

商品住宅价格批量评估研究： 以海口市海甸岛为例

韩念龙^{1,2}, 罗燕², 张非凡², 王德思², 黎兴强²

(1. 惠州学院 地理与旅游学院, 广东 惠州 516007;

2. 海南大学 公共管理学院, 海南 海口 570100, E-mail: 450073618@qq.com)

摘要: 为满足日益增长的房地产评估需求, 批量评估技术正逐渐取代传统的单宗房地产评估。以海口市海甸岛商品住宅小区为研究对象, 通过采集区域的POI数据、小区价格数据和房屋属性特征数据, 基于商品住宅小区的内生、区位和邻里特征, 构建基于GWR的海甸岛商品住宅小区价格模型。基于特定小区的房屋特征, 采用多元回归分析建立小区房屋的比价系数, 以此建立楼盘-房屋二级的批量评估模型。通过一系列检验及校准, 表明模型所构建的批量评估体系能较好反映海甸岛商品住房的特征。在此基础上, 通过样本价格数据采集, 结合批量评估模型可以实现海甸岛全部商品住房价格的批量评估。研究结果可为房地产交易、计税及管理等提供技术支撑。

关键词: 批量评估; GWR模型; 特征价格模型; 多元线性回归; 比价系数

中图分类号: F293.3 文献标识码: A 文章编号: 1674-8859(2023)03-154-05 DOI:10.13991/j.cnki.jem.2023.03.027

Research on Mass Appraisal of Commercial Housing Prices: A Case of Haidian Island in Haikou City

HAN Nianlong^{1,2}, LUO Yan², ZHANG Feifan², WANG Desi², LI Xingqiang²

(1. School of Geography and Tourism, Huizhou University, Huizhou 516007, China;

2. Public Administration College, Hainan University, Haikou 570100, China, E-mail: 450073618@qq.com)

Abstract: To meet the growing demand for real estate evaluation, mass appraisal techniques are gradually replacing traditional single real estate appraisal. This paper took the commercial residential community on Haidian Island in Haikou City, Hainan Province as a research case, and constructed a geographically weighted regression-based (GWR) price model for Haidian Island commercial residential community. The model was built based on the endogenous, locational, and neighborhood characteristics of the commercial residential community by collecting regional point of interest (POI) data, community price data, and housing attribute characteristics data. Meanwhile, based on the housing characteristics of specific neighborhoods, multiple regression analysis was used to establish the price ratio coefficients of neighborhood houses, which was used to establish the two-level mass appraisal model of the building and housing. Through a series of tests and calibration, it is shown that the mass appraisal model can better reflect the characteristics of commercial residence prices on Haidian Island. On this basis, Integrating the sample price data collectiond with the mass appraisal model can realize the mass appraisal of all commercial residential prices in Haidian Island. The research results can provide technical support for real estate transactions, taxation, and management.

Keywords: mass appraisal; GWR model; hedonic price model; multiple linear regression; price ratio coefficient

当前房地产经济活动日益增加, 主要体现在居

民购房、房屋征收补偿及房产出租等需求增多, 因而凸显房地产估价的重要性。批量评估是在传统评估理论基础上, 融合了统计学和地理信息系统等现代技术发展起来的房地产价值评估方法, 较传统评

收稿日期: 2023-01-02.

基金项目: 海南省自然科学基金项目(421RC1034);
海南大学科研启动项目(kyqd(sk)1926).

估方法有诸多优势, 主要体现在评估数量、评估效率提升及评估过程客观等方面^[1]。现阶段我国正稳步推进房地产税的立法及房产税城市试点工作。房产税城市试点是针对整个城市范围的房地产进行征税, 评估的房屋数量巨大, 传统的单宗房地产评估已无法满足现状要求。同时, 当前多个城市为落实“房住不炒”政策, 相继出台二手房交易参考价用于调控房地产市场, 且取得积极成效。这些政策的落实都需要城市整体房地产的评估价格作为支撑依据。

