

昆明市土地利用变化趋势的多模型对比分析

许安泽¹, 张述清², 朱大明¹, 裘木兰¹, 保红³

〔1.昆明理工大学 国土资源工程学院, 云南 昆明 650093; 2.云南省国土资源规划设计研究院, 云南 昆明 650216; 3.昆明市不动产权籍调查中心(昆明市国土规划勘察测绘研究院) 650200, 云南 昆明〕

摘要: [目的] 对比分析逻辑回归(LogReg)、多层神经网络(MLP)和相似度加权学习(SimWeight)3 种土地利用变化趋势模型在多地类变化分析中的模拟效果与土地利用预测精度,为云贵高原地区的国土空间规划、水土保持和生态修复等提供参考。[方法] 以 2000—2020 年昆明市土地利用变化为例,分别采用 3 种模型对该区域建模,并使用受试者工作特征曲线、ROC 曲线下面积和 Kappa 系数等多种方法评估精度。[结果] 对于大多数地类的变化趋势,MLP 和 SimWeight 模型的模拟效果要好于 LogReg 模型,特别是在未利用地的变化分析中 MLP 和 SimWeight 模型的 AUC 值均大于 0.9。在整体土地利用变化预测上,LogReg,MLP 和 SimWeight3 种模型的 Kappa 值分别为 0.9066,0.904 1,0.925 3,整体预测结果表现接近,但 SimWeight 模型略微优于其他模型。[结论] 对于昆明市 LUCC 建模的模型选择,若为追求更高精度可选择 SimWeight 模型,若更在意运算速度则优先选择 MLP 模型,若需要进一步分析驱动因子与土地变化的关系应选择 LogReg 模型。

关键词: 土地利用变化; 逻辑回归; 多层神经网络; 相似度加权学习

文献标识码: B

文章编号: 1000-288X(2023)01-0141-08

中图分类号: F301.2

文献参数: 许安泽, 张述清, 朱大明, 等. 昆明市土地利用变化趋势的多模型对比分析[J]. 水土保持通报, 2023, 43(1):141-148. DOI:10.13961/j.cnki.stbctb.2023.01.017; Xu Anze, Zhang Shuqing, Zhu Daming, et al. Comparative analysis of three land use transition models in Kunming City [J]. Bulletin of Soil and Water Conservation, 2023, 43(1):141-148.

Comparative Analysis of Three Land Use Transition Models in Kunming City

Xu Anze¹, Zhang Shuqing², Zhu Daming¹, Qiu Mulan¹, Baohong³

〔1.Kunming University of Science and Technology, Kunming, Yunnan 650093,

China; 2.Land and Resources Planning and Design Institute of Yunnan Province,

Kunming, Yunnan 650216, China; 3.Kunming Real Estate Ownership Investigation Center

(Kunming Institute of Surveying and Mapping of Land Planning and Prospecting), Kunming, Yunnan 650200, China〕

Abstract: [Objective] The simulation effects and land use prediction accuracy of three land use transition models (LogReg, MLP, and SimWeight) for multi-land change analysis were compared and analyzed in order to provide a reference for land space planning, soil and water conservation, and ecological restoration in the Yunnan-Guizhou Plateau. [Methods] Three models were used to simulate land use changes for Kunming City from 2000 to 2020. Various methods such as ROC curve, AUC coefficient, and kappa coefficient were used to evaluate model performance. [Results] The simulations of change trends for most land use classes by MLP and SimWeight were better than the simulations by LogReg, especially with regard to the change analysis of unused land where the AUC values for MLP and SimWeight were greater than 0.9. The kappa values for LogReg, MLP, and SimWeight were 0.906 6, 0.904 1, and 0.925 3, respectively. The overall prediction results for the three models were similar, and SimWeight was slightly better than the two other models. [Conclusion] For LUCC modeling of Kunming City, SimWeight is recommended when the user desires higher accuracy, MLP is recommended when the user cares more about calculation speed, and LogReg is recommended when the

收稿日期:2022-06-16

修回日期:2022-08-02

资助项目:国家自然科学基金项目“滇中城市群国土格局多尺度演化模拟及优化配置”(41761081)

第一作者:许安泽(1994—),男(汉族),福建省泉州市人,硕士研究生,研究方向为 GIS 应用。Email:andrewbull@foxmail.com。

通信作者:张述清(1965—),男(汉族),重庆市人,教授级高级工程师,硕士,主要从事国土空间规划研究。Email:1163041121@qq.com。

user wants to further analyze the relationship between driving factors and land use changes.

Keywords: LUCC; LogReg; MLP; SimWeight

土地利用变化(land used/cover change, LUCC)在全球物质循环和能量流动中起着关键的作用,是导致全球环境变化的主要组成部分和深入理解人类与环境复杂关系的重要手段^[1-2]。长期以来,学者们构建大量的 LUCC 模型来研究土地利用变化,试图通过 LUCC 模型揭示土地变化的内在机制与原因,并依此预测未来土地变化的方向^[3]。近年常见的 LUCC 模型有小尺度土地利用变化及其空间效应模型(conversion of land use and its effects at small region extent, CLUE-S)^[4-5]、多智体模型(agent based model, ABM)^[6]、元胞自动机—马尔可夫链(CA-Markov)^[7-8]、系统动力学(system dynamics, SD)^[9]、土地系统动态模拟模型(dynamics of land system, DLS)^[10-11]等。虽然不同的 LUCC 模型的结构不尽相同,但大部分 LUCC 模型都采用驱动因子与土地变化关系计算而来的土地变化趋势图集为基础实现土地变化的预测。当前常用的土地变化趋势图计算模型有逻辑回归(LogReg)、相似度加权学习(SimWeight)和多层神经网络(MLP)。LogReg 模型是被应用最广泛的模型,具有原理简单操作方便等诸多优点,但因其不能处理共线的问题一直饱受诟病^[12-13]。MLP 模型属于机器学习的范畴,是能够同时处理多因子、多地类变化和共线的模型,但该方法计算内部不可见^[14-15]。SimWeight 模型也属于机器学习,是一种基于距离加权的方法,因此计算比 MLP 模型要简单,但其对异构跃迁建模的适用性有限^[16]。

