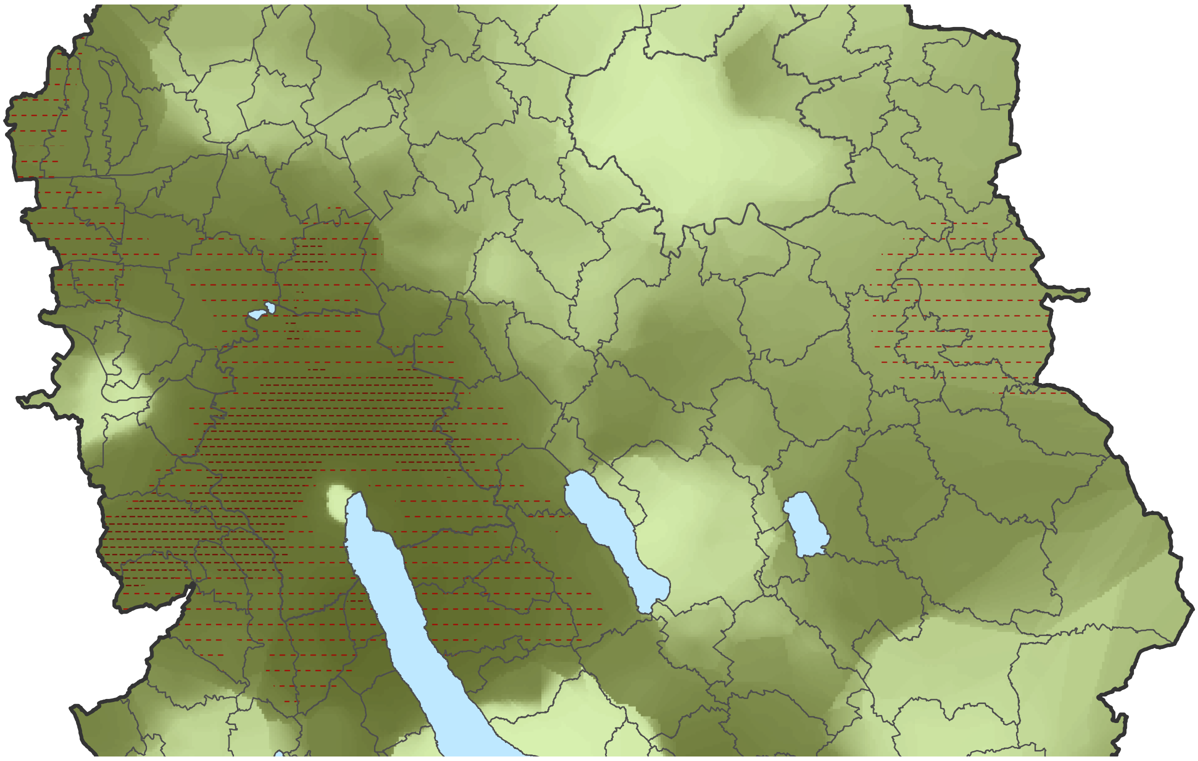


A Hedonic Rental Price Model for the Canton Zurich

Raphael Fuhrer

Supervision:
Prof. Dr. Kay W. Axhausen
Prof. Dr. Adrienne Grêt-Regamey
Dr. Balz Bodenmann
Dr. Ricardo Crespo

Master's Thesis
Spatial Development and Infrastructure Systems **August 2012**

Institut für Verkehrsplanung und Transportsysteme
Institute for Transport Planning and Systems

PLUS
PLANNING OF LANDSCAPE AND URBAN SYSTEMS

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Acknowledgement

I would like to thank Kay W. Axhausen and Adrienne Grêt-Regamey for their support, guidance and valuable comments and suggestions. After every meeting, feedback or information on literature and concepts, I could continue my thesis with more clarity of thought. I am also thankful to Balz Bodenmann whose supervision throughout the Master's Thesis has been very inspiring and encouraging. His wide experience was a great help for me. I thank also Ricardo Crespo for his support, especially with the model estimations.

Many thanks to Patrick Schirmer for his support in the ideation process and data provisioning. Chrisof Zöllig, Ionut Iosifescu and Urs Waldener also supported me in data collection. Finally, I thank Michael Löchl, Sven-Erik Rabe, Noemi Neuenschwander and Silke Rendigs for their helpful recommendations.

Table of Contents

List of Tables

List of Figures

Abbreviations

BAFU	Bundesamt für Umwelt
BfS	Bundesamt für Statistik
CBD	Central business district
CHF	Swiss francs
dB(A)	Decibel, considering the sensibility of the human ear
env.	Environmental
ES	Ecosystem services
GIS	Geographic information system
GWR	Geographically weighted regression
LL	Log-likelihood
SAR	Spatially autoregressive (model)
SSE	Sum of squared errors
STSLS	Spatial two stage least squares (model)
WTP	Willingness to pay

Master of Science Thesis

A Hedonic Rental Price Model for the Canton Zurich

Raphael Fuhrer
Furkastrasse 78
4054 Basel

raphael.fuhrer@gmail.com

August 2012

Abstract

Two types of hedonic rental price models are applied to the Canton Zurich in this thesis. The first one is a spatially autoregressive model, the second one a geographically weighted regression. The focus is on environmental variables including different distance and area based land use related variables as well as visibility, noise and air pollution. In addition, an attempt to introduce qualitative environmental aspects is made. Results show that environmental variables are important and variables related to lakes (distance, area in neighbourhood and visibility), parks (distance), panorama of mountains and noise are most important amongst environmental variables. Tenants are willing to pay higher monthly rents for larger flats (18 CHF/m^2), flats built after 1991 (185 CHF), proximity to parks (100 CHF/km) and lakes (13 CHF/km), for street (2.50 CHF/dB(A)) and railway (3.50 CHF/dB(A)) noise reduction and for visibility of lake surface (0.72 CHF/ha) and mountaintops (10.30 CHF/top). In order to measure ecosystem services, more accurate qualitative data is needed, as the approaches used in this thesis suggest. Nevertheless, results show that improvements in environmental conditions may help to tackle trends such as increasing living space per person or urban sprawl and the corresponding increase in travel distance per person and day.

Keywords

Rental prices, hedonic model, geographically weighted regression, environmental variables

Preferred Citation Style

Fuhrer, R. (2012) A Hedonic Rental Price Model for the Canton Zurich, *Master of Science*, Institute for Transport Planning and Systems (IVT), ETH Zurich, Zurich.

Master of Science Thesis

Ein hedonisches Mietpreismodell für den Kanton Zürich

Raphael Fuhrer
Furkastrasse 78
4054 Basel

raphael.fuhrer@gmail.com

Juni 2012

Zusammenfassung

In dieser Arbeit werden zwei Typen von hedonischen Mietpreismodellen auf den Kanton Zürich angewandt. Der erste Typ ist ein räumliches autoregressives Modell, der zweite Typ eine geographisch gewichtete Regression. Der Fokus liegt auf Umweltvariablen, die verschiedene distanz- und flächenbasierte Landnutzungsvariablen, Sichtbarkeits-, Lärm- und Luftqualitätsvariablen umfassen. Zusätzlich wird versucht, qualitative Umweltaspekte in solche Modelle zu integrieren. Die Resultate zeigen, dass Umweltvariablen grundsätzlich wichtig sind, im Speziellen Variablen in Zusammenhang mit Seen (Distanz, Fläche im Umfeld der Wohnung und Sichtbarkeit), Pärke (Distanz), Bergpanorama sowie Lärm. MieterInnen sind bereit, höhere Monatsmieten zu bezahlen für mehr Wohnfläche (18 CHF/m^2), Wohnungen gebaut nach 1991 (185 CHF), Nähe zu einem Park (100 CHF/km) und einem See (13 CHF/km), für eine Reduktion des Strassen- (2.50 CHF/dB(A)) und Schienenlärms (3.50 CHF/dB(A)) und für die Sichtbarkeit von Seeoberfläche (0.72 CHF/ha) und Bergen (10.30 CHF/Gipfel). Wie in den Konzepten dieser Arbeit sichtbar, sind exaktere qualitative Daten nötig, um Ökosystemdienstleistungen zu integrieren. Es kann trotzdem aufgezeigt werden, dass Verbesserungen in den Umweltbedingungen helfen können, den Trends wie der Zunahme an Wohnfläche pro Person oder der Zersiedelung mit der dazugehöriger Zunahme an Reisedistanzen pro Person und Tag entgegenzuwirken.

Schlagworte

Mietpreise, hedonisches Modell, geographisch gewichtete Regression, Umweltvariablen

Zitierungsvorschlag

Fuhrer, R. (2012) Ein hedonisches Mietpreismodell für den Kanton Zürich, *Masterarbeit*, Institut für Verkehrsplanung und Transportsysteme (IVT), ETH Zürich, Zürich.

1 Introduction

Hedonic rent price models provide information on how different factors, for example floor area, accessibility or view to a lake influence the rent. Thus, such models reveal people's preferences and the attractiveness of different locations. These insights are useful for planners, since the spatial distribution of residential locations is an important factor in traffic demand. They also help to understand the dynamics in settlement, urbanisation and densification. It is possible to derive quantitative indications in order to steer these different processes in space by planning measures.

In this thesis, the focus is on the group of environmental variables. While several studies show their importance in the built environment (e.g. Grêt-Regamey, Neuenschwander, Bakhaus, Wissen Hayek and Tobias (2011), Jim and Chen (2009) and others), they are often not considered in hedonic models (Sirmans, Macpherson and Zietz, 2005). It is the aim of this study to design and integrate environmental variables into a hedonic model for the Canton Zurich. The results should help to identify the importance of environmental variables and visualise hierarchy amongst this variable group, if there is one. Suitable variable specification should be tested and attempts to integrate qualitative information of environmental elements are made. This integration of qualitative information is a prerequisite for the valuation of the more complex ecosystem services. In this study, two types of models are applied. First, spatially autoregressive models as global models and second geographically weighted regression models that provide local coefficient estimates.

In chapter 2, the most important concepts and findings of previous studies are presented. In the next chapter, the selection, design and processing of variables and corresponding data base is explained in detail. In chapter 4, results of both model approaches are given. These results are discussed in chapter 5 and implications are derived.

2 Literature Review

This section gives an overview of different subjects that are relevant for this research. It is important to understand that housing market modelling is based on spatial data, which methods can be applied and what sort of tradition they come from. Furthermore, the terms of ecosystem services and accessibility are introduced and the study area, the Canton Zurich, is presented.

2.1 Spatial Analysis and Modelling

This thesis' research goals and questions concern spatial phenomena. Spatial processes will be analysed and spatial effects modelled. It is thus worthwhile to address this required type of data.

2.1.1 Spatial Data

Spatial data is information that is linked to a pair or a set of X and Y coordinates, or in words of Haining (2009): "Spatially referenced data – that is data where each case has some form of locational co-ordinate attached to it." This widespread type of data occurs not only in disciplines with high affinity towards space, such as geography or planning, it occurs in almost any fields. The daily traffic congestion messages broadcasted on radio, the weather forecast, results of national parliamentary elections, details on the risk of being bitten by a tick, GNP growth comparison between European countries and many other examples all do belong to the type of spatial data, as they all are somehow spatially located. Since new technology makes it increasingly easy to provide collected data with spatial references, spatial data is likely to become even more common (Fotheringham and Rogerson, 2009). When dealing with spatial data, the following three points should be considered (Haining, 2009). First as with all measured data, there might be errors related to false or insufficiently precise measuring as well as to biased capture. The latter leads to systematic errors in results derived from the data set. Systematic in the sense that the measurement process is guided by restrictions, for example economic restrictions. That means it might be considered reasonable to survey in densely populated areas, but not so in sparsely populated regions due to disproportionately higher effort. Second, spatial data usually contradicts the conditions for correct statistical analysis and modelling, namely independently and randomly distributed error sets. Instead, spatial data exhibits autocorrelation. Which means that the variation of a variable in space depends on the distance between the data points (Fortin and Dale, 2009). Data points near to each other tend therefore to be more similar than points farer from each other. Third, information in space can be represented as data in two ways, as a continuous field over which the variable varies or as an object (point, line, polygon) which carries one value of the variable. In both cases, it is important to note the degree of aggregation. High aggregation, in the first case this is large raster cells and in the second case large polygons or rarely segmented lines, leads to intra-areal unit heterogeneity and inter-areal unit heteroscedasticity (Haining, 2009). The higher data aggregation is the smoother data variation becomes. Especially in socio-economic data, say income levels or crime rates, the aggregation level tends to be high, often for privacy or other reasons.

Spatial analysis has to accommodate the peculiarities of this type of data. The first mentioned problem requires special sampling and kriging methods (Delmelle (2009),Yoo and Kyriakidis (2009)). Regarding the other two problems, there are two ways to cope with according to Haining (2009): either working with global models that are fitted to all data or using local statistics which use geographically defined subsets. An example for the latter approach is geographically weighted regression (GWR).

The evolution of spatial analysis has been coupled with progress in quantitative data and methodologies. But unlike in other research areas, there is no precisely identified starting point. Fotheringham and Rogerson (2009) suggest the late 1950s and early 1960s as the period when researches began systematically and analytically to deal with spatial data. Löchl (2010) concludes that since spatial data is so manifold, the approaches having evolved since then are manifold as well and consequently a classification is difficult. By referring to Fischer (2006), he provides the purpose of analysing and modelling spatial data: “Spatial data analysis focuses on detecting patterns and exploring and modelling relationships between such patterns in order to understand processes responsible for observed patterns. In spatial modelling, model outcomes are dependent on the form of spatial interaction between objects in the model, or spatial relationships or the geographical positioning of objects within the model.” It is thus important to recognise patterns and to understand the underlying reasons for them as well as to derive parameters in order to model spatial processes. One of such approaches seeking this goals – in fact the one that is applied in this thesis – is the modelling of housing prices by means of a hedonic pricing model and GWR.

2.1.2 Modelling Housing Markets in the Context of Planning Disciplines

There are different motivations to model housing markets. There can be private sector orientated reasons, such as for banking and insurance companies. But it might well be an issue of public interest, for example when it comes to compensation payments due to zoning, a legislation process to set or alter taxation of properties or in the context of assessing the levels of rental prices (cf. Geiger (2000)). Finally, questions related to urban development may be analysed as described in the subsequent paragraphs.

Modelling housing prices provides information on people's willingness to pay (WTP) for a property at a determined location and may point out the conditions for the observed WTP. Many of these conditions are results of or at least are controllable by planning activities, especially transport, environmental and urban planning. Therefore, modelling can act as a revealed preference method to quantify effects of those planning activities. Vice versa, revealed parameters are valuable findings to justify decisions in relevant planning disciplines.

Recent research focused on improving integrated land use models. These models allow combining transport models and urban simulation. Wegener (2004) gives an overview of different approaches within this research topic. Progressive urbanisation, sustainability concerns, new insights and technical improvements give reason for a recent increase in research activities (Löchl, Bürgle and Axhausen, 2007). There are high expectations to better understand the interactions between spatial development and transport, since this issue has been neglected for quite some time (Bürgle, Löchl, Waldner and Axhausen, 2005). It is also the very goal of housing price modelling, in fact this usually is one module of the integrated models, to describe and decrypt the interactions between urban development, transport and as third environmental aspects.

2.2 Hedonic Pricing Method

The rationale of the hedonic pricing method is to divide a good's price into different components related to the corresponding characteristics of this good. For example, when we pay the rent of an apartment, we do not only pay the property, but also for a certain level of comfort, the degree of accessibility, quietness, etc. In this context, the term «capitalisation» is used (Hilber, 1999) to describe the fact, that the values of various aspects of a consumer good manifest themselves and are summed up in the market price. Vice versa, it is possible to derive monetary values for characteristics, such as quietness, air quality, view, supply of public transport and so forth, from the existing market prices of a consumer good set.

In 1966, Kelvin Lancaster laid the microeconomic foundations for «A New Approach to Consumer Theory» by stating his hypothesis on hedonic or implicit prices. He thought the demand of goods in a new way. Consumers do not demand products as such, but their useful features. Therefore, the demand of goods is based on the utility that is linked to a good's characteristics. The following three assumptions summarise this concept (Lancaster, 1966):

1. *«The good, per se, does not give utility to the consumer; it possesses characteristics, and these characteristics give rise to utility.*
2. *In general, a good will possess more than one characteristic, and many characteristics will be shared by more than one good.*
3. *Goods in combination may possess characteristics different from those pertaining to the goods separately.»*

This implies that, for each person, goods bear a bundle of characteristics in fixed proportions, resulting in an amount of utility, which is based on each person's valuation of these characteristics. In this design, the consumer's preferences and the good's attributes are connected in a linear way.

Sherwin Rosen, who is said to be the second pioneer in paving the way for hedonic pricing, progressed the initial ideas into a theory of hedonic prices and markets (Rosen, 1974). Beneath the interplay of preferences and characteristics, the effects of the supply's structure of products and consumer's budget constraints are taken into account as well. Furthermore, feasible econometric models and techniques to deal with non-linearity are introduced. Following this means, he concludes that «when goods can be treated as tied packages of characteristics, observed market prices are also comparable on those terms» (Rosen, 1974).

The basic form of a hedonic model is the product price regressed on the product's characteristics:

$$P = \alpha \, + \, \beta_1 x_1 + \, \beta_2 x_2 + \cdots + \, \beta_n x_n + \, \varepsilon$$

With the product price, P, a constant, α, the coefficient, β_j, to the characteristic, x_j, and an error term, ε, (adapted from Taylor (2008)). This formula holds true for a certain person i at a time t. The β-parameter can be interpreted as the WTP for a characteristic x (Taylor, 2008). Of course, this simple model can be expanded and transformed to a more complex form aiming in a better fit and higher flexibility (Cheshire and Sheppard, 1995). It is for example common to use the semi-logarithmic form, which leads to non-linear prices of x_n. The natural logarithm of the price is used (Malpezzi, 2002):

$$\ln(P) = \alpha \, + \, \beta_1 x_1 + \, \beta_2 x_2 + \cdots + \, \beta_n x_n + \, \varepsilon$$

Cheshire and Sheppard (1995) used a more elaborate transformation, a Box-Cox model, that considers the type of the characteristics and stabilises variance.

After establishment of the hedonic pricing method by Lancaster and Rosen, this method became widely-used, especially in the real-estate sector (Löchl, 2010). In this context of the built environment, researchers began to integrate models of spatial and urban development. Alonso (1964) presented a model that describes location choice as a function of land price, income, costs of commuting and a factor covering all other expenditures. The price of land and commuting costs both depend on the distance from the city centre. The model assumes a mono-centric city structure, where income is generated in the centre and living takes place in the suburbs. Since then, many researches have applied and consequently merged the classic hedonic pricing method developed by Lancaster and Rosen with Alonso's land use model (Ahlfeldt, 2011a). Hedonic pricing models became spatial when those distance dependent variables, such as accessibility and location, were introduced. Applied hedonic pricing studies focusing on transport and accessibility issues are discussed in more detail in section 2.4.2.

Over the past decades, hedonic pricing method has been applied for a wide range of purposes. As mentioned above, research was done in the field of transport and accessibility. Another important research branch is the valuation of non-market products and services, often directly linked to ecosystem services (section 2.4.1). It is interesting to see for which other diverse reasons hedonic modelling was used. This also points out the flexibility, adaptability and consistency of hedonic price models. Cheshire and Sheppard (2002) used hedonic regression to quantify benefits and costs amongst different income groups in England of land use regulation and could show that public open spaces have a positive effect by reducing income equality. In two recently published papers (Crespo and Grêt-Regamey (2011), Grêt-Regamey and Crespo (2011)), the authors present the implementation of hedonic models in inverse modelling. This approach seeks to define first the desired spatial development in future and then to set out ways of steering it towards the desired state. An other example is the combination of a geographic information system and a hedonic price model (Lake, Lovett, Bateman and Langford, 1998). Hedonic models can also serve as a proposed extension of an existent planning and land use regulating system in order to incorporate price signals in planning decisions (Cheshire and Sheppard, 2005). There have been many other studies conducted using hedonic pricing methods and as a consequence some meta-studies were published recently (Malpezzi (2002), Chau and Chin (2002), Sirmans et al. (2005), Waltert and Schläpfer (2010)).

Despite its frequent and widespread usage, the hedonic pricing method is still faced with objections. According to Malpezzi (2002) they can be sorted in three categories:

1. It is argued that the theoretical footing is too weak; especially the assumption of market equilibrium in housing markets is doubted.

2. The model specification, that is the variable set, the mathematical form and the proper definition of a market, are very likely to be imperfect.
3. It is questioned whether the model fits the purpose.

Following up point two, Taylor (2008) shows in her study the influence of differently designed mathematical forms of the hedonic pricing model. Another important point is the geographic distribution of completed applications of hedonic modelling. During research for this literature review, the bulk of investigated publications reported on studies done in the Anglo-Saxon area, only few in Continental Europe or other parts of the world.

However though objections are there, today the hedonic pricing method is commonly accepted as the method to deal with this research questions, and there have been done many improvements and refinements in this method over the last decades (Malpezzi (2002), Sirmans et al. (2005), Lehner (2011)). One of them is the geographically weighted regression introduced in the next section.

2.3 Geographically Weighted Regression

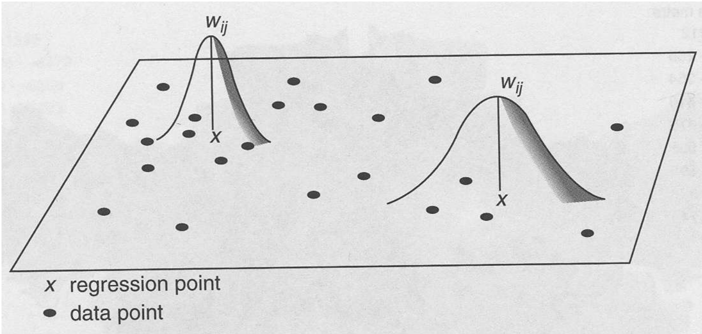
A. S. Fotheringham, C. Brundson and M. Charlton developed geographically weighted regression (GWR) in the late 1990s. At the beginning, there was the observation of non-stationarity in spatial data. And therefore, local forms of spatial analysis were investigated (Fotheringham, Brunsdon and Charlton, 2000). Promoters of local models mention two main reasons for this approach. First, the behaviour and preferences of individuals may change over space. Second, local models give more and precise insights into spatial processes and relationships that, too, may change over space. GWR as a local model incorporates these facts and provides the advantage that it is based on the well-known usual regression framework, in which it integrates local spatial patterns in a intuitive and explicit way (Fotheringham, Brunsdon and Charlton, 2002). According to the authors, the regression is to be rewritten as follows, when geographically weighted:

$$P_i = \alpha(u_i, v_i) \; + \; \beta_1(u_i, v_i)\, x_1 + \; \beta_2(u_i, v_i)\, x_2 + \cdots + \; \beta_n(u_i, v_i)\, x_n + \; \varepsilon_i$$

where all parameters, namely the constant α and the coefficients β_n , are deterministic functions of u and v – the coordinates in space; and the term covers all points i in space, and so builds a continuous surface. This means, at each point i as a pair of u and v, there is a location specific estimation of the parameters α, β_n and ε and the corresponding value of P. With this in mind, the usual regression term (the global regression) is a special case where all β_n are assumed to be constant over space.

Two main questions evolve that are interlinked. Fotheringham et al. (2002) especially address the question of calibrating the model, since there are usually more points to be estimated than measured points, and the choice of a spatial weighting function. The design of such a function is the crucial step in a GWR.

