

Die approbierte Originalversion dieser Dissertation ist in der Hauptbibliothek der Technischen Universität Wien aufgestellt und zugänglich.

<http://www.ub.tuwien.ac.at>



The approved original version of this thesis is available at the main library of the Vienna University of Technology.

<http://www.ub.tuwien.ac.at/eng>



TECHNISCHE
UNIVERSITÄT
WIEN
Vienna University of Technology

Dissertation

Modellierung von wertrelevanten Mikrolageparametern für die automatisierte Immobilienbewertung

ausgeführt zum Zwecke der Erlangung des akademischen Grades eines
Doktors der technischen Wissenschaften unter der Leitung

von

ao.Univ.-Prof. Dipl.-Ing. Dr.techn. Wolfgang Feilmayr (Erstbegutachter)
Fachbereich Stadt- und Regionalforschung (E280/2)
Department für Raumplanung

und

Univ.-Prof. Mag. Dr. Michael Getzner (Zweitbegutachter)
Fachbereich Finanzwissenschaft und Infrastrukturpolitik (E280/3)
Department für Raumplanung

eingereicht an der Technischen Universität Wien
der Fakultät für Architektur und Raumplanung

von

Dipl.-Ing. Ronald S. Weberndorfer
Matr.Nr.: 0125493
1160 Wien, Austria

Wien, am 1. August 2013

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst habe. Es wurden nur die in der Arbeit ausdrücklich benannten Quellen und Hilfsmittel benutzt. Wörtlich oder sinngemäß übernommenes Gedankengut habe ich als solches kenntlich gemacht.

Ich versichere, dass ich dieses Thema bisher weder im In- noch im Ausland in irgendeiner Form als Prüfungsarbeit vorgelegt habe.

Wien, am 1. August 2013

Ronald S. Weberndorfer

Danksagung

Diese Arbeit bietet mir neben der Dokumentation der Modellierungstätigkeiten aus den letzten Jahren auch die Möglichkeit, den Menschen zu danken, die mich bei der Anfertigung meiner Dissertation unterstützt haben.

Aufrichtigen Dank an meinen Doktorvater ao.Univ.Prof. Dipl.-Ing. Dr.techn. Wolfgang Feilmayr, der für die unterschiedlichsten wissenschaftlichen Fragestellungen immer die richtigen Ratschläge parat hatte und mir dadurch zu neuen Perspektiven verhelfen konnte. Ich danke auch meinem Kollegen Dipl.-Ing.(FH) Mag. Dr. Wolfgang Brunauer bei der Immobilien Rating GmbH. Er hat mich durch sein Interesse und seine kritischen Fragen immer wieder motivieren können, um diese Arbeit zum Abschluss zu bringen. Weiters bedanke ich mich bei meinem Arbeitgeber, der Immobilien Rating GmbH (Ein Unternehmen der Bank Austria UniCredit), für die Bereitstellung der Datengrundlagen und der für die Analysen notwendigen EDV-Infrastruktur.

Besonderer Dank gebührt außerdem meiner Lebensgefährtin Lenka sowie meiner Familie und meinen Freunden, die in den letzten Monaten immer ein offenes Ohr für mich hatten und somit ebenfalls zum Gelingen dieser Arbeit beigetragen haben.

Wien, am 1. August 2013

Ronald S. Weberndorfer

Kurzfassung

Der Hauptfokus dieser Arbeit liegt auf der Modellierung von kleinräumigen (mikro) Lageeffekten in einem Geographischen Informationssystem (GIS) sowie der Auswirkungsanalyse dieser Einflüsse auf Immobilienpreise bei Standardwohnobjekten (Einfamilienhaus und Eigentumswohnung).

Für die Modellierung der Mikrolagevariablen in der Analyseregion Österreich wird hauptsächlich auf Rastertechnik mit einer Auflösungsgenauigkeit von bis zu 30m gesetzt. Durch die mittels Python automatisierte GIS-gestützte Generierung von neuen Mikrolagefaktoren – wie beispielsweise Wohnumfeld, Lärm, Erreichbarkeit oder Höhenlage – ist es möglich, ein hochauflösendes Datenmodell aufzubauen.

Mithilfe von semiparametrischen Regressionsmodellen wurden diese Variablen analysiert und festgestellt, dass durch die exakt verorteten Mikrolagevariablen erhebliche Effekte in einem hedonischen Preismodell nachweisbar sind. Die Verwendung von hochauflösenden Raster-Mikrolagevariablen wie z.B. Lärm erklären Wertveränderungen deutlich und verbessern hierdurch in statistisch-räumlichen Analysemodellen die Schätzqualität signifikant.

Schlüsselwörter: Geographisches Informationssystem (GIS), Generalisierte Additive Modelle (GAM), nichtlineare Regression, hybrides Modell, statistisches Modell, Standort, Lage, Mikrolage, Lagefaktoren, Immobilienpreise, hedonische Preise, Raster, Python, ArcGIS, R, Lagemodell, Lärm, Höhenmodell, Erreichbarkeitsmodell, Regionalstatistischer Raster, Gebäudedaten, Variablenselektion, Vereinfachtes Backward Stepwise, Modellvergleiche

Abstract

The main focus of this work concentrates on the modeling of small-scaled (micro) location effects in geographic information systems (GIS) and the impact analysis of these influences on real estate prices for standard residential properties (houses and flats).

The modeling of micro-location variables is mainly based on grid technology with a resolution accuracy of up to 30m all over the analysis region Austria. By using Python the automated GIS-based generation of new micro-location factors – such as residential environment, noise, accessibility or elevation – allows to construct a high-resolution data model.

These variables are analyzed with semi-parametric regression models. It was determined that through the exact rooted micro-location variables in a hedonic price model significant effects are detectable. The use of high resolution micro-location variables e.g. noise declare value changes clearly and improve the estimation quality in statistical spatial analysis models significantly.

Keywords: geographic information systems (GIS), generalized additive models (GAM), nonlinear regression, hybrid model, statistical model, location, micro-location, location factors, house prices, hedonic prices, grid, raster, Python, ArcGIS, R, location model, noise, elevation model, accessibility model, regional statistical grid, building data, variable selection, simplified backward stepwise, model comparison

Inhaltsverzeichnis

Eidesstattliche Erklärung	i
Danksagung	iii
Kurzfassung	v
Abstract	vii
Inhaltsverzeichnis	xiii
1 Einleitung	1
1.1 Motivation, Zielsetzung der Arbeit	2
1.2 Einordnung der Arbeit	3
1.2.1 Forschungsfrage	3
1.2.2 Abgrenzung	4
1.3 Ablauf der Arbeit	4
2 Theoretische Grundlagen	7
2.1 Hedonische Preismodelle	7
2.1.1 Regressionsmodelle	9
2.1.2 Lineares Modell	10
2.1.3 Generalisierte Additive Modelle (GAM)	10
2.1.3.1 Praktische Umsetzung	11
2.1.4 Literaturübersicht	11
2.1.4.1 Abgrenzung zu aktuellen Modellen	12

2.2	GIS Grundlagen	13
2.2.1	Datenmodelle	13
2.2.2	Map Algebra	14
2.2.3	Automatisierung von GIS-Prozessen	14
2.2.3.1	Programmiersprache Python	15
3	Generierung der Mikrolagevariablen	17
3.1	Landnutzungsmodell	17
3.1.1	Automatisierte Erstellung der Landnutzungsoberfläche	21
3.1.1.1	Datenimport in GIS-Datenbank	22
3.1.1.2	Erzeugung von Nutzungsflächen aus dem Straßengraphen	23
3.1.1.3	Zusammenführung aller Datensätze	26
3.1.2	Landnutzungsdatenschicht	27
3.1.2.1	Plausibilisierung der Datenschicht	29
3.2	Höhenmodell	31
3.2.1	Ausrichtung	33
3.2.2	Geländeneigung	34
3.2.3	Geländekrümmung	36
3.2.4	Geländeschummerung	40
3.3	Lärmmodell	45
3.3.1	Berechnungsmethode der Lärmdatenschicht	47
3.3.2	Lärm-Basisformel	49
3.3.2.1	Dämpfungskonstante	50
3.3.2.2	Pegelveränderung - Bebauung	51
3.3.2.3	Pegelveränderung - Steigung	53
3.3.2.4	Pegelveränderung - Flächendämpfung	54
3.3.3	Zusammenführung der Lärmvariablen	57
3.3.4	Anmerkung zur technischen Umsetzung	61
3.4	Erreichbarkeitsmodell	61
3.4.1	Datenvorbereitung	62

3.4.2	Konvertierung des Straßengraphen in einen Raster	63
3.4.3	Reisegeschwindigkeit	64
3.4.4	Wahl der Zielpunkte	67
3.4.5	Reisekostenberechnung	68
3.4.5.1	Validierung der ermittelten Reisekosten	72
3.4.6	Erreichbarkeit	76
3.5	Sonstige Variablen	83
3.5.1	Regionalstatistischer Raster	83
3.5.2	Lebensphasen	85
3.5.3	Gebäudedaten	85
3.5.4	Klimadaten	85
3.6	Verknüpfung der Variablen mit den Immobiliendaten	86
4	Validierung der Mikrolagevariablen mit einem Generalisierten Additiven Modell (GAM)	89
4.1	Basismodell	89
4.1.1	Datenbeschreibung der Immobiliendaten	90
4.1.2	Modellaufbau	92
4.2	Variablenselektion	96
4.2.1	Vereinfachtes Backward Stepwise	97
4.2.1.1	Technische Umsetzung	98
4.2.2	Praktische Umsetzung der Selektion	99
4.3	Finale Modelle und Schlussfolgerungen	100
4.3.1	Ergebnisse aus den linearen Funktionen	100
4.3.2	Ergebnisse der nichtlinearen Funktionen	102
4.3.3	Mikrolagevariablen im Modell Einfamilienhaus (EFH)	105
4.3.3.1	Lineare Mikrolagevariablen (EFH)	105
4.3.3.1.1	Ausländeranteil	105
4.3.3.1.2	Geländekrümmung	106
4.3.3.1.3	Akademikeranteil	107

4.3.3.2	Nichtlineare Mikrolagevariablen (EFH)	109
4.3.3.2.1	Erreichbarkeit	109
4.3.3.2.2	Lärm	110
4.3.3.3	Modellqualität EFH	112
4.3.4	Mikrolagevariablen im Modell Eigentumswohnung (ETW)	113
4.3.4.1	Lineare Mikrolagevariablen (ETW)	114
4.3.4.1.1	Ausländeranteil	114
4.3.4.1.2	Arbeitsstätten pro Einwohner	114
4.3.4.1.3	Anzahl der Wohngebäude	115
4.3.4.1.4	Speckgürtel rund um Städte	116
4.3.4.2	Nichtlineare Mikrolagevariablen (ETW)	118
4.3.4.2.1	Akademikeranteil	118
4.3.4.2.2	Lärm	118
4.3.4.2.3	Erreichbarkeit	119
4.3.4.3	Modellqualität ETW	120
4.4	Modellvergleiche	121
4.4.1	EFH	122
4.4.2	ETW	123
4.4.3	Allgemeines zum Modellvergleich	123
5	Schlussbetrachtungen	125
5.1	Zusammenfassung der Ergebnisse	125
5.1.1	Modell EFH	127
5.1.2	Modell ETW	129
5.1.3	Modell EFH vs. ETW	130
5.2	Diskussion und Ausblick	132
5.2.1	Praktische Verwendung der Ergebnisse in einer Bewertungssoftware	133
Anhang A	Verzeichnisse	135
A.1	Abkürzungsverzeichnis	135
A.2	Literaturverzeichnis	137

A.3	Internet- und sonstige elektronische Quellen	140
A.4	Rechtsquellen	143
A.5	Abbildungsverzeichnis	144
A.6	Tabellenverzeichnis	147
A.7	Verwendete Software	148
 Anhang B Tabellen und Grafiken		 149
B.1	Erreichbarkeit	149
B.2	Projektion im GIS	152
B.2.1	Zielprojektion	152
B.2.2	Importprojektionen	153
B.3	Modellvariablen	154
B.4	Anzahl im Grundbuch erfasster Transaktionen	158
B.4.1	Transaktion pro Bundesland	159
B.4.2	Transaktion pro Jahr	159
B.5	Vergleichsmodelle	160
B.5.1	Das Basismodell nicht hybrid, ohne Raster	160
B.5.1.1	EFH	160
B.5.1.2	ETW	162
B.5.2	Das Basismodell nicht hybrid, mit Raster	164
B.5.2.1	EFH	164
B.5.2.2	ETW	167
B.5.3	Das Basismodell hybrid, ohne Raster	170
B.5.3.1	EFH	170
B.5.3.2	ETW	173

Kapitel 1

Einleitung

In der Immobilienwirtschaft wird seit Jahren immer wieder der Slogan „Lage, Lage, Lage“ zitiert. Egal ob Makler, Immobilieninvestoren, private oder sonstige Akteure am Immobilienmarkt – alle sind sich einig – die Lage eines Objekts ist der wichtigste Faktor für eine Kaufentscheidung und dementsprechend auch für den erzielbaren Preis an einem Standort.

Es drängen sich somit zwei Fragen auf:

- Was ist eigentlich eine „gute“ oder „schlechte“ Lage und wie bzw. durch welche Lagefaktoren kann dies definiert werden?
- Wie hoch ist der Wert einer geographischen Lage?

Die Antwort darauf scheint für die agierenden Individuen auf den ersten Blick relativ einfach zu sein, da jeder seine Bedürfnisse und Anforderungen an die gewünschte Liegenschaft kennt und die Kriterien für seine eigene „optimale Lage“ entsprechend abwägen kann. Es ist klar, dass derartige Lageeinschätzungen/-entscheidungen von privaten und professionellen Akteuren am Immobilienmarkt täglich getroffen werden und darauf auch die Kaufentscheidungen basieren. Will man nun aber die Lage allgemein beschreiben, um die Kaufentscheidungen einzelner Individuen nachvollziehbar zu machen, wird dies schnell kompliziert, da man keine Verallgemeinerung für die verschiedenen Immobilientypen treffen kann. Beispielsweise wird für einen Gewerbe- oder Industriebetrieb eine ruhige Wohnlage wohl kaum ein Kriterium für gute Lage sein, hier sind je nach Betriebstyp eher Themen wie Bahn-/ Autobahnanschluss oder die Emissionsmöglichkeit

von Lärm oder Geruch die wichtigen Faktoren. Eine ruhige Wohnlage kann in diesem Fall sogar negativ behaftet sein, da sich dadurch Nutzungskonflikte mit Nachbarn ergeben können. Dementsprechend bedarf es – um eine qualifizierte Aussage darüber zu treffen, welche Faktoren zu einer guten bzw. schlechten Lage führen – zuerst einer Festlegung, welches Immobiliensegment man diesbezüglich beschreiben will. Diese Arbeit beschränkt sich auf die Beschreibung der Lagefaktoren am Wohnimmobilienmarkt.

Damit diese Fragestellung allgemeingültig beantwortet werden kann, ist es im wissenschaftlichen Sinn nur in Bezug auf geostatistische Einheiten möglich. In Österreich gibt es dazu auf verschiedenen hierarchischen Ebenen (Bundesland, Bezirk, Gemeinde und Zählsprengel) und seit kurzem auch als Raster soziodemographische Daten, die für eine derartige Analyse herangezogen werden können. Mit diesen Daten ist es möglich, eine Aussage über den Einfluss der soziodemographischen Faktoren bis auf die Zählsprengel (ZSP)- oder Rasterebene herunter zu treffen. Dies bedeutet aber auch, dass Mikrolagevariablen wie z.B. Lärm, Aussicht oder Entfernung bis zur nächsten Bushaltestelle etc. damit nicht ausreichend beantwortet werden können, sondern nur jene Variablen, welche für die jeweiligen hierarchischen Ebenen vorhanden sind, wie etwa Kaufkraft, Bevölkerungsveränderung, Bildung, Erreichbarkeit, etc..

1.1 Motivation, Zielsetzung der Arbeit

In der klassischen Immobilienbewertung übernimmt der Gutachter die Einschätzung aller Lageeigenschaften, und es kann somit eine sehr kleinräumige und präzise Lagebewertung durchgeführt werden, da dieser die Objekte auch besichtigt und dadurch natürlich auch das kleinräumige Umfeld gut einschätzen kann.

Bei der automatisierten Immobilienbewertung ist dies anders, hier wird ein hedonisches Preismodell zur Bewertung von Immobilien herangezogen. Dadurch ist es nicht möglich, wie bei einer konventionellen Bewertung alle Lageeigenschaften manuell erfassen zu lassen. Das System (Bewertungsmodell) sollte diese Fragen automatisch aus einer Datenbank beantworten und ohne weiteres Zutun dem Bewerter einen realistischen Wert liefern. Grundsätzlich funktioniert dies auch vom Preisniveau her gut, und

es kann ein Mittelwert der kleinsten räumlichen Einheit (Zählsprenkel) ausgewiesen werden. Es gibt allerdings derzeit noch die Einschränkung, dass die Mikrolagefaktoren noch nicht ausreichend erklärt werden, da für diese keine bzw. zu wenige passende Datenquellen verfügbar sind.

Die Herausforderung in dieser Arbeit wird es sein, verschiedene immobilienwertbeeinflussende Mikrolagefaktoren, welche derzeit nicht verfügbar sind, mit einem GIS zu modellieren, um mit diesen kleinräumigen Lagevariablen auch die Mikrolageeinflüsse auf den Immobilienwert mit einem hedonischen Preismodell simulieren zu können. Die Hauptaufgabe in dieser Arbeit ist es demnach, detaillierte Lageinformationen zusammenzutragen, diese Daten aufzubereiten und einer „modelltechnischen“ Veredelung zuzuführen und in einem zweiten Schritt diese Geodaten mittels eines hedonischen Preismodells auf ihre Wertrelevanz hin zu überprüfen.

Als Ergebnis wäre es wünschenswert, dass man zu den unterschiedlichen modellierbaren Mikrolagevariablen eine Aussage treffen könnte, in welcher Art diese eine wertbeeinflussende Wirkung auf den Marktwert einer Immobilie haben.

1.2 Einordnung der Arbeit

1.2.1 Forschungsfrage

Wie können hochauflösende Mikrolagevariablen mit einem GIS flächendeckend modelliert und in ein (bestehendes) hedonisches Preismodell eingebunden werden, um damit bislang unberücksichtigte Mikrolageeffekte abzubilden?

Daraus leiten sich folgende Sub-Fragen ab:

- Mit welchen Basisdaten können die dafür notwendigen Mikrolagevariablen erzeugt werden?
- Wie funktioniert die GIS-technische Umsetzung und Modellbildung bei der Variablengenerierung?

- Wie können die Mikrolagedaten in das hedonische Modell eingebunden werden, um eine höchstmögliche Genauigkeit zu erreichen?
- Welche Variablen aus diesem umfangreichen Lagedatenpool liefern valide Ergebnisse im hedonischen Preismodell?
- Welche Vorteile bringt die Verwendung von lagegenauen GIS-Variablen (Raster- und Polygondaten) in hedonischen Preismodellen?

1.2.2 Abgrenzung

In der allgemein bekannten wissenschaftlichen Literatur¹ wurden bisher in hedonischen Preismodellen „nur“ jene Lagevariablen verwendet, die als Basis eine räumlich-statistische Bezugseinheit (Politische Grenzen, Zählbezirke, etc.) aufweisen, oder für eine spezielle Fragestellung in einer Arbeit manuell erfasst wurden. Damit die durch die fehlenden Mikrolagefaktoren bedingten Unsicherheiten auch Berücksichtigung bei der Schätzung in einem Preismodell finden können, müssen diese flächendeckend und mit homogener Qualität generiert werden.

Der Fokus dieser Arbeit liegt demnach nicht vordringlich in der Konzeption eines ökonomischen Modells, sondern in der automatisierten GIS-technischen Umsetzung der in Kapitel 3 beschriebenen Erzeugung von Mikrolagevariablen für Österreich. Um die Relevanz der erzeugten Mikrolagevariablen belegen zu können, wird danach anhand eines vereinfachten hedonischen Preismodells die wertbeeinflussende Wirkung der im GIS modellierten Variablen überprüft.

1.3 Ablauf der Arbeit

Nach der Einleitung (Kapitel 1) und Definition der Forschungsfrage wird im Kapitel 2 neben den theoretischen Grundlagen auch das grundlegende Forschungsdesign erläutert.

Im 3. Kapitel wird die mittels Python automatisierte GIS-gestützte Generierung der

¹ Siehe dazu Kapitel 2.1.4.

neuen Mikrolagevariablen (z.B.: Lärm, Erreichbarkeit, etc.) beschrieben.

Um alle Variablen aus diesem Lagedatenpool analysieren zu können, werden mithilfe von semiparametrischen Regressionsmodellen (Kapitel 4) alle im Rahmen dieser Arbeit erzeugten Lagevariablen auf ihre Immobilienpreisrelevanz überprüft.

Im abschließenden Kapitel 5 werden die Ergebnisse aus den zwei hedonischen Immobilienpreismodellen (EFH und ETW) visuell dargestellt und die durch Mikrolagevariablen hervorgerufenen Modellverbesserungen diskutiert.

Kapitel 2

Theoretische Grundlagen

Wie bereits im einführenden Kapitel erläutert, ist die Problemstellung dieser Arbeit die Analyse von im GIS modellierbaren Mikrolagevariablen bezüglich ihrer Auswirkungen auf den Immobilienwert. Grundsätzlich ist es für die Erreichung des Analyseziels in dieser Arbeit notwendig, mittels GIS die notwendigen Geodaten zusammenzuführen und aufzubereiten, um diese dann in weiterer Folge mit statistischen Methoden auf ihre Relevanz hin zu überprüfen.

In diesem Kapitel wird erläutert, mit welchen Methoden die Immobilienpreisanalyse erfolgen wird und wie grundsätzlich die als Basis benötigten Mikrolagevariablen mit GIS modelliert werden können.

2.1 Hedonische Preismodelle

In der Immobilienbewertung gibt es in Österreich neben den klassischen Bewertungsverfahren wie Vergleichs-, Ertrags- und Sachwertverfahren gemäß §§ 4 - 6 LBG auch noch die Möglichkeit, Wertermittlungsverfahren nach dem Stand der Wissenschaft (LBG § 7), wie es die auf der Regressionsanalyse basierenden hedonischen Preismodelle sind, zu verwenden. Die Relevanz dieser Methode hat sich in letzter Zeit deutlich erhöht und es ist auch in Zukunft absehbar, dass sich diese Bewertungsmethode weiter in der Bewertungsbranche durchsetzen wird.

„Hedonische Preise gelten als die virtuellen Preise, die ein Käufer für die einzelnen

Eigenschaften eines Gutes, hier der jeweiligen Immobilie, zahlt. Wenn die hedonischen Preise der Gutseigenschaften als Durchschnittspreise auf dem Markt betrachtet (und entsprechend ermittelt) werden, dann ergibt sich im Besonderen der virtuelle Preis für eine bestimmte Immobilie aus der Summe der hedonischen Preise ihrer Eigenschaften. Als Eigenschaften gelten hier einerseits die Objektattribute der einzelnen Immobilien (z.B. Zustand und Größe einer Wohnung) und andererseits die externen Lageeigenschaften des Standortes der Immobilie.²

Zusammenfassend kann man also festhalten, dass die Hauptkomponenten eines hedonischen Preismodells wie folgt aufgeteilt werden können:

- Objektkriterien
 - Flächen: Grundstück (Garten), Wohnfläche, Balkon, Loggia, Terrasse, Garage, etc.
 - Ausstattung: Bad, Heizung, Garage, Außenanlagen, überdurchschnittliche Eigenschaften, etc.
 - Zustand: Allgemeiner Wartungszustand, Sanierungen, etc.
 - Baujahr

- Lagekriterien
 - Aggregierte Zählsprengeldaten:
 - * Marktindikatoren: Kaufpreisindizes (Grundbuchdaten, Ametanedaten), Marktaktivität
 - Koordinatenverortete Daten:
 - * Demographische Daten: Bevölkerungsdichte, Bildungsniveau (z.B.: Akademikeranteil), Ausländeranteil, Beschäftigte, Arbeitslose, etc.
 - * Erreichbarkeit: Motorisierter Individualverkehr (MIV), Öffentlicher Personennahverkehr (ÖPNV), Entfernung Flughafen, etc.
 - * Klassische Mikrolagevariablen: Lärm, Wohnumfeld (Parkanlagen, Gewässer, etc.), Hangneigung, Geländeausrichtung, etc.

² BÖKEMANN und FEILMAYR 2005, S. 1.

Auf Basis der oben beispielhaft angeführten Immobilieneigenschaften werden mithilfe der Kaufpreisinformationen der Objekte sogenannte implizite Preise pro Eigenschaft ermittelt, aus welchen in weiterer Folge der „explizite“ Preis einer Immobilie abgeleitet werden kann.

Um die Funktionsweise des verwendeten Analysemodells kurz zu erläutern, wird in den nachstehenden Kapiteln 2.1.1 bis 2.1.3 ausgehend von einem allgemeinen Regressionsmodell erklärt, welche Modellart in dieser Arbeit zur Überprüfung der Wertrelevanz der unterschiedlichen Mikrolagevariablen (Lagekriterien) zur Anwendung kommt.

An dieser Stelle ist noch anzumerken, dass abweichend von der Darstellung in der obigen Liste derzeit alle Lagekriterien nur auf Zählsprengelzebene vorliegen. Im Zuge dieser Arbeit werden die koordinatengenauen GIS-Daten (teilweise Raster) für die Analyse der Immobilienpreisdaten herangezogen, um kleinräumige Lageeffekte miteinzubeziehen. Bisher haben die hochaggregierten Daten von z.B. Ausländeranteil oder Hangneigung in hedonischen Preismodellen noch zu keinen validen Ergebnissen geführt, da diesen die räumliche Granularität fehlt, welche in dieser Arbeit aber nun vorhanden sein wird.

2.1.1 Regressionsmodelle

Die Regressionsanalyse kann als Standardverfahren für statistische Fragestellungen angesehen werden. Es wird hierbei die Beziehung zwischen einer quantitativen abhängigen Variable (Zielvariable) und einer oder mehreren unabhängigen Variablen (erklärende Variablen) geschätzt. Das „Hauptziel der Regressionsanalyse besteht somit darin, den Einfluss der erklärenden Variablen auf den Mittelwert der Zielgröße zu untersuchen.“³ Im aktuellen Anwendungsfall werden demnach für die unabhängigen Variablen die bereits oben beschriebenen Objekt- und Lagekriterien verwendet und als abhängige Variable der Kaufpreis der Immobilien eingesetzt. Im Allgemeinen wird die Funktion (f) für mehrere unabhängige Variablen (x) anhand der abhängigen Variable (Y), wie in For-

³ FAHRMEIR, KNEIB und LANG 2007, S. 19.

mel 2.1 dargestellt, gesucht. Der Modellfehler, auch Residuum (ε) genannt, beschreibt die Abweichung des geschätzten Ergebnisses vom Ursprungswert.

$$Y = f(x_1, x_1, \dots, x_n) + \varepsilon \quad (2.1)$$

2.1.2 Lineares Modell

Das lineare Regressionsmodell ist ein spezieller Fall der Regressionsanalyse, welcher „insbesondere dann sinnvoll einsetzbar ist, wenn die Zielvariable Y stetig und wenn möglich approximativ normalverteilt ist.“⁴ In Formel 2.2 ist die lineare Regression allgemein formuliert dargestellt, wobei X die Designmatrix (Matrix aller im Modell zu parametrisierenden Variablen) ist, β für den Parametervektor steht, n die Anzahl der Beobachtungen ist und $p - 1$ die Anzahl der Variablen ergibt.

$$\begin{aligned} X &\in \mathbb{R}^{n \times p} \\ \beta &\in \mathbb{R}^p \\ Y &= X\beta + \varepsilon \end{aligned} \quad (2.2)$$

Bei Verwendung von mehreren Regressoren (unabhängigen Variablen), wie es in dieser Arbeit der Fall ist, spricht man von einem multiplen linearen Modell.

2.1.3 Generalisierte Additive Modelle (GAM)

Ein lineares Regressionsmodell führt aufgrund der Nichtlinearität von Variableneffekten in Immobilienpreisdaten zu einem nicht optimalen Modellergebnis. Damit auch nicht-lineare Variablen wie beispielsweise die Grundstücksfläche (der Preis pro m^2 nimmt mit zunehmender Grundstücksfläche ab) oder das Gebäudealter (Nachkriegsbauten sind weniger wert als Gründerzeithäuser) bestmöglich im Modell abgebildet werden können, wird in dieser Arbeit GAM verwendet. Dieses semiparametrische Regressionsmodell verwendet für die Bildung der nichtlinearen Terme kubische Splines, welche

⁴ FAHRMEIR, KNEIB und LANG 2007, S. 20.

aus miteinander verbundenen Abschnitten von kubischen Polynomen bestehen.⁵ Wie in Formel 2.3 dargestellt, wird bei einem GAM im Vergleich zu einem linearen Modell nun eine zusätzliche Funktion f_q für jeden nichtlinearen Parameter z_q in der Gleichung ergänzt.

$$Y \sim \text{exponentiale Familie (auch Normalverteilung)}$$
$$g(Y) = X\beta + f_1(z_1) + f_2(z_2) + \dots + f_q(z_q) + \varepsilon \quad (2.3)$$

Im aktuellen Anwendungsfall wird aus verteilungstechnischen Gründen auf die logarithmierten Kaufpreise regressiert, und deshalb als Linkfunktion der Logarithmus verwendet.

$$g(x) = \log(x)$$
$$g() = \text{Linkfunktion} \quad (2.4)$$

2.1.3.1 Praktische Umsetzung

Die Umsetzung des Modells erfolgt in der syntaxbasierten Statistiksoftware R. Es wird hierfür das „mgcv“ Package verwendet, welches es ermöglicht, die unterschiedlichen Formen von GAM zu berechnen. Detailliertere Informationen bezüglich der verwendeten Methodik von GAM sind unter WOOD 2006a, S. 121ff zu finden.⁶

2.1.4 Literaturübersicht

Nachdem im vorangegangenen Teil dieses Kapitels die technische Funktionsweise des zur Anwendung kommenden Modells erläutert wurde, wird nachstehend ein kompakter Überblick der grundlegenden Literatur zu hedonischen Modellen angeführt.

Die ersten bekannten hedonischen Modelle (COURT 1939, GRILICHES 1961) wurden in der Autoindustrie verwendet, um den Preis der Autos durch ihre Eigenschaften wie beispielsweise Leistung, Farbe etc. zu erklären.⁷ Nach den Veröffentlichungen der heute

⁵ Vgl.: WOOD 2006a, S. 122f.

⁶ Siehe dazu auch WOOD 2006c und WOOD 2006b.

⁷ Vgl.: PALMQUIST 1999, S.765.

relevanten hedonischen Literatur (ROSEN 1974, BARTIK und V. SMITH 1987, PALMQUIST 1991, FREEMAN 2003, etc.) werden neben der Bewertung von Umweltproblemen auch weitere neue Fragestellungen mit dieser Methode in der Forschung beantwortet. Neben der klassischen Verwendung von einem hedonischen Modell für Immobilienpreise (Weitere Informationen dazu sind z.B. in BRUNAUER 2010 zu finden.) gibt es demnach eine Vielzahl von möglichen Anwendungsgebieten wie z.B. Luftverschmutzung, Gefährdungspotential durch Deponien, Autobahn- oder Flughafenlärm, Erdbebenrisiko, Erosion, etc. für diese Methode, sofern bei diesen der Preis eines Gutes oder eines Produktionsfaktors durch die Umwelt beeinflusst wird.⁸

2.1.4.1 Abgrenzung zu aktuellen Modellen

Um die technischen Neuerungen dieser Arbeit mit den derzeit geläufigen Publikationen und Studien vergleichen zu können, wird in der nachfolgenden Auflistung unter Punkt 1 der Stand der Technik in Bezug auf Mikrolagevariablen bei hedonischen Preismodellen dargestellt. Unter Punkt 2 werden die aus dieser Arbeit hervorgehenden Neuerungen in Bezug auf die verwendeten Mikrolagevariablen aufgelistet.

1. Bisher sind in hedonischen Preismodellen:

- Lagevariablen an räumlich-statistische Bezugseinheiten gebunden,
- keine kleinräumigen Schätzungen in großen Regionen möglich,
- Mikrolagebedingte Schätzunsicherheiten unberücksichtigt geblieben,
- keine kleinräumigen Wertbestimmungen automatisiert möglich.

2. Neu in dieser Arbeit ist, dass für(mit) ein(em) hedonisches Preismodell:

- Mikrolagevariablen durch GIS-technische Automatisierung standardisiert generiert werden,
- eine flächendeckende (österreichweite) Mikrolagevariablenbank erzeugt wird, welche über eine sehr hohe Lagegenauigkeit (Raster: 50m Auflösung,

⁸ Vgl.: PALMQUIST 1999, S.772-773.

Polygon: koordinatengenau) verfügt,

- eine Vielzahl von neuen Mikrolagefaktoren wie z.B. Lärm global auf ihre Wertrelevanz überprüft werden können,
- eine technische Schnittstelle zu GIS-Daten geschaffen wurde, welche es ermöglicht, auch ohne räumlich-statistische Bezugseinheiten (Grenzen) Mikrolagedaten in einem Preismodell zu verwenden,
- Mikrolagevariablen auch bei der Visualisierung der geschätzten Modellergebnisse Verwendung finden und somit die Ergebnisse auch lagegenau abgebildet werden können.

