

RESEARCH

Open Access



Impact of rail transit station proximity to commercial property prices: utilizing big data in urban real estate

Mohammed Ali Berawi^{1*}, Perdana Miraj^{2,3} , Gunawan Saroji³ and Mustika Sari³

*Correspondence:
maberawi@eng.ui.ac.id
¹ Department of Civil
Engineering, Universitas
Indonesia, Depok 16424,
Indonesia
Full list of author information
is available at the end of the
article

Abstract

Urban transport investments have contributed to the exponential increase of value from land and properties around the built infrastructure. Although literature had shown evidence of value uplift from the residential property due to transit development, little is known about the impact on commercial property values. This paper aims to evaluate the impact of rail transit proximity to the commercial property market by taking into account the pre-operation of the LRT project in Jakarta, Indonesia. This study utilizes a big data approach to accelerate research data processing by employing data mining method, as well as geographical information system (GIS) and hedonic price modeling (HPM) to investigate the property prices to formulate empirical evidence for the research. The results show opposite evidence compared to previous studies which argued that accessibility may contribute to the property price when closer to the transit station. The findings indicate that proximity to rail transit has an insignificant impact on the commercial property price compared to other variables such as building size, number of rooms, location, and hospitals. A comparison of the findings between this research with other studies was discussed and recommendation for further research development was proposed.

Keywords: Commercial property, Hedonic price model, Jakarta indonesia, Proximity premium, Rail transit, Value capture

Introduction

Investment in rail transit does not only offer benefits to the public interest, such as increasing physical activity [1], releasing traffic congestion [2, 3], reducing emission [4], and expanding consumer amenities [5], but it also enables property owners nearby transit stations to capitalize their assets for greater benefits due to the land value uplift as the effect of accessibility improvement. The increases in land value indicate the positive impact of rail transit to the local economy, people's quality of life, and the overall property market. The capitalization of real estate market is a combination of structural characteristics, surrounding environment, physical attributes, and accessibility which created the market price of an area [6, 7].

Research topics related to the relationship between rail transit and property market have been extensively investigated in the past decades and have taken different types of rail transit into account, including mass rapid transit (MRT), light rail transit (LRT), and heavy rail. Debrezion, Pels, & Rietveld [8] and Mohammad, Graham, Melo, & Anderson [9] have contributed an extensive literature review that covers approximately 200 publications on the topic of rail transit and property values, while some recent studies have also discussed this topic in various cities worldwide such as Missouri, USA [10], Hamburg, Germany [11], Beijing, China [12], and Queensland, Australia [13]. However, most of them focused on the residential properties, and only a few studies acknowledged commercial or offices property values, such as those located in Arizona, USA [14], Australia [15], Seoul, Korea [16], Dubai, UAE [17]. Moreover, none of them elaborates evidence of commercial property prices due to transit proximity from developing countries such as Indonesia or Southeast Asian countries, whereas commercial properties are believed as a significant function to attract business activities that can accelerate the economic growth of a city.

Currently, most of the property sale techniques are shifting from newspapers or offline brokers to online advertisements on both local and international portals offering services such as selling and renting numerous types of buildings and lands for interested buyers. Large databases of property are advertised daily with useful information and various types of listing on these websites. This massive data source has been utilized by social scientists to evaluate urban vacation rentals or rental housing markets using web scraping techniques to understand land use patterns and market activities [18, 19]. However, there is limited evidence that further investigates this technique into the transport-property market issue by considering online advertisement or marketplace listings. Although primary and secondary data collection are still playing a significant role in understanding the real data pattern, the utilization of data mining and other big data technical advances in handling and analyzing wide range data dimensions may serve as an alternative approach that can minimize exhaustive primary survey process as well as to accelerate the investigation to large-scale online data sources.

The research objectives of this study are twofold. The paper expects to explore website data that are being under-utilized by employing web scraping technique to property listings data gathered from online real estate marketplaces. Subsequently, the article attempts to evaluate the impact of rail transit proximity to the commercial property market by taking the pre-operation of the LRT project in Jakarta, Indonesia, as the study case. Understanding the potential effect of value uplift from rail transit to property value will not only support local governments to establish appropriate policy and regulation but also decision-makers both in domestic and international to dealing with land value mechanism and rail transit investment. The findings are also expected to make contribution to the body of knowledge of urban planning and regional development by proposing arguments on the benefits of adopting digital technologies such as big data to process and analyze a large urban dataset for a greater purpose to the academics and researchers.

Literature review

Rail transit gauges a significant impact on the increased value of land that yields benefits for the property owners nearby transit stations. The stations are believed as the core of transport network that offers better accessibility and mobility to connect destinations

and commuters within walking distance [20]. Numerous research proposes buffer areas in the radial distance to measure travel time needed by users to reach their point of interest [21, 22]. The proximity of these properties to the transit station is debatable among researchers and academics. Some of them measured the acceptable proximity based on walking distance ranging from 300, 600, 800 m, and even less than 1 km depends on the building typology (residential, offices, or commercials) or by walking time – mostly less than 5 min- walk. In general, there is no conclusive evidence to determine the best measurement of proximity but heavily relies on the characteristics of cities and citizens when accessing their destination. A compact city may only require 300–600 m to access the building from the station, while other cities may require beyond that distance due to the opposite local condition.

The impact of rail transit proximity to commercial property prices has been investigated despite remains very limited. Cervero and Duncan [22] found a significant capitalization of commercial land located within walking distance from light rail transit (LRT) and commuter rail stations in Santa Clara County, California, for about 23% and 120% respectively. Although the commercial land gained a positive impact from transit proximity, the result may differ from one city to another [16]. A study in Dubai, United Arab Emirates, showed that commercial properties gained the highest effect within 700–900 m from the metro station [17]. However, a similar study in Wuhan, China, showed commercial property closer to the rail station, approximately 300 m, gained the highest value uplift [23]. This study is confirmed by a recent publication from Phoenix, Arizona, which reported property price might increase to 1.8 times higher when the commercial property is located within the 300 m of LRT station [14]. Few studies presented the negative impact of rail station to property [24, 25], but overall, accessibility positively correlates with the increasing value of a property. Table 1 presents a summary of selected empirical studies on the impact of rail transit on commercial properties.