Figure 1 Basic idea of a geographical weighting function



Source: Fotheringham et al. (2002)

The value of *P* at point *i* is derived from measured values around *i*. In other words, the spatial distance from point *i* to another point is converted into a function that defines the similarity of these values between the point of interest, *i*, and the values of the observed points around *i*. There are different possible variants for this function as described in Fotheringham et al. (2002), applying either a strict and so tighter or a looser and so larger 'buffer' around *i*. Now, the reference to calibrating the model becomes clear: When calibrating the model, this is to fill in the gaps between measured points, the tolerance and operation of the weighting function must be decided. Thereby a trade-off between biased values and a high standard error. The larger the subset, chosen by a weighting function, is the lower the standard error becomes, but also the more likelier bias occur. A large amount of measured points within the subset makes the weighting more robust, for the ratio between the number of unknown points (this number is always 1 – since only one point *i* is estimated per subset) to the number of measured points is better. But at the same time, the chance becomes higher that measured points which do not belong to the same spatial pattern for point *i* are included in the weighting for point *i*. Of course, the reverse effect, high standard error and less bias, holds true as well. Such problems are addressed in extensions to the basic GWR model.

As many authors (Fotheringham et al. (2002), Coulson and McMillen (2007), Löchl (2010) Fortin and Dale (2009), Lehner (2011)) state, GWR has its problems as every local model does. Therefore, there are extensions to the GWR as such, as well as for the geographical weighting. A mixed GWR for example allows some of the β_n to be global and some of them to be locally estimated. Regarding weighting, different types of spatial data require their own way of weighting. A count of individuals or items per location for instance is not the same as a deviation from the median income in a set of neighbourhoods, since the first, unlike the latter, can not take a negative sign. This difference in spatial data affects the rationale of the distributional model within the weighting function for the values of a variable given a location in geographical space (Fotheringham et al., 2002).

Though GWR is a relatively new modelling option, there have been some applications recently, predominately in the fields of ecology, wealth and epidemiology. However, application in urban and transport studies is rare (Wang, Kockelman and Wang, 2011). As mentioned above, GWR builds on the classic regression and hence a combination of hedonic pricing methods and GWR seems promising and feasible.

2.4 Relevant Variables

When applying models, one aim is to explain the dependent variable as well as possible. The variable set should include the most powerful explanatory variables. Researchers can also be interested in one particular aspect, but then still it is important to know, which other variables are relevant.

The question of finding relevant variables came up very soon after the establishment of theories of residential location. Richardson, Vipond and Furbey (1973) compared different models representing different compositions of variables. The first model consists of spatial variables, such as distance to city centre, the direction (radial segments of the city), and land value and house price gradient. The second model covers mainly accessibility attributes including public transport. The third one includes housing characteristics only. The last model is compiled of environmental or area preference attributes, for example social class, average age of population, presence of industry. All models range between 0.45 and 0.60 in their R^2, with the forth model at the top end. The authors recommend using mixed models rather than mono causal models by referring to the newly established hedonic approach.

A very comprehensive review of recently published hedonic pricing studies in the USA was done by Sirmans et al. (2005). Publications published over a decade were investigated in order to compile the most common used and statistically significant variables.

Table 1 Top variables by category from previous hedonic pricing studies

Category	Variable	# app.	# posit.	# negat.	# not sig.
1	**Construction and structure**				
	Lot size	52	45	0	7
	Square meter	69	62	4	3
	Age	78	7	63	8
	Number of bathrooms	40	34	1	5
	Bedrooms	40	21	9	10
2	**House internal features**				
	Full baths	37	31	1	5
	Half baths	7	6	0	1
	Fireplace	57	43	3	11
	Air conditioning	37	34	1	2
	Hardwood floors	7	5	0	2
	Basement	21	15	1	5
3	**House external features**				
	Garage spaces	61	48	0	13
	Deck	12	10	0	2
	Pool	31	27	0	2
	Porch	9	5	0	4
	Carport	4	1	1	2
	Garage	4	3	0	1
4	**Environmental-natural**				
	Lake view	5	5	0	0
	Lake front	5	5	0	0
	Ocean view	4	4	0	0
	“good view”	4	3	0	1
5	**Environmental-neighbourhood & location**				
	Location	9	7	2	0
	Crime	7	1	4	2
	Distance to CBD	15	5	5	2
	Golf course	9	9	0	0

	Trees	6	6	0	0
6	**Environmental-public services**				
	School district	10	3	7	0
	% ethnic minority within school district	7	0	5	2
	Public sewer	2	1	1	0
7	**Marketing, occupancy & selling factors**				
	Assessors quality	6	5	0	1
	Assessed condition	8	7	0	1
	Vacant	10	0	9	1
	Owner-occupied	6	4	0	2
	Time on market	18	1	8	9
	Trend	13	2	3	3
8	**Financing**				
	FHA financing	3	0	3	0
	VA financing	3	0	3	0
	Foreclosure	5	0	5	0
	Favourable financing	3	0	0	3
	Property tax	3	0	1	2

Source: Sirmans et al. (2005) with minor adaptions; # app. = number of appearances, # posit. = number of times positive sign, # negat. = number of times negative sign, # not sig. = number of times not significant; FHA = Federal Housing Administration, VA = Veterans Affairs Department

Though the studies did not examine rent prices, but house selling prices in the USA, some general results can be retrieved. Variables directly related to the property are most often applied. The variables related to neighbourhood, public services and accessibility are considered in lower frequency. Particularly, environmental attributes are very rarely mentioned, but they show high significance and consistency in sign, when present.

The attractiveness of specific variables for modellers seems to be partially dependent on the regional context. The variable “central heating” for example is frequently used in England (e.g. Fotheringham et al. (2002), Ahlfeldt (2011b) and others), but not so in Switzerland (Baranzini, Ramirez, Schaerer and Thalmann, 2008). Chau and Chin (2002) point out the cultural differences between Western and Asian housing valuation. After all, however, it is the aim of any hedonic pricing model to find out the different degree of importance within the variable set and, when combined with GWR, to show regional effects. Thus, a priori all variables are relevant. Nevertheless, literature review reveals that some structural variables are of high explanatory power, whereas regarding environmental attributes there is a research gap.

2.4.1 Valuing Environmental and Ecosystem Services

In the United Nations millennium ecosystem assessment (Mooney et al., 2005), ecosystem services are defined as "the benefits people obtain from ecosystems [and furthermore] the human species, while buffered against environmental changes by culture and technology, are fundamentally dependent on the flow of ecosystem services". It is important to recognise the lack of any market for most ecosystem services. Thereby, while ecosystem services are frequently used, often no prices have to be paid. Consequently, no value in monetary amount can be quantified or attached to ecosystems and their services.

According to Engel and Veronesi (2011) ecosystem services can be categorised in four groups. First, provisioning services include all products which can be harvested, collected or extracted from the ecosystem, such as fresh water, food, constructing materials or fuel. Second, benefits such as temperature, climate and flood regulation or water purification are referred to as regulating services. Third, cultural services cover recreational, spiritual, aesthetic, etc. services. Fourth, the category of supporting services, such as soil formation or primary production, is usually called ecosystem functions, as the authors state that the impact on humans is indirect. This structure is helpful to determine whether an ecosystem service should be included in a hedonic model or GWR and which valuing method is accurate. In the context of property rent modelling, items of the second and third group are of interest.

Until recently, there has been little research in valuing ecosystem services in a purely quantitative manner. Most results were generated in a qualitative way. Burgess, Harrison and Limb (1988) for example used several interview und survey techniques to value the urban greens in London. According to their findings they play a central role in cities and improve citizens' life quality. Thereby important and preferred designs of natural settings and social facilities could be identified. In 1992, the value of woodland in Britain was estimated by using a hedonic approach (Garrod and Willis). The fundamental variables described distances between housing areas and certain types of forest. The model results suggested a positive and statistically significant effect on house prices. However, the authors stated that the theoretical foundations for the link from house prices to the hedonic model to the value of woodland was too weak. A major improvement in hedonic modelling concerning ecosystem services happened when GIS data became more easily accessible. One of the first users of GIS were Lake et al. (1998), who applied a GIS-based hedonic model to value noise and visual intrusion caused by new road development. They quantified that each additional dB(A) noise depressed prices by

1.07% and if roads were visible from the front of the house prices were on average 2.5% lower. Boyle and Kiel (2001) provided a literature review of the first wave of hedonic models which measured the impact of environmental externalities. Often, the models combined qualitative and quantitative methods. Regarding air quality, coefficients were often insignificant and sensitive to other included variables. The reasons can be that official measures of air quality may differ from perception of homeowners and probably their lagged and/or weighted information from the past overlaps current information. Regarding water quality, coefficients were statistically significant and of correct sign. But estimated monetary values for good water quality were inconsistent. Best results could be generated, when water quality was assessed on easily perceivable variables, such as clarity. Regarding undesirable land usages, coefficients were statistically significant and of estimated sign. However, estimated WTP for distance from such usages varied very heavily. In this context, information to homeowners on land usage had a considerable impact. When multiple ecosystem services were modelled in one approach, coefficients generally were statistically significant and of expected sign. Though not reported, problems with multicollinearity might be the price for less bias due to omitted variables, the authors concluded.

In the second wave, more types of ecosystem services are addressed. Scenic beauty (example of cultural services) is often valued by using view variables. Fleischer (2012) compared different holiday offers for hotels around the Mediterranean Sea on the internet. According to her model, view on the Sea from your room lead to 10% higher prices. Baranzini and Schaerer (2011) designed a hedonic model to quantify the capitalisation of different land uses around dwellings and the view form dwellings in Geneva. Results showed that both had an impact. Mean values suggested though relatively low impacts of up to 1.08% price increase; but a maximum of 3.15% and 56.7% for *surface* of water-covered area and *view* on water-covered area respectively was observed. In an earlier study, Baranzini and Ramirez (2005) modelled the impact of different noise sources, including aircraft noise, using different noise indexes in their models. Results revealed a 0.7% decrease in rental prices per dB(A).

There are also studies which investigate several benefits of a specific entity. In a Chinese meta-study (Jim and Chen, 2009) effects of trees in urban areas were examined including regulating (air quality, micro climate, water) cultural, and supporting (O_2 generation or CO_2 sequestration) services. Hedonic pricing models played a minor role; results were widely based on bio-physical assumptions and abatement costs. Donovan and Butry (2011) however used a hedonic regression to quantify the effect of urban trees on the monthly rental price of single-family homes in Oregon. The regression provided values of \$5.62 for each additional tree within the lot and \$21 for an additional street tree. Phaneuf, Smith, Palmquist and Pope (2008) combined recreation site choices and property sales prices data to design a hedonic pricing model which models capitalisation effects of urban watershed services. In both respects the capitalisation is statistically significant, it however varies over space and in magnitude. More research is needed.

Recent studies used hedonic methods and examined interrelation to other research questions, mainly concerning the decisions and perceptions of the data generating persons. Binckebanck, Hettenbach, Schwanke and Werner (2011) applied a survey amongst users of online property sites to find out more on the impacts of ecological attributes. They identified location, price and economic efficiency as important variables. Ecological attributes are midrange and depend on the segment (rental versus sale market). They draw the conclusion that ecological attributes are popular and principally valued, but concerning WTP only few people are willing to pay. Ma and Swinton (2011) investigated whether ecosystem services from rural landscapes were capitalised in land prices using hedonics in Michigan. Ecosystem services that support direct use – for example recreational and aesthetic services – are capitalised. Regulating ecosystem services are partially capitalised and other ecosystem services are likely not to be capitalised. The authors said that this is due to lack of awareness, lack of private incentive, or small perceived values. In an other study, the same authors concluded that, regarding land prices, sale prices are superior to appraised values when they are applied in hedonic pricing methods in order to value environmental amenities (Ma and Swinton, 2012). Amrusch and Feilmayr (2009) used a hedonic model to reveal people's WTP for urban features in Vienna, such as parks, plant species diversity and inner-city open space in general. Influences of income and preference structures are explained and showed. Furthermore, the question is addressed whether the existence value[1] of a species, which is not human-held or preference related, can be estimated.

2.4.2 Studies Related to Transport and Accessibility

Beneath and since Alonso's approaches (1964) there have been a variety of studies based on variables from the transport and accessibility sector. Gibbons and Machin (2004) valued rail access using a transport innovation, namely the extensions of the Jubilee Line and Dockland Light Railway, in London in the late 1990s. Some households in the areas of interest were affected by changes in accessibility for new and nearer stations; some others were not affected. Results of the hedonic model show that "household value rail access and that these valuations are larger compared to the valuations of other local amenities" (Gibbons and Machin, 2004). Similarly, Debrezion, Pels and Rietveld (2010) measured the effect of railway access on house prices in the Netherlands. Their hedonic model was designed to control for structural,

[1] The existence value is the benefit people gain for simply knowing that a species exists and is not extinct. The

spatial (the degree of urbanity) and environmental factors. The authors identified the following variables as relevant: the distance to a rail station, the distance to a railway line, and the railway station service quality (intervals, position in network). Their study was divided into two models, the first one considered the nearest station, the second one the most frequently chosen station; both models were designed semi logarithmically and the log-values of access variables were used. The authors drew the conclusion that the second model outperformed the first one, especially in more urbanised regions. In another study, the general impact of accessibility on land price was examined (Ahlfeldt, 2011a). The intention of the researcher was to test and compare different forms of accessibility in terms of definition and modelling. Mainly mono-centric versus poly-centric models using, in the latter case, gravity employment accessibility measures were compared in Berlin. Alonso’s classic mono-centric model, which basically describes the trade-off between the negative land price gradient versus higher commuting costs, satisfies for general model purposes. However, gravity-based models explain the observed land price gradient more accurately, since effects such as congestion or network design are taken into account. Eventually, the transport geography must be carefully modelled though.

Atkinson-Palombo (2010) point out the importance of how the public transport station is defined. The main distinction was between walk-and-ride (WAR) and park-and-ride (PAR) stations. By using a hedonic model to show effects from light rail transit supply on house prices, they described the interaction of station type (WAR vs. PAR) with the overlaying zoning plan determining the land use and city structure within a certain neighbourhood. The results of a case study in Phoenix describe capitalisation effects; in WAR neighbourhoods price increases for single family houses and condos of 6% and 20% respectively could observed, in PAR neighbourhoods no effects occurred. Therefore, the authors conclude, the effects depend on the land use type (residential, mixed use, etc.) and finally on zoning. Thus when new infrastructure is built, they recommend matching zoning in order to steer development in adjacent areas. There was also a case study done in and around Manchester about 1990, when a light rail system improved the situation for commuters. There, Forrest, Glen and Ward (1996) applied a hedonic property price method which suggested that train stations did not have an effect on property prices. In some cases, stations even seemed to have a negative effect. The research team reduced these facts to the special situation in Manchester area and data correlation in their model. Henneberry (1997) investigated the effects of a tramway introduction in Sheffield. A hedonic model was designed for price observations some years prior to the tramway construction, during construction, and short time after commencement of operation. Prior to construction, slightly higher prices alongside the planned corridor were observed and can be explained due to expectations of improved accessibility. Shortly prior to operation, the opposite effect was measured. According to the author this happened due to awareness of noise issues during construction. Shortly after commencement of operation, hardly any effect was present and it could be shown that capitalisation did not depend on the distance to stations, but followed a rather uniform price effect. The conclusion is that capitalisation of transport innovations needs time, especially the spatial adjustment.

Other models dealt with car accessibility. Chernobai, Reibel and Carney (2011) modelled effects of new highways, focusing on price development variation over space and time. Results revealed non-linearities in spatial and temporal gradients. Maximum house price appreciation occurred at moderate distances from the highway. Lower price increases were observed nearer and farther away. This pattern fades away in the following years after commencement of operation.Any effects were observed prior to construction. The authors state that therefore the housing market is not fully efficient, because the information of the forthcoming construction was not incorporated in sales. Boarnet and Chalermpong (2001) used hedonics to measure impacts of the construction of toll highways in California. They found strong evidence for accessibility improvement premiums amongst households. Such WTP influences urban structures and leads to induced traffic, concluded the researchers. This means, higher highway capacity and improved accessibility boost house building at affected locations, thus more traffic occurs. Yiu and Wong (2004) analysed whether expectations of a new tunnel – meaning a high improvement of accessibility – are reflected in payments for housing prior to completion of the structure by applying a hedonic regression. Unlike Henneberry (1997) they found large expectation effects being present and presented a land selling scheme operated by the government, which would allows the government to finance such a tunnel project in advance.

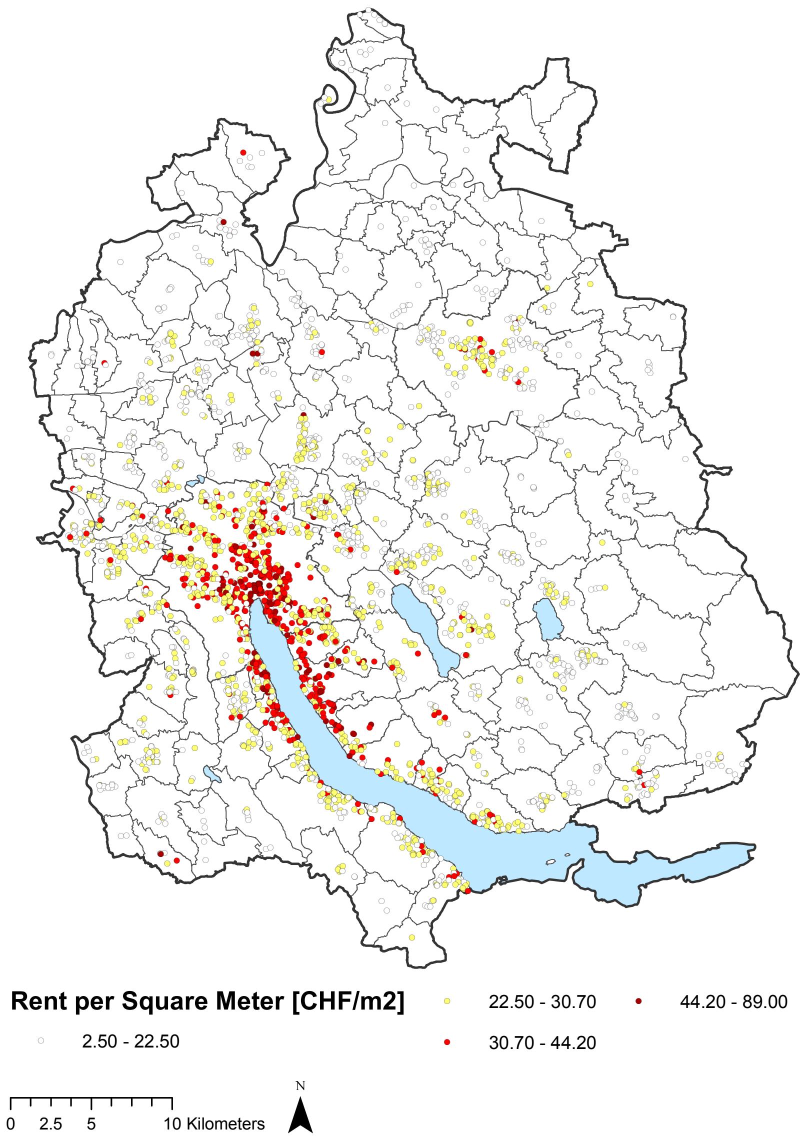
In some cases hedonic pricing methods serve as a component in a more extensive model. Ahlfeldt (2011b) aimed in the development of a better model to predict property price effects due to transport innovations. This model considers gravity-based labour market accessibility and a transport decision model relying on urban rail network and mode switching parameters. It was applied in London when the Jubilee Line and Dockland Light Railway were extended. The hedonic approach was used in order to check for the model's performance and to identify the effect of transport costs. Batty et al. (2011) developed SIMULACRA, a framework for modelling different urban development scenarios. Basic components are amongst others urban economy, transport cost, income and house price. The hedonic is connected with a location choice model. Its application is fast and easy as it is designed in a desktop user interface way. Gibbons and Machin (2008) conducted a meta-study on hedonic models with focus on school quality, transport infrastructure and crime. It was shown that all three variable groups have a large impact on house prices and are highly policy relevant in the UK. Therefore, a carefully researched and elaborately designed hedonic study can offer credible estimates and guidance for policy makers, the authors concluded.

2.5 The study area: The Canton Zurich

2.5.1 The Market Situation in the Canton Zurich

The following information are taken from the official statistics by the Canton Zurich (Hofer et al., 2011). The Canton Zurich is the largest canton in Switzerland with 1.37 million inhabitants by the end of the year 2010. Forecasts predict a further growth in population, workplaces as well as in wages. 42% of the canton's area is used for agriculture, 30% is forest and 22% settlement area. 1.8m^2 of soil is sealed every second. This means, the consumption of land is much higher than in the Swiss average. On average, a person has 45m^2 living area, tenants have 41m^2. In 2009, 7118 new flats were built. With its high dynamic, Zurich is Switzerland's biggest city and the most important economic centre. The second largest city in the canton is Winterthur, other important areas are the Limmattal (northwest of Zurich), the Glattal (northeast of Zurich) including the airport and the two lakesides of Lake Zurich. Figure 2 shows the amount of rent at the data points available for this study. The aforementioned densely populated regions can be observed in the figure. There are also very good and efficient transport infrastructures and several university locations throughout the canton.

Figure 2 Rent per square meter at data points locations



2.5.2 Previous Studies for the Canton Zurich

A hedonic study in the market segment of home ownership was published by the cantonal bank of Zurich (Salvi et al., 2004). The authors investigate the topic from the real estate business perspective and information about price building effects as the basis for an investment analysis is aimed. They define three category of variables: characteristics of the house, macro location (e.g. travel time to Zurich), micro location (e.g. distance to shopping centre, noise).

Only a few ecosystem related variables are considered. The results show that floor area is the most important variable. Other important characteristics of the house are whether the house is constructed according to Minergie (+9%), with a cellar (+6%) and with a modern kitchen (+5%). On the macro level, travel time to Zurich has a strong impact; 15min for example lead to a 20% decrease in price of a property. On the micro level, results are based on a hectare grid. At good locations, these are locations from where many hectare lake surface can be seen, an increase of prices of 11% can be observed; modest views do not have any effect. Street noise was observed as well and at locations with daily average exposure above 55dB(A), one additional dB(A) noise causes a decrease of 0.66%.

Another hedonic model was estimated for the Canton Zurich by Löchl (2010) with the aim to find a feasible and suitable modelling approach of real estate price data for the city simulation tool UrbanSim. Especially, the performance of spatially autoregressive models and GWR models relative to each other was of interest. The data was collected by a survey, evaluation of online property listings and available GIS data. In the models, the rent price was regressed on about ten structural variables, on twelve spatial variables, of which solar exposure, noise and view can be linked to ecosystem services, and on time variables, which indicates when the advertisement was published. The total number of observation was 8592. As a result the highest standardised beta coefficients belong to the variables floor area and car travel time to the centre of Zurich. Regarding the modelling, the author favours spatial lag error models (SARerr) over GWR and spatial mixed model (SARmix). The former perform better regarding correlation problems of the residuals and significance and they are easier to handle and more reliable in the UrbanSim context.

