IDENTIFYING THE IMPACT OF TREE VOLUME ON HOUSE PRICES

A hedonic price method approach in Amsterdam

Job Last

Student number 2033879

Research Project 2014

Master Earth Sciences and Economics

Vrije Universiteit Amsterdam

Supervisor: dr. Eric Koomen (VU University)

Co-supervisor: Azarakhsh Rafiee (Geodan BV)

Second assessor: dr. Mark Bokhorst (VU University)

ABSTRACT

The urban forest is a hot topic in scientific literature and has proven environmental, cultural, economic and health benefits. For policy reasons it is interesting to be able to quantify the value of urban trees. This is possible by carrying out a hedonic price method, in which the impact of trees on house prices is being measured. By decomposing houses in different characteristics, the added value of trees to house prices can be computed by executing an Ordinary Least Squares (OLS) model.

This study performs such an analysis, based on an extensive tree volume dataset and a detailed house transaction dataset, in Amsterdam for the period 2009-2013. Moreover, the quality of the tree volume dataset is being researched in a sub study. We find a relatively high positive impact of trees on house prices for trees located within 10 and 25 metres of the house, which is statistically significant. We also find a possible overestimation of tree volume by the tree volume dataset. This could cause an ambiguous change in the impact of trees on house prices. To obtain more robust results, a spatial error model has been used to eliminate spatial dependence. The results from this specific model were consistent with the OLS model.

* Cover picture: The tree house (1759 production services, 2014)

TABLE OF CONTENTS

1. INTRODUCTION

The urban forest has documented benefits on cities and their citizens. Literature on this subject shows that the urban forest not only has environmental benefits, but also carries out a positive impact on health, cultural and economic aspects as summarized by Rafiee et al (2013).

For policy reasons it is helpful to be able to quantify the positive effects of urban trees in order to estimate their value. Cities could revise their urban landscape planning and public greenery policies when knowing the specific value of urban trees. A widely used approach is the hedonic price method (HPM) which is often used to estimate the value of environmental amenities on house prices. The marginal implicit price of a variable that could have an impact on the house price is estimated, by holding all other attributes constant. Using the HPM to quantify the value of individual trees with respect to house prices has been done in various case studies around the world. Donovan and Butry (2010) used a hedonic price model to estimate the effects of street trees on the sales price and time-on-market of houses in Portland, Oregon. A study performed on houses in Ramsey and Dakota counties, Minnesota, showed the effect of tree cover on house prices by using satellite data (Sander, Polasky, & Haight, 2010). In Perth, Western Australia, Pandit et al (2013) studied the effect of broad-leaved trees on property prices. Their following study focused on the location specific effect of tree cover (Pandit, Polyakov, & Sadler, 2014). Moreover, Luttik (2000), Morancho (2003) and Payton et al (2008) have studied the effect of the urban forest on house prices.

Since 1996 the Dutch government has taken the initiative to create a precise height dataset of the Netherlands through laser altimetry (LiDAR). In 2007 the newest version of this dataset begun being acquired, having a point density of 6 to 10 points per square meter. With a maximum 5 centimetre systematic and stochastic error, it is safe to say the dataset is quite detailed (Rafiee et al, 2013). With the AHN dataset as input data, Jan Clement from research institute Alterra created a tree crown dataset (Wageningen UR, 2014), in which 3D tree models are derived from the Silvistar tree model. Rafiee et al (2013), and Geodan BV, Neo BV and Wageningen UR as partners, then calculated the volume of the 3D tree models. Based on the very accurate and detailed data, they have provided an accurate estimation of 3D tree volume, fully covering the Netherlands. These data is publicly available through the website www.boomregister.nl.

The public-private collaboration between the organizations Geodan BV, Neo BV and Wageningen UR is looking for potential applications of this uniquely detailed data. This study therefore explores two aspects of the dataset:

- The accuracy of the tree volume values.
- Describe the impact of tree volume on the transaction prices of nearby houses.

The latter analysis could be of great importance to policy makers by urban landscape planning. However, despite the existence of this uniquely detailed individual tree volume dataset, no hedonic price model has yet been applied to house prices in the Netherlands using these data. In fact, there is no literature at all on the value added to house prices by individual trees in the Netherlands, let alone tree volume.

This research will try to identify the impact of tree volume on house prices by answering the following research question: “Based on an extensive tree dataset and a house characteristics dataset, what is the impact of tree volume on house prices, as caused by the presence of trees?”. Also, a quality assessment of the tree volume values will be carried out.

First the study area and data will be described in Chapter 2, then the sub study on the accuracy of the 3D tree volume dataset will be performed in Chapter 3. Chapter 4 will explain the hedonic price model, followed by the results in Chapter 5 and finalized by the conclusion and discussion.

2. STUDY AREA AND DATA

2.1 STUDY AREA

The capital of the Netherlands, Amsterdam, has been chosen to serve as study area. The city consists of a great variety of types of houses and neighbourhoods, making it interesting to see whether tree volume makes a difference in the sale price of houses. Since trees are relatively scarce in the city, they could be of great value. The results will answer the question whether trees add a significant amount of money to house prices.

In Figure 1 the study area is visually displayed, the big area within the dashed line is Amsterdam, the smaller area is the city district Amsterdam Zuidoost. The population of Amsterdam has in 2014 exceeded 800,000 inhabitants, accounting for nearly 450,000 households. In this study, over 30,000 sold houses in the period 2009-2013 have been analysed. These observations are dispersed over the entire study area as the next section will further go into this.

2.2 DATA

In order to perform a hedonic price method analysis, which will be further elaborated in chapter 3, different types of explanatory variables are needed to fully explain sale prices of houses. We distinguish these variables into two categories, house characteristics variables and spatial variables. The first category describes structural characteristics of houses, while the second category accounts for neighbourhood characteristics and other spatial aspects. All variables and their descriptive statistics are listed in Table 1, where the first variable is the dependent variable, the variables between the two dashed lines the house characteristics variables and the variables beneath the second dashed line the spatial variables, including the tree volume variables.

2.2.1 NVM DATASET

For this study the Dutch Association of Real Estate Brokers (NVM) has kindly provided a detailed dataset of almost all the houses sold in the period January 2009 – June 2013. The dataset consists of more than 130,000 sold houses, including over fifty characteristics per house, and have been given a geographical location using X- and Y-coordinates (Figure 1). However, most of the housed lacked information of one or more characteristics, reducing the amount of houses to 30,095. Transaction data and prices are also provided, which made it possible to construct the variable *YEAR_SOLD* and the dependent variable, *HOUSE_PRICE*. Other house characteristic variables were selected because of their expected relation with house prices following earlier work of Donovan and Butry (2010), Sander et al (2010) and Pandit et al (2013, 2014). Most of all, the paper of Dekkers and Koomen (2008), studying the valuation of open space on urban areas, has been intensively studied, since they also used this NVM dataset in similar research.

Variable	Description	N	Min	Max	Mean	Std. Dev.
HOUSE PRICE	Price in € for which house is sold	30095	64000	4500000	298076	237213
BUILD_AGE	Building age in 9 categories	30095	1	9	4.05	2.77
SURFACE_AREA	Surface area in m^2	30095	26	1006	87.17	49.54
HOUSE_TYPE	Type of house in 14 categories	30095	2	27	20.39	5.34
YEAR_SOLD	Year in which house is sold	30095	2009	2013	2010.70	1.29
N_ROOMS	Number of rooms	30095	1	20	3.32	1.48
N_BATHROOMS	Number of bathrooms	30095	0	6	.98	.42
GARAGE	Presence of garage (1=present)	30095	0	1	.03	.17
GARDEN	Presence of garden (1=present)	30095	0	1	.98	.14
CONDITION_HOUSE	Condition of house on scale of 1-9	30095	1.0	9.0	7.27	.96
GAS_HEATER	Presence of gas heater (1=present)	30095	0	1	.10	.29
MONUMENT	Monument status of house (1=monument)	30095	0	1	.03	.17
DIST_CENTRE	Distance to centre in km	30095	.000	9.440	1.874	1.941
DIST_OPENSPACE	Distance to open space in km	30095	.000	4.755	2.766	1.157
DIST_PARKSLARGE	Distance to large park in km	30095	.000	8.475	.795	.631
DIST_HIGHWAYRAMP	Distance to highway ramp in km	30095	.016	6.351	1.701	.799
DIST_RAILWAYSTATION	Distance to railway station in km	30095	.097	9.617	1.639	.848
PRES_HIGHWAY_100M	Presence of highway within 100m (1=present)	30095	0	1	.01	.09
PRES_RAILWAY_100M	Presence of railway within 100m (1=present)	30095	0	1	.03	.17
URBAN_ATTRACTIVITY	Urban attractivity index on scale of 0-99	30095	0	87	30.85	23.34
P_N_W_ALLOCH	Non-western inhabitants per neighbourhood %	30095	4	75	29.55	17.63
TREE_VOLUME_10M	Total tree volume in radius 0-10	30095	0	5967	84	241
TREE_VOLUME_25M	Total tree volume in radius 0-25	30095	0	41245	2405	2932
TREE_VOLUME_50M	Total tree volume in radius 0-50	30095	0	189713	12412	13001
TREE_VOLUME_100M	Total tree volume in radius 0-100	30095	34	548346	54324	47954
TREE_VOL_RING_10M	Total tree volume in radius 0-10	30095	0	5967	84	241
TREE_VOL_RING_25M	Total tree volume in radius 10-25	30095	0	40744	2321	2822
TREE_VOL_RING_50M	Total tree volume in radius 25-50	30095	0	149921	10007	11044
TREE_VOL_RING_100M	Total tree volume in radius 50-100	30095	17	504898	41912	38008

