University of Tartu

Faculty of Science and Technology

Institute of Ecology and Earth Sciences

Department of Geography

Master Thesis in Geoinformatics for Urbanised Society

Investigating Relationships Between Environmental Variables and Sold Land Prices in Hiiu County, Estonia

Alex Jarvis

Supervisors: PhD Age Poom,

PhD Evelyn Uuemaa

Approved for defence:

Supervisor:

Head of Department:

Tartu 2019

Abstract

This thesis titled 'Investigating Relationships Between Environmental Variables and Sold Land Prices in Hiiu County, Estonia' looked profit-yielding and residential land plots in Hiiu County, Estonia. With official sold land data from the Estonian Land Board, the county's environmental variables were measured and used in a multiple linear regression model to see how much they could predict the indexed sold land prices (€/m^2) of land designated as profit-yielding or residential. The results show that the indexed price (€/m^2) of profit-yielding land cannot be predicted to an adequate level with only the use of physical environmental variables, whereas the indexed price of residential land (€/m^2) can be predicted to a moderate level in Hiiu County, just with the use of physical environmental variables. The results suggest, that were the findings combined with economic data, then they could increase the accuracy of a model to determine land prices.

Keywords: Estonia, Hiiumaa, residential land, profit-yielding land, environmental factors, land price.

CERCS code: P510

Annotatsioon

Käesolevas magistritöös hinnatakse seoseid tulundusmaa ja elamumaa müügihindade vahel Eestis, Hiiu maakonnas. Kasutades Eesti Maa-ameti ametlikke andmeid maade müügi kohta ja sama organisatsiooni geoportaali ruumiandmeid erinevate keskkonnategurite kohta, koostati suur andmekogu, kus igale müüdud maatulundusmaa ja elamumaa krundile omistati vastavate keskkonnategurite väärtused. Korrelatsiooni- ja lineaarse regressioonanalüüsi abil hinnati maa müügihinna ja keskkonnategurite vahelisi seoseid ning kas koostatud mudelit saab kasutada müümata maatükkide hinna ennustamiseks.

Märksõnad: Eesti, Hiiumaa, hinnaennustus, lineaarregressioon, maa hinna kujunemine, maatulundusmaa, elamumaa, keskkonnategurid.

CERCS kood: P510

Contents

Introduction

Physically, land embraces all the stable or predictable attributes of the biosphere above and below the earth's surface, including those of the atmosphere, the soil and underlying geology, hydrology, plant and animal populations, and the results of past and present human activity, to the extent that those exert a significant influence on present and future uses of the land by man (Verheye, 2009). These attributes and processes are reflected in the physical state of the land's surface cover and determine human land use (Malczewski, 2004).

The three main drivers of land value can be described as productive value, consumption value and speculative value. Productive value refers to the obtainable financial return of the land, either from rents (Střeleček et al., 2010), profit or subsidies (Quiroga et al., 2019); consumption value refers to amenity factors and ideological, intangible quality-of-life factors (Cellmer et al., 2012); and speculative value refers to the potential financial return of the land (Journeaux, 2016). The three drivers of land value are informed by a number of factors including its productivity, internal quality, location (Choumert & Phélinas, 2015) and use (Maddison, 2000).

Typically, land price modelling focuses on a particular land use type or set of variables to research their impact on the dependent variables. This study chose two different land use types to assess which variables were able to determine their official sold prices. As opposed to limiting the number of input variables, the aim was to test a range of variables that were relevant and could be physically measured within the study site. The applied methodology was based on the measuring of spatial data using geostatistical techniques and applied statistical modelling. The thesis looked for relationships between selected variables and official sold land prices (€/m^2) for profit-yielding and residential land parcels in Hiiu County, Estonia, sold between 01.01.2013 and 31.12.2017.

The study benefitted from the author being provided with the official notary sales information during his internship at the Estonian Land Board (in Estonian: Maa-amet). The state organisation is in the process of conducting a new mass evaluation of the country's land for taxation purposes and potentially, sales valuation (A. Juss, personal communication, 2019), and therefore, the thesis was designed to have practical value for them. The task afforded the author flexibility in his approach and the decision was made to consider Hiiu County as the study site due to it being a closed physical space that had a variety of different measurable factors.

The research questions were as follows:

1. What physical environmental factors have an association with the indexed price (€/m^2) of sold land parcels for profit-yielding purposes and for residential purposes in Hiiu County, Estonia?
2. Can the indexed transaction price (€/m^2) of sold land parcels for profit-yielding purposes and for residential purposes be predicted using physical environmental variables in Hiiu County, Estonia?

1. Theoretical Background

1.1. Estonia’s Land Management

The Estonian Land Board are responsible for the management of state-owned land that is under the administration of the Ministry of the Environment. This includes responsibility for the authorization of the use of land. They administer the national land cadastre, which is the register of land parcels. This contains data includes the plot measurements, location, land surface type and designated land. The designated land use is determined by the Land Board. The cadastral information is available as a shapefile, which represents every land parcel nationwide by a polygon. They also provide the Estonian Topographic Database (ETAK), which contains surveyed layers of the various physical and zonal features of the nation.

The Land Board generate revenue through the sale of state-owned land. The privatisation of land is a European-wide approach (Quiroga et al., 2019) that requires purchasers to abide with the land plots legal regulations, which are determined by designated land use type, and by the specific plot. The main category of land in Estonia is listed as profit-yielding. Land designated as profit-yielding includes cultivated land, forest land, natural grassland, yard land and other lands (Riigi Teataja, 2012). The last mass Estonian land evaluation was carried out in 2001 and this is the valuation which is still in use today for taxation purposes and land reform. However, due to the time gap and limited methodology involved, which used 40 evaluators and was performed manually, it is not accurate enough to value land for sale (A. Juss, personal communication, 2019). Despite some alterations and updates to the 2001 re-evaluation, the need to create a new model is a priority of the Land Board (A. Juss, personal communication, 2018). This will need to be relevant to the wide number of changes of social and economic structure, land use, urban change, local wealth and other that have occurred in Estonia.

The main participants of the previous land evaluation process were the National Land Board and the Bureau of Land Valuation. The variables used in the previous re-evaluation were real estate transactions, immovable estate transactions (e.g. houses), rent, the local economy, and soil productivity (UNECE, 2001). Since the 2001 re-evaluation, there has been the introduction of several special value zones in Estonia that recognise the key property value influencer of location. These subdivided zones for taxation are generally located on various coastal strips and in urban areas. Between 2001 and 2018, there was a housing price index rise until 2007, before an economic recession and property price crash. The index score reached its highest ever level in January 2018, after increasing every quarter between the start of 2013 and the end of 2017. The indexes increased by approximately 40% for ‘unimproved’ land and 45% for residential land between 2013 and 2017 (Maa-amet, 2019)

1.2. Factors Affecting the Price of Land

In economics, value is the esteem in which something is held or can be exchanged under current market conditions (Verheye, 2009). Land belongs to the type of object of value that can be subject to a deal and thus exchange or sales value (ibid). With commodities; the greater the exchange value, the greater the demand for the desired object. This also means that for some land, it might be considered of no worth. There is a fundamental difference between price and value. Market price designates what a property might be sold for at a specific period in time, whereas value designates a property’s actual worth in relation to other similar properties (Ewert, 1979; as cited in Verheye, 2009).

There are two main approaches to the valuation of land. The first approach is focused on the production potential of the land, with some minor adjustments for socio-economic considerations. This is currently applied in rural areas where there is either no functional market and/or where there are very few land sales (Verheye, 2009). The second approach, which is commonly used in developed countries is inspired by the economic value of the land in comparison with recent sales of similar plots under similar conditions. This is the sales comparison approach, which, therefore involves a strong temporal aspect (Střeleček et al., 2010). This is the methodology that the Estonian Land Board currently use when selling off state land and determining the auction starting price (A. Juss, personal communication, 2019). The price of other land parcels, sold in Estonia between private buyers and sellers, can be set at the discretion of the two parties, provided a mortgage is not unrequired (ibid). The Land Board, through its taxation department, has access to all the nation’s notary information of land sales.

Land has the condition of being difficult to obtain (Verheye, 2009). In many countries, due to population pressure and the finite amount of the resource available, this is increasingly the case. Demand drives cost in free-market economics and the price of a land parcel is subject to change, for reasons of internal physical changes that may occur to it, or external changes that affect it. These may also be physical or they may be non-physical, such a political change, administrative change or social changes (ibid). Market value is defined by Tegova (1997) as the ‘estimated amount for which an asset should exchange on the date of valuation between a willing buyer and a willing seller in an arms-length transaction after proper marketing wherein the parties had each acted knowledgeably, prudently and without compulsion’.

There is a growing interest in land as an investment. In the US, land has traditionally been considered a good investment with above-inflation returns (Scott, 1983). Land can provide earnings, in the form of agricultural rent, government subsidies (Põllumäe et al., 2014) or capital gains. In the UK, land ownership brings certain tax-planning benefits and is seen as a ‘safe haven’ investment (Jadevicius et al., 2015). As land value increased, to avoid conflict, land property rights were installed. This initially led to the establishment of a national land cadastre services and then to systems of land taxation (Verheye, 2009). Land values dictate property market price changes. Land use policy, including the changes in land designation, purpose, developments, and taxation is dictated by land price (Albouy & Ehrlich, 2015).

Land planning decisions are made by assessing the results of the analysis of quantitative and qualitative evaluations and classifications of land surveys, combined with economic and social analysis. Along with analysing the past and contemporary land usages, the land use potential for change is assessed. Change is considered an essential property of landscape and in turn, landscape reflects the natural and socio-economic processes (Antrop, 2000). It has been, and still is, a core focus of attention for geographers, with von Humdoldt implying that regional diversification is expressed by landscape, which should be considered as a holistic phenomenon that is perceived by humans (Antrop, 2000). Changes brought about by the move in the former Soviet Bloc towards privatisation and policies that can see it divided further upon the sharing of inheritance that have led to the increasingly fragmented ownership of land (Sklenicka et al., 2014). Mixed ownership creates patterns of patches of mixed land usages, levels of management and sustainability (ibid). Patterns of open arable land or grasslands and closed, forested land that reduce the visual homogeneity of landscape, affect, amongst other things, light and microclimate, biodiversity and vista. Settlements lead to intensification of this process, acting as a ‘control’ centre for the territory of the social group living there, that organises space around it according to ecological, economic, social, cultural and psychological rules (Antrop, 2000). A clear example of this is the marking of territories by fences and enclosures. Such territories are usually located close to services, and utilities, either in a linear or cluster pattern, and in a free market economy, these amenities affect marketed goods such as land or real estate (Pyykkönen, 2006).

Agricultural land

Traditionally, land value primarily came about through agricultural market competition for the most productive lands that produced the best crop yields. Agricultural land with high soil quality and high yields can command higher rent for the owners, and demand for a particular crop can increase prices (Scott, 1983). Land is surveyed to evaluate its attributes, and its potential affects price. Soil quality is therefore a central determinate in the price of agricultural land (Drescher et al., 2001; Maddison, 2000).Increasingly, soil quality is seen as an indicator of sustainable land management practices (Herrick, 2000). Soil quality can be defined as the capacity of soils to function, both within its ecosystem boundaries and within the environment external to that system. Soils function is as a medium for plant growth; a partition and regulator the flow of water in the environment, and an environmental buffer (Verheye, 2009). Basic soil quality indicators include physical characterises, such as clay, sand, silt and rock content, and biological characteristics. Along with air and water, soil quality affects the health and productivity of a given ecosystem and therefore the environment within and surrounding it (Doran & Zeiss, 2000). However, soil is not the only determinate of agricultural land price. Studies assessing the price of farmland often seek to analyse the causal relationships between the fluctuation in the price and land price determinants. More recent studies have assumed that price does have an internal cause of origin and have sought to quantify what these are. The commonly used explanatory variables can be grouped into the categories related to measures of government programs, measures of net return to agriculture, measures of land quantity and measures of financial and macroeconomic activity (Awokuse & Duke, 2006). All these categories are underpinned by location, with environmental location determining yield and market access determining logistical costs (Dirgasova et al., 2017). As knowledge of land increased and assessment methods have improved, other factors have been taken consideration when valuing it (Maddison, 2000; Rubinfeld & Harrison Jr, 1978; Xiao et al., 2017). These include other physical environmental variables and socio-economic factors which determine other potential forms of land utilisation (Verheye, 2009).

Forest land

There are certain environmental factors that can be considered important for various land use types. Forested land can hold great economic value, with positive relationships found between the price of Swedish forest land and the proportion of productive forest land, the mean standing volume and the mean site productivity (Roos, 1996). As has been shown in other studies (Huber et al., 2017; Toppinen et al., 2005; Vedeld et al., 2007), forests can generate income, self-sufficiency or subsidies (Quiroga et al., 2019). The financial incentives of forest ownership make owning privatised land a good short-term investment, along with a potential long term one, should the country's population change and demand for land raise the price.

Forests across Central and Eastern Europe are still going through a process of privatisation, for which EU subsidies are often available, to encourage good management (Quiroga et al., 2019). The acquisition of private forest and agricultural land has little restriction in Estonia (Teder et al., 2015), since the 2012 act the 'Restrictions on Acquisition of Immovables Act', which states that EU and EEA citizens have the right to acquire immovables that contain agricultural or forested land without restrictions (Teder et al., 2015). Additionally, the land type is not subjected to inheritance laws, which in the UK, add value, due to the tax breaks provided (Profeta et al., 2014). In the Baltic states, there have been large private investments in the sawmilling industry, due to increased demand for saw log or pulpwood products (Toppinen et al., 2005) from the coniferous trees, which thrive in low quality, acidic soils. Fast growing pine and spruce plantation forests, if well managed, can bring great productive value to land owners. For some, forests have great ideological value. Whilst private forest land parcels are purchased for different reasons, from self-sufficiency for firewood, to long term investment, there have been concerns about the sale of forest logging rights leading to increasing clear-cutting. Additionally, the large number of forest owners provides challenges about coordinated management, such as ensuring road access, drainage, the prevention of disease and bark beetle (Põllumäe et al., 2014).

The effect of access

Road access is an important consideration for any land buyer. A surfaced road means year-round access for most vehicles and rapid transit from rural areas to urban areas and the services located there. A non-surfaced road is likely to have an inferior surface which is unsuited to most vehicles and potentially impassable during certain weather situations. However, they are a key means of rural infrastructure and market access (Jacoby, 2000) The planning of roads is not merely an economic task, but increasingly involves considering ecological and scenic aspects (Antrop, 2004). Despite this, roads tend to be detrimental to landscape and life quality. In terms of noise pollution, there are various sub-variables within the noise variable that will affect noise from a road on a land parcel, including elevation, tunnels, trees, surface type and vehicle type. However, in terms of correlations, in cities such as Seoul, noise has sometimes been found to have a positive correlation with property price (Kim et al., 2007). Noise and pollution from roads depend on the number and type of vehicles using it.

Industry

Industrial sites and their heavy goods vehicles require access. Industry can have a negative effect on land prices (Bloodworth et al., 2009). For some industries, such as a quarry, a mine or a landfill, the effects are clear and long-lasting. Despite the economic and employment benefits, they reduce the local land price and have negative public perceptions, due to various types of pollution produced whilst in operation. After closure, if well restored, former pits can be turned into biodiversity hotspots, assets in flood management or sites of carbon sequestration (Fourie & Brent, 2006; Rhoades et al., 2001). In this way, their local environmental effects can be positive. However, if poorly restored, large pit lakes can be left, which are unsuitable for leisure activities and can lead to the evaporative loss of groundwater (ibid). Similarly, depending on the standards involved, landfills can be environmentally hazardous. However, the real issue in developed countries, where landfills are properly sealed, is the ideological one. Unless viewed by the public as essentially benign, they have been shown to negatively influence prices (Nelson et al., 1992). Land price has been shown to be affected by risk aversion (Thomas, 1999), and another industry that has been shown to reduce average land prices compared to other comparable areas is that of nuclear energy (Folland & Hough, 1991). for part-ideological reasons is nuclear energy. and in terms of this, average land prices the location around a nuclear power plant were found to be significantly lower than areas without , although the choice to construct the industry there was influenced by the already cheap land.

