DETERMINING VEGETATION EFFECT IN URBAN LAND PRICES USING FREE ONLINE RESOURCES

Florence Galeon [1]
[1]Assistant Professor, Department of Geodetic Engineering, University of the Philippines
Room 324 Melchor Hall, UP College of Engineering, Diliman, Quezon City
Email: fagaleon@up.edu.ph

KEY WORDS: Land Price, Bivariate Linear Regression, Correlation, P-Value

ABSTRACT: Land is a primary factor of production and with its importance in the economy of a nation, the value determinants of real estate have long been studied and published in real estate journals. In Metro Manila today, there is growing and increasing demand for green areas as a result of growth in population and rising environmental concerns. Vegetation attributes have often been studied with a positive contribution to property values. The aim of this study is to measure whether vegetation attributes, particularly area covered by trees and distances of sample lots to the nearest tree cluster, are significant to land price or not utilizing the application of online free software and resources. QGIS, a free and open-source GIS application, and Google Earth Pro, which allows visualization, assessment, overlay, and creation of geospatial data, were the primary software used. Remotely-sensed images from Google provided information of area covered by trees and distances of sample lots to the nearest tree cluster. There are 125 samples considered in Quezon City and 16 samples for Mandaluyong City. These sample points are imported as coordinates of latitude and longitude in a .csv file then plotted in QGIS. Using the vector geoprocessing tool, a 300-meter radius buffer is established for the plotted points and considered for the visual inspection of the area of vegetation cover and nearest vegetation cluster. A set of statistics calculators are available online, with features that includes a variety of tests of significance plus correlation, to run a bivariate linear regression. The results for Quezon City samples show that vegetation area has a moderate correlation to land price with $R = -0.477431117$ and $R^2 = 0.227940471$ with P-value $< .00001$ which is significant. This means vegetation area is a significant factor in land valuation at 95% confidence level but with unexpected inverse relationship. The vegetation distance also shows a moderate correlation to land price with $R = 0.317672797$ and $R^2 = 0.100916006$ with P-value = .000306 which is also significant but with another surprising direct relationship. Testing the vegetation area effect for residential lot samples result to $R = -0.418890442$ and P-Value = .000054 which is significant. The commercial lot samples result to $R = -0.574141669$ and P-Value = .000164 which is also significant. Vegetation distance proved to be not significant for both residential and commercial lot samples. The results for the fewer Mandaluyong City samples show, in general, a weak correlation of both vegetation area and vegetation distance to land price and that they are not significant. This study shows that, at 95% level statistical significance, the amount of vegetation cover is significant factor to land price in Quezon City but with an unexpected inverse effect. Vegetation distance proved to be a non-factor to land prices. This study also prove that very expensive proprietary software and data can be avoided and replaced by free online resources in doing geo-spatial research applications.

1. INTRODUCTION

Land is a primary factor of production and considered to be the asset with the longest life span. With its importance in the economy of a nation, the value determinants of real estate have long been studied and published in real estate journals. Real estate property, as a composite goods, is dependent on many unique bundles of attributes for its value (Rosen, 1974; Sirmans et al., 2005). Knowing the factors affecting urban land value is very important in determining the future of urban development and anticipating potential changes. A lot, as an asset, is a collection of various components which add value to the whole.

In Metro Manila today, there is growing and increasing demand for green areas as a result of growth in population and rising environmental concerns. Green spaces or vegetation, undoubtedly, play a major role in the quality of urban life that we experience. The benefits of urban trees are recognized by many which include amenities that are aesthetic, ecological, and economic in nature, as well as those that have a physical or psychological effect on human health. Most of the benefits attributed to urban vegetation are difficult to translate into economic terms. However, the value of at least some of these benefits may be captured in the property values for the land on which the trees stand. Theoretically, any difference in price between two properties that are identical except for their tree cover within their vicinity should be due to the trees themselves.

1.1 Literature Review

Vegetation attributes, principally trees, have often been studied with a positive contribution to property values. Luttik (2000) associated tree presence to increase in property value ranging from 3% to 8%. Seila and Anderson (1982, 1984) found that builders unanimously reported homes on wooded lots sell for an average or estimated 7% more than equivalent houses on unwooded lots. Payne (1973) explored the value added by trees to developed property, controlling the comparability of houses directly using hypothetical values. His results showed that hypothetical sales prices were increased by an average of 7% for houses landscaped with trees. His equation also indicates that trees are associated with an increase in sales price, hardwoods are worth slightly more than pines, and it is only intermediate or large trees that contribute significantly to sales price. To summarize, the average house in his sample sold for 3.5-4.5% more for having five trees in its front yard.