国际税收评估人员协会 (International Association of Assessing Officers, IAAO) 将批量评估定义为利用共同的数据、标准化的方法和统计检验技术评估一组财产确定日期价值的活动^[2]。由于批量评估使用标准化方法, 保障评估过程的一致及公平性, 因此得以广泛应用于房地产税基评估、房地产定价及调控等方面^[3, 4]。早年我国学者聚焦批量评估面临的问题^[5], 结合我国房地产市场现状和国外先进经验, 为批量评估的完善提出对策^[6, 7]。当下, 新技术及方法不断出现及发展, 如从多元回归、支持向量机、随机森林、支持向量机、人工神经网络等机器学习算法^[8~12], 为批量评估奠定了技术基础。同时, 随着房地产市场调控不断加码及房地产税制改革进程的加快, 应用案例不断丰富, 如北京、武汉等城市均已加入批量评估应用行列^[13, 14]。在批量评估中特征价格模型被广泛应用^[15], 传统基于回归分析的特征因素量化忽视了空间效应对房地产价值的影响, 因此引入空间计量模型对于批量评估的精度提高具有重要意义。纪益成等^[16]证明空间误差模型能有效提高批量评估的精度。袁誉萌^[17]研究发现空间自回归模型优于空间误差模型和传统特征价格模型。

上述研究多集中于一、二线的发达城市, 鲜有中小城市的研究, 且多从空间依赖性角度考虑空间效应对房价的影响。海南作为旅游大省, 其房地产兼具旅游地产的特征, 但目前尚无对旅游地产特征及批量评估研究。诸多模型中, 由于地理加权回归 (Geographically Weighted Regression, GWR) 模型能够量化反映房屋价值与其空间特征的相关关系而被广泛应用于房价研究^[18, 19]。因此, 本文以海口市海甸岛为例, 引入GWR模型及特征价格模型, 构建商品住宅小区及其房屋的比价体系的批量评估体系, 以期为房地产市场管理, 房地产调控及房产税改革实施提供数据基础。

1 研究区及数据

1.1 研究区概况

海甸岛位于海口市最北端, 面积约14km², 人口约19.38万人, 处于南渡江的入海口, 与紧邻的新埠岛一起形成南渡江三角洲的中心。海甸岛东临横沟河, 南临海甸溪, 西、北临琼州海峡, 与雷州半岛隔海相望, 海河景观资源丰富。整体环境优越, 且拥有开放性的市级休闲娱乐公园——白沙门公园。海甸岛交通十分便利, 通过和平桥、人民桥及世纪大桥与海口市区的主要道路相连接。教育资源丰富, 岛上有海南省唯一的211大学海南大学, 同时海口市华侨中学、海口市实验中学等学校均坐落于岛上, 海口市人民医院 (三级甲等医院) 建于此地。此外, 鹏辉广场、恒福居商业广场、海甸城等大型商场聚集此地, 生活服务设施完善。综上, 海甸岛是集教育、医疗、居住地以及娱乐休闲服务为一体的综合性功能城市片区 (见图1)。海甸岛上的商品住宅小区有一百多个, 商品住宅近5万套, 区域内房地产交易信息相对集中, 有利于收集数据和进行相关模型考证, 这为批量评估提供了良好样本, 因此本文选择海甸岛作为研究区域。



图1 研究区示意图

1.2 数据处理

本文通过链家、安居客、搜房网及58同城等网站采集海甸岛商品住宅小区价格及属性特征数据, 包含小区名称、建筑年代、容积率、绿化率、挂牌或成交价格等信息, 部分房屋信息通过实地调查, 如向物业管理处及入户调查补充完善。基于高德地图获取POI数据, 其中包含住宅小区、教育用地、三甲医院、便利店和公交站点等数据共计5919条。对获取数据进行初步清理, 主要包括删除重复数据、去除信息缺失或者异常的数据, 以保证基础数据的准确性。经筛选后的海甸岛商品住宅小区共计139个, 且样本数据基本满足正态分布的要求。根据中国指数研究院发布的统计数据, 对案例交易价