Hedonic models can play a vital role in inverse modelling that was applied to the Canton Zurich (Crespo and Grêt-Regamey (2011) and Grêt-Regamey and Crespo (2011). Hedonic price models are used as if…then models, but the desired state to be reached is defined first and, via the hedonic model, the conditions on how this aim can be reached can be identified. This planning approach is presented using a GWR analysis for different clusters within the canton. Using the same data as Löchl (2010), the authors are able to show and explain trade-offs and compensation scheme effects amongst different variables at different locations. The results help to overcome the problems along with densification and give answers on how future transformations of cities and agglomerations can correspond to the inhabitant's living standards.

3 Method and Data

3.1 Method

The general aim is to carry out a hedonic pricing model for rent prices in the Canton Zurich, considering structural, accessibility, environmental and neighbourhood variables. This classification is suggested by Fujita (1989) and is adopted in this thesis. The basic steps are first, defining relevant variables and the modelling approach; second, collect and process data in order to obtain variables; third, estimate a global model; fourth, expand the model to a local model (GWR); interpret the results and derive knowledge.

Predominately, the software R (R-Development-Core-Team, 2012) with several packages[2] and in some cases arcGIS (ESRI, 2011) is used.

3.2 Variable Set

As a result of previous studies (Löchl (2010), Belart (2011), Schirmer, Belart and Axhausen (2011)) and since this thesis is integrated in the research project «SustainCity»[3], many accessibility and neighbourhood variables are already present. They need only minor changes to be implemented into this thesis' models, whereas a major task is to construct environmental variables.

The choice of the variable set is based on literature study, on the framework of environmental variables (section 3.2.5) and in order to do comparisons with the results by Löchl (2010).Hence, the most important structural, accessibility and neighbourhood variables are selected and all environmental variables generated. In the following sections each variable is described.

Rental Price

This variable is the dependent variable. It is the monthly gross rent as it is indicated in the advertisement of each flat in the property listing. It represents the amount of money people are

---

[2] Packages include: foreign, sp, raster, MASS, lattice, geoR, maptools, ggplot2, ez, plyr, stringr and rgeos. rgeos enables to use Geographic Engine Open Source GIS software in R.

[3] http://www.sustaincity.org/

willing to pay for a specific flat with specific attributes. The rental prices in the advertisement are self-reported by the landlords. Lehner (2011) could identify a difference in reported and observed sale and rental prices in Singapore. It can be assumed that this effect is small in this study, as data contains only rental prices.

3.2.1 Structural Variables

Floor Area

This variable represents the amount of living space a flat provides. It is frequently used in hedonic models and acts as an important variable. The information about the square meters available is extracted from the description of each flat.

Number of Rooms

This variable represents the structure of the flat. Bedrooms, living room and kitchen are included. Unlike floor area, not only is it related to living space, but also indicates how space is structured. The number of rooms is frequently used in hedonic models as well. The information is extracted from the description of each flat.

Construction Date

This variable indicates at which date the house of the flat was built. A certain style of construction and comfort can be attached to some time periods. However, information about renovations is missing and not taken into account. The information is extracted from the description of each flat. Unfortunately, some records are missing the construction date. Since a removal of all affected records would cause a loss of approximately a third of the data, the affected rows are matched with observations by Löchl and Axhausen (2010). Consequently, a dummy variable design is adapted: age1 – built before 1920, age2 – built between 1921 and 1930, age_ref – built between 1931 and 1980 (reference dummy; excluded from models), age3 – built between 1981 and 1990, age4 – built between 1991 and 2011.

3.2.2 Neighbourhood Variables

Income Level

This variable indicates the average personal annual income within the municipality in which the property lies. It refers to the taxable income of natural persons. The Zurich Statistical Office provides the data. There is a natural relation between rental price and income, since tenants have to be able to afford a flat. Additionally, it is an aim to analyse the results regarding socio-economic parameters, of which income is an important one.

Population Density

This variable specifies the population density within a circular buffer of 1km radius around the property. The data is retrieved from the national population census based on a hectare grid.

Tax Level

This variable indicates the municipality tax base on income for individuals. It is calculated as percentage of the state income tax and reflects the attractiveness of a municipality from a fiscal point of view. Even though several studies showed that when using the tax level best results are obtained when tax level is confronted with the public service level (Zodrow, 1983), the tax level alone is still reported to be important for location choices (Binckebanck et al. (2011), Belart (2011)). The tax bases for each municipality are made online available by the Zurich Statistical Office and is matched to each record.

Rent Vacancy Rate

This variable indicates the percentage of available flats compared to the total number of flats within a municipality. It is selected since it may be argued that vacancy rates show the attractiveness to live at a location. Furthermore, it can point out the market situation, i.e. demand versus supply. Both might be reflected in rental prices. The rent vacancy rate for each municipality is made online available by the Zurich Statistical Office and is matched to each data record.

Proportion of Foreigners

This variable gives the percentage of inhabitants whose first language is not German. As for most societal parameters segregation effects can be observed (Epstein and Axtell, 1996). Over the past years, Swiss society has been holding a major debate on the desirable proportion of foreigners. This variable is thus selected in order to test its importance. The proportion of foreigners for each municipality is made online available by the Zurich Statistical Office and is matched to each record.

Education Level

This variable gives the percentage of persons who hold a university degree. Again the argument of social clustering theory can be added (Epstein and Axtell, 1996) and the variable is of interest in order to examine correlation between socio-economic and other variables. The information is retrieved from the population census, which provides values based on a hectare grid.

3.2.3 Accessibility Variables

Accessibility index calculations are based on the potential approach as described by Bodenmann (2003). According to the author, the accessibility of a location increases with an increasing number of activities at various other locations, that are accessible at as low as possible travel costs. The total accessibility of a location is the sum of the calculated accessibility value to each other location. Increasing travel costs diminish the accessibility along a negative exponential curve.

$$Accessibility_i = \sum_{j=1}^{j=n} A_j * \exp\left(-\beta * c_{ij}\right)$$

With location i and all other locations j ; the activity possibilities A in j ; and the generalised travel costs c between i and j, that are weighted by β.

Public Transport Accessibility

This variable gives the accessibility regarding public transport (Tschopp, Fröhlich and Axhausen, 2006).

Private Transport Accessibility

This variable gives the accessibility regarding private transport (Tschopp et al., 2006).

Generic Accessibility

This variable is a combination of the two variables above. It takes both public and private transport accessibility into account by taking the mean of the two untransformed values of the two original variables. The new generic variable highly correlates with the two original variables.

Distance to Railway Station

This variable gives the linear distance between the property and the nearest railway station. It approximates the level of access to the regional and national railway network, since it is related to the access costs (effort to reach the station). The Zurich Cantonal Centre for GIS Data provides the station locations. Since the Zurich railway network is dense and extensive, it plays an important role in commuting and other travel purposes and, consequently, is included in the variable set.

Distance to Highway Access

This variable gives the distance between the property and the nearest highway access as the crow flies. It approximates the level of access to the network, since it is related to the access costs (time to reach the highway access). As demonstrated in the literature review (2.4.2), the construction of a highway access does have an impact on the prices of houses within reach. The Zurich Cantonal Centre for GIS Data provides the highway access locations.

Distance to Zurich Centre

This variable gives the distance between the property and the centre of Zurich (Bürkliplatz) as the crow flies. The distance calculation is based on the land registry by the Cantonal Office for Cadastral Surveying.

Distance to Winterthur Centre

This variable gives the distance between the property and the centre of Winterthur as the crow flies. The distance calculation is based on the land registry by the Cantonal Office for Cadastral Surveying.

3.2.4 Environmental Variables

Generally speaking, distance variables focus on access to environmental services, whereas area variables focus on the amount. Variables are based on different radii, since the scale of environmental variable’s capitalisation is still subject to scientific debate (Abbott and Klaiber, 2009).

Noise Exposure: Railway

This variable describes the extent of noise exposure caused by rail traffic during day and night. The information is gathered from SonBase (Höin, Ingold, Köpfli and Minder, 2009), a

GIS noise data base by the Federal Office for Environment, which provides the measured and modelled noise at hectare level. Since the data contains separate files for day and night, it is possible to calculate the day-night average noise index. This index combines exposure by day and night in a standardised and frequently applied way (Nelson, 1982). A 10dB penalty is added to the night value before calculating the mean of the two values. Although studies showed that the bulk of the disturbing noise impact in Switzerland is caused by street noise, rail noise can not be neglected and is supposed to be perceived as more disturbing than street noise (Hartmann, 2011).

Noise Exposure: Street

This variable describes the extent of noise exposure caused by road traffic during day and night. The information is gathered from SonBase (Höin et al., 2009), a GIS noise data base by the Federal Office for Environment, which provides the measured and modelled noise at hectare level. Since the data contains separate files for day and night, it is possible to calculate the day-night average noise index. This index combines exposure by day and night in a standardised and frequently applied way (Nelson, 1982). A 10dB penalty is added to the night value before calculating the mean of the two values. Street noise is generally accepted as an impact factor on housing prices (cf. Nelson (1982), Lake et al. (1998), Chernih and Sherris (2004), Baranzini and Ramirez (2005)).

Noise Exposure: Aircraft

This variable describes the extent of noise exposure caused by air traffic during day and night. The information is prepared in a GIS which is provided by the Zurich Cantonal Centre for GIS Data, the data base is property of Zurich Airport. Other types, such as civil aviation or helicopter flights are not included in the data. As there are no longer any military airbases in the Canton Zurich, the data is still representative. Since the data contains separate files for day and night, it is possible to calculate the day-night average noise index. This index combines exposure by day and night in a standardised and frequently applied way (Nelson, 1982). A 10dB penalty is added to the night value before calculating the mean of the two values. There are various studies about the impact of aircraft noise on housing prices. According to one opinion, aircraft noise has a stronger impact on housing prices (Baranzini and Ramirez, 2005) than other noise sources. Hartmann (2011) reports that aircraft noise has a weaker impact. The case study by Weigt (2009) concludes that different market segments and regions show different reactions to air traffic noise exposure. Since there is a lot of air traffic at Zurich Airport and the airport is centrally located within the canton, it is important to include this variable to examine the reaction pattern in this case.

Noise Exposure: Shooting Range

This variable describes the extent of noise exposure caused by firing on shooting ranges. In the Canton Zurich, firing shooting is an issue because there are 160 or so shooting ranges in operation and shooting takes place regularly by day. The information is prepared and provided in a GIS by the Zurich Cantonal Centre for GIS Data. The noise is modelled in a simplified way: an inner zone around the shooting range where noise levels between 65 and 100 dB occur and an outer zone with values between 60 and 65 dB. To my knowledge, this noise source has been excluded from hedonic house price studies so far and it is thus interesting to test the importance of this type of noise.

Air Quality

This variable gives the air quality measured by the concentration of NO_2 in the air. The information is prepared and provided in a GIS by the Zurich Cantonal Centre for GIS Data. Values of NO_2 concentration are provided on a hectare grid. A negative impact of traffic-caused air pollution on property prices could already be shown by Borjans (1983). Since then air quality in general has gradually improved in Western Europe. Nevertheless, it is worth to include this variable, since there is a major spatial variation across the canton regarding NO_2 emissions.

Distance to Forest

This variable gives the distance from a property to the nearest border of a forest area as the crow flies. It approximates the access to a forest. The distance calculation is based on the land registry of the Cantonal Office for Cadastral Surveying.

Distance to Lake

This variable gives the distance from a property to the nearest border of a lake as the crow flies. It approximates the access to lakes. The distance calculation is based on the land registry of the Cantonal Office for Cadastral Surveying.

Distance to River

This variable gives the distance from a property to the nearest river as the crow flies. Culverted rivers and non-perennial rivers are excluded from consideration. The variable approximates the access to rivers. The distance calculation is based on the land registry of the Cantonal Office for Cadastral Surveying.

Distance to Agricultural Area

This variable gives the distance from a property to the nearest border of an agriculturally used area as the crow flies. These areas are all covered with humus and consist of fields, meadows and pastures. The variable approximates the access to agricultural area. The distance calculation is based on the land registry of the Cantonal Office for Cadastral Surveying.

Distance to Park

This variable gives the distance from a property to the nearest border of a park, an unsealed sports facility or gardens as the crow flies. The variable approximates the access to parks. The distance calculation is based on the land registry of the Cantonal Office for Cadastral Surveying.

Area of Forest within 400m

This variable gives the amount of forest area within a circular buffer around the property. The buffer has a radius of 400m which is equal to an approximately 10 min walk. Up to this distance and duration respectively, people assign features in their neighbourhood as within reach (Weber, Winkler, Graf and Bähni, 2006). The areas of interest are clipped to the circular buffer and summed up. The area calculation is based on the land registry of the Cantonal Office for Cadastral Surveying.

Area of Water within 400m

This variable gives the amount of water area within a circular buffer around the property. Polygons of river and lake surfaces are included in this variable. The buffer has a radius of 400m which is equal to an approximately 10 min walk. Up to this distance and duration respectively, people assign features in their neighbourhood as within reach (Weber et al., 2006). The areas of interest are clipped to the circular buffer and summed up. The area calculation is based on the land registry of the Cantonal Office for Cadastral Surveying.

Area of Agricultural Land within 400m

This variable gives the amount of agricultural area within a circular buffer around the property. The area consists of fields, meadows and pastures. The buffer has a radius of 400m which is equal to an approximately 10 min walk. Up to this distance and duration respectively, people assign features in their neighbourhood as within reach (Weber et al., 2006). The areas of interest are clipped to the circular buffer and summed up.

Area of Parks within 400m

This variable gives the amount of park area within a circular buffer around the property. The area consists of parks, unsealed sports facilities or gardens. The buffer has a radius of 400m which is equal to an approximately 10 min walk. Up to this distance and duration respectively, people assign features in their neighbourhood as within reach (Weber et al., 2006). The areas of interest are clipped to the circular buffer and summed up. The area calculation is based on the land registry of the Cantonal Office for Cadastral Surveying.

Area of All Favourable Land Uses within 400m

This variable gives the amount of all aforementioned surface types (forest, water, agricultural area and parks) area within a circular buffer around the property.

Area of Favourable Land Usages within 50m

This variable gives the amount of favourable surface types area within a circular buffer around the property. Basically all unsealed surface types are included: forest; lakes; rivers; fields, meadows, pastures; parks, unsealed sports facilities, gardens; vines; other unsealed groves; unsealed property surface; moors; unsealed slopes; reeds; Wytweiden (wooded pastures) and other unsealed areas. Unlike the 400m buffer variables this 50m buffer refers to the individual rather than the common supply of certain favourable areas, since the close neighbourhood is considered only. The area calculation is based on the land registry of the Cantonal Office for Cadastral Surveying.

Panorama of Mountains

This variable represents the panorama from the property on mountains belonging to the Alps. A data set of the highest mountaintops is used, which is provided by swisstopo. Each top represents a whole massif (“HGipfel”) and is the highest top of each massif. From this data set 15 tops are selected that are distributed in the range from the Eastern via the Central to the Western Alps. A selection is necessary, since test runs revealed a processing time of one day per mountaintop. In this manner, 15 mountaintops approximate the alpine chain and properties with a wide-angle panorama receive a higher variable value than properties with a limited panorama. Based on the mountaintop selection and the surface model, a visibility analysis is performed. Previous research suggests that panorama in general leads to higher rental prices (e.g. Salvi et al. (2004) or Fleischer (2012)) and view on mountains is thus included in the variable set.

View to Lakes

This variable represents the view on lakes within the Canton Zurich, which have a surface larger than 100'000m^2. Lakes are cut into *n* 100'000m^2 area pieces and a residual area. Afterwards, each area is reduced to a point. Therefore, *n+1* points with one *n* equal to 100'000m^2 approximate a lake. This conversion is required in order to reduce the processing time for the visibility analysis to below one week. Based on the generated points and the surface model, a visibility analysis is performed. Properties with direct view to many points receive a higher variable value than those with direct view to only a few ones. It is not checked if two or more lakes are visible or whether all visible points lie within one lake. Completed studies showed the impact of lake view on property prices (e.g. Sirmans et al. (2005) or Ma and Swinton (2011)); this variable is thus included. Unlike the view variables by Löchl (2010), this visibility analysis (as well as mountain panorama) is based on a digital surface model, which, unlike the elevation model, also represents the land uses, such as houses, trees and hedges.

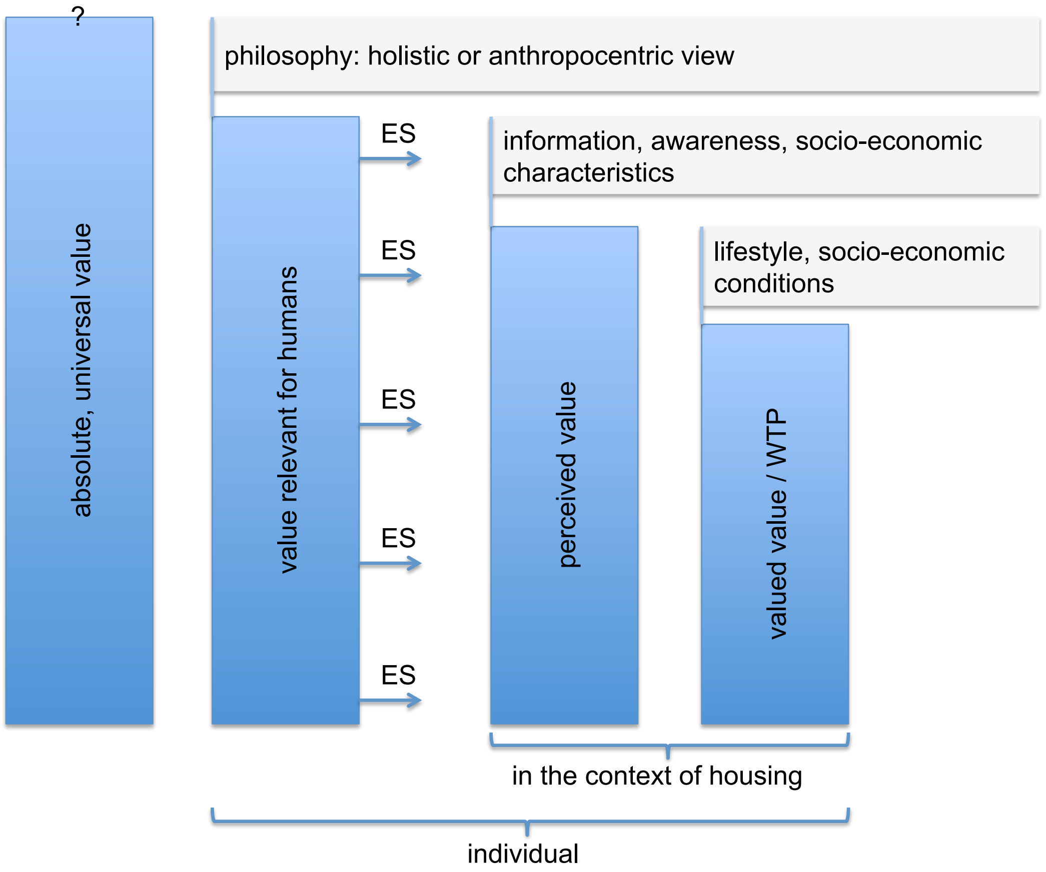
Indicator for Ecology and Identity

This variable is designed in order to assess the ecological and identification value of a property's direct neighbourhood. According to Moretti et al. (2010) these two factors often coincide in urban landscapes. Put simply, this means that inhabitants welcome in ecological terms positive attributes, too. Most important factors are structural complexity and the age of vegetation. The index is based on four data subsets: A) Environmental protection zones and items; B) areas according to the Ecological Quality Ordinance (Ökoqualitätsverordnung, ÖQV; agricultural areas and meadows); C) the ecological potential of open and wooded areas; D) and the local ratio of boundary divided by area of different land usages. The information for A, B and C is provided by the Zurich Cantonal Centre for GIS Data, the information for D by the Cantonal Office for Cadastral Surveying. All subparts of the indicator are related to the ecological and identification value in general, A specifically to the age of vegetation and D specifically to the complexity. In the following, there are some explanations for every subpart. A) For the age of vegetation, many ways of variable calculation were tested, such as making use of inventories, observation of vegetation height, retrieving the age from mapped vector data or from aerial photos. However, none of them were feasible (spatial limitations, implausibility, ineffectiveness and too short time series respectively) on a level as the cantonal one. The age of vegetation must be described by other means. The following heuristic is applied: Environmental protection zones and protected items (boulders, trees, hedges, etc.) were established decades ago. Thus, they assure a certain age of vegetation. Further, Weber (2009) explains that at those times environmental protection actions were lead by «conservation movements», which were the founding organisations of the current environmental associations. They campaigned for the protection of areas or items that people identified with or were especially valuable due to their uniqueness or age. The vegetation and trees cannot be cleared if protected. But management may require some selective clearing interventions. However generally speaking, this vegetation is older and causes higher identification than surrounding areas and landscape items. The polygons of these areas and items are transformed into a 1/0 raster. B) The area polygons are transformed into a 1/0 raster; 1 for cells within the polygon, else 0. C) The layers containing information on ecological potential for open and wooded areas are combined and transformed into a 1/0 raster; 1 for all potentials over 60%, else 0. D) Within a circular buffer with radius 50m around each property, the total area of unsealed land usages and the corresponding total length of boundaries are identified. The total length of boundaries is divided by the total area and in this way higher values represent a higher small-scale variation of different land usages, that is its complexity. The values are transformed into a 1/0 raster; 1 for all values above the mean of the sample, else 0. Since there is a considerable proportion of «other land usages», the Shannon Index, which is usually used and which takes the different land usage types into account, would lead to unreliable results. Having all of this subparts ready, the four raster layers are added and for each property the value is selected that occurs most frequently within a circular buffer of 50m around the property. Thus, the index has a range from 0 to 4, the higher the higher is the ecological and identification value.

3.2.5 Linking Ecosystem Services and Hedonic Variables

A certain combination of biotic and abiotic factors defines the type of an ecosystem. In the wider system, it performs a task according to its type, from which humans can derive benefits, too. In order to find out, how and to what extent these ecosystem services are valued in the housing market, appropriate variables must be designed. The following two figures inform about the fundamental principles of this step.