Several studies have involved experimental manipulation of tree cover on properties. Payne and Strom (1975) manipulated the “tree cover” on a landscape architect’s scale model and found that trees added approximately 30% to the value of undeveloped land. Morales et al. (1976) examined 60 homes in Manchester, Connecticut, relating sales price to several factors including location, date of sale, size of house number of rooms, and other features. They found that tree cover added an average of $2686 to the price of the houses, or 6% of the total sales price. Tyrvainen and Miettinen's work (2000) also showed that the premium associated with proximity to urban forest is significant up to 600 m, or within walking distance and the effect of distance to selling prices was strongest up to 300 meters, Tyrvainen and Vaananen (1998).

1.2 Objectives and Significance of the Study

Many studies have shown that the number of trees has a fairly strong positive correlation with selling price. However, most if not all of these literatures are conducted outside the country and may not be necessarily true in the local setting. So, it is time to test this hypothesis in the Philippine real estate market. The aim of this study is to measure whether vegetation attributes, particularly area covered of vegetation and distances of sample lots to the nearest vegetated area, are significant to land price or not. Part of the study is to find answers, not using very expensive proprietary software and data, but using free online resources.

Previous works by the author includes a study to find the effect of a variety of characteristics on the price of the land and identify which factors or parameters are significant. The regression results indicate how much the value or price of lands will change for a given change in each characteristic, holding other characteristics constant. With sufficient data, this hedonic tool allows us to estimate the individual effects of different land attributes on land prices. A regression analysis can then be calculated to determine the correlation for each of the characteristics measured against the transaction price. Those correlation measurements are then used to create a hedonic price model for land. As an alternate real estate valuation method, hedonic price modeling can be used by professional valuers and other stakeholders to determine which land characteristics add significant value to the potential transaction price. The results produced can provide important information for future decisions and help each party better understand the economics surrounding each asset.

1.3 Data and Study Area

The original data is a collection of more than eight hundred (800) samples of land values and specific data that influences the land values of sample lots selected from all over the North-South Commuter Railway (NSCR) North Line from Tutuban to Clark International Airport, the NSCR South Alignment from Solis Station to Los Banos Depot and areas several kilometers away from the railways around Metro Manila and neighboring provinces of Bulacan, Pampanga, Rizal, Laguna and Cavite. The data for this study is composed of one hundred twenty-five (125) samples of land values and specific vegetation attributes that influence the land values of sample lots located in Quezon City. Another set composed of fifteen (15) samples of land values and specific vegetation attributes are collected from Mandaluyong City. The smaller extra set is used for confirmation for the result that we get from the primary set. Shown are the subset areas for this particular study.

Figure 1. Map of Quezon City and Mandaluyong

2. METHODOLOGY

Hedonic pricing is a method identifying price factors according to the premise that price is determined both by internal characteristics of the good being sold and external factors affecting it. This method uses observed data of actual preferences which is standard, accepted economic techniques. To test the data, land values of Quezon City and Mandaluyong properties are isolated. Six (6) major steps are needed to accomplish this study. The general process flow is shown in Figure 2 below.