当前对于这 3 种常见的土地利用变化趋势算法的对比研究,国外有学者 Mozumder 等^[17]将这 3 种模型应用于印度东北部古瓦哈提市建设用地的变化,对比分析不同模型对城市扩张模拟的优劣。Parsamehr 等^[18]在研究伊朗赫卡尼亚森林减少砍伐和森林退化导致温室气体排放的项目中对比这 3 种模型在该项目选址上的不同性能效果。国外已有的对比研究都是针对单一地类变化,没有考虑多种地类变化的对比研究,而国内则尚未见相关研究。昆明市作为中国西南地区重要的城市之一,是云贵高原中部具有代表性的城市,同时也是滇中城市群的核心城市,土地利用变化多样且复杂。因此本文以昆明市为例,利用 TerrSet 软件建立 3 种不同土地利用变化趋势模型,对比分析 3 种模型在该地区多地类变化的效果,为云贵高原地区的国土空间规划、水土保持和生态修复等提供参考。

1 研究区概况与数据来源

1.1 研究区概况

昆明市位于东经 102°10′—103°40′,北纬 24°23′—26°33′,总面积约 21 012.54 km²。总体地势由北向南呈阶梯状逐渐降低,平均海拔约为 2 000 m,属于低纬度高原山地季风气候,年平均气温 15 ℃,享有“春城”的美誉。第七次全国人口普查常住人口 8 460 088 人,2020 年生产总值 6.73×10¹¹ 元。昆明市区位独特,处于东盟“10+1”自由贸易区经济圈、大湄公河次区域经济合作圈、“泛珠三角”区域经济合作圈的交汇点,是我国面向南亚、东南亚的前沿和门户。

1.2 数据来源与预处理

土地利用数据来源于中国科学院地理科学与资源研究所美国陆地卫星 Landsat 遥感影像作为主要信息源解译的中国土地利用/土地覆盖类型空间分布图,以及第二次全国土地调查和第三次全国国土调查。在综合考虑昆明市土地利用分布和本研究的重点前提下,根据《第二次全国土地调查技术规程》《第三次全国国土调查技术规程》《土地利用现状分类》等技术标准将地类划分为 6 类:耕地、林地、草地、水域、建设用地、未利用地(图 1)。土地利用变化驱动因子数据主要来源于 2000—2020 年昆明市统计年鉴、中科院资源环境科学与数据中心(<http://www.resdc.cn>),包括人口分布、社会生产总值、年平均气温、高程、土壤类型分布等。对所有驱动因子数据进行处理,使它们具有一致的空间坐标、像元大小、像元单位和处理范围。同时考虑到驱动因子数据中不同的特征量纲不一致,数值间的差别可能很大,对具有连续特征的驱动因子作归一化处理,使数据映射到 0~1 的区间内。

1.3 土地利用变化趋势模型

1.3.1 逻辑回归模型(LogReg) 逻辑回归(logistic regression, LogReg)是一种非常经典的算法,常被应用于二分类的问题,因其形式简单、易于建模、可解释性强等优点而被广泛使用^[19]。土地利用变化可以认为是一个二分类问题,即单位土地用地类型发生变化的可能性只能是发生变化或保持不变,LogReg 模型试图通过属性的线性结合函数来预测该变化^[20]。LogReg 模型函数如公式(1)所示,只需要确定式中和的值,整个模型就得以确定。计算时通常先使用对数几率函数作为替代函数,在一定的转换后得到公式(2),然后利用极大似然估计的方法来估计 ω 和 b 。

$$P(y=1|x) = \omega_1 x_1 + \omega_2 x_2 \cdots \omega_d x_d + b \quad (1)$$

$$P(y=1|x) = \frac{e^{\omega T x + b}}{1 + e^{\omega T x + b}} \quad (2)$$

式中: y 为地类发生变化的可能性, 该值只能是 1 或

0, 表示为发生变化或保不变; $P(y=1|x)$ 表示地类变化的概率; x 为土地利用变化的驱动因子; d 为驱动因子的个数; ω 为驱动因子系数; b 为驱动因子常数项系数。