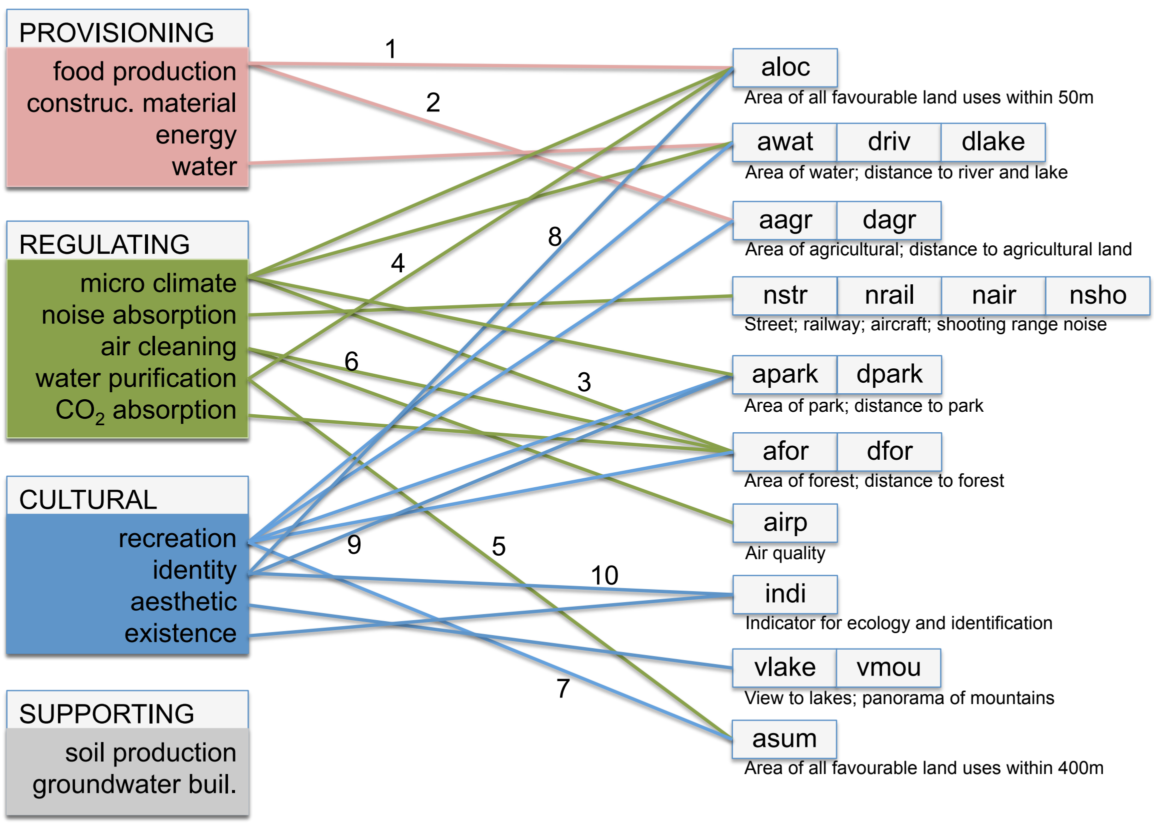
Figure 3 Different values of ecosystem services



Source: Freely adapted and expanded from Huppenbauer (2007). ES = Ecosystem services

The height of the bars differs pursuant to individual characteristics. Using rental prices in regression models, the individual WTP in the context of housing can be estimated.

Figure 4 Ecosystem services and possible base variables



As shown in Figure 4, the most important groups are regulating and cultural ecosystem services in the context of housing. Supporting functions are not relevant and provisioning services play a minor role in the housing market context. Most of the connections are intuitive; some are briefly explained in the following sentences. 1) The possibility to have an own garden. 2) Farm-gate sales. 3) Evaporation and shadow by trees (Nowak et al., 1998). 4) and 5) Natural seepage due to unsealed surface. 6) Air filtering by trees (Nowak, 1994). 7) Well-being and health depend on the design of the urban environment (Kaplan, 1993). 8) Identification with direct natural home environment (Kaufmann et al., 2008). 9) Sense of belonging through social interactions (Bühler, Kaspar and Ostermann, 2008). 10) cf. variable «Indicator for Ecology and Identity». Variables such as air pollution and noise are not directly linked to a service, but ecosystem services contribute to mitigate those problems. A comprehensive explanation of ecosystem services in suburban spaces is given by Grêt-Regamey et al. (2011).

However, it must be clearly stated at this point, that in order to link the environmental variables (variables on the right side in Figure 4) to ecosystem services, there must be information available that qualifies environmental variables in a detailed way. Except the indicator variable, no qualitative specifications can be done due to lack of required data. Therefore, these variables should be still regarded as environmental variables, while bearing in mind that they have the potential to be elaborated into ecosystem services variables, on condition that such data will be available for the whole Canton Zurich.

3.2.6 Variable Description

The tables in this section give an overview of the different variables that are used in the regression models. In Table 2, it is shown how they are coded and the statistics are given. As the models are estimated from a pedestrian point of view, distances are scaled per 100m. Areas are per Are ($100m^2$) and income is in 1000 CHF. The relation between the different variables and the obtained data base is presented in Table 3

Table 2 Descriptive Statistics of the Variables

Variable	Unit	Min	Max	Mean
rent – rental price	CHF per month	52.00	57940.00	2195.00
area – floor area	square metres	9.00	702.00	102.10
room – number of rooms	count	1.00	14.00	3.87
age1 – constructed before 1920	dummy	0.00	1.00	
age2 – constructed1921 to 1930	dummy	0.00	1.00	
age3 – constructed 1981 to 1990	dummy	0.00	1.00	
age4 – constructed 1991 to 2011	dummy	0.00	1.00	
accp – public transport accessibility	ln of acc. index	6.67	11.23	10.12
accm – private transport accessibility	ln of acc. index	9.60	12.53	11.24
accs – combination of accp and accm	ln of acc. index	7.22	10.83	9.49
drst – distance to railway station	100 metres	0.70	45.85	9.63
dhig – distance to highway access	100 metres	0.34	109.99	22.76
dcbd – distance to Zurich centre	100 metres	317.80	371.17	118.41
dwcbd – distance to Winterthur centre	100 metres	1.46	368.75	191.47
popd – population density	capita per hectare	0.23	142.03	26.46
tax – tax level	per cent	58.72	103.81	92.21
inco – income level	kCHF per year	59.41	169.91	80.3
uni – education level	per cent	1.20	19.23	8.38
nonf – proportion of foreigners	per cent	1.63	28.71	16.28
vacr – rent vacancy rate	per cent	0.00	12.20	0.99
nrail – noise exposure: railway	dB(A)	0.00	65.00	14.97
nstr – noise exposure: street	dB(A)	0.00	64.00	40.51
nair – noise exposure: aircraft	dB(A)	0.00	70.00	8.56
nsho – noise exposure: shooting range	dB(A)	0.00	100.00	1.14
airp – air pollution	µg NO_2 per m^3	11846.00	44737.00	21452.00
dfor – distance to forest	100 metres	0.00	18.27	3.37
dlake – distance to lake	100 metres	0.00	227.70	59.98
driv – distance to river	100 metres	1.00	14.36	2.48
dagr – distance to agricultural area	100 metres	0.00	13.48	1.50
dpark – distance to park	100 metres	0.00	27.68	3.51
afor – area of forest within 400m	are	0.00	23827.92	1439.54
awat – area of water within 400m	are	0.00	2045.68	21.93
aagr – area of agricultural land within 400	are	0.00	990.90	40.30
apark – area of parks within 400m	are	0.00	562.55	8.07
asum – area of all fav. land uses within 400m	are	0.00	553.81	90.48
aloc – area of all fav. land uses within 50m	are	0.00	59.06	11.16
vmou – panorama of mountains	count	0.00	12.00	1.23
vlake – view to lakes	10ha	0.00	55.00	1.50
indi – indicator for ecology and identity	count	0.00	3.00	0.34

Table 3 Relation of variables and data base

	Property listing	Statistics: Income	Stat.: Pop. density	Stat.: Tax level	Stat.: % of foreigners	Sta.: % of uni degree	Stat: Vacancy rate	Accessibility index	Airc. & shoot. noise	Street & railw. noise	Air quality	Land registry	Ecological potentials	Env. protection zones	Dig. surface model	Dig. elevation model	Data M.Löchl (2010)	Data B.Belart (2011)
rent	X																	
area	X																	
room	X																	
age	X																X	
accp								X									X	
accm								X									X	
drst												X						X
dhig												X						X
dcbd												X					X	
dwcbd												X					X	
popd			X														X	
tax				X													X	
inco		X															X	
uni						X												
nonf					X													
vacr							X											
airp											X							
nair									X									
nsho									X									
nstr										X								
nrail										X								
dlake												X						
driv												X						
dagr												X						
dfor												X						
dpark												X						
apark												X						
afor												X						
aagr												X						
awat												X						
asum												X						
aloc												X						
indi												X	X	X				
vlake												X			X			
vmou												X			X	X		

Blue: Data preparation entirely or partly (age) by Raphael Fuhrer; white: by IVT (cf. 3.2)

3.3 Data

As mentioned in the previous section 3.2, variables are based on different data sources. These are presented in Table 4.

Table 4 Data

Input data	Time	Source	Type
Property listing	2008-2010	Comparis	Spatial point data frame
Statistics: Income	2010	Canton Zurich	Spatial polygon data frame
Statistics: Population density	2000	BfS	Spatial polygon data frame
Statistics: Tax level	2007	Canton Zurich	Spatial polygon data frame
Statistics: Percentage of Foreigners	2000	Canton Zurich	Spatial polygon data frame
Statistics: Percentage university degree	2001	BfS	Spatial polygon data frame
Statistics: Vacancy rate	2010	Canton Zurich	Spatial polygon data frame
Accessibility Index	2005	IVT	Spatial point data frame
Noise: aircraft + shooting	2011	Canton Zurich	Spatial polygon data frame
Noise: railway + street	2009	BAFU	Raster
Air quality	2010	Canton Zurich	Raster
Land registry	2010	Canton Zurich	Spatial polygon data frame
Ecological potentials and env. protection zones	2010-2011	Canton Zurich	Spatial polygon data frame and raster
Digital elevation model	2009	Swisstopo/ETH	Raster
Digital surface model	2008-2010	Swisstopo/ETH	Raster

Except the data extracted from the population census and the accessibility calculations, all data sets are current and correspond with the time period when the information of the properties were collected. Generally speaking, statistical and spatial data is of high quality in Switzerland (Löchl, 2010). However, some inaccuracies in the land registry data are observed during the variable generating process. In some cases different land usage polygons are overlapping and in sampling inspections some errors, e.g. lacking park areas in the City of Zurich, are identified.

As a summary, there is sufficient data available in order to generate a comprehensive variable set. Only minor aspects, such as crime rate and temperature distributions (heat island effect) are missing due to lack of available data. The variable choice pays special attention to environment related variables and data. Some small data errors were detected and corrected where possible and for some variables approximations are carried out due to lack of exact data. Altogether, there is a robust and reliable variable and data set.

3.4 Hypotheses and Research Questions

The main question is whether environmental variables are important or not, and how they can be implemented into hedonic models. Consequently, the following hypotheses and model expectations based on the literature review are proposed in Table 5.

Table 5 Model expectations

Extent \ Sign	Positive	Negative
Large	floor area, number of rooms age: before 1920 age: 1921-1930 age: 1991-2011	age: 1981-1990
Medium	accessibility of public transport accessibility of private transport generic accessibility income level tax level area of parks area of fav. land uses within 50m panorama of mountains view to lakes	distance to Zurich centre population density street noise aircraft noise distance to lake
Small	education level area of forest area of water area of agricultural land area of fav. land uses within 400m indicator for ecology and identity	distance to railway station distance to highway access distance to Winterthur centre percentage of foreigners vacancy rate railway noise shooting range noise air quality distance to forest distance to river distance to agricultural land distance to park

The expectations in the table above hold true for the global model. In the local context, expected sign and impact extent differ over space.

1. As a general rule, distance variables are expected to show a negative sign, as people should prefer closeness to facilities (the less meters they are away, the higher the price).
2. Structure variables are expected to have a larger impact than the other three types of variables.

3. Area variables with small radius (aloc) should have a larger impact, since they can be related to semi-private capitalisation of environment related benefits (building's parcel), whereas larger radius (asum) mean common benefits.
4. It is assumed that there is a hierarchy amongst environmental variables, visibility for example is assumed to be more important than air quality. Variables related to water (awat, dlake, driv) are more important than other land usages variables, since water is rarer than forest or agricultural area surface
5. In cities and suburban municipalities parks are valued higher than in rural areas, since there is higher building density.
6. Environment related variables are correlated with socioeconomic parameters. A recent study by Diekmann and Meyer (2010) shows that environmental burdens are not equally distributed among social classes in Bern. It is reasonable to assume, that this phenomena occurs in the Zurich area as well.
7. Regarding model specification, it should be analysed whether area or distance related environmental variables perform better in hedonic models.

3.5 Models

First, global models are estimated, where spatially autoregressive models are applied with the package spdep (Bivand, 2012). Second, local models are estimated, where GWR is calculated with the R package spgwr (Bivand and Yu, 2012). For both categories, there are two aims: One is to find a model for if…then scenarios with the most important variables and the other is to have a model in order to test the hypotheses. There are no mono-causal models estimated, since they perform not as well as multi-causal models (Richardson et al., 1973). For all models the 3084 observations are used, which have no NA in any variable.

3.5.1 Global Models

First, the spatial weighting matrix is calculated using a radius of 4000 meters. Since the usual SARlag model (result in Appendix A 1, Table 11) shows difficulties with heteroscedasticity, a generalised spatial two stage least squares (STSLS) model is applied (Kelejian and Prucha, 1998). This is a spatially autoregressive model that can control heteroscedasticity by fitting a spatial lag model by two stage least squares method (Piras, 2010). The disadvantages are that it is not based on maximum log-likelihood and that it computes more conservative and thus lower significance levels of its results (cf. Appendix A 1, Table 11).

In a second step, the problem of multicollinearity is addressed. A factor analysis reveals four basic variable groups. The first group move on the gradient urban-rural (e.g. population density, air pollution etc.), the second group on the gradient strong-weak municipality (e.g. income, tax level etc.), the third group consists of the two variables «number of rooms» and «floor area», and the fourth group includes all «independent» variables (noise, visibility, age, indicator, park, area of water and agricultural land). From among variable sets, which highly correlate with each other, only one variable is included in the model. The two accessibility variables accp and accm are combined into one generic variable, since it seems not justifiable to randomly drop one of them, as accp is more important in cities and accm in rural regions. Regarding environmental variables, one can argue that some of them partially represent similar services (e.g. forest and agricultural areas/pastures can both be linked to recreation) and regarding neighbourhood variables analogous concepts hold true. The resulting models after this optimisation are models 1 and 2 in Table 6 in the next chapter.

In a next step, the number of variables is reduced to a set of the most important ones. This model performs best for if…then analysis. The result of this 13 variable model is given in Table 8.

About 40 variants and sub-models have been designed and tested. One aim is to overcome the problem of many estimate's low significance, which is present in most models. Also, a model with rent per floor area is estimated, but does not lead to good results – all estimates are highly insignificant (Appendix A 1, Table 14). Though some variable transformations, such as taking the logarithm of all or some variables and other means, do improve model performance, they are not progressed, as most of them lack of conceptual justification (examples in Appendix A 1, Table 13). However, the semi-logarithmic approach is presented together with the linear one, a comparison of area and distance related environmental variables is given at Table 7.

3.5.2 Local Models

The model with all variables possible without multicollinearity effects and the reduced model including the most important variables are expanded into a GWR. For both GWR models an adaptive bandwidth is used. This means that the spatial weighting depends on the density of observed data points (Figure 1). In order to obtain a continuous β-surface as a model output, the following parameters are set: the cell size is 100 by 100 meters and the colour bar depends on the quantile distribution, which stresses the differences rather than it smoothens the surface, as the differences over space are of interest. The resulting maps and values are presented in section 4.2.

4 Results

This section provides the results of global and local models. All coefficient estimates are based on standardised values. As the analysis is done from a pedestrian point of view, distances are scaled per 100m in global models. Areas are per are ($100m^2$) and income is in 1000 CHF. Unstandardised results can be found in the appendix (Table 15 to Table 17).

4.1 Global Models: Spatially Autoregressive Models

In Table 6 it is shown that it has an effect whether the semi-log (model 1) or the linear (model 2) model is applied. The sign of small betas, beta values as well as the significance levels are influenced. Approximately half (model 1) and most (model 2) of the estimates are not significant. The variables «area», «age4» (built between 1991 and 2011) and «uni» (percentage of people holding a university degree) are most important in both models. A few variable's signs are against expectation, this is the case for aircraft noise, area of agricultural land and the ecology and identity indicator. This is particularly disturbing regarding aircraft noise and the indicator, since both are significant. In a few other cases, signs of the two model designs contradict. Environment related beta values are smaller and more often insignificant compared to structural, neighbourhood and accessibility variables. However, some of them have similar ranges as the latter two groups. Rho, which indicates the spatial autocorrelation of the dependent variable, is in the range of what can be expected for rental prices in housing markets.

In Table 7 the sum of squared errors (SSE) is of particular interest. The lower the SSE, the better the model fit. The SSE of both models is very similar. This means that model 3 and model 4 perform equally. Regarding significance, model 3 tends to perform better than model 4. However, most environment related variables in both models show unexpected signs.

In Table 8 all variables show expected signs. The SSE is higher than in model 1 and 2 respectively, because the number of variables is reduced compared to the previous models. In general, the significance levels are good in model 5 and 6. But there are problems with the street noise variable and the age dummy variables in the linear model regarding significance. Distance to lakes is the most important one amongst environment related variables.

All results in this section are based on standardised variables in order to be able to compare the different coefficients more easily.

Table 6 Estimated STSLS parameters (N=3084; standardised estimates)

	Model 1: ln(rent) ~ variables			Model 2: rent ~ variables		
Variable	Estimate	Stand. error	P-Value	Estimate	Stand. error	P-Value
(Intercept)	-0.094	0.022	0.000	-0.033	0.029	0.253
area	0.727	0.057	0.000	0.555	0.055	0.000
age1	0.085	0.052	0.099	0.019	0.083	0.823
age2	0.089	0.076	0.241	0.028	0.117	0.808
age3	0.065	0.033	0.047	0.007	0.036	0.851
age4	0.266	0.037	0.000	0.107	0.036	0.003
ln(accs)	0.050	0.016	0.002	0.027	0.032	0.387
popd	0.047	0.023	0.044	0.091	0.078	0.241
uni	0.191	0.014	0.000	0.150	0.071	0.034
vacr	-0.006	0.010	0.559	-0.005	0.008	0.533
nair	0.020	0.007	0.005	0.020	0.006	0.000
nsho	-0.005	0.007	0.487	-0.002	0.006	0.667
nstr	-0.001	0.010	0.944	-0.014	0.017	0.391
nrail	-0.017	0.011	0.109	-0.029	0.012	0.014
dlake	-0.056	0.009	0.000	-0.043	0.017	0.010
driv	-0.007	0.012	0.543	0.003	0.012	0.839
dpark	0.002	0.009	0.822	-0.009	0.010	0.325
apark	0.004	0.008	0.666	-0.004	0.009	0.659
afor	0.005	0.008	0.560	0.004	0.010	0.654
aagr	-0.007	0.007	0.338	-0.001	0.010	0.928
awat	0.023	0.012	0.064	0.045	0.022	0.043
indi	-0.029	0.010	0.003	-0.026	0.012	0.038
vlake	0.038	0.010	0.000	0.015	0.018	0.413
vmou	0.014	0.012	0.233	-0.002	0.011	0.857
Rho	0.179	0.042	0.000	0.241	0.314	0.442
SSE	857			1729		

Table 7 Comparison: Distance and area related env. variables (N=3084; stand. estimates)

	Model 3: rent ~ variables, with env.=distance			Model4: rent ~ variables, env.=area		
Variable	Estimate	Stand. error	P-Value	Estimate	Stand. error	P-Value
(Intercept)	-0.024	0.029	0.414	-0.027	0.028	0.335
area	0.554	0.055	0.000	0.551	0.054	0.000
age1	0.004	0.069	0.950	0.057	0.070	0.413
age2	0.039	0.108	0.719	0.070	0.108	0.516
age3	-0.001	0.037	0.988	-0.003	0.039	0.937
age4	0.083	0.037	0.025	0.079	0.036	0.026
ln(accs)	0.046	0.012	0.000	0.078	0.015	0.000
uni	0.161	0.068	0.017	0.195	0.079	0.013
dlake	-0.050	0.020	0.013			
driv	0.000	0.014	0.993			
dagr	0.055	0.026	0.037			
dfor	0.016	0.017	0.335			
dpark	-0.012	0.008	0.147			
awat				0.039	0.021	0.072
aagr				-0.009	0.010	0.405
afor				-0.005	0.013	0.693
apark				-0.004	0.008	0.636
aloc				-0.001	0.009	0.878
Rho	0.226	0.303	0.455	0.217	0.296	0.463
SSE	1740			1750		

Table 8 Models with the most important variables (N=3084; standardised estimates)

	Model 5: ln(rent) ~ variables			Model 6: rent ~ variables		
Variable	Estimate	Stand. error	P-Value	Estimate	Stand. error	P-Value
(Intercept)	-0.091	0.022	0.000	-0.032	0.028	0.251
area	0.725	0.057	0.000	0.551	0.052	0.000
age1	0.090	0.048	0.059	0.055	0.066	0.404
age2	0.098	0.075	0.188	0.063	0.105	0.547
age3	0.062	0.033	0.059	0.000	0.036	0.992
age4	0.255	0.037	0.000	0.096	0.037	0.009
ln(accs)	0.081	0.014	0.000	0.085	0.018	0.000
uni	0.203	0.014	0.000	0.174	0.062	0.005
nstr	-0.002	0.010	0.873	-0.014	0.017	0.408
nrail	-0.021	0.011	0.049	-0.032	0.011	0.002
dlake	-0.057	0.009	0.000	-0.044	0.015	0.003
dpakr	-0.005	0.009	0.594	-0.021	0.010	0.031
awat	0.019	0.012	0.117	0.039	0.022	0.071
vlake	0.038	0.010	0.000	0.012	0.017	0.495
Rho	0.158	0.043	0.000	0.196	0.267	0.463
SSE	867			2612		

4.2 Local Models: Geographically Weighted Regression

Models 2 and 6 are extended into a GWR model approach; these are model 7 and model 8. It should be mentioned that the following maps are an exploratory representation of the GWR results. They show trends, but they are not intended for single pixel analysis. For many variables, beta estimates are only significant in subareas. Significance (t-value higher than 4; α=0.95) is marked with dense strips, low significance (t-value higher than 3.8; α=0.91) is marked with sparse strips. More information related to local t-values can be found in Fotheringham et al. (2002), chapter 7.

In Figure 5 the model fit of the model with the most important variables (model 8) are low in western and north-western parts of the canton, especially in the western part of the City of Zurich and in the Limmattal. However, very high R^2 values occur locally. A comparison with Figure 2 reveals no correlation between dense data points and high R^2. The plotted residuals of the same model identify some areas in the western and soutwestern part of Zurich, where larger residuals occur. This phenomena can be observed in other parts of Zurich and at the eastern lake bank. In Figure 6 high coefficient estimates for «area» can be observed in the city of Zurich and on both banks of Lake Zurich; low but still clearly positive estimates are observed around Winterthur and in the region of Bülach. The coefficients are extensively significant. Significance of the education level variable indicates that it is important in the South of Zurich and the southern regions of Zurich with high coefficient estimates, whereas the eastern and north-eastern parts of the canton have medium values. In Figure 7 the railway noise variable is generally insignificant. Significant areas show clearly negative estimates. Regarding accessibility the values in the region between Greifensee and Lake Zurich as well as in the Knonaueramt are significant and positive with relatively high values, especially in the first mentioned area. In Figure 8 the influence of distance to a park is only significant in Zurich-Enge and indicates a high negative effect. In the City of Zurich the distance to lake estimates are significant, as they are in the agglomeration, in the very North and in the East of the canton. In the City of Zurich and its agglomeration the coefficients are more negative. In Figure 9 coefficient estimates for area of parks are signigicant in the northern parts of the Canton Zurich, around Winterthur and in Bülach area, in the airport region and in a small zone in the West of the City of Zurich. In contrast to the latter, coefficient estimates are positive. Significant coefficient estimates of the variable area of water show moderate values and are distributed over the canton from South West to East. In Figure 10 coefficient estimates of the panorama of mountains are very locally significant at the West of Zurich and in the municipalities behind the Üetliberg. The values in the city are higher than on the countryside. Regarding view to lakes, significant areas are, in general, the North except Winterthur and a larger region in the North East. Areas in the Glattal and in the South West are significant, too. Estimates nearer to Winterthur are higher than the others; the lowest are those one in Zurich.