Table 1: Descriptive statistics all variables

Specifying the most important house characteristics is necessary, since not all variables can be used as input data for the model, due to multicollinearity between variables. The variables that are expected to be most explanatory are *BUILD_AGE, SURFACE_AREA, HOUSE_TYPE, N_ROOMS, N_BATHROOMS, GARAGE, GARDEN, CONDITION_HOUSE, GAS_HEATER* and *MONUMENT*.

2.2.2 SPATIAL DATASETS

Also the most explanatory spatial variables are specified using the aforementioned method. However, since these variables are not included in the NVM dataset, they were constructed by using the Geographical Information Software (GIS), ArcGIS. All houses are given a geographical location through X- and Y-coordinates. Using ArcGIS, the Euclidean distance between each house and the (edge of) several spatial attributes is calculated.

For the variable *DIST_CENTRE*, a digital geometry dataset of neighbourhoods in the Netherlands is used (CBS, 2014). From this dataset, the neighbourhood 'Centrum' of Amsterdam was selected to define the city centre. The variable *DIST_OPENSPACE* is based on a detailed vector land use dataset (CBS, 2008). From this dataset, open space is defined as the types of land use: nature, agriculture, water and recreation. All the small polygons within the built areas of the cities are removed, in order to only account for open space outside the city. *DIST_PARKS_LARGE* is based on the same land use dataset, where thirteen of the biggest and well-known parks have been selected, see Figure 2. Not all parks are selected, because it is expected only the greatest parks will significantly contribute to the house price. The Top10NL dataset contains detailed topographic information and is ideal to define the input data for the variables *DIST_HIGHWAYRAMP*, *DIST_RAILWAYSTATION*, *PRES_HIGHWAY_100M* and *PRES_RAILWAY_100M* (Kadaster, 2014). Highway ramps and railway stations are believed to positively contribute to house prices, because of improved accessibility, whereas the presence of a highway or railway within 100 metres of a house is expected to negatively contribute. To account for house price differences between neighbourhoods, the variables *URBAN_ATTRACTIVITY* and *P_N_W_ALLOCH* are constructed. The Urban Attractivity Index is an attractivity measure of the urban environment, calculated using number of cultural facilities, number of hotels and catering outlets, number of shops, etc. This raster makes differences in number of amenities between neighbourhoods tangible (VU, 2012). Another way to measure differences between neighbourhoods is the indirect effect of the percentage of non-western foreign inhabitants per neighbourhood. These percentages are provided by the Central Statistics Office (CBS) of the Netherlands in their key figures of neighbourhoods publication (CBS, 2014) and are constructed in the variable *P_N_W_ALLOCH*.

Most of these spatial variables have been visualized in Figure 2.

2.2.3 3D TREE VOLUME DATASET

In Chapter 1 the 3D tree volume dataset has been introduced. Although it is called a 3D tree volume dataset, for constructing the tree variables a 2D dataset was used that consisted of polygons of trees, as can be seen in Figure 3. Each polygon contains information about the specific tree, such as: tree volume, tree area and an identification number (Rafiee et al, 2013). To construct the variables, a buffer could be created around each house to select the trees within the buffer and sum up their volumes. However, due

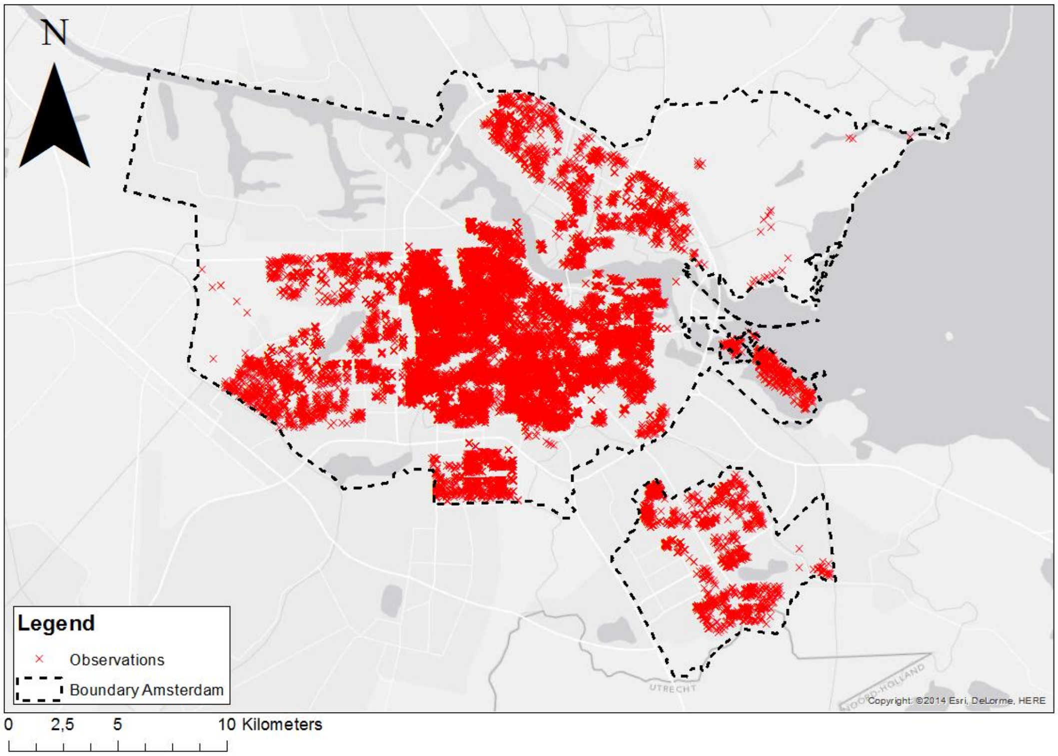

Figure 1: Study area, consisting of the city of Amsterdam, and the observations, consisting of the NVM houses sold in the period 2009-2013