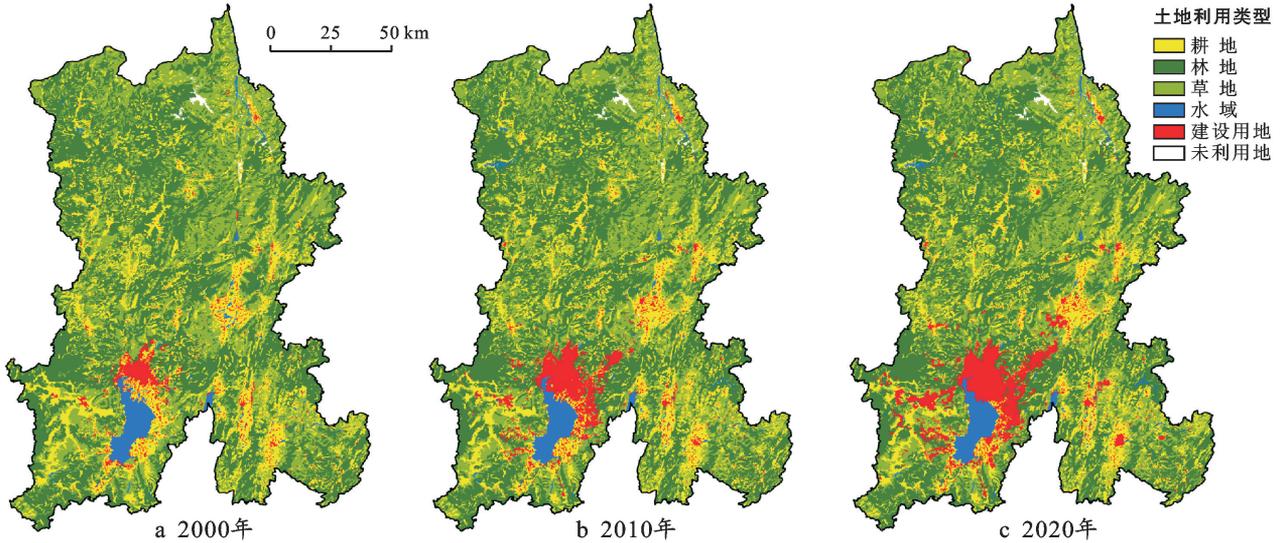


图 1 2000—2020 年昆明市土地利用变化特征

Fig.1 Characteristics of land use change of Kunming City from 2000 to 2020

1.3.2 多层神经网络模型(MLP) 神经网络方面的研究在二十世纪五六十年代就已开始, 最早的研究是从感知机开始的, 该模型结构仅由两层神经元组成。简单的感知机模型学习能力非常有限, 甚至不能解决异或这样的非线性问题, 因此提出了多层神经网络模型(multilayer perceptron, MLP)的概念^[21]。MLP模型的结构由输入层、隐藏层和输出层组成, 输入层和隐藏层都拥有激活函数的功能神经元, 每层神经元与下一层神经元完全互连, 同层之间的神经元不存在相连, 也不存在跨层连接。MLP模型通过训练数据调整神经元之间的连接权重以及每个功能神经元的阈值, 整个过程是由计算机自动完成学习。在土地利用变化研究中建立的 MLP 模型, 输入层即对应土地变化的驱动因子, 输出层则代表不同地类的变化可能性^[22]。因为 MLP 模型结构的特点, 相较于 LogReg 模型与 SimWeight 模型而言, MLP 模型可以同时处理多种地类之间的变化。

1.3.3 距离加权最近邻模型(SimWeight) MLP 模型虽然有很多优点, 但对其进行分析需要调整许多参数, 复杂的参数使得 MLP 模型的内部变得不可见。因此 Sangermano 等^[16] 提出距离加权最近邻模型(similarity weighted, SimWeight), 该模型的底层逻辑来源于 k -近邻算法(k -nearest neighbor, KNN)^[23], 是一种基于周围土地状态的计算模型。对于要预测

的单位土地, SimWeight 方法首先提取 k 个邻近的土地单元, 然后计算它们到预测单元的距离, 最后将其带入指数加权函数中, 其计算公式如下:

$$M = \frac{\sum_{i=1}^c (1 - \frac{1}{1 + e^{1/d_i}})}{k} \quad (c \leq k) \quad (3)$$

式中: M 为土地的变化概率; d_i 为邻近变化土地的距离(m); k 为选定邻近单位土地的数量; c 为邻近单位土地中发生变化的数量。

1.4 驱动因子选择

土地利用变化驱动因子是指导致土地利用方式和目的变化的主要自然因素和社会经济因素, 驱动因子的选择是研究土地利用变化的核心问题, 也是构建土地利用变化趋势模型的基础。为选择合适的驱动因子, 使用克莱姆相关系数(Cramer's V)^[24] 量化驱动因子与土地利用变化的相关性。Cramer's V 是双变量相关分析的一种方法, 专门用于衡量分类数据之间的相关程度, 适合初步评价土地利用的驱动因子, 其计算式为:

$$\varphi_c = \sqrt{\frac{\chi^2}{N(k-1)}} \quad (4)$$

式中: φ_c 为克莱姆相关系数; χ^2 为皮尔森卡方检验系数; N 为样本数量; k 为任一变量最小的类别数。

如表 1 所示, 列举所选驱动因子的 Cramer's V 系数值, 取值范围介于 0 到 1 之间, 0 表示两个变量完

全无关,1 表示完全相关。为确保 LogReg,MLP 和 SimWeight 模型评价的准确性,3 种模型均选取表中

所有的驱动因子参与建模分析,以控制其他变量的一致性。

表 1 昆明市土地利用变化驱动力分析

Table 1 Drivers analysis of land use change in Kunming City

Cramer's V	耕地	林地	草地	水域	建设用地	未利用地	所有地类
高程	0.231 6	0.290 5	0.160 1	0.239 2	0.262 4	0.477 8	0.302 3
坡度等级	0.292 8	0.388 5	0.175 3	0.441 8	0.373 0	0.034 8	0.315 8
土壤分类	0.237 2	0.274 3	0.168 4	0.789 8	0.331 6	0.170 3	0.409 4
河流距离	0.060 9	0.067 1	0.034 5	0.066 9	0.027 9	0.023 1	0.048 9
主要道路距离	0.129 2	0.084 6	0.025 8	0.186 2	0.012 6	0.098 2	0.110 9
坝区范围距离	0.172 4	0.218 6	0.106 5	0.121 6	0.176 3	0.162 2	0.160 4
夜光数据	0.173 5	0.272 1	0.107 8	0.170 4	0.641 6	0.028 8	0.315 1
人口密度	0.082 9	0.158 5	0.073 4	0.012 2	0.534 9	0.017 7	0.243 9
GDP	0.079 4	0.126 7	0.092 4	0.115 1	0.316 6	0.031 3	0.161 2
年降水	0.116 3	0.122 8	0.151 5	0.253 8	0.251 7	0.441 1	0.270 0
年平均气温	0.215 7	0.270 3	0.167 3	0.293 1	0.242 4	0.447 0	0.295 5