2.2 GIS Grundlagen

Bei einem Geographischem Informationssystem (GIS, oder auch Geoinformationssystem genannt) handelt es sich im Wesentlichen um ein elektronisches Informationssystem, welches die Verknüpfung von räumlichen Geometrieobjekten (Punkt-, Linien-, Polygon- und Rasterdaten) mit einer dazugehörigen Datenbank ermöglicht.⁹

In dieser Arbeit werden alle GIS-technischen Berechnungen und Analyseschritte in ArcMAP von ESRI durchgeführt. Mit dieser Software ist es neben den klassischen GIS-Anwendungen – wie etwa Visualisierung von Geodaten – möglich, auch hochauflösende Rasteranalysen und sonstige polygonbasierende „Georechenoperationen“ durchzuführen. Für Rasteranalysen werden die in der Softwareerweiterung „Spatial Analyst“ enthaltenen Werkzeuge verwendet.

2.2.1 Datenmodelle

Grundsätzlich wird im GIS zwischen raster- und vektorbasierten Daten unterschieden.¹⁰ Der Datentyp wird durch die Eigenschaften der vorhandenen Daten bzw. vom Anspruch an die Lagegenauigkeit bestimmt. Möchte man eine möglichst lagegenaue

⁹ Weitere Informationen sind unter WEBERNDORFER 2009, S. 11ff zu finden.

¹⁰ Siehe dazu: WEBERNDORFER 2009, S. 14ff.

Verortung von Daten erreichen, werden Vektordaten verwendet, um beispielsweise den Verlauf einer Straße oder den Standort eines Gebäudes mit den dazugehörigen Eigenschaften darzustellen. Rasterdaten werden üblicherweise für kontinuierliche Daten wie z.B. bei einem Höhenmodell verwendet. Für räumliche Analysen wie Entfernungs-, Dichte- und Statistikberechnungen (Lokal, Zonal, Fokal und Global)¹¹ oder auch für logische Abfragen, Interpolationen, etc. wird bei Rasterdaten die sogenannte Map Algebra herangezogen.

2.2.2 Map Algebra

„Bei Map Algebra handelt es sich um eine einfache und leistungsstarke Algebra, mit der alle Spatial Analyst (SA)-Werkzeuge, -Operatoren und -Funktionen für geographische Analysen ausgeführt werden können. Die Stärke von Map Algebra liegt in der Fähigkeit, mehrere Werkzeuge in einer einzigen Anweisung zu kombinieren, um komplexe Interaktionen darzustellen. Da Map Algebra in Python integriert wurde, stehen alle Funktionen von Python und ArcPy sowie deren Erweiterungen (Module, Klassen, Funktionen und Eigenschaften) bei der Programmierung zur Verfügung.“¹²

2.2.3 Automatisierung von GIS-Prozessen

Um im ESRI ArcGis wiederkehrende Prozesse oder sehr große Analysen zusammenfügen bzw. mehrmals ausführen zu können, gibt es seit der Version 9.0 den ModelBuilder, welcher es ermöglicht, Abläufe mit Geoverarbeitungswerkzeugen in Prozessketten zu verbinden.

Um die in dieser Arbeit notwendigen, umfangreichen Geoverarbeitungsoperationen realisieren zu können, wird die ebenfalls mit Version 9.0 eingeführte Pythonschnittstelle (mit Version 10.0 wurde die Schnittstelle deutlich erweitert und nennt sich nun ArcPy) des ArcGis verwendet, welche es ermöglicht, auch sehr umfangreiche Geooperationen durchzuführen.

¹¹ Eine genaue Definition ist unter WEBERNDORFER 2009, S. 17f zu finden.

¹² ESRI 2012, Kapitel: Was ist Map Algebra?

2.2.3.1 Programmiersprache Python

Python ist eine sogenannte höhere Programmiersprache, welche über eine freie Lizenz verfügt. Da zusammenhängende Codeblöcke (Funktionen, If-Abfragen, Iterationen, etc.) durch Tabulator-Einrückung anstelle von Klammern wie in den meisten anderen Programmiersprachen gekennzeichnet werden, ist die Syntax sehr gut lesbar und immer klar strukturiert und deshalb auch für Programmierneinsteiger gut geeignet. Es besteht neben der klassischen Syntaxprogrammierung auch die Möglichkeit, objektorientierte und funktionale Codes zu erzeugen. Die sehr umfangreichen Standard-Bibliotheken und diverse Module von Drittanbietern stellen für eine Vielzahl von Aufgaben vorgefertigte Lösungen bereit. Weitere Informationen zur Programmiersprache sowie die offizielle Bibliotheksreferenz ist unter PYTHON 2013 zu finden.

Kapitel 3

Generierung der Mikrolagevariablen

In diesem Kapitel wird die Generierung von unterschiedlichen Mikrolagevariablen, welche in weiterer Folge für das hedonische Preismodell verwendet werden, beschrieben. Als Basis für alle weiteren GIS-Modellierungen wird ein Lagemodell (Kapitel 3.1) erzeugt, welches als Grundlage für das Lärm- (Kapitel 3.3) bzw. Erreichbarkeitsmodell (Kapitel 3.4) dient. Damit die Mikrolagefaktoren mit der höchstmöglichen Genauigkeit im GIS abgebildet werden können, wird neben vektorbasierten Daten (Lagemodell) bei allen kontinuierlichen Datenstrukturen (z.B.: Lärm) auf die Verwendung von Rasterdaten gesetzt. Die Auflösung der unterschiedlichen Rasterdaten variiert je nach Datenquelle und in Abhängigkeit von EDV-technischen Performancerestriktionen zwischen ca. 30m beim Höhenmodell und 250m beim Bevölkerungsraster der Statistik Austria. In Kapitel 3.5 werden jene Variablen beschrieben, die nicht extra für diese Arbeit erstellt wurden, sondern direkt von externen Datenanbietern übernommen werden.

3.1 Landnutzungsmodell

Um im ganzen Untersuchungsraum (Österreich) die verschiedenen geographischen Lageeigenschaften eines Standortes bestimmen zu können, ist eine GIS-Datenschicht von Nöten, welche es ermöglicht, alle Nachbarschaftskriterien in homogener Form zu beschreiben. Zum derzeitigen Zeitpunkt sind grundsätzlich zwei Datenschichten, die für

eine solche Fragestellung geeignet sind, bekannt:

1. Die Nutzungsdaten des Landinformationssystems Österreich, welche beim Fonds zur Förderung der wissenschaftlichen Forschung in den Jahren 1994-1999 erstellt wurden.¹³ Diese sind wegen des Alters und aus Kostengründen auszuschneiden.
2. Das aktuellere Projekt Land Information System Austria (LISA) verfügt über eine sehr detaillierte Erhebungsstruktur, wobei zwischen Landbedeckung und Landnutzung unterschieden wird.¹⁴ Diese Daten sind jedoch derzeit nur für einige Testregionen verfügbar und damit für die Fragestellung in dieser Arbeit ebenfalls nicht geeignet.

Aus diesem Grund wird eine österreichweite segmentierte Nutzungsdatenschicht als neue Datengrundlage erstellt.

Um mit diesem Landnutzungsdatensatz die höchstmögliche Genauigkeit zu erreichen, werden verschiedene Datenquellen (Tabelle: 3.1) bei der Erzeugung herangezogen.

Datenquelle	Inhalt	Kürzel ¹⁵
Teleatlas	Straßenkarten inkl. Landnutzungsinformationen	tele
OpenStreetMap (OSM)	Landnutzungsinformationen	osm
CORINE Land Cover	Landbedeckungsinformationen	corine
Adressraster	Anzahl der vorhandenen Adressen je Rasterzelle	bebaut

Tabelle 3.1: Datenquellen

Da jeder dieser GIS-Datensätze über eine Vielzahl von Merkmalsausprägungen verfügt, wurden diese, wie in Tabelle 3.2 dargestellt, zusammengeführt, um eine einheitliche Datenstruktur zu schaffen. Anhand von gemeinsamen Eigenschaften wurden Obergruppen gebildet, welche thematisch zusammenpassende Datensätze beinhalten. Anhand der Spalte Datenquelle ist ersichtlich, woher die Daten stammen und wie diese aufbereitet wurden. Die in der Spalte Datenquelle verwendeten Namen bestehen zu Beginn immer aus dem Namens Kürzel der Datenquelle (Tabelle 3.1) und an zweiter Stelle (falls

¹³ Siehe für weitere Infos: SEGER 1999.

¹⁴ Siehe für weitere Infos: LISA 2012.

¹⁵ Siehe Datenquelle in Tabelle 3.2.

vorhanden) wird das Inputfile der Quelle beschrieben. Beispiel: „tele_snt“ stammt aus der Datenquelle Teleatlas und der Datei „snt“, welche die Straßengraphen beinhaltet. Im Kapitel 3.1.1 wird neben der weiteren Aufschlüsselung dieser Bezeichnungen die genaue Aufbereitungsmethode erläutert. Die Spalte „Prio“ beschreibt den Vor- oder Nachrang der jeweiligen Datenschichten, wobei X als „nicht zu verwenden“ interpretiert wird und NV „nicht vorhanden“ bedeutet. Die nicht verwendeten Daten fließen größtenteils in anderer Form später ein und werden deshalb in der Tabelle 3.2 belassen. Die Farbeinteilungen in der Spalte Nutzungskey (NK) sind die Legende für alle folgenden Grafiken zu diesem Thema. Es wurde bewusst darauf verzichtet, alle Einzelklassen mit einer eigenen Farbe zu versehen, da die Karte in vereinfachter Form besser lesbar ist.

Nutzungskey	Obergruppe / Inhalt	Datenquelle	Prio
107, 108 Einkaufszone			
10701	Fußgängerzonen	tele_ped	X
10801	Einkaufszentrum	tele_lus	7
109 Plätze			
10901	Plätze	tele_adr	6
10902	Zentrumsplätze	tele_snt_lage	12
10903	Schlossplatz	tele_snt_lage	12
10904	Plätze	tele_snt_lage	12
10905	Kirchplätze	tele_snt_lage	12
110 Stadtgebiet (dichte Bebauung)			
11001	Universitäts- oder Fachhochschulgelände	tele_lus	21
11002	Krankenhaus	tele_lus	21
11003	Gebäude (Parkplätze und -häuser)	tele_lus	X
11004	durchgängig städtische Prägung	corine	25
11005	Rehabilitationszentrum	tele_snt_lage	X
111 Stadtgebiet (mitteldichte Bebauung)			
11101	nicht durchgängig städtische Prägung	corine	26
112 Stadtgebiet (lockere Bebauung)			
11201	bebaute Flächen	tele_bup	X
11202	Kleingarten	osm_landuse	19
11203	Adressraster	bebaut	X
11204	Wohnpark	tele_snt_lage	13
11205	Kleingarten	tele_snt_lage	13
113 Industrie und Gewerbe			
11301	Firmengelände	tele_lus	10
11302	Industriegebiet	tele_lus	10
11303	Parkplatz	tele_lus	10
11304	Industriegebiet	osm_landuse	18
11305	Industrie- und Gewerbeflächen	corine	28
11306	Gewerbegebiet	tele_snt_lage	16
11307	Leichtindustrie	tele_snt_lage	16
11308	Messegebiete	tele_snt_lage	16

KAPITEL 3. GENERIERUNG DER MIKROLAGEVARIABLEN

Nutzungskey	Obergruppe / Inhalt	Datenquelle	Prio
 11309	Schwerindustrie	tele_snt_lage	16
114 Verkehrsflächen (Straße)			
11401	Plätze	tele_snt	X
 11402	Autobahn	tele_snt	2_1
 11403	Autobahn unbefestigt	tele_snt	NV
 11404	Fernstraße	tele_snt	2_2
 11405	Fernstraße (unbefestigt)	tele_snt	2_4
 11406	Hauptstraße	tele_snt	2_5
 11407	Hauptstraße (unbefestigt)	tele_snt	2_7
 11421	Nebenstraße Fußgängerzone	tele_snt	2_11
 11408	Nebenstraße	tele_snt	2_8
 11409	Nebenstraße2	tele_snt	2_10
 11410	Nebenstraße unbefestigt	tele_snt	2_12
 11411	Fernstraße (Einbahn)	tele_owy	2_3
 11412	Hauptstraße (Einbahn)	tele_owy	2_6
 11413	Nebenstraße (Einbahn)	tele_owy	2_9
114 Verkehrsflächen (Bahn)			
 11414	Lokalbahn / Sonstige	tele_rwy_neu	1_4
 11415	Hochleistungsstrecke	tele_rwy_neu	1_1
 11416	Fernstrecke	tele_rwy_neu	1_2
 11417	Nebenbahn	tele_rwy_neu	1_3
 11418	Plattform (gr. Bahnhaltstellen)	tele_rwy_neu	1_6
 11419	Straßenbahn	tele_rwy_neu	1_5
11420	Ubahn	tele_rwy_neu	X
 11422	Straßen und Eisenbahn (funktionell zugeordnete Flächen)	corine	30
115 Hafengebiete			
 11501	industrielles Hafengebiet	tele_lus	11
 11502	Hafengebiete	corine	29
116 Flughafen			
 11601	Flughafengelände (Rollbahn)	tele_lus	3
 11602	Flughafengelände	tele_lus	3
 11603	Flughafengelände	corine	24
117 Abbauflächen, Deponien und Baustellen			
 11701	Abbauflächen	corine	33
 11702	Deponien, Abraumhalden	corine	33
 11703	Baustellen	corine	33
118 Städtische Grünflächen, Parks und Gärten			
 11801	Park/Garten	tele_lus	9
 11802	Friedhof	tele_lus	9
 11803	Friedhof	osm_landuse	17
 11804	Park	osm_landuse	17
 11805	städtische Grünflächen	corine	31
 11806	Park	tele_snt_lage	15
 11807	unbefestigte Nebenstraßen	snt_neben	20
119 Sport- und Freizeitanlagen			
 11901	Golfplatz	tele_lus	8
 11902	Stadiongelände	tele_lus	4
 11903	Sport- und Freizeitanlagen	corine	32
 11904	Sport- und Freizeitanlagen	tele_snt_lage	14

Nutzungskey	Obergruppe / Inhalt	Datenquelle	Prio
11905	Sportplatz	tele_snt_lage	14
11906	Mobilheimplatz	tele_snt_lage	14
11907	Golfplatz	tele_snt_lage	14
120 Landwirtschaftlich intensiv genutzte Flächen			
12001	nicht bewässertes Ackerland	corine	34
12002	Weinbauflächen	corine	34
12003	komplexe Parzellenstruktur	corine	34
12004	landwirtschaftlich genutztes Land mit Flächen natürlicher Vegetation von signifikanter Größe	corine	34
121 Wiesen und Weiden			
12101	Wiesen und Weiden	corine	35
122 Wald			
12201	Wald	tele_lco	40
12202	Laubwälder	corine	36
12203	Nadelwälder	corine	36
12204	Mischwälder	corine	36
123 Kraut- und Strauchvegetation			
12301	Moor/Heide	tele_lco	22
12302	natürliches Grünland	corine	37
12303	Heiden und Moorheiden	corine	37
12304	Wald/Strauch Übergangsstadien	corine	37
124 Geringe Vegetation			
12401	Geröll	tele_lco	23
12402	Felsflächen ohne Vegetation	corine	38
12403	Flächen mit spärlicher Vegetation	corine	38
12404	Gletscher/Dauerschneegebiet	corine	38
125 Feuchtgebiete			
12501	Sümpfe	corine	39
12502	Torfmoore	corine	39
12503	Salzwiesen	corine	NV
126 Wasser			
12601	Gewässser	tele_wea	5_1
12602	Gewässser	osm_wat	5_2
12603	Gewässerläufe	corine	27
12604	Wasserflächen	corine	27

Tabelle: 3.2: Nutzungsdatenschicht

3.1.1 Automatisierte Erstellung der Landnutzungsoberfläche

Die Zusammenführung und Harmonisierung der bereits oben beschriebenen Datensätze erfolgt mit einem mehrstufigen Verfahren. Nachstehend werden die wichtigsten Funktionen gemäß ihrem Ablauf im Pythonskript benannt und ihre Arbeitsweise erklärt.

3.1.1.1 Datenimport in GIS-Datenbank

Beim Laden der Daten sind folgende Aufgaben zu bewältigen:

1. Laden der unterschiedlichen Geo-Daten und Zerlegung dieser in einzelne Datenschichten je Nutzungskategorie.
2. Umprojizieren der Geodaten auf eine einheitliche Projektionseinstellung.
3. Vergabe eines eindeutigen „Nutzungskeys“ (NK) für jede Datenschicht, um für alle weiteren Schritte einen automatisierten Zugriff zu gewährleisten.
4. Systematisches Ablegen der generierten Features in einer File-Geodatabase (FGDB)¹⁶, um für spätere Berechnungen ein performantes und einheitliches Fileformat verwenden zu können.

Da dieser Importprozess aufgrund der großen Datenmenge kaum manuell zu bewältigen wäre, wurde dafür eigens ein Importprogramm entwickelt. Der dynamische Aufbau des Programms ermöglicht, dass für jede Nutzungskategorie separate Eingabespezifikationen gesetzt werden können und somit jede Inputdatei damit bearbeitet werden kann. Die wichtigsten Aufrufparameter sind die programmtechnischen Übersetzungen der geladenen Datenstruktur je Importdatei zu den in der obiger Tabelle 3.2 definierten Nutzungskeys. Die Schwierigkeit bei dieser Übersetzungslogik war, dass alle Datenquellen unterschiedlich aufgebaut sind und dementsprechend jede der vorliegenden Datenstrukturen vom Programm automatisiert erkannt werden muss. Daraus folgt, dass die Eingabeparameter bei jeder Datenschicht in Abhängigkeit von den Rohdaten unterschiedlich gesetzt werden müssen. Grundsätzlich kann man davon ausgehen, dass gewisse Inputklassifikationen in eine standardisierte Outputklassifikation übersetzt werden.

¹⁶ Ist eine vom Betriebssystem unabhängige Dateistruktur, welche keine Größenbeschränkungen hat.

3.1.1.2 Erzeugung von Nutzungsflächen aus dem Straßengraphen

Da in den vorhandenen Datenquellen, trotz der mittlerweile recht hohen Informationsdichte der GIS-Daten nicht alle vorhandenen Nutzungen wie z.B. Plätze, Parks oder Gewerbegebiete abgebildet sind, werden zusätzliche Informationen aus dem Straßengraphen ausgelesen, um weitere Flächenausprägungen abzuleiten. Dazu wurden die folgenden Schritte für den kompletten Straßengraphen von Österreich im Programm abgearbeitet:

1. Eine Keywordliste (z.B.: Park, Platz, EKZ, Gewerbe, etc.) wird mit den Straßennamen verglichen. Die Liniensegmente, in denen die Straßennamen Keywords enthalten, bleiben bestehen, alle anderen werden gelöscht.
2. Es laufen Interaktionen durch die Ausschlusslisten (z.B.: Straße, Siedlung, etc.), welche weitere Liniensegmente automatisiert entfernen.
3. Sondernamen die z.B. „am“ enthalten – wie etwa „Am Sportplatz“ – werden ebenfalls extra berücksichtigt, um keine Daten zu verlieren.
4. Danach folgen noch weitere manuelle Löschungen, die vom Skript selbst nicht automatisch abgefangen werden. Durch das Suchwort „Messe“ gelangt beispielsweise die Auswahl „In den Messerern“ in den Datensatz und muss somit manuell gelöscht werden.

Die Keywordliste sowie die Ausschlusslisten sind durch mehrmaliges Aufrufen des Programmes und durch die manuelle Prüfung soweit verbessert worden, dass im Datensatz keine fehlerhaften Flächen mehr vorhanden sind. Dadurch sind nun neue Linien-Features entstanden, welche mittels einer eigenen Übersetzungstabelle, wie bereits im Kapitel 3.1.1.1 beschrieben, neben den normalen Straßen ebenfalls automatisiert in der FGDB abgelegt werden. Diese neuen Datensätze sind in der Spalte Datenquelle (Tabelle 3.2) als „tele_snt_lage“ bezeichnet.

5. Nachdem nun der Ladevorgang abgeschlossen ist, werden die identifizierten Liniensegmente in Flächen umgewandelt. Diese Funktion ermöglicht es beispielsweise,

einen Platz, der an einer oder mehreren Seiten mit einer Straße begrenzt ist, als Fläche darzustellen.

Für die Erzeugung der Flächen wird zunächst für alle aneinandergrenzenden Linien, welche den gleichen Straßennamen haben, eine Pufferoberfläche mit einem Meter Radius errechnet und dann generalisiert, um etwaige Überlappungen auszuschließen. Basierend auf der Tabelle 3.2 (*NKT*) mit den Nutzungskeys wird mit diesen bereinigten Oberflächen die minimale Begrenzungsgeometrie¹⁷ jeweils für das Rechteck der kleinsten Fläche (*RE_{Min. Fläche}*) und das Rechteck mit der kleinsten Breite (*RE_{Min. Breite}*) berechnet. Aus diesen zwei Rechtecken wird dann die Schnittmenge (*SM_i*) gebildet und ein optionaler Puffer (*P_i*) hinzugefügt (Formel 3.1).

$$\begin{aligned}
 NK_Liste &= NKT[Datenquelle == „tele_snt_lage“] \\
 \text{Für } i \in \{NK_Liste\} \text{ gilt :} \\
 SM_i &= (RE_{Min. \text{ Fläche, } i} \cap RE_{Min. \text{ Breite, } i}) \\
 P_i &= f(SM_i, Breite_{SM_i}, Länge_{SM_i}) \\
 \text{Linie zu Fläche}_i &= SM_i + P_i
 \end{aligned} \tag{3.1}$$

Der Radius dieses optionalen Puffers (*P_i*) wird in Abhängigkeit vom Flächentyp, der Größe und der Längen- und Breitenverhältnisse der Rechtecke berechnet (Beispiel in Algorithmus 1).

In der folgenden Abbildung 3.1(a) ist ein Beispiel für die ermittelten Flächen dargestellt, wobei dies als „netto“ betrachtet werden kann, da hier bereits die im folgenden Kapitel 3.1.1.3 beschriebene Zusammenführung geschehen ist und somit nur mehr jene Flächen zu sehen sind, welche auch nach der Schnittlogik übrigbleiben. Demgegenüber steht in der Abbildung 3.1(b) der Ausschnitt mit allen Nutzungskeys und man kann somit erkennen, dass durch die Erzeugung der Flächeninformationen eine Vielzahl an neuen Informationen aus den Straßennamen extrahiert werden konnte. Was im innerstädtischen Bereich auffällt, ist, dass es sich bei den neu entstandenen Flächen hauptsächlich um Plätze, Park- und Grünanlagen handelt. Im ländlichen Gebiet hingegen treten ver-

¹⁷ ESRI 2012, Tool: Minimum Bounding Geometry (Data Management (DM)).

```

def puffer(breite, laenge):
    verhaeltnis = laenge * 0.2
    if verhaeltnis >= breite:
        if laenge <= 25:
            rueckgabe = (laenge * 0.75 - breite) / 2
        elif laenge > 25 and laenge <= 50:
            rueckgabe = (laenge * 0.5 - breite) / 2
        elif laenge > 50 and laenge <= 75:
            rueckgabe = (laenge * 0.35 - breite) / 2
        elif laenge > 75:
            rueckgabe = (laenge * 0.2 - breite) / 2
        return rueckgabe
    else:
        return 0

```

Algorithmus 1: Beispiel Pufferberechnung - 109 Plätze (Python)

mehrt neue Gewerbe- und Industriegebiete sowie Sport- und Freizeiteinrichtungen auf. In dieser Datenschicht sind viele violette Linien zu erkennen, welche die Fußgängerzonen kennzeichnen und ebenfalls in diesem Bearbeitungsschritt (aus Straßengraphen) erzeugt wurden.



(a) Hinzugefügte Flächen



(b) Alle Flächentypen

Abbildung 3.1: Flächeninformation aus Straßennamen (Beispiel: Wien, 1. Bezirk)

6. Im nächsten Schritt werden für alle vorhandenen Verkehrsflächen (Obergruppe (OG) 114) ebenfalls Pufferflächen berechnet, damit auch diese Liniensegmente als Flächen verfügbar sind. Jeder NK der Verkehrsflächen wird mit einer entsprechend angepassten Durchschnittsbreite berechnet.

3.1.1.3 Zusammenführung aller Datensätze

Der abschließende Schritt zur Erzeugung der Landnutzungsoberfläche ist sogleich auch der aufwendigste. Es werden nun alle einzelnen Nutzungskeys priorisiert und sodann gemäß dieser Reihenfolge voneinander ausgeschnitten, um am Ende eine eindeutige Zuordnung zu erreichen. Die Priorisierung wurde anhand logischer Überlegungen an- gestellt, wobei diese im GIS jeweils visuell überprüft wurden, um das bestmögliche Ergebnis zu gewährleisten. Die Spalte „Prio“ in der Tabelle 3.2 enthält die Ergebnisse dieser Reihung.

Die OG 114 (Verkehrsflächen) ist etwas differenzierter zu sehen (Beispiel: 2_9). Die erste Zahl ist die allgemeine Priorität, und die zweite Zahl wird für die interne Rei- henfolge verwendet. Wenn bei den Prioritäten mehrmals die gleichen Zahlen auftreten, liegt dies daran, dass diese Daten vom gleichen Rohdatensatz kommen und demen- sprechend nicht überlappen und somit gleichbehandelt werden können.

Wie in der Formel 3.2 dargestellt ist, werden bei der programmtechnischen Umsetzung in einem iterativen Prozess von jedem NK alle höher priorisierten Flächen (siehe Ta- belle 3.2 (*NKT*)) abgezogen und sodann zu einer lückenlosen Gesamtoberfläche (*GO*) zusammengesetzt (Beispiel: Von der Fläche mit Prio 1_3 werden alle Flächen mit Prio 1_1 und 1_2 abgezogen.).

$$\begin{aligned}
 & \text{Für } i \in \{\text{sort}(\text{NKT}[\text{Prio}])\} \text{ gilt :} \\
 & \text{Fläche}_0 = \emptyset \\
 & \text{Fläche}_i = \text{Fläche}_i \setminus \left(\bigcup_{j=0}^{i-1} \text{Fläche}_j \right) \\
 & \text{GO} = \bigcup_{i=0}^{\text{NKT}[\text{Prio}]} \text{Fläche}_i \qquad (3.2)
 \end{aligned}$$

Bei derart großen Polygonoperationen (Analyseregion Österreich) erreicht man im GIS bereits die Grenze des technisch Machbaren. Um dieses Berechnungsproblem zu bewäl- tigen, gibt es mehrere Möglichkeiten:

1. Die Daten werden in eine generalisierte Rasterdatenschicht umgewandelt. Damit

können auch sehr große Regionen relativ genau und performant berechnet werden, wobei allerdings ein von der Rastergröße abhängiger Genauigkeitsverlust in Kauf genommen werden muss.

2. Die Features werden im GIS generalisiert. Diese Operation ist ebenfalls sehr rechenintensiv, und es kann auch je nach Methode zu Datenverlusten kommen.
3. Die Daten werden in Teilregionen des Analysegebietes aufgeteilt. Die Berechnung wird jeweils für diese Teile durchgeführt, und am Ende werden die Daten wieder zusammengefügt. Diese Methode ist zwar in der Programmierung etwas aufwendiger, führt aber zu einem verlustfreien Ergebnis.

Da im aktuellen Anwendungsfall das Zurückgreifen auf einen Raster aus Genauigkeitsgründen ausscheidet, wurde die dritte Variante gewählt, wobei als Teilregionen die Bundesländer herangezogen wurden. Die Berechnung für ganz Österreich ist aus performancetechnischen Gründen nicht möglich.¹⁸

3.1.2 Landnutzungsdatenschicht

Als Ergebnis erhält man nun die in Abbildung 3.2(a) beispielhaft dargestellte Nutzungsdatschicht für ganz Österreich. Sieht man sich diese Abbildung 3.2 genauer an, kann man feststellen, dass die Nutzungsdatschicht – obwohl hier gemäß der Legende nicht alle Informationen angezeigt werden (siehe Tabelle 3.2) – weitaus mehr Informationen als Google-(3.2(c)) oder BingMaps(3.2(b)) enthält.

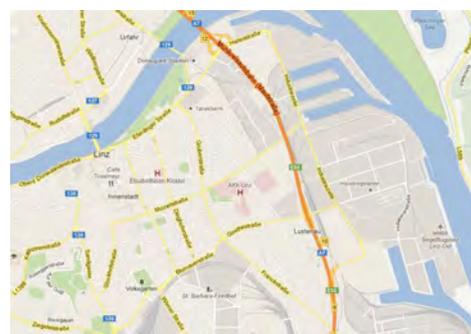
¹⁸ Anmerkung: Da das Testen von Python-Programmierungen für die GIS-Schnittstelle mit großen Datensätzen sehr zeitintensiv ist, wurde während der Programmierung nahezu immer mit kleinen Testdatensätzen vor der finalen Berechnung gearbeitet.



(a) Eigene Nutzungsdatenschicht



(b) Bing Maps



(c) Google Maps

Abbildung 3.2: Vergleiche Nutzungsdaten mit Bing und Google (Beispiel: Linz)

Es sind nun neben der rein optischen Darstellung der Umgebung auch nutzungsbezogene Informationen der Landoberfläche vorhanden. Wie man in der Abbildung 3.2(a) erkennen kann, liegt es zum Einen daran, dass bei der Erzeugung mehrere Datenquellen harmonisiert wurden und deshalb nun eine höhere Informationsdichte vorhanden ist und zum Anderen wurden durch die Verwendung der Informationen vom Straßengraphen (Siehe: Kapitel 3.1.1.2) ebenfalls viele neue Inhalte gewonnen.

3.1.2.1 Plausibilisierung der Datenschicht

Betrachtet man nun das Beispiel von Linz (Abbildung 3.2(a)), kann man in der östlichen Abbildungshälfte das Gelände der Voestalpine (Industrie), das Hafengebiet und die Leichtindustrie mit Gewerbebezonen erkennen. Auch die höhere Bebauungsdichte im Stadtzentrum (südwestlich) sowie die unterschiedlichen Nutzungen werden erkennbar. Im Süden befindet sich das Bahnhofsareal sowie eine größere Grünanlage (Friedhof) entlang der Westbahn.

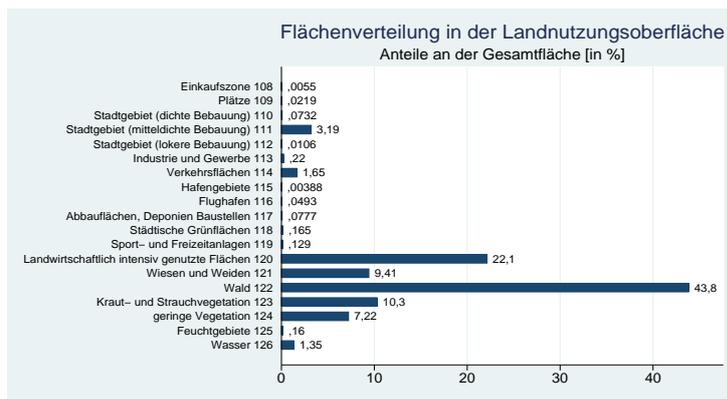
Neben dieser beispielhaft beschriebenen manuellen Kontrolle der Nutzungsdatenschicht wurde diese natürlich auch in größerem Ausmaß auf ihre Richtigkeit überprüft. Zunächst wurden Kontrollregionen in Österreich definiert – es sind darin alle Landeshauptstädte, sowie je Bundesland ein bis vier weitere zufällig gewählte Standorte enthalten. In diesen Kontrollregionen wurden dann die Luftbilder von Bing-¹⁹ und GoogleMaps,²⁰ sowie die Beschriftungen in diesen Karten mit der Nutzungsdatenschicht visuell verglichen und kontrolliert, ob es abweichende Informationen gibt. Das Ergebnis dieser manuellen Kontrolle ergibt, dass die Flächen generell immer korrekt zugewiesen sind, allerdings teilweise speziell in Neubaugebieten gewisse Informationen fehlen. In seltenen Fällen kann es vorkommen, dass die Flächenerzeugungen aus den Straßennameninformationen eine geringfügig abweichende Größe haben und somit auch das Nachbargrundstück mit dieser Flächeninformation versehen ist. Dieser kleine Fehler ist allerdings als vernachlässigbar einzustufen. Die gesamte Datenschicht ist von gleichmäßiger Datenqualität, wobei generell in Städten eine höhere Datendichte besteht. Neben dieser manuellen Qualitätskontrolle wurde auch anhand des Adressrasters (NK 11203) überprüft, ob alle Flächen die gemäß der OG als bebaut einzustufen sind, auch tatsächlich über postalische Adressen verfügen. Hierbei kam es nur zu minimalen Abweichungen. Deshalb wurde darauf verzichtet, diese Flächen aus dem bebauten Gebiet zu extrahieren.

Durch die Erstellung der Nutzungsdatenschicht ist es möglich, die Anteile der verschiedenen Flächen über ganz Österreich zu ermitteln. Wenn man diese Verteilung in

¹⁹ Siehe: BING MAPS 2012.

²⁰ Siehe: GOOGLE MAPS 2012.