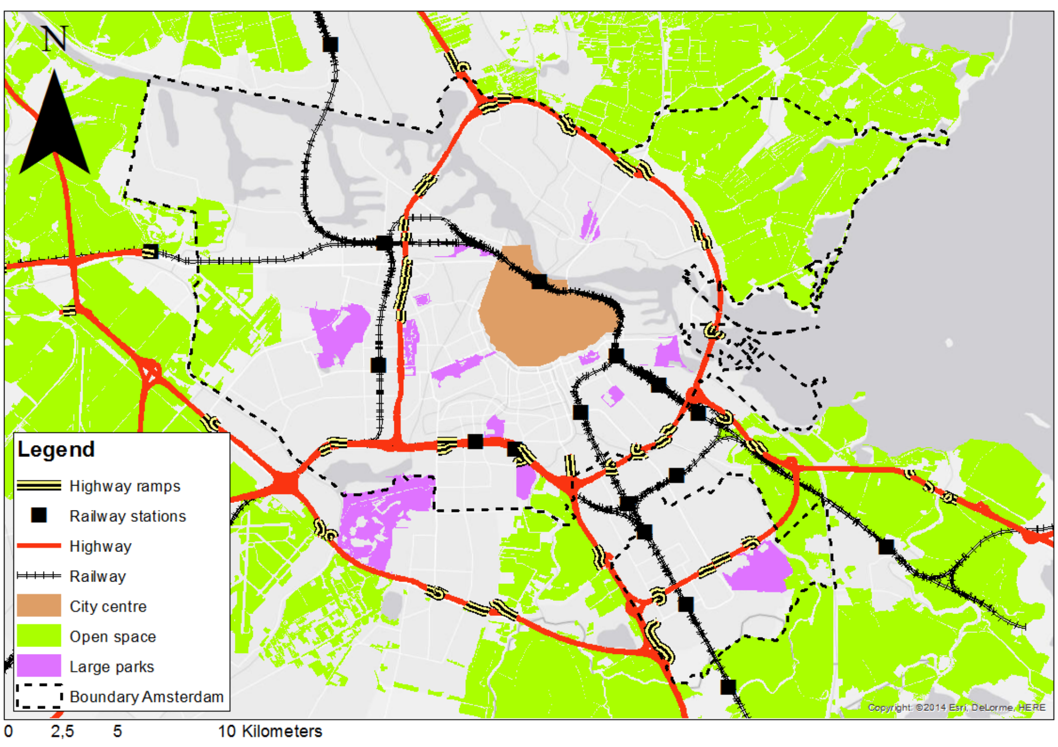

Figure 2: Spatial variables, constructed out of several datasets

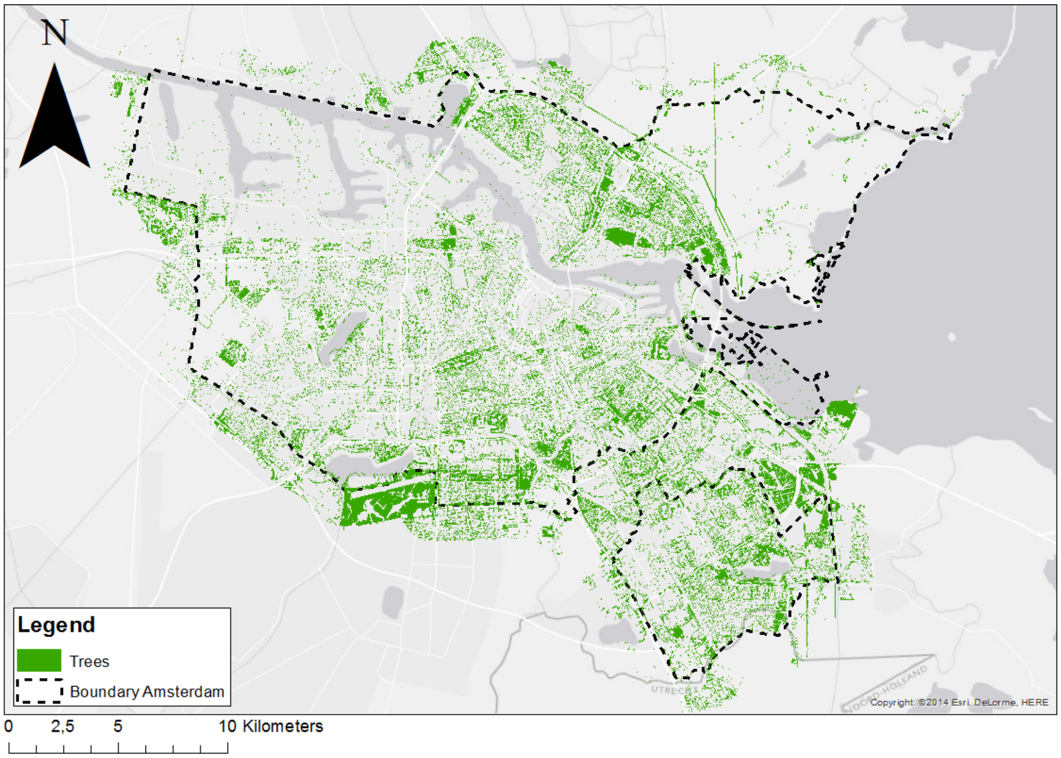

Figure 3: Tree volume dataset in the study area with a 1 kilometre buffer

to an imperfection of the dataset, sometimes trees are clustered in one polygon. Especially in parks and forests it is too hard for the algorithm to distinguish trees, leading to clustering of up to hundreds of trees in one polygon. This becomes a problem, when such a polygon is included in the buffer of a house, because the volume of trees which were originally located outside of the buffer are now included.

For this reason a method has been developed, which calculates the percentage of the area of each polygon falling into a certain buffer. This percentage is being taken from the total tree volume, in order to retrieve more precise data. Figure 4 visually explains this method.

For the variables *TREE_VOLUME_10M*, *TREE_VOLUME_25M*, *TREE_VOLUME_50M* and *TREE_VOLUME_100M* buffers are created of respectively 10 metres, 25 metres, 50 metres and 100 metres around each house. All trees within these buffers are calculated and summed up. For each of these variables a separate regression has to be performed, because variables would overlap each other otherwise. To put all trees within 100 metres in one regression, but label them as different distances, the variables *TREE_VOL_RING_10M*, *TREE_VOL_RING_25M*, *TREE_VOL_RING_50M* and *TREE_VOL_RING_100M* are created. These variables are simply created by taking the specific *TREE_VOLUME* value per distance range, minus the value of the included smaller distance range variables.

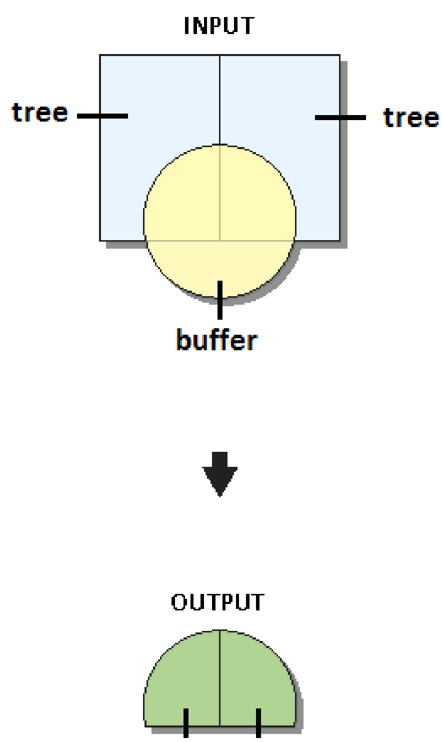

Figure 4: In this figure the buffer would originally select both trees, while most of the trees are located outside of the buffer. The method calculates the tree volume percentages, to retrieve more precise data. N.B. the trees are out of proportion.

3. QUALITY 3D TREE VOLUME DATASET

The 3D tree volume dataset created by Jan Clement from Alterra and Rafiee et al (2013) is a uniquely detailed dataset. However, its quality has never been defined before. Therefore, this study tests the quality of the data. However, the data can only be checked in such a way that we are able to say if the data over- or underestimates reality, due to the limited scope of this sub study.

In order to define the size of the possible errors of the data, first three types of error have to be distinguished:

- Area, position and existence: The error in area, position and existence of the original tree crown data by Jan Clement. When the outline of a tree contains an error, the area and/or position of the resulting polygon is not correct and neither is the volume. Moreover, trees may not be recognized by the tree recognize algorithm of Jan Clement.
- Height: The error in height of the original tree crown data. When the height of a tree contains an error, the polygon outline may be correct, but the volume of the specific tree will be biased.
- Volume: The error in the 3D tree models used to calculate the volume out of area and height input data by Rafiee et al (2013). These could be biased, resulting in over- or underestimated tree volumes.

The goal of this chapter is to find out whether the size and volume of trees in the 3D tree dataset approaches values that are comparable to reality. If this is not the case, it is desirable to find out if the data over- or underestimates.

However, it has to be kept in mind, this chapter is only a sub focus of this study. It has turned out to be impossible to define the rate of the error in height, since no simple techniques are known to check this type of error. For the error in volume a photogrammetric technique has been used to measure a more accurate volume estimation of a tree, but producing significant results has not been achievable.

3.1 AREA, POSITION AND EXISTENCE

The error in area, position and existence represents the differences between the trees derived from the original LiDAR data and the trees truly existing. In order to compare the trees in the dataset with the trees in reality, Google Earth is used, assuming Google Earth represents reality. By visually overlaying the dataset in GIS with Google Earth, it is possible to distinguish trees in four possible types of area, position and existence error:

- True positive: The tree in the dataset is well represented compared to Google Earth, i.e. the tree is likely to exist in reality and the area of the tree in the dataset matches the area in Google Earth.
- False positive: The tree does not exist in Google Earth. This may be possible due to recognizing for instance a lightning pole as a tree by the tree recognition algorithm of Jan Clement, or the cutting of the tree in the time period between creating the data and Google Earths imagery date.
- False area: The tree in the dataset is wrong represented compared to Google Earth, i.e. the tree is likely to exist, but its area in the dataset does not match its area in Google Earth. This could

occur when the algorithm incorrectly recognizes the area of the tree or, concerning a young tree, the tree could have grown in the time period between creating the dataset and Google Earth imagery date.

