利用程度变化率分别只有1.35%、1.37%以及0.66%;

第二梯队包括江岸、汉阳以及洪山,虽然在研究的初始时期(2005年)土地利用程度相对于第一梯队偏低,但在研究时期(2005-2020年)呈现快速增长的趋势,平均土地利用程度变化率分别为5.09%、9.01%以及8.15%,其中,江岸土地利用程度在2020年已突破300,跨入较高土地利用程度的范畴;

第三梯队包括东西湖、蔡甸、江夏以及汉南,四处区域的土地利用程度变化趋势相近,这些区域主要为武汉市的外围行政区,相比主城区城镇化发展缓慢,这些区域在研究的

初始时期(2005年)土地利用程度均低于140,且在研究时期(2005-2020年)变化幅度不大,2005-2020年间蔡甸、江夏以及汉南的平均土地利用程度变化率分别为5.45%、4.80%以及4.72%,虽然东西湖呈现相对较高的平均土地利用程度增长率,为7.49%,但是由于其研究的初始时期(2005年)土地利用程度较低,所以仍划分为第三梯队;

S4.2、从景观格局变化视角对研究区域进行空间划分:选取最大斑块指数(LPI)和边缘密度(ED)作为评估指标,计算得到武汉市2005、2010年、2015年以及2020年的景观格局变化程度,如图2所示,武汉市耕地、林地、水体以及建设用地的边缘密度变化最为明显,且主城区和远城区呈现明显的时空分异特征,具体来说,主城区的耕地和建设用地的最大斑块指数随时间变化最为明显,远城区的优势用地仍为耕地,其最大斑块指数随着时间变化呈现先减小后增大的趋势,但仍远大于其他类型用地。

[0022] S5、选取驱动因素,包括以下步骤:

S5.1、判断驱动因素是否同时具有易获取性、全面性、时空信息一致性、空间差异性以及可量化,当同时具有时,对该驱动因素进行选取;

易获取性:在综合考虑可能对土地利用变化产生影响的驱动力因素后,首先应该利用驱动因素的易获取性对土地利用驱动力进行筛选,驱动因素必须有可靠的获取途径,并且可以应用到研究中,对获取的驱动因素应该检查其准确度以及可用性;

全面性:影响土地利用变化的驱动因素众多,但各个驱动因素对土地利用变化的影响程度各异,各驱动因素从不同的方面对土地利用变化产生影响,在驱动因素选取过程中应综合选取社会、自然以及经济等各方面典型的影响因素,以全面地反映研究区域的土地利用变化;

时空信息一致性:各驱动因素必须在同一时空尺度下,研究才具有科学性,若各驱动因素处于不同的时间空间尺度下,则失去了模拟研究的必要性,要保持各驱动因素具有相同的时间信息和空间坐标信息,驱动因素的时空信息一致由模型特性和研究的科学性决定;对于选取的驱动因素应当保持空间尺度信息的一致,否则分析难以继续进行,所以只有驱动因素在同一时间尺度下,研究才具有可行性;

空间差异性:栅格上驱动因素所决定地类分布概率来决定栅格的最终土地利用类型,若各驱动因素不具备空间差异性,在每个栅格的值相同,则会得出每个栅格结果一致,导致研究结果没有任何意义,因此驱动因素在空间分布上必须具有差异性;

可量化:模拟过程利用卷积网络对驱动因素与地类变化之间的关系进行测算,驱动因素对土地利用变化的影响需要输入至卷积网络中,进行定量分析才能得出模拟结果,因此驱动因素必须可量化;在驱动因素选取过程中,法律等因素虽然对土地利用变化影响较大,但难以定量表示,所以一般不将此类因素纳入驱动力系统。

[0023] S5.2、构建驱动因素数据集:土地利用变化受到自然地理和社会经济的多种驱动因素的影响,已有研究表明,坡度、高程等被认为是影响土地利用变化的重要自然地理驱动因素,同时到河流和湖泊的距离也会影响耕地和建设用地的空间分布;从社会经济方面来看,由于人类活动是影响土地利用变化变化的决定性因素,所以人口密度、国内生产总值以及到政府的距离在城镇化进程中发挥着至关重要的作用。