Abbildung 3.3(a) betrachtet, kann man feststellen, dass es sich beim größten Teil der vorhandenen Flächen (94,4%) um unbesiedeltes Land (OG ≥ 120) handelt. Anzumerken ist hierzu noch, dass Gebäude in peripheren Streulagen und alle anderen Bauwerke (z.B.: landwirtschaftliche Objekte) außerhalb der geschlossenen Siedlungsstruktur in der OG 120 zu finden sind. Diese wurden bei der Erstellung der Landnutzungsoberfläche nicht separat erfasst. Es wurde zwar mit dem Adressraster (NK 12203) versucht, diese ländlichen locker bebauten Flächen zu ermitteln, jedoch waren die Ergebnisse zu ungenau, weshalb diese Flächen nicht extra ausgewiesen worden sind. Für die späteren Analysen haben diese eine untergeordnete Rolle, da sich die meisten Bewertungsobjekte in städtischen Agglomerationen befinden. Die Verkehrsflächen (OG 114-116) nehmen in Summe 1,7% ein und so bleiben für die besiedelten Flächen (OG ≤ 113) tatsächlich nur 3,5% (entspricht 2959 km²) der österreichischen Gesamtfläche übrig. Betrachtet man nun vergleichend die in Abbildung 3.3(a) angeführte Aufstellung mit den Daten aus der Grundstücksdatenbank (Abbildung 3.3(b)) kann man feststellen, dass obwohl bei diesem Kreisdiagramm etwas andere Klassen vorhanden sind, sich das Gesamtbild wie bei der erzeugten Landnutzungsoberfläche verhält und keine großen Abweichungen auftreten.



(a) Flächenanteile der Obergruppen



(b) Regionalinformationen der Grundstücksdatenbank (BEV)²²

Abbildung 3.3: Flächenanteile Österreich

Zusammenfassend lässt sich zur Erstellung der Landnutzungsoberfläche also festhalten,

²² Eigene Abbildung basierend auf UMWELTBUNDESAMT 2012.

dass die gewählte Methode zum gewünschten validen Ergebnis geführt hat und nun für alle weiteren Analyseschritte eine vereinheitlichte und für Österreich flächendeckende GIS-Datengrundlage vorhanden ist.

3.2 Höhenmodell

Um Informationen über die allgemeine Geländestruktur automatisch verarbeitbar zu machen, kommt der Datensatz Global Digital Elevation Map (GDEM) von Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) zur Anwendung. Diese Höhenschicht basiert dem Namen entsprechend auf Satellitendaten und wurde vom Ministry of Economy, Trade, and Industry (METI) in Japan und der National Aeronautics and Space Administration (NASA) in den Vereinigten Staaten im Zuge einer wissenschaftlichen Kooperation erstellt. Die Pixelauflösung des Höhenrasters beträgt eine Gradsekunde auf der World Geodetic System 1984 (WGS84) Projektion, was 30m entspricht. Auf der folgenden Homepage <http://asterweb.jpl.nasa.gov/gdem.asp> kann diese Datenschicht kostenlos heruntergeladen werden. Weitere Informationen zu diesem Datensatz sind unter METI 2012 und TACHIKAWA u. a. 2011 zu finden. Die Rasterkacheln können nur mit einer Größe von $1^\circ \text{ mal } 1^\circ$ ²³ heruntergeladen werden und müssen dementsprechend für die weitere Modellierung zuerst zu einem Gesamtraster zusammengeführt werden. Es ist zu beachten, dass sich die nebeneinanderliegenden Raster überlappen und dementsprechend die korrekte Append-Funktion im GIS zur Anwendung kommt. Abschließend wurde die zusammengeführte Rasterschicht des Höhenmodells für Österreich aus performancetechnischen Gründen noch mit den österreichischen Außengrenzen + 2,5 km abgeschnitten, damit nur in Österreich und der Grenzregion Werte vorhanden sind und alle restlichen Werte somit auf „NoData“²⁴ gesetzt werden. Beim Betrachten des Rasters sind noch einige Datenfehler aufgetreten, die unter Zuhilfenahme von diversen MAP-Algebra-Funktionen bereinigt wurden (beispielsweise sind auf dem Dachsteingletscher 38 Rasterzellen, die einen negativen Höhenwert von bis zu -281m aufweisen, zu finden). Mit dem fertig aufbereiteten Hö-

²³ In WGS84 Projektion.

²⁴ Entspricht dem NULL(Nullwert) in der Informatik, welcher besagt, dass kein Wert existiert.

henraster ist es nun möglich, unter Zuhilfenahme der Oberflächenwerkzeuge aus dem GIS, das vorhandene Terrain auf seine Ausprägungen hin zu analysieren. In Abbildung 3.4 ist der Höhenraster entsprechend der Höhenstufen visualisiert.

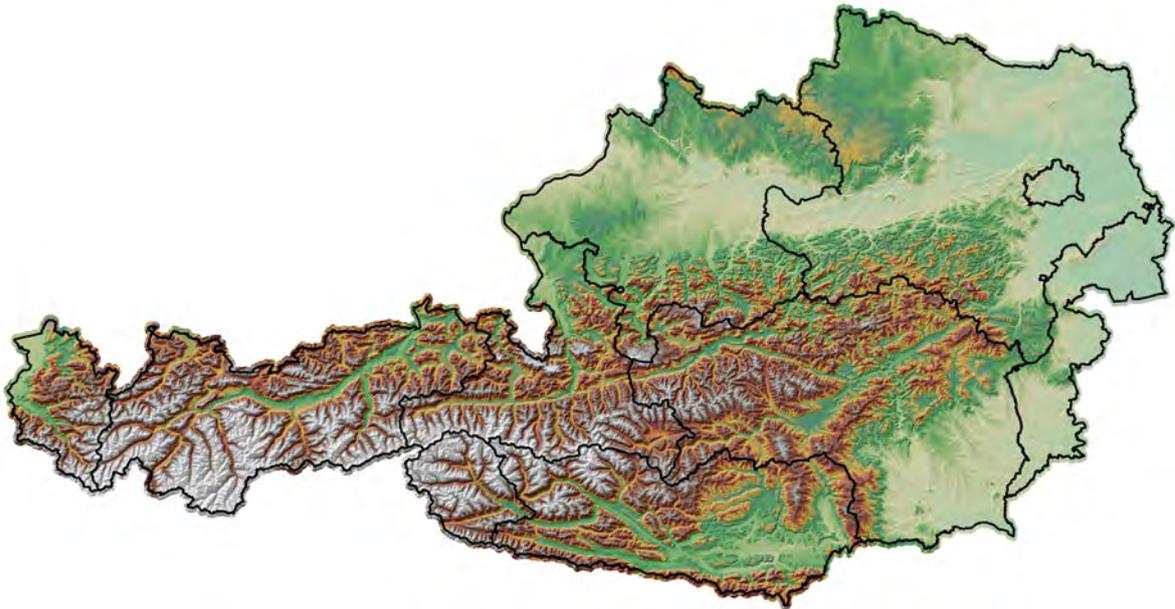


Abbildung 3.4: Höhenmodell Österreich²⁵

Bei Verwendung des ASTER Höhenrasters ist bei allen Analyseschritten darauf zu achten, dass zuerst alle Analysen mit dem Raster in der vorhandenen Projektion vorgenommen werden. Erst dann werden die Rasterschichten in die gewünschte Projektion übergeführt, weil das Umprojizieren von Rasterdaten nur mit Verlusten möglich ist. Würde man die Ermittlung z.B. der Neigung oder Krümmung nach dem Umprojizieren durchführen, hätte man durch die Projektionsänderung verzerrte Werte, da bei der Transformation von WGS84 in Lambert Austria (LAM) die Rasterkacheln von Quadraten in aufgestellte Rechtecke verwandelt werden und demnach die Höhenverhältnisse nicht mehr mit dem ursprünglichen Raster übereinstimmen.

²⁵ Eigene Abbildung der GDEM.

3.2.1 Ausrichtung

Auf Basis des Höhenrasters wird nun die Orientierung aller Hänge in Österreich ermittelt. Die Ausrichtung des Geländes wird über die maximale Neigungsrichtung zu den jeweiligen Nachbarzellen definiert. Man könnte die Ausrichtung also auch als Neigungsrichtung des Geländes bezeichnen. Die Ausgabewerte beziehen sich auf die jeweilige Kompassrichtung, in welche das Gelände abfällt. Demnach sind Werte von 0 bis 360 Grad möglich. Da ebene Flächen über keinerlei Neigung verfügen, ist es auch nicht möglich für diese eine Ausrichtung zu ermitteln, deshalb wird ebenen Flächen der Wert „-1“ zugewiesen.²⁶

Um jede Zelle eines Rasters mit einem Ausrichtungswert versehen zu können, sieht der Berechnungsalgorithmus²⁷ vor, dass über jede einzelne Zelle der Höhendatenschicht ein Nachbarschaftsraster mit der Größe von 3×3 Rasterzellen gelegt (siehe Abbildung 3.5) wird.

<i>A</i>	<i>B</i>	<i>C</i>
<i>D</i>	E	<i>F</i>
<i>G</i>	<i>H</i>	<i>I</i>

Abbildung 3.5: Nachbarschaftsraster²⁸

Dieser Nachbarschaftsraster iteriert dann über alle Rasterzellen und berechnet anhand der Formel 3.3²⁹ die Ausrichtung für die zentrale Zelle *E*. Die Änderungsraten $[dz/dx]_{Ausr}$ und $[dz/dy]_{Ausr}$ beschreiben dabei die Veränderung des Geländes gegenüber der Zelle *E* in Richtung X (Nord–Süd) bzw. Y (Ost–West).

$$\begin{aligned}
 [dz/dx]_{Ausr} &= \frac{(C + 2F + I) - (A + 2D + G)}{8} \\
 [dz/dy]_{Ausr} &= \frac{(G + 2H + I) - (A + 2B + C)}{8} \\
 \text{Ausrichtung}_{(E)} &= 57,29578 \times \arctan 2 \left([dz/dy]_{Ausr}, -[dz/dx]_{Ausr} \right) \quad (3.3)
 \end{aligned}$$

Der ermittelte Ausrichtungswert in Zelle *E* wird nun noch durch folgende Bedingungen

²⁶ Vgl. dazu: ESRI 2012, Tool: Aspect (SA).

²⁷ Siehe auch: BURROUGH und MCDONNELL 1998, S. 190ff

²⁸ Eigene Abbildung nach ESRI 2012, Tool: Aspect (SA).

²⁹ ESRI 2012, Tool: Aspect (SA).

in Formel 3.4³⁰ in Gradwerte umgerechnet.

$$\begin{aligned}
 & \text{if } Ausrichtung_{(E)} < 0 : E = 90 - Ausrichtung_{(E)} \\
 & \text{elseif } Ausrichtung_{(E)} > 90 : E = 360 - Ausrichtung_{(E)} + 90 \\
 & \text{else :} & E = 90 - Ausrichtung_{(E)} \qquad (3.4)
 \end{aligned}$$

Als Ergebnis dieser Berechnung erhält man, wie in Abbildung 3.6 ersichtlich, einen Raster, welcher die bereits oben beschriebenen Gradwerte enthält.

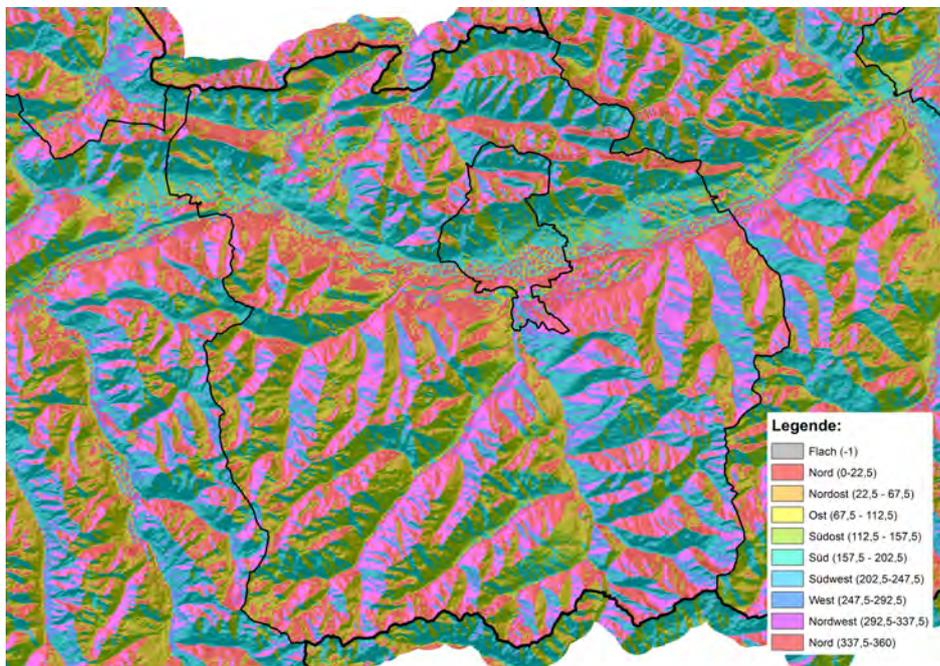


Abbildung 3.6: Ausrichtung (Bezirke Innsbruck Stadt und Land)

3.2.2 Geländeneigung

Die Bestimmung der Geländeneigung erfolgt mit dem „Slope“ SA Werkzeug, welches für jede Rasterzelle die maximale Änderungsrate des Z-Wertes (Höhenwertes) festlegt. Wie bei der Ausrichtung (Kapitel: 3.2.1) wird auch hier ein Wert, bezogen auf die 8 angrenzenden Nachbarzellen (siehe dazu Abbildung 3.5), ermittelt. Als Ergebnis erhält man einen Neigungswert, der in Grad oder Prozent ausgegeben werden kann. Wählt man die Berechnung für Prozent, wird die Geländeneigung mit dem Algorithmus der

³⁰ ESRI 2012, Tool: Aspect (SA).

Formel 3.5³¹ ermittelt.³²

$$\begin{aligned}
 [dz/dx]_{Neig} &= \frac{(C + 2F + I) - (A + 2D + G)}{8 \times ZellgrößeX} \\
 [dz/dy]_{Neig} &= \frac{(G + 2H + I) - (A + 2B + C)}{8 \times ZellgrößeY} \\
 Neigung_{[in\%]} &= \arctan \left(\sqrt{[dz/dx]_{Neig}^2 + [dz/dy]_{Neig}^2} \right) \quad (3.5)
 \end{aligned}$$

Es ist zu beachten, dass in der WGS84 Projektion des Digitalen Höhenmodells (DHM) von ASTER die Neigung nur als Prozentwert korrekt ausgegeben werden kann. Die Ermittlung der Gradwerte würde aufgrund der Höhenangabe in Meter und der Positionsbeschreibung in Längen- und Breitengraden zu falschen Ergebnissen, mit der im GIS implementierten Funktion, führen. Da die Interpretierbarkeit von Gradwerten einfacher ist, werden die mit dem „Slope“ SA Werkzeug ermittelten Prozentwerte nach der Formel 3.6 in Grad umgerechnet.

$$Neigung_{[in^\circ]} = \frac{360}{2 \times \pi} \times \arctan \left(\frac{Neigung_{[in\%]}}{10^7} \right) \quad (3.6)$$

Alternativ könnte man natürlich anstelle der manuellen Umrechnung mit Formel 3.6 auch einfach den $Z_{Höhen}Faktor$ ³³ und die Gradeinstellung verwenden und man würde grundsätzlich auf das gleiche Ergebnis kommen. Allerdings hätte man hier eine Verzerrung (Nord/Süd) der Werte in der Zielprojektion, da im GIS nur ein globaler $Z_{Höhen}Faktor$ möglich ist. Der Output der manuellen Berechnung gemäß Formel 3.6 ist in Abbildung 3.7 ersichtlich und enthält die Neigungswerte in Grad. Etwas nordwestlich vom Zentrum in Galtür ist beispielsweise der berüchtigte Lawinengang, welcher 1999 zu einem großen Schadensereignis geführt hat, ersichtlich. Dieser Hang hat gemäß der errechneten Neigung an den steilsten Stellen bis zu ca. 55° und bewegt sich im Allgemeinen zwischen etwa 30 - 45°, was die Neigung mit der höchsten Lawinengefahr³⁴ ist. Neben den vielen steilen Hängen in dieser Region sind auch die flachen Täler in der Abbildung gut zu erkennen.

³¹ ESRI 2012, Tool: Slope (SA).

³² Vgl. BURROUGH und McDONNELL 1998, S. 190ff.

³³ Siehe dazu: Kapitel 3.2.4.

³⁴ Weitere Informationen dazu: EDLINGER, STAUBE und STOCK 2003, S. 11.

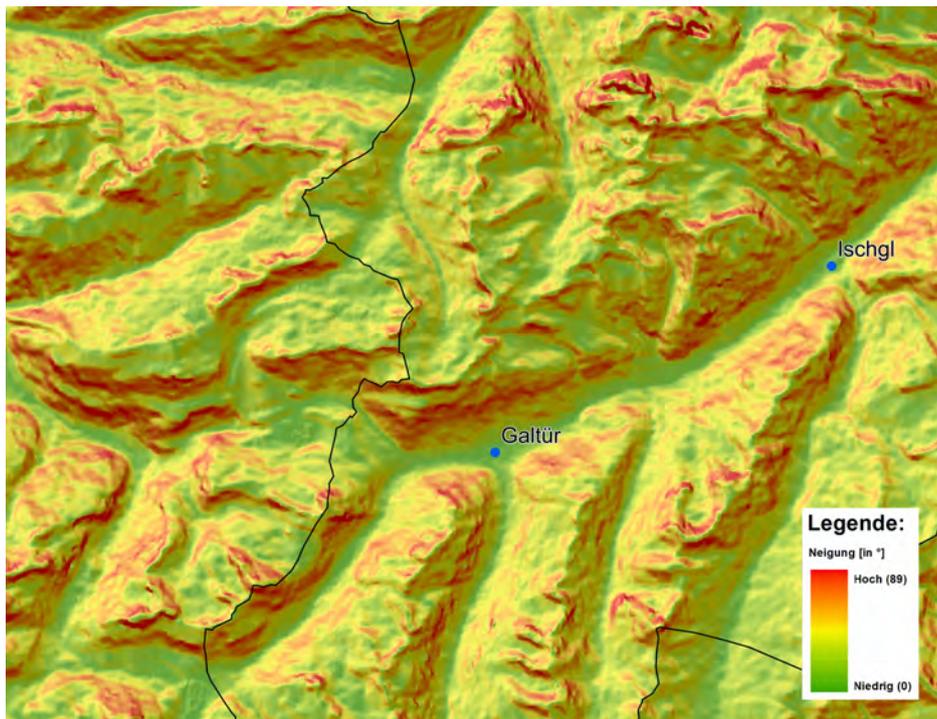


Abbildung 3.7: Neigung [in°] (Bezirke Bludenz und Landeck)

3.2.3 Gelände­krümmung

Die Krümmung eines Geländes gibt eine genaue Auskunft über die umliegenden Höhenverhältnisse. Es soll etwa die Fragestellung, ob man sich in einer Mulden- oder Kuppenlage befindet, beantwortet werden. Das im GIS eingebaute Werkzeug „Curvature“ SA funktioniert ebenfalls wieder unter Zuhilfenahme der 8 angrenzenden Rasterzellen (siehe dazu Abbildung 3.5). Bei dieser Funktion wird für jede Zelle ein Polynom Z vierter Ordnung (Formel 3.7³⁵) errechnet.

$$Z = Ax^2y^2 + Bx^2y + Cxy^2 + Dx^2 + Ey^2 + Fxy + Gx + Hy + I \quad (3.7)$$

Auf Basis dieser Polynomoberfläche können nun für jede der neun Zellen Koeffizienten errechnet werden. Wenn man diese nun in Beziehung mit den neun Höhenwerten setzt, ergibt sich daraus für D und E die Formel 3.8.³⁶ Alle weiteren Zellbeziehungen sind unter ESRI 2012 Tool: Curvature SA angeführt, weiterführende Informationen dazu

³⁵ ESRI 2012, Tool: Curvature (SA).

³⁶ Formel nach: ESRI 2012, Tool: Curvature (SA).

sind auch unter MOORE, GRAYSON und LADSON 1991, 9ff und ZEVEBERGEN und THORNE 1987 zu finden.

$$D = \frac{\left(\frac{Z_D + Z_F}{2 - Z_E}\right)}{\text{Zellgröße}^2}$$

$$E = \frac{\left(\frac{Z_B + Z_H}{2 - Z_E}\right)}{\text{Zellgröße}^2} \quad (3.8)$$

Der Output des SA Krümmungswerkzeuges ist dann die zweite Ableitung daraus, die sich wie in Formel 3.9³⁷ darstellt.

$$\text{Krümmung} = -2(D + E) \times 100 \quad (3.9)$$

Der $Z_{\text{Höhen}}\text{Faktor}$ ³⁸ wird bei der Ermittlung der Gelände­krümmung nicht verwendet, da dieser Wert für die Ermittlung nicht unbedingt notwendig ist. Diese mathematisch doch recht aufwendige Berechnung der Krümmung führt bei visueller Kontrolle leider nur bedingt zum gewünschten Ergebnis. Dies ist vermutlich darauf zurückzuführen, dass immer nur die 8 angrenzenden Rasterzellen für die Ermittlung der Geländeform herangezogen werden. Möchte man nun eine etwas großräumigere/generalisiertere Auswertung der Gelände­krümmung ermitteln, kann man mit der Formel 3.11³⁹ eine differenziertere konvex/konkave-Oberfläche kk unter Einbeziehung von mehr als 8 Nachbarzellen berechnen. Hierbei ist DHM der komplette Höhenraster und n die Anzahl der definierten Rasterzellen. Die Funktion „Focal Statistics“ SA ermöglicht es für alle Eingabezellen eine Zellstatistik über die definierte Nachbarschaft zu berechnen. Die Nachbarschaft wurde in diesem Fall als Kreisfläche K (siehe Abbildung: 3.8) gewählt, welche ausgehend vom Berechnungspunkt (Processing cell) mit einem Radius von R_n Rasterzellen ermittelt wird.

³⁷ ESRI 2012, Tool: Curvature (SA).

³⁸ Siehe dazu: Kapitel 3.2.4.

³⁹ Eigene Formel nach Grundidee von: RIEDL und KALASEK 2008, 49ff.

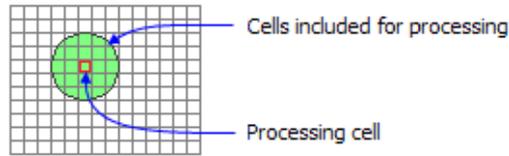


Abbildung 3.8: Kreisnachbarschaft (Focal Statistics)⁴⁰

Die zu berechnende Statistik wird über den Typ St_{Typ} beschrieben, wobei bei der Krümmungsanalyse entweder das arithmetische Mittel \bar{x} oder der Median \tilde{x} zur Anwendung kommt (Formeln 3.10). Die beiden Statistiktypen werden in Abhängigkeit von den zu berechnenden Daten gewählt. Dabei kommt es beim Digitalen Höhenmodell (DHM) durch die durchwegs kontinuierliche Oberfläche bei kleinen Radien R_n nur zu geringen Unterschieden. Je größer der Radius und somit das Analysegebiet für die Krümmung des Geländes wird, desto größer wird auch der Unterschied zwischen den Statistiktypen. Dementsprechend wird aufgrund der größeren Robustheit gegenüber Ausreißern der Median \tilde{x} bei allen Radien über zwei Zellen verwendet. RIEDL und KALASEK 2008 verwenden nur \bar{x} , was bei großen Radien allerdings teilweise zu etwas verfälschten Ergebnissen führt.

$$\begin{aligned} \bar{x}_{arithm} &= \left(\frac{1}{n} \sum_{i=1}^N x_i \right) \\ \tilde{x}_{median} &= \begin{cases} x_{\frac{n+1}{2}} & n \text{ ungerade,} \\ \frac{1}{2} (x_{\frac{n}{2}} + x_{\frac{n}{2}+1}) & n \text{ gerade.} \end{cases} \end{aligned} \quad (3.10)$$

Alle Zellen, die im Digitalen Höhenmodell (DHM) „NoData“ sind, werden bei der Berechnung ignoriert, es sollten aber grundsätzlich keine fehlenden Werte im Höhenmodell vorhanden sein. Die Ermittlung der Krümmung erfolgt dann nach der Formel 3.11, in welcher für alle Rasterzellen N , die innerhalb von Österreich liegen (der Digitale Höhenmodell (DHM) Raster ist größer als Österreich, deshalb kann hier die Randthematik außer Acht gelassen werden) vom Höhenraster DHM die soeben beschriebenen „Focal

⁴⁰ Grafikquelle: ESRI 2012 Tool: Focal Statistics SA.

Statistics“ für die jeweilige Indexzelle i subtrahiert werden.

$$\begin{aligned}
 kk_{R_n}^i &= DHM_i - Focal\ Statistics_i \left(DHM_i, K, R_n, St_{Typ} = \begin{cases} \bar{x}_{arithm} \\ \tilde{x}_{median} \end{cases} \right) \\
 kk_{R_n} &= \sum_{i=1}^N kk_{R_n}^i
 \end{aligned} \tag{3.11}$$

In Abbildung 3.9 sind die Ergebnisse für die verschiedenen Radien im Vergleich mit der im GIS implementierten Curvature-Funktion (SA) dargestellt. Betrachtet man nun die Abbildung 3.9(e) etwas genauer, kann man feststellen, dass dieser Algorithmus an einigen Stellen zu nicht erwarteten Ergebnissen führt. In der Mitte der Abbildung 3.9(e) und auch an einigen anderen Stellen entstehen gitterförmige Anordnungen der Rasterzellen, welche mit den tatsächlich vorkommenden Geländestrukturen nichts zu tun haben. Weiters kann auf den Hängen häufig ein sehr unruhiger Wechsel der Krümmungseigenschaften beobachtet werden, wobei hier kaum der Geländeverlauf sichtbar wird, sondern lediglich geringe Geländeunebenheiten abgebildet werden. Betrachtet man hingegen etwa die Ergebnisse in Abbildung 3.9(d) bzw. auch alle anderen Radien in Abbildung 3.9, kann man erkennen, dass die vorherrschende Geländestruktur bei diesen deutlich besser abgebildet wird. In Abhängigkeit vom Radius R_n wird bei kleinen Radien die Veränderung zu den direkten Nachbarn und bei großen eine weiträumige konvex-konkave Geländeoberfläche visuell fassbar. Die hier gewählten Radien beziehen sich auf die Rasterzellen des Höhenrasters, demzufolge entspricht eine Rasterzelle 30m. Beim größten abgebildeten Radius ($R = 15$ in Abbildung 3.9(h)) beträgt der Analyseradius also $15 \times 30 = 450m$ und beschreibt somit in einer sehr weiträumigen Nachbarschaftsabgrenzung die Gelände­krümmung.

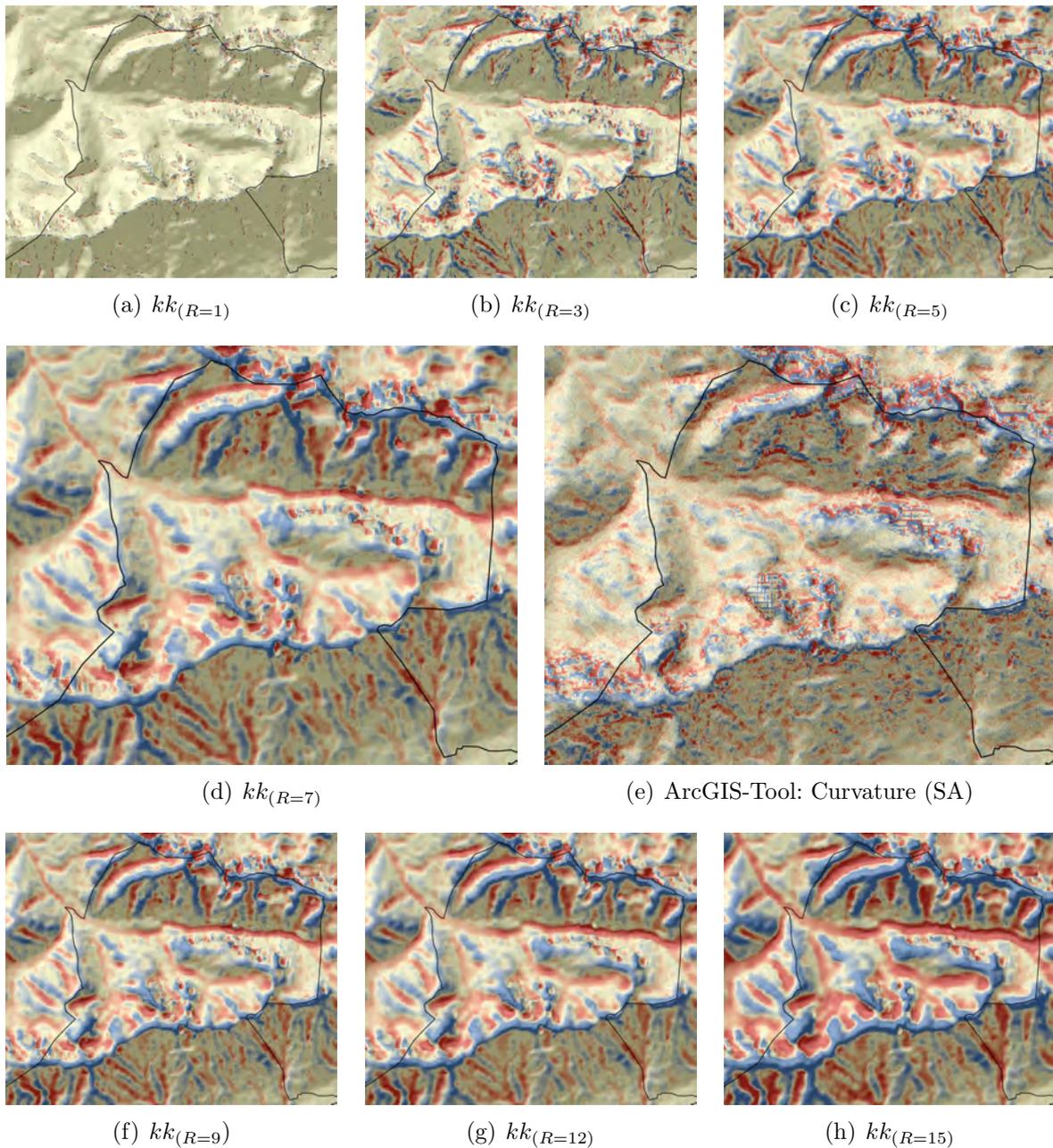


Abbildung 3.9: Gelände­krümmung (Beispiel: Nördlicher Teil des Bezirks Innsbruck(Stadt))

3.2.4 Geländeschummerung

Mit dem ESRI 2012 Tool Hillshade SA ist es möglich, unter Berücksichtigung des Beleuchtungswinkels eines Standortes eine Schummerungsoberfläche zu errechnen. Weiters kann neben dem Beleuchtungswinkel auch jede Zelle, die im Schatten einer anderen

liegt, ermittelt werden.

Da sich bei der WGS84 Projektion die Latitude (Breitengrad) (LAT) vom Äquator (Nulllinie-0°) ausgehend zu den Polen hin immer weiter verkleinert, ist es bei der Berechnung der Schummerung notwendig, in Abhängigkeit von der geografischen Lage einen $Z_{Höhen}Faktor$ in die Berechnung einfließen zu lassen. Da sich Österreich in etwa von $46^{\circ}22' - 49^{\circ}01' N$ und von $9^{\circ}17' - 17^{\circ}09' E$ erstreckt, wurde als Lagemittelwert $AUT_{\bar{x}} = 47^{\circ}41' N, 13^{\circ}20' E$ (südwestlich vom Wolfgangsee im Salzkammergut) bestimmt. Ausgehend von $AUT_{\bar{x}}$ wurde der $Z_{Höhen}Faktor$ von $1,343 E-5$ unter Zuhilfenahme der Latitude/Z-Faktor Tabelle in ESRI 2012 Tool: Hillshade SA gefunden. Beim Berechnen der Schummerung wurden für die Erzeugung der Reliefkarte die voreingestellten Werte bei Azimut ($\alpha = 315^{\circ} NW$)⁴¹ und Elevation ($\sphericalangle_{Sun} = 45^{\circ}$)⁴² für die fiktive Sonne verwendet. Für die Erzeugung der Schummerungsdatenschicht wurden für den bereits beschriebenen Standort $AUT_{\bar{x}}$ unter Zuhilfenahme der Website von SUNEARTHTOOLS 2012⁴³ die in Tabelle 3.3 dargestellten Werte ermittelt. Die Berechnung der Position der Sonne am Himmel nach SUNEARTHTOOLS 2012 erfolgt in Abhängigkeit vom Standort auf der Erde unter Bezugnahme auf Datum und Uhrzeit. In Abbildung 3.10 ist beispielsweise der Sonnenverlauf am 01.01.2012 eingezeichnet, und man kann hier erkennen, dass die Sonne um 12:00 Uhr (gelbe Line) bei $\alpha = 163^{\circ}$ und $\sphericalangle = 17^{\circ}$ steht. Die exakten Berechnungswerte, welche in Tabelle 3.3 dargestellt sind, wurden in der „Sun-Tabelle“ auf der SUNEARTHTOOLS 2012 Homepage abgelesen.

⁴¹ Die Winkelrichtung der Sonne.