2 模型评价方法

2.1 模型评价技术路线

以 2000—2010 年昆明市土地利用分布图为基础,建立 3 种不同的土地利用变化趋势模型,生成一系列不同地类间的趋势图集。使用概率并集 (LogicOR) 方法叠加趋势图集,产生 3 幅不同模型的软预测图,即表示未来单位土地变化的概率。并采用 ROC 曲线和 AUC 对比 3 种模型产生的土地利用变化图集与软预测图。同时使用多标准评价方法 (mult-criteria evaluation, MCE) 对 2020 年昆明市的土地利用分布进行预测,其结果称为硬预测,并采用 Kappa 系数和误差分析评价不同模型的硬预测结果。3 种模型的对比研究路线见图 2。

2.2 土地利用变化趋势评价方法

土地利用变化趋势图集以及由它叠加产生的软预测图皆为概率预测,此类预测通常可以采用 ROC 曲线泛化其性能。受试者工作特征曲线 (receiver operating characteristic curve, ROC) 是一种常用于机器学习领域的评估工具,它是根据不同的分类阈值计算分类结果的混淆矩阵,然后依次算出不同阈值下分类的假正例率 (false positive rate, FPR) 和真正例率 (true positive rate, TPR),分别对应 ROC 曲线的横轴和纵轴^[25]。在进行对比时,若一条曲线完全包住另一条曲线,则可断言该曲线的模型性能优于后者。如果两条 ROC 曲线发生交叉,则难以断言两者的优劣。此时若仍然要进行比较,较为合理的判断方法是比较 ROC 曲线下面积 (area under ROC curve, AUC)^[26-27]。通常,AUC 的值介于 0.5 到 1.0 之间,

0.5 表示随机分类,较大的 AUC 代表模型拥有更好的性能^[27]。

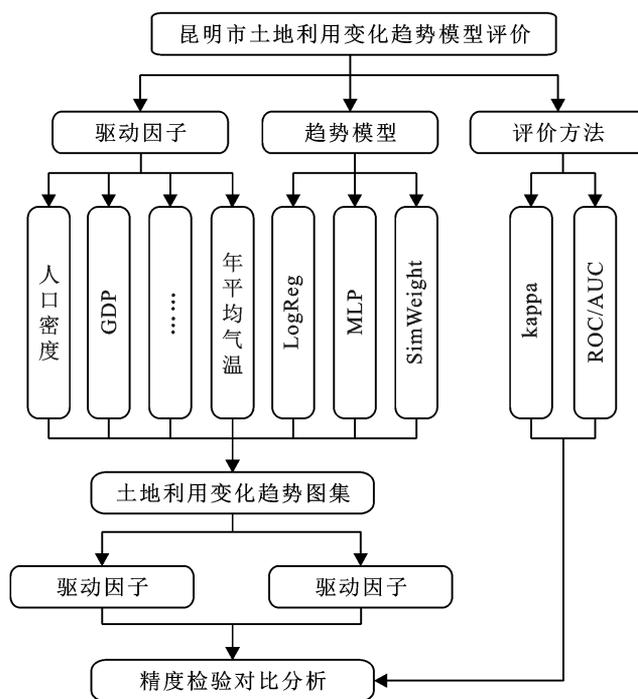


图 2 土地利用变化趋势模型评价流程

Fig.2 Evaluation process of land use change trend model

2.3 土地利用硬预测评价方法

硬预测的结果是具有明确地类分类的土地利用覆盖图,对于该类预测的评价,我们借用遥感分类中常用的精度检验 kappa 系数作为整体预测结果评价标准,该系数是一个用于一致性检验的指标,也可以用于衡量分类的效果^[28]。kappa 系数基于误差表计

算,取值为-1到1之间,且通常大于0。kappa值的区间为0~0.20表示极低的一致性,0.21~0.40表示一般的一致性,0.41~0.60表示中等的一致性,0.61~0.80表示高度的一致性,0.81~1表示几乎完全一致。kappa计算式为:

$$\text{kappa} = \frac{P_o - P_c}{P_p - P_c} \quad (5)$$

式中: P_o 为土地利用分布模拟正确模拟单元格的比例; P_c 为随机模拟正确单元格的比例; P_p 为理想模拟正确单元格的比例即100%。

3 模型精度评价分析

3.1 变化趋势图集评价

土地利用变化趋势图是一系列地类变化的概率图,也是土地利用变化趋势模型计算结果的直接产物,作为LUCC预测模型的重要参与者,图集的好坏直接影响土地利用变化预测的效果。比对LogReg,MLP和SimWeight3种模型计算得到的土地利用趋

势图集的AUC如图3所示。每幅子图的横坐标表示研究初期的土地利用类型,纵坐标表示AUC计算值,子图名表示研究末期的土地利用类型,并以不同的颜色标示土地利用变化趋势模型。观察发现,不同的地类变化对应的3种不同模型计算结果表现不尽相同,例如在水域—耕地的变化上SimWeight模型的AUC值为0.913明显大于另两个模型,而在未利用地—建设用地的变化上AUC值最大的则是MLP模型为0.990。但大部分地类变化使用LogReg模型的效果都要差于MLP和SimWeight模型,特别是在图3f未利用地的表现上LogReg模型AUC值接近于0.5,几乎无异于随机分类,而MLP和SimWeight模型的AUC值均大于0.9。同时观察到图3e,在建设用地的扩张上MLP模型表现十分优异,AUC值均保持在0.8以上,除未利用地—建设用地外SimWeight模型在这方面的表现也较好。但是在建设用地—林地、建设用地—草地这两种地类变化的潜力分析中3种模型的表现都较差。