There are some reasons that might be the main challenge as to why research on the effect of rail accessibility on the commercial market from developing countries remains limited, which include the lack of investment from both state and private investors, and scattered and lack of reliable data for analysis [26]. Although research in rail transit and property have emerged in recent years from developing countries such as Malaysia [27], Turkey [28], and Thailand [26], none of them discussed the impact of rail proximity to the commercial market nearby. Therefore, research regarding this issue in the context of non-developed countries may contribute to the gap of knowledge and lesson learned for countries currently developing their railway sector worldwide.

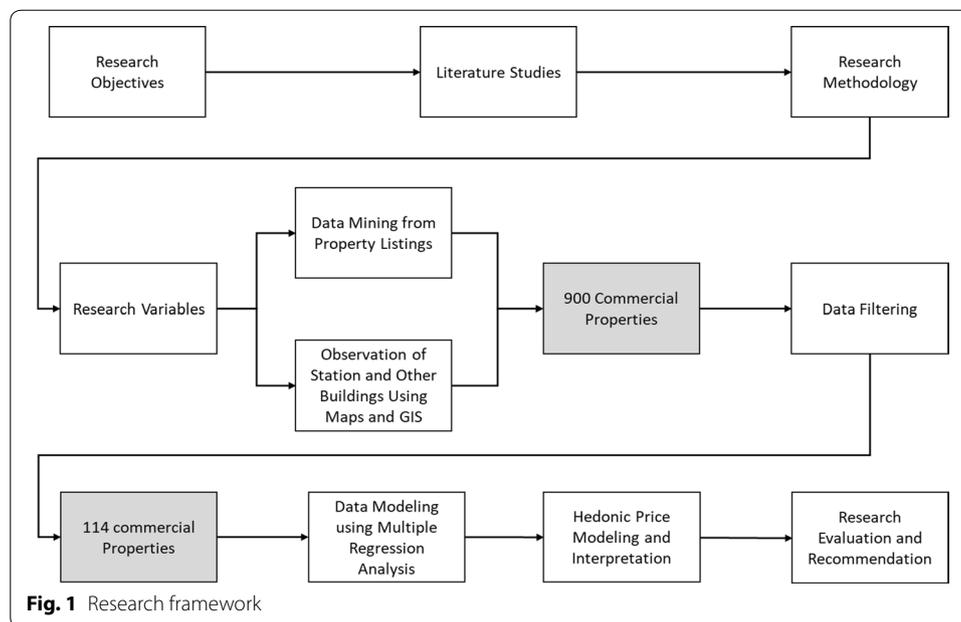
In contrast with residential properties that are easier to extract either through sales prices or rental rates, non-residential properties are still facing difficulties in obtaining data to measure property values [8, 29]. Some challenges in making the data in non-residential properties are fewer than residential properties in terms of quality and quantity. Unlike residential properties which experience a high turnover for listings, the sales of non-residential properties are rather limited. The real estate association is believed to have this kind of information as they gather sales of each member in order to expose the annual property landscape, particularly in developing countries such as Indonesia. However, it is not easy to obtain these data due to information confidentiality or when available, the records release on the incomplete manner for analysis.

Table 1 Selected studies on the impact of rail transit on commercial properties

Authors	Location	Rail type	Measure of property value	Findings
Cervero and Duncan [22]	Santa Clara County, USA	LRT and commuter rail	Hedonic price	LRT station contributes to the increase of commercial land values on average by more than US\$ 4
Kim and Zhang [16]	Seoul, Korea	MRT	Hedonic price	The transit station located near CBD has a premium of US\$7.54/meter
Mohammad, Graham, & Melo [17]	Dubai, UAE	MRT	Hedonic price and difference – in – difference (DID)	The impact of metro on the value of commercials within 701 to 900 m is 76%
Zhang and Xu [23]	Wuhan, China	MRT	Hedonic price	Commercial property gained the highest value compared to other types of properties concentrate of 300 m from transit station about 14.7%
Seo, Salon, Kuby, & Golub [14]	Arizona, USA	LRT	Hedonic price	Light rail stations have a positive impact on commercial property values

The limitation to measure the values of the property made academics and researchers elaborate on the alternative sources for data scarcity by taking into account advertising/asking price [29, 30]. These types of price—despite limited in numbers, are advertised in various means of media such as magazines that can be easily accessed by potential buyers or tenants in real estate marketplaces. This measurement strategy might experience a gap of price as the asking price did not express contractual price and was believed as less accurate to show the equilibrium price on the real estate market. However, some researchers have shown that asking price also can be used for property analysis and market prediction [31, 32] and therefore employed in many property-related studies.

Asking price using real estate marketplaces is categorized as a part of volunteered geographic information (VGI) uploaded by the users in voluntary-basis tagged with geolocation [18]. The asking price using this approach is argued as one of the sources of big data characterized by its large samples with a variety of contents both in a structured and unstructured manner, and the speed of data processing [33]. These types of datasets may promote better planning in urban and transportation areas by leaving subjectivity in data collection from personal interests. Big data is argued to have a huge potential in reshaping and directing social science research towards digitalization [34]. Research dealing with a large dataset such as property market research and development is predicted to become the first that gained the impact of the fast-growing sources of big data. As rail transit in developing countries is experiencing a massive development with more urban rails are currently under construction, the commercial property market continues to grow in urban areas, and internet-based resources are tremendously rising, this research attempts to provide empirical evidence and body of knowledge in transport



policy and urban studies by taking into account large scale listings of the property market.

Methodology

This research conducted a three-stage approach from data mining, data cleaning filtering, and data modeling by taking into account web scraping, geographical information system (GIS), and hedonic price modeling (HPM) to investigate the property market around the transit station. The research framework can be seen in Fig. 1. The researchers collected more than 900 property listings from two online real estate marketplaces (Lamudi.co.id and Rumah.com) in Indonesia from September 2018 to February 2019. It is crucial to note that the property listings in both marketplaces are the advertised price, which then may be different from the transaction price [35]. This price suggests closer to the price in legal contracts compared to a door-to-door survey to the targeted properties or to the sale value of the tax object (*Nilai Jual Objek Pajak/NJOP*) issued by the government. The latter has been used to determine property prices in developed countries, while in Indonesia, the property price tends to surpass up to three times of NJOP depends on the location, property condition, and many other considerations. Nonetheless, real estate marketplaces offer invaluable data sources for research development in the property and housing sectors.