⁴² Der Winkel der Sonne über dem Horizont.

⁴³ Link: http://www.sunearthtools.com/dp/tools/pos_sun.php

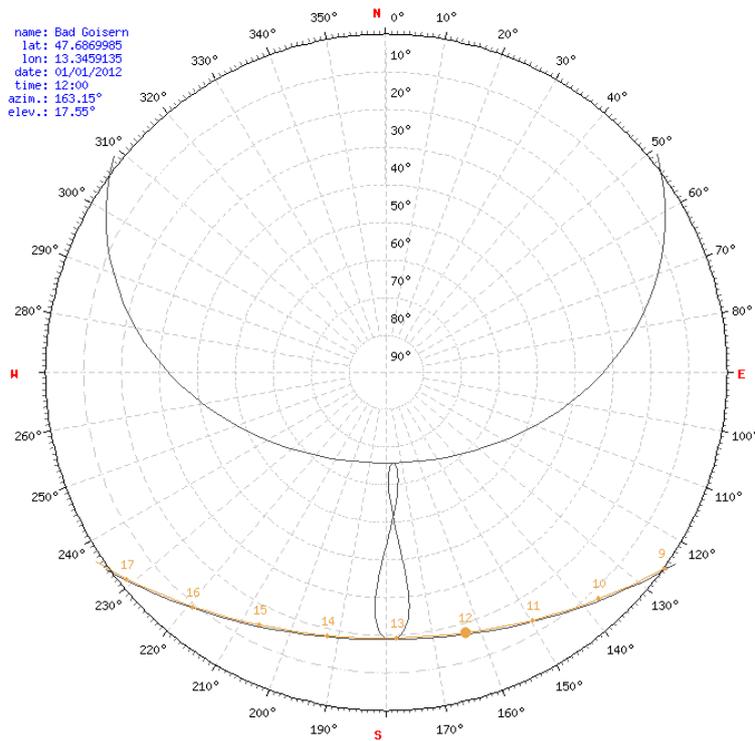


Abbildung 3.10: Azimut α und Elevation(\sphericalangle) der Sonne in $AUT_{\bar{x}}$, am 01.01.2012, 12:00 Uhr⁴⁵

Datum, Uhrzeit	Azimut α	Elevation(\sphericalangle)
31.03.2012, 12:00	155,06°	44,12°
30.06.2012, 12:00	144,06°	61,63°
30.09.2012, 12:00	162,24°	37,43°
31.12.2012, 12:00	163,18°	17,54°

Tabelle 3.3: Sonnenposition am Standort $AUT_{\bar{x}}$ ⁴⁶

Mit den ermittelten Werten wird nun jeweils eine Schummerungsberechnung durchgeführt, und man erhält dadurch einen Kennwert, welcher als Strahlungsintensität einer definierten Lichtquelle interpretiert werden kann. Es wird demnach für jeden definierten Zeitpunkt in der Tabelle 3.3 eine hypothetische Beleuchtung auf der Rasteroberfläche nach der Formel 3.16 errechnet. Um diese Berechnung durchführen zu können, werden vom ESRI 2012 Tool: Hillshade SA die nachfolgenden Zwischenschritte (Formeln 3.12

⁴⁵ Grafik wurde mit SUNEARTHTOOLS 2012 erstellt.

⁴⁶ Eigene Zusammenstellung nach SUNEARTHTOOLS 2012.

bis 3.14)⁴⁷ gemacht.⁴⁸

Es wäre auch möglich, anstelle dieser hypothetischen Beleuchtung die direkte Sonneneinstrahlung in Wattstunden pro Quadratmeter mit der in ArcGIS eingebauten „Area Solar Radiation“-Funktion zu berechnen, allerdings würde diese Berechnung aufgrund des dort verwendeten Algorithmus⁴⁹ für die gesamte Analyseregion zu lange dauern. Außerdem ist für die Fragestellung in dieser Arbeit die Genauigkeit der verwendeten Methode mehr als ausreichend. Zunächst wird der Höhenwinkel \sphericalangle in einen Zenitwinkel umgerechnet und das Ergebnis in Radiant dargestellt (Formel 3.12).

$$Zenith_{rad} = (90 - \sphericalangle) \times \frac{\pi}{180} \quad (3.12)$$

Im nächsten Schritt wird der Azimutwinkel α aus der Tabelle 3.3 in eine mathematische Einheit transformiert und dann in einen Radianten konvertiert (Formel 3.13).

$$\begin{aligned} \alpha_{math} &= 360 - \alpha + 90 \\ & \text{if } \alpha_{math} \geq 360 : \alpha_{math} = \alpha_{math} - 360 \\ \alpha_{rad} &= \alpha_{math} \times \frac{\pi}{180} \end{aligned} \quad (3.13)$$

„Die Richtung der steilsten Neigungsrichtung ist die Ausrichtung. Die Ausrichtung in Radianten wird im Bereich von 0 bis 2 Pi definiert, mit 0 nach Osten“⁵⁰ bemessen und wie in der Formel 3.14 beschrieben, errechnet. Die Definitionen von $[dz/dy]_{Neig}$ und $[dz/dx]_{Neig}$ sind im Kapitel 3.2.2 zu finden.

⁴⁷ Eigene Formeldarstellungen nach: ESRI 2012, Tool: Hillshade SA.

⁴⁸ Vgl.: BURROUGH und MCDONNELL 1998, 190ff.

⁴⁹ Informationen dazu sind unter: ESRI 2012, Tool: Area Solar Radiation SA zu finden.

⁵⁰ ESRI 2012, Tool: Hillshade SA.

$$\begin{aligned}
 \text{if } [dz/dx]_{Neig} <> 0 : \\
 \quad Ausrichtung_{rad} &= \arctan 2\left([dz/dy]_{Neig}, -[dz/dx]_{Neig}\right) \\
 \text{if } Ausrichtung_{rad} < 0 : \quad Ausrichtung_{rad} &= 2\pi + Ausrichtung_{rad} \\
 \text{if } [dz/dx]_{Neig} == 0 : \\
 \quad \text{if } [dz/dy]_{Neig} > 0 : \quad Ausrichtung_{rad} &= \frac{\pi}{2} \\
 \quad \text{elseif } [dz/dy]_{Neig} < 0 : \quad Ausrichtung_{rad} &= 2\pi - \frac{\pi}{2} \\
 \quad \text{else : } \quad Ausrichtung_{rad} &= Ausrichtung_{rad} \quad (3.14)
 \end{aligned}$$

Wie in Formel 3.15⁵¹ ersichtlich ist, wird die Neigung hier nun geringfügig anders berechnet, als dies in Kapitel 3.2.2 beschrieben wurde. Es wird nun auch der $Z_{Höhen}Faktor$ in die Berechnung miteinbezogen.

$$Neigung_{rad} = \arctan \left(Z_{Höhen}Faktor \times \sqrt{[dz/dx]_{Neig}^2 + [dz/dy]_{Neig}^2} \right) \quad (3.15)$$

Nachdem nun alle vorbereitenden Schritte abgeschlossen sind, kann nun mit dem folgenden Algorithmus (Formel 3.16⁵²) die endgültige Schummerung ermittelt werden.

$$\begin{aligned}
 Schummerung &= 255.0 \times \left((\cos(Zenith_{rad}) \times \cos(Neigung_{rad})) + (\sin(Zenith_{rad}) \right. \\
 &\quad \left. \times \sin(Neigung_{rad}) \times \cos(\alpha_{rad} - Ausrichtung_{rad}) \right) \quad (3.16)
 \end{aligned}$$

In der Abbildung 3.11 sind die Ergebnisse der für die vier Sonnenpositionen (Tabelle: 3.3) ermittelten hypothetischen Strahlungswerte visualisiert. Man kann hierbei gut erkennen, dass wie zu erwarten, im Juni die höchsten Strahlungswerte auftreten.

⁵¹ Eigene Formeldarstellung nach: ESRI 2012, Tool: Hillshade SA.

⁵² ESRI 2012, Tool: Hillshade SA.

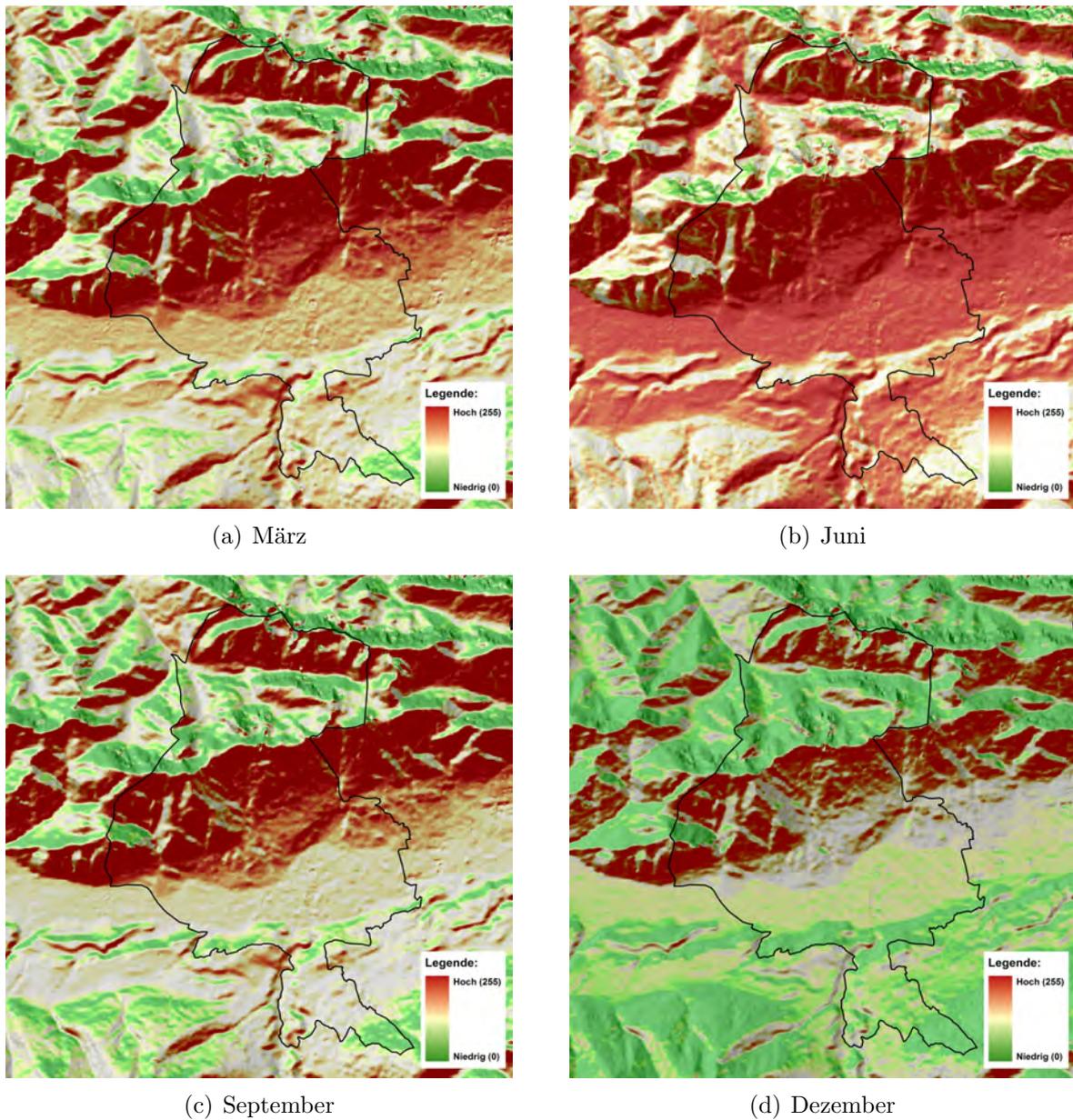


Abbildung 3.11: Hypothetische Beleuchtung (Innsbruck Stadt)

3.3 Lärmmodell

Um die Beeinträchtigung durch Lärm modellieren zu können, muss zuerst festgestellt werden, welche Arten von Lärmquellen überhaupt relevant sind. In der Mikrozensusserhebung „Umweltbedingungen, Umweltverhalten“ 2007⁵³ wird neben der Anzahl der

⁵³ WEGSCHEIDER-PICHLER 2009, S. 153-164.

Bevölkerung die Höhe der Belastung und der Anteil der jeweiligen Lärmquelle erhoben. Demnach wird Lärmstörung durch die Bevölkerung, wie in Abbildung 3.12 dargestellt, wahrgenommen.⁵⁴ Wie man erkennen kann, wird von rund 64 Prozent der Personen Verkehrslärm als Hauptlärmquelle angegeben. Die restlichen 36 Prozent werden durch sonstige Lärmquellen emittiert.

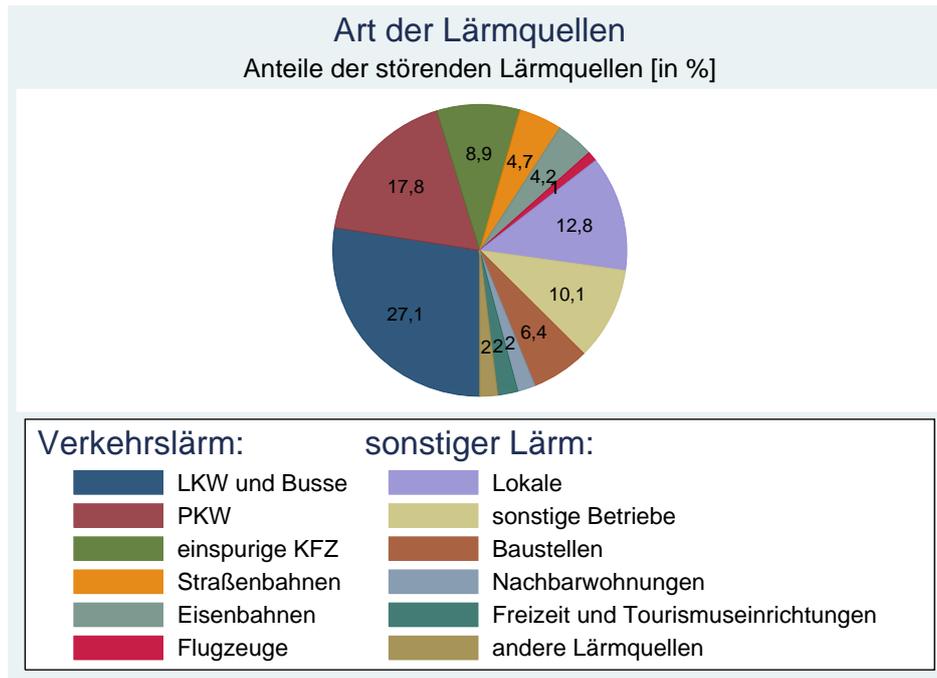


Abbildung 3.12: Verteilung der Lärmtypen⁵⁵

Für die Modellierung von Lärm wird dementsprechend das Hauptaugenmerk auf Verkehrslärm gelegt, da dieser die relevanteste Lärmquelle darstellt. In der nachstehenden Tabelle 3.4 sind alle Hauptlärmquellen, welche in die GIS-Modellierung einfließen, dargestellt. Die Lärmquellen wurden im Vergleich zu der genauen Darstellung im Kreisdiagramm etwas zusammengefasst und gekürzt. Die Lärmbelästigung durch Lokale, Nachbarwohnungen und andere Lärmquellen kann aufgrund fehlender Datengrundlagen nicht mit dem GIS modelliert werden und ist demnach als Objektkriterium im Modell anzusehen.

⁵⁴ Für weitere Informationen dazu siehe auch: WEGSCHEIDER-PICHLER 2009, S. 33-51.

⁵⁵ Eigene Grafik nach Daten von WEGSCHEIDER-PICHLER 2009, S. 157-158.

Verkehrslärm	sonstiger Lärm
Straße	Industrie
Bahn	Gewerbe
Straßenbahn	Hafengebiete
Flugzeug	Sport- und Freizeitanlagen
	Deponien, Baustellen und Abraumhalden

Tabelle 3.4: Modellierbare Lärmquellen

3.3.1 Berechnungsmethode der Lärmdatenschicht

Da Verkehrslärm auf einzelnen in Intervallen wiederkehrenden Schallereignissen beruht, werden diese durch eine fiktive Ersatzschallquelle mit konstanter Belastung ersetzt. Man spricht hier auch vom energieäquivalenten Dauerschallpegel L_{eq} , welcher eine Aussage über den konstanten Schalldruckpegel während eines bestimmten Zeitraumes zulässt. Anders formuliert könnte man auch sagen, es ist der Mittelungspegel über alle im Zeitraum gemessenen Schallereignisse. In der österreichischen Bundes-Umgebungslärmschutzverordnung⁵⁶ ist hierfür ein Tag-Abend-Nacht-Lärmindex L_{den} ⁵⁷ vorgesehen, der wie folgt berechnet wird:

$$L_{den} = 10 \log \frac{1}{24} \left(13 \times 10^{\frac{L_{Tag}}{10}} + 3 \times 10^{\frac{L_{Abend}+5}{10}} + 8 \times 10^{\frac{L_{Nacht}+10}{10}} \right) \quad (3.17)$$

Der Beurteilungszeitraum für den Lärmindex beträgt ein Jahr, und es wird – wie in der Formel 3.17 ersichtlich ist – jeweils nach 3 Tageszeiten (Tag, Abend und Nacht) unterschieden. Da es im Rahmen dieser Arbeit zu aufwendig wäre, die gesetzlich vorgeschriebene Ermittlungsvariante zu verwenden und die dafür benötigten Daten (z.B.: genaue Bebauungsinformationen, durchschnittliche tägliche Verkehrsstärke, Informationen über Lärmschutzverbauungen, etc.) nicht flächendeckend zur Verfügung stehen, wird hier auf Durchschnittswerte je Lärmquelle für den energieäquivalenten Dauerschallpegel (siehe Tabelle 3.5) zurückgegriffen.

⁵⁶ Siehe dazu: BUNDES-LÄRMV 2006, § 3 Abs. 1.

⁵⁷ Lärmindex (day/evening/night) Die Berechnung erfolgt gemäß ISO 1996-2: 1987.

Lärmgruppen (A–D) (Einzellärmquellen)	Nutzungskey ⁵⁸	geschätzter L_{eq} [in dB]
A: Straßelärm (Linienquelle: $10 \log \frac{A_r}{A_{25}}$)		
Hauptstraße (Einbahn)	11412	70
Hauptstraße und Hauptstraße (unbefestigt)	11406, 11407	73
Fernstraße (Einbahn)	11411	76
Fernstraße und Fernstraße (unbefestigt)	11404, 11405	78
Autobahn	11402	85
Flughafengelände (Rollbahn)	11601	105
B: Bahnlärm (Linienquelle: $17 \log \frac{A_r}{A_{25}}$)		
Straßenbahn	11419	65
Lokalbahn / Sonstige	11414	70
Nebenbahn	11417	75
Bahn (Fernstrecke)	11416	78
Hochleistungstrecke	11415	85
C: Verkehrslärm (Flächenquelle: $10 \log \left(\frac{A_r}{A_{25}} \right)^2$)		
Parkplatz	11303	60
Straßen und Eisenbahn (funktionell zugeordnete Flächen)	11422	65
Plattform (gr. Bahnhaltestellen)	11418	70
Flughafengelände	11602, 11603	80
D: sonstiger Lärm (Flächenquelle: $10 \log \left(\frac{A_r}{A_{25}} \right)^2$)		
Einkaufszentrum	10801	65
Messegebiete	11308	65
Sportplatz	11905	65
Firmengelände, Gewerbegebiet	11301, 11306	70
Sport- und Freizeitanlagen	11903, 11904	70
Industrie- und Gewerbeflächen	11305	75
Leichtindustrie	11307	75
Hafengebiete	11502	75
Abbauflächen	11701	75
Deponien, Abraumhalden	11702	75
Baustellen	11703	75
Stadiongelände	11902	75
Industriegebiet	11302, 11304	80
Schwerindustrie	11309	80
Industrielles Hafengebiet	11501	80

 Tabelle 3.5: Angenommene L_{eq} [in dB] pro Lärmquelle

Da für die unterschiedlichen Lärmquellen keine geeignete Kennwerttabelle vorhanden

⁵⁸ Siehe in Tabelle 3.2.

ist, wurde L_{eq} auf Basis der Werte in den folgenden Quellen KRAWIETZ und HEIMKE 2007, S. 166; WILLEMS, SCHILD und DINTER 2006, Kap. 7.16ff (man beachte: die Tabellen in Kap.7.52f zu den vereinfachten Schätzverfahren); FORUM SCHALL 2006, S. 1–2; FALLAST und SAMMER 1994, S. 10; SENGPIEL 2012; KALIVODA und BUKOVNIK 2010, 29ff und BAFU, BUNDESAMT FÜR UMWELT 2008, 10ff als eine Art allgemeingültiger Durchschnittswert⁵⁹ pro berechnungsrelevanter Lärmquelle ermittelt. Dieser „Durchschnittswert“ (L_{eq}) darf nicht mit dem oben beschriebenen L_{den} verglichen/verwechselt werden, da sich dieser auf die direkte Lärmemission von verschiedenen Lärmquellen bezieht und kein Tagesmittel ist. Da für die Straßen keine flächendeckenden Angaben bezüglich der durchschnittlichen täglichen Verkehrsstärke (DTV) vorliegen, wurde diese den Straßenkategorien entsprechend eingeschätzt, wobei die Kategorie Nebenstraße⁶⁰ nicht für die Lärmmodellierung verwendet wird. Bei der Bahn wurden alle Kategorien verwendet, und es wurde auf die Dezibelverringerung bezüglich des Schienenbonus gemäß § 2 Abs. 4 SCHIV 2012 verzichtet.

3.3.2 Lärm-Basisformel

Die Modellierung des Lärmrasters erfolgt auf Basis folgender Grundformel:⁶¹

$$L_r = L_{25} - (\log_{Typ} SA_{Typ}) - DK + \sum_{\substack{a \in \{RB, \\ Steigung, Fl\}}} \Delta P_a \quad (3.18)$$

Die Formel 3.18 beschreibt die Ermittlung der Lärmbelastung L in einer Entfernung r zur Lärmquelle, wobei L_{25} als energieäquivalenter Dauerschallpegel in 25m Entfernung zur Lärmemissionsquelle für die Lärmbestimmung verwendet wird. Neben dem $10 \log$ kann in Bodennähe und bebautem Gelände auch $17 \log$ verwendet werden, in der Formel wird dies als \log_{Typ} dargestellt. SA_{Typ} ist die Schallabnahme je Entfernung, welche in Abhängigkeit von der Schallquelle unterschiedlich ermittelt wird. Demnach wird bei Punkt- oder Flächenschallquellen, wie etwa bei Parkplätzen oder Industrieanlagen, eine kugelförmige Schallausbreitung SA_{Kugel} und bei linienförmigen Schallquellen, was im

⁵⁹ Es wurde aus Performancegründen im GIS auf ganze Zahlen gerundet.

⁶⁰ Siehe dazu in Tabelle 3.2 NK: 11408–11410 und 11413.

⁶¹ Vgl. KRAWIETZ und HEIMKE 2007, S. 165–167.

Wesentlichen den Verkehrslärm betrifft, eine zylindrische Schallausbreitung $SA_{Zylinder}$ verwendet. Theoretisch müsste man bei der Ermittlung von flächenförmigen Schallquellen eine etwas differenziertere Berechnung zur Anwendung⁶² bringen. Es müsste in diesem Fall für jede L_{eq} -Rasterzelle der Schall zu jeder Rasterzelle von einer Flächenquelle berechnet werden und dieser dann jeweils für eine Fläche addiert werden. Da nicht bekannt ist, an welcher Stelle eine flächenförmige Lärmquelle (z.B. Industriegebiet) tatsächlich emittiert wird, kann im GIS die Annahme getroffen werden, dass es sich bei Flächen um eine Vielzahl punktueller Lärmquellen handelt und somit auf die Summierung des Flächenschalles verzichtet werden. Durch diese Annahme wird die Ausbreitung von flächigen Schallquellen geringfügig gedämpft.

$$SA_{Kugel} = \frac{A_r}{A_{25}} = \left(\frac{r}{25}\right)^2 \quad (3.19)$$

$$SA_{Zylinder} = \frac{A_r}{A_{25}} = \frac{r}{25} \quad (3.20)$$

3.3.2.1 Dämpfungskonstante

Neben der soeben beschriebenen Methode zur Berechnung der Schallabnahme wird des Weiteren eine Dämpfungskonstante $DK = 10\mu lge$ (Einheit: $[DK] = dBm^{-1}$)⁶³ für zusätzliche Schallpegelabnahmen berücksichtigt (Formel 3.21).

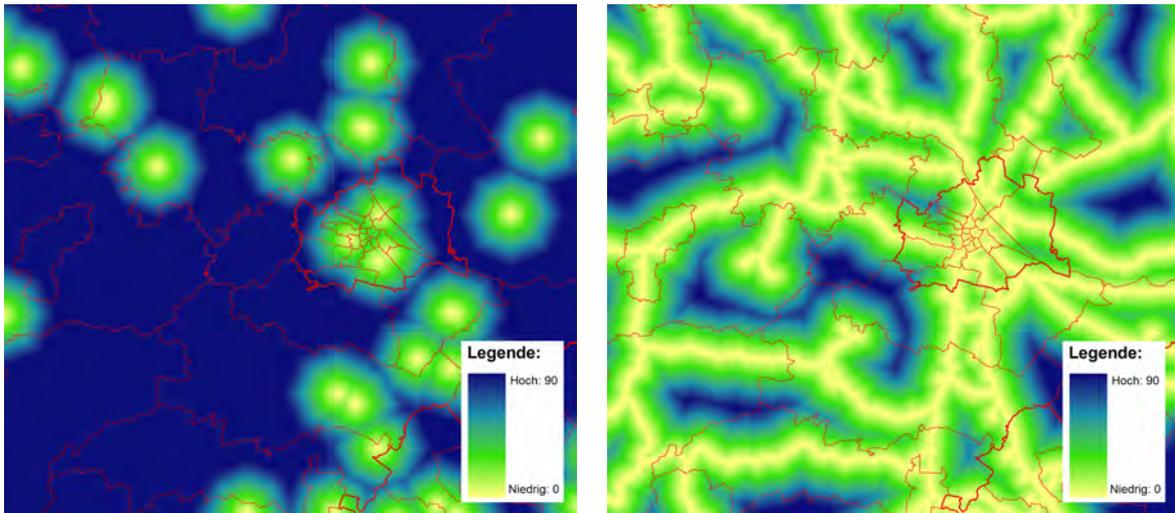
$$DK_i = 100^{-1} (A_{r,i} - A_{25})$$

$$DK = \sum_{i=1}^N DK_i \quad (3.21)$$

Betrachtet man die beispielhafte Darstellung der ermittelten Dämpfungskonstanten in Abbildung 3.13, kann man sofort erkennen, dass diese in direktem Zusammenhang mit der Entfernung von der Lärmquelle steht. Die Dämpfungskonstante DK nimmt somit mit steigender Entfernung zur Lärmquelle gleichmäßig zu. Die Berechnung der DK erfolgt für alle verwendeten Lärmquellen gleich.

⁶² Siehe dazu: HÄUPL 2008, S. 460f.

⁶³ Für weitere Information dazu siehe: KRAWIETZ und HEIMKE 2007, S. 167.



(a) DK Firmengelände, Gewerbegebiet (NK=11306)

(b) DK Fernstraße (NK=11404)

Abbildung 3.13: Dämpfungskonstante DK (Beispiel: Wien und Umgebung)

3.3.2.2 Pegelveränderung - Bebauung

Nachdem nun die den Ausbreitungspegel beschreibenden Variablen erklärt sind, wird noch die Pegelveränderung ΔP ergänzt. Bei einer sehr genauen Lärmpegelermittlung würde man mehrere unterschiedliche Pegelveränderungen miteinbeziehen. Da es sich hier um ein sehr vereinfachtes Modell handelt, wird auf eine genaue Differenzierung wie z.B. durch unterschiedliche Straßenbeläge, Witterungseinflüsse oder Bodenbeschaffenheiten im Umfeld,⁶⁴ etc. mangels geeigneter Daten verzichtet.

Eine allerdings gut umsetzbare Pegelerhöhung ist jene, die durch beidseitige Randbebauung (ΔP_{RB}) einer Straße hervorgerufen wird. Dies kann, wie in der Formel 3.22 abgebildet, ebenfalls aus $10 \log$ berechnet werden, wobei hier demnach das Verhältnis von h (mittlere Gebäudehöhe) und b (Hausfluchtenabstand) unterschieden wird.

$$\Delta P_{RB,i} = \left(10 \log \left(1 + \frac{h_i}{b_i} \right) + 3 \right) - \left(\log_{Typ} = \begin{cases} 10 \log & SA_{Typ} = \left(\frac{A_{r,i}}{A_{25}} \right)^2 \\ 17 \log & \left(\frac{A_{r,i}}{A_{25}} \right) \end{cases} \right)$$

$$\Delta P_{RB} = \sum_{i=1}^N \Delta P_{RB,i} \quad (3.22)$$

⁶⁴ Vgl. HEUTSCHI und ROSENHECK 1997, S. 33.

Das Verhältnis wird auf Basis der Nutzungsdatenschicht im GIS anhand von visuellen Vergleichen mit dem Luftbild eingeschätzt. Die Einschätzung wird nur für die bebauten Flächentypen in der Nutzungsdatenschicht vorgenommen, wobei diese Flächen zuvor auf gleichmäßige Verbauung hin geprüft wurden. Da derartige Pegelerhöhungen hauptsächlich für Verkehrslärm relevant sind und sich diese direkt an der Lärmquelle auswirken, wurde auf die Einschätzung von Flächentypen wie z.B. Hafen- und Industriegebieten oder Sport- und Freizeitanlagen bewusst verzichtet. Die Ermittlung von Durchschnittswerten für diese Flächentypen wäre des Weiteren äußerst problematisch, da diese Flächen nicht unbedingt über eine homogene Bebauung verfügen. Die ermittelten Durchschnittswerte für das Verhältnis zwischen mittlerer Gebäudehöhe und dem Hausfluchtenabstand sind in Tabelle 3.6 visualisiert.

Flächentyp	Nutzungskey ⁶⁵	Verhältnisschätzung $\frac{h}{b}$
Universitäts- oder Fachhochschulgelände	11001	9/30 = 0,3
Kleingarten	11202, 11205	6/20 = 0,3
Einkaufszentrum	10801	9/30 = 0,3
Industrie- und Gewerbeflächen, Gewerbegebiet, Leichtindustrie	11305–11307	9/30 = 0,3
Messegebiete	11308	9/30 = 0,3
Firmengelände	11301	9/25 = 0,36
Wohnpark	11204	9/20 = 0,45
nicht durchgängig städtische Prägung	11101	12/20 = 0,6
Krankenhaus	11002	21/30 = 0,7
durchgängig städtische Prägung	11004	18/20 = 0,9

Tabelle 3.6: Angenommenes $\frac{h}{b}$ Verhältnis pro Lärmquelle

Das Ergebnis der Pegelveränderung ΔP_{RB} aus der Formel 3.22, welches anhand des $\frac{h}{b}$ Verhältnisses errechnet wurde, kann in Abbildung 3.14 betrachtet werden. Hier ist deutlich erkennbar, dass es in Abhängigkeit von der Bebauungsstruktur eine Pegelerhöhung gibt und diese dementsprechend im Stadtzentrum höher ausfällt als in den ländlichen Regionen rund um Wien. Im Grünland oder sonstigen unbebauten Gebieten wird keine Pegelveränderung vorgenommen (Gelb in der Abbildung 3.14). Die

⁶⁵ Siehe in Tabelle 3.2.

Pegeleränderung durch die Entfernung von der Lärmquelle wird an dieser Stelle nicht dargestellt.

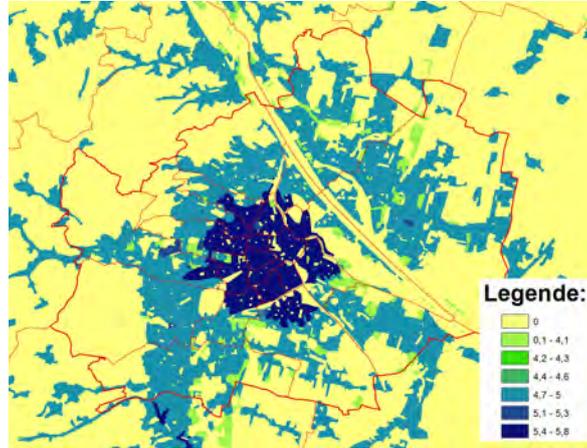


Abbildung 3.14: Pegelveränderung $\Delta P_{RB}[\text{indB}]$ (Beispiel: Wien)

3.3.2.3 Pegelveränderung - Steigung

Der Einfluss der Geländesteigung auf den Lärmemissionsausstoß bei verkehrsbedingten Lärmquellen ist weitläufig bekannt. Deshalb werden alle Lärmquellen, die gemäß Tabelle 3.5 der Lärmgruppe A angehören (ausgenommen Flughafengelände NK 11601) mit einer Pegelerhöhung, welche die Steigung im Gelände berücksichtigt, errechnet. Für die Geländeneigung wird die in Kapitel 3.2.2 beschriebene Datenschicht herangezogen. Die Steigung g [in %] wird, wie in Formel 3.23 dargestellt, als Pegelerhöhung im Modell berücksichtigt, wobei diese Erhöhung nur zutrifft, wenn die Steigung größer als 5 Prozent ist.⁶⁶

$$\Delta P_{\text{Steigung},i, g>5[\%]} = (g_i[\%] - 5) \times 0,6[\text{dB}]$$

$$\Delta P_{\text{Steigung}} = \sum_{i=1}^N \Delta P_{\text{Steigung},i, g>5[\%]} \quad (3.23)$$

Um bei etwaigen Datenfehlern oder Ungenauigkeiten im Höhenmodell nicht ausufernde Werte zu erhalten, wurde die Pegelerhöhung mit der maximalen Steigung von 20 Prozent gedeckelt, dementsprechend ($0,6 \times 20 = 12\text{dB}$) kann der Lärmpegel durch Stei-

⁶⁶ Siehe dazu: HEUTSCHI und ROSENHECK 1997, S. 36.

gung um maximal $12dB$ erhöht werden. In der Abbildung 3.15 ist die Pegelerhöhung basierend auf der Straßensteigung für die Westautobahn (A1) und die Wiener Außenringautobahn (A21) visualisiert. Es ist gut zu erkennen, dass diese zwei Autobahnen, welche sich westlich von Wien durch den Wienerwald schlängeln durchaus einige erhebliche Steigungen aufweisen und eben an diesen Stellen eine höhere Lärmbelastung ermittelt wird.