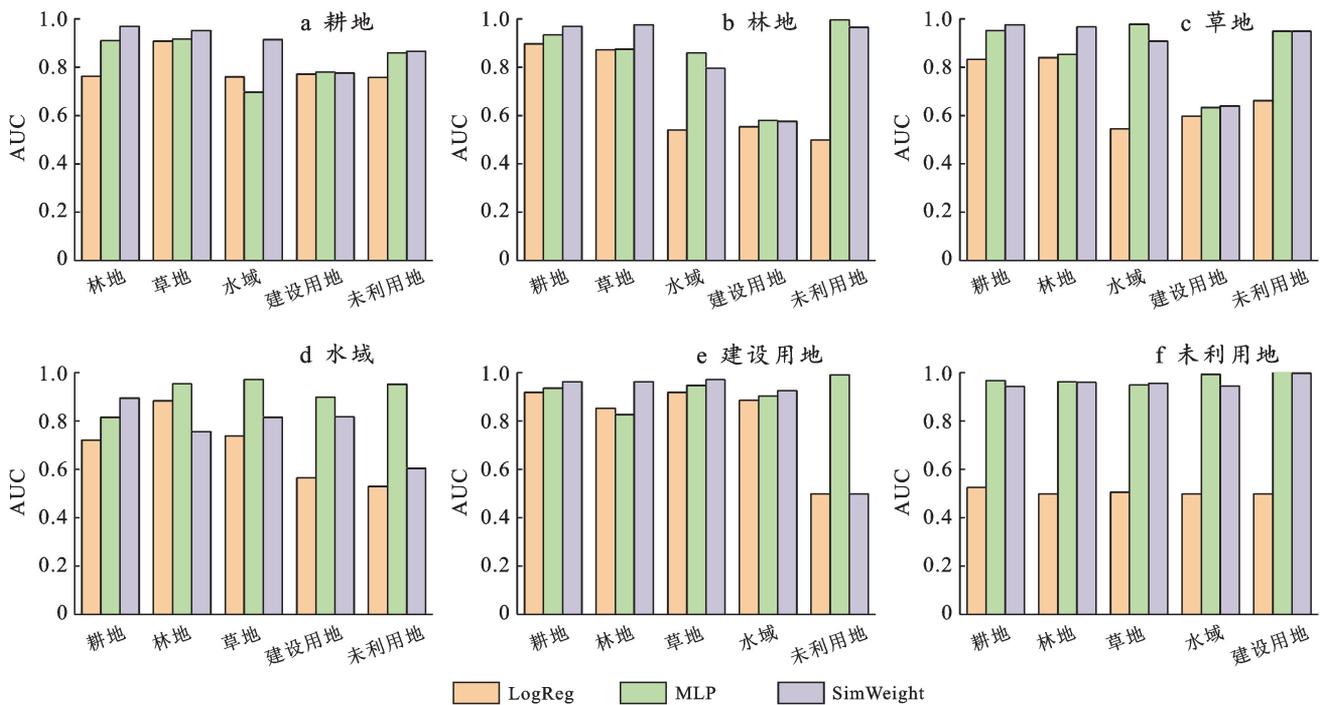


图3 土地利用变化趋势曲线下面积(AUC)

Fig.3 Area under curve of transition potentials of land use change

3.2 软预测精度评价

从土地利用变化趋势图集对比来看,LogReg,MLP和SimWeight这3种模型在不同地类变化的趋势预测上各有优势,为了研究模型在整体土地利用变化趋势预测的效果。以不同模型计算得到的2000—2020年昆明市土地利用变化趋势图集为基础,采用概率并集(LogicOR)方法^[29]整合土地利用趋势图集,

即将不同地类变化的概率看成是独立的,求这些事件的并集。概率并集计算式为:

$$P(a \cup b) = P(a) + P(b) - P(ab) \quad (6)$$

式中: $P(a \cup b)$ 表示土地利用变化概率; $P(a)$, $P(b)$ 分别表示两种不同地类的变化概率。

如图4所示,使用LogicOR方法计算出3种不同模型的软预测图,软预测的结果仍然是一种概率

图,表示单位土地发生变化的可能性。观察 3 幅不同模型计算的软预测图发现,3 种模型预测出的整体趋势是比较相近的,例如在滇池及滇池北部已建成的老城区的预测上,3 种模型都给出的较低的变化概率,与实际情况相符。但在部分区域上 3 种模型给出的概率存在差距,为量化软预测的预测效果,分别作 3 种不同模型软预测的 ROC 曲线并计算它们的 AUC

值。由图 5 可知,3 条 ROC 曲线均发生交叉,不能直接断言某个模型具有更大优势,因此可采用 AUC 作为量化精度的评价标准。SimWeight 模型 AUC 值为 0.837,LogReg 和 MLP 模型的 AUC 值分别为 0.783,0.789。软预测 AUC 精度验证结果显示 LogReg 和 MLP 模型的效果接近,而 SimWeight 模型具有更好的性能。