[0024] 此外,交通要素是评估空间相互作用强度与广度的重要维度,而可达性作为交通系统的主要产品,直接决定了各区域的区位优势,因而与区域经济发展水平、社会福祉以及

生态环境密不可分；现有研究为了量化交通要素对土地利用变化的影响，普遍采用交通可达性的方式进行计算，如到公路的距离、到城市中心的距离等。

[0025] S5.3、评估驱动因素对土地利用变化的贡献度：定量描述土地利用变化建模中各驱动因素的贡献度，有助于城市规划者进一步了解这些因素是如何影响土地利用的演变，对于制定土地利用规划政策至关重要；空间分区以步骤S4.2得到的武汉市主城区和远城区分区图为例，基于卷积网络，分别定量描述主城区和远城区内的各用地驱动因素的贡献度。

[0026] SHAP是一种通过计算每个特征对模型的贡献度，对机器学习网络架构进行启发式解释的有效方法，本发明通过计算研究分区（主城区与远城区）各土地利用类型的驱动因素总体SHAP值，评估驱动因素对土地利用变化的贡献度。步骤S4.2中定量描述第一梯队、第二梯队以及第三梯队中各用地驱动因素的贡献度过程与步骤S5.3中空间分区各用地驱动因素的贡献度计算过程类似，此处不再赘述。

[0027] 如图3所示，可以看出武汉市主城区和远城区的土地利用变化驱动机制具有差异性，主城区中耕地空间分布受人口和GDP的影响较大，SHAP值分别为10.27%和9.85%；自然地理驱动因素的影响次之，距离水体的距离、坡度以及高程对耕地的影响较为接近，SHAP值分别为6.75%、6.61%以及6.56%；相比主城区，远城区中自然地理驱动因素对耕地空间分布情况具有更大的贡献度，远城区中高程和坡度的SHAP值分别提高了3.98%和2.70%，水体的贡献度则没有较大变化。

[0028] 此外，林地的SHAP值在主城区和远城区基本呈现相近的变化趋势，以高程为代表的自然地理驱动因素对林地空间分布更具影响力，远超交通可达性、人口以及GDP；以林地为例，主城区和远城区的高程SHAP值分别为2.79%和3.33%，而社会经济驱动因素的SHAP值，如人口、GDP以及距离主干道的距离等，均小于0.5%；总体上，农业和林地的演变主要受自然环境的影响，因此自然地理驱动因素对这两种土地利用变化的贡献度最大。

[0029] S6、基于CNN网络构建土地利用驱动机制，配合混合元胞自动机，对用地适宜性概率进行训练，训练方法包括以下步骤：

S6.1、卷积层：卷积网络模型输入数据首先经过卷积层，通过卷积核学习输入变量的特征表示；卷积层由多个用于提取输入数据不同特征的卷积核组成，是卷积神经网络的核心，卷积网络的特征映射输出，需要通过卷积核进行卷积操作，经过非线性激活函数，得到特征提取输出；卷积网络模型通过多个卷积层，逐步解析图像不同维度的抽象特征，获取更深层的特征表示；此外，卷积层通过权重共享机制减少了运算的参数数量，有效降低了网络计算的时间复杂度和空间复杂度，卷积层的数学形式如下所示：

$$x_j^l = f \left(\sum_{i=1}^n k_j^l \times x_i^{l-1} + b_j^l \right)$$

式中， x_i^{l-1} 表示第1层卷积层的第i维输入变量， k_j^l 和 b_j^l 分别表示第1层卷积层中第j个卷积核组的权重向量和偏置项， x_j^l 表示第1层卷积层中第j个卷积核对应的输出变量，n表示第1层卷积层输入变量的总维数，f表示非线性激活函数；

S6.2、池化层：卷积层操作完成后，参数量级虽然得到降低，但数据量仍然很大，所以继续对卷积提取后的数据特征进行选择，根据信息量的不同筛选出符合要求的特征，对

不重要特征进行剔除,同时降低数据维度并避免过拟合,对池化区域内的数据特征计算函数值,该函数值因参数的不同而有所差异,所以通过均值池化将区域中所有元素的算术平均值作为输出,通过最大值池化提取局域特征平面的最显著的特征值作为输出,以用于浅层特征的提取,池化层的数学形式如下所示:

$$x_j^l = \beta_i^l \times \text{down}(x_i^{l-1}) + b_j^l$$

式中, x_i^{l-1} 表示第1池化层的第i维输入变量, β_i^l 和 b_j^l 分别表示第1池化层中第j维变量的乘法偏置项和加法偏置项, x_j^l 表示第1池化层中第j维变量的输出变量, down表示池化函数;