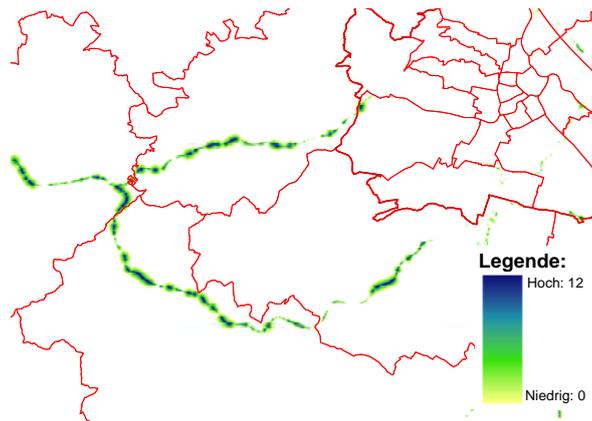


Abbildung 3.15: Pegelveränderung $\Delta P_{Steigung}$ (Beispiel: Autobahn NK)11402

3.3.2.4 Pegelveränderung - Flächendämpfung

Eine weitere äußerst wichtige Pegelminderung ist die durch Flächenausprägungen hervorgerufene Dämpfung. Hierfür werden aus der Nutzungsdatenschicht alle relevanten Flächentypen selektiert und für diese eine zuvor festgelegte Dämpfung errechnet. Die Grundidee für die Berechnungslogik stammt aus der Syntax der SPreAD-GIS Programmierung⁶⁷, welche unter Zuhilfenahme einer fiktiven Dämpfungskonstante f_{DK} eine fortschreibbare Verminderung des Lärmpegels für jeden definierten Flächentyp (Dämpfungsfläche) ermöglicht. Die dort gewählte Vorgehensweise wurde etwas adaptiert und es ist somit möglich, Pegelverminderungen ΔP_{Typ} nicht nur für Laub- und Bodenbedeckung zu berechnen, sondern auch die Veränderung für andere Flächentypen zu ermitteln. Die Pegelminderungen für Bewuchs wurden von REED, BOGGS und

⁶⁷ Die Programmierung kann unter <http://warnercnr.colostate.edu/~sereed/research/SPreAD-GIS.html> geladen werden. Das Benutzerhandbuch (REED, BOGGS und MANN 2010) enthält weitere Informationen dazu.

MANN 2010 übernommen. Die angenommene Minderung deckt sich gut mit der gängigen Literatur, wo in verschiedenen Quellen immer wieder von 5–10 dB⁶⁸ je 100m Tiefe gesprochen wird. Für die neu eingeführten Flächentypen wurde ebenfalls anhand von Angaben in der Literatur f_{DK} festgelegt.⁶⁹ Beispielsweise kann hier für Bebauung in Abhängigkeit von der Bebauungsdichte zwischen 10–15 dB Pegelverminderung angenommen werden. In der nachstehenden Formel 3.24 ist in vereinfachter Form ersichtlich, wie die Berechnung vonstatten geht. Es wird zunächst manuell der Kostenraster KR für die folgende Berechnung mit dem ESRI 2012 Tool: Path Distance (SA) ermittelt. Der KR wird danach als eine Kostenoberfläche⁷⁰ im Sinne des ESRI 2012 Tool: Cost Distance (SA) in die Berechnung mit aufgenommen. Somit kann bei der fortlaufenden Ausbreitung des Schalls einer Lärmquelle LQ jede Schallbarriere in den Ausbreitungsalgorithmus einfließen, indem von der Path Distance die Distanz zur jeweiligen Schallquelle subtrahiert wird.

$$\Delta P_{Fl,i} = Path\ Distance_i \left(LQ_i, KR_i = \left(\frac{f_{DK}}{100} + 1 \right), DHM_i, DB_{Max,i} \right) - A_{r,i}$$

$$\Delta P_{Fl} = - \sum_{i=1}^N \Delta P_{Fl,i} \quad (3.24)$$

In der Tabelle 3.7 sind alle für die Flächendämpfungen verwendeten Flächentypen angeführt. Die Basiswerte wurden, wie bereits beschrieben, der Literatur entnommen und dann mittels eines „trial and error“ Testverfahrens für die einzelnen Flächentypen im GIS ermittelt.

⁶⁸ In Abhängigkeit von den Baumarten.

⁶⁹ Vgl. MÜLLER und MÖSER (HRSG.) 2003, S. 199–200.

⁷⁰ Für weitere Informationen dazu siehe Kapitel 3.4.5.

Dämpfungsfläche	Nutzungskey	modifizierte Dämpfungs-konstante f_{DK}^{71}
Laubwälder	12202	660
Wald, Mischwälder	12201, 12204	550
Nadelwälder, Krankenhaus, durchgängig städtische Prägung	12203, 11002, 11004	500
Universitäts- oder Fachhochschulgelände, nicht durchgängig städtische Prägung	11001, 11101	400
Wohnpark	11204	350
Einkaufszentrum, Kleingärten	10801, 11202, 11205	300
Messegebiete, Gewerbegebiet, Firmengelände, Industrie- und Gewerbeflächen, Leichtindustrie	11308, 11306, 11301, 11305, 11307	250
Stadiongelände, Park/Garten, städtische Grünflächen	11902, 11801, 11804, 11805, 11806	200
Industriegebiet, Schwerindustrie, industrielles Hafengebiet	11302, 11304, 11309, 11501	100
Wald/Strauch Übergangsstadien	12304	50

Tabelle 3.7: Angenommene Dämpfungskonstante f_{DK} pro Lärmquelle

Als Ergebnis der Berechnung der Flächendämpfung erhält man für alle Lärmquellen eine eigene Datenschicht mit der Flächendämpfung ΔP_{Fl} , welche exemplarisch in Abbildung 3.16 für eine Flächen- und Linienlärmquelle visualisiert ist. Bei der Flächenlärmquelle fällt auf, dass die Flächendämpfung in einer gewissen Entfernung zur Quelle plötzlich aufhört und nur mehr ein weißer Hintergrund zu erkennen ist. Dies hat den Hintergrund, dass in der Programmierung aus Performancegründen nur jene Flächen berechnet wurden, die bei einer normalen Schallabnahme gemäß Formel 3.18 (allerdings ohne ΔP_{Fl}) auch noch einen DB -Wert aufweisen würden und alle anderen Flächen – demnach mit DB_{Max} – ausgeschlossen werden müssten. Bei NK 11306 wird gut sichtbar, dass ausgehend von einer Lärmquelle je nach Bewuchs oder Bebauung die Flächendämpfung in alle Richtungen variieren kann. Bei der Fernstraße auf der rechten Seite darf man sich nicht von den sehr hohen Werten im Westen von Wien täuschen lassen – diese kommen nur durch große Entfernungen zwischen den verschiedenen Straßen

⁷¹ Nach REED, BOGGS und MANN 2010.

zustande.

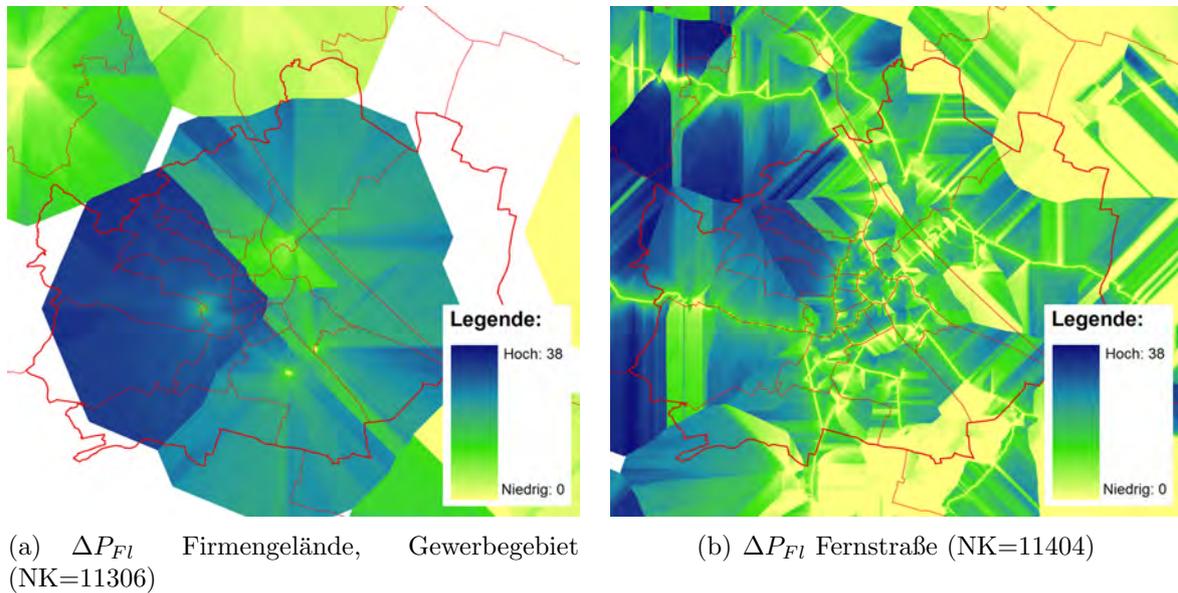


Abbildung 3.16: Flächendämpfung ΔP_{Fl} (Beispiel: Wien)

3.3.3 Zusammenführung der Lärmvariablen

Wenn man nun die einzelnen Komponenten in einer kompletten Formel für das GIS zusammenfasst, sieht dies in verallgemeinerter⁷² Form für alle Einzellärmquellen L_{Elq} , wie in Formel 3.25 dargestellt, aus.

⁷² Die nachfolgend beschriebenen Ausnahmen sind in der Formel nicht berücksichtigt.

$$\begin{aligned}
 L_{Elq,i} = L_{25,i} - & \left(\log_{Typ} = \begin{cases} 10 \log & SA_{Typ} = \left\{ \left(\frac{A_{r,i}}{A_{25}} \right)^2 \right. \\ 17 \log & \left. \left(\frac{A_{r,i}}{A_{25}} \right) \right\} \end{cases} \right) \\
 - & \left(100^{-1} (A_{r,i} - A_{25}) \right) + \left(\begin{cases} (g_i[\%] - 5) \times 0,6[dB], & g > 5[\%] \\ 0, & g \leq 5[\%] \end{cases} \right) \\
 + & \left(\left(10 \log \left(1 + \frac{h_i}{b_i} \right) + 3 \right) - \left(\log_{Typ} = \begin{cases} 10 \log & SA_{Typ} = \left\{ \left(\frac{A_{r,i}}{A_{25}} \right)^2 \right. \\ 17 \log & \left. \left(\frac{A_{r,i}}{A_{25}} \right) \right\} \right) \right) \\
 - & Path\ Distance_i \left(LQ_i, \left(\frac{fDK}{100} + 1 \right)_i, DHM_i, DB_{Max,i} \right) - A_{r,i} \\
 L_{Elq} = & \sum_{i=1}^N L_{Elq,i} \tag{3.25}
 \end{aligned}$$

Die Modellierung der unterschiedlichen Lärmquellen wird gemäß Tabelle 3.5 durchgeführt. Die Umsetzung von Bahnlärm funktioniert ähnlich wie Straßenlärm, wobei hier allerdings $\Delta P_{Steigung}$ nicht berücksichtigt wird. Weiters wird anstelle von $\log 10$ der $\log 17$ verwendet, um wie bereits oben beschrieben, einen etwas schnelleren Lärmabfall zu modellieren. Die Verwendung der Pegelminderung nach REED, BOGGS und MANN 2010, welche die verstärkte Schallabnahme (ΔP_{Typ}) durch Hindernisse wie z.B. Gebäude berücksichtigt, wird – ausgenommen Fluglärm – für alle Lärmquellen angewendet. Bei Fluglärm wird des Weiteren $\Delta P_{Steigung}$ und $\Delta P_{Randbebauung}$ nicht miteinbezogen, da diese Größen bei diesem nicht relevant sind. Die Ausnahme bei Fluglärm gilt nur für die großen Flughäfen Wien, Linz, Salzburg, Klagenfurt, Graz und Innsbruck, welche zusätzlich zum Gelände auch noch über eine Spezifikation der Rollbahn (NK 11601) verfügen. Alle anderen kleinen Sportflughäfen werden nur als Flächenlärmquelle erfasst⁷³.

Bei der Modellierung im GIS wird gemäß Tabelle 3.5 für jede einzelne Lärmquelle der geschätzte L_{eq} [in dB] herangezogen und für diesen unter Verwendung der oben beschriebenen Formeln, je nach Lärmquellentyp eine Lärmrasterschicht errechnet.

⁷³ Vgl. dazu auch Tabelle 3.5.

Anschließend werden alle diese Rasterdatenschichten mit den dB -Werten gemäß der Formel 3.26⁷⁴ aufsummiert.

$$L_{eq,(Gesamt)} = 10 \log \sum_{j=1}^N 10^{0,1 \times L_{eq,j}} \quad (3.26)$$

Addiert man nun beispielsweise 70 und 70 dB wird dies nicht zu 140 dB sondern nur zu 73,01 dB .

In der praktischen Umsetzung im GIS wird für alle linienförmigen Lärmquellen das Maximum innerhalb der jeweiligen Lärmgruppe A und B (siehe dazu Tabelle 3.5) gebildet. Innerhalb dieser Lärmgruppen erfolgt keine Summierung, da sich diese Verkehrsschallquellen an den Enden überlappen können und deshalb beispielsweise beim Addieren von „Hauptstraße (Einbahn)“ und „Hauptstraße“ am Punkt des Zusammenstoßens ein erhöhter Wert entstehen würde.

Den durch die Bahnlinien erzeugten Lärmpegel kann man in der Abbildung 3.17(a) deutlich erkennen. Es sind in den innerstädtischen Bereichen auch Straßenbahnlinien erkennbar, welche eine deutlich geringere Lärmbelastung im Vergleich mit z.B. der West- oder Südbahn in Wien aufweisen. Die in Abbildung 3.17(b) visualisierten Flächenlärmquellen (Lärmgruppe C und D) zeigen neben den funktionell zugeordneten Flächen für Eisenbahn (NK 11422) – etwa am Hauptbahnhofgelände (ehemaliger Südbahnhof in Wien) – die vielen unterschiedlichen Industrie- und Gewerbezone in und rund um Wien.

Die zwei Maximalwerte aus den Lärmgruppen A und B werden dann mit allen anderen Lärmquellen aus den Gruppen C und D gemäß der Formel 3.26 addiert. Als Ergebnis erhält man somit die in Abbildung 3.17(c) visualisierte Datenschicht, welche als die simulierte Summe aller verwendeten Umgebungslärmquellen (3.5) interpretiert werden kann. In der finalen Abbildung ist nun gut erkennbar, wie die Lärmbelastung über die Region verteilt ist. In diesem Beispiel sind die Rollbahnen am Wiener Flughafen mit den höchsten DB -Werten (Dunkelrot) erkennbar. In den bewaldeten Gegenden im Westen von Wien sind die geringsten Lärmbelastungen (Dunkelgrün) zu erkennen. Wie zu erwarten, ist entlang höherrangiger Verkehrsinfrastruktur sowie in den Industriezo-

⁷⁴ WILLEMS, SCHILD und DINTER 2006, Kap. 7.16.

nen mit erhöhtem Lärmpegel zu rechnen. Eine interessante Tatsache ist, dass bereits in diesem großen Maßstab die guten/teuren Wohngegenden Wiens (z.B.: äußerer 18., 19. oder 13. Bezirk), welche über große Flächen mit geringer Lärmbelastung verfügen, erkennbar sind.

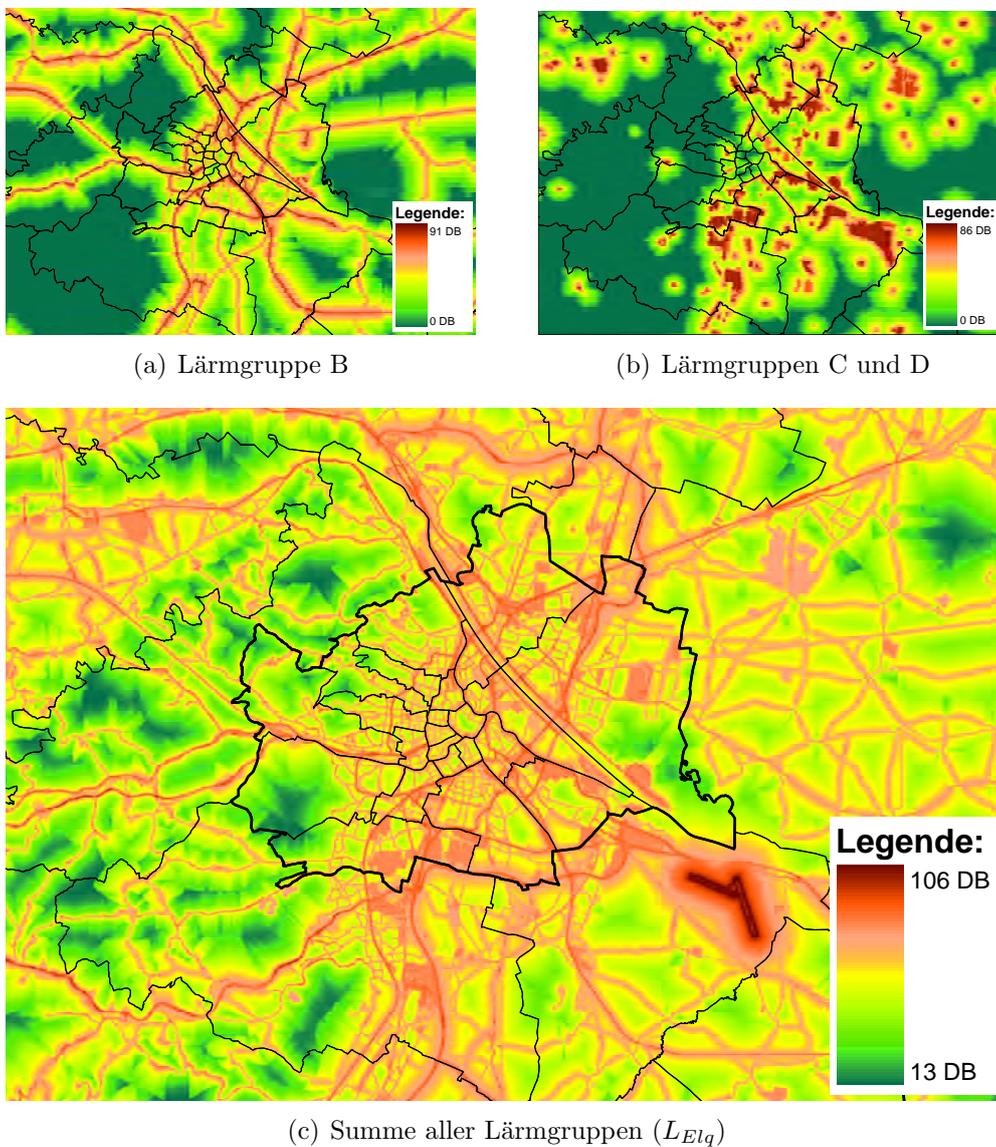


Abbildung 3.17: Simulierter Umgebungslärm (Beispiel Wien und Umgebung)

3.3.4 Anmerkung zur technischen Umsetzung

Die Entfernungsermittlung für Lärm ist im GIS immer mit der Funktion Path Distance⁷⁵ unter Zuhilfenahme des Höhenrasters erfolgt und somit ist immer die reale Entfernung verwendet worden. Weiteres wird die Zellengröße und Ausdehnung des Höhenrasters immer als Basis für die Berechnung verwendet. Demnach ist die Berechnungsgenauigkeit auf Rasterzellen von ca. 26 m beschränkt. Eine höhere Auflösung für die Berechnung des Lärmrasters in ganz Österreich wäre aus performancetechnischen Gründen nicht möglich.

3.4 Erreichbarkeitsmodell

Da in dieser Arbeit das Hauptaugenmerk auf die Bewertung von Standardwohnimmobilien gelegt wurde und dabei im Speziellen die neu generierten Lagevariablen bei der Bewertung von Einfamilienhäusern in peripherer Lage helfen sollen, wird das Erreichbarkeitsmodell nur für den motorisierten Individualverkehr (MIV) erstellt und der öffentliche Personennahverkehr (ÖPNV) bei der GIS-Modellierung außer Acht gelassen. Der ÖPNV wird in weiterer Folge im hedonischen Modell mit einer externen Datenquelle Berücksichtigung finden. Mit dem MIV Erreichbarkeitsmodell wird für jeden möglichen Standort im Analysegebiet Österreich eine homogene flächendeckende Rasterdatenschicht erzeugt, welche einen Kennwert für die Erreichbarkeit der definierten Zentren enthält.

Üblicherweise wird für die Berechnung einer Entfernung im GIS eine Netzwerkberechnung gemäß graphentheoretischer Methoden angewandt. Da es in dieser Arbeit für alle weiteren Analysen notwendig ist, über Rasterdaten zu verfügen und ein derartiges Erreichbarkeitsmodell auch mit Graphen bei der Berechnung performancetechnisch äußerst problematisch werden könnte, wurde entschieden, das Erreichbarkeitsmodell mit dem SA auf Rasterbasis im GIS zu berechnen. Der dadurch entstehende Genauigkeitsverlust bei der Berechnung spielt für die weiteren Analysen keine Rolle, da es

⁷⁵ Weitere Erklärungen dazu sind Kapitel 3.4.5 zu finden.

in erster Linie um eine Einschätzung der generellen Erreichbarkeit an allen Standorten in Österreich geht und eine „Sekundengenauigkeit“ (wie etwa beim Routing in einem Navigationssystem) nicht notwendig ist.

3.4.1 Datenvorbereitung

Da der grundlegende Straßengraph von Teleatlas nur für Österreich zur Verfügung steht, wurden, um dem Grenzproblem etwas entgegenzuwirken, die höherrangigen Verkehrsverbindungen (Autobahnen und wichtige Fern- und Hauptstraßen) sowie die 22 angrenzenden Städte mit mehr als 50.000 Einwohnern und Vaduz rund um Österreich im GIS anhand der Bing Kartenmaterialien manuell nacherfasst⁷⁶. So kann nun beispielsweise das Große bzw. Kleine Deutsche Eck auch in die Erreichbarkeitsberechnung mit einfließen, und es gibt demnach in den betroffenen Grenzregionen nur geringfügige Abweichungen von den tatsächlichen Gegebenheiten. Es ist auch möglich, Städte wie München oder Bratislava bei der Erreichbarkeitsberechnung mit einfließen zu lassen.

Die in der folgenden Tabelle 3.8 dargestellten Straßentypen wurden bereits in der Landnutzungsoberfläche verwendet. Der Unterschied hier ist allerdings, dass nun der Straßengraph als Linie herangezogen wird und nicht die extrapolierte Fläche, wie bei der Nutzungsdatenschicht in Kapitel 3.1. Da das Erreichbarkeitsmodell vor der Nutzungsdatenschicht aufgebaut wurde, ist hier beim Straßengraphen noch ein älterer Datenstand (2010) in Verwendung. Aus diesem Grund passen die NK's nicht ganz zusammen, es wurde allerdings – wie in der Tabelle 3.8 dargestellt – die alte Datenstruktur so gut wie möglich eingegliedert und für die Modellierung ein eigener Straßenkey (SK) vergeben. Hier werden auch die Straßenliniensegmente, welche sich mit Plätzen schneiden, verwendet, um an diesen eine geringere Geschwindigkeit simulieren zu können. Für alle Autobahnen in Österreich wurden mithilfe eines Navigationsgerätes ergänzend differenzierte Streckengeschwindigkeiten (60, 80, 100 und 130 km/h) manuell digitalisiert. Deshalb ist nun beispielsweise die A23 in Wien als 80 km/h Zone im Datensatz vorhanden. Die Einbahnstraßen bleiben in der Erreichbarkeitsmodellierung unberücksichtigt, da diese mit den vorliegenden Daten im gegenständlichen Rastermodell nicht abgebil-

⁷⁶ Eine Darstellung ist in Abbildung 3.25(c) ersichtlich.

det werden können und der dadurch entstehende Fehler so gering ist, dass dies für das Gesamtmodell vernachlässigt werden kann. Grundsätzlich gäbe es mittels Kostenoberflächen für alle Einbahnstraßen die Möglichkeit diese Problemstellung auch mit dem Raster zu lösen. Wie dies theoretisch funktioniert, wird in Kapitel 3.4.5 erläutert.

Straßenkey (SK)	Nutzungskey (NK)	Straßentyp (<i>km/h</i>)	Konstante GK_j
101	11402	Autobahn (60)	72
102	11402	Autobahn (80)	96
103	11402	Autobahn (100)	120
104	11402	Autobahn (130)	156
200	11404	Fernstraße	100
300	11406	Hauptstraße	75
400	11408	Nebenstraße	50
500	11410, 11405, 11407	unbefestigte Straßen, weitere Kategorien	20
600	10901	Plätze	15
700	11409	Fähren, Tunnelschleusen, etc.	5

Tabelle 3.8: Verwendete Straßentypen

3.4.2 Konvertierung des Straßengraphen in einen Raster

Im ersten Schritt (Formel 3.27) wird aus dem vorbereiteten Straßengraphen (SQ) jeder Straßentyp einzeln selektiert und in einen Straßenraster (SR) konvertiert. Hierbei ist zu beachten, dass neben den Standardeinstellungen Projektion und Ausdehnung immer auch ein Fangraster und die Zellgröße für alle Flächentypen gleich gesetzt werden, um keine fehlerhaften Zuordnungen wegen verschobenen Rasterbezügen bei den überlappenden Straßentypen zu erhalten. Die entstandene Rasterschicht SQ wird nun noch mit dem ESRI 2012 Tool: Reclassify (SA) reklassifiziert und die Werte im bestehenden Raster durch definierte Regeln mit neuen Werten versehen (Beispiel: $RK_{[0=1]}(H)$ alle Werte die im Raster H zuvor 0 waren, werden mit 1 überschrieben). In weiterer Folge wird dann für jeden Straßentyp, wie in Formel 3.28 dargestellt, ein Straßendummyraster SDR erzeugt, welcher in hierarchischer Abfolge gemäß der Reihenfolge in Tabelle 3.8 an einer Rasterzelle immer nur den Straßentyp mit der höchstmöglichen Verkehrs-

durchflussmenge aufweist. Wenn beispielsweise eine Hauptstraße mit einer Nebenstraße kreuzt, wird in der Kreuzungsfläche immer der Wert der Hauptstraße herangezogen und die Nebenstraße an dieser Stelle außer Acht gelassen.

$$SK_{all} = \{101, 102, 103, 104, 200, 300, 400, 500, 600, 700\}$$

Für $j \in \{SK_{all}\}$ gilt :

$$SR_j = RK_{\left[\begin{smallmatrix} 1-10=1, \\ NoData=0 \end{smallmatrix} \right]} \left(GridConvert_{SQ} \left(set_{\substack{Zellgröße, \\ Fangraster}}, selekt(Typ = j) \right) \right) \quad (3.27)$$

$$SR_{j(Alt)} = \begin{cases} SR_{j-1(First)} \\ RK_{\left[\begin{smallmatrix} 1-10=1, \\ NoData=0 \end{smallmatrix} \right]} (SR_j + SR_{j-1}) \end{cases}$$

$$SDR_j = RK_{[NoData=0]} \left(RK_{[1-10=1]} \left(SR_{j(Alt)} \&_{(Boolean)} SR_j \right) - SR_j \right)$$

$$SDR = \sum_j SDR_j \quad (3.28)$$

3.4.3 Reisegeschwindigkeit

Um bei der Berechnung der Reisekosten (-zeiten) die Geschwindigkeit – auch ohne genaue Informationen über Geschwindigkeitsbeschränkungen – modellieren zu können, wird in den Stadt- und Ortsgebieten sowie beim Übergang zwischen diesen eine Art zusätzliche „Brems-/Beschleunigungszone“ (B) berechnet, welche dann in Abhängigkeit von der Entfernung als Geschwindigkeitsreduktion in das Modell einfließt. Hierfür werden die bebauten Flächen (BF) gemäß Nutzungsdatschicht (Alle NK's aus den Obergruppen von 107 bis OG ≤ 113 und OG = 118) als Basisraster und alle Straßenflächen (SF) als Kostenraster mit dem gleichen Widerstandswert von 1 herangezogen, um damit entlang der Straßenverläufe eine Path Distance (SA) Entfernungsberechnung durchführen zu können. Die euklidische Distanzberechnung (SA) kann in diesem Fall nicht gewählt werden, da es mit dieser nicht möglich ist, eine Distanz entlang einer vorgegeben Strecke (die Straßen) zu berechnen. Die ermittelten Distanzen werden mittels Fuzzyfizierung⁷⁷ auf einen Wertebereich zwischen 0 und 1 gebracht, negiert und in

⁷⁷ Für weitere Informationen zu Fuzzyfizierung siehe: RIEDL 1999, S. 50ff.

weiterer Folge um 2 erhöht und reklassifiziert, um dann für alle weiteren Reisekostenberechnungen über einen Bremsraster (BR) mit Werten zwischen 1 und 2 zu verfügen, wobei 2 die Straßen in bebautem Gebiet und 1 als außer Orts – also maximale Geschwindigkeit – darstellt. Die ermittelte Rasteroberfläche, welche in Abbildung 3.18(a) visualisiert ist, bewirkt demnach den Effekt, dass die Geschwindigkeit in bebauten Gebieten auf die Hälfte reduziert wird, da sich hier der Widerstandswert verdoppelt. Eine schematische Darstellung dieser Berechnung ist in Formel 3.29⁷⁸ beschrieben.

$$\begin{aligned}
 SF &= RK_{[1-max=1]}(SDR) \\
 B &= PathDistance \left(RK_{\left[\begin{smallmatrix} 0=NoData, \\ 1-n=1 \end{smallmatrix} \right]}(BF), RK_{\left[\begin{smallmatrix} 0=NoData, \\ 1-n=1 \end{smallmatrix} \right]}(SF), Distanz_{Max=355m} \right) \\
 BR &= RK_{[NoData=1]} \left(\left(\left(Fuzzy_{\substack{B_0=0, \\ B_1=355}} = \begin{cases} 0, & B \leq 0 \\ \frac{B-B_0}{B_1-B_0}, & B_0 \leq B \leq B_1 \\ 1, & B \geq 1 \end{cases} \right) \times (-1) + 2 \right) \right)
 \end{aligned} \tag{3.29}$$

Wie in Formel 3.30 ersichtlich, wird nun mit dem Straßendummyraster SDR gemäß der Tabelle 3.8 anhand der Geschwindigkeitskonstanten GK_j je Straßentyp ein gewichteter Raster (SDR_{GK}) gebildet, welcher somit die unterschiedlichen Reisegeschwindigkeiten im Straßennetz abbildet. Die Geschwindigkeitskonstanten GK_j in Tabelle 3.8 wurden zuerst anhand der zulässigen Maximalgeschwindigkeit festgelegt, dann allerdings nach einigen Tests der Berechnungen etwas weiter gespannt, um mit diesen realistische Ergebnisse zu generieren. Mit dem maximalen Wert aus diesem Raster wird nun eine Verhältniszahl ermittelt, welche mit dem Bremsraster BR_{rein} multipliziert zum endgültigen Kostenraster KR (Formel 3.31) führt. Es entsteht somit für die Distanzberechnung eine geeignete Reisekostenwiderstandsoberfläche, welche neben den unterschiedlichen Straßentypen eben auch die oben beschriebenen „Bremswirkungen“ in bebautem Gebiet berücksichtigt. Wie bereits aus der Formel 3.31 ersichtlich ist, werden beim Bremsraster (BR_{rein}) alle SDR ausgeschlossen, welche dem Typ Autobahn angehören, da bei diesem Straßentyp keine Geschwindigkeitsreduktionen in bebautem Gebiet

⁷⁸ Vgl. bezüglich Fuzzy RIEDL 1999, S. 52.

herangezogen werden, sondern dafür auf die bereits oben beschriebene Digitalisierung zurückgegriffen wird.

$$SDR_{GK} = \sum_{\substack{j \in \{SK_{all}\} \\ GK_j \in [0,156]}} GK_j \times SDR_j \quad (3.30)$$

$$KR = \frac{Max(SDR_{GK})}{SDR_{GK}} \times \left(BR_{rein} = \begin{cases} 0, & SDR_{j \leq 104} = 1 \\ BR, & SDR_{j \leq 104} = 0 \end{cases} \right) \quad (3.31)$$

Der in Abbildung 3.18(b) visualisierte Kostenraster KR zeigt deutlich, dass höherrangige Straßen wie etwa die Autobahn im südlichen Bereich der Karte den geringsten Kostenwiderstand aufweisen und – im Gegensatz dazu – die Güterstraßen im nordwestlichen Bereich deutlich höhere Reisekosten aufweisen. Weiters ist ersichtlich, dass sich die bereits in Abbildung 3.18(a) dargestellten bebauten Gebiete ebenfalls werterhöhend auf den Kostenraster auswirken und somit alle Orte eine deutlich reduzierte Reisegeschwindigkeit bei der Simulation der Erreichbarkeit aufweisen werden. Beachtenswert ist hierbei noch, dass sich der Mittelwert über dem österreichweiten Raster bei 4,6 bewegt und dementsprechend die Legende bei dieser Karte an das interessante untere Ende der Skala verschoben wurde. Dies ist darauf zurückzuführen, dass nur wenige Rasterzellen den „langsamen“ Kategorien gemäß der Konstante GK_j in der Tabelle 3.8 zugeordnet sind.