格进行时间修正，修正至统一时间时点，保证价格的可比性。

2 模型及算法

本文基于各商品住宅小区的价格因素影响特征，采用GWR模型构建小区的价格模型，建立小区的比价系数。考虑到同一小区内的房屋区位特征和邻里特征往往相同，建筑特征对于房屋价格的影响更为显著。因此，可通过构建房屋特征价格模型来评估各房屋价格，以此构建小区-房屋二级比价系数体系，对海口市海甸岛全部商品住宅实现批量评估。

2.1 GWR 模型

地理加权回归分析技术(GWR)以随着空间位置不同而变化的参数估计结果，量化反映空间数据关系中的异质性或非平稳性特征^[20]。不同于 OLS 全局回归模型，GWR 模型是由 Fotheringham 等^[21]提出的一种空间变系数回归模型，更能刻画出房价影响因素的空间分异规律，其模型具体形式为：

$$y_i = \beta_0(\mu_i, v_i) + \sum_{k=1}^K \beta_k(\mu_i, v_i) x_{ik} + \varepsilon_i \quad (1)$$

式中， (μ_i, v_i) 是 i 的地理区位坐标； $\beta_0(\mu_i, v_i)$ 是常数项； $\beta_k(\mu_i, v_i)$ 表示第 i 个样本点的第 k 个回归参数； ε_i 为随机误差项。

本文中 GWR 模型核类型选择默认的 FIXED (固定距离法)，以便于形成更加平滑的核平面。影响因素对房价的影响程度受带宽影响，确定带宽对于 GWR 模型来说十分重要。为取得最优带宽，本文采用赤池信息准则(AIC)，即 AICc 值最小。

2.2 特征价格模型

房地产异质性较强，具有多方面的不同特征，其价格的形成与这些特征密不可分。特征价格模型(Hedonic Price Model)是分析异质产品差异特征与产品价格间关系的模型^[22]，被广泛应用于房地产价格研究当中。其基本形式如下：

$$f(x) = f(x_1, x_2, x_3, \dots, x_n) + \varepsilon \quad (2)$$

式中， $f(x)$ 为房地产价格； $x_1, x_2, x_3, \dots, x_n$ 为影响房屋价格的属性因子，即特征变量； ε 是随机干扰项。

特征价格模型一般有 3 种表现形式，即线性模型、对数模型和半对数模型。考虑到线性函数形式简单，能对自变量和因变量之间的关系进行清晰的说明，而非线性回归存在着操作复杂，计算成本高等劣势^[23]，故本文选取线性形式构建特征价格模型。

3 批量评估实证

3.1 房价空间分布及自相关分析

基于房价的克里金插值显示(见图 2)，海甸岛房价空间分异特征明显，商品住宅均价高值主要分布在西部、北部及南部，整体上呈现南北高中间低的空间格局。这是由于海甸岛三面环海，尤其是西部及北部区域的海景景观资源丰富，高端别墅住宅区、白沙门公园及高尔夫球场等设施均分布于此。东南部分布有福安上村、广益新村等城中村，区域建筑密度大，商品住房价格相对偏低。南部海甸溪一带环境优美，教育、商业及绿道设施配套完善，也是海甸岛房价的高值区。

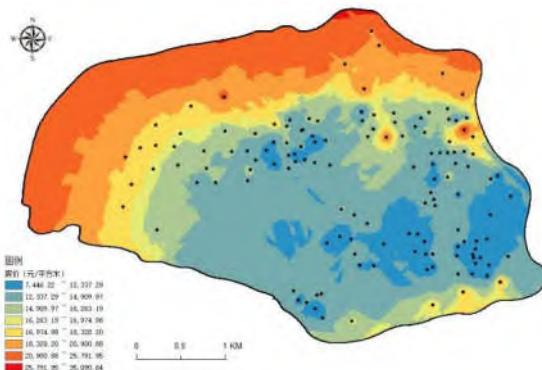


图 2 海甸岛商品房房价空间分异图

对房地产进行空间自相关分析是进行批量评估的基础性工作。通过房价的空间自相关分析发现，全局莫兰指数(Moran's I) 为 0.9329，Z 值和 P 值均通过了检验，表明海甸岛商品楼盘均价存在空间正相关。局部空间自相关分析(Anselin Local Moran's I)结果显示部分楼盘均价存在高低或低高聚集(见图 3)，因此可采用 GWR 模型对其分异性进一步探究。