- False negative: There is no tree existing in the dataset, while it should be according to Google Earth. Possible reasons being the algorithm not recognizing the tree or concerning a newly planted tree in the discussed time period.

Investigating the accuracy of the area, position and existence of the tree crowns, aims at coming up with a percentage representing the difference in tree surface area between the tree dataset and the actual tree surface area in real life. Due to the small scale of this particular quality assessment, the percentage can only be used to indicate whether the dataset overestimates reality, approaches or underestimates reality.

Prior to visually overlaying the dataset with Google Earth, it is necessary to come up with a method of appointing samples. One could take a random tree from the dataset, compare it to Google Earth, document the type of area, position and existence error and go one with the next tree. This method however is quite inefficient and time-consuming. For this research first of all four different land use types were distinguished, being residential areas, industrial areas, agricultural areas and parks. The distinction was made because differences may arise between land use types. The second step is randomly assigning a certain number of points to each land use type. Next, a buffer of 50 metres for each point is created and the trees within this buffer are selected. Then the types of area, position and existence error per tree are documented and for the false area type, the new surface area per tree is measured. Finally, the percentages are being calculated.

While executing this method, problems arose concerning the land use types industrial areas, agricultural areas and parks. For each land use type the problems caused difficulties in being able to document the types. Buffers in industrial areas usually lacked trees and trees did not match in size due to their usually relative young age. Buffers in agricultural were also lacking trees, causing locations in these two types of land use to be too specific. For instance, appointing a location which by chance contained a group of trees, would seriously distort the results for these land use types. In parks however usually a lot of trees are present, causing another type of problem. While the dataset mostly identifies the edges of groups of trees well, it seems to be hard to identify each individual tree. This problem makes it difficult to document the codes well in order not to distort the results.

Since a major part of the investigated houses is located in the fourth land use type, residential areas, it was decided to only investigate this type of land use. It is expected these results are far more credible in the relatively small scope of this sub study.

	True positive	False positive	False area	Total dataset	False negative
Number of trees	187	102	44	333	3
Percentage (%)	56.2	30.6	13.2		
Tree surface area (m^2)	6940.3	915.5	1635.3	9491.1	157.4
Percentage (%)	73.2	9.6	17.2		

Table 2: Area, position and existence error results

A total of ten random locations in the city of Amsterdam was chosen in which 333 trees of the dataset were selected. These 333 trees have a total surface area of 9491.1 m^2. Table 2 shows the results of the quality assessment, with the probably most surprising result being 30.4% of the trees denoted as false positive. This percentage implies 102 of the 333 trees in the dataset do not exist in reality. However, when the surface area of these trees is being calculated, just 9.6% remains. This great difference is caused by the model recognizing lighting poles, traffic signs and other small objects as tiny trees. The false negative type shows that only three trees in the specific buffers are not recognized by the algorithm. These two observations probably imply the model overestimates the total tree surface area compared to reality.

False area	Positive error	Negative error	Total
Number of trees	28	16	44
Percentage (%)	63.6	34.4	
Tree surface area (m^2)	1159.25	-700.50	458.75

Table 3: False area results

The false area type needed to be deeper looked into. As stated, false area trees represent trees which do exist, but their area does not match the area in Google Earth. Therefore, the size error can be positive when the actual size is greater, as well as negative when the actual size is smaller. The results in the table show the overall error is positive, with an average of over 10 m^2 per tree.

	Total dataset	Total reality	Difference (%)
Tree surface area (m^2)	9491.1	8276.1	-12.8

Table 4: Error in area, position and existence, conclusive results

When accounting for the changes the false positive, false area and false negative trees imply, the total tree surface area shrinks from 9491.1 m^2 to 8276.1 m^2. This means that in Google Earth the total tree surface area is 12.8% smaller compared to the tree dataset.

As mentioned before in this report, checking the quality of the tree dataset has a sub focus in this study. However, it is necessary in order to being able to say whether the tree dataset overestimates, approaches or underestimates reality. Since the difference in total tree surface area between the tree dataset and reality, assuming Google Earth represents reality, is -12.8%, it can be stated the tree dataset slightly overestimates reality. This statement will be taken into account in the final conclusion of the report.

3.2 VOLUME

No tools or methods are readily available to obtain the actual volume of a tree that exists in reality, therefore a photogrammetric technique was used. This technique generates 3D models using stereo photos. The open source program 123DCatch is able to create 3D models out of small objects. This can be done by taking series of up to seventy photos from different angles of an object. The software is able to find reference points in the pictures and by using these points it is able of creating a 3D model (Autodesk, 2014a).

To be able to properly measure an object, it needs to meet some requirements. It may not move, it cannot be transparent, plain or reflective and it, or the setting, needs to have a minimum of three distinctive reference points. Trees usually do not meet any of these requirements, therefore a list of demands has been put together which helps the tree become a better object:

- To get reference points, either three objects needs to be put around the tree before taking pictures, or the tree must be located in an urban environment with distinctive building surrounding it;
- It cannot be too windy, causing the tree to move, or too sunny, causing shadows;
- It has to be a freestanding tree, so photographs from about fifty meters all around the tree can be taken;
- The crown of the tree cannot be transparent.

These demands make it rather difficult and time-consuming to collect a significant sample of tree volumes, explaining the unfeasibility of producing results in this study. However, when these demands are met and a sufficient amount of pictures from all around the tree is taken, it is possible to create a 3D model of a tree (Figure 5). Scaling can be done when another object in the setting is measured and scaled, subsequently the height can be measured, referring back to Section 3.2. The tree will contain a few gaps, since it is usually impossible to take photographs of the top of the tree. Also, the software does not contain the option to calculate any volumes.

For both problems a solution exists, by using another program from software developer Autodesk, called Meshmixer. This program is designed for editing 3D models and is able to fill up the gaps and calculate the volume (Autodesk, 2014b). First, in 123DCatch, all objects surrounding the tree need to be removed including the stem of the tree, since we want to know the volume of the crown. Filling up the gaps is done by a function that connects the outlines of the gaps. It is assumed this is a good solution for the gaps and will not significantly bias the volume calculation. Since the object is already scaled, the volume is calculated easily. An example is visualized in Figure 6.

This method has been tested on two trees, resulting in volume values that were close to the values in the 3D tree dataset. The method described was applied to a large number of trees by VU University Amsterdam student Ryan Beij (2014) as part of his GI-minor internship at Geodan.