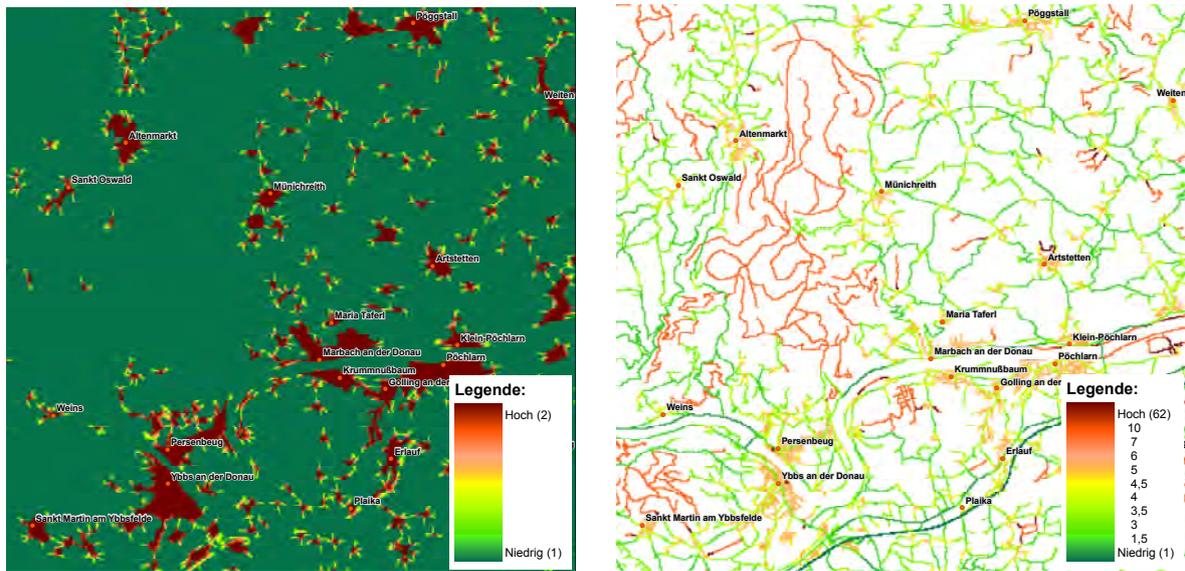
(a) Bremsraster (*BR*)(b) Reisekostenraster (*KR*)

Abbildung 3.18: Reisegeschwindigkeiten (Beispiel: Ybbs an der Donau)

3.4.4 Wahl der Zielpunkte

Bei Erreichbarkeitsmodellen gibt es in der Wissenschaft eine Vielzahl von verschiedenen Ansätzen. Einige verwenden für die Zielpunkte der Berechnungen die klassische „Theorie der zentralen Orte“⁷⁹ und die daraus resultierenden Orte mit einer hierarchischen Zuordnung. Andere versuchen unter Zuhilfenahme der Arbeitsplatzpotentiale die Zielorte für das Erreichbarkeitsmodell zu eruieren. Oftmals ist die Zieldefinition auch problemlos möglich, wenn beispielsweise die Filialen einer Firma auf ihre allgemeine Erreichbarkeit hin überprüft werden sollen. Im Allgemeinen kann man also feststellen, dass je nach Problemstellung bzw. Sinn und Zweck des Erreichbarkeitsmodells die Zielorte gewählt werden. In dieser Arbeit werden anhand von Ortsmittelpunkten (*OP*) die folgenden Zieldefinitionen⁸⁰ herangezogen, um dadurch ein differenziertes Erreichbarkeitsmodell zu erhalten. Es gibt für die zweite und dritte Klasse jeweils eine Unterkategorie, welche auch die angrenzenden ausländischen Städte miteinbezieht, um auch die Bedeutung von Grenzstädten mit dem Erreichbarkeitsmodell untersuchen zu

⁷⁹ Siehe dazu: CHRISTALLER 1980.

⁸⁰ Die komplette Liste ist im Anhang in Tabelle B.1 zu finden.

können.

- Klasse 1 ($Kl_1 = OP_{\{1-9\}}$) : Landeshauptstädte (9)
- Klasse 2 ($Kl_2 = OP_{\{2-47\}}$): Alle Wiener Bezirke und Städte mit mehr als 20.000 Einwohnern (46)
 - Klasse 2b ($Kl_{2b} = OP_{\{2-68\}}$): Klasse 2 inklusive der ausländischen Grenzstädte (67)
- Klasse 3 ($Kl_3 = OP_{\{1-47, 69-156\}}$): Landes- und Bezirkshauptstädte und Städte mit über 10.000 Einwohnern (134)
 - Klasse 3b ($Kl_{3b} = OP_{\{1-156\}}$): Klasse 3 inklusive der ausländischen Grenzstädte (156)

Für die definierten Klassen⁸¹(Kl_n) gilt grundsätzlich $Kl_1 \subset Kl_2^{82} \subset Kl_{2b} \subset Kl_3 \subset Kl_{3b}$, wonach in jeder Teilmenge jeweils die höherrangige Klasse enthalten ist und somit diese Ziele mehrmals herangezogen werden.

3.4.5 Reisekostenberechnung

Die Berechnung der Reisekostenoberfläche RKo_i wird in diesem Modell (Formel: 3.32) zunächst für alle 156 verwendeten Zielorte (ZO_i) einzeln durchgeführt. Die Durchführung erfolgt mit dem ESRI 2012 Tool: Path Distance (SA), welches bereits zuvor mehrmals zur Anwendung gekommen ist.

$$RKo_i = PathDistance(ZO_i, KR, DHM, VF_{Neigung}) \quad , i \in \{1 - 156\} \quad (3.32)$$

Das Path Distance Tool „berechnet für jede Zelle die Entfernung mit den geringsten akkumulativen Kosten zur nächstgelegenen Quelle, wobei die Oberflächenentfernung und horizontale und vertikale Kostenfaktoren berücksichtigt werden.“⁸³

Zum besseren Verständnis wird zunächst die Funktionsweise anhand der reinen Cost

⁸¹ Siehe dazu Formeln 3.39 bis 3.41 und Kapitel 3.4.6.

⁸² Ausnahme: Eisenstadt (ca. 11500 Einwohner).

⁸³ ESRI 2012, Tool: Path Distance (SA).

Distance Funktion erklärt. Die akkumulativen Kosten werden mit dem ESRI 2012 Cost Distance Werkzeug für jede einzelne Zelle ausgehend von der nächstgelegenen Quellzelle bestimmt. Der verwendete, auf der Graphentheorie basierende Algorithmus beschreibt die Mittelpunkte der Rasterzellen als Knoten, welche über Kanten mit den Nachbarzellen verbunden sind. Jede Kante ist mit der zugeordneten Kostenoberfläche verknüpft, womit es möglich wird, die Reisekosten für gewisse Zellen zu erhöhen, da die Zellgröße mit dem Kostenwert multipliziert wird. Wenn beispielsweise im Kostenraster KR eine Landstraße den Wert 2,5 hat und dieser mit der Zellgröße von 50m multipliziert wird, ergibt sich eine Reisedistanz von 125m anstelle der ursprünglichen Distanz von nur 50m. Dadurch wird der Raster quasi in einen fiktiven Graphen umgebaut und es ist dadurch möglich eben dieses Erreichbarkeitsmodell auf Rasterbasis mit sehr ähnlichen Ergebnissen, wie bei einer klassischen Netzwerkberechnung, zu erstellen. In der Abbildung 3.19 ist passend zu den Formeln⁸⁴ 3.33 bis 3.35 die Berechnung der Kostenentfernung für die benachbarten (a) und diagonalen (c) sowie jeweils die akkumulativen (b, c) Knotenkosten daraus dargestellt. Die Gesamtkosten der Kante $a1$ werden bei den direkt verbundenen Nachbarzellen (a) aus den Kosten der beiden Zellen (Startpunkt: $cost1$, Endpunkt: $cost2$) ermittelt. Um die akkumulativen Kosten (b) zu erhalten, werden die Kantenkosten ($a1, a2$), welche sich zwischen Start- ($cost1$) und Endpunkt ($cost3$) befinden, ebenfalls addiert. Bei den diagonalen Kosten (c) wird dasselbe Prinzip wie vorher verwendet, mit dem Unterschied, dass sich die Wegstrecke und somit auch die Kantenkosten ($a1_{diag}, a2_{diag}$) gemäß der Schrägdistanz der jeweiligen Rasterzellen erhöht.

⁸⁴ Formeln nach: ESRI 2012, Tool: Cost Distance (SA).

$$(a) \text{ Kosten zw. Nachbarzellen : } a1 = \frac{cost1 + cost2}{2} \quad (3.33)$$

$$(b) \text{ Akkumulative Kosten : } a2 = \frac{cost2 + cost3}{2}$$

$$AK = a1 + a2 \quad (3.34)$$

$$(c) \text{ Diagonale Kosten : } a1_{diag} = \sqrt{2} \times \left(\frac{cost1 + cost2}{2} \right)$$

$$a2_{diag} = \sqrt{2} \times \left(\frac{cost2 + cost3}{2} \right)$$

$$AK_{diag} = a1_{diag} + a2_{diag} \quad (3.35)$$

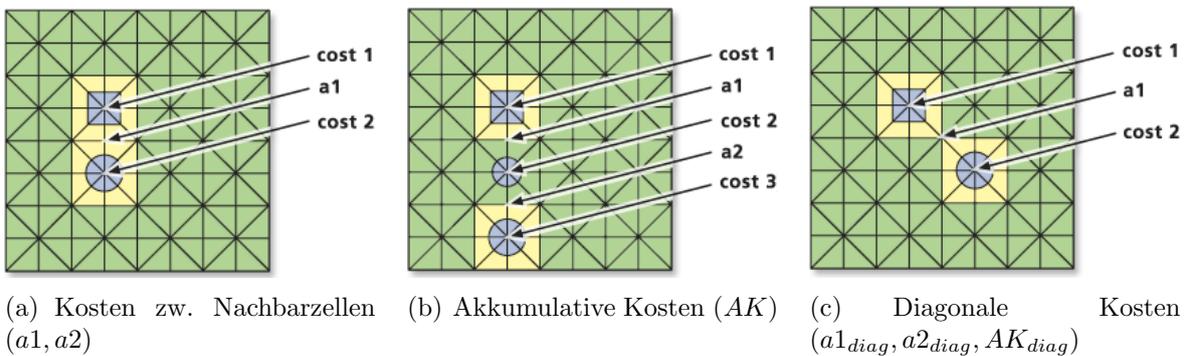


Abbildung 3.19: Cost Distance - Knotenreisekosten⁸⁵

Um mittels der oben beschriebenen Logik nun die akkumulativen Kosten für einen gesamten Raster ($\sum AK$) zu errechnen, wird ausgehend von den Quellzellen ZO_i für alle Nachbarzellen der geringste akkumulative Kostenwert ermittelt und dieser dann als Basis für den nächsten Berechnungsdurchlauf herangezogen. Wenn sich die Kosten von zwei Quellzellen überlappen, wird an diesen Rasterpositionen immer der kleinste Wert herangezogen. Diese Vorgehensweise wird in einem iterativen Prozess so lange wiederholt, bis alle der zur Verfügung stehenden Rasterzellen mit einem akkumulativen Kostenwert befüllt sind und somit die endgültige Reisekostenoberfläche (Rko) ermittelt wurde.⁸⁶

⁸⁵ Grafikquellen: ESRI 2012 Tool: Cost Distance (SA).

⁸⁶ Detaillierte Informationen dazu sind unter: ESRI 2012, Tool: Cost Distance (SA) zu finden.

Der Unterschied zwischen Path Distance (SA) und der reinen Cost Distance (SA) besteht darin, dass bei der Path Distance (SA) zusätzlich zum Kostenraster auch der Raster eines Höhenmodells zur Ermittlung der tatsächlichen Oberflächenentfernung ($DHM_{Distanz}$) herangezogen werden kann. Des Weiteren ist es möglich, einen horizontalen HF (Himmelsrichtung) bzw. vertikalen (Geländeneigung) Faktor VF zu vergeben, welche jeweils in Kombination mit einem geeigneten Kostenraster ($cost_{a,b}$) bei der Berechnung der Gesamtreisekosten einfließen. In den nachstehenden Formeln⁸⁷ 3.36 und 3.36 werden, für das bereits oben angeführte Beispiel, die Kantenkosten gemäß der Path Distance Funktion definiert. Die hier ermittelten Kostendistanzen ($cost1_{Dist}$, $cost2_{Dist}$, $AKcost_{Dist}$) ersetzen bei Verwendung der Path Distance Funktion die Kantenkosten ($cost1$, $cost2$, AK) in den oberen Formeln 3.33 und 3.34, ansonsten ist die Berechnungsmethode ident.

$$(a) \text{ Kostendistanz zw. Nachbarzellen : } cost1_{Dist} = KR \times DHM_{Distanz} \quad (3.36) \\ \times \left(\frac{cost_a \times HF_a + cost_b \times HF_b}{2} \right) \times VF$$

$$(b) \text{ Akkumulative Kostendistanz : } cost2_{Dist} = KR \times DHM_{Distanz} \\ \times \left(\frac{cost_b \times HF_b + cost_c \times HF_c}{2} \right) \times VF \\ AKcost_{Dist} = cost1_{Dist} + cost2_{Dist} \quad (3.37)$$

Im aktuellen Anwendungsfall wird neben der tatsächlichen Oberflächenentfernung (DHM) der Vertikale Faktor $VF_{Neigung}$, mittels dem in Abbildung 3.20 dargestellten Verlauf mit Bezug auf die aus dem Höhenraster ermittelte Neigung (vertikaler relativer Bewegungswinkel in Abbildung), für jede Rasterzelle gebildet. Somit ist gewährleistet, dass sich bei einer Steigung die Reisekosten erhöhen und somit die Geschwindigkeit an diesen Rasterzellen verringert wird. Bei einem Gefälle hingegen wird die Geschwindigkeit in diesem Modell nicht erhöht, da immer nur die dem Straßentypen entsprechende Höchstgeschwindigkeit zugelassen ist. Würde man beispielsweise den Benzinverbrauch modellieren, könnte man die Reisekosten bergab verringern.

⁸⁷ Formeln nach: ESRI 2012, Tool: Path Distance (SA).

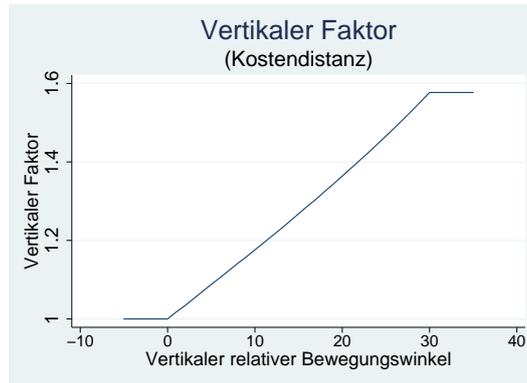


Abbildung 3.20: Vertikaler Faktor ($VF_{Neigung}$)

Exkurs Einbahnstraßen: Grundsätzlich könnte man einen Raster erzeugen, welcher für alle Zellen, die Einbahnstraßen sind, die Richtung des Straßenverlaufes als Gradwerte der Himmelsrichtung und alle anderen Straßen als Ebene, also mit „-1“ definieren. Für diesen Raster würde man dann den horizontalen Faktor bei allen horizontalen relativen Bewegungswinkeln die größer als 0 sind auf unendlich setzen und somit wäre gewährleistet, dass die Einbahnen nur in die gewünschte Richtung befahren werden, da ansonsten der Kostenwert für das Befahren dieser Zellen in abweichende Richtungen unendlich hoch wäre.⁸⁸

3.4.5.1 Validierung der ermittelten Reisekosten

Theoretisch führt die oben beschriebene Ermittlung der Reisekosten RK_{o_i} mit der Rastertechnik zu einem hinreichend genauen Ergebnis. Um dieses Ergebnis mit einem herkömmlichen Routing vergleichen zu können, wurde in einem Vergleichsbeispiel ausgehend vom Zielort (ZO_L) Linz, für 122 zufällig gewählte Ortszentren (OZ_i) im geographischen Einzugsbereich, mittels GoogleMaps-Routing,⁸⁹ für die schnellste Route jeweils die Fahrzeit und die Distanz ermittelt. Damit es möglich ist, die Kilometer und Zeitangaben von GoogleMaps mit den Reisekosten aus dem Erreichbarkeitsmodell vergleichen zu können, wurde wie in Formel 3.38 dargestellt, die Reisedistanz (RD) aus dem Reisekostenraster von Linz (RK_{o_L}) zurückgerechnet. Ausgehend von den ver-

⁸⁸ Weitere Informationen dazu sind unter ESRI 2012 Tool: Path Distance (SA) zu finden.

⁸⁹ Siehe: GOOGLE MAPS 2012.

wendeten Ortszentren (OZ_i) werden anhand der Reisekosten für Linz (RK_{o_L}) und dem zugehörigen Knotenrückverknüpfungsraster (KnR_L) – welcher „die Richtung definiert, oder die nächste angrenzende Zelle (die nachfolgende Zelle) entlang dem kleinsten akkumulativen Kostenpfad von einer Zelle zu ihrer kostengünstigsten Quelle identifiziert“⁹⁰ und ein Nebenprodukt der Kostenberechnung ist – unter Zuhilfenahme des ESRI 2012 Cost Path (SA) Werkzeuges, zunächst die kostengünstigsten Routen (CP_L) von und zum Zielort Linz (ZO_L) ermittelt. Im nächsten Schritt wird dann erneut die PathDistance ausgehend vom Zielort entlang der ermittelten schnellsten Route, welche reklassifiziert wird, berechnet und man erhält somit die Reisedistanz der schnellsten Route.

$$\begin{aligned}
 CP_L &= CostPath(OZ_i, RK_{o_L}, KnR_L) && , i \in \{1 - 122\} \\
 RD &= PathDistance(ZO_L, RK_{[1-122=NoData]}(CP_L)) && (3.38)
 \end{aligned}$$

In Abbildung 3.21 sind für die beispielhaft gewählten Ortszentren (OZ_i) die vom Erreichbarkeitsmodell verwendeten Routen sowie die Entfernung in Kilometer visualisiert. Des Weiteren kann man im Hintergrund anhand der Farbskala die entsprechenden Reisekosten aus dem Modell entlang aller Routen ablesen. Betrachtet man die gewählten Routen etwas genauer, kann man feststellen, dass bei längeren Routen wie etwa nach Wels (38,7 km) der Routenverlauf gleich wie bei einem herkömmlichen Routing mit beispielsweise GoogleMaps⁹¹ über die höherrangigen Straßen – in diesem Fall die Autobahnen A7, A1 und A25 – ist, und die ermittelten Reisekosten entsprechend realistisch sind.

⁹⁰ ESRI 2012, Tool: Cost Path (SA).

⁹¹ Siehe online dazu: GOOGLE MAPS 2012.

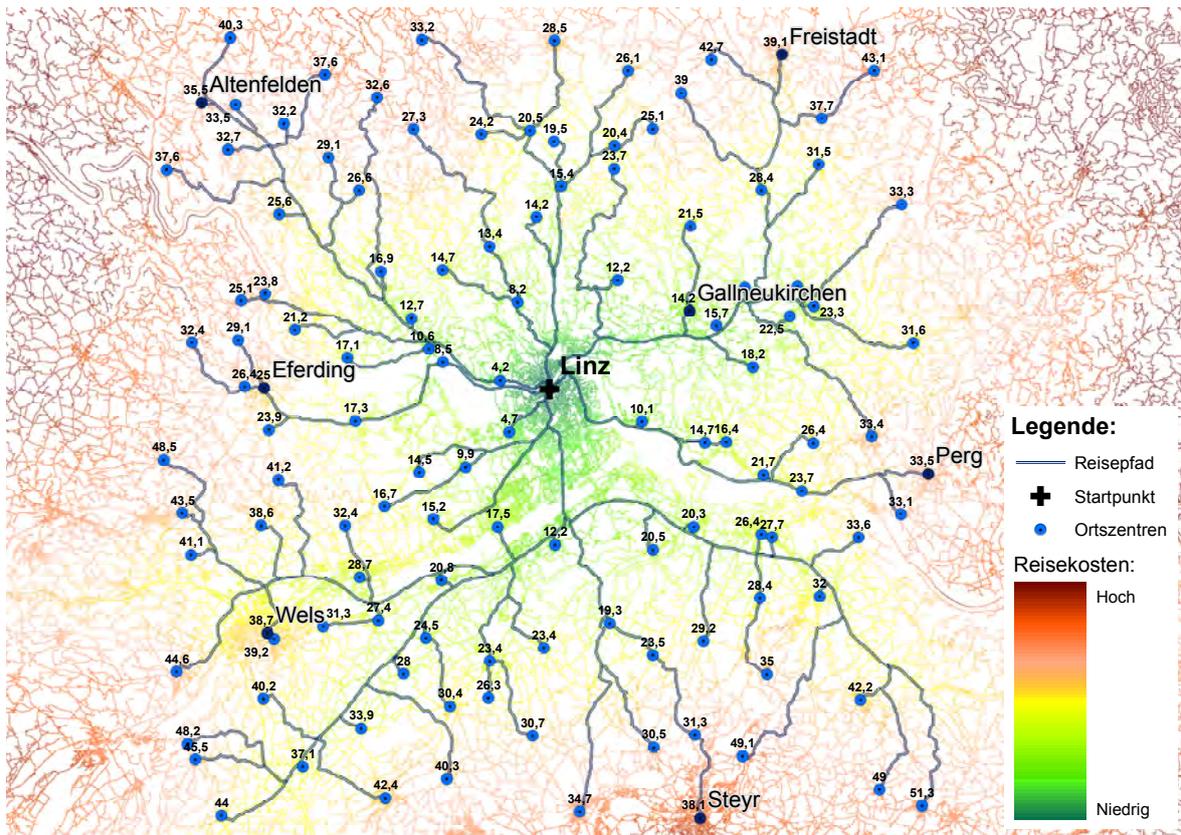


Abbildung 3.21: Reisepfad, Reisedistanz und Reisekosten am Beispiel Linz

Um neben der visuellen Qualitätsprüfung der Routenwahl im Erreichbarkeitsmodell auch eine quantitative Aussage treffen zu können, wurden in der folgenden Grafik 3.22 die Reisekosten RKo aus dem Modell mit der manuell erfassten Reisezeit von Google Maps Routing anhand des oben beschriebenen Beispiels gegenübergestellt. Um die Reisekosten mit der Reisezeit auf der gleichen Skala darstellen zu können, wurde diese durch den jeweiligen Mittelwert dividiert. Wie man deutlich erkennen kann, liegen die Kurven als auch die 95% Konfidenzintervalle fast in der ganzen Grafik direkt übereinander. Daraus kann man folgern, dass die mit dem Rastermodell ermittelten Reisekosten von gleicher Qualität wie die Fahrzeiten mittels konventioneller Routingsysteme wie GoogleMaps sind. Die kleine Abweichung der Kurve des Erreichbarkeitsmodells im oberen Bereich der Skala kann außer Acht gelassen werden, da hier, wie man in der Grafik erkennen kann, nur sehr wenige Daten vorhanden sind.

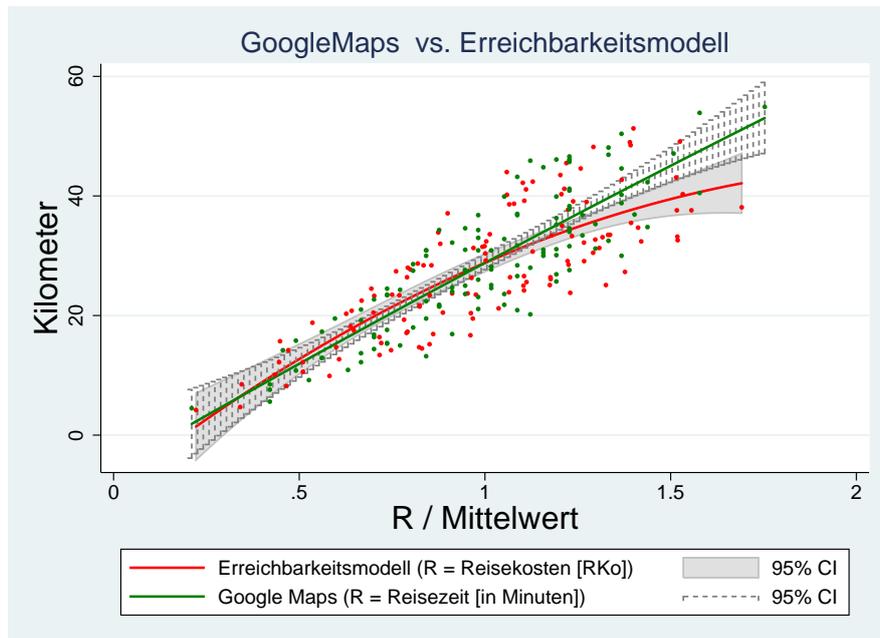


Abbildung 3.22: Vergleiche GoogleMaps-Routing mit Erreichbarkeitsmodell

Eine erwähnenswerte Einschränkung des Rastermodells ist allerdings, dass bei der Routenermittlung an jeder Stelle im Raster auf einen angrenzenden Straßentyp gewechselt werden kann und somit beispielsweise überall, wo die Autobahn eine andere Straße kreuzt, die Route sich auf die Autobahn verlagern kann, obwohl dies verkehrstechnisch eigentlich nicht möglich wäre. Global betrachtet ist dieser „Modellfehler“ allerdings, wie schon die obige Grafik 3.22 zeigt, irrelevant, da es sich nur um abweichende Zufahrtswege zu Autobahnen handelt. Weiters führt diese Routenverschiebung in den meisten Fällen nur zu sehr geringen Unterschieden in der Gesamtroute, da Autobahnen zumeist nur von anderen hochrangigen Straßen, welche zumeist eine Auffahrt haben, gequert werden und das reine Auffahrtsproblem bei 50m Rasterzellen sich ohnedies nur auf einige wenige Zellen beschränkt, wenn man etwa eine „Schleife“ zur Auffahrt fahren muss.

3.4.6 Erreichbarkeit

Da mit diesem Erreichbarkeitsmodell die allgemeine Erreichbarkeit für alle österreichischen Standorte definiert werden soll, werden nun die einzelnen Reisekosten Rko den in Kapitel 3.4.4 definierten Klassen Kl_n zugeordnet. Für jede Klasse wird die tatsächliche Erreichbarkeit aus den einzelnen Reisekosten der zuvor definierten Zielpunkte gemäß der in den Formeln 3.39 bis 3.41 dargestellten Logik abgeleitet. Hierbei wird in einem ersten Schritt (Formel 3.39) für alle zuvor ermittelten Reisekostenraster (RKo_i), welche sich innerhalb der jeweiligen Klasse befinden, das Mini- und Maximum je Rasterzelle ermittelt. Das erste Minimum $Min_{1,j}$ ist somit die Entfernung zur nächstgelegenen Stadt innerhalb der jeweiligen Klasse Kl_n (Beispiel: Kl_1 – Entfernung zur nächstgelegenen Landeshauptstadt.). Um zusätzlich zu dieser (ersten) Entfernung auch weitere Zentralitätsbestimmungen für Standorte festlegen zu können, wird immer auch die zweit- ($Min_{2,j}$) und drittnächstliegende ($Min_{3,j}$) Stadt innerhalb der Klasse ermittelt. Hierfür wird, wie in den Formeln 3.40 und 3.41 dargestellt, jeweils eine neue Berechnungsbasis der Reisekosten ($RKo_{Neu,j}$) ermittelt, in welcher die nächstgelegene Stadt durch den Maximalwert ersetzt wird und somit von der Neuermittlung des Minimums ($Min_{2,j}, Min_{3,j}$) ausgeschlossen ist. Als Ergebnis dieser Berechnung erhält man somit drei Rasterdatenschichten mit den nächstgelegenen Städten für jeden Standort und dies wiederum einzeln für jede der fünf Klassen Kl_n gemäß der Zielpunktdefinition.

Für $j \in \{Kl_1, Kl_2, Kl_{2b}, Kl_3, Kl_{3b}\}$ und für $j_i \in \{Kl_j\}$ gilt :

$$Max_j = Maximum(RKo_{j_i})$$

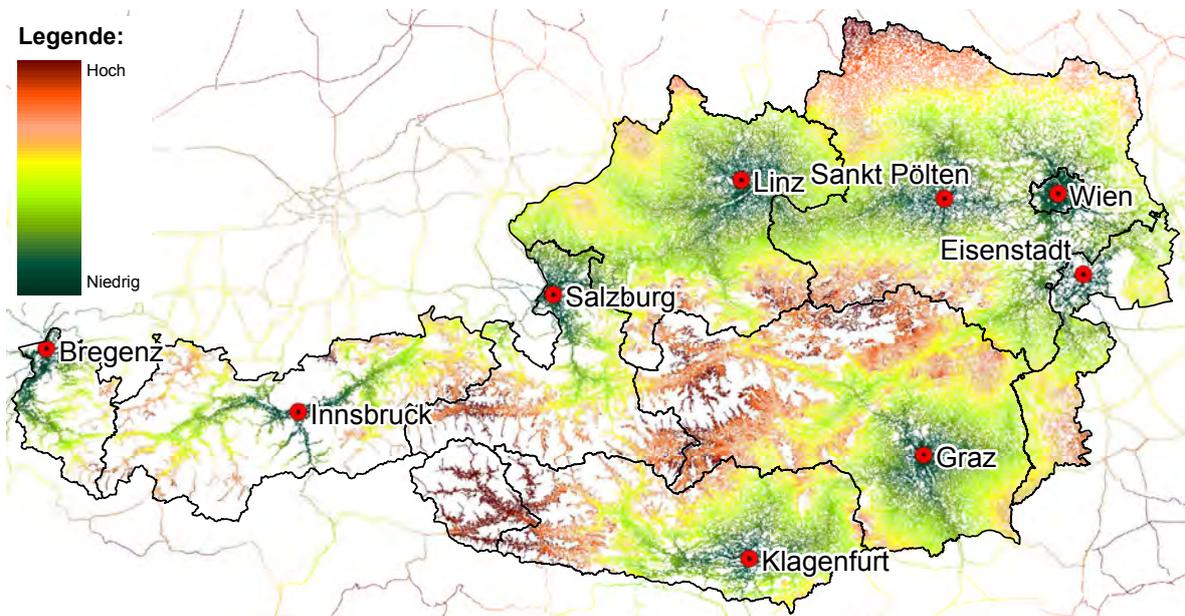
$$Min_{1,j} = Minimum(G_{1,j_i} = (RKo_{j_i})) \quad (3.39)$$

$$Min_{2,j} = Minimum \left(G_{2,j_i} = \left(RKo_{Neu2,j_i} = \begin{cases} RKo_{j_i}, & RKo_{j_i} > Min_{1,j} \\ Max_j, & RKo_{j_i} = Min_{1,j} \end{cases} \right) \right) \quad (3.40)$$

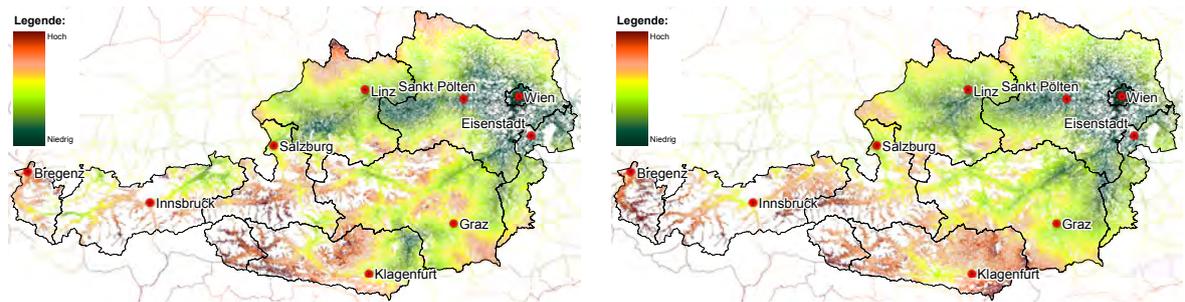
$$G_{3,j_i} = \left(RKo_{Neu3,j_i} = \begin{cases} RKo_{Neu2,j_i}, & RKo_{Neu2,j_i} > Min_{2,j} \\ Max_j, & RKo_{Neu2,j_i} = Min_{2,j} \end{cases} \right)$$

$$Min_{3,j} = Minimum(G_{3,j_i}) \quad (3.41)$$

In der nachstehenden Abbildung 3.23 sind die drei ermittelten Minimalentfernungen zu den Landeshauptstädten visualisiert. In der Abbildung 3.23(a) kann man die Einzugsgebiete der jeweiligen Städte sowie die peripheren, schlecht erreichbaren Regionen – wie etwa das nördliche Waldviertel, Osttirol und Teile der Zentralalpen – deutlich erkennen. Die hier abgebildeten Reisekosten geben demnach an jedem Standort die Auskunft, wie weit es bis zum nächstgelegenen Zentrum ist. Auf eine Gewichtung der Zentren untereinander wurde bewusst verzichtet, da mit dem Modell die Zentrumsnähe und nicht deren Ausstattung bewertet werden soll. Betrachtet man unter dieser Prämisse nun die Abbildung 3.23(b), kann man feststellen, dass – logischerweise – dort, wo die definierten Zentren näher beieinanderliegen, niedrigere Reisekosten gegeben sind, als an Standorten mit nur wenigen Zentren. In der Abbildung 3.23(c) ist die Entfernung zum drittnächsten Zentrumspunkt visualisiert. Hier kann man erkennen, dass sich – gemäß der geographischen Verteilung der Landeshauptstädte in Österreich – nun geringe Reisekosten nur mehr in den nordöstlichen Wirtschaftsregionen außerhalb der Zentralalpen wiederfinden. Man kann also feststellen, dass neben der direkten Distanz zu einem Zentrum mit dem zweit- und drittnächstgelegenen Zentralpunkt eine Art „Agglomerationsmesszahl“ für die Klassen (Kl_n) definiert wird.