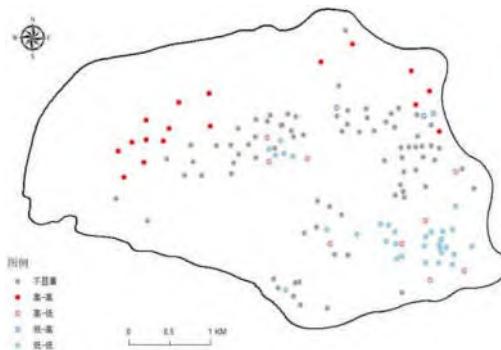


图 3 海甸岛商品房房价 LISA 图

3.2 基于 GWR 模型的小区比价系数构建

将影响小区均价空间分异的因素分为内生因素、区位特征及邻里特征三大类，分别有建筑密度、

房龄、公交车站、主干道、海景、大型商场、商超便利、三甲医院和学校共 9 个影响因素, 建立 GWR 模型的变量集合。根据 GWR 模型的模拟结果得到调整后的 R^2 为 0.525。 R^2 反映模型拟合优度, 越接近 1 表示拟合得越好^[19, 23]。

GWR 的结果是各影响因素的回归系数对房价的影响有正有负, 汇总如表 1 所示。正向影响由大到小的因素依次为学校、商超便利、公交车站和主干道, 负向影响由大到小的因素依次为房龄、建筑密度、大型商场、海景和医院。其中, 主干道和医院对房价的影响有正有负, 说明在不同的空间上, 影响因素的作用具有方向性差异。房龄的标准差最大, 学校其次, 说明这些因素在不同空间上影响程度具有显著差异, 具有明显的空间不稳定性。基于 GWR 模型的回归系数结果, 可构建海甸岛商品住宅小区的比价系数。

表 1 地理加权回归模型结果

	变量	最小值	最大值	平均值	标准差	变量描述
内生 因素	建筑密度	-2.17	-3.76	-2.55	0.42	各小区所在地区的建筑密度/%
	房龄	-217.26	-262.92	-235.65	9.61	房屋竣工至今期限/年
区位 特征	公交站	7.10	2.51	4.18	1.28	距小区最近的公交车站距离/m
	主干道	2.16	-0.57	0.74	0.77	距小区最近的主干道距离/m
	海景	-0.55	-1.01	-0.75	0.10	距小区最近的海景距离/m
邻里 特征	大型商场	-0.45	-1.07	-0.75	0.17	距小区最近的大型商场距离/m
	商超便利	5.26	3.07	4.19	0.52	距小区最近的商超便利距离/m
	三甲医院	0.49	-1.36	-0.54	0.45	距小区最近的三甲医院距离/m
	学校	8.63	0.74	5.79	2.12	距小区最近学校距离/m

3.3 基于特征价格模型的房屋比价系数构建

在房屋层级, 本文通过构建特征价格模型及多元回归分析获取房屋间的比价系数。由于海甸岛海悦国际小区成交样本量较充足, 且房屋价值居中, 因此选取其作为案例。该小区位于海南省海口市美兰区海甸五西路, 东临海南大学北门, 西临别墅区, 南临海南大学校园, 北临海甸五西路, 占地面积为 8304.37m², 容积率为 3.86, 共有住宅 384 套。结合小区实际情况选取朝向、景观、噪音程度、面积、楼层和装修程度 6 个特征变量构建房屋特征属性指标体系如表 2 所示。

以小区的房屋价格为因变量, 朝向、景观、噪

表 2 住宅特征因素的量化

特征价格 因素	变量 符号	划分方法	变量变化
朝向	DRE	综合评价打分法	南北: 5 分; 东南: 4 分; 西南: 3 分; 东北: 2 分; 西北: 1 分
景观	VIE	综合评价打分法	海景及世纪大桥可见: 4 分; 可见海景及街景: 3 分; 仅可见世纪大桥: 2 分; 仅可见街景: 1 分
噪音程度	NOI	虚拟变量法	嘈杂: 3; 较嘈杂: 2; 较安静: 1; 安静: 0
面积	ARE	实测数据法	按搜集的数据值填写
楼层	FLO	实测数据法& 虚拟变量法	1~6 层: 1; 7~12 层: 2; 13~18 层: 3
装修程度	DEC	分等级赋值法	精装修: 4; 高档装修: 3; 中档装修: 2; 简易装修: 1