In Appendix A 2 maps of further variables are presented. They have less significant coefficient estimate areas than the ones in the following figures. It is important to note, that insignificant coefficient estimates generated by GWR are unreliable, since they do not necessarily represent trends, but moreover can show the wrong sign and magnitude (Fotheringham et al., 2002).

Figure 5 Model fit (left) and residuals (right) of the most preferred model (model 8)

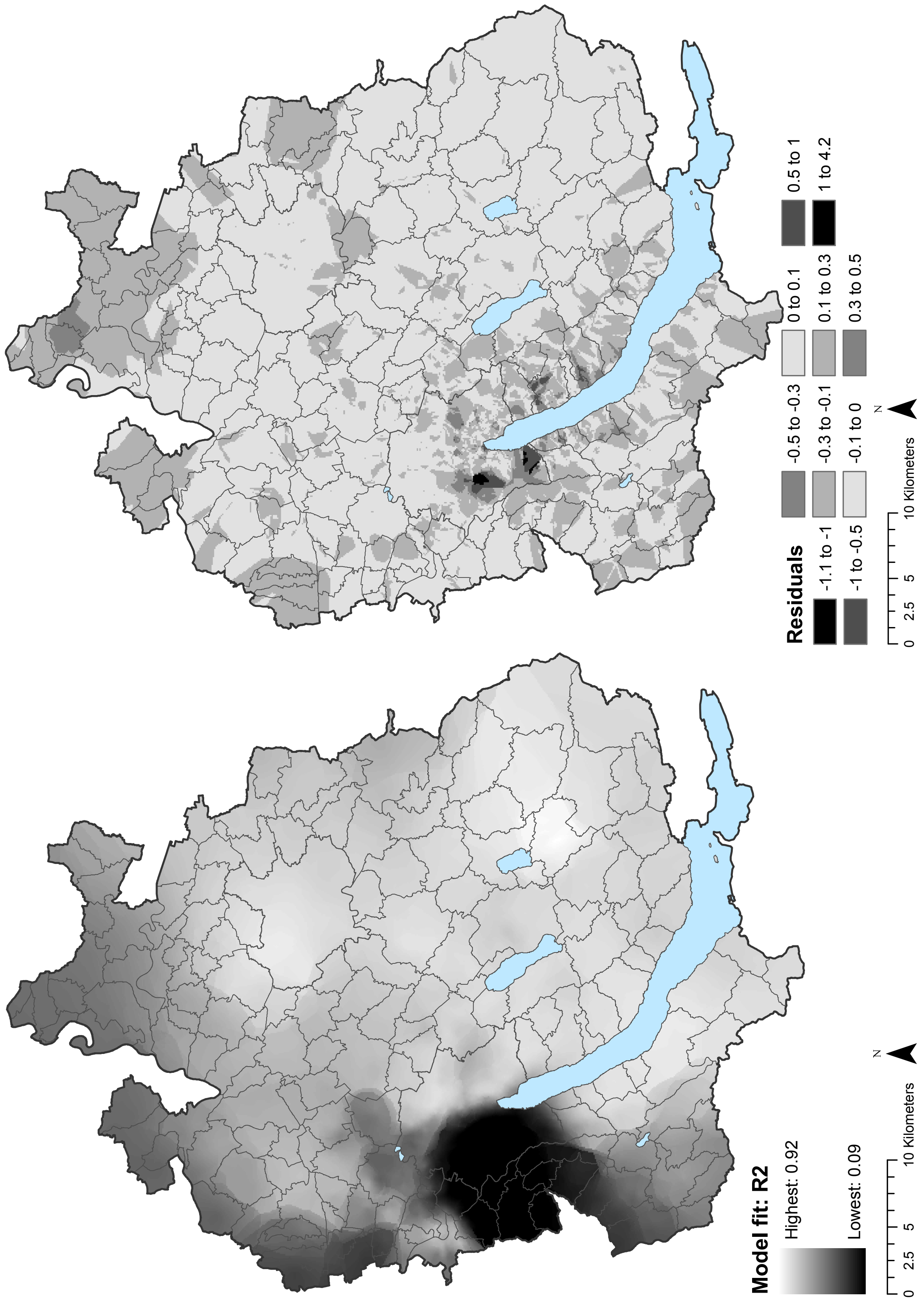


Figure 6 GWR: Floor area and education level (model 8)

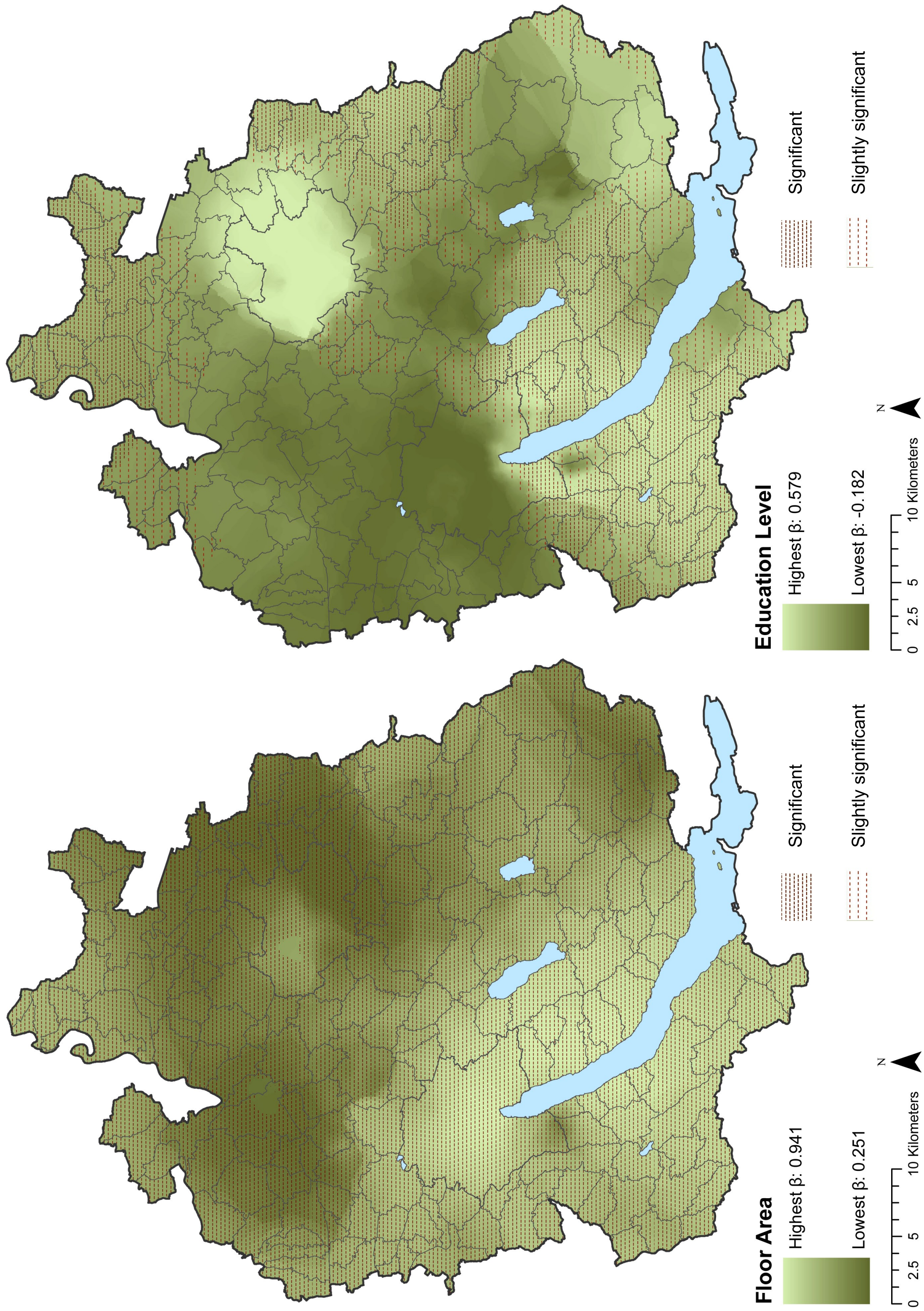


Figure 7 GWR: Railway noise and generic accessibility (model 8)

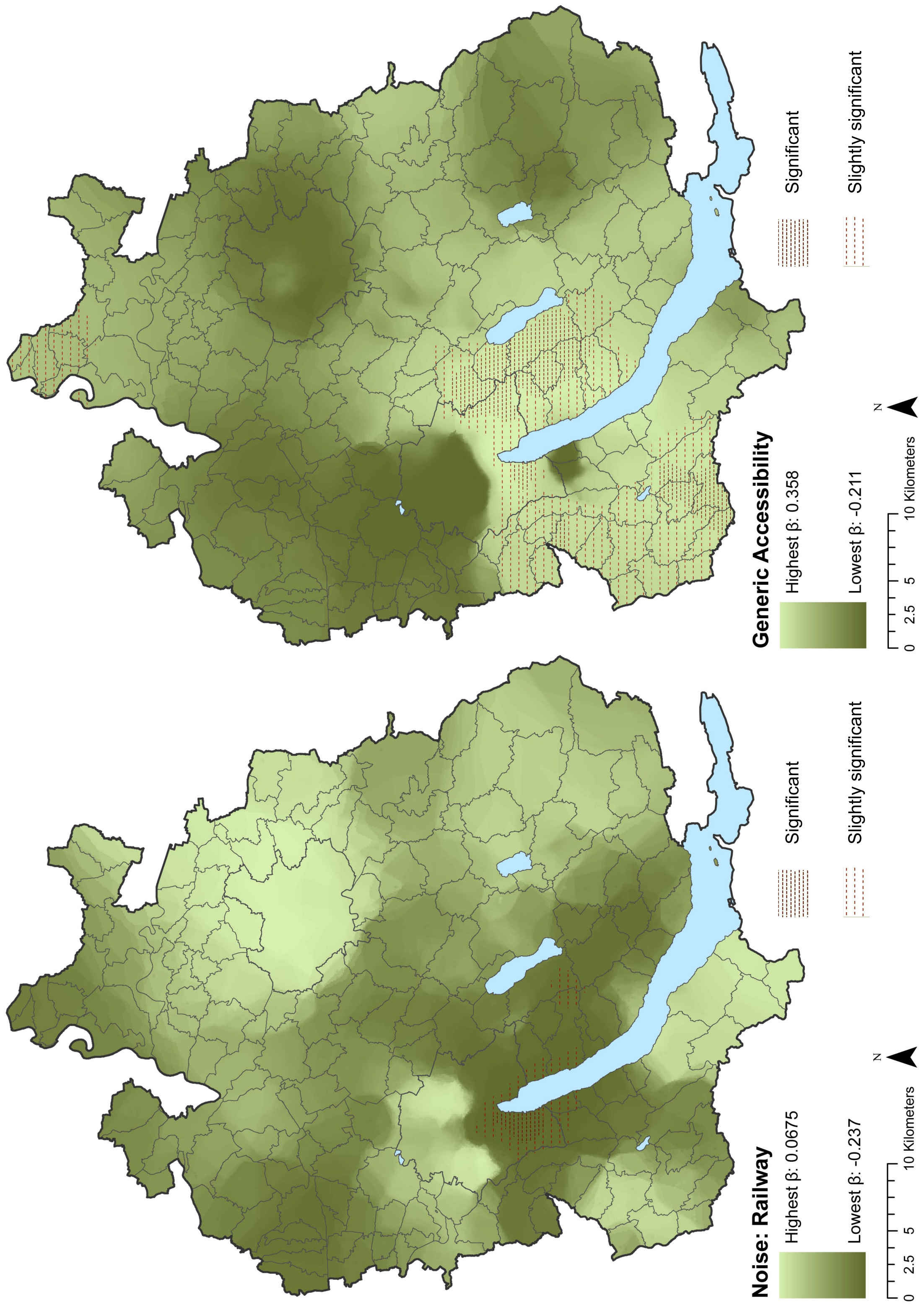


Figure 8 GWR: Distance to a park and distance to a lake (model 8)

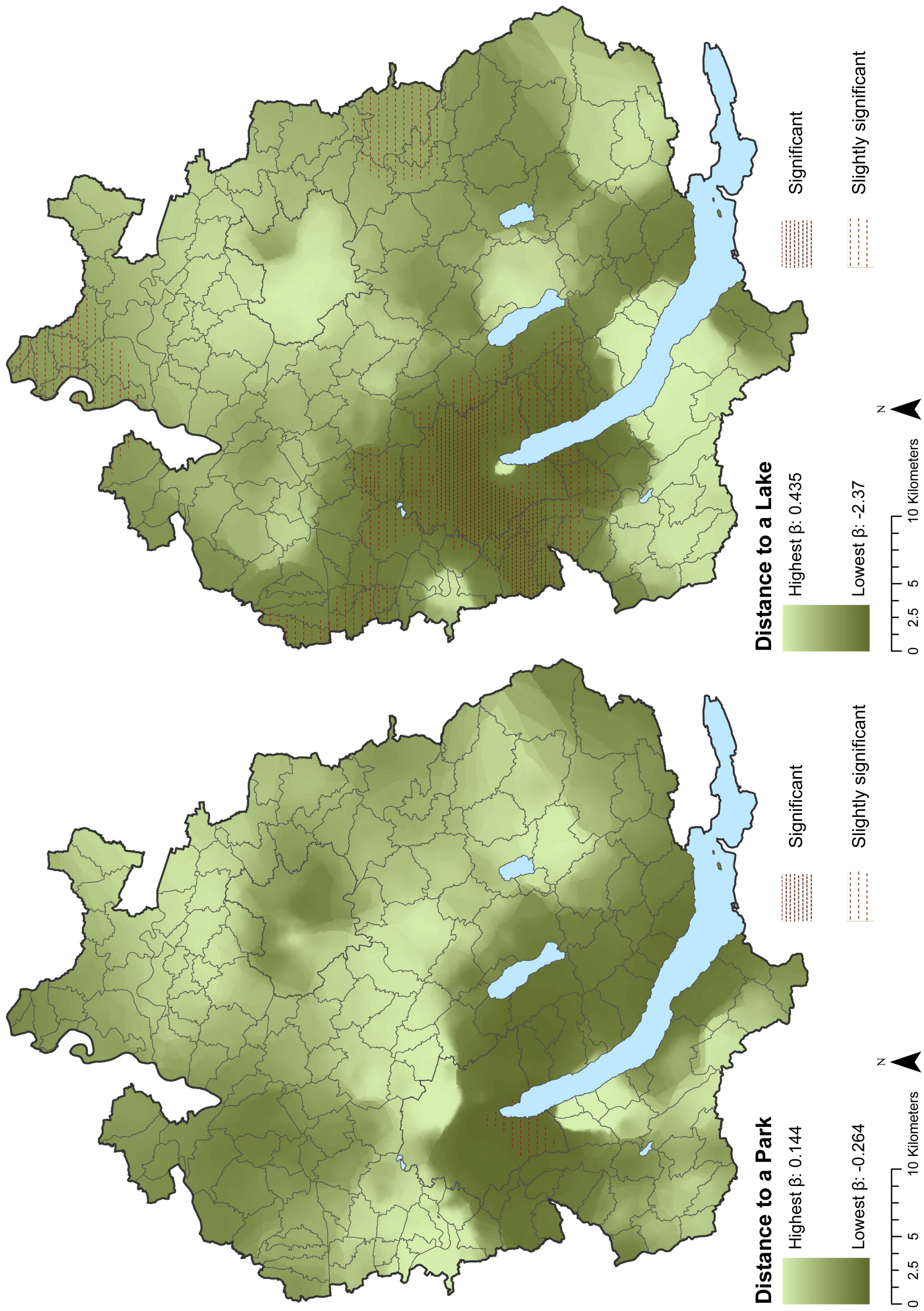


Figure 9 GWR: Area of parks and area of water surface (model 8)

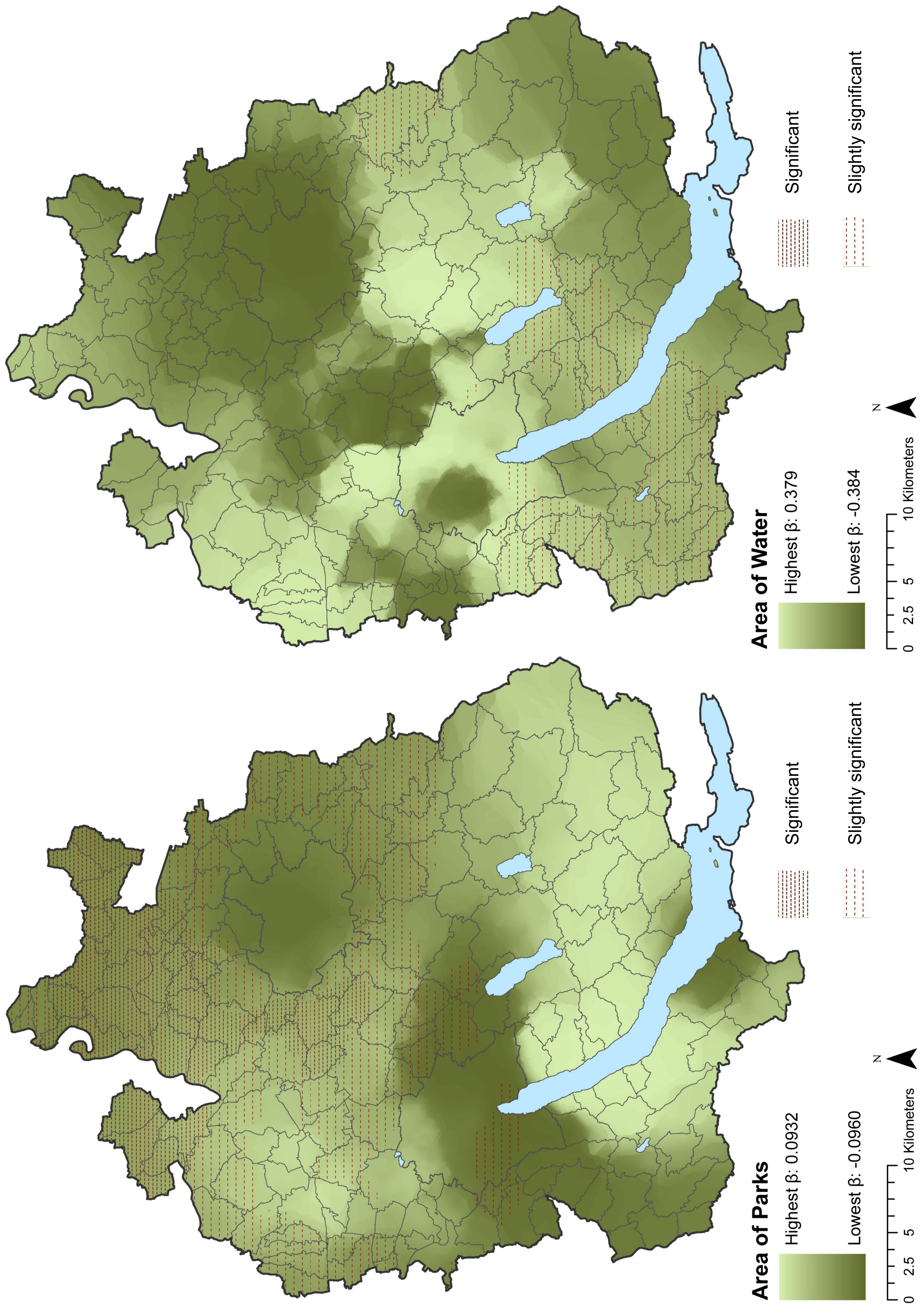
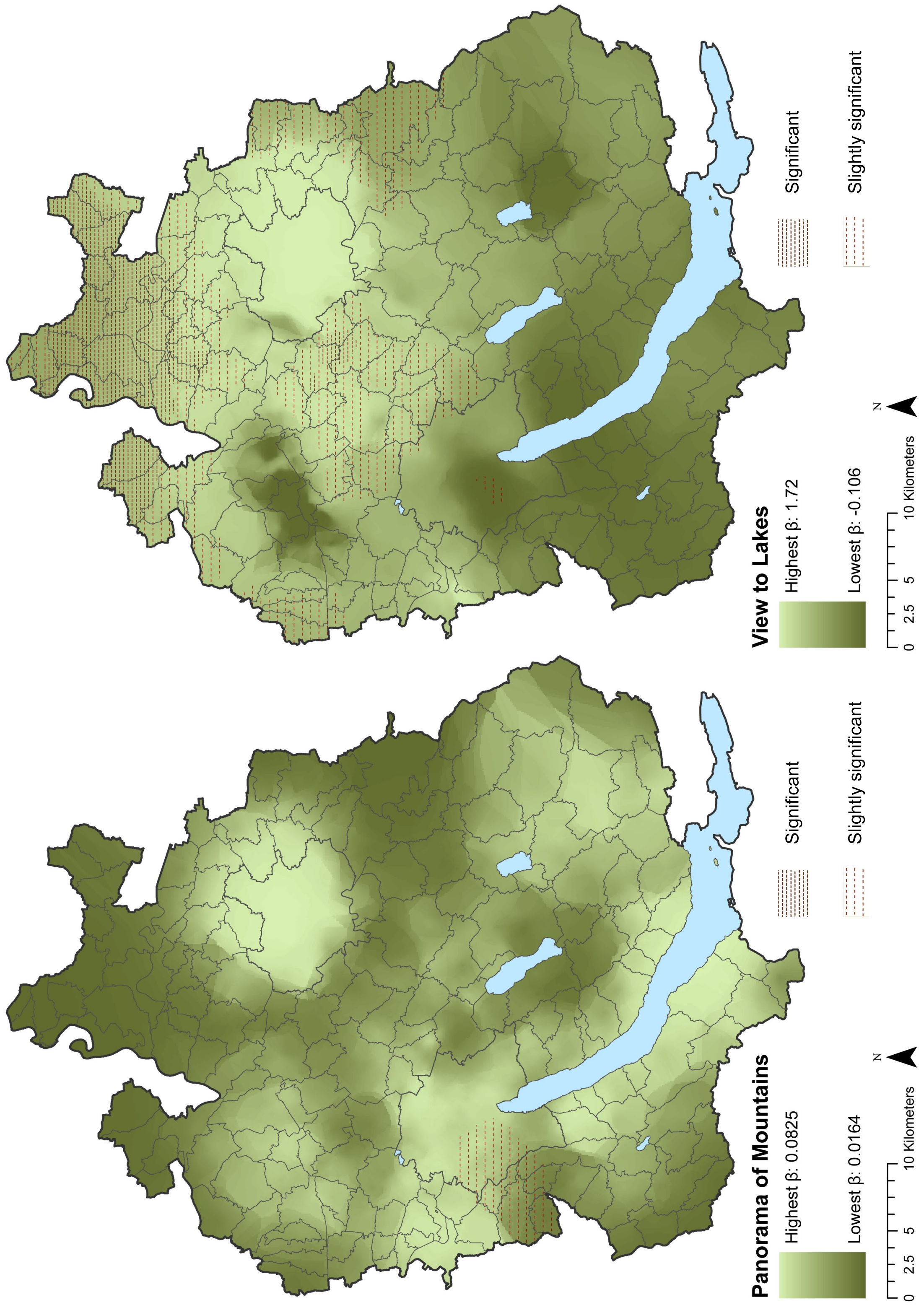


Figure 10 GWR: Panorama of mountains (model 7) and view to lakes (model 8)



5 Discussion and Conclusions

5.1 Analysing the Results

In principle, it was possible to model rental prices in the Canton Zurich with SAR and GWR models based on structural, accessibility, neighbourhood and environmental variables. However, the structures and patterns vary over space and interact with each other in a very complex way. It was thus difficult to address and overcome all of these apparent and underlying effects in the models. Attempts were successful, but at the price of some variables needed to be removed and low significance levels in the results. But even without these necessary measures, lack of significance is one of the main problems in the results (see section 5.2.1). On the other hand, significant results do show very interesting findings and also insignificant results may show trends.