Data mining

Web scraping has overcome the difficulties in collecting huge numbers of database for research on various sectors such as transportation, housing, tourism, and many others. There are many existing tools that can be used for data mining from web pages and offer a customized solution for novice users. These web scraping tools enable users to extract information, transform databases, structure the data generated from web pages

Table 2 Variables measurement

Variable types	Variables	Variable quantification	Data source
Dependent variable	Property price	US\$/m ²	Real estate marketplaces
Independent variables	Structural attributes		
	Number of Rooms	Total number of rooms in property (Unit)	Real estate marketplaces
	Building Size	Total area of a property (m ²)	Real estate marketplaces
	Accessibility attributes		
	Dist. to station	Distance between the property to station (m)	GIS
	Neighborhood attributes		
	Schools	If there is a school within 1 km, the variable is 1, otherwise 0	GIS
	University	If there is a university within 1 km, the variable is 1, otherwise 0	GIS
	Hospital	If there is a hospital within 1 km, the variable is 1, otherwise 0	GIS
	CBD	If there is a CBD within 1 km, the variable is 1, otherwise 0	GIS
	Location attributes		
	Karet Kuningan	Location of transit station that being investigated	GIS
	Dukuh Atas		
Rasuna Said			
Cawang			
Kuningan			
Cikoko			
Ciliwung			

automatically, and store in both clouds and local databases depends on user interest [36]. Some tools were created using a programming language such as R or python for statistical computing and graphics [37]. In this study, a Parsehub – a web scraping tool to extract data in property market development, was used for web-crawling purposes. This tool offers ease of accessibility for non-technical users, flexible to handle complicated websites and easy to use interface for data retrieval. First, the targeted website is selected and inserted into the Parsehub. The tool then used to select attributes of the property, including property price, building size, and the number of bedrooms. Parsehub also enables users to obtain the location of the property by extracting latitude and longitude points by activating coordinate inspection on the website page prior to submission in parsehub. Finally, the web-scrapers save these structured data into a certain format based on CSV or JSON depends on the need for research purpose and further analysis. Web scraping method extracts information from the property website regarding (1) property price, (2) building size, (3) Number of rooms, and (4) geolocation (X, Y coordinates).

Property listings and its attributes generated from previous data mining were added with accessibility and surrounding attributes (see Table 2), including the distance of buildings (central business district/CBD, hospital, school, and university) to the transit station by taking the point of coordinate from google maps. Each building's geolocation, including its latitude and longitude that were retrieved to measure the building's

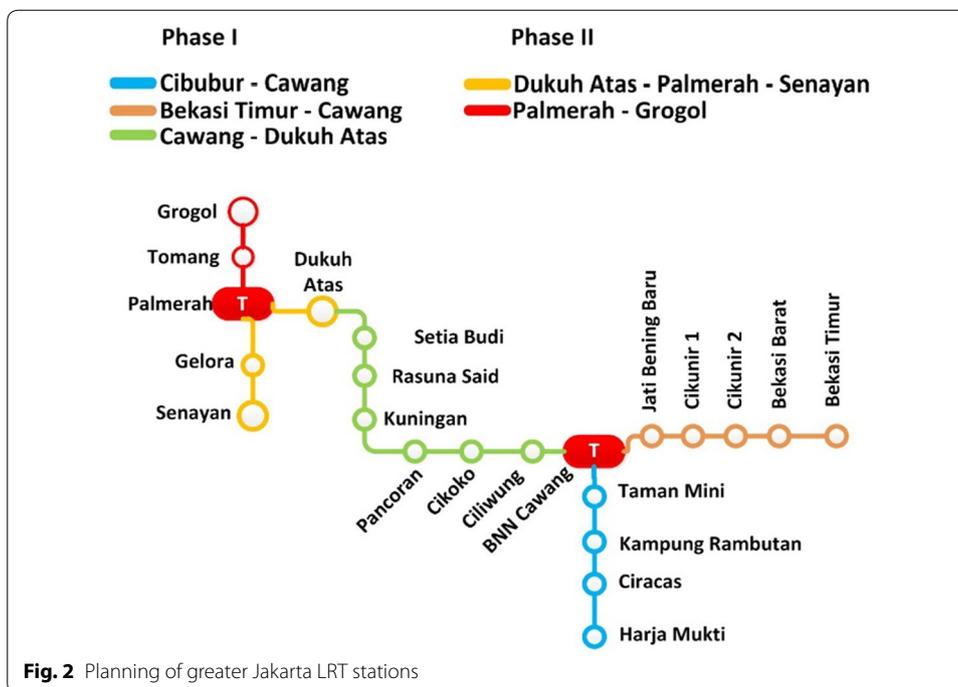


Fig. 2 Planning of greater Jakarta LRT stations

distance to the nearest transit station. Four building typologies were selected due to their capability to attract a significant number of users taking transit mode rather than automobiles for mobility [23, 38, 39]. All of these data will then be plotted in the Geographic information system (GIS).

GIS is a system that allows users to extract, manage, analyze, and assist in decision making based on spatial data. This research used ArcGIS software to evaluate geographic information and interpret factors in determining property prices. The software only included property data located approximately 1 km from the transit station, a distance that provides walkable access for citizens reaching targeted destinations from a transit station [38, 40]. Seven stations of upcoming light rail transit in Jakarta (see Fig. 2) namely Dukuh Atas, Rasuna Said, Karet Kuningan, Kuningan, Cawang, Cikoko Station, and Ciliwung were selected for analysis. By taking into account transit stations, GIS enabled researchers to retain properties close to the transit proximity.

Data cleaning and filtering

When conducting the data mining process from property websites, it is common that the raw data are in a disorderly manner. Initially, more than 900 properties were extracted for potential analysis. Although the websites suggest structured text entry to create listings for users, sometimes these listings databases are incomplete, irrational, double, or unstructured. Therefore, data cleaning and data filtering were required to standardize the original dataset for subsequent analysis.

In the first stage, duplicate listings were identified from the initial data mining process. Since the real estate marketplaces do not prohibit users with the same ID to resubmit their listings multiple times to restore their listings into top search results, therefore, a similar property name will be immediately deleted when it has appeared more than

Table 3 One-sample Kolmogorov–Smirnov test

		Unstandardized residual
N		95
Normal parameters ^{a,b}	Mean	0.0000000
	Std. deviation	0.10808902
Most extreme differences	Absolute	0.074
	Positive	0.074
	Negative	− 0.067
Test statistic		0.074
Asymp. Sig. (2-tailed)		0.200 ^{c,d}

^a Test distribution is normal

^b Calculated from data

^c Lilliefors significance correction

^d This is a lower bound of the true significance

once. Afterward, research team members evaluated listings with unique values regarding property price, building size, and the number of rooms. The marketplaces allow property agents to advertise the same property with different prices or similar depends on their agreement with the owner. Therefore, it is crucial to look upon the similarity of these values and exclude them for analysis.