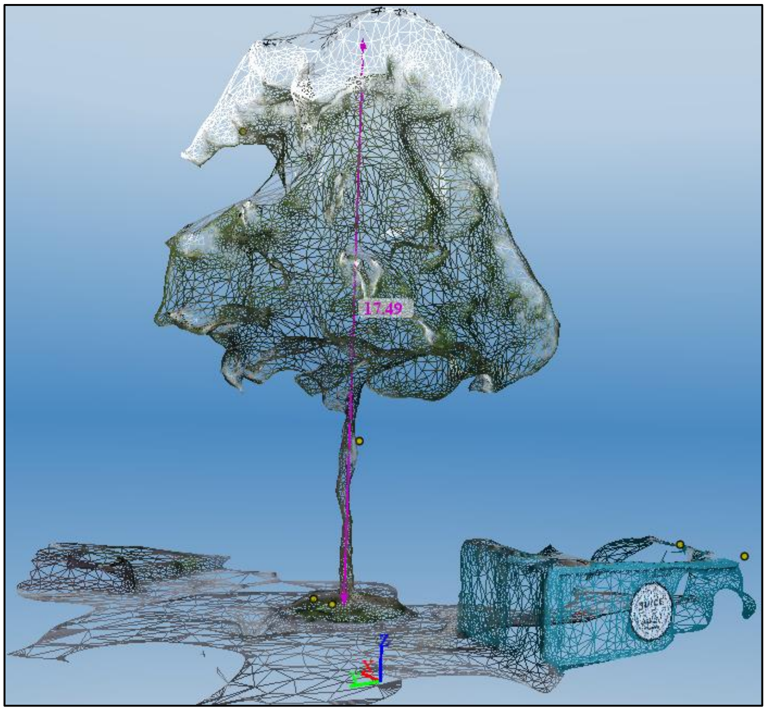

Figure 5: Example of a 3D tree model in 123DCatch

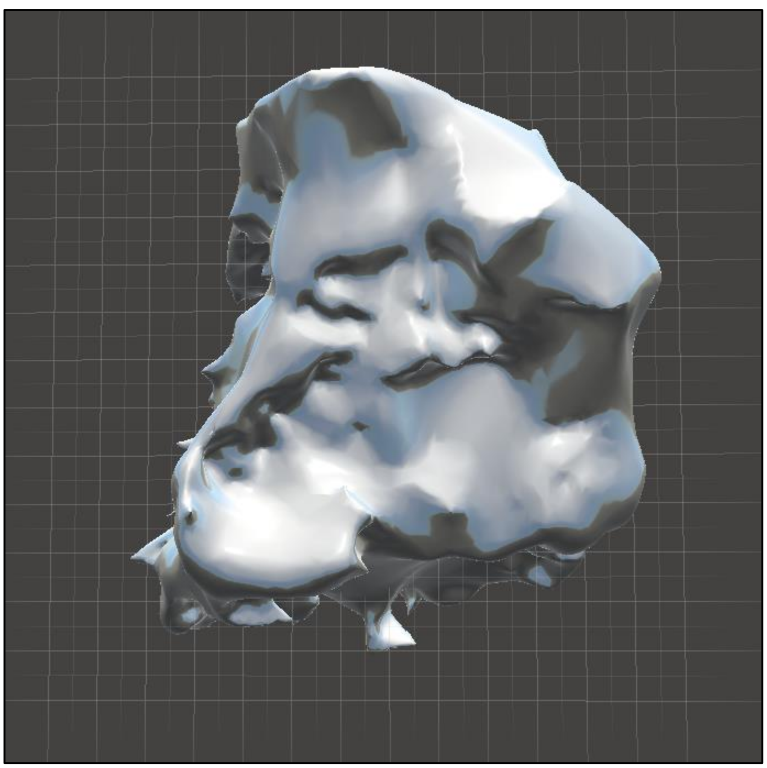

Figure 6: Example of a 3D tree model without stem in Meshmixer

4. HEDONIC PRICE METHOD

The hedonic price method estimates the implicit price of the different components that give a particular good its value. Applied to house prices, it looks at the impact of house characteristic variables, neighbourhood variables and other spatial variables on the total price of a house. The method was described in a general framework by Rosen (1974). The marginal implicit price of the constituting attributes is included in the actual sale price of a house, since an individual shows his willingness to pay for different attributes when buying a house. The marginal implicit price of tree volume can therefore be estimated by holding all other variables constant (Freeman, 2003). An ordinary least squares (OLS) regression will be performed containing the house characteristic variables and spatial variables shown in Table 1. The OLS hedonic pricing method can be written as:

$$\ln(P_i) = \beta_0 + \beta_1 C_i + \beta_2 S_i + \varepsilon_i$$

where $\ln(P_i)$ is the dependent variable and represents the natural log sale price of house i; C_i is a vector of the characteristics of house i; S_i is a vector of the spatial variables for house i; β_0, β_1 and β_2 are the associated coefficients and ε_i is an error term for house i.

Hedonic price models deal however almost always with spatial dependence, because of the spatial character of the particular good. The model has difficulty assessing the impact of spatial relationships, despite the presence of spatial variables. There are two types of spatial dependence model estimates can suffer from: spatial lag and spatial error. The first type occurs when house prices of houses close to each other are correlated. The second type is present when unobserved variables, related to the location of a house, spatially correlate, causing an omitted variable bias (Anselin, 1988).

When taking both spatial lag and spatial error into account, the model can be rewritten as:

$$P_i = \beta_0 + \rho W_i Pi + \beta_1 C_i + \beta_2 S_i + \varepsilon_i$$

where ε_i is equal to:

$$\varepsilon_i = \lambda W_i \varepsilon_i + u_i \quad \text{and} \quad u \sim N(0,\sigma^2)$$

In this model, W_i represents a row-standardized spatial weight matrix and u_i is the uncorrelated error term. The coefficients ρ and λ define the magnitude of the spatial lag and spatial error respectively. The model will be tested for both types of spatial dependence by computing the values of ρ and λ and accordingly to those results, a spatial lag model or spatial error model will be computed.

5. RESULTS

The statistical results for the hedonic price method used are shown in Table 5. In this table the regression results for the case of tree volume within 10 metres, within 25 metres, within 50 metres and within 100 metres are present. Also the single regression with the separate volumes in rings is denoted in this table.

The adjusted R^2 for each regression, given in the table description, tells us nearly 90 percent of the observed variance in house prices is explained. This percentage is very high, due to the high amount of observations and variables. However, multicollinearity between variables is likely to be present, suggesting multiple variables explain the same parts of house prices. To limit this issue, some variables were eliminated from the regression, as they correlated too much with other variables and proved to be weaker when performing a bivariate correlation test between all variables. The variables that were eliminated are: *N_ROOMS*, which correlated with *SURFACE_AREA*, *DIST_OPENSPACE*, which correlated with *DIST_CENTER* and *P_N_W_ALLOCH*, which correlated with *URBAN_ATTRACTIVITY*. Also the weaker variables *DIST_HIGHWAYRAMP* and *PRES_RAILWAY_100M* were excluded.

The table presents satisfying results, since almost all variables proof to significantly contribute to the house price in a positive or negative way. The most distinctive results for the different characteristics will be shortly discussed.

- First, the transaction year shows a negative effect on the house price in 2013 and 2014, compared to the year 2009. This is as expected, house prices lowered the past few years due to the economic crisis.
- From the house characteristics, surface area and house type prove to be important variables. Doubling the surface area of a house almost doubles the house price and most of the house types turn out to have positive effects on the house price, compared to a standard apartment. Especially manor houses have a strong effect, as expected.
- Spatial characteristics turn out to be not such strong predictors of the house price, although they are all significant and have the right sign; as distance to city centre, large parks and railway stations grows, the house price declines.

The most important category for this study however, is the one of the tree volume characteristics. It is not straightforward to interpret these values, therefore their marginal implicit prices have been calculated later in this chapter. The marginal implicit price shows how much the specific variable contributes to the average house price in euros. For now we conclude that all of the tree volume characteristics variables are significant and their effect declines as the radius increases, which is also as expected. There is one value which was not expected, the result of the variable *TREE_VOL_RING_50M*. This value is negative, indicating tree volume between 25 to 50 metres from a house would have a negative effect on the house price. However, it is believed this negative effect is due to autocorrelation. The regression analysis probably corrected for correlation in the third ring, *TREE_VOL_RING_50M*, with a negative value.