(a) Entfernung zur nächstgelegenen Stadt (Min_{1,Kl_1})



(b) Entfernung zur 2. nächsten Stadt (Min_{2,Kl_1}) (c) Entfernung zur 3. nächsten Stadt (Min_{3,Kl_1})

Abbildung 3.23: Reisekosten zu den Landeshauptstädten (Kl_1)

Wie man in der oben stehenden Abbildung 3.23 erkennen kann, wurden nur für Rasterzellen, die entlang der Straßen liegen, die Reisekosten errechnet. Um auch für alle restlichen Gebiete Reisekosten zu erhalten, werden wie in Formel 3.42 dargestellt, mittels ESRI 2012 Tool: Path Distance Allocation (SA), unter Zuhilfenahme des Höhenrasters die vorhandenen Reisekosten entlang der Straßen auf alle übrigen Flächen extrapoliert. Die Ermittlung erfolgt unter Zuhilfenahme des vertikalen Faktors ($VF_{Neigung}$)⁹², der hier als eine inverse lineare Funktion, die ausgehend vom vertikalen relativen Bewegungswinkel eine symmetrische Form aufweist,⁹³ verwendet wird. Somit ist es möglich,

⁹² Weitere Informationen dazu sind in Kapitel 3.4.5 und in Abbildung 3.20 zu finden.

⁹³ Vgl. dazu ESRI 2012, Tool: Path Distance Allocation (SA).

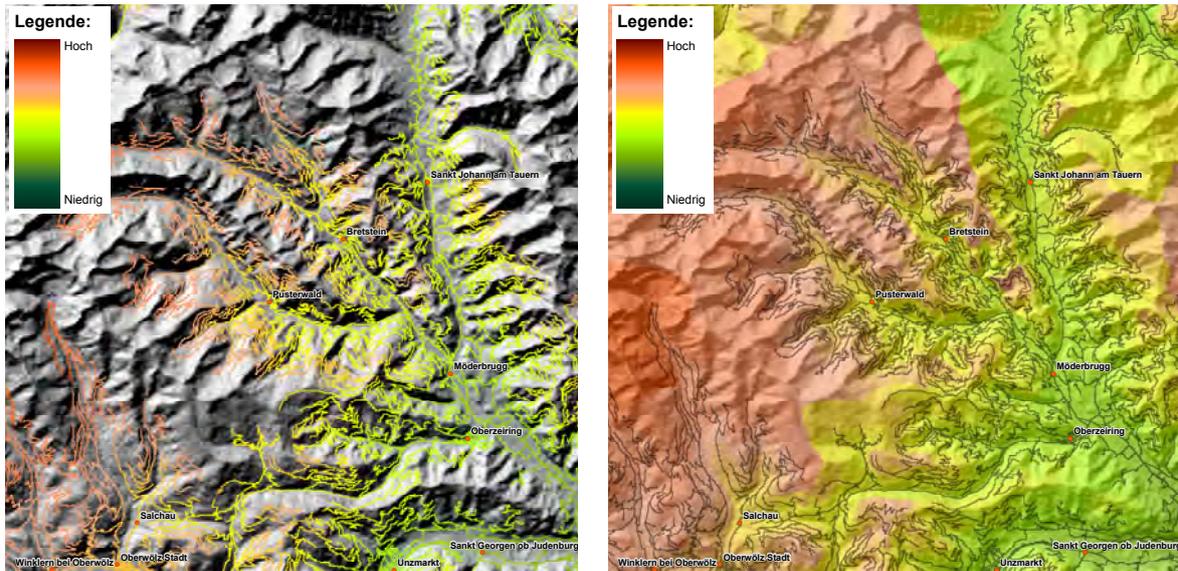
dass z.B. auch Neubaugebiete, die noch über keine Erschließung mit Straßen verfügen, trotzdem mit einem Wert für Reisekosten versehen werden können.

$$VF_{Neigung} = SymInverseLinear \left(1, -30, 30, \frac{1}{30} \right)$$

$$AkRK_{o_{m_j}} = PathAllocation (Min_{m_j}, DHM, VF_{Neigung}) \quad , m \in \{1 - 3\} \quad (3.42)$$

In der Abbildung 3.24(a) sind exemplarisch die minimalen Reisekosten ($Min_{m,j}$), welche entlang der Straße ermittelt wurden, zu erkennen. Demgegenüber steht in Abbildung 3.24(b) die Visualisierung der daraus extrapolierten Reisekosten ($AkRK_{o_{m,j}}$) auf die gesamte Fläche. Gut zu erkennen ist der Einfluss des vertikalen Faktors und des Höhenmodells (DHM) auf die Zonengrenzen, welche sich zumeist entlang der Gebirgsgrate oder an sonstigen Erhebungen orientieren und somit die Reisekosten auch in der flächigen Ausbreitung gut zugeordnet werden. Bei diesem Verfahren zur Flächenextrapolation mittels Allokation wird der Rasterwert der Reisekosten mit zunehmender Entfernung von der Straße nicht erhöht. Anstelle dessen wird für jede leere Zelle der nächstgelegene Quellzellwert (Reisekosten des Erreichbarkeitsmodells ($Min_{m,j}$)) mit den geringsten akkumulativen Kosten ermittelt und dieser dann auf die leere Zelle übertragen.⁹⁴

⁹⁴ Weitere Informationen dazu: ESRI 2012, Tool: Path Distance Allocation und Cost Allocation (SA).



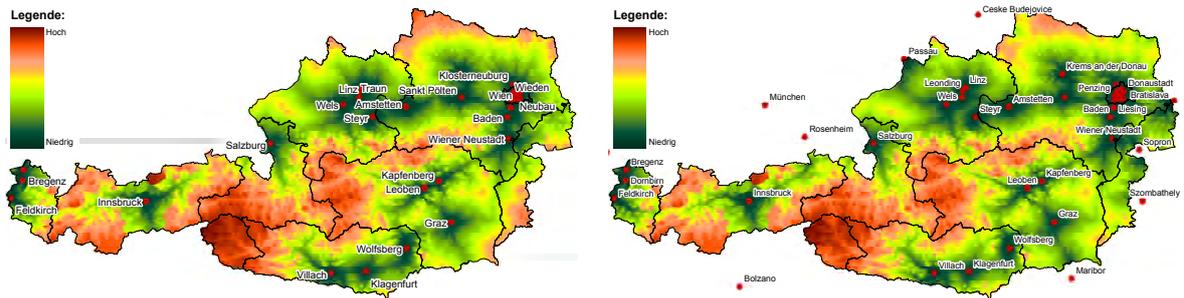
(a) Reisekosten an der Straßen (Min_1, Kl_2)

(b) Extrapolierte Reisekosten ($AkRKo_1, Kl_2$)

Abbildung 3.24: Allokation der Reisekosten (Beispiel: nordwestliche Steiermark)

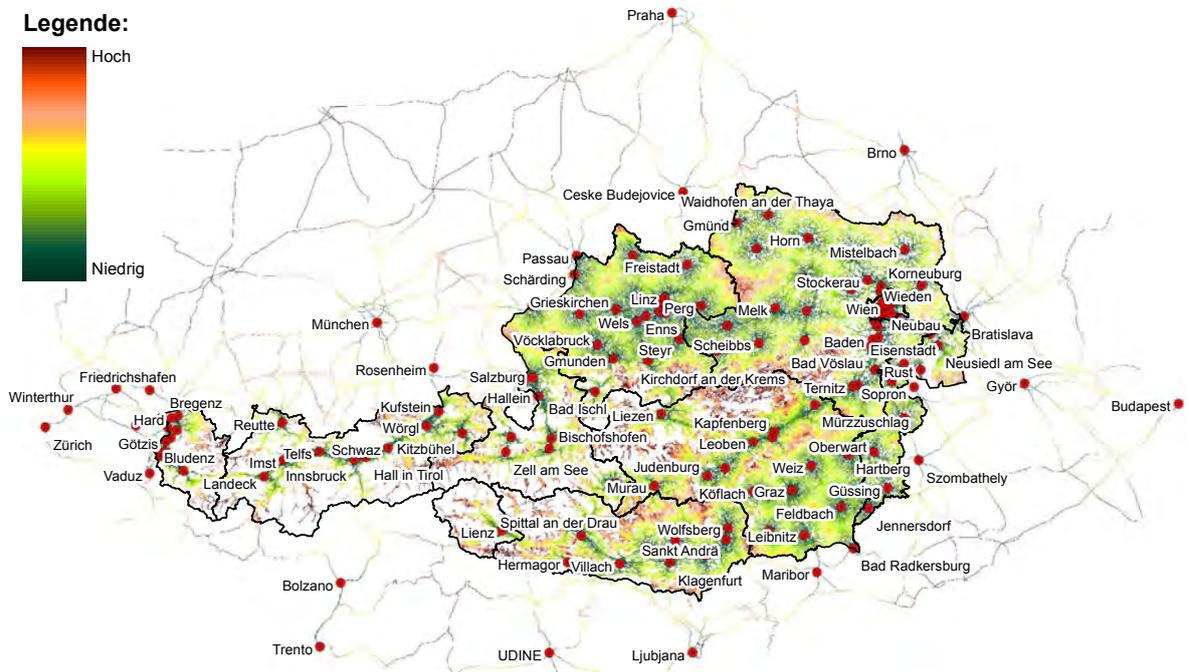
In der Abbildung 3.25 sind die Zielpunkte mit den nächstgelegenen Zentren (Min_1) für die Klasse 2 und 3 dargestellt. Bei der dritten Klasse (Kl_{3b}) sind die bereits in Kapitel 3.4.1 beschriebenen zusätzlich erfassten Daten (außerhalb von Österreich) erkennbar. Bei der Berechnung fließen die Straßen außerhalb von Österreich bei allen Klassen ein. Die ausländischen Zielorte finden allerdings nur bei den Klassen 2b und 3b Verwendung. Auswirkungen der ausländischen Zielorte auf die Reisekosten innerhalb von Österreich sind bei Klasse 3b mit Ausnahme von Bratislava und Passau kaum vorhanden. Bei Klasse 2 stellt sich dies etwas anders dar, hier wirken die Zielorte aufgrund der sehr eingeschränkten Anzahl der Zielpunkte auch deutlich auf österreichischem Gebiet, dementsprechend wird diese Klasse vermutlich bei der weiteren Verwendung der Daten im hedonischen Modell keine oder eine nur sehr eingeschränkte Verwendung⁹⁵ finden.

⁹⁵ Weitere Informationen dazu sind im Kapitel 4.3 zu finden.



(a) Reisekosten zu Städten ($AkRK_{01,Kl_2}$)

(b) Reisekosten zu Städten (inkl. Ausland) ($AkRK_{01,Kl_{2b}}$)



(c) Reisekosten zu Städten > 10T Einwohner (inkl. Ausland) (Kl_{3b})

Abbildung 3.25: Entfernung zu den nächstgelegenen Zentren (Min_1)

Um die bereits oben angesprochenen Auswirkungen der ausländischen Zielorte auf die Erreichbarkeit in Österreich bei den Klassen 2b und 3b darstellen zu können, werden gemäß der Formel 3.43 analog zur bereits beschriebenen Flächenextrapolationsmethode auch die Einzugsgebiete ($AkEZG_{m,j}$) der definierten Zielorte ermittelt. Unter Einzugsgebieten wird in diesem Zusammenhang für jede Rasterzelle die kostengünstigste Reiseroute zum nächsten Zentrum verstanden, und es ist demnach jeder Rasterzelle immer nur ein Zentrum zugewiesen. Um auch Aussagen über die Nähe weiterer Zentren treffen zu können, wurden, wie bereits oben beschrieben, auch die Distanzen zum

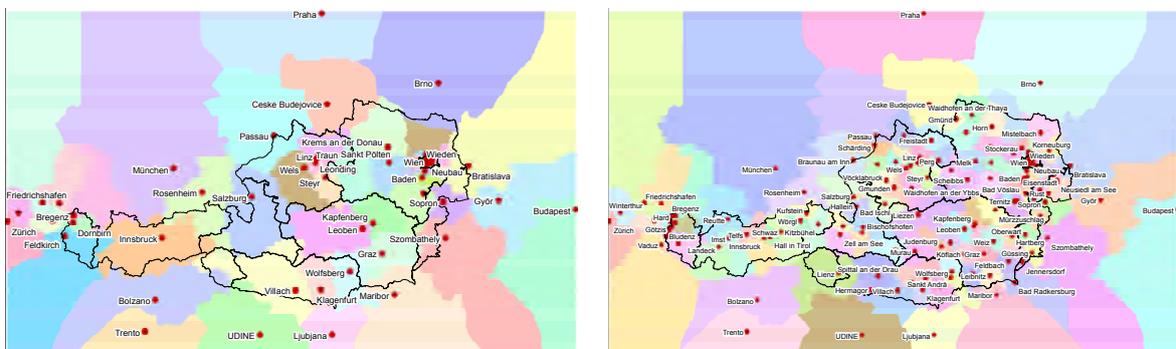
zweit- und drittnächsten Zentrum ermittelt. Die Basisdaten ($G_{m,j}$) für die Berechnung der jeweiligen Einzugsgebiete stammen aus den Formeln 3.39 bis 3.41 und sind somit ein Nebenprodukt der Erreichbarkeitsberechnung. In der Programmierung wird mittels Reklassifizierung die ID der Zielorte in den Raster geschrieben, dadurch erhält man für alle Klassen und nächstgelegenen Punkte das jeweilige Einzugsgebiet des definierten Zielortes.

Für $m \in \{1 - 3\}$ gilt :

$$EZG_{m_j} = RK_{m_j[0=NoData]} \left(RK_{m_j \begin{bmatrix} 0=OrtID, \\ 0-1=0, \\ NoData=0 \end{bmatrix}} (G_{m_j}) \right)$$

$$AkEZG_{m_j} = PathAllocation (EZG_{m_j}, DHM, VF_{Neigung}) \quad (3.43)$$

In Abbildung 3.26 sind für die Klassen 2b und 3b die Einzugsgebiete visualisiert. Wie man erkennen kann, ist auf der linken Seite (a) ein deutlicher Einfluss der Städte in Nachbarländern wie etwa von Brno, Bratislava, Maribor, Rosenheim etc. bis weit in das Landesinnere von Österreich zu erkennen. Dies ist darauf zurückzuführen, dass die Klasse 2b weniger definierte Zielpunkte als die Klasse 3 aufweist, bei der nur ein geringer Einfluss von ausländischen Orten erkennbar ist. Inwieweit auch das Erreichbarkeitsmodell mit ausländischen Zielorten Verwendung im hedonischen Preismodell findet, bzw. ob dies überhaupt zu sinnvollen Ergebnissen führt, wird in Kapitel 4.3 beschrieben.



(a) Einzugsgebiete ($AkEZG_{1,Kl_{2b}}$)

(b) Einzugsgebiete ($AkEZG_{1,Kl_{3b}}$)

Abbildung 3.26: Einzugsgebiete der definierten Zentren

3.5 Sonstige Variablen

In diesem Kapitel werden alle weiteren Daten, die von externen Datenanbietern für das Modell herangezogen werden, angeführt und kurz beschrieben. Eine komplette Aufstellung der Variablen aus den nachstehend angeführten Datensätzen ist im Anhang B.3 in der Tabelle B.2 zu finden. Dort sind auch die jeweiligen Eigenschaften dieser Variablen detailliert erläutert. Aus den externen Daten werden für das Modell noch unterschiedliche neue Variablen (beispielsweise wird aus der Anzahl der Ausländer ein Anteil an der Bevölkerung ermittelt) errechnet. Diese sind ebenfalls in der bereits oben angeführten Tabelle beschrieben.

3.5.1 Regionalstatistischer Raster

Von Statistik Austria sind zurück bis in das Jahr 2001 diverse regionalstatistische Daten auch auf Rasterbasis verfügbar. Es sind bei diesen Daten verschiedene Rasterweiten zwischen 100 und 10.000m verfügbar. Für diese Arbeit wird der 250m Raster verwendet, da aufgrund von Datenschutzgründen nur dann alle Merkmale in allen Rasterzellen enthalten sind, wenn eine gewisse Anzahl an Merkmalsausprägungen (Personen/Gebäuden/Wohnungen) in der jeweiligen Zelle vorhanden ist. Bei den älteren Daten bis zum 31.10.2011 ist diese Grenze mit 31 Personen bzw. 4 Gebäuden/Wohnungen festgelegt, ab diesem Datum wurde diese Grenzen aufgrund eines neuen Nymisierungsverfahrens der Statistik Austria bei Personen auf 4 und bei Gebäuden/Wohnungen auf 3 abgesenkt. In der nachstehenden Tabelle 3.9 ist beispielhaft für die Einwohner dargestellt, dass bei den älteren Daten nur rund 23% der Rasterzellen, was 81% der Einwohner entspricht, über genaue Merkmalsausprägungen verfügen. Bei den neuen Daten von Statistik Austria ist aufgrund der bereits erwähnten Änderung der Grenzen eine durchaus gute Abdeckung des ganzen Landes erreicht worden.

	2011	2012
Anzahl Einwohner	8.384.973	8.435.035
Einwohner mit Merkmalen	6.786.278	8.341.503
Anteile der Einwohner mit Merkmalen	80,9%	98,9%
Anzahl Rasterzellen	236.054	236.423
Rasterzellen mit Merkmalen	54.781	191.327
Anteile der Rasterzellen mit Merkmalen	23,2%	80,9%

Tabelle 3.9: Merkmale bei Einwohnerzahlen im Regionalstatistischen Raster

Da es für die Verwendung der Daten in dieser Arbeit dennoch unerlässlich ist, dass alle Merkmale des Statistikerasters vorhanden sind, um diese im hedonischen Preismodell verwenden zu können, werden jene Rasterzellen, die über keine genauen Inhalte bei den Merkmalen verfügen, mittels dem nachfolgend beschriebenen Verfahren komplettiert. Zunächst werden, wie in Formel 3.44 beispielhaft für Einwohner 2012 dargestellt, alle Zellen, die über der jeweiligen Datenschutzgrenze liegen, selektiert und dann für diese, wie bereits zuvor im Kapitel 3.4.6 erläutert, eine Path Distance Allokation berechnet. Basierend auf dieser Zonierung werden mittels Combine⁹⁶ die ursprünglichen Raster (UR) mit den distanzbasierten Allokationszonen (PA) kombiniert, wodurch man eine Zuordnungstabelle für die fehlenden Informationen erhält. Mithilfe dieser Tabelle kann man anhand der Einwohnerzahlen die Werte der nächstgelegenen Zelle übernehmen, diese anteilig auf die vorhandene Bevölkerung verteilen und somit die Informationen über alle Merkmale auch in diesen Bereichen erhalten.

$$\begin{aligned}
 PA &= PathAllocation(Selekt(Einwohner \geq 4), DHM, VF_{Neigung}) \\
 ZoT &= Combine(UR, PA)
 \end{aligned}
 \tag{3.44}$$

Diese Methode wurde gewählt, da dadurch die vorherrschende räumliche Struktur, nicht wie etwa bei einer möglichen zonalen Mittelung (z.B.: über Zählsprengel) der Werte, grundsätzlich nicht verändert, sondern an den Grenzen fortgeschrieben wird. Die Summe der Merkmale (Bevölkerung, Gebäude, etc.) wird mit diesem Verfahren nicht angetastet, man hat allerdings dadurch im Datensatz halbe Merkmale, was etwa bei den Einwohnern etwas unschön erscheint, aber für die weiteren Analysen kein Problem

⁹⁶ Siehe dazu: ESRI 2012, Tool: Combine (SA).

darstellt.

3.5.2 Lebensphasen

Der Datenproduzent vom Mikroraster Lebensphasen ist die ProfileAddress Direktmarketing GmbH, welcher sich ebenfalls auf die 250m Rasterkacheln von Statistik Austria bezieht. Datenstand ist 2012. Es sind in den Rasterzellen folgende Inhalte zu finden:

- Anteil von Haushaltstypen (Singlehaushalt, Wohngemeinschaften, Familien nach Größe, etc.)
- Akademikeranteil
- Anteil der Entscheider (Entscheidungsträger in Großunternehmen, Führungskräfte, Vorstandsmitglieder, etc.)

Für die gegenständliche Arbeit ist der Anteil der Akademiker bzw. Entscheider von Interesse, von den übrigen Eigenschaften wird nur ein geringer Einfluss erwartet.

3.5.3 Gebäudedaten

Dieser Datensatz basiert ebenfalls auf dem 250m Raster von Statistik Austria und wurde vom Bundesamt für Eich- und Vermessungswesen (BEV) in Kooperation mit Geomarketing GmbH erstellt. Es ist darin die Anzahl der Gebäude nach Typ (Büro, Handel, Industrie, Wohnen, etc.) pro Rasterzelle zu finden.

3.5.4 Klimadaten

Die Klimadaten stammen von der Zentralanstalt für Meteorologie und Geodynamik (ZAMG) und beziehen sich auf das langjährige Mittel aus den Jahren 1961-1990. Diese Daten liegen als Raster mit 250m Maschenweite vor und mussten bei der Aufbereitung lediglich in die korrekte Projektion gebracht werden. Weitere Informationen, sowie die Variablenbezeichnung im Modell sind der bereits erwähnten Tabelle B.2 im Anhang zu entnehmen.

3.6 Verknüpfung der Variablen mit den Immobiliendaten

Alle im Kapitel 3 beschriebenen Variablen liegen nun in einer einheitlichen Projektion (Lambert Austria) vor und sind nach Themengebieten getrennt in File Geodatabases (FGDB) abgespeichert. Um diese nun in der Statistiksoftware R verwenden zu können, wird für alle Rastervariablen der jeweilige Rasterwert anhand der Koordinaten aus den Immobiliendaten extrahiert und kann somit den entsprechenden Kaufpreisinformationen zugewiesen werden. Bei den Rasterdaten werden in der Analyse die konkreten Zellwerte herangezogen, um die Auswirkungen auf den Immobilienpreis zu untersuchen. Hierbei handelt es sich immer um kontinuierliche Phänomene wie etwa Lärm, Hangneigung oder Erreichbarkeit.

Bei der Nutzungsdatenschicht⁹⁷ die im Vektorformat vorliegt, werden zunächst ebenfalls die Koordinaten der Immobilie herangezogen, um den exakten Standort zu lokalisieren. Ausgehend von diesen Standorten werden, wie in Abbildung 3.27 dargestellt, allerdings 3 Pufferringe mit 50, 200 und 500m Radius erzeugt, welche dann jeweils einzeln mit der Nutzungsdatenschicht verschnitten werden. Durch diese Verschneidung kann man ermitteln, wie groß die Flächenanteile der jeweiligen NK's je Pufferfläche sind und somit eine Aussage über die räumliche Strukturierung des direkten Umfeldes treffen.

⁹⁷ Siehe dazu Kapitel 3.1.



Abbildung 3.27: Verknüpfung mit NK's (Beispiel: Graz Nord)

In der Variablenbeschreibung (Tabelle B.2) sind nur speziell behandelte Variablen (beispielsweise wenn die Anteile kategorisiert wurden oder zusätzliche Gruppenvariablen erzeugt wurden) aus der Nutzungsdatenschicht angeführt. Die Namen bzw. Zuordnungen der Standardvariablen aus dieser können allerdings gemäß der folgenden Logik unter Zuhilfenahme der NK's in Tabelle 3.2 nachvollzogen werden. Für alle 3 Pufferringe wird je NK der Anteil der darin vorhanden Flächen ermittelt. Der Variablenname setzt sich dann aus „nu_area“, der Pufferdistanz dem NK und „-ant“ zusammen (Beispiel für 500m Puffer mit NK 10902 (Zentrumsplätze): „nu_area_500_10902_ant“). Für die Gesamtanteile der Obergruppen wird im Prinzip gleich vorgegangen, wobei hier noch „og“ im Namen eingefügt wird (Beispiel für OG 109 (Plätze): „nu_area_500_og_109“). Da bei einigen Variablen aus der Nutzungsdatenschicht die Anteile nicht gut interpretierbar sind, gibt es für diese Fälle auch eine dummycodierte Version dieser Variablen. In diesem Fall wird dem Namen „-dum“ hintangestellt (Beispiel: „nu_area_500_10902_ant_dum“).

Kapitel 4

Validierung der Mikrolagevariablen mit einem Generalisierten Additiven Modell (GAM)

Wie man bereits aus dem vorherigen Kapitel 3 schließen kann, ist neben den Immobiliendaten (Objekteigenschaften) eine Vielzahl an neuen Lagevariablen vorhanden. Um nun alle Variablen aus diesem umfangreichen Lagedatenpool analysieren zu können, wird zunächst ein reduziertes hedonisches Preismodell (Basismodell) mit Objektkriterien und zwei Indexvariablen aufgestellt.

Ausgehend von diesem Basismodell werden sodann alle im Rahmen dieser Arbeit erzeugten Lagevariablen durchgetestet. Ziel ist es herauszufinden, welche dieser Lagevariablen eine Immobilienpreisrelevanz aufweisen und wie hoch die Ausprägung dieser ist. Es wird insbesondere darauf geachtet, dass die Variablen auch bei verschiedensten Modellkonstellationen immer wieder zu gleichen Ergebnissen führen, um am Schluss der Analyse die validen Variablen identifizieren zu können.

4.1 Basismodell

Beim Aufstellen des Basismodells wurde eine möglichst schlanke Modellstruktur generiert, um für die folgenden Tests mit den rasterbasierten Lagevariablen eine stabile Modellgrundlage zu haben. Prinzipiell sind die zwei Modelle für ETW und EFH nahezu

gleich aufgebaut, um die Lagevariablen vergleichbar für die beiden Objektkategorien auswerten zu können.⁹⁸ In beiden Modellen werden neben den klassischen Objektkriterien⁹⁹ Dummies, welche die Kaufpreisinformationen beschreiben, sowie zwei Indexvariablen, welche die räumliche Variation der Kaufpreise beschreiben, verwendet.

Diese zwei konstruierten Indexvariablen basieren auf dem in BRUNAUER, LANG und FEILMAYR 2013 beschriebenen hybriden Modell. Diese Art von Modellen erlaubt es, „zwei heterogene Datensätze für räumliche Hauspreisvorhersagen nutzbar“¹⁰⁰ zu machen. Der Vorteil dieser Methode ist, dass man Datensätze, die eine Vielzahl an Kaufpreisinformationen enthalten, jedoch über nur wenige Objekteigenschaften verfügen (ein langer Datensatz) mit Datensätzen, die zwar über eine geringere Fallzahl, aber dafür eine genaue Objektbeschreibung (ein breiter Datensatz) haben, zusammenführen und somit die Qualitäten/Informationen beider Datensätze bestmöglich nutzen kann.¹⁰¹ Auf die Verwendung der in BRUNAUER, LANG und FEILMAYR 2013 beschriebenen hierarchischen Version von Strukturierten Additiven Regressionsmodellen (STAR) wird in dieser Arbeit verzichtet, da das Testen der Rastervariablen mit dieser komplexen Modellform aus Performancegründen in der Statistiksoftware BayesX nicht möglich ist. Ein weiterer Unterschied ist, dass in dieser Arbeit zwei so genannte lange Datensätze und ein breiten Datensatz verwendet werden und somit die räumlichen Lageeffekte durch den weiteren Informationsgewinn aus den zusätzlichen Daten noch detaillierter erklärt werden können.

4.1.1 Datenbeschreibung der Immobiliendaten

Wie bereits erwähnt, kommen drei Datensätze mit Immobilienpreisinformationen aus unterschiedlichen Quellen zur Anwendung:

1. UniCredit Bank Austria AG – Kauf- und Angebotspreise: Diese Daten werden im Zuge der Hypothekenbesicherung (Projekt Liegenschaftsbewertung (LIEBE))

⁹⁸ Weitere Details zu den Modellen sind in Kapitel 4.1.2 zu finden.

⁹⁹ Weitere Informationen dazu sind in Kapitel 2.1 zu finden.

¹⁰⁰ BRUNAUER, LANG und FEILMAYR 2013.

¹⁰¹ Vgl.: BRUNAUER, LANG und FEILMAYR 2013.

gesammelt und verfügen bezüglich der Objekteigenschaften über einen hohen Detaillierungsgrad. Weiteres verfügt diese Datensammlung über eine koordinatengenaue Verortung und ist dementsprechend bestens für die anstehenden Analysen der generierten Mikrolagevariablen geeignet.

2. IMMOUnited (IU) – Kaufpreise:¹⁰² Diese nach Objektart getrennten Kaufpreise werden aus der Kaufvertragssammlung des österreichischen Grundbuches erhoben. Man kann hier beinahe von einer Vollerfassung sprechen, da hier alle verbücherten Kauftransaktionen enthalten sind. Neben dem Kaufpreisdatum sind in diesem Datensatz als Objekteigenschaften nur die Grundstücksfläche bzw. teilweise die Wohnnutzfläche bei ETW vorhanden. Dementsprechend kann mit diesen Kaufpreisinformationen lediglich die räumliche Variation der Preisniveaus erklärt werden.
3. Ametanet Real Estate Software GmbH – Angebotspreise: Es handelt sich hier um Angebotsdaten aus einer Maklerdatenbank, die ebenfalls eine gute Abdeckung aufweisen. Objekteigenschaften wie Zustand, Flächen (Wohnnutzfläche, Terrasse, Loggia, Balkon), Zimmeranzahl, Heizungstyp, Badezimmertyp, Stockwerk, etc. sind in diesem Datensatz grundsätzlich vorhanden, wobei die Qualität schwankt und die Datenerfassung nur teilweise vollständig vorhanden ist.

In der nachstehenden Tabelle 4.1 sind zu den drei oben beschriebenen Datensätzen einige Eckdaten angeführt.

Datensatz	Datensatztyp ¹⁰³	Verfügbarkeitszeitraum		Anzahl plausibler Datensätze	
		von	bis	ETW	EFH
LIEBE (1)	breiter	Sep. 2008	Jän. 2013	3.294	3.842
IU (2)	langer	Jän. 2008	Sep. 2012	34.492	35.826
Ametanet (3)	langer	Jän. 2007	Sep. 2012	21.095	14.600

Tabelle 4.1: Informationen zu den Immobilienpreisdaten

¹⁰² Weitere Informationen zu diesem Datensatz sind im Anhang B.4 zu finden.

¹⁰³ Gemäß der Definition aus Kapitel 4.1.

In der Abbildung 4.1 ist die geographische Lage des LIEBE-Datensatzes getrennt nach EFH und ETW dargestellt. Wie man erkennen kann, sind die Preisdaten bei EFH im östlichen Bereich von Österreich dichter als im Westen und dementsprechend fällt dort auch die Schätzung besser aus. Die Datenlage bei Eigentumswohnungen (ETW) weist grundsätzlich ein ähnliches Ost-Westgefälle auf, wobei hier noch zu erwähnen ist, dass diese tendenziell eher in zentralen Lagen zu finden sind, da in peripheren Regionen hauptsächlich EFH errichtet werden.