From initial data mining with more than 900 properties, the research team members maintained 114 listings. Unclear listings that may include irrational values, spam, and incomplete submission were dropped using data filtering in excel. Last, the listings with explicit longitude and latitude of the property were maintained. Listings with coordinates are argued to offer higher credibility when the given location is accurate [41, 42].

Prior to conduct statistical analysis to form a regression formula, 114 listings were evaluated using a normality test. This cleansing process aims to exclude abnormal data and repair data outliers for better analysis. This research used Kolmogorov–Smirnov due to its practicality and interpretation. Residual value from the regression of dependent and independent variables processed in the Kolmogorov–Smirnov test. Property price becomes the dependent variable, while independent variables consist of distance to station, number of rooms, building size, neighborhood attributes, and location attributes. The four earlier variables are in the form of log normal. The sample has a normal distribution (see Table 3) when sig. (2-tailed) shows a value higher than 0.05 and vice versa. Subsequently, the final listings included for analysis are 95 listings.

Data and modelling

This research adopts the hedonic price modeling (HPM) method to investigate the effects of property values and measure how property values in regard to rail station proximity. This method has been extensively used by academics and researchers worldwide and in diverse sectors. HPM can be used to measure the effect of attributes that affect the overall transaction price. Multiple regression analysis allows property values, among other variables, to be determined based on structural characteristics, neighborhood characteristics, accessibility, and land-use types.

HPM can be used in different types of functional forms, including linear, log–log, and semi-log. The linear form assumed the effect of property price in the constant term in regard to the distance to the rail station. Unlike semi-log, which only transformed dependent variables, log–log form converts all variables into the logarithmic model. Log–log form predicts a constant elasticity of property price to the distance to rail station [23]. The theory does not guide how to select the functional form of HPM. However, some studies have attempted to shed light on the benefits and limitations of each form [43], however, there is no evidence showing that one form is dominant over the others [44].

In general, the specification of the model depends on the relationship between the dependent and the independent variables. The selection of functional form suggests following the data availability, the nature of the analysis, and case studies to generate optimum results [45]. This research used the semi-log model due to its practicality in showing the price elasticity of associated characteristics and its contribution to the property prices. The models are estimated based on the ordinary least square (OLS) method to estimate independent variables and are used for subsequent regression analysis. The model denotes the following formula.

$$Y = \alpha + \beta_1.x_1 + \beta_2.x_2 + \beta_3.x_3 + \beta_4.x_4 + \dots + e \quad (1)$$

where,

Y = Dependent variable (Property prices).

β = Regression Coefficient.

x = Independent Variables.

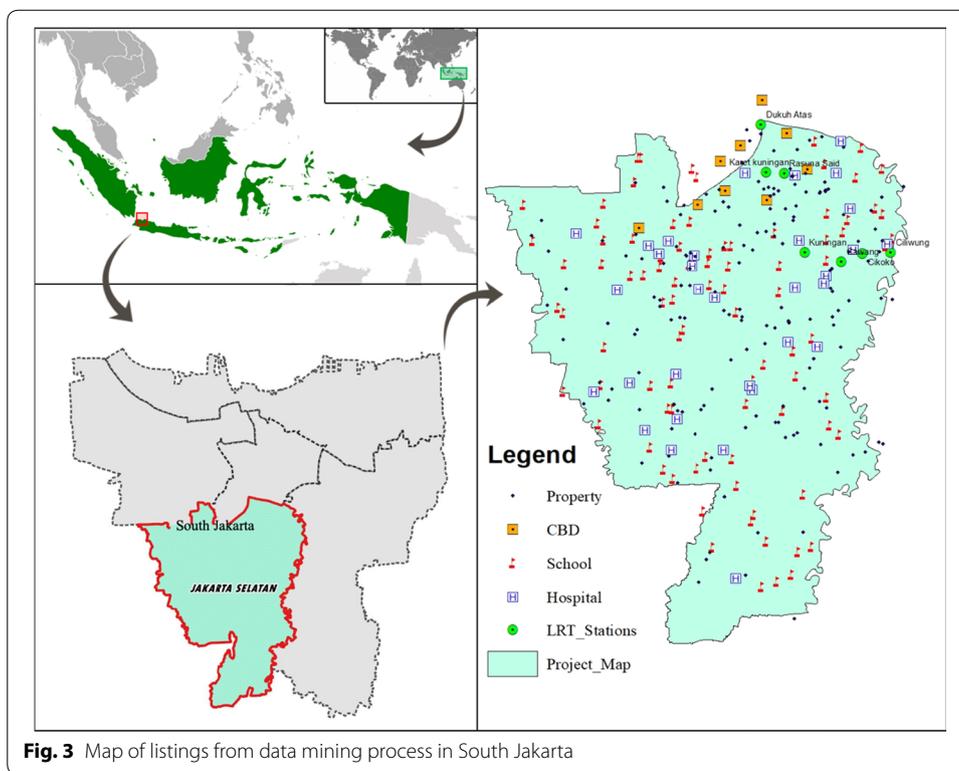
e = Error value (5%).

This research uses two data sources to generate values for each identified variable in the HPM models. The first data source was generated from real-estate marketplaces operated in Indonesia obtained through the data mining process using Parsehub. This data consists of property price (US\$), building size (square meter), and the number of rooms (unit). The second data source regarding the distance, buildings, and location is obtained from a combination of google maps and GIS. Both have been thoroughly discussed in the previous section.

Result and analysis

Geographical distribution of property

Urban planners and practitioners require data for analysis from the local scale that represents the population and geographical diversity. Instead of using manual questionnaire surveys, extracting data from real estate marketplaces is believed to gain similar results with limited effort use during the research data collection. This data provides useful indicators depends on the interest of the researchers regarding components for evaluation. However, this mapping only shows the property location without taking into account the pattern of the property price. This mapping also can not directly specify the location of the property based on the street or transit proximity due to the limitation of data crawling and real estate marketplaces availability. The map in Fig. 3 shows listings in Jakarta Selatan – an administrative city in the capital city of Indonesia, where seven investigated stations in this research are located.