Water

Freshwater from rivers, drainage channels, lakes or ground sources is a necessary resource for most land usages. Groundwater contamination, if it occurs can seriously impact on land price (Page & Rabinowitz, 1993). Surface water is considered one of the most important factors for residential property prices. Water bodies are considered a key determinate of scenic environmental value (Cellmer et al., 2012). In urban zones, it is said that green spaces, water bodies and healthy environments provide amenities and services that raise the quality of life. Quality of life, sentiment, ideology and scenic quality, although very influential on the property price, are not tangible and are therefore difficult to assess and quantify (Jim & Chen, 2006). There is a lack of academic literature that considers these subjective values, although some studies have attempted to look for such relationships (Cellmer et al., 2012). There is, however, a lot of research that goes into the negative effects of water as a risk on land and property prices. In the event of a flood or drought, or even the increased probability of one, land prices can fall (Daniel et al., 2009) For coastal zones, whilst sea view is one of the most desirable attributes of a residential property (Shi Ming & Chee Hian, 2005), and a desirable sandy beach in a warm climate provides a unique pull-factor for leisure and tourism, low-lying coastal zones are at risk of flooding. Wetlands, which also benefit eco-tourism, have recently been the focus of numerous economic studies and conservation drives. Scientists have attempted to quantify their value as carbon and nitrogen sinks (Mitsch & Gosselink, 2005), and natural wastewater treatment and flood defence tools. Their proximity to properties has also been measured, with open wetlands in Portland, US, shown to have a positive relationship, whilst other types of wetlands had either a negative or non-significant relationship (Bin, 2005).

Opportunity Cost

For many investors in profit-yielding designated land, the land's potential use will likely have been factored into the rationale behind the purchase. This is the 'opportunity cost'; the cost of the most appropriate alternative use (Verheye, 2009), as opposed to the growth potential of the money invested, compared to if it had been invested elsewhere. One way of maximising the land's price without any major alterations would be by the sound maintenance of it (Põllumäe et al., 2014). A change in the land's designation, such as a decision by the Land Board to encourage new suburban developments by changing the designation from rural to urban will also be implemented with the aim of increasing the real estate value of an area (Roka & Palmquist, 1997). Private construction companies will be encouraged to physically alter the land as the land use changes from profit yielding to residential. Likewise, brownfield sites, such as industrial zones can become zones of redevelopment, which can improve the livability of an area (Ruelle et al., 2013). In some greenfield developments, a series of planned surfaced roads and utilities, such as electricity lines are laid out in advance of the construction of homes. Owners can also add value to the land themselves. Structural improvements may also affect the land's taxation rate (Verheye, 2009).

Recent research concludes that homebuyers are willing to pay for a better environment (Montero et al., 2018). Land is, in Western countries, considered a prerequisite of achieving individual freedom (Verheye, 2009). The selection of location and environment is not neutral, because humans aspire to live in areas where they can enjoy a high quality of life (Cellmer et al., 2012).Across Europe, there is a growing demand for green areas for living, due to population growth, increased leisure time and increasing knowledge and awareness of the environment (Tyrväinen & Miettinen, 2000). Counter urbanisation, including to distant rural areas has been an academic focus in some former Soviet Bloc countries (Šimon, 2014). There is a desire to meet one of man's basic needs by returning to live in a landscape where there is the presence of greenery, forests and water (Cellmer et al., 2012), something made possible by telecommuting work. For studying price in residential land, the assumption that residential land is purchased in order to construct a dwelling, live in or renovate a property that is already in place, hedonic theory can be applied (Orford, 2017; Xiao et al., 2017). This means not only predicting the value of the property based on the known variables related to the building or land in which the property is or will be located, but also a set of variables related to the neighbourhood, economy (census tract data and income), local politics, as yet-unmapped information or variables that have not been studied but which affect the demand for a property (such as noise levels) (Kim et al., 2007). Some of these features may or may not affect the property in question. However, the assumption can be made that residential land is selected for work location or for pleasant environmental factors (Cellmer et al., 2012). Therefore, residential land is usually located close to services, such as schools and leisure facilities, either in a linear or cluster pattern. Land use patterns are therefore not random. They are organised based on the factors that add value. Differently from agricultural land, the areas surrounding residential land plots, particularly those associated with pleasant natural or managed environments, such as and beaches, parks, golf courses , are often improved or purchased because of their quality, and reduced distance to these leisure locations has been seen to increase property prices (Bolitzer & Netusil, 2000). Human improvement and change are considered an essential property of landscape (Antrop, 2004). Land ownership creates patterns of patches of mixed land usages. Patterns of open and closed, arable or grasslands and forested land affect the homogeneity of landscape and affect, amongst other things, light and microclimate, biodiversity and vista. Each human settlement has been described as a 'control' centre for the territory of the social group living there. It organises space around it according to ecological, economic, social, cultural and psychological rules (Antrop, 2000) A clear example of this is the marking of territories by fences and enclosures. Residential land change has been measured by regression analysis, with researchers looking at factors such as the number of bedrooms, the spatial density and size of properties to assess the movement of boundaries between urban and rural zones, with administrative methods such as the implementation of split taxation zones used to control these changes (Banzhaf & Lavery, 2010).

2. Data and Methods

2.1. Study Area

Hiiu County was chosen as the investigation site for this thesis. Having the aforementioned environmental qualities and a mix of land types, urban spaces, industrial zones, forest and protected zones, it can be considered an ideal as a microcosm the whole of Estonia. The county is a group of nearly 200 islands to the west of the mainland, which is made up almost entirely by the 989 km² island of Hiiumaa. Hiiu County is by far the least populated county in Estonia, with 9387 (as of 2018) inhabitants and a population density of 9.1 inhabitants per km². The county centre is the city of Kärdla, on the northern shore of Hiiumaa.

Hiiumaa is accessible year-round, either by the ferry port, the harbour port, the airport or by temporary ice-road. Internally, the island has well maintained main roads, and numerous secondary roads, including unsurfaced tracks. With its proximity to the capital, Tallinn, and the well-regarded nature reserves, sand beaches and a milder climate, it is a popular summer destination and location for second homeowners, of whom there has been a relative increase in recent years (Hiiu County Government, 2015).

Hiiu County can be considered as the wildest county in Estonia, with nearly 70% of the main island covered with different types of woody vegetation, from scrub to ancient coniferous and deciduous forests. These land types have increased as the reprivatisation of land at the beginning of the 1990s saw the demise of collective agriculture and the natural rewilding of large parts of the county. The 'saving grace' of the natural landscapes was the young and mostly thin and stony limestone soils, which have limited intensive agriculture (Kaasik et al., 2011). In the modern day, there are still various open spaces too, from arable land and agricultural fields to natural grasslands, but 24% of the territory is environmentally protected. These zones include the alvars and the still significant wetland environments (although twentieth century land reclamation reduced the number significantly (ibid)). Nowadays, saturated land accounts for 7% of the island's surface area and vary from bogs, quaking bogs and fens to shoreline reeds (Hiiu County Government, 2015). There are various short and small water streams, with most merging with the rows of drainage ditches that cover the southern half of Hiiumaa. Extensive land improvement work has significantly improved the flow of the rivers and streams and their seasonal levels (ibid).

Most Hiiu County's soils are poor quality for arable purposes (Reintam et al., 2005). In terms of industry and production land, the forestry industry that is set around the extensive biomass resources is central to the local economy. This industry provides energy sources for the country, and for export. Other zones of economic resource include quarries. Other major economic sectors are tourism and manufacturing, which is located within production sites (Hiiu County Government, 2015).

In the years 2013 to 2017, there were 2945 real estate transactions in Hiiu County, representing 1% of the total nationwide sales, 2.2% of the total nationwide territory sold, and 0.48% of the total value of nationwide sales. There were 1964 transactions of immovables without buildings over the 5-year period. 80% of these transactions were purchase and sale transactions (Maa-amet, 2019). The average price per square metre over the 5-year period was €0.16 (figure 1) and remained steady.

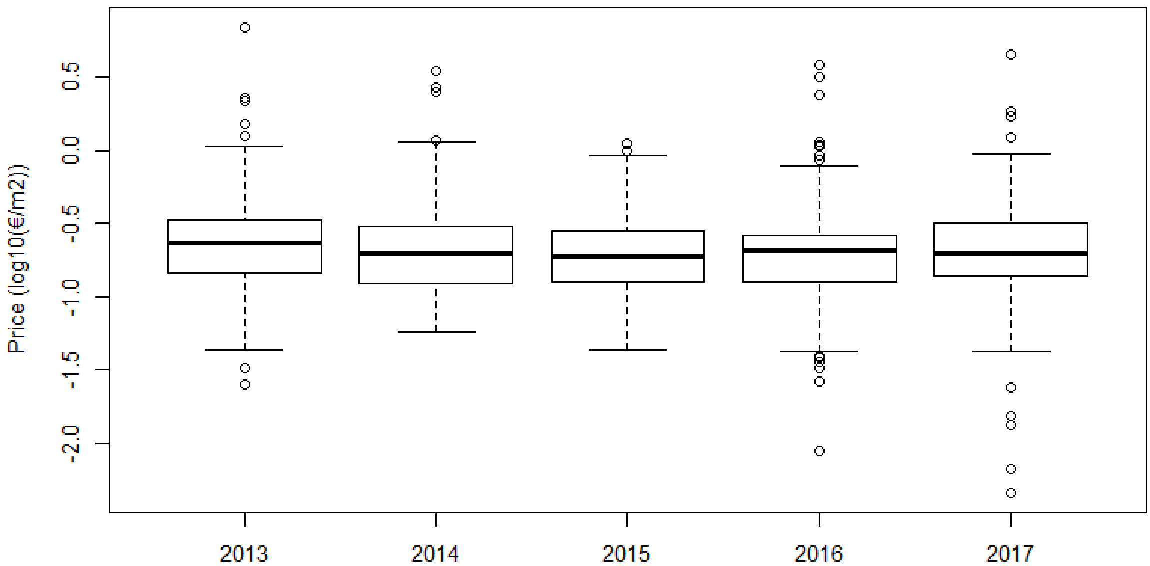

Figure 1. Yearly median, quartiles and range of Sold Profit-Yielding Land Price (log10(€/m^2)) for Hiiu County (2013-2017)).

The price of sold residential land (€/m^2) in Hiiumaa, where one land parcel was sold per transaction, decreased slightly between 2013 and 2014, before increasing considerably in 2015. The average price dropped again in 2016, before recovering in 2017 (figure 2). The average price per square metre over the 5-year period was €0.54.

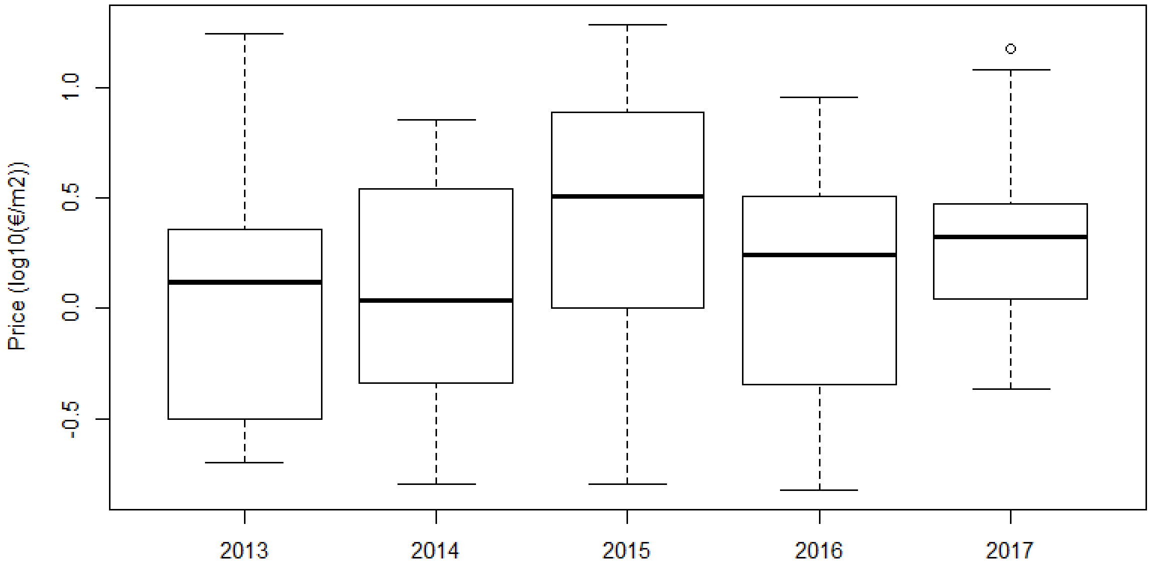

Figure 2. Yearly median, quartiles and range of Sold Residential Land Price (log10(€/m^2)) for Hiiu County (2013-2017).

2.2. Data

There were two datasets used for the dependent variable throughout the statistical analysis of land parcel prices for 1) profit-yielding (in Estonian: maatulundusmaa) land parcels and 2) the residential land (in Estonian: elamumaa) parcels. The dependent variable was provided by the Estonian Land Board. The provided dataset contained the sales of all land parcels without buildings in Hiiu County, both by private sellers and the Land Board themselves, for the 5-year period between 01.01.2013 and 31.01.2017. The sold land prices had been precalculated by the indexated price for 31.12.2017 and divided by the size of the land parcel to give the indexed price per square metre (€/m^2). The independent variables can be grouped into three:

1) the proximity of each land parcel to selected environmental factors: distance in metres from the land parcel to the environmental factor 2) the internal land type of the land parcels: the percentage of each land type per land parcel 3) the soil parameters of the land parcels.

The information of the first two groups were obtained from the Estonian National Topographic database (ETAK) (1:10 000) end the third group from Estonian Soil Map (1:10 000). A dataset of the dependent variables and independent variables was created. Only land parcels designated as 100% profit-yielding, or 100% residential purpose were considered, as were sales where there was only one cadastre registered per sale. Table 1 lists all the 48 variables in each of the two investigation datasets and the abbreviations used in the figures in Annexes 1-2.

Table 1. The measured independent variables and the dependent variable.

Abbreviation	Full Variable Name	Unit
AREA	The size of the land parcel	m^2
D_AIR	Distance to airport	m
D_ARA	Distance to arable land	m
D_BUI	Distance to any building (buildings, foundations, ruins, outside buildings)	m
D_BOG	Distance to bog	m
D_COA	Distance to coast	m
D_DRA	Distance to drainage feature (ditches and dykes)	m
D_ELE	Distance to electrical line	m
D_ENV	Distance to zone of environmental protection	m
D_FEN	Distance to fen	m
D_FOR	Distance to forest	m
D_GRA	Distance to grassland	m
D_GVE	Distance to graveyard	m
D_HAR	Distance to harbour (including ferry port)	m
D_LDF	Distance to landfill	m
D_LCB	Distance to building for living (living and communal buildings)	m
D_ORD	Distance to non-main roads	m
D_PAT	Distance to pathways	m
D_PRO	Distance to production land	m
D_QUA	Distance to quarry	m
D_QKM	Distance to quaking bog	m
D_RIV	Distance to rivers	m
D_RDS	Distance to main roads	m
D_SAN	Distance to sandy beach (sandy land within 100m of the coast)	m
D_SCH	Distance to school	m
D_SHO	Distance to shoreline reeds (reeds within 100m of the coast)	m
D_SPO	Distance to sporting venues	m
D_SPR	Distance to spring (water source)	m
D_WAT	Distance to waterbodies (lakes and ponds)	m
D_WET	Distance to wetlands (any type)	m
D_WLL	Distance to waterlogged land	m
D_WDV	Distance to woody vegetation (excluding forest)	m
P_ARA	Percentage of land parcel with arable land	%
P_BUI	Percentage of land parcel with any building (see D_BUI)	%
P_FIE	Percentage of land parcel with fields	%
P_GRA	Percentage of land parcel with grassland	%
P_HOR	Percentage of land parcel with horticultural land	%
P_PRI	Percentage of land parcel with private land	%

	…continued	
P_PRO	Percentage of land parcel with production land	%
P_WAT	Percentage of land parcel with water bodies	%
P_WET	Percentage of land parcel with wetlands (any type)	%
P_WOO	Percentage of land parcel with woodlands (forest % woody vegetation)	%
AV_CLA	Average percentage of clay soil type in land parcel	%
AV_ROC	Average percentage of rocky soil type in land parcel	%
AV_SAN	Average percentage of sandy soil type in land parcel	%
AV_SIL	Average percentage of silty soil type in land parcel	%
AV_P_FER	Average percentage of soil fertility level in land parcel	%
AV_PR_M	Sold Land Parcel Price (DEPENDENT VARIABLE)	€/m^2

The dependent variable was joined to the land cadastral shapefile, which contained data for land use type. Non profit-yielding and residential land parcels were removed, along with land parcels without soil survey data. This left a joined shapefile and data table containing 10254 100% profit-yielding purpose land parcels, of which 913 had at least one sale registered, and 2693 100% residential purpose land parcels, of which 131 of the residential land parcels had at least one sale registered. From the total of 15930 registered land parcels (cadastral units) in Hiiu County at the end of 2017 (covering 98.6% of the territory (Maa-amet, 2017)), the 100% profit-yielding land parcels counted for 64% and the 100% residential land parcels counted for 17%. Figure 3 shows the location of each land parcel type in the study site.