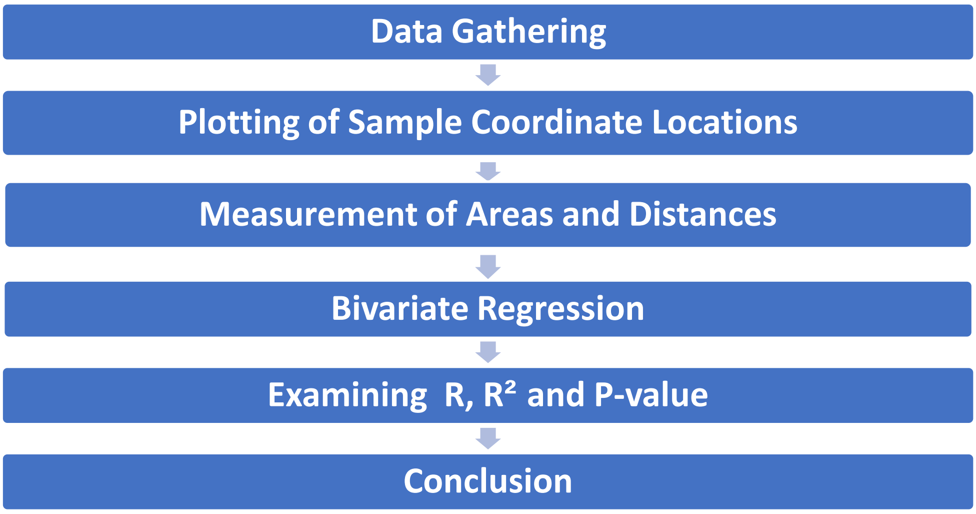

Figure 2. General Process Flow of the Study

There are 141 samples considered, 125 for Quezon City and 16 for Mandaluyong. Google Earth Pro and Google Maps features were utilized to determine distances, areas, and boundaries. According to google, Google Earth Pro is a free software that, albeit not a true GIS, allows visualization, assessment, overlay, and creation of geospatial data. All these capabilities are good enough for an average appraiser, who needs only the very basic geo-spatial technology for a land appraisal work. This user-friendly resource is often a useful intermediary for learners who are interested in learning more about GIS and want to start with more basic processes and tools. Google Maps is a web mapping platform and consumer application offered by Google. It offers satellite imagery, aerial photography, street maps, 360° interactive panoramic views of streets, real-time traffic conditions, and route planning for traveling by foot, car, bike, air and public transportation.

With these 141 lots identified for the analysis, the next step is to run a bivariate linear regression. This analysis examines the effect of any property characteristic individually without controlling for the effect of any other property characteristic. The results will show the coefficient that is the best predictor of sale price for each parameter, the percentage of the sale price accounted for by the perimeter individually, and whether the estimate is statistically significant at 95% level. Parameters or variables that are not statistically significant at least at 95% threshold do not reliably account for differences in the sales price of the examined lots.

3. RESULTS AND DISCUSSIONS

The hypothesis in this study is that lots in an attractive setting attract a premium over lots in a moderate setting. Green areas water bodies, open space and attractive landscape types are aspects of an attractive setting. In valuing urban forests using hedonic price method one of the main problems in empirical applications is that the amenities are difficult to define and measure. In most empirical studies done by (Laakso, 1997: Vainio, 1995: Willis et al., 1993) the variables measuring urban forest amenities are either missing or approximative.

3.1 GoogleEarth

GoogleEarth is first tried to plot the sample points, using their given Latitude and Longitude coordinates, to visualize their relative locations. Once the sample point has been located the vicinity is observed and analized. Particularly of interest are the locations and amount of vegetation cover in the subject area. To limit the working area a radius of 300 meters is set, with reference to Joke Luttik (2000) and Tyrvainen & Vaananen, (1998). The work is tedious and challenging, as there are so many small patches to be considered, ending up estimating most of the time. Next thing to measure is distance of the nearest significant vegetation cover from the sample point. This involves scourging the area and finding a minimum area of 500 m². All these functions are readily available in Google Earth Pro. However, there are instances of blurred images that makes it difficult to map vegetation cover. Google Maps, which can display both map and satellite images, provides alternative remotely sensed data for this study. Figure 3 below shows the Google Earth Pro interface and a sample point with a 300 meters radius.