S6.3、全连接层:全连接层通常位于卷积网络的末尾,将前一层网络的所有输出变量进行连接;卷积网络通过逐层的特征提取和过滤,将原始输入数据的抽象表达输入到全连接层中进行特征融合,并形成与目标空间对应的一维特征向量,全连接层的数学形式如下所示:

$$x_j^l = f\left(\sum_{i=1}^n k_{j,i}^l \times x_i^{l-1} + b_{j,i}^l\right)$$

式中, x_i^{l-1} 表示第1层全连接层的第i维神经元, $k_{j,i}^l$ 表示第1层卷积层中神经元i到神经元j的权重, $b_{j,i}^l$ 表示第1层卷积层中神经元i到神经元j的偏置项, x_j^l 表示第1层卷积层中第j个卷积核对应的输出变量, n表示全连接层输入神经元的总数, f表示非线性激活函数;

S6.4、激活层:卷积网络通常在卷积层和全连接层后加一个非线性的激活函数,以模仿人类神经元的抑制行为,从而提高网络的图像识别能力;常见的激活函数包括Sigmoid激活函数、Tanh激活函数以及Relu激活函数, Sigmoid作为使用范围最广的激活函数,将输出映射到(0,1)之间;本发明兼顾混合元胞,在卷积网络适宜性训练过程中,将混合元胞的用地覆盖比作为卷积网络训练目标,全连接层采用的激活函数设置为Sigmoid激活函数,其数学形式如下所示:

$$\text{sigmoid}(x_j^l) = \frac{1}{1 + e^{-x_j^l}}$$

式中, x_j^l 表示激活层的输入,通过Sigmoid函数能将输入数据压缩到0-1的范围内。

[0030] S7、计算适宜性概率:各用地的选择概率和其适宜性概率成比例,适宜性概率越大,用地竞争力越强,用地被选作元胞下一时刻的转换状态的概率也越大;其数学形式如下所示:

$$P(x_i) = \frac{f(x_i)}{\sum_{i=1}^N f(x_j)}$$

式中, X_i 表示元胞状态,即用地类型, $P(x_i)$ 表示元胞转换为 X_i 的概率, $f(x_i)$ 表示元胞状态 X_i 的适宜性概率, $\sum_{i=1}^N f(x_j)$ 表示累积概率。

[0031] 为获取元胞自动机模型转换规则的土地利用变化适宜性概率,本发明参考

LetNet-5模型构建一个具有七层网络结构的卷积神经网络模型,该模型以土地利用变化驱动因素作为输入,各用地覆盖比为输出,具体网络结构如图4所示:

以研究区域土地利用分类图中每个像素为中心单元,从研究区域左上角开始选取N种驱动因素的中心单元及空间邻域的三维张量;结合LetNet-5模型,本发明选取33个元胞作为卷积网络输入数据的高度和宽度,即 $N \times 35 \times 35$ 的元胞三维张量,N表示驱动因素的个数, $N=12$ 。

[0032] 该模型包括三层卷积层、一层池化层以及两层全连接层,第一层由 $12 \times 5 \times 5$ 的卷积核组成,输出 $12 \times 29 \times 29$ 的特征空间;第二层为 2×2 的池化层,输出 $12 \times 14 \times 14$ 的特征空间;第三层为 $16 \times 5 \times 5$ 的卷积核;第四层为 2×2 的池化层,得到 $16 \times 5 \times 5$ 的特征空间;第五层为 $120 \times 5 \times 5$ 的卷积核;第六层为全连接层,包含120个神经元;第七层为包含84个神经元的连接层;最后采用Sigmoid激活函数,计算各用地的转换适宜性概率。