Having a comparison between the expected coefficient estimates and the results (see Table 9), the coefficient estimates basically correspond to the expected coefficients.

Table 9 Variable hierarchy according to the results

Extent \ Sign	Positive	Negative
Large	floor area, number of room age: 1991-2011 education level income level	
Medium	age: before 1920 age: 1921-1930 generic accessibility accessibility of public transport accessibility of private transport *population density* *aircraft noise* area of water view to lakes *distance to forest* *distance to agricultural land* *percentage of foreigners* tax level	street noise railway noise distance to lake distance to park *indicator for ecology and identity* distance to Zurich centre
Small	*age: 1981-1990* area of forest panorama of mountains all fav. land uses within 50m all fav. land uses within 400m *air quality* distance to Winterthur centre	vacancy rate distance to river *area of parks* *area of agricultural land* distance to railway station distance to highway access

Variables that clearly contradict the expectations are in italic.

But there are some exceptions as explained in the following. The age dummy variable «Built 1991 until 2011» (age4) has a larger impact than expected and also than the other age variables. People value the fact that the flat has been recently built in a strongly positive way. Another reason for the result among the age variables may be the fact that for age4 the real age of the flat correspond with the construction date, whereas the other age dummies may not always represent the age of the current standard, since renovations could have taken place. And thus those other age dummy variables may be less accurate, which is supported by the lower significance levels. Similar, the approximation as described in section 3.2.1 predominately used for age1 to age3 may have an influence. Unlike expectation, the education level (uni) seems to be the most important neighbourhood variable. This is interesting since municipalities usually focus on their tax policy, supply of housing (variable «vacancy rate») and public services, such as schools and public transport. Measures as such in order to increase the percentage of people with university degree are usually not taken and it would be difficult to do so, since this group is heterogeneous. The variable «uni» is rather a variable that somehow summarises other neighbourhood variables up to a certain extent in addition to its own representation.

The positive coefficient estimate of aircraft noise is difficult to explain, as rental prices increase by 1.50 CHF for an additional dB(A) noise (cf. Table 16). Other studies (e.g. Baranzini and Ramirez (2005) found negative estimates. A possible, but quite unlikely explanation is that people prefer to be near to the airport and therefore live at locations that are exposed to aircraft noise. According to Weigt (2009) aircraft noise does not automatically lead to lower rental prices; in each airport region prices react uniquely. Under some conditions, for example low vacancy rates – which is the case in Zurich and its agglomeration –, positive price effects can be possible. Then, tenants value the proximity to the airport, but cannot discriminate noisy locations due to lack of alternatives. It might also occur the effect, that certain social classes find flats in noisy areas only, since competition with other applicants is smaller there. Landlords can take advantage of this situation and do not lower the rents. All other noise variables show expected sign. But railway noise is valued more strongly than street noise. The reasons might be the expansion of the S-Bahn and thus shorter intervals and railway noise during night, which both can be perceived as more annoying. Furthermore, it was confirmed that the variable «distance to a lake» is amongst the environmental distance variables the most important one. This might be because, compared to forest or agricultural areas (which though are excluded due to multicollinearity), the variable is less randomly distributed and thus more distinct. Locations near to the lakes are preferred compared to rivers and parks as well as forest as the impression from model variants (that are not reported in the text) implies. Unfortunately, many other environmental distance and area variables needed to be excluded or are insignificant. The problem of insignificance is discussed in the next section 5.2. The ecology and identification indicator (indi), which is highly correlated with the variable «aloc», shows a significant negative sign. A positive and small to medium impact was actually expected, since very local environment related benefits accrue on the parcel and become partly semi-private. When «indi» is replaced by «aloc», the result looks similar. The calculation method for the indicator is elaborated and, though carefully done, error-prone. But as the similar results of «indi» and «aloc» show, the reason lies not in an error during calculations.

Indeed, Donovan and Butry (2011) report similar results that contradict traditional concepts of public goods in environmental economics. They showed that an additional tree on the pavement has a larger impact on the rental price as a tree on the lot. However, the impact of trees on the lot was smaller, but still positive. Regarding the visibility variables, it is a surprise that their results are not equal, apparently people value view to a lake higher than a panorama of mountains. After discussing the expectations for the different variables, the hypotheses are treated. The first two hypotheses are very general and rely on experience from previous studies, they have the function of general indicators whether the model makes sense or not.

1. *Distance variables show negative signs.* All significant distance variables show this pattern. It can be assumed that the model is designed correctly in its principles.

2. *Structure variables have the largest impact.* The results give no reason to reject the hypothesis. This too gives certainty that the model is correctly established and we can rely on significant results.

3. *Small area variables have a larger impact extent than large area variables.* As mentioned above, results do not support this hypothesis. Environment related variables that have a character of private or club goods («indi» and «aloc») do not have higher positive estimates compared to clearly public goods as for example area of lakes, forest etc. It can be argued that view is somehow a club good, since it is in a way consumed at home and other people are excluded from this particular view as they do not live in the flat. In this context, the hypothesis does not have to be rejected, since the two view variables have larger coefficient estimates as most other environmental variables. But regarding the area variables, this principle does not hold true.

4. *Environment variables related to visibility and water are most important.* The hypothesis holds largely true. Most important are variables related to lakes, including distance, visibility and area. First, lake surface is more scarce and less randomly distributed than other surfaces, for example forest. Second, services provided by lakes might be more obvious than provided by other ecosystems. Most regulating services provided by a forest for instance are not directly visible, concerning recreation a lake and its bank tends to be more functional than pastures and forest. However, this is not always the case. The aesthetic value of a lake seems also to be higher than other land usages, as real estate companies do not advertise with view to pastures but view to lakes.

5. *In cities parks are valued higher than in rural areas.* Since the GWR result for the variable «distance to a park» is mainly insignificant, no conclusive statements can be

made to this hypothesis. When referring to the variable «area of parks», the results show that park area is valued more negatively in Zurich than in the countryside and the hypothesis must be rejected. These results are in line with the findings of Baranzini and Schaerer (2011) who argue that parks are valued positively in general, but negatively in the direct neighbourhood as parks might generate negative externalities (noise, unrest etc.). This conflict is more accented in cities than in rural areas, which may be an explanation for negative coefficient estimates.

6. *Environment related variables are correlated with socioeconomic parameters.* Factor analysis and correlation matrix show correlation between some socioeconomic variables and environmental variables. However, since most socioeconomic variables are on municipality level and environmental variables on a much lower level (parcel or ha), it is difficult to extract reliable results and thus these are not presented in the result section. As an example the correlation between «percentage of people with non-first language German» and «air pollution» is 0.71. Though there are no systematic results, it seems that the trend shows correlation between the two variable groups in some cases.

7. *There is a difference in model performance between environmental area and distance variables.* The two models perform equally. Apparently, trade-offs such as a near but small park versus a far but large park do not affect the results or the two approximations, caused by such trade-off effects, differ from reality in about the same amount. As a conclusion of the different models, it makes nevertheless sense to mix the two approaches among variables in one model in order to reduce correlation problems.

The low R^2 values in the western part of Zurich and the western agglomeration of Zurich as reported in Figure 5 are striking, since the data point density is high. In this region, the variance of rental prices is comparatively high as a glance at this local data subset shows. The plotted residuals point out that the model has difficulties to fit to the data points at certain locations in this part of the city. One probable explanation may be the influence of housing co-operatives, of which many are in these areas. Floor area is valued highest in Zurich and the two lake sides. This can be due to the high demand for flats on the one hand while space is limited on the other hand, so living area becomes more valuable. Two further conditions facilitate the process. First, it is an economically strong region. Second, there are more old buildings in urban areas, which usually provide smaller flats. Looking at the map of education level, we can conclude that an increase in the percentage of people graduated from university leads to a disproportionate increase in rental prices at locations on the northern lake banks. Education level plays as an umbrella variable for positive neighbourhood characteristics.

People at these areas particularly seem to pay attention on those neighbourhood characteristics. In Figure 8 the significant results suggest that proximity to a lake is valued higher in urban areas than on the countryside. This might be the case, because natural elements are scarcer in cities. Living directly at the lake bank means to live exquisitely, which may also have an effect on the results in these areas. Sine Zurich and agglomeration shows high preference for proximity to lakes, inhabitants would appreciate improvements in accessibility to the lake (direct paths, banks open to public). The panorama of mountains (Figure 10) is valued higher inside the city than in its suburban municipalities. Also, this is a part within the city where view to mountains is very limited and we can assume that at these locations the rare flats with a panorama of mountains are more popular. Similar concepts may explain the spatial variance of the estimates for view to lakes. In northern and eastern parts of the canton, where view is very limited since there are no large lakes, the coefficients are higher.

An attempt is made to improve model fit by including an evening sunshine index (calculation see Löchl (2007)). As the results in Table 12 (Appendix A 1) suggest, evening sunshine is an important explanatory variable. But it cannot solve the problems with unexpected signs, e.g. aircraft noise. Thus, GWR results are not expected to be improved by the inclusion of this index.

A comparison with results by Löchl and Axhausen (2010) is difficult since they used more structural and less environmental variables and different model specifications. But coefficient estimates for floor area and for the variable «built before 1920» are similar. Regarding the variable «1991 until today» the values in this study are much higher. This is also the case for visibility of lakes. On the other hand, accessibility – which is not generic in the other model – leads to slightly lower coefficients in this study. Population density shows a positive sign in this study, whereas Löchl and Axhausen (2010) report a negative sign. The signs are interchanged in the other direction for coefficients of air noise, foreigners and tax level as well. Except for the difference in air noise, a possible explanation can be the still increasing demand for flats in urban areas. Regarding SSE, the models by Löchl and Axhausen (2010) perform better.

5.2 Gained Knowledge

The validity of the results and their implications are discussed in this section.

5.2.1 Restrictions

Though the majority in the Canton Zurich is formed by tenants (Hofer et al., 2011), the results are limited to the rental sector, since the home ownership sector may react differently. Probably, there might be a slight bias towards urban areas, as the proportion of rental to ownership homes is said to be higher in cities. The number of data points is, compared to other studies, rather small. Löchl (2010) for example had 8592 data points and Lehner (2011) between 6351 and 45'792 in Singapore, depending on the model. Increasing data points is discussed in the next section in combination with the significance problem. Another point is the assumption of market equilibrium. The models show a small impact of the rent vacancy rate. However, some areas in the Canton Zurich are well known for their housing shortage. In this form the models deliver reliable results. But when the data will be updated in the future, a correction of the rental prices by the vacancy rate or something similar may be considered. Concerning data update, it should be repeated that the population census data is not current. Current data is expected to be published very soon. In this context, it would be a considerable improvement of the models, if socio-economic data was available in higher resolution, e.g. in a hectare grid or other resolution below the current municipality level. Due to multicollinearity many variables were excluded. In order to avoid that and to improve significance levels, many transformations have been tested. This process was partly successful, but not always justifiable. In some cases, variables perform better if transformed logarithmically; in other cases transformations do not improve the model. Examples are given in Appendix A 1, Table 13 and Table 14. Due to this variable exclusion, less knowledge could be derived. Another way is the combination of variables though. This works for accessibility, but the variable «asum», which adds all different environmental area variables (forest, agricultural areas, water, parks), performed not satisfactorily. An attempt with variable «rent per square meter» (rent/ area) was made, but the results are not better compared to the normal models. A factor that was ignored is the influence of life-style of the tenants, as investigated by Belart (2011).

Problem low significance

The reasons for low significance are based on the number of data points. In rural area they are sparse, in urban areas the variance in the dependent variable «rent» is high. These two facts are identified as the main reasons for low significance. Another possible source for this problem could be the combination of different aggregation levels (level house, 400m buffer, hectare grid, municipality). While this can be handled in analysing the SAR models, it is especially a problem in GWR results. The data point density in this study is 1.73 data points per square kilometre; higher values occur in urban areas, lower in rural areas. An increase in point density would lead to very time-consuming data collection and processing of variables.

Particularly, calculations such as visibility analysis would take at least several weeks. These problems could be avoided when the area was reduced to a test field. This test field should be chosen in a way that it starts in the centre of Zurich and extends as far as the very rural areas. In this way the main spatial patterns could be modelled, as for example the urban-rural gradient or economically strong vs. weak municipalities. But the price is the loss of a cantonwide model. It would also be possible to include more accurate environmental variables, which is a prerequisite for modelling the influence of ecosystem services services.

Environmental and ecosystem services variables

All environmental variables depend on quantitative units – distances, areas and counts. Noise and air quality variables are only indirectly linked to environmental entities, as environmental entities abate such problems (e.g. absorption of noise by vegetation, air cleaning by trees). In order to assess ecosystem services, qualitative information is needed for specifying adequate and accurate indicators within these simplified and uniform current environmental variables. The indicator for ecology and identity is an attempt in this direction, but the results do not convince. In that light, the following questions arise. What kind of environmental aspects do tenants perceive? Results show, that the more obvious environmental variables are, the higher valued they are, which is in line with findings by Boyle and Kiel (2001). Can these findings give hints to answer the same question for the more complex ecosystem services? And, having the answer to that, which data is needed? Today, it is very difficult or impossible to find data that provides information, such as the age of vegetation, the composition of species, the recreational value etc., on an extensive base. The level canton might be too high for such high-resolution variables that ecosystem services valuation requires. According to the results in this study, it is most important to find reliable ecosystem variables for lakes, rivers and parks.

5.2.2 Further research

Environmental variables should be converted into ecosystem services variables in order to obtain results for specific WTP and trade-off analysis with ecosystem services. An approach could be the weighting of the environmental variable units by qualitative information related to specific ecosystem services. This implies further data on qualitative variation within different land uses on a small-scale base. Further research can also help to improve significance in GWR models. For both points, a test field might be a solution, since effort for data can be kept reasonable, but the basic rationale of such variables can be tested. In this context, the question of privately or semi-privately versus commonly used services preference can be addressed, as there seems to be no consistent answer yet. A comparison between the rental and ownership sector may provide more clarity, as these two sectors likely lead to different results. Furthermore, models can be improved by integrating updated census data and higher resolution of socio-economic data, as soon as they are available. Also, an integration of housing cooperation (Genossenschaftswohnungen) as a factor to the rental price can improve the models. Finally, analysis of self-selection effects will deliver valuable results concerning decisions by tenants within a certain spatial structure.

5.2.3 Implications for Transport, Environmental and Spatial Planning

In Table 10 derived WTP for some selected variables are presented. The values are based on the evaluation of different global models. Though, we should interpret the values with caution, especially results on the right column. Based on these numbers, trade off analysis with implications for transport and environmental planning are presented. It would have been interesting to do similar analysis with GWR results. But unfortunately the lack of significance in many environmental variables makes this impossible.

Table 10 Willingness to pay for some selected variables in CHF and in per cent of rent (%)

WTP CHF	Variable	WTP %	WTP CHF	Variable	WTP %
184.00	number of rooms [/#]	12.00	0.20	accessibility [/unit]	0.01
18.00	floor area [/m^2]	0.75	2.45	pop.density [/%]	0.10
110.00	age:<1920 [dummy]	4.50	0.90	air noise [/dB(A)]	0.04
185.00	age:>1991 [dummy]	12.01	-78.00	indicator [/#]	-3.00
56.00	education level [/%]	2.27			
-2.50	street noise [/dB(A)]	-0.10			
-3.50	rail noise [/dB(A)]	-0.14			
-1.25	distance lake [/100m]	-0.05			
-10.00	distance park [/100m]	-0.41			
0.60	water area [/are]	0.02			
7.15	view lake [/10ha]	0.29			
10.30	view mountain [/# top]	0.24			

Left side: Robust results – different models show similar (significant) values. Right side: Only trends for WTP.

From a transport planning point of view, the implications for traffic demand are important, since the spatial structure of settlement and transport infrastructure interact. The coefficient of the variable that gives the distance to Zurich centre is negative, which means centrally located flats are preferred in general. Nevertheless, there have to be reasons that people choose a flat at a remote location. People would move 1km outwards, if accessibility was improved by 26 absolute units or the outer flat was 0.42m^2 larger than the inner one, with the assumption that the two flats are identical in all other characteristics. Again with the assumption of having two identical flats, one of them could be 24km more remote than the other one, if the first one was built after 1991 and the second one between 1931 and 1980. In a nutshell, improved accessibility, larger flats and newly built flats are sound reasons for people to live at remote locations. The impacts of the three characteristics are surprisingly strong and need to be kept in mind when planning measures are taken, as they may occur as side effects.

From an environmental planner's point of view, it might be of interest how environmental aspects interact with the preferences of tenants. As the results show, floor area is the most important variable, particularly in urban areas. Since space is limited, the ratio of living area per person cannot increase endlessly and densification must be guided in a way that is accepted by population. People would give up 1m^2 of their flat, if…

- the nearest park was 180m nearer to their flat.
- the nearest lake was 1.46km nearer to their flat.
- 30 ares of water surface were present in their neighbourhood.
- street noise was reduced by 2dB(A) or railway noise by 1.6dB(A).
- 25ha lake surface were visible from their flat.
- 1.74 mountaintops were visible (of which every top represents a whole massif).
- the accessibility was improved by 90 absolute units.

Parks, lakes and so one cannot be relocated of course. But their accessibility can be improved and the information can be useful to identify locations that are suitable for densification due to their environmental characteristics. Furthermore, artificially built environmental features can contribute to a planning measure. By improving and expanding the supply of environmental services, disadvantages due to unpopular planning measures can be compensated.

6 References

Abbott, J. K. and H. A. Klaiber (2009) An Embarrassment of Riches: Confronting Omitted Variable Bias and Multiscale Capitalization in Hedonic Models, *The Review of Economics and Statistics,* **93** (4) 1331-1342.

Ahlfeldt, G. (2011a) If Alonso Was Right: Modeling Accessibility and Explaining the Residential Land Gradient, *Journal of Regional Science,* **51** (2) 318-338.

Ahlfeldt, G. (2011b) If We Build, Will They Pay? Predicting Property Price Effects of Transport Innovations, *SERC Discussion Paper,* Department of Geography & Environment, London School of Economics, London.

Alonso, W. (1964) *Location and Land Use*, Harvard University Press, Cambridge, Massachusetts.

Amrusch, P. and W. Feilmayr (2009) Nonmarket valuation of inner-city ecological values, in C. A. Brebbia and E. Tiezzi (Eds.) *Ecosystems and Sustainable Development VII,* 415-424, Wessex Institute of Technology Press, Southampton.

Atkinson-Palombo, C. (2010) Comparing the Capitalisation Benefits of Light-rail Transit and Overlay Zoning for Single-family Houses and Condos by Neighbourhood Type in Metropolitan Phoenix, Arizona, *Urban Studies,* **47** (11) 2409-2426.

Baranzini, A. and J. V. Ramirez (2005) Paying for Quietness: The Impact of Noise on Geneva Rents, *Urban Studies,* **42** (4) 633-646.

Baranzini, A., J. V. Ramirez, C. Schaerer and P. Thalmann (2008) Applying Hedonics in the Swiss Housing Marekts, *Swiss Journal of Economics and Statistics,* **144** (4) 543-559.

Baranzini, A. and C. Schaerer (2011) A sight for sore eyes: Assessing the value of view and land use in the housing market, *Journal of Housing Economics,* **20** (3) 191-199.

Batty, M., C. Vargas, D. Smith, J. Serras, J. Reades and A. Johansson (2011) Visually-Intelligible Land Use Transportation Models for the Rapid Assessment of Urban Futures, *UCL Working Papers Series,* **163**, Centre for Advanced Spatial Analysis, UCL, London.

Belart, B. C. (2011) *Wohnstandortwahl im Grossraum Zürich,* Master Thesis, Swiss Federal Institute of Technology, Zürich. (305)

Binckebanck, L., P. Hettenbach, M. Schwanke and A. Werner (2011) Ökologische Aspekte bei der Wahl von Wohnimmobilien, *Arbeitspapiere der Nordakademie,* **2011-02**, Nordakademie - Hochschule der Wirtschaft, Elmshorn.

Bivand, R. (2012) Package 'spdep': Spatial dependence: Weighting schemes, statistics and models, *Version 0.5-46,* Wien.

Bivand, R. and D. Yu (2012) Package 'spgwr': Geographically Weighted Regression, *Version 0.6-14,* Wien.

Boarnet, M. G. and S. Chalermpong (2001) New Highways, House Prices, and Urban Development: A Case Study of Toll Roads on Orange County, CA, *Housing Policy Debate,* **12** (3) 575-605.

Bodenmann, B. (2003) *Zusammenhänge zwischen Raumnutzung und Erreichbarkeit: Das Beispiel der Region St. Gallen zwischen 1950 und 2020,* Diploma Thesis, IVT, ETH Zürich, Zürich.

Borjans, R. (1983) Immobilienpreise als Indikatoren der Umweltbelastungen durch den städtischen Kraftverkehr, *Buchreihe des Instituts für Verkehrswissenschaften an der Universität zu Köln,* **44**, Verkehrsverlag J. Fischer, Düsseldorf.

Boyle, M. A. and K. A. Kiel (2001) A Survey of House Price Hedonic Studies of the Impact of Environmental Externalities, *Journal of Real Estate Literature,* **9** (2) 117-144.

Bühler, E., H. Kaspar and F. Ostermann (2008) Nachhaltiges Entwerfen, Bewirtschaften und Aneignen städtischer Parkanlagen, *NFP 54,* SNF, Bern.

Burgess, J., C. M. Harrison and M. Limb (1988) People, Parks and the Urban Green: A Study of Popular Meanings and Values for Open Spaces in the City, *Urban Studies,* **25** (6) 455-473.

Bürgle, M., M. Löchl, U. Waldner and K. W. Axhausen (2005) Land use and transport simulation: Applying UrbanSim in the Greater Zürich area, *Conference Report,* **197**, Computers in Urban Planning and Urban Management Conference (CUPUM), London.

Chau, K. W. and T. L. Chin (2002) A Critical Review of the Literature on the Hedonic Pricing Model and Its Aplication to the Housing Market in Penang, *The Seventh Asian Real Estate Society Conference,* Seoul, Korea.

Chernih, A. and M. Sherris. (2004). Geoadditive Hedonic Pricing Models. Retrieved from http://www.asb.unsw.edu.au/schools/actuarialstudies/Documents/A.%20Chernih%20and%20M.%20Sherris%20-%20Geoadditive%20Hedonic%20Pricing%20Models.pdf, March 2012.

Chernobai, E., M. Reibel and M. Carney (2011) Nonlinear Spatial and Temporal Effects of Highway Construction on House Prices, *Journal of Real Estate Finance and Economics,* **42** (3) 348-370.

Cheshire, P. and S. Sheppard (1995) On the price of land and the valuation of amenities, *Economia,* **62** (5) 247-267.

Cheshire, P. and S. Sheppard (2002) The welfare economics of land use planning, *Journal of Urban Economics,* **52** (2) 242-269.

Cheshire, P. and S. Sheppard (2005) The Introduction of Price Signals into Land Use Planning: Decision-making: A Proposal, *Urban Studies,* **42** (4) 467-663.