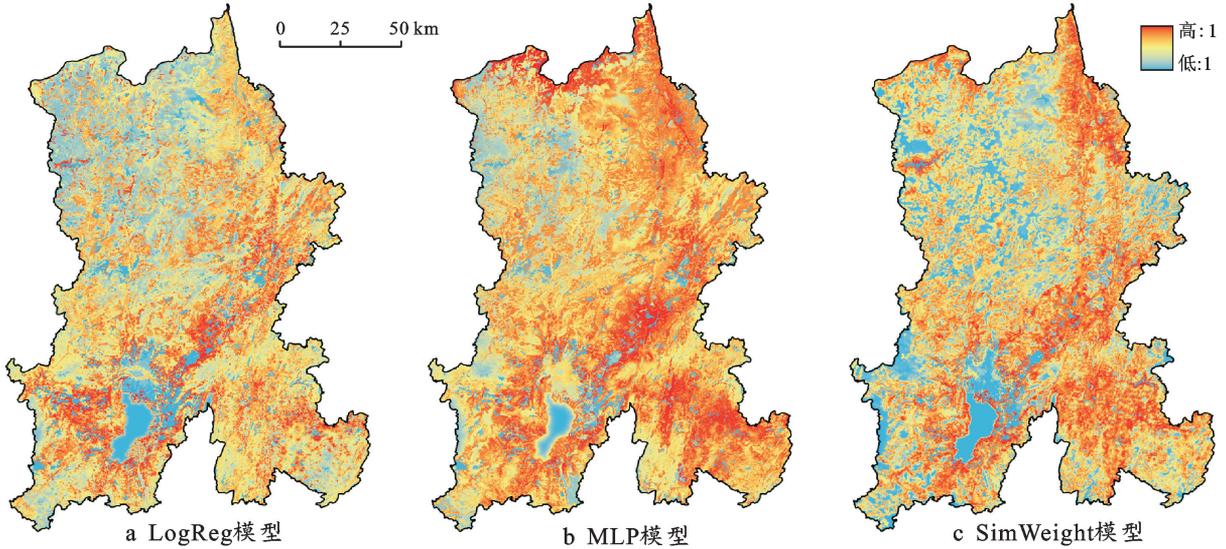


图 4 2020 年昆明市土地利用变化软预测

Fig.4 Soft prediction of land use change of Kunming City in 2020

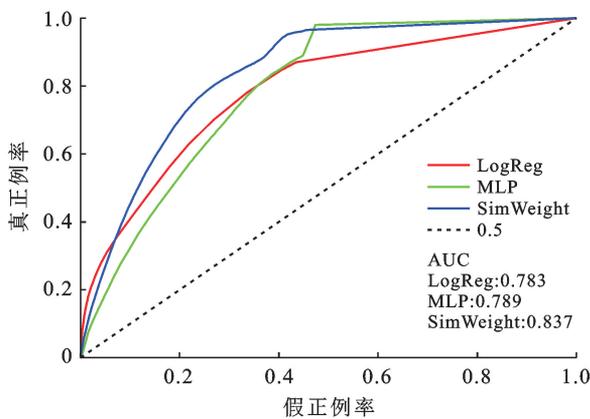


图 5 软预测受试者特征曲线(ROC)

Fig.5 Receive operating characteristic curve of soft prediction

3.3 硬预测精度评价

土地利用变化的预测,往往更希望得到一幅具有明确地类分类的土地利用分布图作为最终的预测结果图,常使用土地利用变化趋势图集结合多标准评价方法作为该预测的手段。多标准评价法(Multicriteria Evaluation, MCE)是一种基于多种条件的评估决策系统,其结果是一幅预测时间节点的土地利用分布图,称为硬预测^[30-31]。MCE 决策条件被称为标准,这些标准可能是限制条件或激励因子,限制条件一般表

现为二值化的规则,它的作用是排除某些变化趋势,比如在生态保护区与耕地红线内禁止建设用地的扩张。激励因子是连续的数据,体现某些变化的相对适宜性,比如已有规划的道路周围有更大的可能会出现新的建设用地。该方法的计算步骤首先查询土地利用变化趋势图上单位土地变化概率,并为单位土地创建一个地类流出与流入概率表,然后依据转移概率矩阵计算的变化需求量与保护区等限制条件重新分配单位土地的地类,最后将重新分配的地类替换原有地类^[32]。

以 LogReg,MLP 和 SimWeight 3 种模型计算的土地利用变化趋势图集为数据基础,使用 MCE 方法,控制除趋势图集以外的其他变量相等,预测 2020 年昆明市的土地利用分布(图 6)。将 3 种模型的预测结果与图 1c 的 2020 年真实土地利用作对比发现,MLP 模型在东南地区的石林县和宜良县的预测效果比较差,该模型错误地预测更多的建设用地,在嵩明县建设用地扩张的预测上,3 种模型对建设用地的扩张与真实情况有出入。在滇池周边建设用地的预测上 3 种模型都具有较好表现,它们都较为正确地预测昆明市主城区的扩张。在非建设用地的其他地类上,3 种模型的性能也较好,特别是对昆明市北部山区的预测。