OLS model		10m radius	25m radius	50m radius	100m radius	Tree rings
		Coef. St.Err.	Coef. St.Err.	Coef. St.Err.	Coef. St.Err.	Coef. St.Err.
	Constant	7.878 (0.018) ***	7.882 (0.018) ***	7.885 (0.018) ***	7.884 (0.018) ***	7.876 (0.018) ***
Transaction characteristics	YEAR_2010	0.014 (0.003) ***	0.014 (0.003) ***	0.014 (0.003) ***	0.014 (0.003) ***	0.014 (0.003) ***
	YEAR_2011	0.010 (0.003) ***	0.010 (0.003) ***	0.010 (0.003) ***	0.010 (0.003) ***	0.01 (0.003) ***
	YEAR_2012	-0.043 (0.003) ***	-0.044 (0.003) ***	-0.043 (0.003) ***	-0.043 (0.003) ***	-0.044 (0.003) ***
	YEAR_2013	-0.067 (0.004) ***	-0.067 (0.004) ***	-0.067 (0.004) ***	-0.067 (0.004) ***	-0.067 (0.004) ***
House characteristics	BUILD YEAR 1500-1905	-0.008 (0.004) **	-0.008 (0.004) **	-0.009 (0.004) **	-0.009 (0.004) **	-0.008 (0.004) **
	BUILD YEAR 1931-1944	-0.026 (0.004) ***	-0.025 (0.004) ***	-0.025 (0.004) ***	-0.025 (0.004) ***	-0.026 (0.004) ***
	BUILD YEAR 1945-1959	-0.141 (0.006) ***	-0.143 (0.006) ***	-0.143 (0.006) ***	-0.143 (0.006) ***	-0.144 (0.006) ***
	BUILD YEAR 1960-1970	-0.168 (0.005) ***	-0.171 (0.005) ***	-0.175 (0.005) ***	-0.179 (0.005) ***	-0.174 (0.005) ***
	BUILD YEAR 1971-1980	-0.251 (0.007) ***	-0.258 (0.007) ***	-0.261 (0.007) ***	-0.266 (0.007) ***	-0.262 (0.007) ***
	BUILD YEAR 1981-1990	-0.173 (0.004) ***	-0.172 (0.004) ***	-0.174 (0.004) ***	-0.174 (0.004) ***	-0.173 (0.004) ***
	BUILD YEAR 1991-2000	-0.045 (0.005) ***	-0.039 (0.005) ***	-0.043 (0.005) ***	-0.042 (0.005) ***	-0.037 (0.005) ***
	BUILD YEAR ≥ 2001	-0.015 (0.005) ***	-0.009 (0.005) *	-0.012 (0.005) ***	-0.010 (0.005) **	-0.006 (0.005)
	SIMPLE HOUSE	0.086 (0.017) ***	0.090 (0.017) ***	0.091 (0.017) ***	0.090 (0.017) ***	0.088 (0.017) ***
	SINGLE-FAMILY HOUSE	0.050 (0.005) ***	0.055 (0.005) ***	0.056 (0.005) ***	0.057 (0.005) ***	0.053 (0.005) ***
	CANAL HOUSE	0.103 (0.015) ***	0.108 (0.015) ***	0.109 (0.015) ***	0.110 (0.015) ***	0.107 (0.015) ***
	MANSION	0.108 (0.008) ***	0.109 (0.008) ***	0.111 (0.008) ***	0.110 (0.008) ***	0.107 (0.008) ***
	FARM HOUSE	0.644 (0.083) ***	0.675 (0.083) ***	0.689 (0.083) ***	0.694 (0.083) ***	0.661 (0.083) ***
	BUNGALOW	0.405 (0.032) ***	0.418 (0.032) ***	0.423 (0.032) ***	0.420 (0.032) ***	0.405 (0.032) ***
	VILLA	0.408 (0.019) ***	0.409 (0.019) ***	0.416 (0.019) ***	0.417 (0.019) ***	0.404 (0.019) ***
	MANOR HOUSE	0.825 (0.107) ***	0.820 (0.107) ***	0.825 (0.107) ***	0.838 (0.107) ***	0.837 (0.107) ***
	GROUND FLOOR APART	0.064 (0.003) ***	0.064 (0.003) ***	0.065 (0.003) ***	0.065 (0.003) ***	0.064 (0.003) ***
	MAISONNETTE	-0.069 (0.006) ***	-0.066 (0.006) ***	-0.067 (0.006) ***	-0.066 (0.006) ***	-0.066 (0.006) ***
	PORCH APART.	-0.019 (0.004) ***	-0.019 (0.004) ***	-0.020 (0.004) ***	-0.021 (0.004) ***	-0.02 (0.004) ***
	GALLERY APART.	-0.084 (0.006)***	-0.082 (0.006) ***	-0.085 (0.006) ***	-0.085 (0.006) ***	-0.082 (0.006) ***
	TWO FLOOR APART.	0.039 (0.012) ***	0.039 (0.012) ***	0.039 (0.012) ***	0.040 (0.012) ***	0.039 (0.012) ***
	LN_SURFACE	0.945 (0.003) ***	0.942 (0.003) ***	0.942 (0.003) ***	0.941 (0.003) ***	0.942 (0.003) ***
	N_BATHROOMS	0.020 (0.003) ***	0.020 (0.003) ***	0.020 (0.003) ***	0.021 (0.003) ***	0.020 (0.003) ***
	GARAGE	0.092 (0.007) ***	0.093 (0.007) ***	0.093 (0.007) ***	0.093 (0.007) ***	0.093 (0.007) ***
	GARDEN	0.050 (0.008) ***	0.049 (0.008) ***	0.050 (0.008) ***	0.050 (0.008) ***	0.050 (0.008) ***
	CONDITION_HOUSE	0.060 (0.001) ***	0.060 (0.001) ***	0.060 (0.001) ***	0.060 (0.001) ***	0.060 (0.001) ***
	GAS_HEATER	-0.043 (0.004) ***	-0.043 (0.004) ***	-0.043 (0.004) ***	-0.043 (0.004) ***	-0.043 (0.004) ***
	MONUMENT	0.047 (0.007) ***	0.048 (0.007) ***	0.047 (0.007) ***	0.047 (0.007) ***	0.048 (0.007) ***
Spatial characteristics	DIST_CENTRE	-0.045 (0.001) ***	-0.045 (0.001) ***	-0.045 (0.001) ***	-0.045 (0.001) ***	-0.045 (0.001) ***
	DIST_PARKSLARGE	-0.023 (0.002) ***	-0.023 (0.002) ***	-0.024 (0.002) ***	-0.023 (0.002) ***	-0.022 (0.002) ***
	DIST_RAILWAYSTATION	-0.024 (0.001) ***	-0.025 (0.001) ***	-0.025 (0.001) ***	-0.025 (0.001) ***	-0.025 (0.001) ***
	PRES_HIGHWAY_100M	-0.100 (0.012) ***	-0.099 (0.012) ***	-0.101 (0.012) ***	-0.102 (0.012) ***	-0.099 (0.012) ***
	URBAN_ATTRACTIVITY	0.004 (0.000) ***	0.004 (0.000) ***	0.004 (0.000) ***	0.004 (0.000) ***	0.004 (0.000) ***
Tree volume characteristics	TREE_VOLUME_10M	4.237E-005 (0.000) ***				
	TREE_VOLUME_25M		4.760E-006 (0.000) ***			
	TREE_VOLUME_50M			7.032E-007 (0.000) ***		
	TREE_VOLUME_100M				2.682E-007 (0.000) ***	
	TREE_VOL_RING_10M					2.327E-005 (0.000) ***
	TREE_VOL_RING_25M					4.302E-006 (0.000) ***
	TREE_VOL_RING_50M					-9.404E-007 (0.000) ***
	TREE_VOL_RING_100M					3.653E-007 (0.000) ***

Table 5: Estimation results for five regressions using an OLS model (period 2009-2013). All regressions are based upon 30,095 observations. Removed variables, and thus reference value, are *YEAR_2009*, *BUILD YEAR 1906-1930* and *STANDARD APART*. Adjusted R^2 respectively: 0.876; 0.876; 0.876; 0.876; 0.877. Note: * = significant at 0.01 (blue); ** = significant at 0.05; * = significant at 0.10.**

5.1 SPATIAL DEPENDENCE

Before we calculate the actual prices in euros and thus give conclusive results, we test for the expected spatial dependence. Five tests have been performed using a spatial weight matrix with a threshold value of 500 metres, since it is unlikely houses located more than 500 metres from each other show a strong correlation in transaction prices. The results of the tests are shown in Table 6. First, the Moran's I is calculated. Its positive and significant value proves the regression results are biased, since a positive correlation is found between the residuals of houses that are located within 500 metres of each other. The four Lagrange Multiplier tests in the table compute the coefficients ρ and λ and their significance shows both spatial lag and spatial error are present. Since the λ values are much greater than the ρ values, it is appropriate to use a spatial error model, to correct for the important unobserved characteristics on local scale. This spatial error model is estimated using the open source software Geoda and the results are shown in Table 7 and discussed in Section 5.2.

	MI/DF	Value
Moran's I (error)	0.3119	627 ***
Lagrange Multiplier (lag, ρ)	1	12293 ***
Robust Lagrange Multiplier (lag, ρ)	1	4011 ***
Lagrange Multiplier (error, λ)	1	355696 ***
Robust Lagrange Multiplier (error, λ)	1	347415 ***

Table 6: Spatial dependence tests, where MI denotes the Moran's I test-value and DF indicates the degrees of freedom in the Lagrange Multiplier test. Note: * = significant at 0.01; ** = significant at 0.05; * = significant at 0.10.**

5.2 SPATIAL ERROR MODEL

Taking a look at the corrected coefficients of the transaction, house and spatial characteristics, we see nothing much has changed (Table 7). The estimated coefficients differ somewhat, but their significance is maintained. Comparing the tree characteristic results of the spatial error model with the original OLS model, we see also these results remain significant with a slight change in their values. Based on these observations the conclusion can be drawn that the results are not strongly influenced by spatial error. In other words, any variables that may have been unobserved are not believed to have a great effect on the regression results.

However, the spatial error model has produced results that are likely to be more accurate than the results produced by the OLS model. Therefore we calculate the marginal implicit prices based on the results of the spatial error model (Table 8).