Coulson, N. E. and D. P. McMillen (2007) The Dynamics of Interurban Quantile House Price Indexes, *Urban Studies,* **44** (8) 1517-1537.

Crespo, R. and A. Grêt-Regamey (2011) Spatially explicit inverse modeling for urban planning, *Applied Geography,* **34** (2012) 47-56.

Debrezion, G., E. Pels and P. Rietveld (2010) The Impact of Rail Transport on Real Estate Prices: An Empirical Analysis of the Dutch Housing Market, *Urban Studies,* **48** (5) 997-1015.

Delmelle, E. (2009) Spatial Sampling, in A. S. Fotheringham and P. A. Rogerson (Eds.) *The SAGE Handbook of Spatial Analysis,* 183-206, SAGE Publications Ltd, Thousand Oaks.

Diekmann, A. and R. Meyer (2010) Demokratischer Smog? Eine empirische Untersuchung zum Zusammenhang zwischen Sozialschicht und Umweltbelastung, *Kölner Zeitschrift für Soziologie und Sozialphsychologie,* **62** (3) 437-457.

Donovan, G. H. and D. T. Butry (2011) The effect of urban trees on the rental price of single-family homes in Portland, Oregon, *Urban Forestry & Urban Greening,* **10** (3) 163-168.

Engel, S. and M. Veronesi (2011) Policy and Economics of Ecosystem Services, *Presentation in February 2011,* Institute for Environmental Decisions, ETH, Zürich.

Epstein, J. M. and R. Axtell (1996) *Growing Artificial Societies: Social Science from the Bottom Up*, MIT Press, Cambridge.

ESRI (2011) ArcGIS Desktop: Release 10, Environmental Systems Research Institute, Redlands CA.

Fischer, M. M. (2006) *Spatial Analysis and Geocumputations: Selected Essays*, Springer, Berlin.

Fleischer, A. (2012) A room with a view - A valuation of the Mediterranean Sea view, *Tourism Management,* **33** (3) 598-602.

Forrest, D., J. Glen and R. Ward (1996) The Impact of a Light Rail System on the Structure of House Prices, *Journal of Transport Economics and Policy,* **30** (1) 15-29.

Fortin, M.-J. and M. R. T. Dale (2009) Spatial Autocorrelation, in A. S. Fotheringham and P. A. Rogerson (Eds.) *The SAGE Handbood of Spatial Analysis,* 89-103, SAGE Publications Ltd, Thousand Oaks.

Fotheringham, A. S., C. Brunsdon and M. Charlton (2000) *Quantitative Geography*, SAGE Publications, Thousand Oaks.

Fotheringham, A. S., C. Brunsdon and M. Charlton (2002) *Geographically Weighted Regression*, John Wiley & Sons, Chichester.

Fotheringham, A. S. and P. A. Rogerson (2009) Introduction, in A. S. Fotheringham and P. A. Rogerson (Eds.) *The SAGE Handbook of Spatial Analysis,* 1-4, SAGE Publications Ltd, Thousand Oaks.

Fujita, M. (1989) *Urban Economic Theory: Land Use and City Size*, Cambridge University Press, Cambridge.

Garrod, G. and K. Willis (2006) The environmental economic impact of woodland: a two-stage hedonic price model of the amenity value of forestry in Britain, *Journal of Applied Economics,* **24** (7) 715-728.

Geiger, M. (2000) SNL-Mietpreisstruktur bestehender Mietverhältnisse, *Schlussbericht des Forschungsauftrages "SNL-Mietpreisstruktur bestehender Mietpreisverhältnisse" im Rahmen der Diskussion um die Teilrevision des Mietrechts,* **F-8358**, Bundesamt für Wohnungswesen, Grenchen.

Gibbons, S. and S. Machin (2004) Valuing rail access using transport innovations, *Journal of Urban Economics,* **57** (1) 148-169.

Gibbons, S. and S. Machin (2008) Valuing school quality, better transport, and lower crime: Evidence from house prices, *Oxford Review of Economic Policy,* **24** (1) 99-119.

Grêt-Regamey, A. and R. Crespo (2011) Planning from a future vision: Inverse modeling in spatial planning, *Environment and Planning B: Planning and Design,* **38** (6) 979-994.

Grêt-Regamey, A., N. Neuenschwander, N. Bakhaus, U. Wissen Hayek and S. Tobias (2011) Fokusstudie Landschaft, *Synthesebericht des Nationalen Forschungsprogrammes "Nachhaltige Siedlungs- und Infrastrukturentwicklung,* Schweizerischer Nationalfonds, Bern.

Haining, R. (2009) The Special Nature of Spatial Data, in A. S. Fotheringham and P. A. Rogerson (Eds.) *The SAGE Handbook of Spatial Analysis,* 6-23, SAGE Publications Ltd, Thousand Oaks.

Hartmann, S. (2011) Lärm entwertet Mietliegenschaften, *umwelt,* **4** 39-41.

Henneberry, J. (1997) Transport investments and house prices, *Journal of Property Valuation and Investment,* **16** (2) 1544-1158.

Hilber, C. (1999) Standortattraktivität, Liegenschaftspreise und Mieten, *WWZ-Forschungsberichte,* **4/99**, Wirtschaftswissenschaftliches Zentrum der Universität Basel, Basel.

Hofer, T., H.-P. Bucher, R. Gysel, M.-A. Battaglia, C. Bürgi, P. Moser, R. Page and U. Rey (2011) *Kanton Zürich in Zahlen 2011*, Statistisches Amt des Kantons Zürich and ZKB, Zürich.

Höin, R., K. Ingold, M. Köpfli and T. Minder (2009) *SonBase - die GIS-Lärmdatenbank der Schweiz*, Bundesamt für Umwelt, Bern.

Huppenbauer, M. (2007) Einführung in die Umweltethik, *Presentation in October 2007,* UFSP Ethik, University of Zurich, Zürich.

Jim, C. Y. and W. Y. Chen (2009) Ecosystem services and valuation of urban forests in China, *Cities,* **26** (4) 187-194.

Kaplan, R. (1993) The Role of Nature in the Context of the Workplace, *Landscape and Urban Planning,* **26** (4) 193-201.

Kaufmann, V., L. Pattaroni, M.-P. Thomas, S. Brändli, C. Ménétrieux, M. Piraud, S. Piot, L. Ortelli and A. Rabinovich (2008) Habitat urbain durable pour les Familles, *NFP 54,* SNF, Bern.

Kelejian, H. H. and I. R. Prucha (1998) A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances, *Journal of Real Estate Finance and Economics,* **17** (1) 99-121.

Lake, I. R., A. A. Lovett, I. J. Bateman and I. H. Langford (1998) Modelling Environmental Influence on Property Prices in an Urban Environment, *Computers, Environment and Urban Systems,* **22** (2) 121-136.

Lancaster, K. J. (1966) A New Approach to Consumer Theory, *The Journal of Political Economy,* **74** (2) 132-157.

Lehner, M. (2011) *Modelling housing prices in Singapore applying spatial hedonic regression,* Master Thesis, ETH, Zürich. (307)

Löchl, M. (2007) Considering spatial dependence in hedonic rent price regression, *7th Swiss Transport Research Conference,* Institute of Transport Planning and Systems ETH Zurich, Monte Verità, Ascona.

Löchl, M. (2010) *Application of spatial analysis methods for understanding geographic variation of prices, demand and market success,* Doctor of Science, ETH, Zürich. (18911)

Löchl, M. and K. W. Axhausen (2010) Modeling hedonic residential rents for land use and transport simulation while considering spatial effects, *The Journal of Transport and Land Use,* **3** (n) 39-63.

Löchl, M., M. Bürgle and K. W. Axhausen (2007), Implementierung des integrierten Flächennutzungsmodells UrbanSim für den Grossraum Zürich, *disP - The Planning Review,* 168, 13-25.

Ma, S. and S. M. Swinton (2011) Valuation of ecosystem services from rural landscapes using agricultural land prices, *Ecological Economics,* **70** (9) 1649-1659.

Ma, S. and S. M. Swinton (2012) Hedonic Valuation of Farmland Using Sale Prices versus Appraised Values, *Land Economics,* **88** (1) 1-15.

Malpezzi, S. (2002) Hedonic Pricing Models: A Selective and Applied Review, *Housing Economics: Essays in Honor of Duncan Maclenman,* The Center for Urban Land Economics Research, Madison, Wisconsin.

Mooney, H. A., A. Cropper, W. V. Reid, D. Capistrano, S. R. Carpenter, K. Chopra, P. Dasgupta, T. Dietz, A. K. Duraiappah, R. Hassan, R. Kasperson, R. Leemans, R. M. May, T. McMichael, P. Pingali, C. n. Samper, R. Scholes, R. T. Watson, A. H. Zakri, Z. Shidong, N. J. Ash, E. Bennett, P. Kumar, M. J. Lee, C. Raudsepp-Hearne, H. Simons, J. Thonell and M. B. Zurek (2005) Ecosystems and Human Well-being: Synthesis, *Millenium Ecosystem Assessment,* Washington DC.

Moretti, M., S. Gloor, N. Bauer, F. Bontadina, P. D. Bruna, P. Duelli, R. Home, M. Hunziker, M. Obrist, T. Sattler and S. Fontana (2010) BiodiverCity: Biodiversität im Siedlungsraum (Zusammenfassung), *unpublished,* Bundesamt für Umwelt, Bern.

Nelson, J. P. (1982) Highway Noise and Property Values, *Journal of Transport Economics and Policy,* **16** (2) 117-138.

Nowak, D. J. (1994) Air Pollution Removal by Chicago's Urban Forest, *Chicago's Urban Forest Ecosystem: Results of the Chicago Urban Forest Climate Project,* Chicago.

Nowak, D. J., P. j. McHale, M. Ibarra, D. Crane, J. C. Stevens and C. J. Luley (1998) *Modeling the Effect of Urban Vegetation on Air Pollution*, Plenum Press, New York.

Phaneuf, D. J., V. K. Smith, R. B. Palmquist and J. C. Pope (2008) Integrating Property Value and Local Recreation Models to Value Ecosystem Services in Urban Watersheds, *Journal of Land Economics,* **84** (3) 361-381.

Piras, G. (2010) sphet: Spatial Models with Heteroskedastic Innovations in R, *Journal of Statistical Software,* **35** (1) 1-21.

R-Development-Core-Team (2012) R: A language and environment for statistical computing, R Foundation for Statistical Computing, Wien.

Richardson, H. W., J. Vipond and R. A. Furbey (1973) Determinants of Urban House Prices, *Urban Studies,* **11** 189-199.

Rosen, S. (1974) Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition, *The Journal of Political Economy,* **82** (1) 34-55.

Salvi, M., P. Schellenberger, H. Schmidt, M. Richter, R. Müri-Leupp and O. Köchle (2004) *Preise, Mieten und Renditen*, Zürcher Kantonalbank, Zürich.

Schirmer, P., B. C. Belart and K. W. Axhausen (2011) Location choice in greater Zurich Area - an intermediate report, *11th Swiss Transport Research Conference,* Institute of Transport Planning and Systems ETH Zurich, Monte Verità, Ascona.

Sirmans, G. S., D. A. Macpherson and E. N. Zietz (2005) The composition of hedonic pricing models: A review of the literature, *Journal of Real Estate Literature,* **13** (3) 3-43.

Taylor, L. O. (2008) Theoretical Foundations and Empirical Developments in Hedonic Modeling, in A. Baranzini, J. V. Ramirez, C. Schaerer and P. Thalmann (Eds.) *Hedonic Methods in Housing Markets,* 15-37, Springer Science+Business Media, New York.

Tschopp, M., P. Fröhlich and K. W. Axhausen (2006) Verkehrsinfrastruktur und räumliche Entwicklung - eine ökonometrische Analyse, *Arbeitsberichte Verkehrs- und Raumplanung,* **352**, IVT, ETH Zürich, Zürich.

Waltert, F. and F. Schläpfer (2010) Landscape amenities and local development: A review of migration, regional economic and hedonic pricing studies, *Ecological Economics,* **70** (2) 141-152.

Wang, Y., K. M. Kockelman and X. C. Wang (2011) Anticipation of Land Use Change Through Use of Geographically Weighted Regression Models for Discrete Response, *Journal of the Transportation Research Board,* **2** (2245) 111-123.

Weber, C., E. Winkler, S. Graf and I. Bähni (2006) *Das Grünbuch der Stadt Zürich*, Grün Stadt Zürich, Zürich.

Weber, R. (2009) Das politische Auftreten von Pro Natura, *Portfolio 100 Jahre Pro Natura,* **3**, Pro Natura, Basel.

Wegener, M. (2004) Overview of land use transport models, in D. Hensher, K. J. Button, K. E. Haynes and P. R. Stopher (Eds.) *Handbook of Transport Geography and Spatial Systems,* 127-146, Elsevier, Oxford.

Weigt, D. (2009) Die Reaktion des Immobilienmarkts auf Fluglärm, *PLANERIN,* **4** (61-62).

Yiu, C. Y. and S. K. Wong (2004) The Effects of Expected Transport Improvements on Housing Prices, *Urban Studies,* **42** (1) 113-125.

Yoo, E.-H. and P. C. Kyriakidis (2009) Area-to-point Kriging in spatial hedonic pricing models, *Journal of Geographical Systems,* **11** (4) 381-400.

Zodrow, G. (1983) *Local Provision of Public Services: The Tiebout Model After Twenty-Five Years*, Academic Press, New York.

A 1 Global Models: Further Results

Table 11 Comparison of STSLS and SARlag estimates (*N*=3084; standardised estimates)

	Model: STSLS			Model: SARlag		
Variable	Estimate	Stand. error	P-Value	Estimate	Stand. error	P-Value
(Intercept)	0.031	0.027	0.247	0.034	0.018	0.053
room	0.147	0.067	0.028	0.144	0.022	0.000
area	0.437	0.093	0.000	0.453	0.023	0.000
age1	0.021	0.073	0.772	0.032	0.061	0.599
age2	0.013	0.109	0.907	0.023	0.085	0.791
age3	-0.010	0.037	0.789	-0.020	0.044	0.651
age4	0.110	0.035	0.002	0.117	0.032	0.000
dhig	-0.004	0.010	0.712	-0.002	0.006	0.755
dcbd	-0.010	0.028	0.727	-0.021	0.026	0.433
dwcbd	0.005	0.009	0.565	0.007	0.006	0.263
drst	-0.003	0.011	0.813	0.000	0.000	0.433
ln(accp)	0.021	0.017	0.224	0.023	0.020	0.261
ln(accm)	-0.022	0.032	0.496	-0.018	0.021	0.388
popd	0.081	0.104	0.437	0.099	0.025	0.000
tax	0.027	0.023	0.234	0.035	0.023	0.139
inco	0.094	0.070	0.184	0.122	0.033	0.000
uni	0.087	0.032	0.006	0.096	0.030	0.002
nonf	0.051	0.019	0.009	0.053	0.026	0.043
vacr	-0.004	0.008	0.656	-0.004	0.006	0.477
airp	0.019	0.033	0.579	0.007	0.009	0.458
nair	0.009	0.009	0.315	0.008	0.014	0.573
nsho	0.000	0.007	0.947	-0.003	0.004	0.474
nstr	-0.013	0.015	0.388	-0.011	0.010	0.273
nrail	-0.038	0.012	0.002	-0.041	0.015	0.007
dlake	-0.042	0.016	0.010	-0.047	0.017	0.007
driv	-0.005	0.012	0.668	-0.005	0.004	0.313
dagr	0.034	0.036	0.346	0.044	0.021	0.040
dfor	0.023	0.016	0.163	0.024	0.015	0.112

dpark	-0.005	0.009	0.563	-0.006	0.008	0.415
apark	0.002	0.008	0.774	0.002	0.003	0.533
afor	0.012	0.011	0.288	0.009	0.009	0.315
aagr	0.000	0.009	0.965	-0.001	0.001	0.541
awat	0.048	0.023	0.040	0.052	0.014	0.000
aloc	0.004	0.009	0.625	0.007	0.007	0.314
indi	-0.019	0.011	0.076	-0.017	0.010	0.084
vlake	0.016	0.021	0.445	0.024	0.013	0.066
vmou	0.000	0.010	0.981	0.001	0.001	0.488
Rho	0.264	0.353	0.455			
SSE / LL	1698			-3469	38 deg. of freedom	

Note: In Table 11 and Table 15 results of models, in which multicollinearity may be present, are showed.

Table 12 STSLS parameters including sunshine index (*N*=3084; standardised coefficients)

	Model 1': ln(rent) ~ variables			Model 2': rent ~ variables		
Variable	Estimate %	Stand. error	P-Value	Estimate	Stand. error	P-Value
(Intercept)	0.094	0.022	0.000	0.033	0.028	0.251
area	0.725	0.057	0.000	0.553	0.054	0.000
age1	0.085	0.052	0.103	0.018	0.083	0.830
age2	0.080	0.075	0.290	0.020	0.115	0.860
age3	0.065	0.033	0.045	0.007	0.036	0.847
age4	0.267	0.037	0.000	0.108	0.035	0.002
ln(accs)	0.049	0.016	0.003	0.026	0.032	0.412
popd	0.046	0.023	0.047	0.091	0.077	0.238
uni	0.188	0.014	0.000	0.147	0.068	0.030
vacr	-0.007	0.010	0.462	-0.006	0.008	0.445
nair	0.019	0.007	0.009	0.018	0.006	0.001
nsho	-0.003	0.006	0.654	-0.001	0.005	0.840
nstr	-0.001	0.010	0.901	-0.015	0.016	0.376
nrail	-0.011	0.010	0.291	-0.024	0.010	0.021
dlake	-0.060	0.009	0.000	-0.046	0.017	0.008
driv	-0.008	0.012	0.479	0.002	0.012	0.901
dpark	0.003	0.009	0.743	-0.009	0.009	0.359
apark	0.004	0.008	0.643	-0.004	0.009	0.671
afor	0.003	0.008	0.728	0.003	0.010	0.781
aagr	-0.006	0.007	0.375	0.000	0.010	0.961
awat	0.023	0.012	0.062	0.045	0.022	0.042
indi	-0.028	0.010	0.005	-0.024	0.013	0.051
vlake	0.041	0.010	0.000	0.017	0.019	0.352
vmou	0.007	0.012	0.575	-0.008	0.010	0.412
sunshine	0.035	0.011	0.002	0.029	0.018	0.115
Rho	0.173	0.041	0.000	0.235	0.309	0.446
SSE	855				1727	

Table 13 STSLS examples of model variants (*N*=3084;standardised coefficients)

	Model 2a: rent ~ variables			Model 2b: rent ~ variables		
Variable	Estimate %	Stand. error	P-Value	Estimate	Stand. error	P-Value
(Intercept)	0.028	0.028	0.314	0.033	0.027	0.220
ln(area)	0.518	0.042	0.000	0.514	0.038	0.000
age1	0.073	0.093	0.432	0.117	0.083	0.161
age2	0.028	0.130	0.828	0.052	0.122	0.669
age3	-0.026	0.042	0.546	-0.026	0.041	0.523
age4	0.090	0.031	0.004	0.095	0.032	0.003
ln(accs)	0.023	0.034	0.492	0.039	0.017	0.023
ln(popd) [a]	0.099	0.083	0.230	0.070	0.034	0.041
ln(uni) [b]	0.142	0.090	0.114	0.174	0.098	0.076
vacr	-0.012	0.009	0.202	-0.005	0.007	0.505
nair	0.019	0.006	0.003	0.011	0.007	0.120
nsho	0.000	0.009	0.994	0.002	0.008	0.748
nstr	-0.014	0.018	0.422	-0.013	0.018	0.459
nrail	-0.042	0.016	0.008	-0.047	0.017	0.005
dlake	-0.049	0.025	0.047	-0.042	0.022	0.051
ln(driv)	0.011	0.012	0.396	0.016	0.015	0.307
dpark	-0.011	0.011	0.276	-0.009	0.010	0.408
apark	-0.006	0.010	0.537	-0.004	0.010	0.666
afor	0.002	0.011	0.882	-0.011	0.016	0.495
aagr	0.002	0.012	0.893	0.001	0.012	0.954
awat	0.052	0.030	0.078	0.048	0.029	0.093
indi	-0.019	0.013	0.137	-0.011	0.011	0.333
vlake	0.005	0.023	0.815	0.003	0.022	0.902
vmou	-0.008	0.012	0.479	-0.008	0.012	0.533
Rho	0.290	0.433	0.504	0.262	0.417	0.530
SSE	1870				1863	

Note: ln(popd) in model 2a (left side); ln(uni) in model 2b (right side)

Table 14 STSLS parameters of variant rent per square metre (*N*=3084; standardised coeff.)

Model 6a: rent/area ~ variables

Variable	Estimate %	Stand. error	P-Value
(Intercept)	0.026	0.341	0.940
age1	0.048	2.100	0.982
age2	-0.025	1.326	0.985
age3	-0.032	1.215	0.979
age4	0.057	0.434	0.895
ln(accs)	0.029	2.372	0.990
uni	0.056	4.127	0.989
nstr	-0.029	0.430	0.946
nrail	-0.004	1.036	0.997
dlake	-0.008	0.901	0.993
dpark	-0.004	0.511	0.995
awat	0.014	0.032	0.665
vlake	0.004	0.206	0.985
Rho	0.732	20.805	0.972
SSE	3334		

Table 15 STSLS unstandardised coefficient estimates (N=3084)

Model: ln(rent) ~ variables				Model: rent ~ variables		
Variable	Estimate %	Stand. error	P-Value	Estimate	Stand. error	P-Value
(Intercept)	470.000	0.287	0.000	2343.800	433.740	0.000
room	12.110	0.024	0.000	184.650	84.177	0.028
area	0.467	0.001	0.000	15.758	3.365	0.000
age1	4.366	0.022	0.046	35.336	122.040	0.772
age2	1.262	0.034	0.706	21.314	181.790	0.907
age3	1.309	0.012	0.285	-16.464	61.633	0.789
age4	13.033	0.015	0.000	184.030	59.364	0.002
dhig	-0.003	0.000	0.932	-0.394	1.065	0.712
dcbd	-0.031	0.000	0.006	-0.228	0.652	0.727
dwcbd	0.011	0.000	0.069	0.119	0.206	0.565
drst	0.010	0.001	0.876	-0.579	2.442	0.813
ln(accp)	0.000	0.000	0.010	0.002	0.002	0.224
ln(accm)	0.000	0.000	0.774	-0.001	0.001	0.496
popd	0.049	0.001	0.464	6.424	8.264	0.437
tax	0.010	0.001	0.862	3.448	2.899	0.234
inco	0.201	0.001	0.000	7.421	5.586	0.184
uni	1.159	0.002	0.000	32.505	11.845	0.006
nonf	0.441	0.001	0.001	15.973	6.102	0.009
vacr	-0.153	0.004	0.693	-4.965	11.129	0.656
airp	0.000	0.000	0.311	0.007	0.012	0.579
nair	-0.001	0.000	0.937	0.495	0.493	0.315
nsho	-0.017	0.000	0.645	-0.079	1.193	0.947
nstr	0.020	0.001	0.755	-2.964	3.430	0.388
nrail	-0.062	0.000	0.074	-4.422	1.431	0.002
dlake	-0.039	0.000	0.000	-1.299	0.505	0.010
driv	-0.405	0.003	0.114	-4.351	10.160	0.668
dagr	1.780	0.004	0.000	29.938	31.767	0.346
dfor	0.055	0.002	0.787	14.401	10.312	0.163
dpark	0.081	0.001	0.573	-3.234	5.595	0.563

apark	0.029	0.000	0.151	0.193	0.671	0.774
afor	0.000	0.000	0.523	0.007	0.007	0.288
aagr	-0.003	0.000	0.423	0.008	0.180	0.965
awat	0.012	0.000	0.005	0.694	0.337	0.040
aloc	0.066	0.000	0.381	1.218	2.495	0.625
indi	-1.516	0.010	0.114	-64.920	36.648	0.076
vlake	0.304	0.001	0.000	4.420	5.792	0.445
vmou	0.422	0.002	0.073	-0.192	7.946	0.981
Rho	0.191	0.034	0.000	0.264	0.353	0.455
SSE	160.025			4’753’078’082		

Note: In Table 11 and Table 15 results of models, in which multicollinearity may be present, are showed.