[0033] S8、按照空间分区计算各用地基于混淆矩阵的卷积网络训练结果,如图5至6所示,分别显示了主城区和远城区各用地基于混淆矩阵的卷积网络训练结果,总体上混合用地模拟结果与真实混合用地一致性较高,普遍大于0.8;由于草地和未利用地的样本较少,卷积网络训练准确度有所下降;此外,基于卷积网络获取了各土地利用类型的适宜性的空间分布,为耦合模型的土地利用动态演变模拟提供数据支持;

输出各土地利用类型一致的多波段适宜性概率训练结果图,适宜性概率高意味着该用地类型在邻域内易发生扩张,以占据更大土地覆盖面积,因此其他用地在土地利用演变过程中容易向该用地状态转换;如图7所示,为试验各用地适宜性概率训练图,图中可以发现耕地空间分布适宜性主要受水体和建设用地的影响,但在总体上,以人口和GDP为主的社会经济驱动因素及以高程和坡度为主的自然地理驱动因素,对各土地利用类型的空间分布起主导作用。

[0034] 除上述实施例外,本发明还可以有其他实施方式。凡采用等同替换或等效变换形成的技术方案,均落在本发明要求的保护范围。

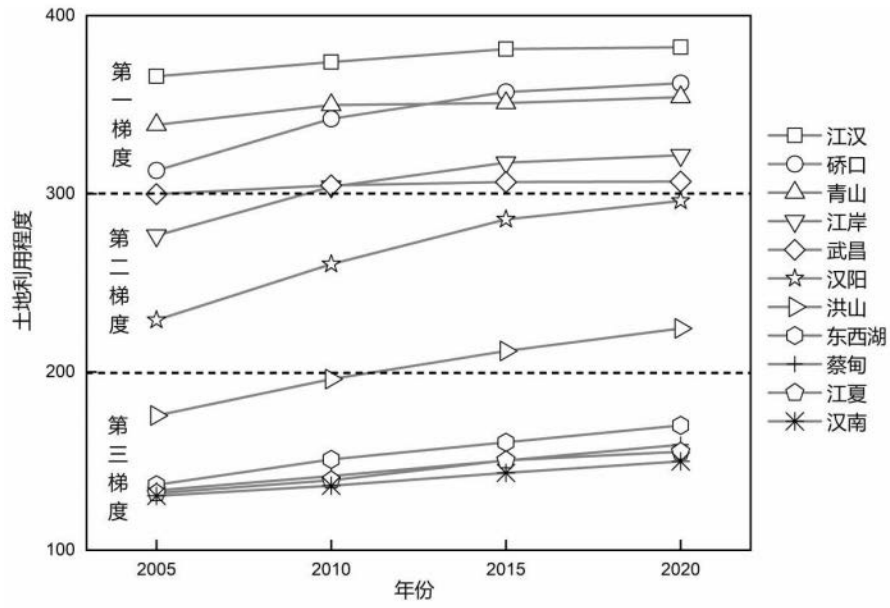


图1

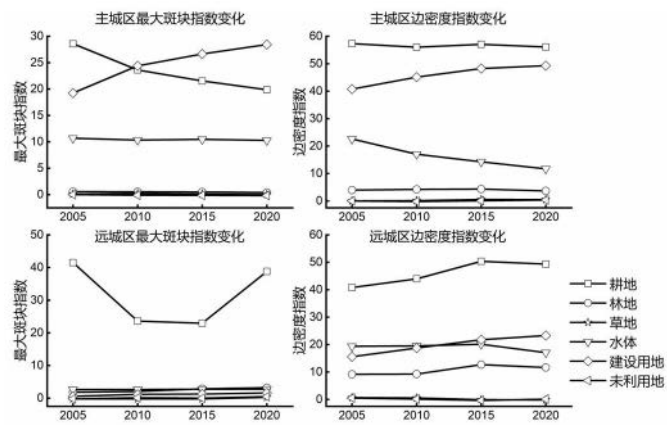


图2

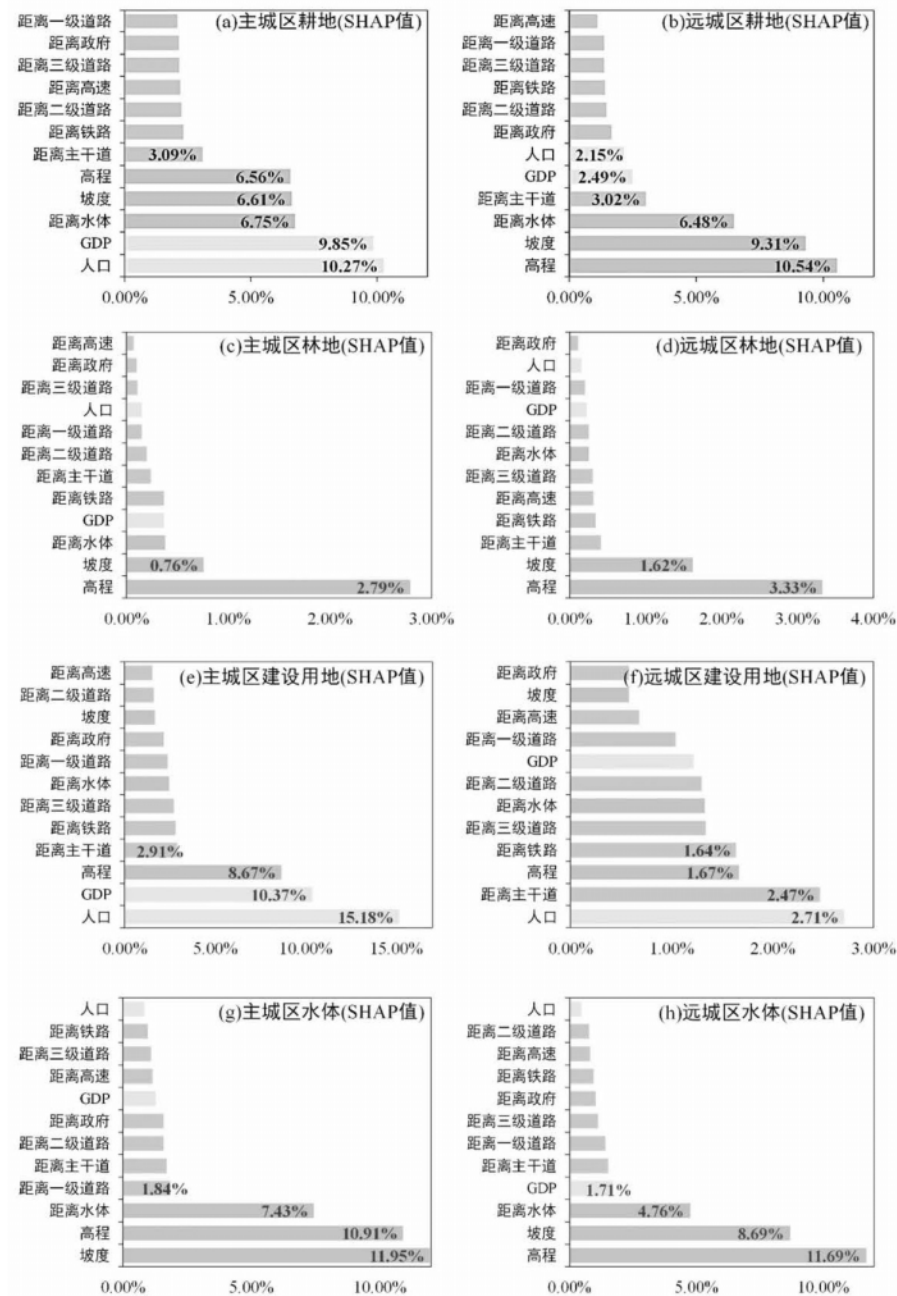