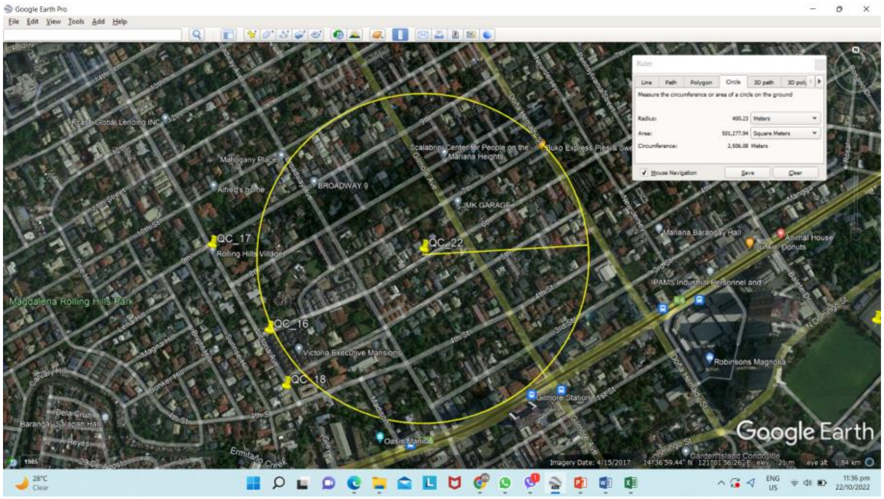

Figure 3. Google Earth Pro Interface Showing a Sample Point and Area of Interest

3.2 Quantum Geographic Information System (QGIS)

Alternative to GoogleEarth in processing data is Quantum Geographic Information System or QGIS. It is a free, open-source software that allows users to create, edit, visualize, analyze, and publish geospatial information. In the web tab there are various base maps to choose from, we chose Google Satellite for this project. We import the sample points as coordinates of latitude and longitude in a .csv file, allowing the plotting of the sample points re-projecting them to EPSG:32651 - WGS 84 / UTM zone 51N. Using the vector geoprocessing tool, create a 300-meter radius buffer for the plotted points. The area within the buffer is considered for the visual inspection of the nearest vegetation cluster and area of vegetation cover. The sample points are shown with the 300-meter radius in Figure 4 below. We delineate the distance of the nearest vegetation cluster by using the vector line tool as shown in Figure 5, the measurement of which is shown in the attribute table and may be exported as a .csv file.

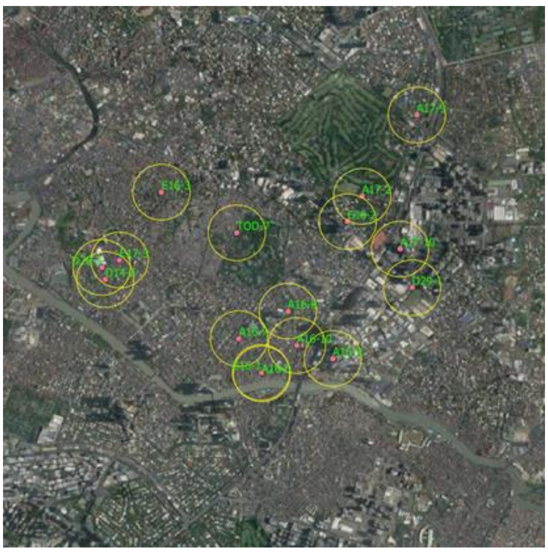

Figure 4. Sample Points with 300-m. Radius

Figure 5. Determining Distance of Vegetation

Digitizing is one of the most common tasks that we can do with QGIS. Often a large amount of GIS time is spent in digitizing raster data to create vector layers that is used in the study, particularly in the analysis part. QGIS has powerful on-screen digitizing and editing capabilities that we used to digitize the area of vegetation cover from online satellite image, as shown in Figure 6 below, by creating a vector polygon shapefile. The total area is shown in the attribute table and may be exported as a *.csv* file.

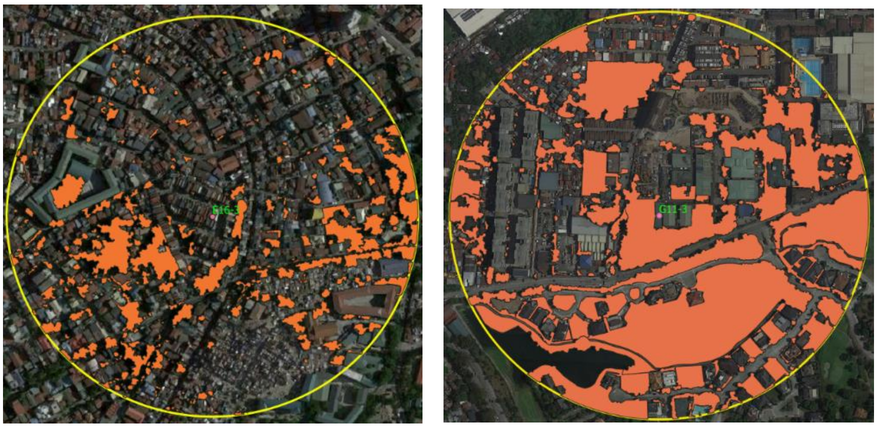

Figure 6. The Digitized Area of Vegetation Cover

3.3 Measurement of Vegetation Areas and Distances to Sample Points

The output of the exercise is a tabulation of the digitized area of vegetation cover in square meters and the distance of vegetation cluster in meters to each one of the 141 sample points. This study is about the possible relationship of the two variables, vegetation area and distance, to the land price. Table 1 below shows the sample points, coordinates in latitude and longitude, vegetation area, vegetation distance, and the land price.