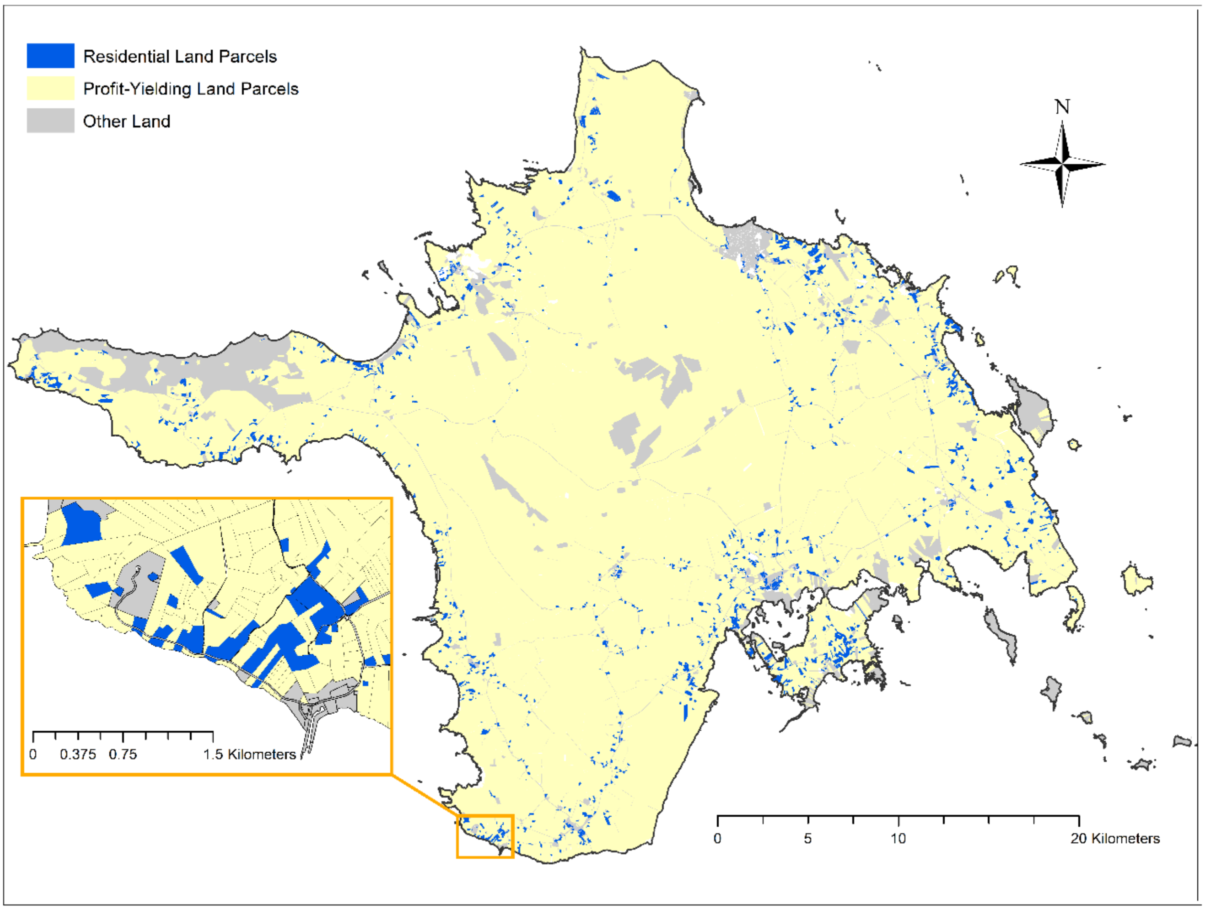

Figure 3. Profit-yielding & residential land parcels in Hiiumaa. (Source: Estonian Land Board).

2.3. Data analysis and preparation

Correlation analyses are used to gain an overview of the variables' relationships (Bolitzer & Netusil, 2000) and focus the analysis and sometimes a bivariate correlation can bring unexpected results (Kim et al., 2007) and refocus the study. However, environmental researchers have to be aware of contradictory findings, which are possibly because of spatial autocorrelation (ibid). In datasets where the variables are non-parametrically distributed, the alternative procedure of Spearman's rank can be used to measure the correlation coefficient. Because environmental data is often non-normally distributed, then Spearman's rank is preferred to study land values. This method was used to further investigate the effects of land use regulation restrictiveness on house and vacant land prices (Ihlanfeldt, 2007).

For profit-yielding land, the strongest overall correlation with sold indexed land price (€/m^2) was distance to arable land (Table 2), with a positive correlation coefficient of 0.27. The relationship indicated that indexed sold land price (€/m^2) increased with increasing distance from arable land. The second strongest correlation with indexed sold land price (€/m^2) was percentage of fields. The relationship indicated that indexed sold land price (€/m^2) increased with a reduced internal percentage of fields. The other variables with statistically significant correlations were only very weakly correlated with indexed sold land price (€/m^2).

Table 2. Statistically significant ($p<0.05$) Spearman correlations with indexed sold land price (€/m^2) for profit-yielding land.

Positive Correlation Coefficients (r)	**Independent Variable**	**Negative Correlation Coefficients (r)**	**Independent Variable**
0.27	Distance to Arable Land	-0.26	Percentage of Fields
0.18	Percentage of Woodland	-0.15	Average Percentage of Soil Fertility
0.14	Distance to Drainage Feature	-0.13	Distance to Coast
0.14	Distance to Landfill	-0.13	Size of Plot
0.13	Distance to Airport	-0.1	Distance to Harbour
0.1	Distance to Production Land	-0.09	Average Percentage of Silty Soils
0.1	Distance to Quaking bog	-0.08	Average Percentage of Clay Soils
0.1	Distance to River	-0.08	Distance to Shoreline Reeds
0.07	Distance to Electric Line	-0.07	Distance to Forest
0.07	Distance to Waterbody		

For residential land, the strongest overall correlation with indexed sold land price (€/m^2) was size of plot, with a negative correlation coefficient of -0.55 (Table 3). This can be considered a moderate relationship which indicated that indexed sold land price (€/m^2) increased with decreasing sold land plot size. The distance variables with weak to moderate positive correlations were distance to drainage feature, distance to arable land and distance to grassland. The direction of these relationships, although weak, indicated that indexed sold land price (€/m^2) increased with increasing distance from them. The distance variables with weak positive correlations were distance to waterbody, distance to landfill, distance to schools and distance to spring. The direction of these relationships, although very weak, indicated that indexed sold land price (€/m^2) increased with increasing distance from them. The distance variables with weak to moderate negative correlations were distance to sandy beach and distance to coast. The direction of these relationships, although weak, indicated that indexed sold land price (€/m^2) increased with reduced distance from them. The variables with weak negative correlations were distance to living or communal building and distance to building. The direction of these relationships, although very weak, indicated that indexed sold land price (€/m^2) increased with reduced distance from them. The other variables with statistically significant correlations were only very weakly correlated with indexed sold land price (€/m^2).

The percentage variable with a weak positive correlation was percentage of woodland. The direction of this relationship, although very weak, indicated that indexed (€/m^2) increased with a larger internal percentage of woodland. The percentage variables with weak negative correlations were percentage of fields and percentage of grasslands. The direction of this relationship, although weak, indicated that indexed sold land price (€/m^2) increased with a smaller internal percentage of fields.

The average soil percentage variables with weak negative correlations were average percentage of soil fertility, average percentage of clay soils and average percentage of silty soils. The direction of these relationships, although very weak, indicated that indexed sold land price (€/m^2) increased with a reduced internal average percentage of these soil properties.

Full correlation matrices can be found in Annex A. These show the variables with statistically significant ($p<0.05$) and non-significant correlation coefficients.

Table 3. Statistically significant ($p<0.05$) Spearman correlations with indexed sold land price (€/m^2) for residential land.

Positive Correlation Coefficients (r)	**Independent Variable**	**Negative Correlation Coefficients (r)**	**Independent Variable**
0.44	Distance to Drainage Feature	-0.55	Size of plot
0.32	Distance to Arable Land	-0.41	Distance to Sandy Beach
0.3	Distance to Grassland	-0.33	Distance to Coast
0.28	Distance to Waterbody	-0.3	Average % of Silty Soils
0.25	% of Woodland	-0.27	Distance to Living/Communal Building
0.24	Distance to Landfill	-0.26	Average % of Clay Soils
0.22	Average % of Sandy Soils	-0.25	Average % of Soil Fertility
0.21	Distance to Schools	-0.23	% of Fields
0.21	Distance to Spring	-0.22	% of Grassland
		-0.21	Distance to Building
		-0.19	Distance to Non-Main Roads
		-0.18	Distance to Harbour

Multicollinearity causes issues within linear regression, which requires the careful selection of the geographic domain from which observations are drawn (Heikkila, 1988). Choumert & Phélina's (2015) study on Argentinian farmlands explained that the heterogeneity of farmlands can cause heteroscedasticity in the residuals of the hedonic price estimation. They used variance influence factors (VIF) to check for and find multicollinearity, noting that their models high number of characteristics was one of the main reasons for this issue. Kim et al. (2007), as well as using correlation analysis, checked for VIF in each of the four models (linear, semi-log, inverse semi-log and double log) that they used as the functional forms in their overall hedonic price model.

The results of the correlation analysis between independent variables (Annex A) were used for preparing the linear regression model. Independent variables with correlation coefficients over 0.7 with other independent variable were detected by the correlation analysis. Of the two variables, the variable with the lower individual correlation coefficient value was generally dropped. Where the variables had the same correlation coefficient, if one was statically non-significant, then it was dropped. However, as the models were repeat-tested, the VIF test within the model showed that some variables, such as distance to school, were scored very high for multicollinearity. Therefore, the variable with the lower correlation score with the dependent variable was selected instead.

To fix the non-parametric distributions of many of the independent variables, including the dependent variable, skewness was calculated. Variables with an absolute skew greater than 0.8 were log-transformed using log10+1 (+1 to avoid division by zero issues). This was a quick technique, which was satisfactorily effective for all but a few of the percentage independent variables, which had spreads of observations from 0-100, but were zero-heavy. In these cases, the transformation improved the skewness, but did not reduce it to within the optimum range of -1 to 1. The variable's dependent variable's distribution was normalised, and its skewness score reduced from 9.06 to 0.19 after a log10 transformation. Boxplots displaying the remaining independent and dependent variables after removal for high collinearity, which were then transformed, can be found in Annex B for both datasets. Finally, to ensure a stable convergence of weight and biases, all variables except the dependent variable were normalised and placed into the same range.

2.4. Multiple Linear Regression

Price prediction academic literature is explanatory and at best advisory, rather than instructional. This can be explained by the different study locations and methodical and data constraints (Mulley, 2014). A price investigation typically involves a regression analysis. There are various forms of regression analyses, which vary in complexity and user input. Regression can be described as "the study of how the conditional distribution of y|x changes with the value of x" (Cook & Weisberg, 2009, p. 37). Linear regression is a preferred starting and comparison point for other regression techniques and is associated with the hedonic model. The hedonic model refers to a range of techniques to identify and estimate price factors by measuring the economic values of goods based on the concept that their value comes from peoples' valuation of their characteristics, or the services they produce (Pyykkönen, 2006), rather than the simple physical object or space. One of the most common ways to assess hedonic pricing is with a regression model. This is a revealed-preference method to determine the relative importance of the variable affecting goods or service prices. A typical linear regression model is shown in equation 1.

$$y = a.x1 + b.x2 + \dots + n.xI$$

(1)

where y is the predicted price, x^1, x^2 and x^i are the property attributes and a, b and n indicate the correlation coefficients of each variable in determining property price (Nur et al., 2017).

A common use of a hedonic regression model is to attempt to quantify the effect of one or several similar environmental factors on the price of land. Many counties, such as Estonia and Lithuania, implemented at least one hedonic regression model in their overall methodology of state mass land evaluation (Bagdonavičius & Deveikis, 2011). Hedonic regression models can require a lot of data preparation and can incorporate many variables. Stepwise linear regression, where the variables are either added (forwards) or removed (backwards) is used to reduce a model with many independent input variables down to the few that contributed most to it. It was employed as part of a hedonic study into the Ankara housing market, where, according to the root mean squared error (RMSE) scores, it was shown to perform best out of the applied regression models, and was able to explain 78% of the variance in the residential sales prices (Hayrullahoğlu et al., 2018).

The spatial limitations of linear models are undoubtedly their weakness in this field, Therefore, in recent years, more advanced methodologies, which include spatial variables that account for the issues of autocorrelation, heterogeneity and nonlinearity have been designed, to absorb the unwanted spatial effect of environmental variables on prediction. A successful example of this was with the reduced of the environmental effect on the 10512 housing dataset of Madrid

(Montero et al., 2018). Where investigations become more complex, other regression techniques are implemented. A quantile regression approach allowed the researchers looking at determinants of the agricultural land prices in Sweden with a focus on location-specific factors, to examine the relative importance of explanatory variables at different point of the distribution of the dependent variable. This enabled a more complete picture of price determinants. (Nilsson & Johansson, 2013).

Geographically weighted regression (GWR) is a complimentary approach (Sá et al., 2011) to global spatial analysis methods, which has been used to study the effects of geographically non-stationary phenomenon such as sub-Saharan wildfire (Sá et al., 2011). The GWR method was compared to standard ordinary least squares (OLS) linear regression and was found to be able to replicate the data by 87% (R^2) due to dealing far better with spatial autocorrelation and discovering local patterns in parameter estimates. Such findings provide improved understandings of multiple variable relationships. Alongside the well-documented and statistically analysed effect of a property's internal features and neighbourhood effects, accessibility has been investigated using GWR. There are certain techniques that can be used to account for spatial dependency on price estimation values. GWR was used to add the geographical component to the regression model to discover that transportation accessibility did improve land prices, albeit moderately, and that within closer proximity, the transportation's negative externalities of noise and pollution were detrimental to land price (Mulley, 2014). In addition, it was concluded that flat rate land taxation to fund such a network would cause cost/benefit imbalances.

Other alternatives to regression can be more manually iterative and participatory, such as the mixed method qualitative and quantitative approach of MCA that successfully provided an insight into the opportunities provided by non-wood forest products in Austria and Finland (Huber et al., 2017). Alternative mapping techniques such as kriging and semivariograms can be used to provide price zones, as demonstrated with the investigation into desirable environmental living spaces in Poland (Cellmer et al., 2012). Finally, there are a range of sophisticated modelling alternatives that have kept interest high in attempting to accurately predict land price. An example is the artificial neural network (ANN) method, which compared to the typical hedonic multiple regression technique, was found to be a better alternative for predicting the house prices in Turkey (Selim, 2009). The other alternative is to combine the benefits of multiple models. ANN, regression and GWR were combined to study house prices in the riverside city of Wuhan, China. A stepwise linear regression hedonic model was found to be inferior to the other two models, which were combined to produce the results. The advantage that linear regression has over ANN is its clear interpretability and ability to analyse the importance of each factor (Wu et al., 2018), which is the primary reason for its use in this thesis.

Before modelling, the sold profit-yielding land and sold residential land datasets were split into training and verification datasets, at a ratio of 7:3. The multiple linear regressions were performed several times. Diagnostic steps were performed with the backward elimination of independent variables one at a time. After every elimination and refit, the model's F-statistic score improved. Variance influence factors (VIF) were checked and variables were removed for having high scores above 5. Distance to quaking bog was removed from both the profit-yielding and residential multiple linear regression models for this reason. For residential land, distance to quarry, distance to fen and distance to bog were also removed for VIF. Next, variables with non-statistically significant p-values (>0.05) were removed from the multiple linear regression model. Outliers with leverage were detected using Cook's distance plots. 9 extreme observation were removed from the profit-yielding regression model and 6 extreme observations were removed from the residential regression model before leverage was satisfied. The model was used to predict the prices on the training dataset, before it was verified by predicting the prices on the unseen verification dataset. In both instances, the predicted and real prices were reverse transformed so that the predictions and errors were in €/m^2. From these, predicted and real prices were used to calculate the correlation (R2) between the real and predicted prices, the root mean squared error (RMSE) score, which is the square root of the average squared residual distance, and the mean absolute error (MAE), which is the average of all the absolute residual errors.

3. Results

The results chapter displays the results of multiple linear regression models for the profit-yielding and residential datasets. The trained multiple linear regression model for profit-yielding land was able to describe 15% of the variance in in the dependent variable. When tested on the validation dataset, and the real and predicted prices reverse transformed, the correlation between the real and predicted sold prices (€/m^2) was 8% (R^2).

The trained multiple linear regression linear model for residential land was able to describe 77% of the variance in the dependent variables. When tested on the validation dataset, and the real and predicted prices reverse transformed, the correlation between the real and predicted sold prices (€/m^2) was 44% (R^2).

3.1. Profit-Yielding Land

The best fitting regression model for profit-yielding land explained 15% (adjusted R^2) of the variation in the training data. The model was statistically significant ($p<0.001$). Residual plots can be found in Annex C. The residuals were reasonably balanced around the line of best fit and all within Cook's distance. The residuals minimum value was -3.63, the maximum 2.34 and the median 0.021. The first quartile was 0.36 and the third quartile was 2.34.

Table 4. Independent variable coefficients of the multiple linear regression model for profit-yielding land (adj. R2 of 0.15).