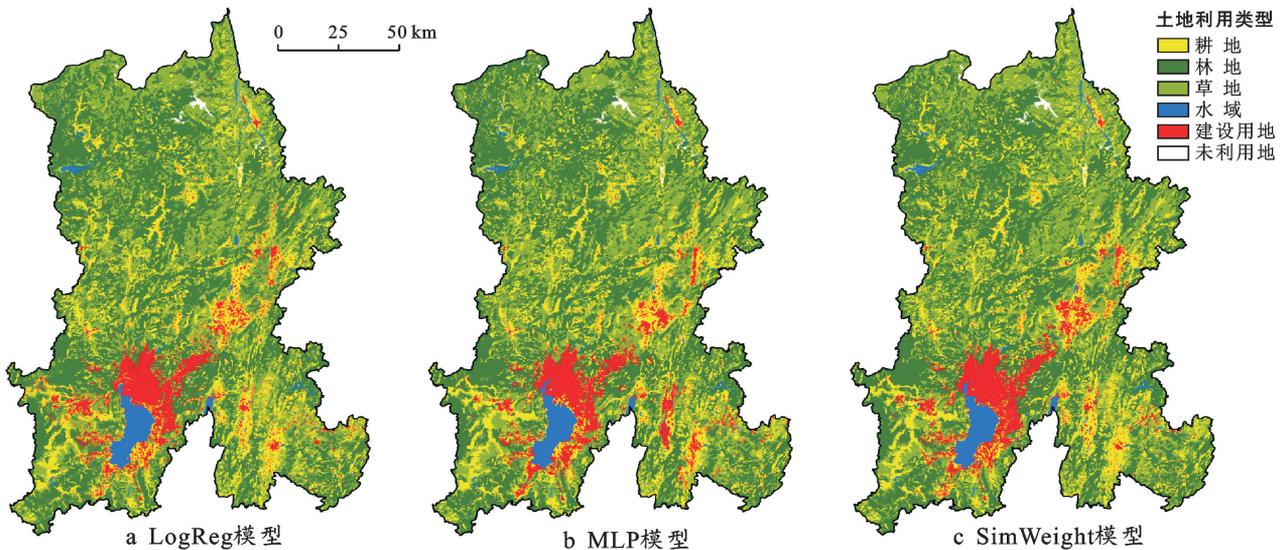


图 6 2020 年昆明市土地利用变化硬预测

Fig.6 Hard prediction of land use change of Kunming City in 2020

为量化基于 3 种模型的硬预测效果,作预测结果与真实地类分布的混淆矩阵,并依此计算预测的误判率和遗漏率(表 2)。误判率表示地类被判断错误的概率,遗漏率表示地类漏概率,两项指标的值都是越小表示预判的效果越好。同时根据公式(5)计算得到 LogReg,MLP 和 SimWeight3 种模型硬预测 kappa 值分别为 0.906 6,0.904 1,0.925 3。表 2 土地利用变化预测误差表对比不同地类预测的误差结果,3 种模型在不同地类的变化预测中都有较高的准确率,特别是林地的预测误判率和遗漏率均保持在较低的水平,在总体预测上,3 种模型的误判率和遗漏率均小于 10%。而 kappa 值则反应整体的预测效果,计算结果显示 3 种模型的 kappa 值均大于 0.9,故就硬预测的整体结果而言 3 种模型都具有很好性能。

表 2 2020 年昆明市土地利用变化硬预测误差

Table 2 Hard prediction error of land use change of Kunming City in 2020

土地类型	误判率			遗漏率		
	LR	MLP	SW	LR	MLP	SW
耕地	8.56	6.56	6.29	8.46	7.08	6.73
林地	3.84	5.10	3.06	4.60	5.29	3.05
草地	6.72	6.96	5.85	6.10	6.76	5.50
水域	12.97	11.32	9.31	11.76	10.36	8.81
建设用地	13.78	12.94	11.07	12.25	11.50	11.72
未利用地	10.75	10.06	8.02	9.66	9.89	6.86
总体	9.44	8.82	7.27	8.81	8.48	7.11

4 结论

分析 LogReg,MLP 和 SimWeight 3 种不同的土地利用变化趋势模型在 2000—2020 年昆明市多地类

变化的应用,对比 3 种模型计算的变化趋势图集以及在此基础上生成的软预测和硬预测。

(1) 对于大部分地类变化的趋势预测 MLP 和 SimWeight 模型的变化趋势图集要好于 LogReg 模型。SimWeight 模型侧重于以往地类变化的距离,而 LogReg 模型则更加关注的是驱动因素与变化之间的线性关系,在具有明显扩张的地类上,SimWeight 模型要好于 LogReg 模型。MLP 模型通过多次迭代产生驱动因子与土地变化的最佳拟合,也能够较好地估计土地利用变化的潜力。

(2) 软预测和硬预测的对比评估显示,3 种模型都具有较高的性能,误判率和遗漏率都小于 10%,kappa 系数均大于 0.9,MLP 和 LogReg 模型的性能相近,SimWeight 模型略优于前两者。但需要注意的是 SimWeight 模型是一种非线性的方法,它的可解释性并没有 LogReg 模型强,但又不如 MLP 模型能够同时对多地类建模。

(3) 综合分析表明,对于昆明市 LUCC 建模的模型选择,若为追求更高精度可选择 SimWeight 模型,若更在意运算速度则优先选择 MLP 模型,若需要进一步分析驱动因子与土地变化的关系应选择 LogReg 模型。值得关注的是 SimWeight 模型对于所选的邻近单位土地数量 k 值非常敏感, k 值的微小变动都可能对结果产生很大的影响,因此之后的研究作者将会把重点放在 SimWeight 模型 k 值选择的研究上。

[参 考 文 献]

[1] Foley J A, Defries R, Asner G P, et al. Global consequences of land use [J]. Science, 2005,309(5734):570-574.
 [2] 刘纪远,张增祥,张树文,等.中国土地利用变化遥感研究