Table 16 STSLS unstandardised coefficient estimates without multicollinearity (N=3084)

Model 1: ln(rent) ~ variables				Model 2: rent ~ variables		
Variable	Estimate %	Stand. error	P-Value	Estimate	Stand. error	P-Value
(Intercept)	493.500	0.280	0.000	1417.500	870.480	0.103
area	0.742	0.001	0.000	20.002	1.979	0.000
age1	4.038	0.024	0.099	31.017	138.950	0.823
age2	4.228	0.036	0.241	47.601	195.530	0.808
age3	3.073	0.015	0.047	11.468	60.920	0.851
age4	12.585	0.017	0.000	179.600	59.546	0.003
ln(accs)	3.378	0.011	0.002	64.887	74.987	0.387
popd	0.105	0.001	0.044	7.211	6.145	0.241
uni	2.012	0.001	0.000	55.659	26.261	0.034
vacr	-0.228	0.004	0.559	-6.843	10.990	0.533
nair	0.031	0.000	0.005	1.053	0.296	0.000
nsho	-0.023	0.000	0.487	-0.425	0.986	0.667
nstr	-0.005	0.001	0.944	-3.247	3.786	0.391
nrail	-0.056	0.000	0.109	-3.379	1.374	0.014
dlake	-0.050	0.000	0.000	-1.331	0.516	0.010
driv	-0.166	0.003	0.543	2.094	10.294	0.839
dpark	0.034	0.001	0.822	-5.644	5.738	0.325
apark	0.008	0.000	0.666	-0.321	0.729	0.659
afor	0.000	0.000	0.560	0.003	0.006	0.654
aagr	-0.004	0.000	0.338	-0.018	0.194	0.928
awat	0.009	0.000	0.064	0.658	0.326	0.043
indi	-2.794	0.010	0.003	-87.132	42.062	0.038
vlake	0.296	0.001	0.000	4.048	4.940	0.413
vmou	0.304	0.003	0.233	-1.499	8.349	0.857
Rho	0.179	0.042	0.000	0.241	0.314	0.442
SSE	192.330				4’845’683’026	

Table 17 STSLS unstandardised estimates for the most important variables (N=3084)

	Model 5: ln(rent) ~ variables			Model 6: rent ~ variables		
Variable	Estimate %	Stand. error	P-Value	Estimate	Stand. error	P-Value
(Intercept)	492.170	0.278	0.000	2457.000	428.860	0.000
area	0.740	0.001	0.000	19.886	1.868	0.000
age1	5.021	0.023	0.027	112.150	105.120	0.286
age2	4.950	0.035	0.157	113.710	173.600	0.512
age3	3.032	0.016	0.052	1.066	59.768	0.986
age4	12.386	0.018	0.000	167.870	59.943	0.005
ln(accs)	5.378	0.009	0.000	200.270	41.393	0.000
uni	2.118	0.001	0.000	64.599	22.549	0.004
nstr	0.004	0.001	0.950	-2.775	3.702	0.454
nrail	-0.067	0.000	0.056	-3.745	1.222	0.002
dlake	-0.051	0.000	0.000	-1.421	0.450	0.002
dpark	-0.074	0.001	0.614	-12.155	5.658	0.032
awat	0.008	0.000	0.113	0.572	0.314	0.069
vlake	0.285	0.001	0.000	2.894	4.695	0.538
Rho	0.159	0.043	0.000	0.194	0.261	0.458
SSE	194.046				4’873’341’116	

A 2 Local Models: Further Results

Figure 11 GWR: age1 and age2 (model 8)

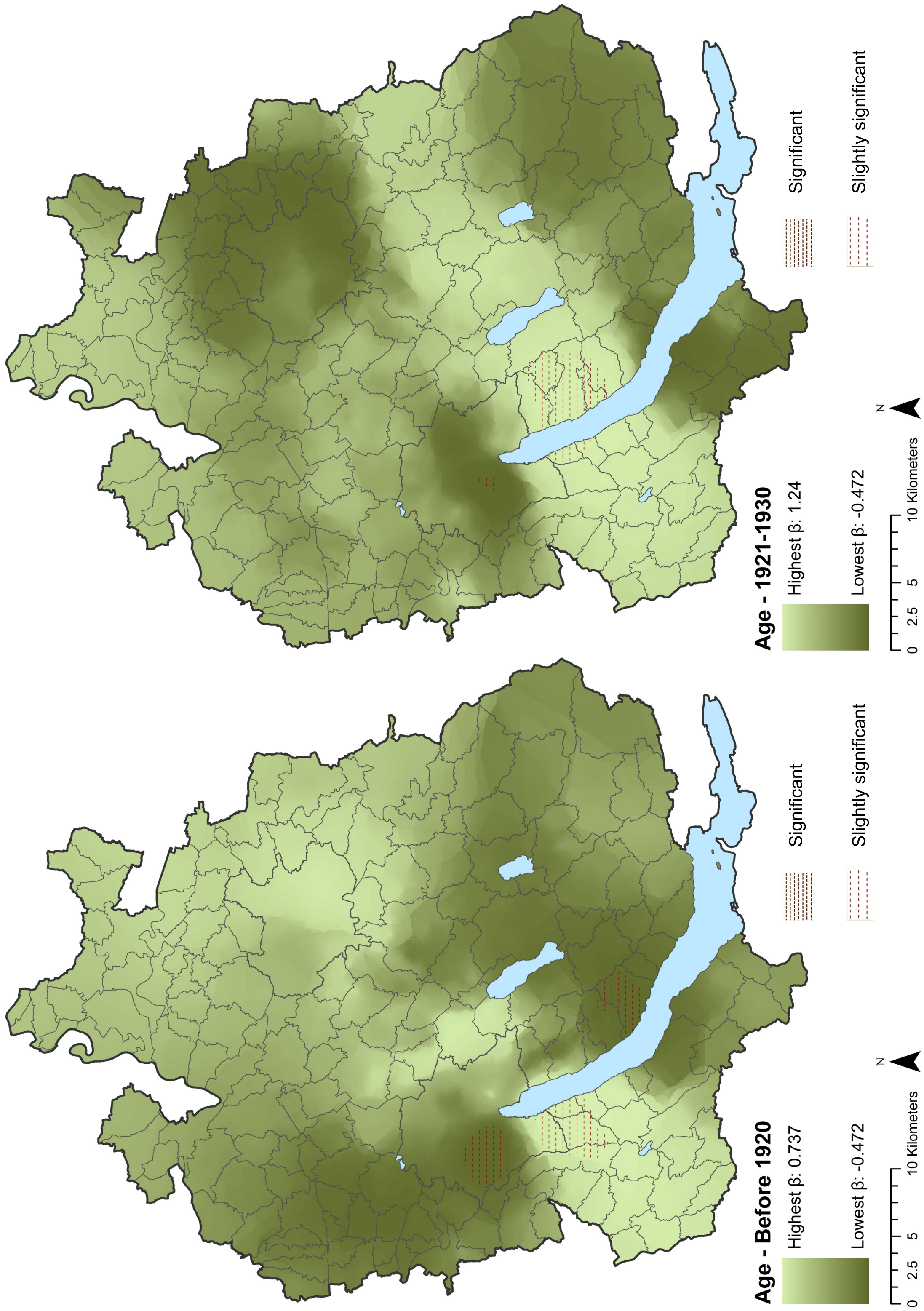


Figure 12 GWR: age3 and age4 (model 8)

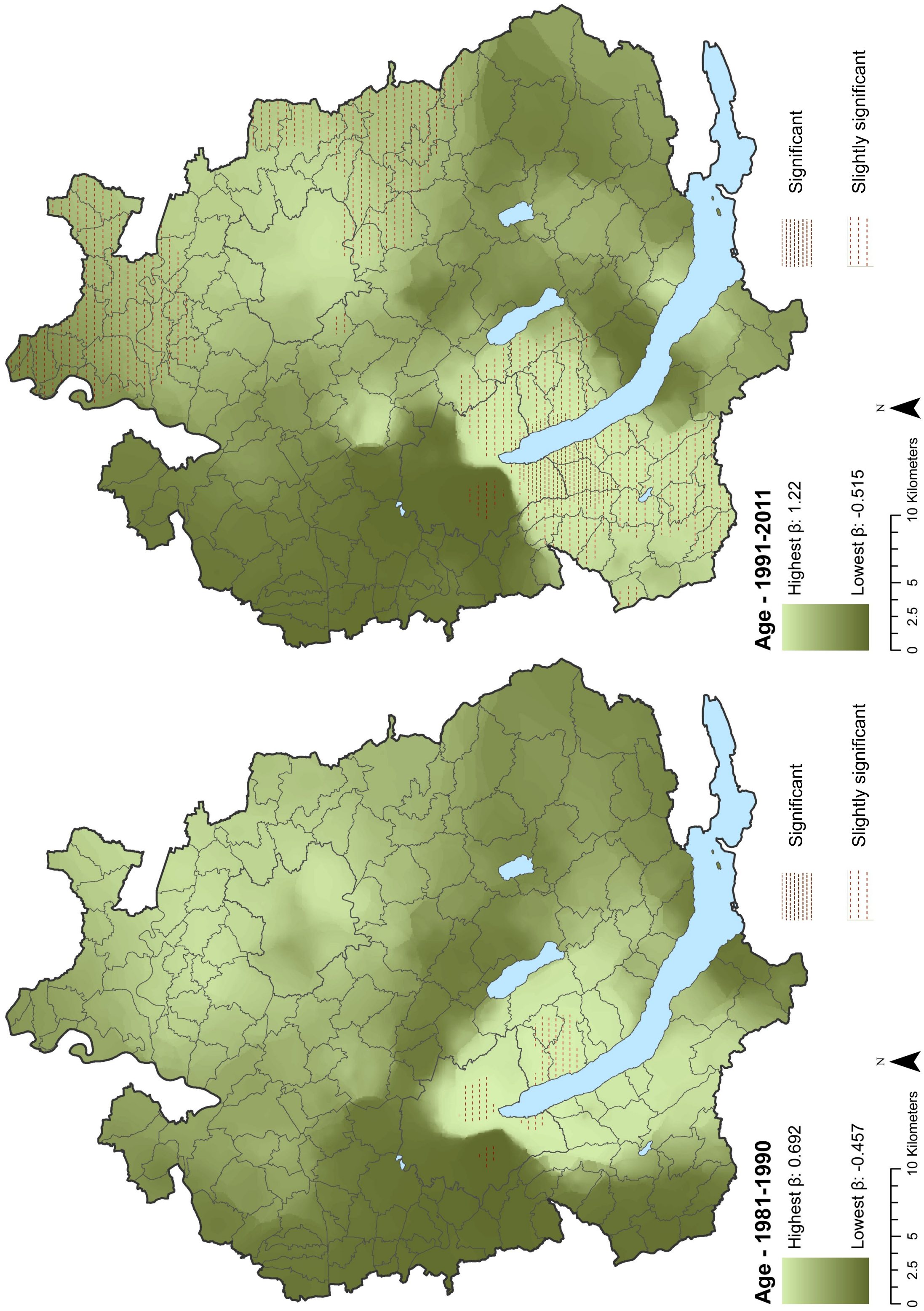


Figure 13 GWR: Population density and vacancy rate (model 7)

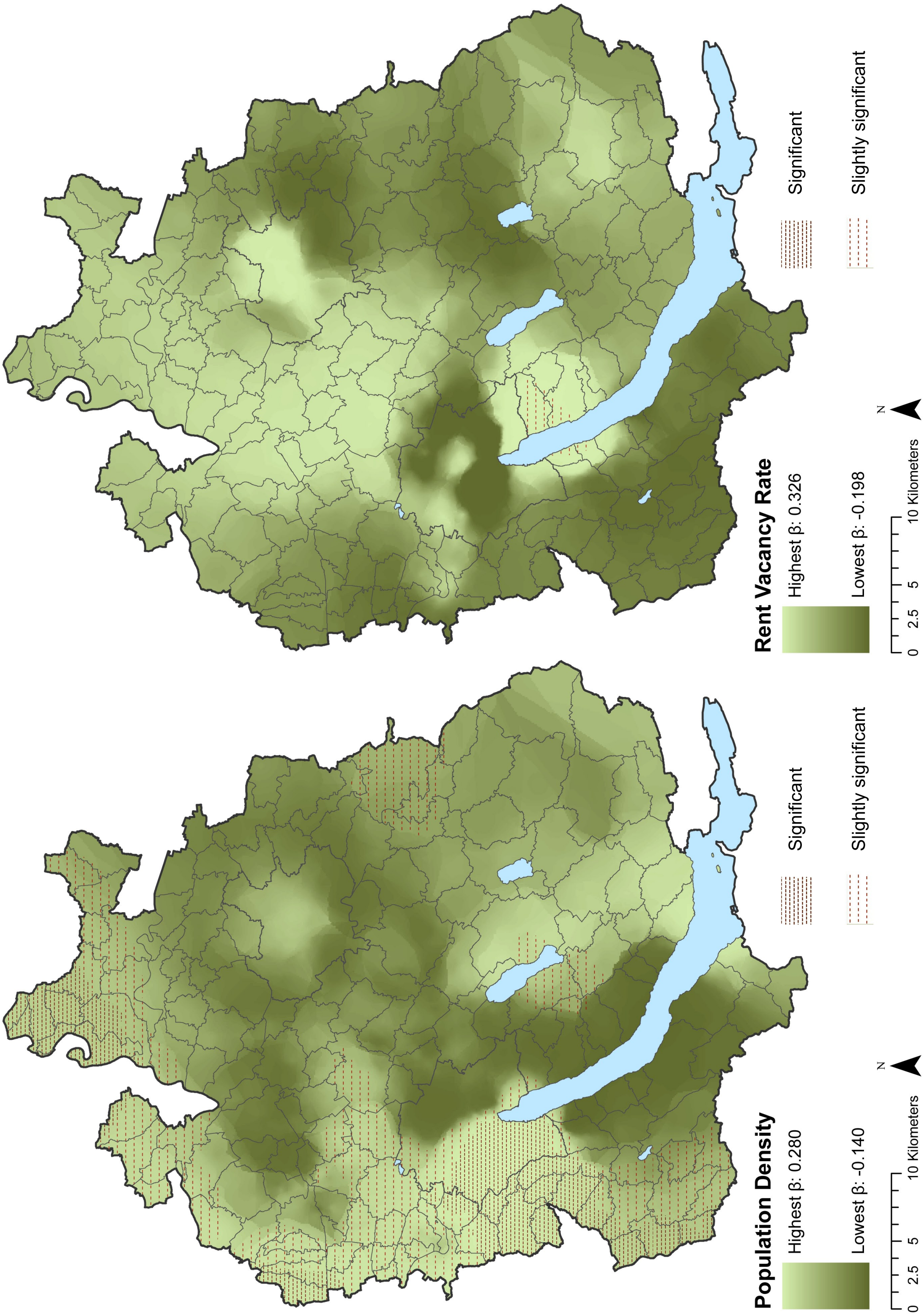


Figure 14 GWR: Aircraft noise and noise by shooting ranges (model 7)

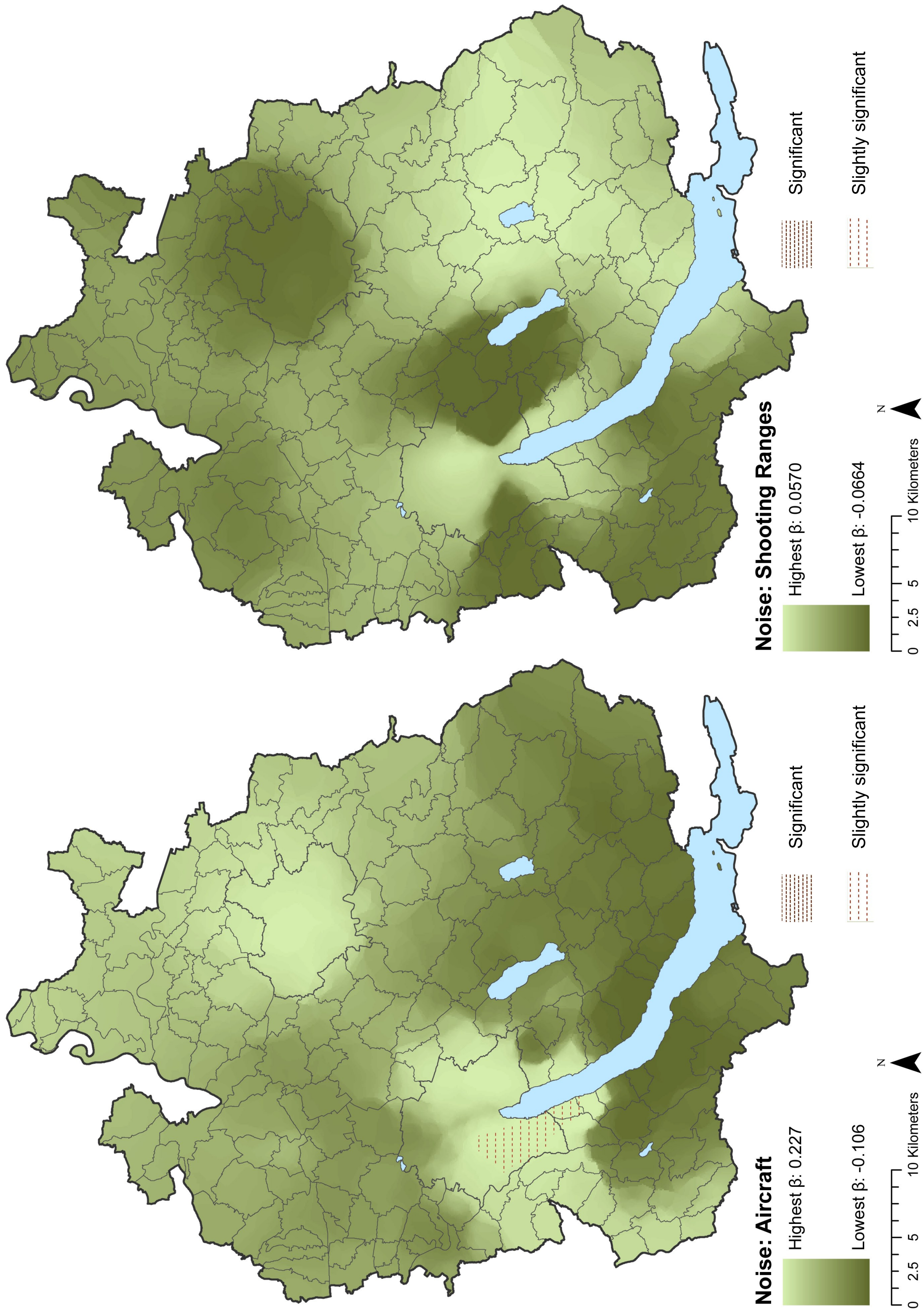


Figure 15 GWR: Street noise (model 8) and distance to a river (model 7)

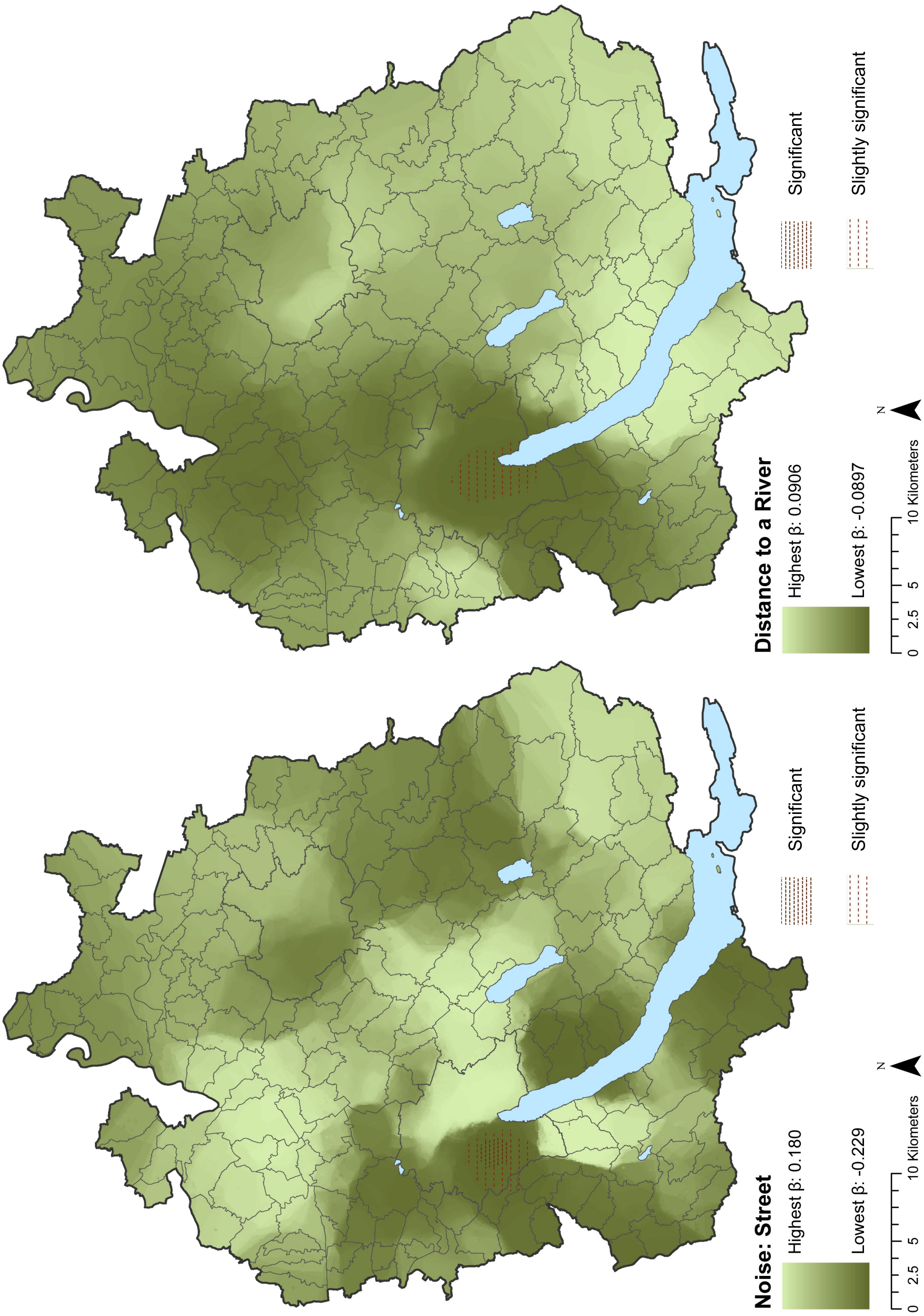


Figure 16 GWR: Area of agricultural land and area of forest (model 7)