图3

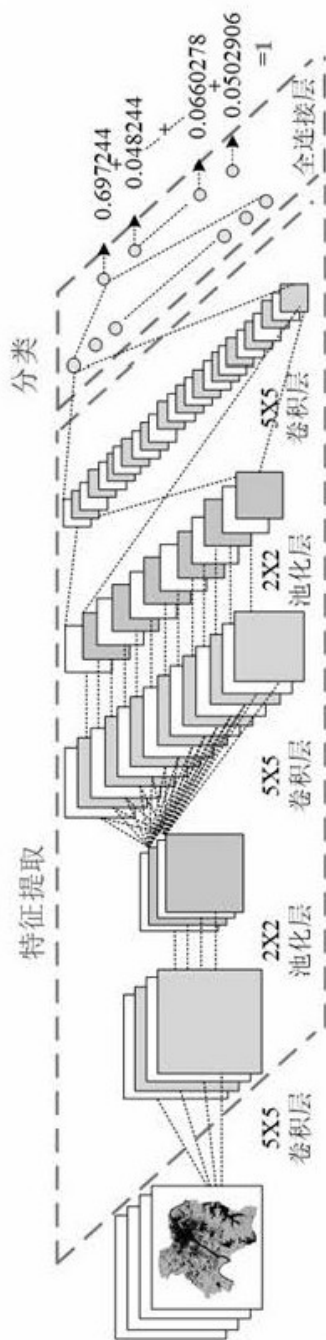


图4

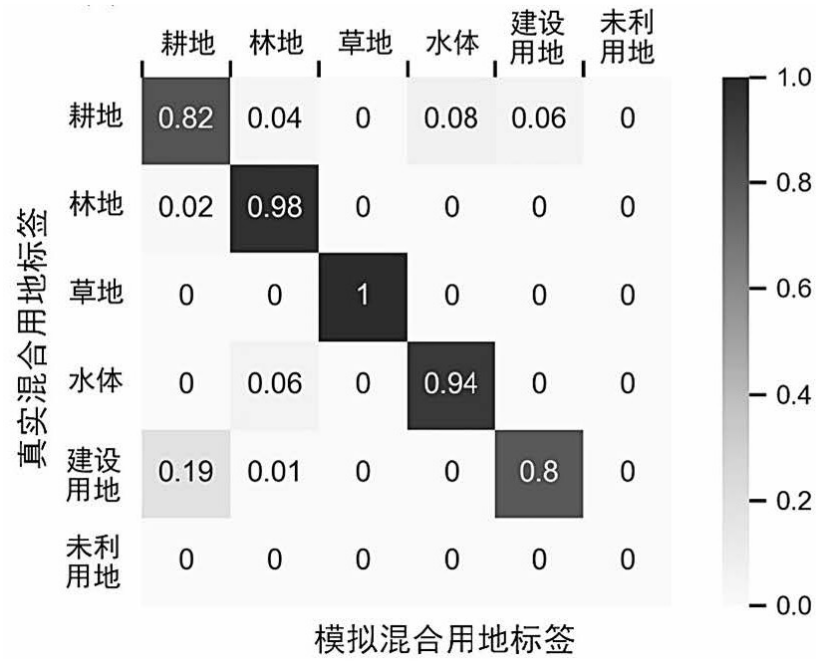


图5

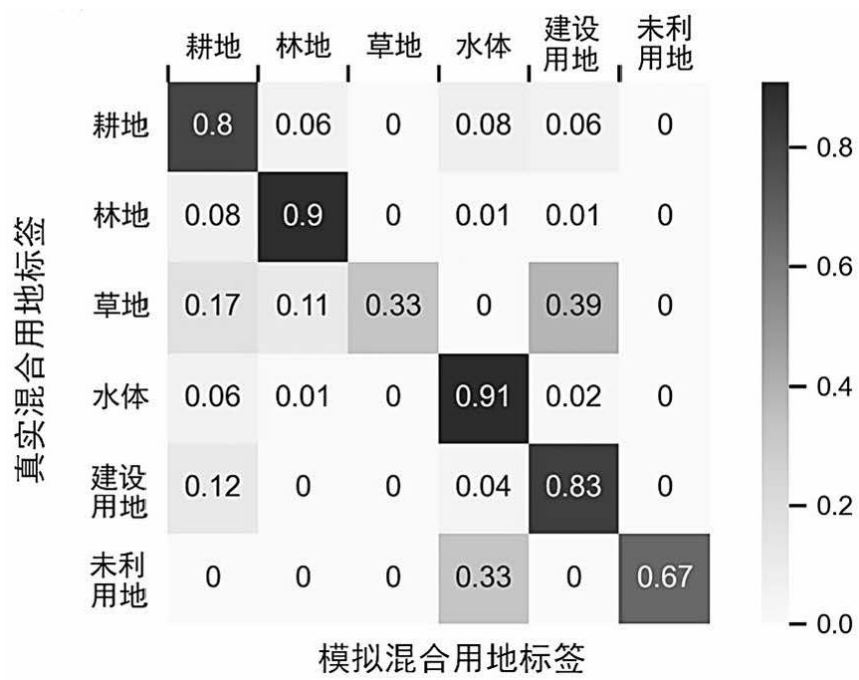


图6

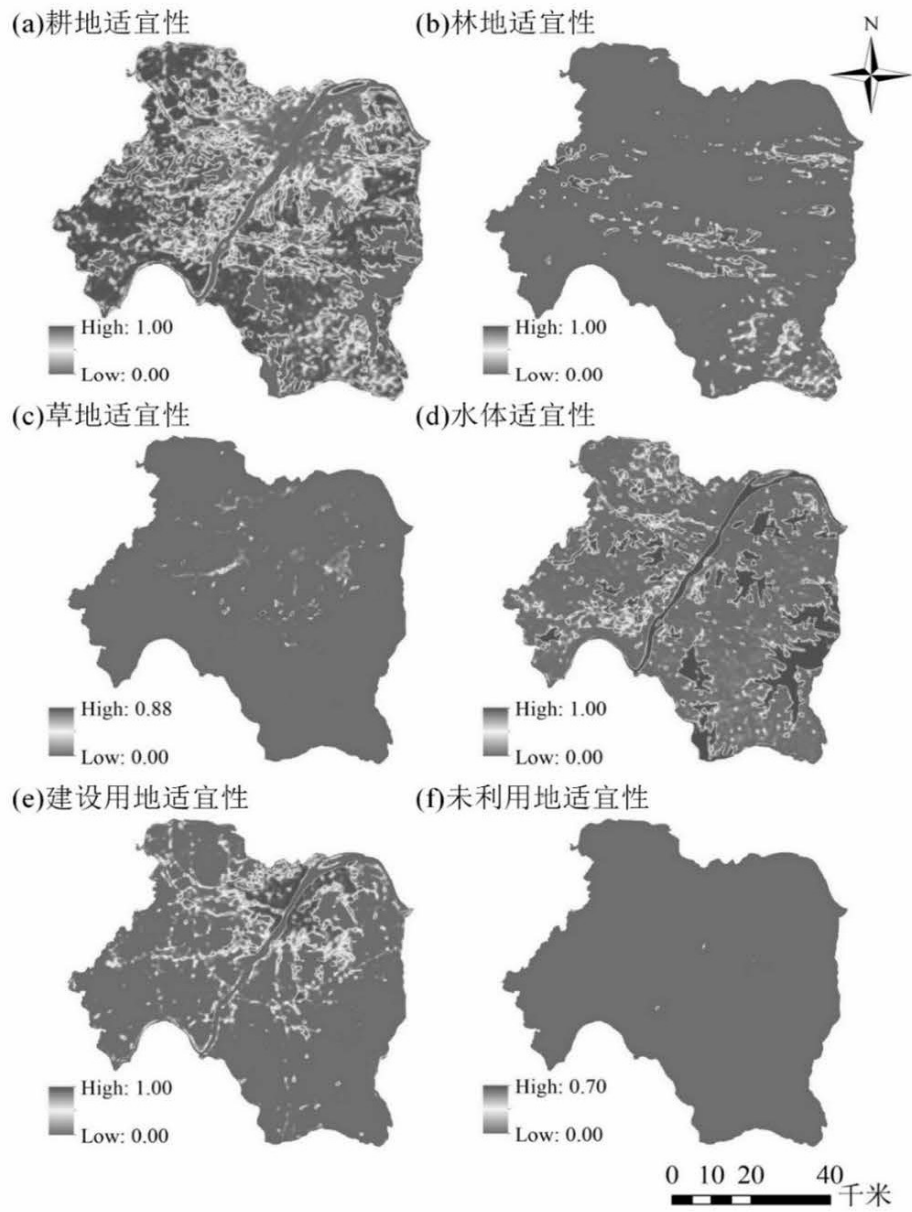


图7