- 的回顾与展望: 基于陈述彭学术思想的引领[J]. 地球信息科学学报, 2020, 22(4): 680-687.
- [3] 肖宝玉. 中国 LUCC 研究特征与趋势: 基于 CiteSpace 的分析[J]. 亚热带资源与环境学报, 2020, 15(1): 61-70.
- [4] Verburg P H, Soepboer W, Veldkamp A, et al. Modeling the spatial dynamics of regional land use: The CLUE-S model [J]. *Environmental Management*, 2002, 30(3): 391-405.
- [5] 张沐锋, 刘万侠, 王健恩, 等. 基于 Clue-S 模型的石马河流域东莞段生态系统服务价值变化情景模拟[J]. 水土保持通报, 2021, 41(1): 152-160.
- [6] Bakker M M, Alam S J, van Dijk J, et al. Land-use change arising from rural land exchange: An agent-based simulation model [J]. *Landscape Ecology*, 2015, 30(2): 273-286.
- [7] 冯丽媛, 米文宝, 马国庆. 基于 CA-Markov 模型的宁夏沿黄生态经济带土地利用变化及模拟研究[J]. 水土保持通报, 2019, 39(5): 218-222.
- [8] Gemitzi A. Predicting land cover changes using a CA Markov model under different shared socioeconomic pathways in Greece [J]. *GIScience & Remote Sensing*, 2021, 58(3): 425-441.
- [9] Shen Qiping, Chen Qing, Tang B S, et al. A system dynamics model for the sustainable land use planning and development [J]. *Habitat International*, 2009, 33(1): 15-25.
- [10] Deng Xiangzheng, Su Hongbo, Zhan Jinyan. Integration of multiple data sources to simulate the dynamics of land systems [J]. *Sensors (Basel, Switzerland)*, 2008, 8(2): 620-634.
- [11] 田益多, 梅昀, 陈银蓉. 基于 Markov-DLS 模型的江西省多情景下土地利用时空演变分析[J]. 水土保持通报, 2021, 41(3): 218-227.
- [12] Rutherford G N, Guisan A, Zimmermann N E. Evaluating sampling strategies and logistic regression methods for modelling complex land cover changes [J]. *Journal of Applied Ecology*, 2007, 44(2): 414-424.
- [13] Lin Yingzhi, Deng Xiangzheng, Li Xing, et al. Comparison of multinomial logistic regression and logistic regression: Which is more efficient in allocating land use? [J]. *Frontiers of Earth Science*, 2014, 8(4): 512-523.
- [14] Ozturk D. Urban growth simulation of atakum (Samsun, Turkey) using cellular automata-Markov chain and multi-layer perceptron-Markov chain models [J]. *Remote Sensing*, 2015, 7(5): 5918-5950.
- [15] Guo Andong, Zhang Yuqing, Hao Qing. Monitoring and simulation of dynamic spatiotemporal land use/cover changes [J]. *Complexity*, 2020, 2020: 3547323.
- [16] Sangermano F, Eastman J R, Zhu Honglei. Similarity weighted instance-based learning for the generation of transition potentials in land use change modeling [J]. *Transactions in GIS*, 2010, 14(5): 569-580.
- [17] Mozumder C, Tripathi N K, Losiri C. Comparing three transition potential models: A case study of built-up transitions in North-East India [J]. *Computers, Environment and Urban Systems*, 2016, 59: 38-49.
- [18] Parsamehr K, Gholamalifard M, Kooch Y. Comparing three transition potential modeling for identifying suitable sites for REDD+ projects [J]. *Spatial Information Research*, 2020, 28(2): 159-171.
- [19] Allwein E, Schapire R, Singer Y. Reducing multiclass to binary: A unifying approach for margin classifiers [J]. *Journal of Machine Learning Research*, 2001, 1(2): 113-141.
- [20] 李强, 任志远. 基于 Logistic 回归分析的土地利用变化空间统计与模拟[J]. 统计与信息论坛, 2012, 27(3): 98-103.
- [21] Rumelhart D E, Hinton G E, Williams R J. Learning representations by back-propagating errors [J]. *Nature*, 1986, 323(6088): 533-536.
- [22] Chan J C W, Chan K P, Yeh A. Detecting the nature of change in an urban environment: A comparison of machine learning algorithms [J]. *Photogrammetric Engineering and Remote Sensing*, 2001, 67(2): 213-225.
- [23] Weinberger K, Saul L. Distance metric learning for large margin nearest neighbor classification [J]. *Journal of Machine Learning Research*, 2009, 10: 207-244.
- [24] Cramér H. *Mathematical methods of statistics (PMS-9)*, volume 9 [J]. Princeton, NJ, USA, 1946.
- [25] Bradley A P. The use of the area under the ROC curve in the evaluation of machine learning algorithms [J]. *Pattern Recognition*, 1997, 30(7): 1145-1159.
- [26] Hand D J, Till R J. A simple generalisation of the area under the ROC curve for multiple class classification problems [J]. *Machine Learning*, 2001, 45(2): 171-186.
- [27] Fawcett T. An introduction to ROC analysis [J]. *Pattern Recognition Letters*, 2006, 27(8): 861-874.
- [28] Pontius R. Quantification error versus location error in comparison of categorical maps [J]. *Photogrammetric Engineering and Remote Sensing*, 2000, 66: 1011-1016.
- [29] 陈希孺. 概率论与数理统计[M]. 安徽 合肥: 中国科学技术大学出版社, 2009.
- [30] Buckley J J. The multiple judge, multiple criteria ranking problem: A fuzzy set approach [J]. *Fuzzy Sets and Systems*, 1984, 13(1): 25-37.
- [31] Eastman J R, Jin Weigen, Kyem P, et al. Raste procedure for multi-criteria/multi-objective decisions [J]. *Photogrammetric Engineering and Remote Sensing*, 1995, 61(5): 539-547.
- [32] 张晓娟, 周启刚, 王兆林, 等. 基于 MCE-CA-Markov 的三峡库区土地利用演变模拟及预测[J]. 农业工程学报, 2017, 33(19): 268-277.