The map reveals not only property location but also other building typologies such as CBD, high school, and hospital represented by different symbols. The maintained 114 listings located within 1 km radius from the transit station are shown in Fig. 4.

Based on the filtered dataset, descriptive analysis for the dependent variable and independent variables are shown in Table 4, including the minimum, maximum, mean, and standard deviation. Property prices extracted from data mining are expressed in Rupiah and then converted to US\$.¹ The mean value of property price is US\$ 3666.86 per square meter and the mean distance of the property to rail station is 604.85 m. The highest property price is US\$ 9578.54, located 800 m from the transit station, while the lowest property price is US\$ 1179.67, located 787 m from the station. Based on these findings, it is argued that proximity to the station is not the only factor that contributes to the property price elasticity near the transit station. Other factors that may affect the price include the number of rooms, building size, and location of the property.

Effect of proximity to transit station on property values

A correlation test was first performed to determine the degree of correlation between independent variables and the dependent variable. A Pearson correlation was used due to the nature of data using parametric (nominal and scale). Kendall or Spearman test is proposed for analysis if non-parametric data was used. The correlation coefficient may vary from - 1 to + 1. A value close to - 1 or + 1 indicates a strong relationship between

¹ 1 US\$ = Rp. 14,500.

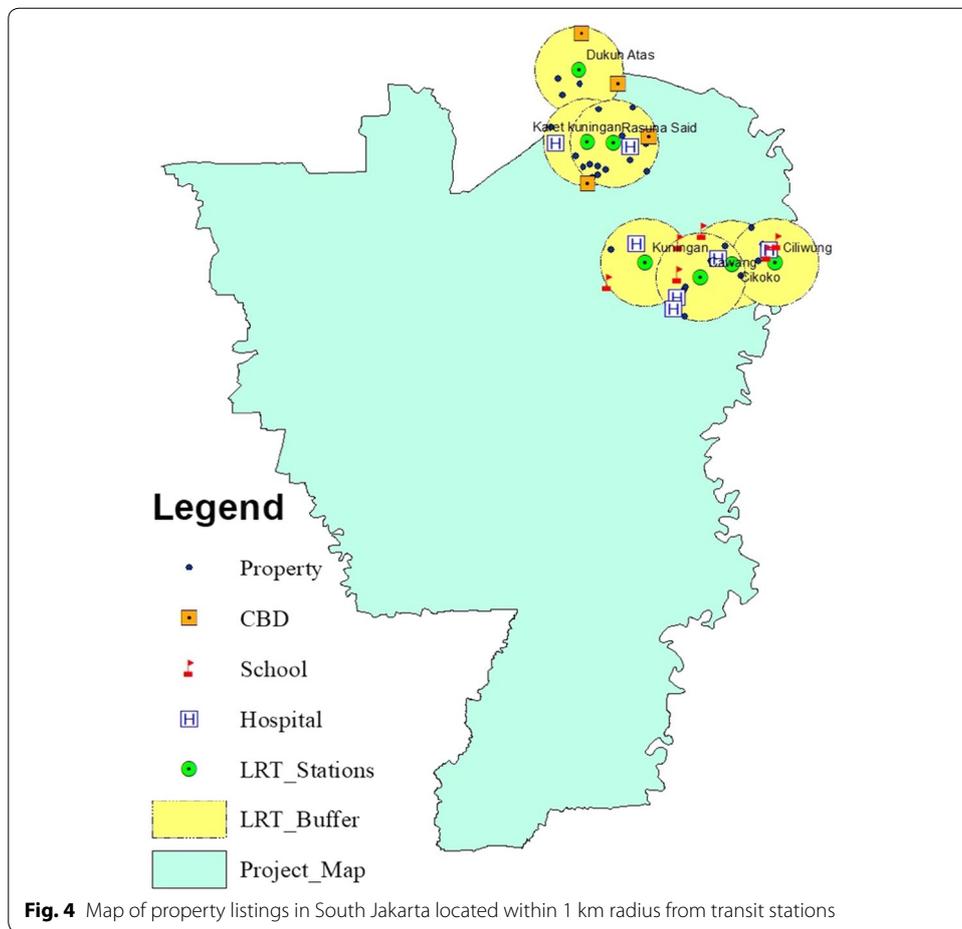


Table 4 Sample descriptive statistics

Variables	N	Min	Max	Mean	Std. deviation
Property price (US\$)	95	1179.67	9578.54	3666.86	1545.25
Number of rooms (unit)	95	1.00	30.00	2.22	3.84
Building size (square meter)	95	100.00	3380.00	967.39	708.39
Karet Kuningan	95	0.00	1.00	0.43	0.50
Dukuh Atas	95	0.00	1.00	0.22	0.42
Rasuna Said	95	0.00	1.00	0.18	0.39
Cawang	95	0.00	1.00	0.04	0.20
Kuningan	95	0.00	1.00	0.01	0.10
Cikoko	95	0.00	1.00	0.10	0.31
Ciliwung	95	0.00	1.00	0.01	0.10
Dist. to station (m)	95	176.90	982.80	604.85	236.30
Schools	95	0.00	1.00	0.35	0.48
University	95	0.00	1.00	0.45	0.50
Hospital	95	0.00	1.00	0.78	0.42
CBD	95	0.00	1.00	0.83	0.38

Table 5 Correlation test between dependent and independent variables

	Ln Price	Sig. (2-tailed)	N
Ln Distance	− 0.091	0.380	95
Ln NRooms	− 0.046	0.659	95
Ln BSize	0.951**	0.000	95
Karet Kuningan	0.026	0.805	95
Dukuh Atas	0.212*	0.039	95
Rasuna Said	− 0.216*	0.035	95
Cawang	− 0.202*	0.049	95
Kuningan	0.135	0.191	95
Cikoko	− 0.008	0.936	95
Ciliwung	0.113	0.276	95
Schools	− 0.212*	0.039	95
University	− 0.044	0.675	95
Hospital	0.212*	0.039	95
CBD	0.048	0.646	95

* Correlation is significant at the 0.05 level (2-tailed)

** Correlation is significant at the 0.01 level (2-tailed)

the two variables and a value close to 0 indicates a weak relationship. On the other hand, positive (+) and negative (-) signs suggest the direction of the relationship between the two variables. If the value is + (positive), then the two variables have a direct relationship. The significance of the variable can be seen from Sig. (2-tailed) when the value lower than 0.05.