Table 1. Sample Points with Vegetation Area, Distance and Land Price

SAMPLE POINT	LATITUDE	DEG	MIN	SEC	LONGITUDE	DEG	MIN	SEC	Vege Area	Distance	Land Price
1	14.590517	14°	35'	25.86"	121.060868	121°	3'	39.12"	44694	55.7	328000
2	14.590701	14°	35'	26.52"	121.060027	121°	3'	36.10"	41976	31.4	261000
3	14.592059	14°	35'	31.41"	121.061078	121°	3'	39.88"	55296	0.0	128000
4	14.591583	14°	35'	29.70"	121.061163	121°	3'	40.19"	51318	0.0	249000
5	14.592750	14°	35'	33.70"	121.059909	121°	3'	35.67"	50994	12.0	228000
6	14.592750	14°	35'	33.70"	121.059909	121°	3'	35.67"	50994	12.0	228000
7	14.592673	14°	35'	33.62"	121.059424	121°	3'	33.93"	48042	0.0	249000
8	14.602053	14°	36'	7.39"	121.076754	121°	4'	36.31"	55971	9.2	100000
9	14.604881	14°	36'	17.57"	121.018794	121°	1'	7.66"	63855	143.9	100000
10	14.608451	14°	36'	30.42"	121.038335	121°	2'	18.01"	68499	32.5	100000
11	14.608934	14°	36'	32.16"	121.035486	121°	2'	7.75"	73161	7.5	65000
12	14.608940	14°	36'	32.18"	121.037831	121°	2'	16.19"	68121	9.9	100000
13	14.616026	14°	36'	57.69"	121.067393	121°	4'	2.61"	71964	0.0	50000
14	14.614321	14°	36'	51.56"	121.046821	121°	2'	48.56"	70038	17.9	65000
15	14.614321	14°	36'	51.56"	121.046821	121°	2'	48.56"	70038	17.9	65000
16	14.614744	14°	36'	53.08"	121.028407	121°	1'	42.27"	77121	17.0	80000
17	14.616586	14°	36'	59.71"	121.027092	121°	1'	37.53"	79623	16.3	80000
18	14.613565	14°	36'	48.83"	121.028815	121°	1'	43.73"	69453	19.7	80000
19	14.616166	14°	36'	58.20"	121.042786	121°	2'	34.03"	83025	1.0	65000
20	14.614947	14°	36'	53.81"	121.041770	121°	2'	30.37"	85203	9.8	100000
21	14.615675	14°	36'	56.43"	121.048409	121°	2'	54.27"	71622	14.5	50000

3.4 Bivariate Regression

In statistics, bivariate data is data on each of two variables, where each value of one of the variables is paired with a value of the other variable. R is the correlation between the predicted values and the observed values of Y. R^2 is the square of this coefficient and indicates the percentage of variation explained by the regression line out of the total variation. With Excel, determining R and R^2 value is very easy with the help of the functions CORREL and RSQ. R^2 is relevant in various fields such as in stock market and mutual funds because it is able to find the probability or present the correlation between two variables. It has also the ability to explain how much of the movement of one variable can explain the trend of another variable.

The correlation coefficient is given by the formula:

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2]\,[n\sum y^2 - (\sum y)^2]}}$$

The formula for R^2 is given by

$$R\ squared = r^2 = \left(\frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2]\,[n\sum y^2 - (\sum y)^2]}} \right)^2$$

3.5 Examining R, R^2 and P-value

Having computed R and R^2, the significance of the results in relation to the null hypothesis must be determined by performing a statistical test. The null hypothesis states that there is no relationship between the two variables being studied, the land value and the particular parameter. It states the results are due to chance and are not significant in terms of supporting the idea being investigated. Thus, the null hypothesis assumes that whatever is being proven did not happen. The alternative hypothesis states that the independent variable did affect the dependent variable which is the land price, and the results are significant in terms of supporting the theory being investigated and not due to chance.

According to McLeod (2019), a p-value, or probability value, is a number describing how likely it is that your data would have occurred by random chance. The level of statistical significance is often expressed as a p-value between 0 and 1. The smaller the p-value, the stronger the evidence that you should reject the null hypothesis. A p-value less than 0.05 is statistically significant. A p-value higher than 0.05 is not statistically significant and indicates strong evidence for the null hypothesis. This means we retain the null hypothesis and reject the alternative hypothesis.