Spatial error model		10m radius	25m radius	50m radius	100m radius	Tree rings
		Coef. St.Err.	Coef. St.Err.	Coef. St.Err.	Coef. St.Err.	Coef. St.Err.
	Constant	7.263 (0.020) ***	7.274 (0.021) ***	7.270 (0.021) ***	7.272 (0.021) ***	7.268 (0.021) ***
Transaction characteristics	YEAR_2010	0.022 (0.004) ***	0.021 (0.004) ***	0.022 (0.004) ***	0.023 (0.004) ***	0.021 (0.004) ***
	YEAR_2011	0.022 (0.004) ***	0.021 (0.004) ***	0.022 (0.004) ***	0.022 (0.004) ***	0.021 (0.004) ***
	YEAR_2012	-0.033 (0.004) ***	-0.034 (0.004) ***	-0.032 (0.004) ***	-0.032 (0.004) ***	-0.034 (0.004) ***
	YEAR_2013	-0.042 (0.005) ***	-0.042 (0.005) ***	-0.041 (0.005) ***	-0.041 (0.005) ***	-0.042 (0.005) ***
House characteristics	BUILD YEAR 1500-1905	-0.013 (0.004) ***	-0.013 (0.004) ***	-0.015 (0.004) ***	-0.014 (0.004) ***	-0.012 (0.004) ***
	BUILD YEAR 1931-1944	-0.062 (0.005) ***	-0.060 (0.005) ***	-0.059 (0.005) ***	-0.059 (0.005) ***	-0.062 (0.005) ***
	BUILD YEAR 1945-1959	-0.200 (0.005) ***	-0.202 (0.005) ***	-0.201 (0.005) ***	-0.202 (0.005) ***	-0.201 (0.005) ***
	BUILD YEAR 1960-1970	-0.176 (0.004) ***	-0.180 (0.004) ***	-0.183 (0.005) ***	-0.189 (0.005) ***	-0.179 (0.005) ***
	BUILD YEAR 1971-1980	-0.355 (0.006) ***	-0.371 (0.006) ***	-0.371 (0.006) ***	-0.379 (0.006) ***	-0.366 (0.006) ***
	BUILD YEAR 1981-1990	-0.245 (0.004) ***	-0.243 (0.004) ***	-0.245 (0.004) ***	-0.246 (0.004) ***	-0.244 (0.004) ***
	BUILD YEAR 1991-2000	-0.091 (0.005) ***	-0.082 (0.005) ***	-0.090 (0.005) ***	-0.087 (0.005) ***	-0.079 (0.005) ***
	BUILD YEAR ≥ 2001	-0.084 (0.005) ***	-0.072 (0.005) ***	-0.081 (0.005) ***	-0.076 (0.005) ***	-0.069 (0.005) ***
	SIMPLE HOUSE	0.019 (0.021)	0.027 (0.021)	0.026 (0.021)	0.024 (0.021)	0.017 (0.021)
	SINGLE-FAMILY HOUSE	-0.022 (0.005) ***	-0.014 (0.005) ***	-0.014 (0.005) ***	-0.013 (0.005) **	-0.021 (0.005) ***
	CANAL HOUSE	-0.040 (0.019) **	-0.028 (0.019)	-0.032 (0.019)	-0.030 (0.019)	-0.031 (0.019)
	MANSION	0.053 (0.010) ***	0.057 (0.010) ***	0.059 (0.010) ***	0.057 (0.010) ***	0.050 (0.010) ***
	FARM HOUSE	0.552 (0.073) ***	0.658 (0.072) ***	0.678 (0.073) ***	0.688 (0.072) ***	0.592 (0.073) ***
	BUNGALOW	0.381 (0.041) ***	0.402 (0.041) ***	0.412 (0.041) ***	0.403 (0.041) ***	0.371 (0.041) ***
	VILLA	0.354 (0.024) ***	0.363 (0.024) ***	0.373 (0.024) ***	0.372 (0.024) ***	0.347 (0.024) ***
	MANOR HOUSE	0.710 (0.140) ***	0.711 (0.140) ***	0.718 (0.140) ***	0.721 (0.140) ***	0.728 (0.139) ***
	GROUND FLOOR APART	0.070 (0.004) ***	0.071 (0.004) ***	0.073 (0.004) ***	0.072 (0.004) ***	0.069 (0.004) ***
	MAISONNETTE	-0.170 (0.008) ***	-0.165 (0.008) ***	-0.168 (0.008) ***	-0.166 (0.008) ***	-0.166 (0.008) ***
	PORCH APART.	-0.034 (0.005) ***	-0.037 (0.005) ***	-0.037 (0.005) ***	-0.039 (0.005) ***	-0.037 (0.005) ***
	GALLERY APART.	-0.186 (0.007)***	-0.181 (0.007)***	-0.186 (0.007)***	-0.187 (0.008)***	-0.183 (0.007) ***
	TWO FLOOR APART.	-0.032 (0.016) **	-0.032 (0.016) **	-0.032 (0.016) **	-0.031 (0.016) **	-0.031 (0.016) **
	LN_SURFACE	1.076 (0.003) ***	1.071 (0.003) ***	1.074 (0.003) ***	1.072 (0.003) ***	1.071 (0.003) ***
	N_BATHROOMS	0.032 (0.004) ***	0.032 (0.004) ***	0.032 (0.004) ***	0.032 (0.004) ***	0.032 (0.004) ***
	GARAGE	0.125 (0.008) ***	0.125 (0.008) ***	0.125 (0.008) ***	0.126 (0.008) ***	0.126 (0.008) ***
	GARDEN	0.066 (0.010) ***	0.065 (0.010) ***	0.066 (0.010) ***	0.067 (0.010) ***	0.066 (0.010) ***
	CONDITION_HOUSE	0.065 (0.002) ***	0.065 (0.002) ***	0.065 (0.002) ***	0.065 (0.002) ***	0.065 (0.002) ***
	GAS_HEATER	-0.028 (0.005) ***	-0.028 (0.005) ***	-0.028 (0.005) ***	-0.028 (0.005) ***	-0.028 (0.005) ***
	MONUMENT	0.052 (0.008) ***	0.056 (0.008) ***	0.053 (0.008) ***	0.053 (0.008) ***	0.055 (0.008) ***
Spatial characteristics	DIST_CENTRE	-0.039 (0.001) ***	-0.039 (0.001) ***	-0.039 (0.001) ***	-0.039 (0.001) ***	-0.039 (0.001) ***
	DIST_PARKSLARGE	-0.014 (0.001) ***	-0.016 (0.001) ***	-0.016 (0.001) ***	-0.016 (0.001) ***	-0.014 (0.001) ***
	DIST_RAILWAYSTATION	-0.025 (0.001) ***	-0.025 (0.001) ***	-0.025 (0.001) ***	-0.025 (0.001) ***	-0.025 (0.001) ***
	PRES_HIGHWAY_100M	-0.171 (0.013) ***	-0.171 (0.013) ***	-0.175 (0.013) ***	-0.176 (0.013) ***	-0.170 (0.013) ***
	URBAN_ATTRACTIVITY	0.004 (0.000) ***	0.004 (0.000) ***	0.004 (0.000) ***	0.004 (0.000) ***	0.004 (0.000) ***
Tree volume characteristics	TREE_VOLUME_10M	8.355E-005 (0.000) ***				
	TREE_VOLUME_25M		6.760E-006 (0.000) ***			
	TREE_VOLUME_50M			7.837E-007 (0.000) ***		
	TREE_VOLUME_100M				2.744E-007 (0.000) ***	
	TREE_VOL_RING_10M					5.530E-005 (0.000) ***
	TREE_VOL_RING_25M					6.378E-006 (0.000) ***
	TREE_VOL_RING_50M					-1.894E-006 (0.000) ***
	TREE_VOL_RING_100M					4.613E-007 (0.000) ***

Table 7: Estimation results for five regressions using a spatial error model (period 2009-2013). All regressions are based upon 30,095 observations. Removed variables, and thus reference value, are *YEAR_2009*, *BUILD YEAR 1906-1930* and *STANDARD APART*. Adjusted R^2 respectively: 0.876; 0.876; 0.876; 0.876; 0.877. Note: * = significant at 0.01 (blue); ** = significant at 0.05; * = significant at 0.10.**

The coefficients can basically be translated into the increase in the natural log of the house price per unit of the specific variable. When we take for example the variable *DIST_CENTRE*, per kilometre 0,039 is being subtracted from *LN_PRICE*. Taking an average house price of €300,000, this subtraction can be translated into euros, as shown in table 8. In this table the marginal implicit price for the variable *DIST_CENTRE* is -€11,475. Thus, moving a house with an average house price 1 kilometre further away from the city centre, results in a €11,475 loss in house price.

The other non-tree variables denoted in the table are examples to get the feeling of indication of the various coefficients. When a house with an average house price of €300,000 suddenly becomes a manor house, it would greatly contribute to the house price with €321,280. When the surface of this average house increases with 1%, €5,755 is added. When the house has no central heating, but a gas heater instead, on average €8,283 is subtracted from the house price, and so on and so forth.