Among three variables (distance, number of rooms, and building size), as shown in Table 5, the variable of building size shows a correlation coefficient close to 1 and sig. A 2-tailed value lower than 0.05. A positive mark means that a larger building size contributes to a higher property price. In terms of location, three of the stations, namely Rasuna Said, Dukuh Atas, and Cawang, are highly correlated and contribute to the property price, respectively. Subsequently, the result suggests excluding CBD and university as both of them do not significantly contribute to the property price.

The HPM equation uses SPSS software by taking into account the stepwise regression method. This method evaluated the input and produced four equation models as the best alternative in the case study. The best equation selected based on the highest R-squared close to 1. It will represent the impact of independent variables on the dependent variable. Based on the result in Table 6, model 4 is the best model due to the highest R-squared with 0.919 and suggests a lower error value with 7.80%, among other variables.

Following the correlation test, another round of tests is performed, including multi-collinearity and heteroscedasticity tests. The multi-collinearity test aims to determine inter-correlation between independent variables in a regression model. The model will proceed for further analysis when any correlation between independent variables was not found. This test was taking into account tolerance and VIF score, which should be higher than 0.10 and less than 10, respectively. The result shows that all four models (Table 6) do not correlate with independent variables. However, the fourth model which

Table 6 Result of multi-collinearity and heteroscedasticity tests

Model		Unstandardized coefficients		Standardized coefficients	t	Sig	Collinearity statistics		R-squared
		B	Std. error				Beta	Tolerance	
1	(Constant)	14.185	0.119		119.283	0.000			0.904
	Ln BSize	0.531	0.018	0.951	29.823	0.000	1.000	1.000	
2	(Constant)	14.295	0.120		119.017	0.000			0.912
	Ln BSize	0.524	0.017	0.940	30.408	0.000	0.984	1.017	
	Hospital	-0.085	0.029	-0.092	-2.965	0.004	0.984	1.017	
3	(Constant)	14.324	0.118		121.557	0.000			0.916
	Ln BSize	0.520	0.017	0.932	30.737	0.000	0.973	1.028	
	Hospital	-0.069	0.029	-0.074	-2.394	0.019	0.929	1.077	
	RasunaSaid	-0.074	0.031	-0.074	-2.370	0.020	0.928	1.078	
4	(Constant)	14.288	0.117		122.273	0.000			0.919
	Ln BSize	0.524	0.017	0.939	31.364	0.000	0.962	1.040	
	Hospital	-0.068	0.028	-0.073	-2.403	0.018	0.929	1.077	
	RasunaSaid	-0.077	0.031	-0.077	-2.510	0.014	0.926	1.080	
	Ln NRooms	0.032	0.015	0.062	2.111	0.038	0.986	1.014	

includes four variables (Building Size, Hospital, Rasuna Said, and Number of Rooms) shows the highest score among others.

Based on the findings, the formula regarding the effect of proximity to transit station on property values in the case study is presented as follows.

$$\begin{aligned} \text{LnPrice} = & 14.288 + 0.524 \text{ Ln_BSize} - 0.068 \text{ Hospital} \\ & - 0.077 \text{ RasunaSaid} + 0.032 \text{ Ln_NRooms} + \%error \end{aligned} \tag{2}$$

where,

Ln Price = Log Natural of property price (US\$).

Ln BSize = Log Natural building size (square meter).

Ln NRooms = Log Natural number of rooms (unit).

Hospital = Location of hospital within 1 km (dummy).

RasunaSaid = Property location within 1 km from Rasuna Said (dummy).

Overall, two variables that consist of building size and the number of rooms significantly contribute to the property price around the transit station. A larger building is suitable for commercials and offices, which requires large space to perform daily activities. The findings showed that expanding the building size every square meter will increase the property price by US\$ 6.45. These findings also found two variables consist of hospital and Rasuna Said—a location of transit station generates negative signs, thus create an opposite correlation. Hospital existence in a 1 km radius of transit station is believed to decrease property price about 6.8% or equal to US\$ 7.29 while the property located 1 km from Rasuna Said station is cheaper US\$ 8.21 than properties beyond the radius.

Discussion

Previous studies showed insightful evidence to create a basis comparison for the present studies. Despite all these findings, the property market on each country has different characteristics and conditions, causing direct comparisons rather difficult. Table 1

illustrates selected studies regarding the impact of rail station on commercial properties worldwide. Unlike previous studies that showed premium accessibility to rail transit, this research found the accessibility attribute to the transit station has an insignificant impact on the property price. In contrast, structural and neighborhood attributes play a critical role in increasing the price of property nearby rail transit.

There are some reasons the variable of distance to transit was out of the equation. First, data mining was conducted starting from September 2018 to March 2019, where the transit stations are yet to operate. At this stage, owners believed transit stations offer less added-value when the users still require additional modes of transportation (bus or motorbike ride-hailing) to reach their property. Accessibility within walking distance is argued to remain unattractive due to various reasons. Citizens of Jakarta tend to have a negative experience when accessing destinations through pedestrian paths since those paths are still lacking in width outside the arterial roads; usually only 1 m wide, lacking in amenities, and lacking in the number of lightings. Moreover, poor condition of walkway pavements, street vendors occupying the sidewalks, and the excessive exposure from vehicle noise and pollution are also among other reasons. Therefore, it is understandable from the viewpoint of property owners to not taking into account better accessibility and mobility as they believe that do not significantly correlate to property prices.

Second, location with strong branding plays a significant role in determining property prices despite the far location from the rail transit station. Rasuna Said, which was included in the formula, is located in the Kuningan area of South Jakarta that provides a balanced lifestyle between working and leisure. This area comprises a variety of office buildings, hotels, embassy offices, and a variety of tourist attractions, including shopping malls, art galleries, public open space, and many others that can be regarded as a place to relax or refresh after a day of work. However, Rasuna Said station in this project is argued to be located in a bad environment where there is a slum area nearby, bad odor from the polluted river, and it is crime-prone due to the incomplete street amenities. Therefore, it is not a surprise when the property in Rasuna Said has a higher price though it is outside of 1 km walkability radius compared to those located with proximity from the rail station.