Coefficient Name	Coefficient Estimate	Coefficient Standard Error	Coefficient T-Value	Coefficient Significance
(Intercept)	-1.51233	0.02942	-51.407	***
Size of Plot	0.12073	0.03018	-4.5100	***
Dist. Arable Land	0.17528	0.02661	6.5870	***
Dist. Harbour	-0.08667	0.02898	-2.9900	**
Dist. Spring	-0.07616	0.02665	-2.8580	**
% Grassland	0.09605	0.03359	2.8600	**
% Waterbody	-0.20144	0.07112	-2.8320	**
Av. % Soil Fertility	-0.11665	0.03590	-3.2500	**

Significance: *** 0.001, ** 0.01, * 0.05

The variables contributing to the model with positive correlations in descending order of significance of contribution (Table 4) were distance to arable land (0.18), size of plot (0.12) and percentage of grassland (0.10). The variables contributing to the model with negative correlations in descending order of significance of contribution were percentage of waterbody (-0.20), average percentage of soil fertility (-0.12), distance to harbour (-0.09) and distance to spring (-0.08).

Table 5 shows the results of the implimentation of the model on the training and verification datasets. The results were taken after both the predicted and real prices had been reverse transformed back to the original measure of €/m^2. The correlation between the real and predicted indexed sold land prices (€/m^2) when the model was used on the training dataset was 15%. The correlation between the real and predicted indexed sold land price (€/m^2) when the model used on the verification dataset to verify the model's ability to predict unseen data was only 8% (Table 4) which is very low predictive ability. The root mean squared error (RMSE) score, which is the square root of the average squared residual distance from the line of best fit was 1.52 for the training dataset and 0.50 for the verification dataset. The mean absolute error (MAE), the average of all the absolute residual errors, was 0.18 for the training dataset and 0.09 for the verification dataset. The RMSE and MAE scores were lower for the verification dataset.

Table 5. Indexed profit-yielding sold land price (€/m^2) prediction results using multiple linear regression model.

Real Price (€/m^2) Vs Predicted Price (€/m^2)	**Training Dataset**	**Verification Dataset**
Correlation (R^2)	0.15	0.08
RMSE	1.52	0.50
MAE	0.18	0.09

3.2. Residential Land

The best fitting regression model for residential land explained 77% (adjusted R^2) of the variation in the training data. The model was statistically significant ($p<0.001$). Residual plots can be found in Annex C. The residuals were reasonably balanced around the line of best fit and all within Cook's distance. The residuals minimum value was -0.54, the maximum 0.45 and the median 0.025. The first quartile was 0.15 and the third quartile was 0.45.

Table 4. Independent variable coefficients of the multiple linear regression model for residential land (adj. R2 of 0.77).

Coefficient Name	**Coefficient Estimate**	**Coefficient Standard Error**	**Coefficient t-value**	**Sig.**
(Intercept)	-0.20163	0.04262	-4.731	***
Dist. Arable Land	0.120730	0.0437	2.762	**
Dist. Coast	-0.303990	0.05239	-5.802	***
Dist. Electric Line	-0.142450	0.04071	-3.499	***
Dist. Zone of Environmental Protection	-0.089380	0.03545	-2.522	*

Dist. Grassland	0.071180	0.03138	2.268	*
Dist. Graveyard	-0.166310	0.04403	-3.777	***
Dist. Harbour	-0.268550	0.03561	-7.54	***
Dist. Main Road	0.160400	0.04143	3.872	***
Dist. Sandy Beach	-0.122080	0.04819	-2.533	*
Dist. Shoreline Reeds	0.152990	0.04898	3.124	**
Dist. Sporting Venue	-0.164850	0.05259	-3.135	**
Dist. Spring	0.204000	0.04323	4.719	***
Dist. Waterbody	0.08484	0.03121	2.719	**
Dist. Woody Vegetation	0.07189	0.03241	2.218	*
Av. % Rocky Soil	0.09165	0.03052	3.003	**
Av. % Soil Fertility	-0.07769	0.02104	-3.693	***

Significance: *** 0.001, ** 0.01, * 0.05

The variables contributing to the model with positive correlations in descending order of significance of contribution (Table 6) were distance to spring (0.2), distance to main road (0.16), distance to shoreline reeds (0.15), distance to arable land (0.12), average percentage of rocks (0.09), distance to waterbody (0.8) and distance to woody vegetation (0.7).

The variables contributing to the model with negative correlations in descending order of significance of contribution (Table 6) were distance to coast (-0.3), distance to harbour (-0.27), distance to graveyard (-0.17), distance to sporting venue (-0.16), distance to electric line (-0.14), distance to sandy beach (-0.12), distance to zone of environmental protection (-0.9), average percentage of soil fertility (-0.8) and distance to grassland (-0.7).

Table 7 shows the results of the implimentation of the model on the training and verification datasets. The results were taken after both the predicted and real prices had been reverse transformed back to the original measure of €/m^2. The correlation between the real and predicted indexed sold land prices (€/m^2) after the model had been used to predict the dependent variable of the training dataset was 81%. The correlation between the real and predicted indexed sold land price (€/m^2) when the model used on the verification dataset to verify the model's ability to predict unseen data was 44% (Table 7), which is still good predictive ability.

The root mean squared error (RMSE) score, 1.37 for the training dataset and 2.2 for the verification dataset. The mean absolute error (MAE) was 0.71 for the training dataset and 1.22 for the verification dataset.

Table 5. Indexed residential sold land price (€/m^2) prediction results using multiple linear regression model.

Real Price (€/m^2) Vs Predicted Price (€/m^2)	**Training Dataset**	**Verification Dataset**
Correlation (R^2)	0.81	0.44
RMSE	1.37	2.2
MAE	0.71	1.22

4. Discussion

4.1. Predicting the indexed transaction price (€/m^2)

The multiple linear regression for the two land types produced very different results. For profit-yielding land, the model's 7 statistically-significant independent input variables that were shown to have an association with the indexed price (€/m^2) of sold profit-yielding land parcels were only able to explain 15% of the variance of the dependent variable of sold indexed land price (€/m^2). When the model was verified, the R^2 score for the correlation between the real sold prices and the predicted prices was only 8%. Therefore, this model was inadequate to predict indexed sold land price (€/m^2) and it was shown to have no potential use for this purpose.

For residential land, the model's 16 statistically-significant independent input variables were shown to have an association with the indexed price (€/m^2) of sold residential land parcels and were able to explain 77% of the variance of dependent variables of sold indexed land price (€/m^2). When the model was verified, the R^2 score for the correlation between the real sold prices and the predicted prices was 44%, which was a lot lower than the testing dataset, but still a good score. Therefore, at this point, the model is merely suggestive, but it has potential for improvement. Therefore, it can be said that the indexed transaction price (€/m^2) of sold land parcels for residential purposes can be partially predicted by using physical environmental variables in Hiiu County.

Despite the disappointing performance of the profit-yielding multiple linear regression model, the he results model were better than the 7% score produced by the first multiple linear regression in Dirgasova et al's (2017) analysis to predict the price of agricultural land in Slovakia. In their study, the variables used also included size of plot, and the distance variables (distance from district city and distance from regional seat). However, when other region-specific factors were added, the model improved significantly (R^2=21%).

The results of the residential multiple linear regression model were better than the 33% R^2 score produced by the multiple linear regression in Cellmer et al's (2012) analysis to predict the price of land zoned for residential or residential development in Poland. However, they only used three independent input variables to gauge scenic value, which were waterbody, forests and elevation.

4.2. Variables association with indexed sold land price (€/m^2)

Coastal

For residential land, correlations were expected between the pleasant environmental variables such as those considered scenic for residential land. The strongest correlation of all input independent variables was distance to coast. The results, which suggest that indexed land price increased as the distance to Hiiu County's famous coastline decreases is no surprise. The coasts are the locations of most of the main island's residences, services and seasonal tourism industry and the variable showed one of the strongest relationships with indexed sold land price (€/m2) for residential land, both in the bivariate Spearman correlations and the linear model. Interestingly, and worth further investigation despite the weak relationship, reduced distance to coast also showed a statistically significant bivariate correlation with increasing indexed sold land price (€/m2) for profit-yielding land. The purpose of such purchases might a financial investment, with coastal land traditionally one of the most valuable (Conroy & Milosch, 2011) With poor soils and heavy woodland, land use is restricted, although as seen in the peripheral coastal zones of urban areas, many permanent and summer residences been constructed since independence (Rivis et al., 2009), and were the land use type to change from profit-yielding to residential, the value of the land should increase substantially (Roka & Palmquist, 1997).

However, such investments are not without risk. Despite the isostatic uplift of Estonia, (which has been modelled to show that Hiiumaa is in the fastest-rising land zone (Kall et al., 2014)), Hiiu County's coastal zones are still predicted to suffer ecological damage, including the shrinkage or disappearance of many of the valuable recreational beaches (Kont et al., 2003). Of course, sand beaches were also positively correlated with indexed sold land price (€/m2) for the residential model. , as well as the bivariate correlations, for which it had a stronger correlation than distance to coast. Thus, unsurprisingly, the benefit of proximity to a sandy beach was not shown to be affected by risk thinking (Thomas, 1999). By contrast, the other major type of coastal frontier, the shoreline reed beds, were shown to negatively affect indexed sold land price (€/m^2) (or do not increase it as much as the easily accessible coastal alternatives). As opposed to the benefits of leisure space (Bolitzer & Netusil, 2000; Cellmer et al., 2012), the coastal reed beds provide a key habitat for breeding fish. Like sandy beaches, they are also threatened and will either retreat inland or disappear along with the complete disappearance of the northern coast's lagoons and orchid-rich calcareous meadows

(Kont et al., 2003). Another potential risk for the Northern and Western coastal regions of Hiiumaa are the drifting oil pollutants and other harmful substances from the busy shipping channels to the island's North into the Baltic Sea (Lehmann et al., 2014). This could prove an interesting further area of research, with beach quality considered an important determinate of coastal property values, alongside beach width, which had the highest coefficient in Pompe & Rinehart's (1995) regression model. However, in comparison to heavily populated beach resort islands, Hiiumaa's coastal villages tend to be located a few kilometres inland from the coast (Kont et al., 2003), with the land parcels connecting with beaches almost mostly profit-yielding woodlands. Therefore, competition can currently be assumed to be high for the intermittent residential land plots with open views of the sea.

Another coastal factor, distance to harbour, was the second strongest correlated of the residential model variables. The variable was also included in the profit-yielding model. Hiiumaa's harbours were variables included in both land types final regression models. Spread around the island, these had a moderate to weak bivariate correlation coefficient with coast. A preparatory investigation into the land prices in Kärdla showed clear indications of higher land prices close to the harbour.

Saturated land

In terms of negative factors, water-related variables were the strongest correlated model variables that were seen to reduce indexed sold land price (€/m^2) for both land types. However, the other water variables had different results. For profit-yielding land, percentage to waterbody was the strongest correlated of all the variables, despite having only shown a very weak but statistically significant positive bivariate relationship with indexed sold land price (€/m^2). Profit-yielding land also had a correlation with distance to spring that showed that reduced distance meant a higher indexed sold land price (€/m^2).

The residential land model also showed that reduced distance to waterbody (the same factor) had an adverse effect on price. These findings contrasted with Snyder et al's (2007) hedonic model of forest land, where waterbody (and river) were found to be positively correlated with price. It also goes against the idea that scenic value or recreational potential adds value to land (Cellmer et al., 2012). However, in a local context, most of Hiiu Counties' waterbodies are small and are located within the forested heartland, which as a closed land type, and removed from view. Additionally, some of the waterbodies are within zones of environmental protection, for which imposed building restrictions could drive down value (Spalatro &

Provencher, 2001). In terms of profit-yielding land, a percentage of waterbody might reduce the forest capacity considerably, for which there is usage potential and financial incentive (Põllumäe et al., 2014; Quiroga et al., 2019).

Distance to spring, which was shown to have an adverse effect on residential indexed sold land price (€/m^2) showed the opposite for the profit-yielding land model. This suggests that fresh water might add value to profit-yielding land, although further investigation would be required to determine the correlation for agricultural land vs forest land.

Woodland vs Arable

For the profit yielding land model, neither of the two main land cover types of wood land and arable land (or variables related to them) were statistically significant in the model and therefore were dropped. This was surprising, given that distance of arable land and percentage of woodland were two of the three highest corelated variables with price. Percentage of fields had the second highest correlation, but this was dropped for high collinearity with both of the other variables. These results are unfortunate and do not allow for any further comparison to be made regarding the value of Hiiu County's agricultural land verses its woodland, and therefore cannot be contextualised within the theory relating to the economic value of woodland. It has been demonstrated across academic literature that price for agricultural or forest land can be modelled with far greater accuracy. Studies that focused on one of the land types and therefore went into greater depth were able to produce far superior results. These can be seen in Maddison's (2000) study into the price of farmland. In this instance, typical hedonic variables such as those related to the size of farmhouse were included and the model was able to account for 60% of the variance in price. Likewise, for forest land, Snyder et al's (2007) study into the determinants of forest land prices was able to determine 53% of the variability in per hectare sale price. In this instance, a range of variables were used that looked at the details of financing methods, road access, river frontage and timber growing volume. Thus, it might be the case that the methodology used in this study was too broad and lacking in detail. There was, however, a positive correlation between profit-yielding indexed sold land price (€/m^2) and percentage of grassland in the model. It was noted that this might be due to the economic benefit that this land type also brings with taxation incentives (EMTA, 2016) . For residential land, the opposite was true, and one can suggest that the restrictions on much of the protected grassland land types drive down price.

Soils

The relationship between average soil fertility percentage is particularly interesting. One might assume, given the history of land price, that higher soil fertility levels would mean higher land prices, due to the higher potential output of the land and therefore return (Verheye, 2009). Additionally, the Estonian Land Board used soil quality as one of the input variables in their hedonic regression model and one would assume that as one of the few variables chosen, there would have been stronger correlations with indexed sold land price (€/m^2) for profit-yielding lands. However, this might have been the case were the soils productive, but where there are poor soil characteristics, including erodibility, wetlands or excessive water permeability, of which Hiiu County has examples of all three, soil suggests agricultural limitations and therefore exerts a negative influence on land price (Roka & Palmquist, 1997). The trend in Hiiu County, as indicated by the results, is that the more expensive plots had lower levels of average soil fertility. Certainly, more local analyses would be required, to investigate whether this is a relationship pattern across the county, or whether it varies by land use type, location or other variables. With most of the investigation area's land type being woody vegetation, and the increasing percentage of this variable having a positive impact on sold land prices, one might assume that the 'poorer' quality soils benefit the woody vegetation, as fast-growing coniferous forests thrive here, which makes the land more valuable for the logging and biomass industries. However, the relationship between increased percentage of woodland and decreasing soil fertility, only shows weak significance. The island's profit-yielding and residential soils, which tend to be around 80% sand, are protected by the trees, including those on the delicate coastal sand dunes (Rivis et al., 2009). The soils may also be affected by privatisation and fragmentation of the land, which is recognised as an issue in Estonia (Jürgenson, 2016) and the Czech Republic, where fragmentation has made smaller plots commercially unviable for owners, who lease them out. There are concerns over neglect and abandonment for soil sustainability from the tenants (Jürgenson, 2016; Sklenicka et al., 2014). In terms of size of plot, Roka & Palmquist (1997) explain that the direction of the trend with price depends on the type of land being analysed. For farmland, one would expect the price per acre to decrease as the land parcel reaches towards a commercially viable size (ibid). However, for this study, size of land plot, which was included in the profit-yielding model showed a positive correlation. This went against the bivariate correlation using Spearman's rank. If the model correlation is more accurate, then the positive relationship could be linked to the potential for profit-yielding land to be converted into another land use type which commands a better price (€/m^2), such as residential. The other soil variable in the residential model was average percentage of rocks. It may be the case that residential land has greater value when firmer, rocky soils are present, since they may be more suitable for construction. This would require further investigation.

Services

The elimination of the capital Kärdla and other urban zones from the dataset, to focus on the areas that had been soil surveyed, likely affected the results for the mostly urban-specific variables: distance to schools, distance to sporting venue, distance to living or communal building, distance to any building, and even distance to main roads. which were expected to show stronger relationships with sold price. However, for residential land, there was still a weak bivariate correlation coefficient with distance to sporting venue. Distance to sporting venue, whilst not showing a statistically significant correlational relationship with indexed sold land price (€/m^2), did provided a significant but weak correlation coefficient in the linear model, showing that it needed to interact with other independent variables to produce a significant relationship. Distance to graveyard, unlike when measured against house prices in Portland (Bolitzer & Netusil, 2000), was statistically significant in the residential land linear model. Finally, given the theory that people wish to live in the countryside, but be able to telecommute and enjoy modern amenities, reduced distance to electric lines unsurprisingly showed a positive relationship with increasing indexed sold land price (€/m^2) for residential land.