Variable	Coefficient	Marginal implicit price
MANOR HOUSE (dummy)	0.728	€ 321,280
LN_SURFACE (%)	1.071	€ 5,755
GARAGE (dummy)	0.126	€ 40,285
GAS_HEATER (dummy)	-0.028	-€ 8,283
DIST_CENTRE (km)	-0.039	-€ 11,475
URBAN_ATTRACTIVITY (scale 1-100)	0.004	€ 1,202
TREE_VOLUME_10M (tree)	8.36E-05	€ 6,267
TREE_VOLUME_25M (tree)	6.76E-06	€ 507
TREE_VOLUME_50M (tree)	7.84E-07	€ 59
TREE_VOLUME_100M (tree)	2.74E-07	€ 21
TREE_VOL_RING_10M (tree)	5.53E-05	€ 4,148
TREE_VOL_RING_25M (tree)	6.38E-06	€ 478
TREE_VOL_RING_50M (tree)	-1.89E-06	-€ 142
TREE_VOL_RING_100M (tree)	4.61E-07	€ 35

Table 8: Marginal implicit prices based upon coefficients in the spatial error model

Assuming these marginal implicit prices are realistic, we can give an actual estimation of the contribution to the house price per tree located at different distances. Since the tree characteristics are denoted in cubic metres, the marginal implicit prices are multiplied by the average tree volume of 250 m^3. This results for a tree within 10 metres of the house in a contribution of €6,267 to the house price. This amount can be seen as rather high, since trees are usually not that expensive or even free, when the tree is public property. When we also take trees up until 25 metres into account, the average price per tree drops to €507. Trees up to 50 and 100 metres force the marginal implicit price per tree even to drop dramatically to respectively €59 and €21.

This trend is also present when the trees are separated into different rings of 0-10 metres, 10-25 metres, 25-50 metres and 50-100 metres. The prices per tree up to 25 metres are somewhat lower, but the result of the price of trees located between 25 and 50 metres of the house is surprising. Also in the spatial error model a negative effect of trees within 25 and 50 metres on the house prices is shown. A negative price of €142 would suggest trees on that specific distance have a negative effect on the house price. Since this is probably due to autocorrelation, we conclude trees at a distance of over 25 metres are arbitrary in their contribution to the house price. Yet, the high prices of trees within 25 metres prove trees nearby houses clearly positively contribute to house prices.

CONCLUSION AND DISCUSSION

In this study I tried to answer the following research question: “Based on an extensive tree dataset and a house characteristics dataset, what is the impact of tree volume on house prices, as caused by the presence of trees?”.

Using a hedonic price method the marginal implicit prices of tree volume have been estimated. According to high contributions found for trees located in a radius of 10 metres and 25 metres (€4,148 and €478 respectively for a house with the average house price of €300,000), it can be concluded trees located nearby houses significantly and positively contribute to house prices. Trees located within 25 metres and 100 metres from houses have, if at all, a less positive effect on trees and therefore their contribution to house prices can be called arbitrary.

This conclusion has however to be reconsidered, when taking into account a possible overestimation of the tree dataset, as concluded in Chapter 3. It is hard to say how the effect per tree would change due to this overestimation, since the change is ambiguous. Less trees in reality could imply a stronger effect per tree, however, less trees in reality could also imply a part of the contribution should be referred to another variable.

Yet, the results have to be put into perspective. The small study of the quality of tree dataset has not all been completed and basically no conclusions can be drawn from it. However, a basis has been provided to fully research the quality of the tree dataset.

Also, despite the use of a spatial error model, it is likely that some spatial dependence and multicollinearity between variables biases the results. The results can therefore only give an estimation and no precise values can be extracted from this study. Yet, the conclusions from this study could give policymakers an incentive to further research the effect of nearby located trees on house prices.

It is hard to compare the results from this study to similar studies for two reasons. The study area, the city of Amsterdam, has a relative high household density and is barely comparable to the U.S. and Australian study areas used in similar research. Moreover, every author uses a different approach when performing an HPM analysis such as this study, with different variables and urban forest units. However, results from similar studies can be generally compared to the values this study found. Donovan and Butry (2010) found that street trees on average add €7,280 to house prices. Sander et al (2010) find that a 10% increase in tree cover within 100 metres adds €1,125 to the house price, whereas Pandit et al (2013) showed that a broad-leaved tree on the street verge adds €11,255 to the median property price. It can be concluded that results in this study are somewhat within the range of the results of the similar studies.

Further studies should be carried out to analyse the difference in impact of private property and public property trees, as literature shows public property has a higher positive impact (Pandit et al, 2014). This can be done by obtaining a Kadaster dataset of the Netherlands, which shows all private properties. Trees can be categorized into private and public property. Moreover, this Kadaster dataset can be used to define the actual house perimeter instead of using point data, as in this study has been done. All of these suggestions will most likely results in more precise estimations.

ACKNOWLEDGEMENTS

I am very thankful for the guidance of my direct supervisors dr. Eric Koomen of the Department of Spatial Economics/SPINlab of the Vrije Universiteit Amsterdam and Azarakhsh Rafiee of Geodan Amsterdam. In addition, I am grateful towards the Dutch Association of Real Estate Brokers (NVM), whom have kindly provided the transaction- and house characteristics dataset of Amsterdam of the years 2009-2013. Lastly, I would like to thank the public-private collaboration between Geodan BV, Neo BV and Wageningen UR for providing the 3D tree volume dataset.

REFERENCES

1759 production services (2014). Retrieved from http://www.1759.co.uk/thelodge/treehouse.htm

Anselin, L. (1988). *Spatial Econometrics: Methods and Models.* Dordrecht: Kluwer Academic Publishers.

Autodesk (2014a, November 24). Retrieved from http://www.123dapp.com/howto/catch

Autodesk (2014b, November 24). Retrieved from http://www.123dapp.com/meshmixer

Beij, R. (2014). *Accuracy assessment on Boomregister.nl.* Amsterdam: VU University.

CBS (2008). *Bestand Bodemgebruik Productbeschrijving.* Voorburg/Heerlen: Centraal Bureau voor de Statistiek.

CBS (2014, November 23). *Centraal Bureau voor de Statistiek*. Retrieved from Kerncijfers wijken en buurten: http://www.cbs.nl/nl-NL/menu/methoden/dataverzameling/kerncijfers-wijkbuurt-kob.htm

CBS (2014). *Toelichting Wijk- en Buurtkaart 2013.* Voorburg/Heerlen: Centraal Bureau voor de Statistiek.

Dekkers, J. E., & Koomen, E. (2008). *Valuation of open space; hedonic house price analyses in the Dutch Randstad region.* Amsterdam: Research Memorandum 24, Faculty of Economics and Business Administration, Vrije Universiteit Amsterdam.

Donovan, G. H., & Butry, D. T. (2010). *Trees in the city: Valuing street trees in Portland, Oregon.* Landscape and Urban Planning 94, 77–83.

Freeman, A. M. (2003). *The Measurement of Environmental and Resource Values, Theory and Method.* Washington: Resources for the future.

Kadaster (2014, November 23). Retrieved from Gedetailleerd topografisch bestand van het Kadaster: https://www.kadaster.nl/web/artikel/productartikel/TOP10NL.htm

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.

Morancho, A. B. (2003). *A hedonic valuation of urban green areas.* Landscape and Urban Planning 66, 35–41.

Pandit, R., Polyakov, M., & Sadler, R. (2014). *Valuing public and private urban tree canopy.* Australian Journal of Agricultural and Resource Economics 58, 453–470.

Pandit, R., Polykov, M., Tapsuwan, S., & Moran, T. (2013). *The effect of street trees on property value in Perth, Western Australia.* Landscape and Urban Planning 110, 134–142.

Payton, S., Lindsey, G., Wilson, J., Ottensmann, J. R., & Man, J. (2008). *Valuing the benefits of the urban forest: a spatial hedonic approach.* Journal of Environmental Planning and Management 51, 717-736.

Rafiee, A., Dias, A., & Koomen, E. (2013). Between Green and Grey: Towards a New Green Volume Indicator for Cities. *Geertman, S., Stillwell, J.C.H. and Toppen, F. (eds.).* Utrecht: Proceedings of CUPUM 2013. The 13th International Conference on Computers in Urban Planning and Urban Management. Planning Support Systems for Sustainable Urban Development.

Rosen, S. (1974). *Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition.* Journal of Political Economy 82, 34-55.

Sander, H., Polasky, S., & Haight, R. C. (2010). *The value of urban tree cover: A hedonic property price model in Ramsey and Dakota Counties, Minnesota, USA.* Ecological Economics 69, 1646–1656.

VU University (2012). *Data taken from Land Use Scanner demonstration version.* Amsterdam: VU University. Retrieved from Data taken from Land Use Scanner demonstration version.

Wageningen UR (2014). *Open boomdata als basis voor webdiensten*. Retrieved from http://www.wageningenur.nl/nl/project/Open-boomdata-als-basis-voor-webdiensten.htm