Conclusion

There is a huge potential to utilize a value capture approach to finance rail transit projects by fully understanding the positive impact of accessibility improvement due to the existence of transit stations on property values. Although capitalization of property market around transit stations has been extensively investigated in developed countries such as in European Countries, the United States, and the East Asian region, research in the same topics in developing countries such as Indonesia remains limited. This research elaborates properties listed in several real estate marketplaces and takes into account seven LRT stations located in Jakarta, Indonesia, that are planned to fully operate in 2021 to fill the research gaps.

Most studies use geographical information system (GIS) and hedonic price modeling (HPM) by taking into account the contractual price of property market. For instance, Barbara et al. [13] elaborated property sales data in Australia and New Zealand from a property business company. On the other hand, Mohammad et al. [17]

used repeated cross-sectional survey data in order to evaluate the value of commercial properties from metro existence. This research argued that the combination of web scraping, geographical information system (GIS), and hedonic price modeling (HPM) has found limited in the literature study. Therefore, this approach is argued to contribute to the body of knowledge from an operational research perspective regarding the transit and property sector.

The paper identifies the effect on commercial property values considering the proximity to rail transit. The area of investigation of rail transit accessibility was within a 1 km radius of road network distance to LRT station. The distance in this research is argued to be similar to the literature that has suggested a distance between 500 to 1500 m [17, 23]. The result of this research shows opposite evidence compared to previous studies, which suggested that accessibility highly contributes to the property price near rail transit [8, 22, 29, 38]. The findings of this research suggest that proximity to rail transit does not have a significant impact on the commercial property price compared to other variables such as building size, number of rooms, location, and hospitals. These other variables contribute to the increasing property price by 52.4, 7.7, and 3.2%, respectively. Building size and the number of rooms suggest a positive effect which means larger building attributed to a number of rooms is highly preferable. On the other hand, property located outside the radius of rail transit (e.g., Rasuna Said station) may experience a higher property price rather than those located closer to the station (see the previous section for more discussion). Hospitals also generate a negative effect on commercial property prices due to several reasons, including pollution, noise, and accessibility. A negative result from the hospital also concurrent with the previous study conducted by Xu, Zhang & Aditjandra [38] where they saw the negativity coming from pollution and cultural factors. The findings of this research contributed as the basis to further examine the value capture scheme from transit operation. It is crucial that policy-makers committed to seeking the best alternative design of land value uplift from different commercial property distance to gain maximum benefit for the public interest.

There are some limitations that should be addressed for future research development. This study was conducted to the under construction transit project, which helps improve the attractiveness of selling commercial property in transit line for property owners. It is suggested that follow-up research takes into account property listings in operated lines such as MRT, BRT or commuter lines in Indonesia to gain more comparative insight into real estate market conditions. The same research also can be performed when the LRT is already operated providing a longitudinal study and suggesting comparable findings. This research also experiences a limitation regarding property values from contractual price due to data scarcity by the real estate industry. Therefore, asking price was used as one of alternative strategies to generate property values for analysis. Future research is also suggested to incorporate a variety of measurements, different spatial models, and techniques to recommend a framework regarding transport-real estate market issues. Furthermore, academics and researchers are also encouraged to involve urban big data and digitalization methods regarding data extraction and analysis to encounter limited data availability. More advanced techniques such as machine learning, predictive modeling, and natural language processing can be adopted in future work to generate more

comprehensive result. Furthermore, the results from this paper are also expected to provide an informative result that can be used by other similar case studies or different domain beyond transportation sector.

Acknowledgements

Not applicable.

Authors' contributions

MAB obtained funding, provided conceptualization, methodology and supervision in the research and suggested revision of manuscript. PM assisted in data mining and assisted in writing the manuscript. GS performed data mining and analysis to the manuscript. MS assisted in analysis and assisted in writing the manuscript. All authors read and approved the final manuscript.

Funding

This research is supported by Directorate of Research and Community Engagement (DRPM), Universitas Indonesia, through Hibah Q1Q2 2019. This work is also partially funded by the Indonesian Ministry of Research, Technology, and Higher Education under WCU Program managed by Institut Teknologi Bandung.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ Department of Civil Engineering, Universitas Indonesia, Depok 16424, Indonesia. ² Department of Civil Engineering, Universitas Pancasila, Jakarta 12640, Indonesia. ³ Center for Sustainable Infrastructure Development, Universitas Indonesia, Depok 16424, Indonesia.

Received: 13 April 2020 Accepted: 14 August 2020

Published online: 02 September 2020

References

- Freeland AL, Banerjee SN, Dannenberg AL, Wendel AM. Walking associated with public transit: moving toward increased physical activity in the United States. *Am J Public Health*. 2013;103:536–42.
- Bhattacharjee S, Goetz AR. Impact of light rail on traffic congestion in Denver. *J Transp Geogr*. 2012;22:262–70.
- Rahman HZ, Berawi MA, Susantono B, Miraj P, Petroceany JS, Maya R. Investigation of an operation and maintenance framework in the railway industry: a case study of the makassar-parepare. *Int J Technol*. 2018;9:549–57.
- Sharma N, Singh A, Dhyani R, Gaur S. Emission reduction from MRTS projects—a case study of Delhi metro. *Atmos Pollut Res*. 2014;5:721–8.
- Zheng S, Hu X, Wang J, Wang R. Subways near the subway: rail transit and neighborhood catering businesses in Beijing. *Transp Policy*. 2016;51:81–92.
- Zhang M, Meng X, Wang L, Xu T. Transit development shaping urbanization: evidence from the housing market in Beijing. *Habitat Int*. 2014;44:545–54.
- Berawi MA, Suwartha N, Kurnia K, Miraj P, Berawi ARB. Forecasting the land value around commuter rail stations using Hedonic price modeling. *Int J Technol*. 2018;9:1329–37.
- Debrezion G, Pels E, Rietveld P. The impact of railway stations on residential and commercial property value: a meta-analysis. *J Real Estate Financ Econ*. 2007;35:161–80.
- Mohammad SI, Graham DJ, Melo PC, Anderson RJ. A meta-analysis of the impact of rail projects on land and property values. *Transp Res Part A Policy Pract*. 2013;50:158–70.
- Tenholder N, Hung MC, Wu YH. Effect of metrolink light rail stations on residential property values in St. Louis county, Missouri, USA. *Int J Geoinformatics*. 2018;14:13–25.
- Brandt S, Maennig W. The impact of rail access on condominium prices in Hamburg. *Transportation (Amst)*. 2012;39:997–1017.
- Sun W, Zheng S, Wang R. The capitalization of subway access in home value: a repeat-rentals model with supply constraints in Beijing. *Transp Res Part A Policy Pract*. 2015;80:104–15.
- Yen BTH, Mulley C, Shearer H, Burke M. Announcement, construction or delivery: When does value uplift occur for residential properties? Evidence from the Gold Coast Light Rail system in Australia. *Land Use Policy*. 2018;73:412–22.
- Seo K, Salon D, Kuby M, Golub A. Hedonic modeling of commercial property values: distance decay from the links and nodes of rail and highway infrastructure. *Transportation (Amst)*. 2019;46:859–82.
- Mulley C, Tsai CH. When and how much does new transport infrastructure add to property values? Evidence from the bus rapid transit system in Sydney, Australia. *Transp Policy*. 2016;51:15–23.
- Kim J, Zhang M. Determining transit's impact on seoul commercial land values: an application of spatial econometrics. *Int Real Estate Rev*. 2005;8:1–26.
- Mohammad SI, Graham DJ, Melo PC. The effect of the Dubai metro on the value of residential and commercial properties. *J Transp Land Use*. 2017;10:263–90.
- Boeing G, Waddell P. New insights into rental housing markets across the United States: web scraping and analyzing craigslist rental listings. *J Plan Educ Res*. 2017;37:457–76.