音程度、面积、楼层和装修程度作为自变量, 采用 SPSS 软件进行多元线性回归分析。分析前需要对自变量和因变量做相关性分析, 以排除混杂干扰项。相关性检验结果表明朝向与因变量之间的皮尔逊系数小于 0.3, 呈现极弱相关, 故将其排除。回归分析中结果显示景观变量存在多重共线性, 因此也将其排除。最终得到 R^2 为 0.799, 调整后的 R^2 为 0.780, 表明模型拟合度较好, 最终的回归模型如下:

$$P = -340 \times NOI - 10 \times ARE - 70 \times FLO + 50 \times DEC + 16920 \quad (3)$$

上式显示, 噪音程度、面积与楼层对房价均产生负向影响。噪音对房价影响最明显, 噪音每上升一级, 房价下降 340 元/m²; 其次是楼层, 每升高一层, 房价下降 70 元/m²; 面积因素对房价呈现偏弱负向影响, 说明该小区的小户型受一定程度的追捧。装修程度对房价呈正向影响, 因为装修程度涉及居住体验, 装修程度提升一个档次, 房价上涨 50 元/m²。通过将房屋的属性值代入回归方程式(3)即可获得小区的所有房屋的评估价, 基于此可构建小区内所有房屋的比价关系。采用 GWR 模型及特征价格模型定量获取商品住房楼盘及楼盘内房屋的比价系数后, 在基于小区-房屋的二级比价关系基础上, 实现了海甸岛商品住房批量评估模型的构建。在实际应用中, 可以通过定期采集研究区内的住房成交案例, 基于构建的批量评估模型, 实现区域内所有商品房屋价格的定期批量评估与监测。

4 结语

本文将多元回归及 GWR 模型应用于海甸岛商品住宅的小区及房屋的比价系数构建。研究结果表明, 海甸岛小区房价存在空间正相关, GWR 模型能较好地揭示区域内住宅小区的空间异质性; 同时,

基于特征价格模型构建住宅小区的房屋特征，并结合多元线性回归分析对房屋特征进行了量化。通过一系列的模型检验及校准，表明两者结合能较好构建商品住房的批量评估体系。根据模型构建的小区-房屋二级比价系数，通过设定相关标准房，采集实际成交案例，实现海甸岛全部商品住房的批量评估。在此基础上，进一步可以拓宽至全市乃至更大范围的商品住房评估。

本文不足之处在于，仅以特定楼盘为例评估房屋比价系数，用以构建研究区批量评估模型。由于房地产的异质性，各楼盘小区的特征因素并不统一，因此需要进一步挖掘各楼盘的特征变量以提升房屋比价系数可靠性；在模型方法方面，需要进一步探索人工神经网络及机器学习等算法以提升评估模型的精度。本文通过挖掘海口市海甸岛的房地产特征，探索具有旅游地产特征的商品住房批量评估应用，可为区域房地产计税评估、房地产出租交易定价及房地产市场调控等方面提供数据及技术支撑。