Computation of P Value from Pearson (R) Calculator is done online for free in https://www.socscistatistics.com/pvalues/pearsondistribution.aspx. The results for Quezon City samples show that vegetation area has a moderate correlation to land price with R = -0.477431117 and R^2 = 0.227940471 with P-value < .00001 which is significant at $p < .05$. This means vegetation area is a significant factor in land valuation at 95% confidence level but with unexpected inverse relationship, as shown in Figure 7 below. The vegetation distance also shows a moderate correlation to land price having R = 0.317672797 and R^2 = 0.100916006 with P-value = .000306 which is also significant at $p < .05$ but with another surprising direct relationship, as shown in Figure 8 below.

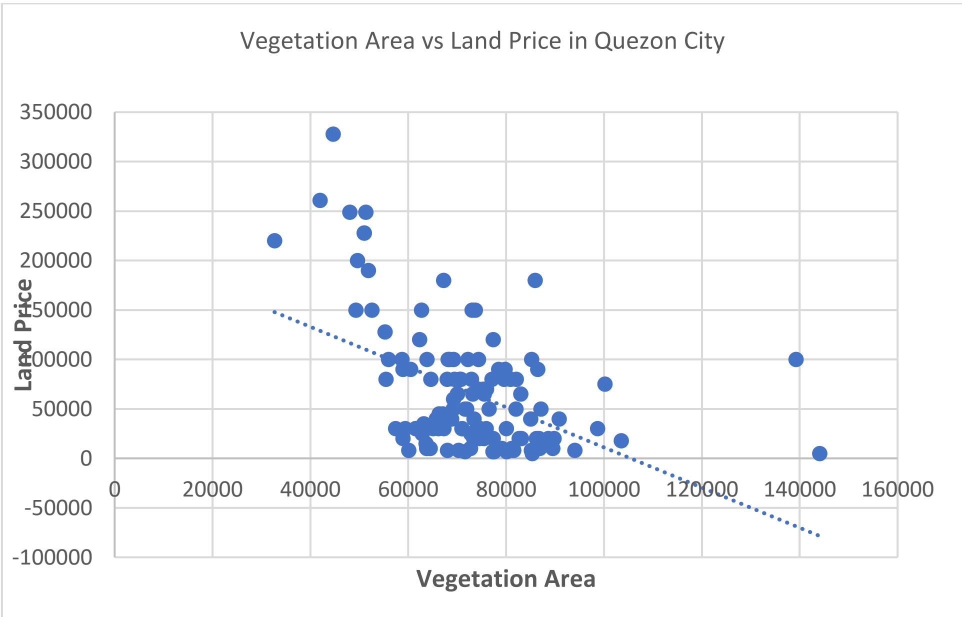

Figure 7. Vegetation Area and Land Price Scatter Plot for Quezon City

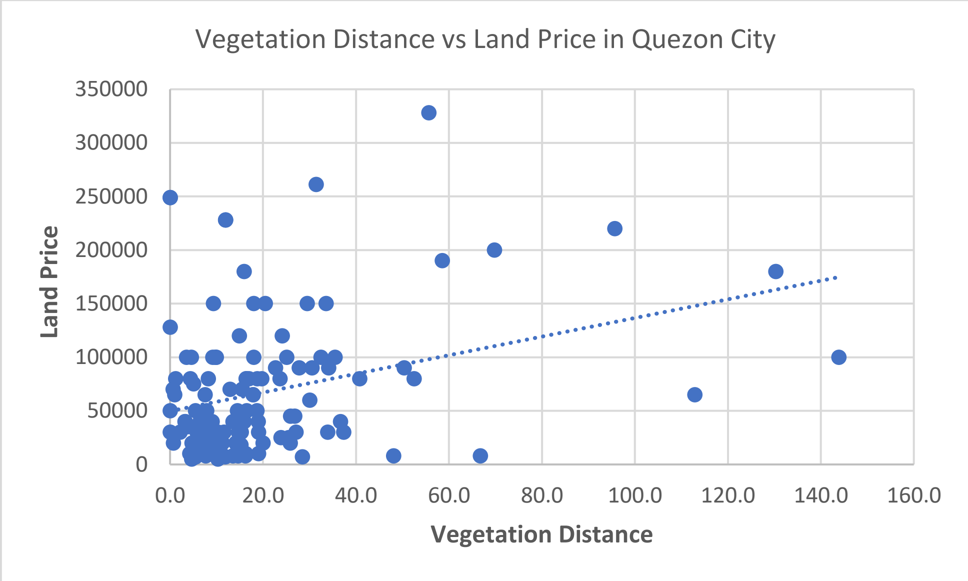

Figure 8. Vegetation Distance and Land Price Scatter Plot for Quezon City

Distinguishing the two categories of samples such as residential and commercial, it is checked if this will alter the results or not. The vegetation area effect for residential lot samples has moderate correlation to land price resulting to R = -0.418890442 and R^2 = 0.175469203 with P-Value = .000054 which is significant at $p < .05$. The vegetation area effect for commercial lot samples also have moderate correlation to land price resulting to R = -0.574141669 and 0.329638656 with P-Value = .000164 which is also significant at $p < .05$. The vegetation distance effect for residential lot samples has weak correlation to land price resulting to R = 0.020899345 and R^2 = 0.000436783 with P-Value = .847702which is not significant at $p < .05$. The vegetation distance also has weak correlation to land price for commercial lots resulting to R = 0.252612418 and R^2 = 0.063813034 with P-Value = .125989 which is not significant at $p < .05$.

The results for Mandaluyong City samples show that vegetation area has a weak correlation to land price with R = -0.043999081 and R^2 = 0.001935919 with P-value = .871753 which is not significant at $p < .05$. The vegetation distance also has weak correlation to land price resulting to R = 0.192149525 and R^2 = 0.03692144 with P-Value = .475907 which is not significant at $p < .05$. The scatter plots are shown in Figure 9 and Figure 10 below.