19. Wegmann J, Jiao J. Taming Airbnb: toward guiding principles for local regulation of urban vacation rentals based on empirical results from five US cities. *Land Use Policy*. 2017;69:494–501.
20. Berawi MA, Ibrahim BE, Miraj P. Developing a conceptual design of transit-oriented development to improve urban land use planning. *J Des Built Environ*. 2019;19:40–8.
21. Hess DB, Almeida TM. Impact of proximity to light rail rapid transit on station-area property values in Buffalo, New York. *Urban Stud*. 2007;44:1041–68.
22. Cervero R, Duncan M. Transit's value-added effects: Light and commuter rail services and commercial land values. *Transp Res Rec*. 2002;1805(1):8–15.
23. Zhang M, Xu T. Uncovering the potential for value capture from rail transit services. *J Urban Plan Dev*. 2017;143:04017006.
24. Bollinger CR, Ihlanfeldt KR, Bowes DR. Spatial variation in office rents within the Atlanta region. *Urban Stud*. 1998;35:1097–118.
25. Bowes DR, Ihlanfeldt KR. Identifying the impacts of rail transit stations on residential property values. *J Urban Econ*. 2001;50:1–25.
26. Chalermpong S. Rail transit and residential land use in developing countries: Hedonic study of residential property prices in Bangkok Thailand. *Transp Res Rec*. 2007;2038(1):111–9.
27. Dziauddin MF. Estimating land value uplift around light rail transit stations in Greater Kuala Lumpur: an empirical study based on geographically weighted regression (GWR). *Res Transp Econ*. 2019;74:10–20.
28. Forouhar A, Hasankhani M. The effect of Tehran metro rail system on residential property values: a comparative analysis between high-income and low-income neighbourhoods. *Urban Stud*. 2018;55:3503–24.
29. Ko K, Cao XJ. The impact of hiawatha light rail on commercial and industrial property values in Minneapolis. *J Public Transp*. 2013;16:47–66.
30. Ryan S. The value of access to highways and light rail transit: evidence for industrial and office firms. *Urban Stud*. 2005;42:751–64.
31. RodrGuez DA, Targa F. Value of accessibility to bogotá's bus rapid transit system. *Transp Rev*. 2004;24:587–610.
32. Deng T, Ma M, Nelson JD. Measuring the impacts of bus rapid transit on residential property values: the Beijing case. *Res Transp Econ*. 2016;60:54–61.
33. Clark SD, Lomax N. A mass-market appraisal of the English housing rental market using a diverse range of modelling techniques. *J Big Data*. 2018. <https://doi.org/10.1186/s40537-018-0154-3>.
34. Batty M, Axhausen KW, Giannotti F, Pozdnoukhov A, Bazzani A, Wachowicz M, et al. Smart cities of the future. *Eur Phys J Spec Top*. 2012;214:481–518.
35. Horowitz JL. The role of the list price in housing markets: theory and an econometric model. *J Appl Econom*. 1992;7:115–29.
36. Castrillo-Fernández O. Web scraping : applications and tools. *Eur Public Sect Inf Platf Top Rep No 2015 / 10*. 2015;1–31.
37. Iacus SM. Automated data collection with—a practical guide to web scraping and text mining. *J Stat Soft*. 2015. <https://doi.org/10.18637/jss.v068.b03>.
38. Xu T, Zhang M, Aditjandra PT. The impact of urban rail transit on commercial property value: new evidence from Wuhan, China. *Transp Res Part A Policy Pract*. 2016;91:223–35.
39. Cordera R, Coppola P, dell'Olio L, Ibeas Á. The impact of accessibility by public transport on real estate values: a comparison between the cities of Rome and Santander. *Transp Res Part A Policy Pract*. 2019;125:308–19.
40. Wang L. Impact of urban rapid transit on residential property values. *Chinese Econ*. 2010;43:33–52.
41. Zipf A, Jöst MM. Location-based services. In: Kresse W, Danko DM, editors. *Springer handbook of geographic information*. Springer: Berlin, Heidelberg; 2012. p. 417–21.
42. Zandbergen PA. A comparison of address point, parcel and street geocoding techniques. *Comput Environ Urban Syst*. 2008;32:214–32.
43. Malaitham S, Fukuda A, Vichiensan V, Wasuntarasook V. Hedonic pricing model of assessed and market land values: a case study in Bangkok metropolitan area, Thailand. *Case Stud Transp Policy*. 2020;8:153–62.
44. Hussain T, Abbas J, Wei Z, Nurunnabi M. The effect of sustainable urban planning and slum disamenity on the value of neighboring residential property: application of the hedonic pricing model in rent price appraisal. *Sustain*. 2019;11(4):1144.
45. Malpezzi S. Hedonic pricing models: a selective and applied review. In: O'Sullivan T, Gibb K, editors. *Housing economics and public policy*. Oxford: Blackwell Science Ltd; 2008. p. 67–89.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.