参考文献：

- [1] 徐爱农. 国外物业税税基评估实践与启示[J]. 经济问题探索, 2011 (8): 101-104.
- [2] 纪益成, 傅传锐. 批量评估: 从价税的税基评估方法[J]. 中国资产评估, 2005 (11): 5-9.
- [3] 焦嫣然. 房地产计税依据确定方法的法律分析[J]. 税务研究, 2019 (4): 43-48.
- [4] 郑江晖. 用聚类与判别分析的房地产特征因素识别与分类方法研究[J]. 福建论坛(人文社会科学版), 2014 (10): 36-41.
- [5] 王瑞玲, 邓晓. 我国房产税征收中批量评估所面临的困难和对策研究[J]. 生产力研究, 2011 (5): 144-145.
- [6] 陈汉明. 房地产税基批量评估技术路径选择研究[J]. 学术论坛, 2012, 35 (7): 158-161.
- [7] 纪益成, 王诚军, 傅传锐. 国外 AVM 技术在批量评估中的应用[J]. 中国资产评估, 2006 (3): 13-18.
- [8] 张思雪, 林汉川, 吕臣, 等. 住宅房地产批量评估模型的构建[J]. 统计与决策, 2016 (3): 69-72.
- [9] Torres-Pruñonosa J, García-Estevez P, Prado-Román C. Artificial neural network, quantile and semi-log regression modelling of mass appraisal in housing[J]. Mathematics, 2021, 9 (7): 783.
- [10] Sefa B S, Murat Y H. Comparison of different machine learning models for mass appraisal of real estate[J]. Survey Review, 2023, 55 (388): 32-43.
- [11] Dimopoulos T, Bakas N. Sensitivity analysis of machine learning models for the mass appraisal of real estate. case study of residential units in nicosia, cyprus[J]. Remote Sensing, 2019, 11 (24): 3047.
- [12] 陈诗沁, 王洪伟. 基于机器学习的房地产批量评估模型[J]. 统计与决策, 2020, 36 (9): 181-185.
- [13] 刘洪玉, 杨振鹏. 基于主体变量的住房价格批量评估[J]. 统计与决策, 2012 (3): 62-66.
- [14] 吴桐仪, 邓大松. 基于粒子群优化算法的房地产税基评估: 以武汉市 10 个区的数据为例[J]. 经济与管理研究, 2015, 36 (2): 136-144.
- [15] Hu L, He S, Han Z, et al. Monitoring housing rental prices based on social media: an integrated approach of machine-learning algorithms and hedonic modeling to inform equitable housing policies[J]. Land Use Policy, 2019, 82: 657-673.
- [16] 纪益成, 马铮, 张然, 等. 基于空间误差模型的房地产批量评估研究[J]. 建筑经济, 2015, 36 (2): 62-67.
- [17] 袁誉萌. 基于空间特征的房产税税基评估方法研究[D]. 济南: 山东大学, 2020.
- [18] 唐红涛, 刘亦鹏, 吴忠才. 基于 POI 数据的房价影响因素空间异质性分析: 以长沙市为例[J]. 城市问题, 2021 (2): 95-103.
- [19] 张非凡, 韩念龙, 纪旭, 等. 基于 GWR 模型的房价空间影响因素研究: 以海口市海甸岛为例[J]. 地理信息世界, 2022, 29 (1): 52-57.
- [20] 卢宾宾, 葛咏, 秦昆, 等. 地理加权回归分析技术综述[J]. 武汉大学学报(信息科学版), 2020, 45 (9): 1356-1366.
- [21] Fotheringham A S, Oshan T M. Geographically weighted regression and multicollinearity: dispelling the myth[J]. Journal of Geographical Systems, 2016, 18 (4): 303-329.
- [22] 黄醇醇, 王晓文, 李琳娜. 城市轨道交通对沿线住宅价格的时空效应: 以福州地铁 1 号线为例[J]. 地理研究, 2021, 40 (10): 2808-2822.
- [23] Anselin L. Under the hood Issues in the specification and interpretation of spatial regression models[J]. Agricultural Economics, 2002, 27: 247-267.
- [24] 何鑫, 田丽慧, 楚尔鸣. 人口流动视角下中国房价波动的空间异质性[J]. 人口与经济, 2017 (6): 43-57.

作者简介:

- 韩念龙 (1983-), 男, 博士, 高级工程师, 研究方向: 土地资源管理;
- 罗燕 (2000-), 女, 本科, 研究方向: 土地资源管理;
- 张非凡 (1999-), 女, 本科, 研究方向: 土地资源管理;
- 王德思 (1998-), 女, 本科, 研究方向: 土地资源管理;
- 黎兴强 (1972-), 通信作者, 男, 硕士, 教授, 研究方向: 旅游开发规划与管理。