Access

In terms of access, the model showed that increased distance from main roads increased price. This goes with the general theory regarding traffic noise (Kim et al., 2007), although there are exceptions, and Hiiumaa does not have a lot of traffic compared to the mainland. In the bivariate relationships, indexed sold residential land price (€/m^2) had a negative relationship with distance to non-main roads. However, this did not appear in the model and so it cannot be said that accessibility appears to affect price, which goes against the notion that it the primary criterion of suburbanisation access and rural settlement merge in Estonia (Roose et al., 2013) and crucial for profit-yielding land access.

Industry

The industrial factors distance to landfill, distance to airport, distance to production land and distance to quaking bog all showed bivariate correlations with profit yielding land indexed sold land price (€/m^2), that unsurprisingly suggested that ideological and environmentally detrimental factors reduce value (Folland & Hough, 1991; Nelson et al., 1992). However, none of these variables made the final linear model. This can be assumed to be due them mostly being monocentric variables (Heikkila et al., 1989) located inland or within relatively close proximity. Distance to landfill, distance to airport and distance to quaking bog all produced instances of multicollinearity. The high correlation between distance to airport and distance to landfill was very clear in the correlation matrices, whereas VIF highlighted distance to quaking bog for removal in the linear model. For residential land, distance to schools had high multicollinearity. This is despite there being 6 schools from which land parcels were measures against and which were not clustered around the island. Spatially, the residential land plots were less spread around the island. They were mostly located around (but not in, due the lack of soil survey map) Kärdla, Käina, Meelste and other coastal areas. This explains why more distance residential variables were removed for multicollinearity than profit-yielding variables (Heikkila, 1988), as they followed similar directional paths between the land clusters to the variables.

4.3. Limitations

The study into profit-yielding land was clearly flawed in terms of its approach. The major issue was the broadness of the topic. If one wishes to be able to confidently explanation the variance in indexed sold land price (€/m^2), then as demonstrated in other studies (Maddison, 2000), the focus should go deeper into one land type and focus on smaller study sites to begin with. For purchase motivations, a multipart model with a participatory approach to find out the profit yielding buyers' motivations, such as MCA, might prove helpful at focusing the dataset, as per the successful study into the opportunities provided by non-wood forest products in Austria and Finland (Huber et al., 2017). One starting place for further research might be the forestry committees as discussed in Põllumäe et al., (2014). To focus a study into Hiiu County's agricultural land, focus could change from the physical farmland soil values, to yields, livestock inventories and a larger range of hedonic variables related to buildings and amenities.

The study into residential land indexed sold land price (€/m^2) was limited due a small dataset, in part due to the quiet economic activity of Hiiumaa. However, whilst it specifically did not consider the typical key drivers of price; housing quality, neighbourhood, societal status, economic status, pollution, political and further legal restrictions, it still obtained positive results. The predictive power without these variables shows that there is a clear correlation between the physical environmental factors and indexed sold land price (€/m^2) in Hiiu County.

This study could have better informed by considering different types of variables, such as social, political, economic, legal, building type and quality, and more environmental. As is the nature of land price modelling, a model cannot simply be borrowed from another spatial location; it needs to be constructed and tuned to the local factors. As discussed, multiple linear modelling is often either a first step in regression modelling or used to compare the results of more powerful models (Wu et al., 2018), all of which come under the process of determining a methodology to provide accurate results for land price prediction. Non-linear models or robust regression models that can better manage non-parametrically distributed variables and zero heavy independent variables are recommended. Likewise, the linear model assumes autocorrelation and doesn't account for the natural variations within the study site. The other major issue was with the dataset. Hiiumaa's low population and relative number of sales compared to the other counties restricted the residential land dataset to 131 sold observations. These were not randomly located around the study site. Most were in coastal clusters on the main island of Hiiumaa. The results would likely have been different and correlations with some of the service variables stronger had Kärdla not been removed from the model due to not having soil survey data. The removal of variables for multicollinearity to improve modelling results could have been avoided with alternative, more spatially aware selections (Heikkila, 1988). Indeed, the models might improve with a smaller selection of independent variables, since environmental characteristics can lead to residual heteroscedasticity (Choumert & Phélinas, 2015).

Alternative regression methods that do not assume autoregression might work better for the selected environmental variables, many of which were heavily skewed and some of which were zero-heavy. The linear model's functional form has been shown be lacking in determining land prices, compared to other model forms. Non-linear perspectives, which don't assume that all influencing factors have a constant influence and ignore spatial location, are often proven to provide better fitting models, which deal better with the dynamic nature of price and which can account for non-linearity of environmental factors (Maddison, 2000; Mulley, 2014; Sklenicka et al., 2014; Wu et al., 2018).

5. Summary

The thesis' findings should be able to inform the Land Board about the associations between the indexed price (€/m^2) of sold residential land parcels and physical environmental variables. If some of the measured variables from this study were included alongside additional variables related to society, local economy, administrative zones, planning restrictions and taxation price, then, there is the potential for this study's dataset to have utilitarian value and aid increase the accuracy of the new valuation zones for residential land. In terms of specific variables, the clear correlation between the indexed sold land price (€/m^2) and the variables distance to the coast and distance to sandy beaches (as opposed to inaccessible reed-bed coastline) could be used to better understand the spatial structure of the demand on land in these regions, and mitigate the environmental damage caused by private residence constructions in fragile environments (Rivis et al., 2009) by administrative and economic measures (Banzhaf & Lavery, 2010). Overall, the study showed that 16 of the environemntal variables, when combined, were able to account for a moderate amount of the variance in the indexed sold residential land prices.

Keskkonnategurite ja maa müügihinna seoste hindamine Hiiumaa näitel

Alex Jarvis

Kokkuvõte

Magistritöö käsitleb keskkonnategurite mõju maa müügihinnale Hiiu maakonnas, EestisVaatluse all olid maaüksused, mille riiklikult registreeritud sihtotstarve on tulundusmaa või elamumaa.

Magistritöö esimene peatükk annab teoreetilise ülevaate keskkonnategurite mõjust maa hinnale. Keskkonnategureid käsitletakse üldises kontekstis, kuid sellises, millel on tähendus neile, kes asuvad uuringu alas st Hiiu maakonnas.

Töö teises peatükis antakse ülevaade uuritavast alast, andmetest ja metoodikast. Uuringus kasutatakse sõltuvate muutujatenas maatulundumaa ja elamumaa müügihinnad ajavahemikus 01.01.2013 kuni 31.12.2017 (€/m^2). Sõltumatud keskkonnategurid arvutati ruumianalüüsi meetodeid kasutades Eesti Topograafilise Andmekogu (ETAK) põhjal kasutades ArcGIS tarkvara. Selle tulemusena loodi andmekogu, mis sisaldab rida iga maatüki kohta ja veergu iga 47 keskkonnamuutuja ning maahinna kohta. Statistilise analüüsi meetodina kasutati korrelatsiooni ja lineaarset regressiooni analüüsi leidmaks seoseid maa hinna ja keskkonnategurite vahel. Statistiline analüüs viidi läbi R-Studios.

Korrelatsioonanalüüsi tulemused näitasid, et kokku 32-l 48-st keskkonnategurist leiti olevat mingi seos müüdud elamumaa hinnaga (€t/m^2). Teisest küljest, kokku ainult 12-l 47-st keskkonnategurist leiti olevat mingi seos müüdud tulundusmaa hinnaga (€t/m^2).

Lineaarse regressiooni tulemused näitaseid, et elamumaa maatükkide puhul moodustas kombinatsioon 16 sõltumatust keskkonnamuutujast statistiliselt olulise mudeli, kus sõltumatute andmetega testides oli R^2 0,44. See tähendab, et mudel kirjeldab ära 44% müüdud elamumaa hinnast (€t/m^2).Maatulundumaa maatükkide puhul moodustas kombinatsioon seitsmest sõltumatust keskkonnamuutujast statistiliselt olulise mudeli, kus sõltumatute andmetega testides oli R^2 0,08 ja seega kirjeldab mudel ära väga väikese osa hinna varieeruvusest.

Lõppkokkuvõttes näitas magistritöö, et Hiiu maakonnas müüdud elamumaa maatükkide hinnal oli palju tugevam seos valitud keskkonnateguritega kui müüdud tulundusmaamaatükkide hinnal.

Acknowledgments

The author would like to thank his supervisors Age Poom and Evelyn Uuemaa for their guidance and considerable support. It has been much appreciated.

The author would like to thank the Estonian Land Board (in Estonian: maa sihtotstarve) and Andres Juss, the Head of Real Estate Valuation Department for providing the dataset containing the dependent variable during his internship at the Estonian Land Board, Tallinn.

Antrop, M. (2000). Geography and landscape science. *Belgeo*, (Special issue), 9–36.

Antrop, Marc. (2004). Landscape change and the urbanization process in Europe. *Landscape and Urban Planning*, *67*(1–4), 9–26.

Bagdonavičius, A., & Deveikis, S. (2011). Mass Appraisal—The Method and an Experience in Lithuania. *FIG (International Federation of Surveyors) Working Week*.

Banzhaf, H. S., & Lavery, N. (2010). Can the land tax help curb urban sprawl? Evidence from growth patterns in Pennsylvania. *Journal of Urban Economics*, *67*(2), 169–179.

Bin, O. (2005). A semiparametric hedonic model for valuing wetlands. *Applied Economics Letters*, *12*(10), 597–601.

Bloodworth, A., Scott, P., & McEvoy, F. (2009). Digging the backyard: Mining and quarrying in the UK and their impact on future land use. *Land Use Policy*, *26*, S317–S325.

Bolitzer, B., & Netusil, N. R. (2000). The impact of open spaces on property values in Portland, Oregon. *Journal of Environmental Management*, *59*(3), 185–193.

Cellmer, R., Senetra, A., & Szczepanska, A. (2012). The Effect of Environmental Factors on Real Estate Value. In *FIG Working Week* (Vol. 20012, pp. 1–13).

Choumert, J., & Phélinas, P. (2015a). Determinants of agricultural land values in Argentina. *Ecological Economics*, *110*, 134–140.

Choumert, J., & Phélinas, P. (2015b). Determinants of agricultural land values in Argentina. *Ecological Economics*, *110*, 134–140.

Cook, R. D., & Weisberg, S. (2009). *Applied regression including computing and graphics* (Vol. 488). John Wiley & Sons.

Daniel, V. E., Florax, R. J., & Rietveld, P. (2009). Flooding risk and housing values: An economic assessment of environmental hazard. *Ecological Economics*, *69*(2), 355–365.

Dirgasova, K., Bandlerova, A., & Lazikova, J. (2017). Factors affecting the price of agricultural land in Slovakia. *Journal of Central European Agriculture*, *18*(2), 291–304.

Doran, J. W., & Zeiss, M. R. (2000). Soil health and sustainability: managing the biotic component of soil quality. *Applied Soil Ecology*, *15*(1), 3–11.

Drescher, K., Henderson, J. R., & McNamara, K. T. (2001). *Farmland price determinants*.

EMTA, Lõõtsa 8a, C. B. |, & tel.: +372 880 0811, 15176 Tallinn |. (2016, February 14). On land tax | Estonian Tax and Customs Board. Retrieved 27 May 2019, from https://www.emta.ee/eng/business-client/excise-duties-assets-gambling/land-tax

Folland, S. T., & Hough, R. R. (1991). Nuclear power plants and the value of agricultural land. *Land Economics*, *67*(1), 30–36.

Fourie, A., & Brent, A. C. (2006). A project-based mine closure model (MCM) for sustainable asset life cycle management. *Journal of Cleaner Production*, *14*(12–13), 1085–1095.

Hayrullahoğlu, G., Aliefendioğlu, Y., Tanrıvermiş, H., & Hayrullahoğlu, A. C. (2018). Estimation of the hedonic valuation model in housing markets: the case of Cukurambar region in Cankaya district of Ankara province. *Ecoforum Journal*, *7*(1).

Heikkila, E. (1988). Multicollinearity in regression models with multiple distance measures. *Journal of Regional Science*, *28*(3), 345–362.

Heikkila, E., Gordon, P., Kim, J. I., Peiser, R. B., Richardson, H. W., & Dale-Johnson, D. (1989). What happened to the CBD-distance gradient?: land values in a policentric city. *Environment and Planning A*, *21*(2), 221–232.

Herrick, J. E. (2000). Soil quality: an indicator of sustainable land management? *Applied Soil Ecology*, *15*(1), 75–83.

Hiiu County Government. (2015). Hiiumaa maakonnaplaneeringu ruumiline analüüs. Retrieved from https://maakonnaplaneering.ee/documents/2845826/19017765/Lisa+4.+MP+Olemasolev+olukord.+Ruumiline+anal%C3%BC%C3%BCs.pdf/ab52152b-f020-4cd6-8964-ff165ca22929

Huber, P., Hujala, T., Kurttila, M., Wolfslehner, B., & Vacik, H. (2017). Application of multi criteria analysis methods for a participatory assessment of non-wood forest products in two European case studies. *Forest Policy and Economics*.

Ihlanfeldt, K. R. (2007). The effect of land use regulation on housing and land prices. *Journal of Urban Economics*, *61*(3), 420–435.

Jacoby, H. G. (2000). Access to markets and the benefits of rural roads. *The Economic Journal, 110*(465), 713–737.

Jim, C. Y., & Chen, W. Y. (2006). Impacts of urban environmental elements on residential housing prices in Guangzhou (China). *Landscape and Urban Planning*, *78*(4), 422–434. https://doi.org/10.1016/j.landurbplan.2005.12.003

Journeaux, P. (2016). *The Effect of Environmental Constraints on Land Prices* (2016 Conference, August 25-26, Nelson, New Zealand No. 260796). New Zealand Agricultural and Resource Economics Society. Retrieved from https://ideas.repec.org/p/ags/nzar16/260796.html

Jürgenson, E. (2016). Land reform, land fragmentation and perspectives for future land consolidation in Estonia. *Land Use Policy*, *57*, 34–43.

Kaasik, A., Sepp, K., Raet, J., & Kuusemets, V. (2011). Transformation of rural landscapes in Hiiumaa since 1956 and the consequences to open and half-open semi-natural habitats.

Ekológia (Bratislava), *30*(2), 257–268.

Kall, T., Oja, T., & Tänavsuu, K. (2014). Postglacial land uplift in Estonia based on four precise levelings. *Tectonophysics*, *610*, 25–38.

Kim, K. S., Park, S. J., & Kweon, Y.-J. (2007a). Highway traffic noise effects on land price in an urban area. *Transportation Research Part D: Transport and Environment*, *12*(4), 275–280.

Kim, K. S., Park, S. J., & Kweon, Y.-J. (2007b). Highway traffic noise effects on land price in an urban area. *Transportation Research Part D: Transport and Environment*, *12*(4), 275–280. https://doi.org/10.1016/j.trd.2007.03.002

Kont, A., Jaagus, J., & Aunap, R. (2003). Climate change scenarios and the effect of sea-level rise for Estonia. *Global and Planetary Change*, *36*(1–2), 1–15.

Lehmann, A., Hinrichsen, H.-H., & Getzlaff, K. (2014). Identifying potentially high risk areas for environmental pollution in the Baltic Sea.

Maa-amet. (2017). Maakatastri statistics | Geoportaal | Maa-amet. Retrieved 15 May 2019, from https://geoportaal.maaamet.ee/index.php?lang_id=1&page_id=506&type=regkat&year=2017&month=12&group=39

Maa-amet. (2019). Statistics on real property transactions | Estonian Land Board. Retrieved 9 May 2019, from https://www.maaamet.ee/en/objectives-activities/land-valuation-and-transactions/statistics-real-property-transactions

Maddison, D. (2000). A hedonic analysis of agricultural land prices in England and Wales. *European Review of Agricultural Economics*, *27*(4), 519–532.

Malczewski, J. (2004). GIS-based land-use suitability analysis: a critical overview. *Progress* *in Planning*, 63.

Mitsch, W. J., & Gosselink, J. G. (2005). *Wetlands* (5th ed.). Wiley.

Montero, J., Fernández-Avilés, G., & Mínguez, R. (2018). Estimating environment impacts on housing prices. *Environmetrics*, *29*(5–6), e2453.

Mulley, C. (2014). Accessibility and residential land value uplift: Identifying spatial variations in the accessibility impacts of a bus transitway. *Urban Studies*, *51*(8), 1707–1724.

Nelson, A. C., Genereux, J., & Genereux, M. (1992). Price Effects of Landfills on House Values. *Land Economics*, *68*(4), 359–365. https://doi.org/10.2307/3146693

Nilsson, P., & Johansson, S. (2013). Location determinants of agricultural land prices. *Jahrbuch Für Regionalwissenschaft*, *33*(1), 1–21.

Nur, A., Ema, R., Taufiq, H., & Firdaus, W. (2017). Modeling House Price Prediction using Regression Analysis and Particle Swarm Optimization Case Study : Malang, East Java, Indonesia. *International Journal of Advanced Computer Science and Applications*, *8*(10). https://doi.org/10.14569/IJACSA.2017.081042

Orford, S. (2017). *Valuing the built environment: GIS and house price analysis*. Routledge.