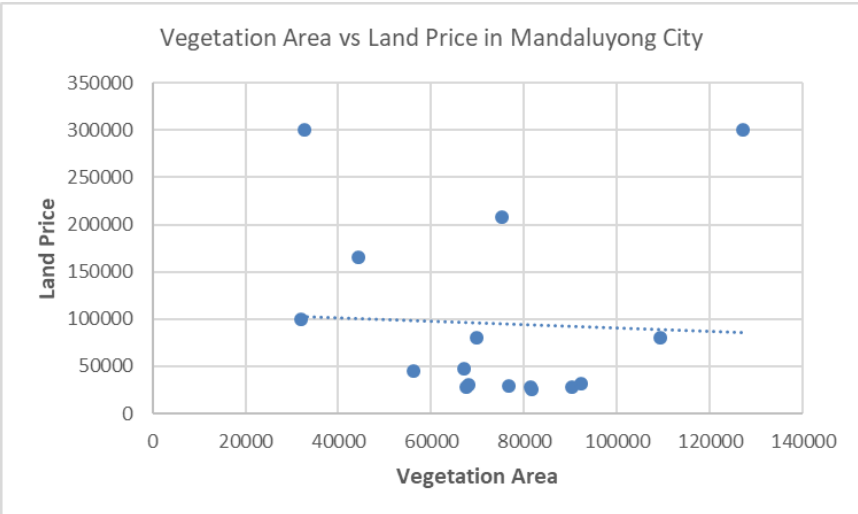

Figure 9. Vegetation Area and Land Price Scatter Plot for Mandaluyong

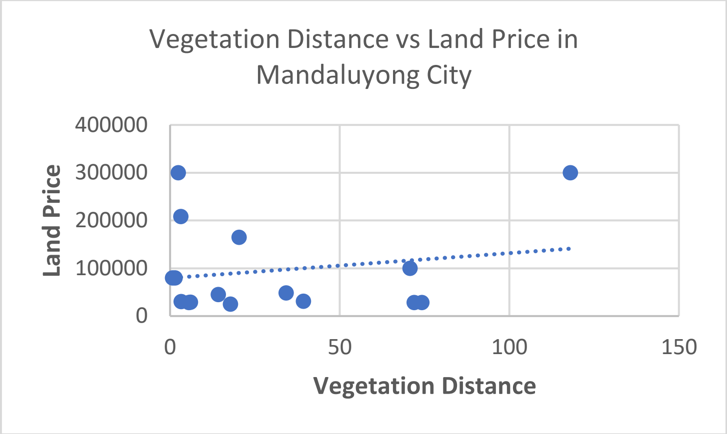

Figure 10. Vegetation Distance and Land Price Scatter Plot for Mandaluyong

Distinguishing the two categories of samples in Mandaluyong, such as residential and commercial lots, did not alter the results. The vegetation area effect for both residential and commercial lot samples has weak correlation to land price indicating that it is not significant at $p < .05$. The vegetation distance effect for both residential and commercial lot samples has weak correlation to land price indicating that it is not significant at $p < .05$.
The inverse relationship between vegetation area and land price in the Quezon City samples, and the direct relationship between vegetation distance and land price, are totally unexpected. This means that land prices are high when there is less vegetation and when trees are far from the subject lot. This is in contrary to the popular belief that trees have strong positive correlation with the selling prices of land.

One possible explanation for this is the fact that some lots located at the heart of commercial centers in Quezon City are the most expensive because of their commercial importance. Examining the physical characteristics of these high-valued lots, they are located where buildings are so closed with each other, high density, concrete, bare or simply deprived of vegetation.

4. CONCLUSION

This study shows that in Quezon City, vegetation area is a significant factor in land pricing, at 95% confidence level, but with unexpected inverse relationship. The vegetation distance is also a significant factor in land pricing, at 95% confidence level, but with another surprising direct relationship. This means that land prices are high when there is less vegetation and when trees are far from the subject lot. The results for Mandaluyong City samples show that vegetation area and distance have weak correlation to land price and they are not significant factors in land pricing. This study also prove that very expensive proprietary software and data can be avoided and replaced by free online resources in doing geo-spatial research applications.

Acknowledgements

The author would like to express gratitude to Cristy R. Hernandez Professorial Chair in Engineering for providing funds in this study, to Aviso Valuation & Advisory Corp. for sharing the data, and to the Office of the Vice Chancellor for Research and Development (OVCRD) of the University of the Philippines for providing financial support to participate and present this paper in the Asian Conference on Remote Sensing (ACRS) 2023.