Page, W., & Rabinowitz, H. (1993). Groundwater contamination: Its effects on property values and cities. *Journal of the American Planning Association*, *59*(4), 473–481.

Põllumäe, P., Korjus, H., Kaimre, P., & Vahter, T. (2014). Motives and Incentives for Joining Forest Owner Associations in Estonia. *Small-Scale Forestry*, *13*(1), 19–33. https://doi.org/10.1007/s11842-013-9237-3

Pompe, J. J., & Rinehart, J. R. (1995). Beach quality and the enhancement of recreational property values. *Journal of Leisure Research*, *27*(2), 143–154.

Profeta, P., Scabrosetti, S., & Winer, S. L. (2014). Wealth transfer taxation: an empirical investigation. *International Tax and Public Finance*, *21*(4), 720–767.

Pyykkönen, P. (2006). *Factors affecting farmland prices in Finland* (Doctoral). University of Helsinki.

Quiroga, S., Suarez, C., Ficko, A., Feliciano, D., Bouriaud, L., Brahic, E., … Lawrence, A. (2019). What influences European private forest owners' affinity for subsidies? *Forest Policy and Economics*, *99*, 136–144.

Reintam, L., Rooma, I., Kull, A., & Kõlli, R. (2005). Soil information and its application in Estonia. *Soil Resources of Europe*, 121–132.

Rhoades, C., Kolka, R., Graves, D., Ringe, J., & Stringer, J. (2001). Carbon sequestration on surface mine lands.

Riigi Teataja. (2012). Katastriüksuse sihtotstarvete liigid ja nende määramise kord – Riigi Teataja. Retrieved 10 May 2019, from https://www.riigiteataja.ee/akt/13058153

Rivis, R., Vilumaa, K., & Ratas, U. (2009). Development of aeolian coastal landscapes on the Island of Hiiumaa, Estonia. *Journal of Coastal Research*, 655–659.

Roka, F. M., & Palmquist, R. B. (1997). Examining the use of national databases in a hedonic analysis of regional farmland values. *American Journal of Agricultural Economics*, *79*(5), 1651–1656.

Roos, A. (1996). A hedonic price function for forest land in Sweden. *Canadian Journal of Forest Research*, *26*(5), 740–746.

Roose, A., Kull, A., Gauk, M., & Tali, T. (2013). Land use policy shocks in the post-communist urban fringe: A case study of Estonia. *Land Use Policy*, *30*(1), 76–83.

Rubinfeld, D. L., & Harrison Jr, D. (1978). Hedonic housing prices and the demand for clean air. *Journal of Environmental Economics and Management*, *5*(1), 81–102.

Ruelle, C., Halleux, J.-M., & Teller, J. (2013). Landscape quality and brownfield regeneration: a community investigation approach inspired by landscape preference studies. *Landscape Research*, *38*(1), 75–99.

Sá, A. C., Pereira, J. M., Charlton, M. E., Mota, B., Barbosa, P. M., & Fotheringham, A. S. (2011). The pyrogeography of sub-Saharan Africa: a study of the spatial non-stationarity of fire–environment relationships using GWR. *Journal of Geographical Systems*, *13*(3), 227–248.

Scott, J. T. (1983). Factors Affecting Land Price Decline. Retrieved from https://www.jstor.org/stable/1240469

Selim, H. (2009). Determinants of house prices in Turkey: Hedonic regression versus artificial neural network. *Expert Systems with Applications*, *36*(2), 2843–2852.

Shi Ming, Y., & Chee Hian, C. (2005). Obstruction of view and its impact on residential apartment prices. *Pacific Rim Property Research Journal*, *11*(3), 299–315.

Šimon, M. (2014). Exploring Counterurbanisation in a Post-Socialist Context: Case of the C zech R epublic. *Sociologia Ruralis*, *54*(2), 117–142.

Sklenicka, P., Janovska, V., Salek, M., Vlasak, J., & Molnarova, K. (2014). The Farmland Rental Paradox: extreme land ownership fragmentation as a new form of land degradation. *Land Use Policy*, *38*, 587–593.

Snyder, S. A., Kilgore, M. A., Hudson, R., & Donnay, J. (2007). Determinants of forest land prices in Northern Minnesota: A hedonic pricing approach. *Forest Science*, *53*(1), 25–36.

Spalatro, F., & Provencher, B. (2001). An analysis of minimum frontage zoning to preserve lakefront amenities. *Land Economics*, *77*(4), 469–481.

Střeleček, F., Lososova, J., & ZDENĚK, R. (2010). The relations between the rent and price of agricultural land in the EU countries. *Agricultural Economics*, *56*(12), 558–568.

Teder, M., Põllumäe, P., & Korjus, H. (2015). Forest Land Ownership Change in Estonia. In *COST Action FP1201-FACESMAP Country Report*. European Forest Institute Central-East and South-East European Regional ….

Thomas, J.-P. C. and A. (1999). A Dynamic Analysis of Land Prices. *American Journal of Agricultural Economics*.

Toppinen, A., Viitanen, J., Leskinen, P., & Toivonen, R. (2005). Dynamics of Roundwood Prices in Estonia, Finland and Lithuania. *BALTIC FORESTRY*, *11*(1), 10.

Tyrväinen, L., & Miettinen, A. (2000). Property prices and urban forest amenities. *Journal of Environmental Economics and Management*, *39*(2), 205–223.

UNECE. (2001). *Land (Real Estate) Mass Valuation Systems for Taxation Purposes in Europe*. Moscow: Federal Land Cadastre Service of Russia on behalf of the UN ECE Working Party on Land Administration.

Vedeld, P., Angelsen, A., Bojö, J., Sjaastad, E., & Berg, G. K. (2007). Forest environmental incomes and the rural poor. *Forest Policy and Economics*, *9*(7), 869–879.

Verheye, W. (2009a). *Land Use, Land Cover and Soil Sciences - Volume IV: Land Use Management and Case Studies*. EOLSS Publications.

Verheye, W. (2009b). The Value and Price of Land. *Land Use, Land Cover and Soil Sciences-Volume III: Land Use Planning*, 82.

Wu, H., Jiao, H., Yu, Y., Li, Z., Peng, Z., Liu, L., & Zeng, Z. (2018). Influence Factors and Regression Model of Urban Housing Prices Based on Internet Open Access Data.

Sustainability, *10*(5), 1676. https://doi.org/10.3390/su10051676

Xiao, Y., Chen, X., Li, Q., Yu, X., Chen, J., & Guo, J. (2017). Exploring determinants of housing prices in Beijing: An enhanced hedonic regression with open access poi data. *ISPRS International Journal of Geo-Information*, *6*(11), 358.

Non-exclusive licence to reproduce thesis and make thesis public

I, Alex Jarvis

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to

reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright,

"Relationships Between Environmental Factors and Sold Land Prices in Hiiu County, Estonia"

supervised by PhD Age Poom and PhD Evelyn Uuemaa.

2. I grant the University of Tartu a permit to make the work specified in p. 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 3.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.

3. I am aware of the fact that the author retains the rights specified in p. 1 and 2.

4. I certify that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Alex Jarvis

Annex A

Table 1. Sold profit-yielding land variables Spearman's rank correlation matrix

AREA																																															
D_AIR	-0.06																																														
D_ARA	0.04	0.2																																													
D_BUI	0.04	0.07	0.31																																												
D_BOG	-0.16	0.43	-0.07	-0.17																																											
D_COA	0.13	-0.25	-0.12	0.35	-0.45																																										
D_DRA	-0.24	0.17	0.29	0.06	0.14	-0.23																																									
D_ELE	0	0.1	0.35	0.53	-0.01	0.18	0.04																																								
D_ENV	-0.02	-0.09	-0.03	0.28	-0.2	0.35	-0.16	0.13																																							
D_FEN	-0.14	0.04	-0.1	-0.15	0.36	-0.23	0	-0.15	0.11																																						
D_FOR	-0.16	-0.08	-0.32	-0.07	0	0.02	0	-0.1	0.02	0.12																																					
D_GRA	-0.02	0.1	0.29	0.69	-0.13	0.33	0.09	0.36	0.22	-0.11	-0.03																																				
D_GVE	0.04	0.09	0.22	0.35	-0.16	0.33	-0.04	0.3	0.19	0.03	-0.09	0.26																																			
D_HAR	0.13	-0.07	0.15	0.15	-0.28	0.34	-0.1	0.11	-0.06	-0.27	-0.1	0.12	0.21																																		
D_LDF	-0.1	0.83	0.23	-0.12	0.62	-0.61	0.27	0.02	-0.31	0.04	-0.11	-0.06	-0.17	-0.17																																	
D_LCB	0.07	-0.01	0.34	0.89	-0.22	0.41	0.04	0.56	0.28	-0.17	-0.08	0.61	0.41	0.21	-0.21																																
D_ORD	-0.11	0.05	0.31	0.39	0	0.15	0.21	0.29	0.11	0	-0.06	0.33	0.15	0.15	0	0.36																															
D_PAT	-0.15	0.16	-0.02	0.31	0.04	0.2	-0.01	0.21	0.23	0.16	0.16	0.3	0.12	-0.02	-0.01	0.28	0.11																														
D_PRO	0.07	0.15	0.48	0.45	-0.38	0.09	0.1	0.36	0.06	-0.4	-0.18	0.32	0.26	0.25	0.1	0.45	0.18	0.02																													
D_QUA	-0.12	0.56	-0.11	0.1	0.56	-0.01	0.04	0.09	0.12	0.37	0.07	0.08	0.16	-0.46	0.4	0.03	0.03	0.28	-0.22																												
D_QKM	-0.19	0.43	-0.07	-0.12	0.62	-0.5	0.19	-0.03	-0.2	0.18	0.08	-0.14	-0.12	-0.33	0.53	-0.17	0.07	-0.02	-0.22	0.42																											
D_RIV	-0.2	0.1	0.17	-0.13	0.25	-0.41	0.18	-0.03	-0.24	0.1	-0.02	-0.02	-0.09	-0.2	0.29	-0.15	0.08	-0.04	0.03	-0.03	0.36																										
D_RDS	0.06	0.06	0.3	0.53	-0.15	0.18	0.04	0.49	0.15	-0.08	-0.13	0.41	0.28	0.16	-0.03	0.55	0.23	0.2	0.34	0.05	-0.12	-0.11																									
D_SAN	-0.05	-0.33	-0.3	0.16	-0.25	0.47	-0.17	0.01	0.32	0.18	0.17	0.1	0.2	-0.18	-0.63	0.19	-0.02	0.22	-0.22	0.25	-0.03	-0.2	0.01																								
D_SCH	0.03	0.33	0.16	0.24	-0.02	0.2	-0.06	0.21	0.13	-0.05	-0.09	0.26	0.35	0.1	0.12	0.22	-0.02	0.24	0.24	0.2	-0.34	-0.11	0.15	0.13																							
D_SHO	0.11	-0.19	-0.01	0.32	-0.45	0.89	-0.17	0.17	0.25	-0.23	-0.01	0.29	0.38	0.41	-0.51	0.39	0.16	0.14	0.17	-0.09	-0.47	-0.36	0.19	0.3	0.14																						
D_SPO	0.04	0.04	0.22	0.1	0.14	-0.13	0.01	0.15	0	-0.12	-0.1	0.03	0.28	-0.05	0.19	0.11	0.06	-0.06	0.26	0	0.18	0.2	0.07	-0.03	0.34	-0.16																					
D_SPR	-0.07	0.28	0.12	0.06	0.21	-0.2	0.04	0.22	0.03	0.08	-0.07	-0.07	0.18	-0.1	0.23	0.08	0.03	0.09	0.14	0.23	0.14	0.05	0.04	-0.13	0.12	-0.2	0.12																				
D_WAT	-0.09	-0.01	0.17	0.54	-0.05	0.18	0.15	0.32	0.19	-0.04	0.01	0.41	0.2	-0.03	-0.07	0.51	0.38	0.16	0.27	0.08	0.07	0.01	0.31	0.16	0.06	0.12	0.1	0.03																			
D_WET	-0.21	0.07	-0.25	0.04	0.18	0.12	-0.07	-0.06	0.27	0.48	0.18	0.08	-0.04	-0.28	-0.03	-0.01	0.05	0.31	-0.24	0.39	0.11	-0.05	-0.04	0.25	-0.01	0.02	-0.18	0.06	0.15																		
D_WLL	-0.16	0.05	-0.18	0.11	-0.01	0.24	-0.08	0.03	0.24	0.16	0.16	0.13	0	-0.22	-0.09	0.09	0.07	0.34	-0.06	0.32	0.05	-0.09	0.03	0.26	0.01	0.13	-0.09	0.11	0.19	0.8																	
D_WDV	0.02	0.16	0.28	0.44	-0.23	0.25	-0.06	0.3	0.11	-0.27	-0.09	0.46	0.25	0.31	0	0.43	0.17	0.12	0.42	-0.11	-0.22	-0.11	0.29	-0.09	0.3	0.26	0.05	0.14	0.28	0.01	0.14																
P_ARA	-0.04	-0.01	-0.07	-0.09	0.02	0	-0.01	-0.03	-0.05	0.05	0.03	-0.03	-0.07	-0.01	0	-0.04	-0.01	0.02	-0.07	0.03	0.02	0.04	-0.04	0.04	-0.03	0	-0.01	-0.01	-0.05	0	0	-0.05															
P_BUI	0.02	-0.03	-0.05	-0.33	0.04	0	-0.04	-0.07	-0.06	0.03	-0.04	-0.2	0.01	0.02	-0.02	-0.18	-0.12	-0.1	-0.08	0.01	0	-0.03	-0.06	-0.02	-0.06	0	-0.03	0.02	-0.12	-0.01	0	-0.08	0.07														
P_FIE	-0.05	-0.18	-0.82	-0.15	0	0.15	-0.26	-0.23	0.03	0.06	0.5	-0.13	-0.18	-0.12	-0.23	-0.19	-0.25	0.12	-0.33	0.08	0.03	-0.14	-0.2	0.28	-0.12	0.06	-0.2	-0.1	-0.08	0.23	0.18	-0.14	0.04	-0.03													
P_GRA	0.01	-0.06	-0.14	-0.48	0.13	-0.28	-0.06	-0.22	-0.21	0.1	-0.05	-0.78	-0.16	-0.08	0.08	-0.41	-0.28	-0.26	-0.2	-0.05	0.09	0.05	-0.26	-0.1	-0.15	-0.24	0.01	0.07	-0.3	-0.1	-0.12	-0.3	0.02	0.28	0												
P_HOR	-0.04	-0.01	-0.07	-0.09	0.02	0	-0.01	-0.03	-0.05	0.05	0.03	-0.03	-0.07	-0.01	0	-0.04	-0.01	0.02	-0.07	0.03	0.02	0.04	-0.04	0.04	-0.03	0	-0.01	-0.01	-0.05	0	0	-0.05	1	0.07	0.04	0.02											
P_PRI	-0.09	0.01	-0.04	-0.25	0.07	-0.05	-0.09	-0.07	-0.07	0.02	-0.03	-0.13	-0.04	-0.02	0.04	-0.25	-0.14	-0.03	-0.1	0	0.01	0.03	-0.07	-0.03	0	-0.07	-0.02	-0.04	-0.15	-0.03	-0.03	-0.08	0.1	0.3	-0.02	0.15	0.1										
P_PRO	0.02	-0.06	-0.07	-0.12	-0.01	0.02	-0.02	-0.1	-0.04	0.02	0.05	-0.05	-0.06	0.06	-0.06	-0.08	-0.05	-0.03	-0.16	-0.04	-0.04	0.02	-0.07	0.01	-0.08	0.05	-0.1	-0.05	-0.13	-0.03	-0.05	-0.05	-0.01	-0.02	0.08	0.01	-0.01	-0.02									
P_WAT	0.05	0.03	0.06	-0.08	-0.03	-0.03	-0.07	0.01	-0.09	-0.04	-0.06	-0.04	-0.05	0.07	0.07	-0.04	-0.06	-0.06	0.02	-0.03	-0.07	-0.01	-0.02	-0.12	0.01	-0.01	-0.01	-0.01	-0.29	-0.06	-0.04	-0.04	-0.01	0.07	-0.07	0.07	-0.01	0.05	0.12								
P_WET	0.19	0.11	0.21	0.02	-0.06	-0.09	0.11	0.05	-0.15	-0.22	-0.08	0.03	0.01	0.12	0.13	0.03	0.01	-0.16	0.15	-0.1	0.02	-0.05	0	-0.16	0	-0.05	0.04	-0.02	-0.03	-0.43	-0.27	0.08	0.05	-0.01	-0.17	-0.01	0.05	-0.01	-0.02	0							
P_WOO	0.09	0.16	0.68	0.32	-0.07	0.06	0.18	0.32	0.11	-0.13	-0.5	0.37	0.2	0.14	0.13	0.33	0.33	0.01	0.36	-0.04	-0.09	0.11	0.3	-0.19	0.15	0.11	0.14	0.05	0.21	-0.09	-0.05	0.23	-0.06	-0.07	-0.81	-0.32	-0.06	-0.05	-0.11	-0.01	-0.02						
S_CLA	0.09	-0.1	-0.11	-0.14	-0.07	-0.04	-0.04	-0.08	-0.18	-0.04	0.09	-0.1	-0.19	0.14	-0.05	-0.13	-0.05	0	-0.07	-0.21	-0.03	0.06	-0.11	-0.06	-0.15	-0.04	-0.14	-0.09	-0.09	-0.07	0.03	-0.04	0.07	0.02	0.14	0.05	0.07	0.03	0.07	-0.05	0.14	-0.17					
S_ROC	0	-0.15	-0.05	0.13	-0.05	0.18	-0.04	0.05	0.13	-0.01	-0.04	0.13	0.06	-0.03	-0.2	0.09	0.04	-0.05	0	0.06	-0.03	-0.2	0.02	0.23	-0.03	0.19	-0.03	-0.11	0.08	0.06	0	0.05	-0.06	0.01	0.02	-0.08	-0.06	-0.03	-0.04	-0.04	-0.1	0.06	-0.42				
S_SAN	-0.06	0.08	0.11	0.18	0.01	0.1	0.02	0.1	0.23	-0.05	-0.09	0.15	0.22	-0.07	0.02	0.18	0.07	0.02	0.12	0.15	-0.01	-0.06	0.14	0.04	0.16	0.1	0.14	0.07	0.12	0.05	-0.02	0.1	-0.07	-0.02	-0.14	-0.13	-0.07	-0.03	-0.08	-0.01	-0.12	0.22	-0.91	0.45			
S_SIL	0.06	-0.13	-0.16	-0.19	-0.02	-0.06	-0.04	-0.13	-0.16	0.07	0.12	-0.15	-0.21	0.08	-0.07	-0.18	-0.06	0.01	-0.16	-0.17	-0.01	0.08	-0.15	-0.02	-0.19	-0.06	-0.16	-0.1	-0.1	-0.02	0.02	-0.13	0.08	0.03	0.19	0.1	0.08	0.05	0.09	-0.06	0.1	-0.22	0.96	-0.41	-0.9		
S_FER	-0.03	-0.09	-0.22	0.12	-0.01	0.18	-0.1	0.02	0.2	0.16	0.13	0.1	0	-0.08	-0.17	0.07	0.1	0.24	-0.16	0.1	0.05	-0.07	0.04	0.19	-0.13	0.08	-0.12	-0.01	0.12	0.29	0.27	-0.06	0.07	-0.06	0.25	-0.16	0.07	-0.02	0	-0.1	-0.18	-0.12	0.27	-0.01	-0.24	0.3	
AV_PR_	-0.13	0.13	0.27	0.03	0.05	-0.13	0.14	0.07	0.03	0.06	-0.07	0.01	0.04	-0.1	0.14	0.04	0.04	0.03	0.1	0.06	0.1	0.1	0.04	-0.06	0.03	-0.08	0.06	-0.03	0.07	-0.02	0	0	0.03	-0.01	-0.26	0.04	0.03	0.01	-0.06	-0.02	-0.01	0.18	-0.08	0.02	0.04	-0.09	-0.15
	AREA	D_AIR	D_ARA	D_BUI	D_BOG	D_COA	D_DRA	D_ELE	D_ENV	D_FEN	D_FOR	D_GRA	D_GVE	D_HAR	D_LDF	D_LCB	D_ORD	D_PAT	D_PRO	D_QUA	D_QKM	D_RIV	D_RDS	D_SAN	D_SCH	D_SHO	D_SPO	D_SPR	D_WAT	D_WET	D_WLL	D_WDV	P_ARA	P_BUI	P_FIE	P_GRA	P_HOR	P_PRI	P_PRO	P_WAT	P_WET	P_WOO	S_CLA	S_ROC	S_SAN	S_SIL	S_FER

Significance: Green cells = p<0.05

Table 2. Sold residential land variables Spearman's rank correlation matrix

	AREA	D_AIR	D_ARA	D_BUI	D_BOG	D_COA	D_DRA	D_ELE	D_ENV	D_FEN	D_FOR	D_GRA	D_GVE	D_HAR	D_LDF	D_LCB	D_ORD	D_PAT	D_PRO	D_QUA	D_QKM	D_RIV	D_RDS	D_SAN	D_SCH	D_SHO	D_SPO	D_SPR	D_WAT	D_WET	D_WLL	D_WDV	P_ARA	P_BUI	P_FIE	P_GRA	P_HOR	P_PRI	P_PRO	P_WAT	P_WET	P_WOO	S_CLA	S_ROC	S_SAN	S_SIL	S_FER
AREA																																															
D_AIR	-0.05																																														
D_ARA	-0.41	0.11																																													
D_BUI	0.08	-0.1	0.17																																												
D_BOG	0.3	0.26	-0.45	-0.04																																											
D_COA	-0.09	-0.39	0.06	0.15	-0.62																																										
D_DRA	-0.18	0.03	0.06	-0.02	0.08	-0.24																																									
D_ELE	-0.13	0.19	0.46	0.22	-0.14	-0.03	-0.11																																								
D_ENV	-0.27	-0.15	0.35	0.09	-0.42	0.4	-0.34	0.16																																							
D_FEN	0.18	-0.04	-0.45	-0.08	0.6	-0.34	0.1	-0.26	-0.21																																						
D_FOR	-0.07	0.05	-0.19	-0.12	0	0.01	0.08	-0.25	-0.09	0.27																																					
D_GRA	-0.44	-0.22	0.62	0.17	-0.44	0.16	0.06	0.4	0.32	-0.39	-0.2																																				
D_GVE	-0.05	0.5	0.1	0.09	0.44	-0.4	0.1	0.17	-0.08	0	-0.06	0.03																																			
D_HAR	-0.02	-0.12	0.14	-0.08	-0.45	0.37	-0.27	0.16	0.11	-0.48	-0.19	0.15	-0.35																																		
D_LDF	-0.15	0.82	0.35	-0.04	0.18	-0.42	0.07	0.36	-0.1	-0.31	-0.22	0.08	0.54	-0.05																																	
D_LCB	0.15	-0.11	0.07	0.81	-0.02	0.22	-0.09	0.04	0.02	-0.06	-0.09	0.03	-0.06	0.03	-0.07																																
D_ORD	-0.01	-0.08	0.21	0.15	-0.18	0.23	-0.29	0.09	0.29	-0.1	-0.01	0.29	-0.07	0.09	-0.01	0.13																															
D_PAT	-0.09	0.07	0.01	0.04	-0.07	0.04	-0.22	0.04	0.07	0.13	0.1	0.16	-0.09	0.14	-0.06	-0.02	0.24																														
D_PRO	-0.17	0.04	0.52	0.04	-0.63	0.24	-0.05	0.49	0.3	-0.59	-0.2	0.52	-0.11	0.46	0.28	-0.07	0.18	0.05																													
D_QUA	0.29	0.34	-0.44	-0.01	0.75	-0.42	0.06	-0.16	-0.18	0.74	0.26	-0.45	0.33	-0.61	0.03	-0.05	-0.09	0.1	-0.56																												
D_QKM	0.23	0.67	-0.36	-0.05	0.71	-0.53	0.18	-0.12	-0.44	0.33	0.1	-0.5	0.42	-0.46	0.47	-0.03	-0.11	-0.02	-0.44	0.67																											
D_RIV	-0.16	0.46	-0.09	0.03	0.39	-0.33	0.26	0.1	-0.26	0.24	-0.01	-0.2	0.27	-0.43	0.46	0.04	-0.04	-0.08	-0.22	0.29	0.6																										
D_RDS	-0.16	-0.02	0.31	0.3	-0.12	0.07	-0.08	0.6	0.13	-0.15	-0.24	0.37	0.08	0.1	0.22	0.29	0.04	0	0.34	-0.25	-0.26	0.16																									
D_SAN	0.33	-0.34	-0.54	0.14	0.25	0.18	0.04	-0.23	-0.14	0.63	0.33	-0.28	-0.28	-0.29	-0.64	0.11	0.03	0.14	-0.41	0.52	0.11	-0.05	-0.13																								
D_SCH	-0.36	0.27	0.51	-0.02	-0.14	-0.09	0.01	0.46	0.13	-0.27	-0.18	0.5	0.4	0.05	0.49	-0.14	0.01	0.02	0.43	-0.27	-0.22	0.14	0.42	-0.47																							
D_SHO	-0.38	0.15	0.4	-0.04	-0.35	0.39	0	0.14	0.27	-0.41	-0.15	0.31	0.26	0.04	0.24	-0.03	0.06	-0.06	0.24	-0.35	-0.12	0.09	0.1	-0.44	0.33																						
D_SPO	-0.11	0.21	0.32	0.14	0.08	-0.07	-0.01	0.42	0.06	-0.22	-0.26	0.36	0.4	-0.11	0.53	0.03	0.08	-0.1	0.36	-0.08	0.03	0.29	0.43	-0.32	0.73	0.21																					
D_SPR	-0.06	-0.04	0.21	0.05	-0.06	-0.16	0.13	0.22	-0.09	0.07	0.02	0.26	0.13	0.13	0.04	-0.07	-0.19	0.05	0.24	-0.04	-0.33	-0.26	0.22	-0.08	0.39	-0.25	0.11																				
D_WAT	-0.23	-0.11	-0.06	0.22	0.04	0.09	0.43	-0.07	-0.07	0.08	-0.01	0.05	0.02	-0.12	-0.02	0.1	-0.17	-0.1	-0.02	0.07	0.02	0.16	0.05	0.18	-0.01	-0.05	0.21	0.11																			
D_WET	-0.2	0.15	-0.07	-0.11	0.07	-0.01	0.06	0.02	0.05	0.27	0.33	-0.07	0.26	-0.39	0	-0.16	-0.02	0	-0.12	0.29	0.14	0.34	-0.03	0.13	0.17	0.26	0.08	-0.03	0.07																		
D_WLL	-0.33	0.14	0.08	-0.09	-0.17	0.17	0.08	0.05	0.13	0.01	0.24	0.09	0.14	-0.26	0.04	-0.11	0.12	0.05	0.03	0.02	0.01	0.39	0.03	0.01	0.25	0.41	0.11	-0.17	0.04	0.86																	
D_WDV	-0.37	-0.01	0.63	0.01	-0.61	0.39	-0.11	0.37	0.33	-0.61	-0.1	0.55	0.06	0.33	0.18	-0.07	0.24	-0.02	0.61	-0.57	-0.49	-0.21	0.23	-0.44	0.54	0.49	0.32	0.21	-0.04	0.1	0.29																
P_ARA	-0.09	-0.13	-0.12	0.03	0.01	0.1	-0.1	0.04	-0.12	-0.14	-0.02	-0.01	-0.12	0.06	-0.12	0.04	-0.07	0.02	-0.01	-0.1	-0.02	-0.14	-0.03	0.08	-0.14	-0.02	-0.14	-0.11	0.06	-0.11	-0.05	-0.03															
P_BUI	0.09	-0.03	-0.19	-0.63	0.11	-0.07	-0.01	-0.22	-0.07	0.16	0.06	-0.2	-0.16	0.07	-0.07	-0.41	-0.04	0	-0.04	0.14	0.04	-0.06	-0.17	0.08	-0.16	-0.09	-0.2	-0.09	-0.13	-0.07	-0.14	-0.18	-0.04														
P_FIE	0.3	-0.22	-0.72	-0.01	0.26	0.12	0.05	-0.27	-0.24	0.34	0.2	-0.31	0	-0.13	-0.37	0.01	-0.11	0.03	-0.34	0.29	0.12	0.02	-0.11	0.51	-0.29	-0.16	-0.17	-0.02	0.21	0.28	0.16	-0.32	0.14	-0.01													
P_GRA	0.35	0.09	-0.47	-0.15	0.26	-0.03	-0.18	-0.24	-0.19	0.28	0.19	-0.83	-0.08	-0.04	-0.12	-0.03	-0.26	-0.12	-0.34	0.31	0.27	0.14	-0.27	0.21	-0.41	-0.25	-0.25	-0.13	-0.01	0.11	-0.04	-0.4	-0.07	0.2	0.21												
P_HOR	-0.09	-0.13	-0.12	0.03	0.01	0.1	-0.1	0.04	-0.12	-0.14	-0.02	-0.01	-0.12	0.06	-0.12	0.04	-0.07	0.02	-0.01	-0.1	-0.02	-0.14	-0.03	0.08	-0.14	-0.02	-0.14	-0.11	0.06	-0.11	-0.05	-0.03	1	-0.04	0.14	-0.07											
P_PRI	-0.07	0.13	-0.05	-0.43	0.03	0.03	-0.03	-0.07	-0.03	0.03	0.07	-0.16	0.03	-0.19	0.04	-0.44	-0.08	-0.01	-0.12	0.12	0.09	0.06	-0.06	0	0.05	0.14	-0.07	-0.07	-0.19	0.19	0.21	0.07	-0.04	0.28	-0.03	0.11	-0.04										
P_PRO	-0.15	-0.1	0.01	-0.07	-0.06	0.11	-0.02	-0.14	0.15	0.07	0.32	-0.01	-0.11	-0.09	-0.14	-0.01	0.13	0.04	-0.15	0.04	-0.07	-0.13	-0.08	0.01	-0.14	0	-0.14	-0.13	0.01	0.09	0.03	0.05	-0.01	-0.04	-0.05	-0.07	-0.01	0.2									
P_WAT	0.17	0.06	0	0.03	-0.1	0.04	-0.07	-0.02	0.08	-0.11	0.05	-0.03	0.03	-0.02	0.02	0.01	0.08	0.01	0.05	-0.04	0.1	-0.07	-0.08	-0.03	-0.06	0.18	-0.09	-0.17	-0.44	-0.05	0.01	-0.01	-0.02	-0.05	-0.01	-0.03	-0.02	-0.03	-0.02								
P_WET	0.14	0.11	0.13	0.18	-0.01	-0.05	0.07	0.05	-0.12	-0.23	-0.05	0	0.13	0.15	0.14	0.15	0.04	0.02	0.08	-0.02	0.1	-0.1	0.02	-0.12	-0.08	-0.02	-0.08	0.02	-0.08	-0.33	-0.26	0.15	-0.02	0	-0.11	-0.02	-0.02	0.03	-0.02	0.26							
P_WOO	-0.34	-0.05	0.66	0.18	-0.34	0.04	0.12	0.28	0.3	-0.35	-0.37	0.68	-0.01	0.13	0.21	0.11	0.15	0.02	0.46	-0.42	-0.37	-0.08	0.32	-0.39	0.45	0.24	0.35	0.17	0.09	-0.24	-0.07	0.39	-0.07	-0.12	-0.61	-0.64	-0.07	-0.27	-0.15	-0.14	-0.1						
S_CLA	0.33	-0.2	-0.47	0.19	0.24	-0.03	0.1	-0.31	-0.23	0.28	0.18	-0.31	-0.1	0.04	-0.32	0.21	-0.05	0.02	-0.2	0.24	0.13	-0.05	-0.08	0.43	-0.52	-0.34	-0.36	-0.02	0.13	-0.18	-0.26	-0.49	0.12	0.08	0.42	0.2	0.12	-0.18	-0.12	0.05	0.17	-0.31					
S_ROC	-0.01	-0.01	-0.16	-0.12	0.19	-0.1	0.18	-0.19	-0.13	0.24	0.09	-0.07	0.14	-0.3	-0.04	-0.11	-0.21	0.03	-0.22	0.2	0.18	0.07	-0.08	0.12	-0.04	0.17	0.11	-0.19	0.15	0.01	-0.03	-0.23	-0.05	0.08	0.09	0.09	-0.05	0.05	0.11	0.08	-0.09	-0.15	-0.03				
S_SAN	-0.37	0.26	0.4	-0.14	-0.22	0.01	-0.02	0.34	0.1	-0.32	-0.21	0.38	0.16	0	0.41	-0.16	-0.08	0.07	0.27	-0.27	-0.09	0.13	0.13	-0.44	0.61	0.39	0.44	0.08	-0.04	0.11	0.23	0.43	-0.13	-0.11	-0.36	-0.27	-0.13	0.05	-0.04	-0.07	-0.18	0.39	-0.84	0.1			
S_SIL	0.36	-0.22	-0.52	0.19	0.26	-0.02	0.08	-0.32	-0.2	0.31	0.17	-0.33	-0.12	0.01	-0.37	0.21	-0.07	0.03	-0.23	0.28	0.14	-0.07	-0.1	0.48	-0.54	-0.36	-0.35	-0.05	0.14	-0.14	-0.24	-0.56	0.15	0.06	0.46	0.23	0.15	-0.2	-0.12	0.06	0.08	-0.32	0.97	0	-0.83		
S_FER	0.06	-0.15	-0.33	0.04	0.16	0.04	-0.22	-0.27	-0.1	0.21	0.27	-0.22	-0.1	0	-0.26	0.1	0.06	0.11	-0.32	0.11	0.07	0.05	-0.12	0.2	-0.27	-0.08	-0.29	-0.06	-0.26	0.04	-0.02	-0.25	0.14	-0.01	0.27	0.17	0.14	-0.06	0.08	0.03	-0.05	-0.29	0.43	-0.06	-0.3	0.4	
AV_PR_	-0.55	0.06	0.32	-0.21	-0.13	-0.33	0.44	0	0.03	-0.12	0.09	0.3	0.09	-0.18	0.24	-0.27	-0.19	-0.04	0.08	-0.16	-0.03	0.15	-0.03	-0.41	0.21	0.13	0.06	0.21	0.28	0.16	0.16	0.16	-0.1	-0.02	-0.23	-0.22	-0.1	0.04	0.11	-0.13	-0.07	0.25	-0.26	0.11	0.22	-0.3	-0.25