References:

[1] Rosen, S., 1974. Hedonic prices and implicit markets: product differentiation in pure competition. J. Political Econ. 82 (1), 34–55.

[2] Sirmans, S.G., Macpherson, D.A., Zietz, E.N., 2005. The composition of hedonic pricing models. J. Real Estate Lit. 13 (1), 1–44.

[3] Luttik J, 2000, ``The value of trees, water and open space as reflected by house prices in the Netherlands" Landscape and Urban Planning 48 161 - 167.

[4] Seila, A F, Anderson, L M, 1982, “Estimating costs of tree preservation on residential lots”, Journal of Arboriculture 8 182 – 185.

[5] Seila, A.F. and .Anderson, L.M.. 1984. Estimating tree preservation costs on urban residential lots in metropolitan Atlanta. Macon, GA: Ga. For. Comm., For. Res. Pap. No. 48. 6 pp.

[6] Payne, B.R., 1973. The twenty-nine tree home Improvement plan. Nat. Hist. 82(Y): 74-75.

[7] Payne, B.R. and Strom, S. 1975. The contribution of trees to the appraised value of unimproved residential land. Valuation. 21(2): 36-35.

[8] Morales D, Boyce B N, Favretti R J, 1976, ``The contribution of trees to residential property value: Manchester, Connecticut'' Valuation 23 26 – 43.

[9] Tyrvainen L, Miettinen A, 2000, ``Property prices and urban forest amenities'' Journal of Environmental Economics and Management 39 205 - 223.

[10] L. Tyrvainen and H. Vaananen, The economic value of urban forest amenities: An application of the contingent valuation method, Landscape Urban Planning 43, 105-118 (1998).

[11] S. Laakso, Urban housing prices and the demand for housing characteristics. A study of housing prices and the willingness to pay for housing characteristics and local public goods in the Helsinki metropolitan area, Dissertation thesis, University of Helsinki (1997).

[12] M. Vainio, Traffic noise and air pollution, valuation of externalities with hedonic price and contingent valuation methods, Helsinki School of Economics and Business Administration, A-102, Helsinki, p. 239 (1995).

[13] K. G. Willis and G. D. Garrod, The contribution of trees and woodland to the value of property, J.

[14] McLeod, S. A. (2019, May 28). Introduction to the normal distribution (bell curve). *Simply psychology*: https://www.simplypsychology.org/normal-distribution.html.

[15] https://www.socscistatistics.com/pvalues/pearsondistribution.aspx.