Significance: Green cells = $p<0.05$

Annex B

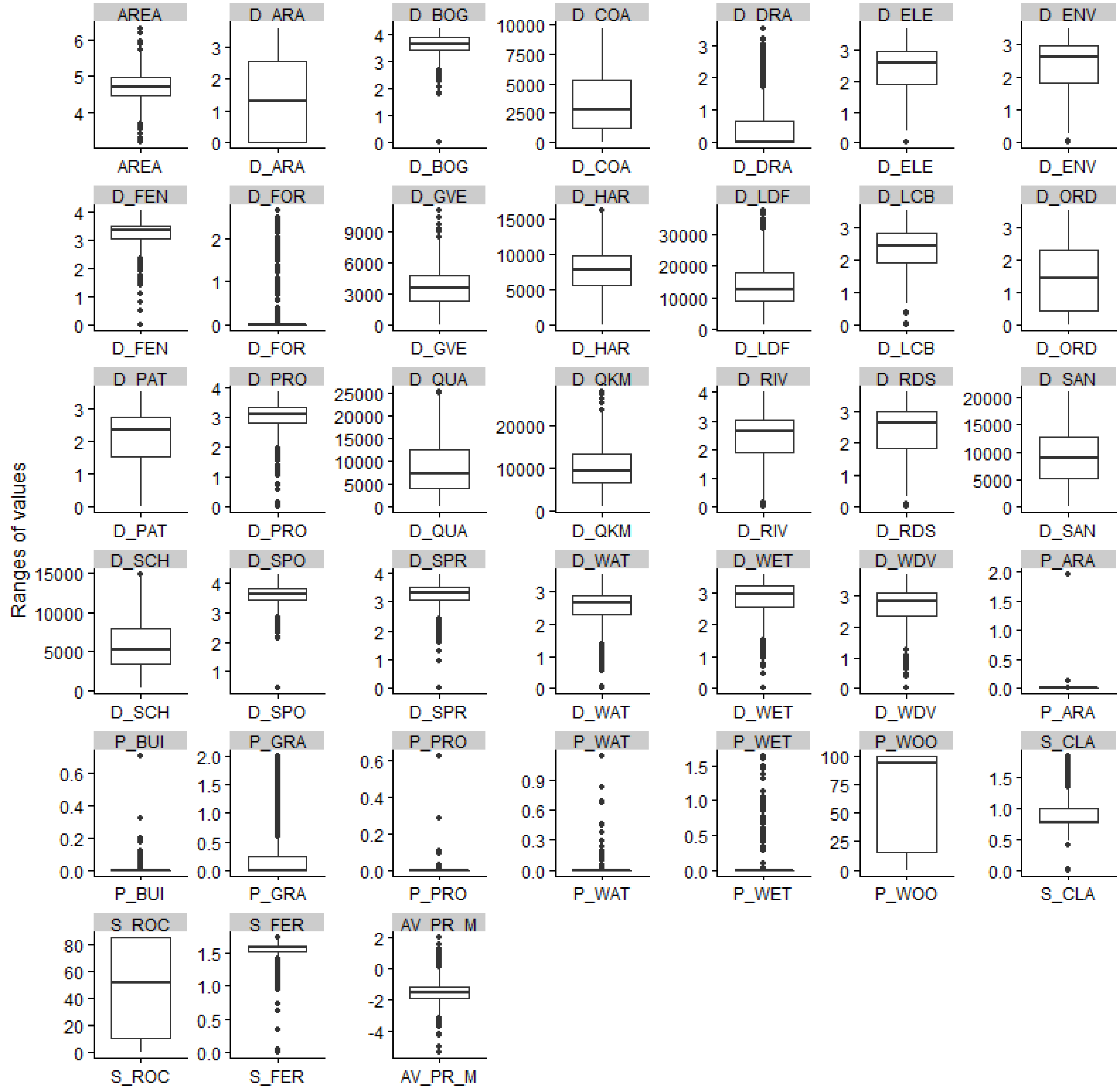

Figure 1. Median, quartiles and range of variables used in multiple linear regression model for profit-yielding land. Those with absolute skew >0.8 had been log10+1 transformed.

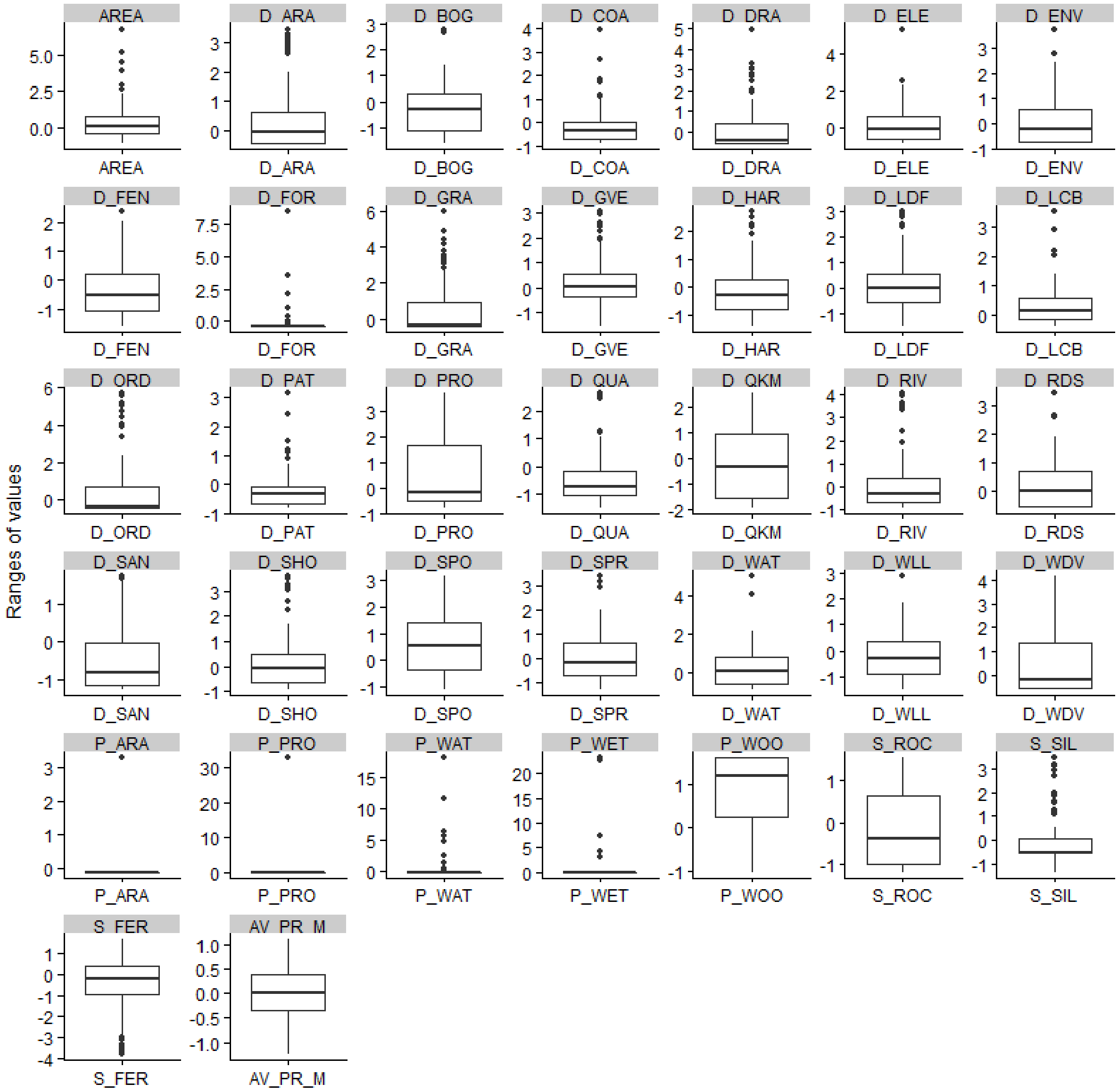

Figure 2. Median, quartiles and range of variables used in multiple linear regression model for residential land. Those with absolute skew >0.8 had been log10+1 transformed.

Annex C

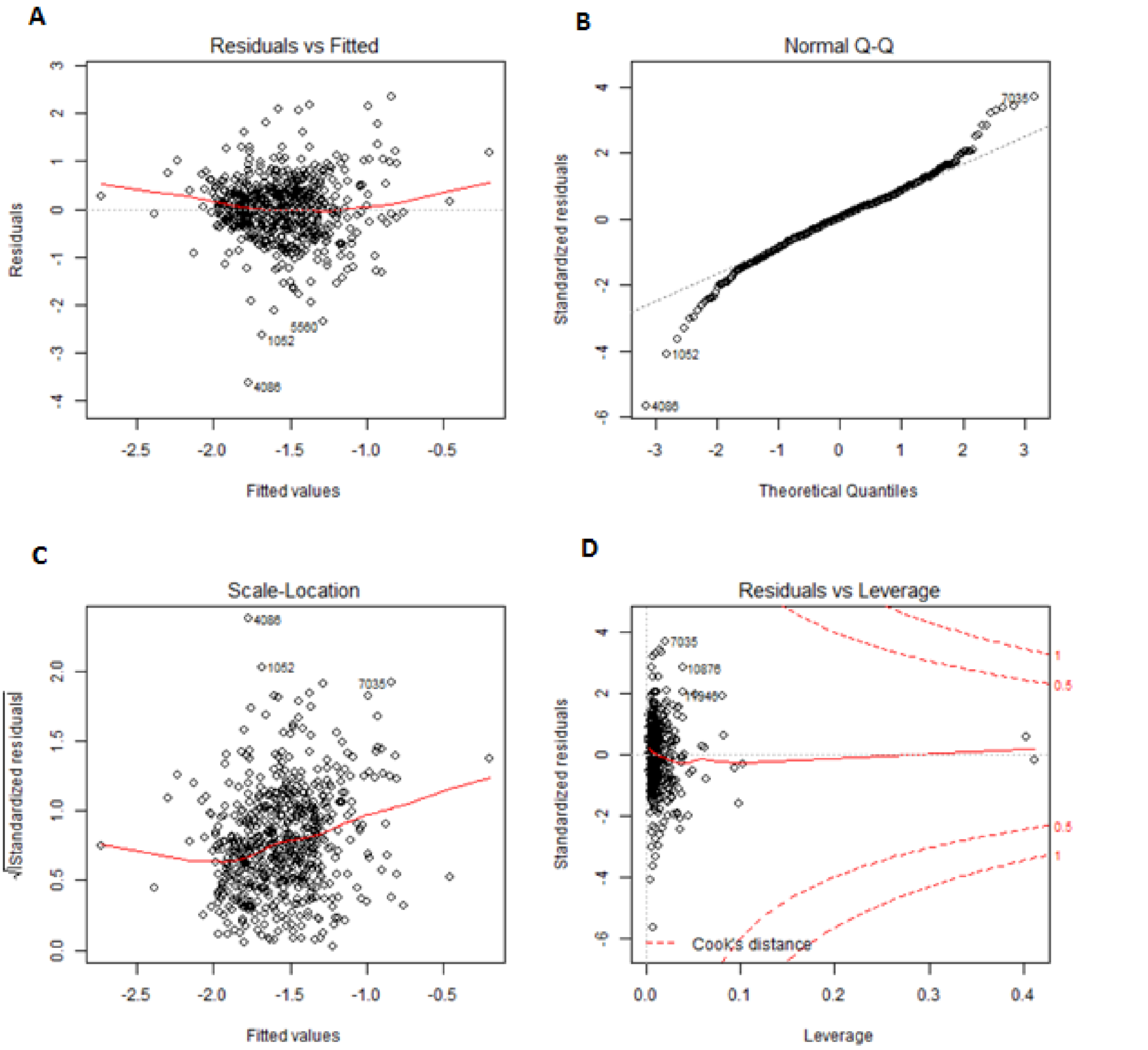

Figure 1. Residual plots for the profit-yielding land multiple regression model. A. the residuals vs the fitted values. B. The residual normal QQ plot. C. The plot of scale-location. D. The residuals within Cooks' distance plot.

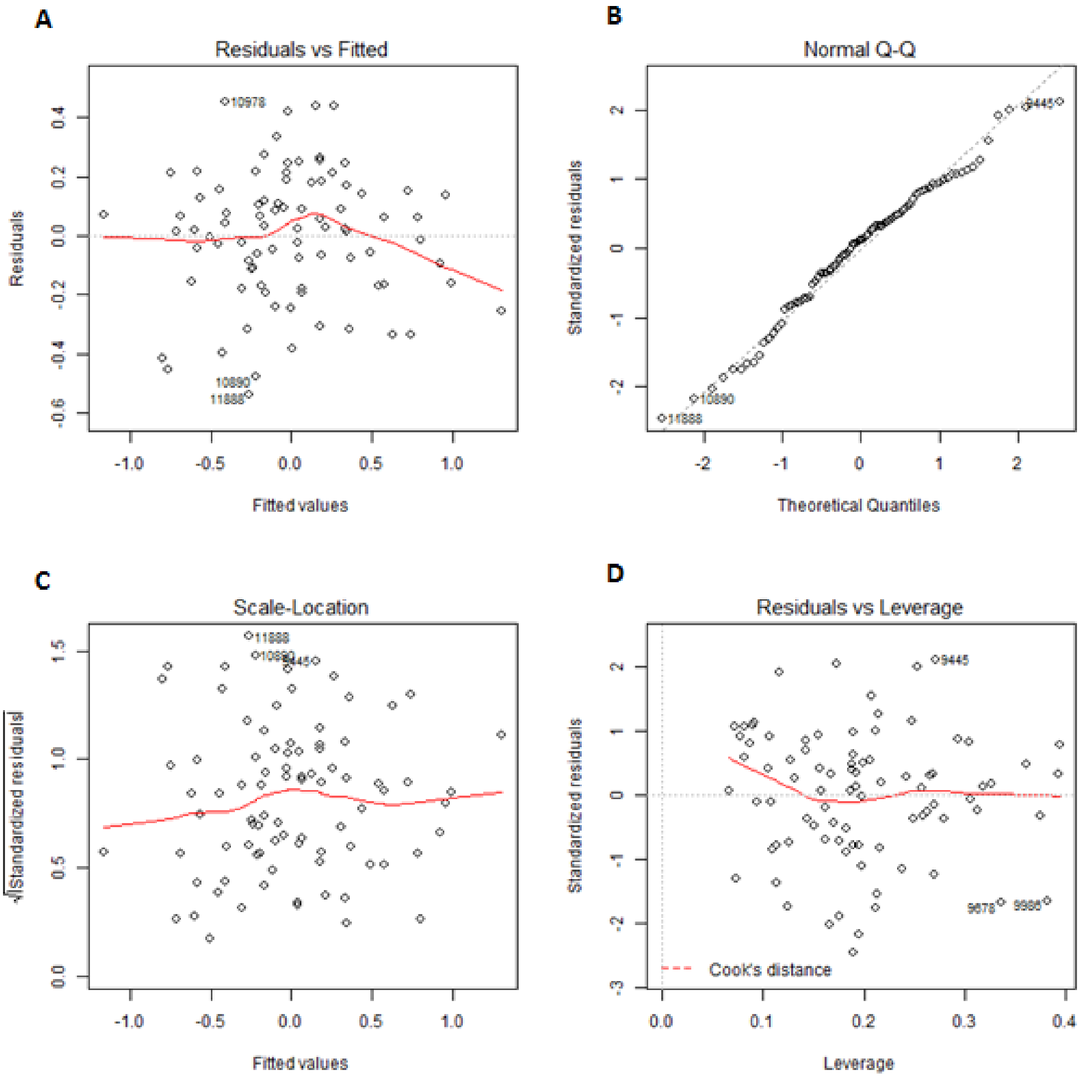

Figure 2. Residual plots for the residential land multiple regression model. A. the residuals vs the fitted values. B. The residual normal QQ plot. C. The plot of scale-location. D. The residuals within Cooks' distance plot.