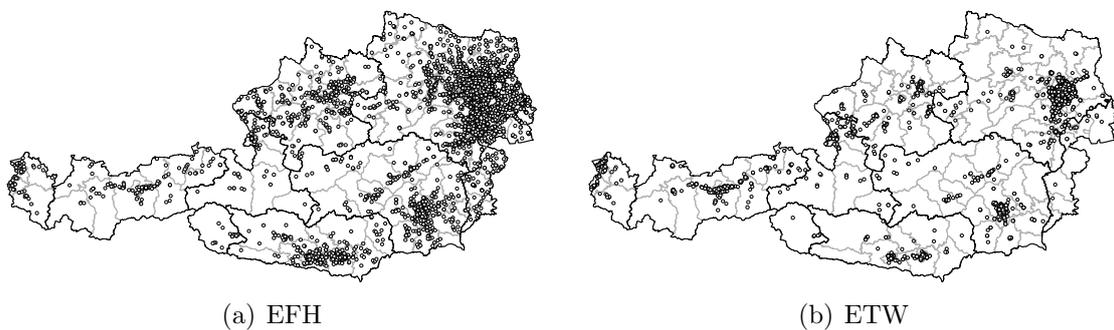


Abbildung 4.1: Verortung der LIEBE-Daten

Der IU und der Ameternetdatensatz sind österreichweit von sehr homogener Struktur, wobei natürlich in den nicht so dicht besiedelten Bereichen (z.B.: Alpen, Waldviertel (Niederösterreich), etc.) weniger Daten vorhanden sind.

4.1.2 Modellaufbau

Für die Regressionsmodelle, wird wie bereits im einführenden Kapitel 2.1.3 erläutert, ein generalisiertes additives Modell (GAM) verwendet, welches mit hybriden Variablen bestückt wird. Die beiden Modelle für ETW und EFH werden so ähnlich wie möglich gestaltet, um für die Validierung der Lagevariablen eine gut vergleichbare Grundlage zu haben. In der nachstehenden Tabelle 4.2 ist eine Variablenbeschreibung für die Basismodelle (EFH und ETW) angeführt.

Variablenname	Inhalt
lnp (EFH & ETW)	Die Zielvariable ist der logarithmierte Kaufpreis.
area_weight (EFH & ETW)	Da der breite LIEBE-Datensatz bei den Nutzflächen der Gebäude bzw. Wohnungen über einen hohen Differenzierungsgrad (Keller, Erdgeschoss, Dachgeschoss, Loggia, Balkon, Wintergarten, Terrasse) verfügt, wurden im ersten Schritt die Gewichte der unterschiedlichen Flächentypen mit einem Modell geschätzt. Die Referenzkategorie in diesem Modell sind die für Wohnzwecke genutzten Flächen mit einem Gewicht von 1. Kellerflächen oder nicht für Wohnzwecke geeignete Dachgeschossflächen werden entsprechend dem geschätzten Gewichtungsmo- dell mit einem geringeren Gewicht eingestuft. Mit dem Output aus diesem Modell wird dann eine über die Flächentypen gewichtete Gesamtnutzfläche ermittelt.
dum_angebot (EFH & ETW)	Dummy für Angebotspreis (EFH: 405, ETW: 360)
dum_gemeinnutz (EFH & ETW)	Dummy für Objekte, die von einem gemeinnützigen Bauträger errichtet wurden (EFH: 31, ETW: 118).
dum_bewerter (EFH & ETW)	Dummy für Bewertereinschätzung bedeutet, dass im Zuge des Kreditvergabeprozesses eine professionelle Immobilienbewertung mit der LIEBE-Software durchgeführt wurde und der dort ausgewiesene Marktwert durch den Immobiliensachverständigen übersteuert wurde (EFH: 2023, ETW: 1084).
rh (EFH)	Reihenhausdummy
antwortid_heiz (EFH & ETW)	Kategoriale für Zentral- (=1) oder Individualheizung (=2)
antwortid_bad_qual (EFH & ETW)	Ausstattungsqualität des Badezimmers in drei Kategorien: 1 = sehr gut, 2 = mittel, 3 = schlecht

KAPITEL 4. VALIDIERUNG DER MIKROLAGEVARIABLEN MIT EINEM GENERALISIERTEN ADDITIVEN MODELL (GAM)

Variablenname	Inhalt
bad_size_kat1 (EFH & ETW)	Die Kategorien für die Größe des Badezimmers sind wie folgt eingeteilt: 1 = bis 4 m ² und 2 = größer als 4 m ² .
zusatzbad (EFH & ETW)	Dieser Dummy beschreibt das Vorhandensein eines zusätzlichen Badezimmers.
terr_wohn_vorh (EFH)	Dummy für eine überdachte Terrasse
balc_vorh (ETW)	Dummy für Balkon
lift_upTRUE (ETW)	Diese Interaktionsvariable prüft, ob ein Lift vorhanden ist und die Wohnung mindestens im 3. Stock liegt.
antwortid_garage_3 (ETW)	Dummy für die Verfügbarkeit einer Garage
antwortid_zustand _haus (EFH & ETW)	Kategoriale Variable, die den Zustand des EFH beschreibt. 1 = sehr gut, 2 = mittel, 3 = schlecht. Bei ETW wird der Zustand des gesamten Wohnhauses damit beschrieben.
antwortid_zustand _wohnung (ETW)	Kategoriale Variable für die Beschreibung des Wohnungszustandes. 1 = Erstbezug nach Generalsanierung, 2 = sehr gut, 3 = mittel, 4 = schlecht
bundes (EFH & ETW)	Eine kategoriale Variable, welche die Bundeslandzugehörigkeit der Beobachtung beschreibt.
baujahr_berech (EFH & ETW)	In diesem Datensatz wird neben dem Baujahr auch erhoben, wann welches Bauteil (bzw. Gewerk) saniert wurde. Daraus kann ein fiktives Baujahr ermittelt werden, welches in diesem Modell anstelle des klassischen Baujahrs verwendet wird.
area_gst (EFH)	Grundstücksfläche
kp_year (EFH & ETW)	Das Transaktionsdatum wird monatsweise dargestellt (Jahr + Monat/12).

Variablenname	Inhalt
zsp_iu_lnefh_2012 zsp_iu_lnetw_2012 (EFH & ETW)	Dies ist eine der im Kapitel 4.1 erläuterten „hybriden“ Variablen, welche die logarithmierten Schätzergebnisse eines darunter liegenden Regressionsmodells aus den IU-Daten enthält. Diese Variable variiert auf Zählsprengelevelniveau.
zsp_am_lnefh_2012 zsp_am_lnetw_2012 (EFH & ETW)	Die zweite „hybride“ Variable stammt aus dem Ametanetdatensatz und enthält ebenfalls die logarithmierten Schätzergebnisse des Ametanetmodells auf Zählsprengelebene.

Tabelle: 4.2: Variablen der Basismodelle

Wenn sich nach dieser Auflistung nun die Frage stellt, warum beim Variablenset der beiden Basismodelle neben den hybriden Variablen aus dem Ametanet- und IU-Datensatz keine weiteren Lagevariablen, welche die räumlichen Variationen der Wertniveaus erklären, vorhanden sind, ist dies wie folgt zu begründen:

In den vorgeschalteten Submodellen kommen, wie bereits erläutert, die zwei langen Datensätze sowie alle wertrelevanten Lageparameter von Bundesland- bis auf Zählsprengelebene zum Einsatz. Diese Lageparameter sind beispielsweise:

- Anteil der Akademiker pro ZSP
- Anzahl der Nächtigungen pro Gemeinde (GEM)
- Bevölkerungspotential pro GEM
- Veränderung der Kaufkraft pro GEM
- Bevölkerungsdichte pro GEM
- Ein Index aus dem Wirtschaftskammer Österreich (WKO)-Immobilienpreisspiegel pro Bezirk (BEZ)
- Kategoriale Variable der Bundesländer

Durch die Verwendung all dieser Lageparameter in den Submodellen werden die räumlichen Variationen durch die hybriden Variablen bereits so gut erklärt, dass diese Va-

riablen im Modell mit dem breiten Datensatz nicht mehr benötigt werden.¹⁰⁴

Im Anhang sind unter B.5.3 in den Abbildungen B.13 und B.16 die Formeln und die Regressionsoutputs der zwei Basismodelle, sowie nachfolgend auch die Funktionen der nichtlinearen Terme (Abbildung B.14 und B.17) dargestellt. Wie man dort erkennen kann, verfügen in diesen Modellen – mit Ausnahme der Bundeslandvariable bei ETW – alle Objekteigenschaften über einen sehr guten P-Wert und auch die Modellkennzahlen wie beispielsweise R^2 und Devianz haben bereits ein sehr gutes Niveau. Generell kann man also festhalten, dass man mit diesem bereits sehr guten Basismodell kaum mehr Verbesserungen bei den klassischen Modellkennzahlen machen wird und deshalb für die Modellvergleiche das Akaike Informationskriterium (AIC) als Hauptmodellwahlkriterien verwendet wird. Weiters werden auch das Bayesianische Informationskriterium (BIC), die erklärte Devianz (EDEV) und die geschätzten Freiheitsgrade (GFG) beim Aufstellen des Modells berücksichtigt.

Da die im Basismodell verwendeten linearen und nichtlinearen Terme auch im finalen Modell die gleiche Berücksichtigung finden, wird die Aufstellung des Modells im Zuge der Diskussion der finalen Modelle in Kapitel 4.3.3 und 4.3.4 noch detailliert erläutert.

4.2 Variablenselektion

In der Tabelle B.2 im Anhang sind alle zur Verfügung stehenden Lagevariablen, welche im Zuge des Kapitels 3 erzeugt wurden, aufgelistet. Zusammengefasst sind dies folgende Gruppen von Mikrolagevariablen:

- Nutzungsdatenschicht ohne Dummies (ca. 300) – Kapitel 3.1
- Höhenmodell (30) – Kapitel 3.2
- Lärmmodell ohne Einzellärmvariablen (6) – Kapitel 3.3
- Erreichbarkeitsmodell (30) – Kapitel 3.4
- Bevölkerungsraster (ca. 70)– Kapitel 3.5.1
- Lebensphasen (20) – Kapitel 3.5.2

¹⁰⁴ Weitere Informationen zu hybriden Modellen sind in BRUNAUER, LANG und FEILMAYR 2013 zu finden.

- Gebäuderaster (20) – Kapitel 3.5.3
- Klimadaten (7) – Kapitel 3.5.4

In der obigen Aufstellung sind abgeleitete, kategoriale, Dummies und logarithmierte Variablen nicht enthalten. Um diese Variablen in weiterer Folge im Modell erkennen zu können, wird bei Dummies ein „_dum“, bei logarithmierten eine „_ln“ und bei kategorialen jeweils die Ausprägung der Kategorie beim Variablennamen ergänzt.

Da für die Modellierung der Preisdaten ein GAM verwendet wird, ergibt sich die Problemstellung, dass für diese Modellfamilie die klassische Stepwisemethode im „mgcv“ R-Package von GAM nicht zur Verfügung steht. Um im Zuge dieser Arbeit allerdings die mehr als 600 Variablen abtesten zu können, ist es unumgänglich eine automatisierte Form der Variablen(vor-)selektion durchzuführen, um zumindest eine sehr grobe Einschränkung auf die relevanten Variablen zu erhalten.

4.2.1 Vereinfachtes Backward Stepwise

Es ergibt sich durch die sehr große Anzahl von Variablen das Problem, dass selbst der Versuch, das Modell für die Variablenselektion vorerst nur linear aufzubauen, um damit die klassische Stepwiseselektion auszuführen, sehr lange dauert und kein erwartetes Ergebnis ermittelt wird. Demnach spielt die Rechenleistung auch bei der Variablenselektion eine wesentliche Rolle. Somit ist es notwendig, eine Methode für eine performante automatisierte Variablen(vor-)selektion zu entwickeln. Optimalerweise würde man eine Selektionsmethode verwenden, die neben der Selektion auch entscheiden kann, ob die Variable als linearer Term oder als nichtlinearer Spline im Modell zur Anwendung kommt. Um nicht Gefahr zu laufen, während der Stepwiseiteration ein Problem mit zu vielen verbrauchten Freiheitsgraden zu bekommen, wird diese Funktionalität nicht in die vereinfachte Lösung eingebaut. Ziel dieser Methode ist es möglichst schnell zu einer vorläufigen Grobselektion zu gelangen, um mit den generierten Daten viele verschiedene Varianten ausprobieren zu können und erst darauf aufbauend die finale Modellzusammenstellung durchzuführen.

4.2.1.1 Technische Umsetzung

Bevor die vereinfachte Variablenselektion begonnen wird, muss ein Modell mit allen für die Selektion vorgesehenen Variablen aufgestellt werden, wobei alle Variablentypen inklusive nichtlinearer Funktionen zulässig sind. Ausgehend von diesem Modell wird anhand der ermittelten P-Werte eine Selektion der Variablen vorgenommen.

Um nun nicht einfach alle Variablen, die einen schlechten P-Wert aufweisen, aus dem Modell zu werfen, werden in einem iterativen Prozess nach und nach die schlechtesten aus dem Modell entfernt und dann das Modell erneut berechnet. Ein beispielhafter Funktionsaufruf der vereinfachten Variablenselektion in der Statistiksoftware R ist, wie in Algorithmus 2 dargestellt, definiert. Mit dem Verkleinerungsfaktor `Vkf` können die Iterationsschritte, welche zur Verkleinerung der P-Wertgrenzen herangezogen werden, spezifiziert werden. Der minimalste P-Wert wird mit `MPw` definiert. Das bedeutet, dass die Iterationen bis zu diesem Wert fortgesetzt werden. Da es aufgrund der teilweise ähnlichen Variablen im Modell durchaus vorkommen kann, dass einige davon kollinear sind, kann mit `KV` eingestellt werden, ob diese Variablen für den nächsten Berechnungsdurchlauf behalten werden oder ausscheiden sollen. Sobald eine oder mehrere Variablen aus dem Modell entfernt werden, startet die Iteration mit der daraus resultierenden neuen Formel wieder am Anfang und fährt gemäß dem eingestellten `Vkf` wieder so lange fort, bis die nächstschlechtere Variable zum Entfernen aus dem Modell gewählt wurde.

```
Modell <- gam(lnp ~ s(baujahr_berech) + s(area_gst) + s(kp_year, bs = 'ps')
              + ...
              ,data = efh)

Vkf <- 0.01
MPw <- 0.1
KV <- "`YES'"
Selektionsmodelle <- my_easy_step(Modell, MPw, Vkf, KV)

Einzelmodell <- Modellerzeugung(Selektionsmodelle,"mod_efh_2012_12_02_142557")
summary(Einzelmodell)
...
```

Algorithmus 2: Aufruf der eigenen Variablenselektion

Als Output eines Selektionsdurchlaufs erhält man dann eine Tabelle

(Selektionsmodelle) mit der Definition (Modellnamen und Formeln) aller automatisch berechneten Modelle sowie die Werte einiger Modellwahlkriterien (AIC, BIC, R^2 , Erklärte Devianz (EDEV), Geschätzte Freiheitsgrade (GFG)). Ausgehend vom Modellnamen in dieser Tabelle kann das automatisch erstellte Modell mit einer weiteren Funktion (Einzelmodell) rekonstruiert werden. Man hat somit die Möglichkeit auch die einzelnen Variablen des spezifizierten Modells zu überprüfen.

4.2.2 Praktische Umsetzung der Selektion

Die zweistufig aufgebaute Variablenselektion wird für EFH und ETW separat ausgeführt, folgt allerdings in beiden Fällen dem nachstehenden Vorgehen:

1. Mittels der in Kapitel 4.2.1 beschriebenen vereinfachten Backward Stepwise Methode wurde eine Vielzahl von unterschiedlichen Berechnungsdurchläufen ausgeführt. Hierbei sind einerseits unterschiedliche Einstellungen (Vkf, MPw, KV) beim Aufruf der Modellerzeugung berechnet worden, andererseits wurde auch die Spezifikation des Startmodells (Modell) mit unterschiedlichen Varianten getestet. Beispielsweise wurden – neben dem Abtesten aller Variablen im jeweiligen Datensatz – die Daten aus der Nutzungsschicht mit Dummycodierung und als Flächenanteile separat getestet, oder es wurden nur Variablen aus bestimmten Themengebieten, wie etwa dem Bevölkerungsraster, getestet. Bei all diesen Modelltestläufen wurde immer dokumentiert, welche Variablen in den erstellten Selektionsmodellen übrig blieben. Aus diesem Variablenset wurden dann jene gewählt, die in unterschiedlichen Modellkonstellationen mehrmals als signifikant identifiziert wurden.
2. Ausgehend von der in Punkt eins beschriebenen groben Vorauswahl der Variablen wurden nun manuell verschiedene Modellvarianten ausgetestet. Die daraus abgeleiteten Ergebnisse werden im folgenden Kapitel 4.3 detailliert erläutert.

4.3 Finale Modelle und Schlussfolgerungen

In diesem Kapitel werden die zwei Modelle EFH (Abbildung 4.2 und 4.4) und ETW (Abbildung 4.7 und 4.8) sowie die Effekte der darin enthaltenen Objektkriterien als auch die in dieser Arbeit neu erstellten Mikrolagevariablen diskutiert. Da bei der Beschreibung des Basismodells (Kapitel 4.1.2) im Wesentlichen nur die Variablen erklärt wurden (Tabelle: 4.2) – nicht aber deren Effekte – wird dies nun in diesem Kapitel nachgeholt. Bei der Interpretation der Modellergebnisse wird grundsätzlich immer vom finalen Modell ausgegangen. Direkte Modellvergleiche werden dann im nachfolgenden Kapitel 4.4 angestellt.

Wie man aus den Regressionsergebnissen für EFH (Abbildung 4.2) und ETW (Abbildung 4.7) erkennen kann, werden bei den Parametern zuerst die linearen Terme ausgegeben. Im zweiten Tabellenblock sind die nichtlinearen Funktionen angeführt. Für die Interpretation der Effekte dieser Funktionen werden die in den Abbildungen 4.4 (EFH) und 4.8 (ETW) dargestellten Funktionskurven herangezogen. Nachfolgend werden nun die linearen (Kapitel 4.3.1) und die nichtlinearen (Kapitel 4.3.2) Variableneffekte aus dem Basismodell anhand der finalen Modellzusammenstellung erläutert.

Die neu generierten Mikrolagevariablen werden danach in den Kapiteln 4.3.3 (EFH) und 4.3.4 (ETW) diskutiert.

4.3.1 Ergebnisse aus den linearen Funktionen

In diesem Kapitel werden die Effekte der linearen Funktionen aller aus dem Basismodell stammenden Variablen für die beiden Objektarten (EFH und ETW) diskutiert.

- *Dummies für Angebot, gemeinnützige Bauträger und Bewerter:* Bei diesen Variablen ist das Niveau bei allen Modellen durchwegs ähnlich. Durch gemeinnützige Bauträger errichtete Objekte weisen immer einen deutlich niedrigeren Kaufpreis auf. Der Unterschied im Vergleich mit dem tatsächlichen Kaufpreis liegt bei EFH mit -10,6 %¹⁰⁵ noch niedriger als bei ETW mit -7,2 % . Bei Angebotspreisen liegt

¹⁰⁵ Für diese und alle folgenden aus den Modellen abgeleiteten Prozentwerte gilt, dass diese entlogarithmiert werden und demnach vom jeweiligen Regressionsoutput abweichen können.

EFH um 9,4 % und ETW um 5,8 % über den tatsächlichen Kaufpreisen. Die Dummyvariable der Bewertereinschätzung ist bei Eigentumswohnungen obsolet, da diese keinen Wert im Modell liefert und dementsprechend die Kaufpreise mit den Bewertereinschätzungen perfekt zusammenpassen. Bei Häusern überschätzen die Bewerter allerdings um 3,5 %. Es lässt sich zu diesen Variablen also festhalten, dass dadurch der Effekt der Datenherkunft im Modell gut erklärt werden kann.

- Dummies für Reihenhäuser: Reihenhäuser sind im Vergleich zum klassischen Einfamilienhaus um 4,6 % teurer. Die Begründung könnte darin liegen, dass Reihenhäuser eher, wie auch Wohnungen, von einem Bauträger errichtet werden und somit die Ersparnis bei der Errichtung durch die sogenannte Nachbarschaftshilfe wegfällt.
- Individualheizung: Wenn keine Zentralheizung vorhanden ist, schlägt sich dies in allen Modellen negativ nieder. Bei EFH sind es -8,8 %, bei ETW ist dieser Effekt mit -5,7 % zu beziffern.
- Ausstattung des Badezimmers: Hat bei EFH eine Bandbreite von knapp 16,6 % zwischenn guter und schlechter Ausstattung, bei ETW beträgt dies 14,3 %.
- Kleines Badezimmer: Badezimmer unter 4 m² sind im Vergleich mit größeren Bädern bei ETW -2,1 % weniger wert, wobei der Effekt bei Einfamilienhäusern mit -4,7 % etwas stärker ausfällt.
- Zusätzliches Badezimmer: Ein zweites Badezimmer wirkt sich bei beiden Objektkategorien positiv aus, wobei der Effekt bei ETW mit 4,3 % beinahe doppelt so hoch wie bei EFH mit 2,7 % ausfällt.
- Zustand des Hauses/Gebäudes: Dieser hat neben Fläche und Baujahr einen der stärksten Effekte bei den Objektkriterien. Bei EFH ist die Spanne deutlich größer, da bei ETW zusätzlich noch der Wohnungszustand hinzukommt. Dass der Gebäudezustand bei ETW dennoch eine sehr hohe Spanne zur Referenzkategorie (3 = schlechter Zustand) aufweist, ist darauf zurückzuführen, dass die Gebäude bei Eigentumswohnungen üblicherweise gut in Stand gehalten werden und deshalb nur selten ein sehr schlechter Zustand auftritt. Dementsprechend greift hier zumeist der nur sehr geringe Unterschied von 7,5 % zwischen gutem und mitt-

lerem Zustand. Bei EFH ist der Unterschied zwischen gut und mittel hingegen 14,9 %.

- Zustand der Wohnung: Bei ETW ist der Wohnungszustand wichtiger als der Gebäudezustand (der Unterschied zwischen sehr gut und schlecht ist 26,6 %).
- Balkon: Bei ETW bringt ein Balkon einen leicht positiven Effekt von 2,6 %.
- Lift: Der Interaktionseffekt für Lift bringt bei Eigentumswohnungen eine deutliche Wertsteigerung von 7,1 %, wobei dieser allerdings auch eine Stockwerkslage ab dem 3. Stock einschließt.
- Garage: Das Vorhandensein einer der Wohnung direkt zugeordneten Garage bringt im Vergleich mit Tiefgaragen bzw. privat verfügbaren Außenabstellplätzen eine vergleichsweise geringe Wertsteigerung von 3,7%.
- Bundesland: Die Referenzkategorie ist bei EFH die Steiermark und bei ETW Wien. Der Bundeslandeffekt stellt die Unterschiede bei den Immobilienpreisen bei beiden Objektarten gut dar. Die günstigsten Wohnimmobilien sind dementsprechend im Burgenland und Kärnten zu finden, gefolgt von der Steiermark, Nieder- und Oberösterreich. Beim Niveau der restlichen Bundesländer ist ein Ost-West Preisgefälle deutlich zu erkennen. Es kommt hier Salzburg und dann Tirol und Vorarlberg. Wien muss separat betrachtet werden, da für Wien bei EFH die Grundstücksfläche und bei ETW die Wohnnutzfläche mit einer eigenen Funktion geschätzt wird und dadurch die singuläre Betrachtung des geschätzten Parameters aus dem Regressionsplot nicht ausreicht. Generell kann man aber festhalten, dass in Wien neben Tirol und Vorarlberg die höchsten Immobilienpreise erzielt werden.

4.3.2 Ergebnisse der nichtlinearen Funktionen

In diesem Kapitel werden die aus dem Basismodell stammenden Funktionen der nichtlinearen Variablen für EFH und ETW erläutert und interpretiert.

- Grundstücksfläche: Der Effekt der Grundstücksfläche wird bei EFH für Wien (Abbildung 4.4(c)) und Restösterreich (Abbildung 4.4(b)) separat geschätzt. Wenn

man bei diesen Grafiken die Skala betrachtet, fällt auf, dass das Niveau in Wien deutlich niedriger ist. Dies hängt damit zusammen, dass der bereits oben erwähnte Bundeslandeffekt in Wien aufgrund der hier separat durchgeführten Schätzung in einen ganz anderen Wertebereich fällt und dementsprechend hier hinzugerechnet werden muss, um eine Aussage über die Unterschiede zwischen Wien und Restösterreich treffen zu können.

Wie in den Abbildungen ersichtlich, weist der Effekt in Österreich bei ca. 900m² einen deutlichen Knick auf, was bedeutet, dass ab diesem Punkt die Zahlungsbereitschaft für einen weiteren m² Grundstücksfläche deutlich abnimmt und dementsprechend größere Grundstücke verhältnismäßig günstiger veräußert werden. In Wien hingegen ist der Effekt als beinahe linear zu beobachten, was bedeutet, dass ein zusätzlicher m² Grund hier immer den gleichen Preis aufweist. In den restlichen Landeshauptstädten oder in einzelnen Gebirgslagen, wo der Siedlungsdruck ähnlich hoch ist, kann dieser Effekt ebenfalls beobachtet werden.

- Gewichtete Nutzfläche: Bei EFH (Abbildung 4.4(a)) zeigt die gewichtete Nutzfläche einen ähnlichen Effekt wie die Grundstücksfläche, wobei hier der erste nennenswerte Knick bei etwa 100m² auftritt. Demnach beginnt ab 100m² die Zahlungsbereitschaft für zusätzliche m² Nutzfläche relativ gleichmäßig zu sinken und ab 300m² ist in diesem Modell quasi keine Wertsteigerung mehr aufgrund des Flächeneffektes zu verzeichnen.

Bei ETW wird die gewichtete Nutzfläche für Wien (Abbildung 4.8(d)) und die restlichen Bundesländer (Abbildung 4.8(c)) separat modelliert. Allerdings ist der Verlauf beider Kurven ähnlich, wobei der Kurvenverlauf für Wien auch hier etwas progressiver ist. Der bereits erwähnte Niveauunterschied durch die codierten Bundeslandeffekte ist auch bei der Interpretation der Niveaus dieser Kurven zu berücksichtigen.

- Transaktionsdatum: Dies kann als Index für die Entwicklung der Immobilienpreise im Beobachtungszeitraum von Jänner 2006 bis Jänner 2013 gesehen werden, wobei hier zu erwähnen ist, dass erst seit dem dritten Quartal 2008 eine ausreichende Anzahl an Immobilienpreisen vorhanden ist und dementsprechend die Schätzung

am unteren Ende der Skala etwas volatiler wird. Folgt man also den Daten, kann man zwischen 9/2008 und 1/2013 eine Wertsteigerung von ca. 10 % bei Einfamilienhäusern (Abbildung 4.4(d)) feststellen. Bei ETW (Abbildung 4.8(a)) ist die Wertsteigerung deutlich höher und kann zwischen 9/2008 und 1/2013 mit beinahe 35 % beziffert werden.

- *Berechnetes Baujahr*: Ein sehr interessanter Effekt, der sich beim Baujahr (Abbildung 4.4(e) (EFH) und 4.8(b) (ETW)) abzeichnet, ist, dass sich in beiden Modellen die absolut tiefsten Werte in den Jahren um 1940 zeigen (statistisches Artefakt, da in dieser Periode keine Häuser gebaut worden sind). Wohnungen in Gebäuden, die um 1900 (Gründerzeithäuser in Wien) erbaut wurden, erzielen heute einen rund 10,5 % höheren Verkaufspreis als jene, die in der Zwischenkriegszeit und auch nach dem Zweiten Weltkrieg errichtet wurden. Vermutlich ist dies auf die billige und schnelle Bauweise nach dem Krieg zurückzuführen. Erst ab dem Jahr 1970 beginnt bei Eigentumswohnungen wieder ein Aufwärtstrend beim Baujahreseffekt. Bei Einfamilienhäusern ist dieser Effekt nicht so stark ausgeprägt, hier ist die Talsohle um 1940 nur ca. 5,1 % niedriger im Vergleich mit den ältesten Häusern im Datensatz.
- *IU-Index*: Diese hybride Indexvariable auf Zählsprengelniveau stellt sich bei EFH (Abbildung 4.4(f)) beinahe linear dar und zeigt einen gleichmäßig steigenden Trendverlauf des Kaufpreisniveaus. Bei ETW (Abbildung 4.8(f)) hat diese Variable ebenfalls einen steigenden Wertverlauf. Der Effekt ist hier allerdings nicht so stark ausgeprägt, da Eigentumswohnungen grundsätzlich eher in zentralen Lagen errichtet werden und demnach die Niveauunterschiede nicht so hoch sind. Der Kurveneinbruch beim Indexwert 12,2 beträgt knapp 2 % und ist in Bezug auf das Gesamtmodell als unbedeutend anzusehen. Nach diesem kleinen Knick folgt der gesamte Kurvenverlauf wieder dem gleichen Aufwärtstrend wie zuvor und kann deshalb ohne Weiteres in dieser Form verwendet werden.
- *Ametanet-Index*: Bei diesem ist, ähnlich wie zuvor beim IU-Index bei EFH (Abbildung 4.4(g)) und auch bei ETW (Abbildung 4.8(e)) ein linearer Trend zu erkennen.

4.3.3 Mikrolagevariablen im Modell Einfamilienhaus (EFH)

In diesem Kapitel sind die Regressionsergebnisse des EFH-Modells dargestellt. Es werden die Effekte der im Zuge dieser Arbeit generierten Mikrolagevariablen diskutiert. Zum finalen Modell (Abbildung 4.2 und 4.4) kann man allgemein festhalten, dass durch die ergänzten Mikrolagevariablen bei den P-Werten der Objektkriterien kaum Veränderungen erkennbar sind und dementsprechend deren Variablenlänge im Modell nicht verändert wird. Das Niveau des Intercepts bleibt im Vergleich mit dem Basismodell nahezu ident. Dies bedeutet, dass die neu hinzugefügten Mikrolagevariablen die Effekte der Variablen aus dem Basismodell zwar deutlich verringern, sich aber dadurch keine wesentlichen Strukturbrüche im Modell ergeben. Was im Basismodell durch Objektkriterien erklärt wurde, wird jetzt verstärkt durch Mikrolagevariablen erklärt.

Beim Effekt der Bundeslandvariablen ist zu erkennen, dass diese nun etwas signifikanter geworden sind und auch die Niveauunterschiede zwischen den Bundesländern deutlicher ausfallen. Die bereits oben (Kapitel 4.3.2) gezeigten Funktionskurven verhalten sich im Basismodell ebenfalls sehr ähnlich wie im finalen Modell mit den Mikrolageeffekten, wobei auch hier die Niveaus etwas nach unten verschoben wurden.

Weitere ausführlichere Modellvergleiche folgen in Kapitel 4.4, dort wird auch der Unterschied zum hybriden Modell nach BRUNAUER, LANG und FEILMAYR 2013 dargestellt.

4.3.3.1 Lineare Mikrolagevariablen (EFH)

Nachfolgend werden die in dieser Arbeit verwendeten Mikrolagevariablen beschrieben und die daraus folgenden Effekte erörtert.

4.3.3.1.1 Ausländeranteil

Wenn man beispielsweise die Preisniveaus in jenen Wiener Gebieten betrachtet, die über einen hohen Ausländeranteil verfügen, kann man feststellen, dass in diesen Gebieten auch die Immobilienpreise ein etwas geringeres Niveau aufweisen. Diese These konnte bislang allerdings mit hedonischen Preismodellen für ganz Österreich nicht nachgewiesen werden. Ursache ist die Verfügbarkeit entsprechender Daten nur auf Zähl-

sprengelbene. Verwendet man hingegen die in dieser Arbeit neu verfügbaren Rasterinformationen, kann die Bevölkerungsstruktur räumlich differenziert ausgewertet werden. Die Daten pro Rasterzelle sind nach Geburtsland, Staatsangehörigkeit und Herkunft gegliedert. In dieser Arbeit wurde der Anteil der Einwohner bezogen auf ihre Herkunft verwendet. Es hat sich bei der näheren Analyse herausgestellt, dass im Wesentlichen die Süd- und Osteuropäer (inkl. Ex-Jugoslawen und Türken) die größte Gruppe von Personen mit ausländischer Herkunft sind. Alle anderen Herkunftsregionen (Amerika, Afrika oder Asien etc.) haben bei den Rasterdatensätzen, welche den Immobiliendaten zugeordnet werden, eine so geringe Fallzahl, dass sich diese bei der Schätzung als nicht signifikant darstellen.

Der lineare Effekt des Ausländeranteils von Süd- und Osteuropäern ist mit -13,1 % sogar stärker ausgefallen als erwartet. Bei diesem linearen Effekt bedeutet dies, dass bei 50 % Ausländeranteil im Umfeld der Immobilienpreis um 6,8 % unter dem in Österreich üblichen Niveau liegt. Selbst wenn nur 10 % der Bevölkerung eine Herkunft aus Süd- und Osteuropa aufweisen, verringert dies den Immobilienwert um 1,4 %.

4.3.3.1.2 Geländekrümmung

Die in Kapitel 3.2.3 beschriebene Geländekrümmung mit einem Radius von 15 Zellen und dem $St_{Typ} = \tilde{x}$ (Abbildung 3.9(h)) kommt im Modell zur Anwendung. Der Wertebereich dieser Variable in der Rasterdatenschicht liegt nach Bereinigung von Datenfehlern zwischen -100 (extreme Tallage) und + 100 (sehr hohe Kuppenlage mit weiter Aussicht). Da extreme Gebirgslagen, wo Werte bis 100 (Beispiel: Schutzhütte vor einer 300 Meter hohen Felswand) auftreten, üblicherweise unbesiedelt sind, ist in den Immobiliendaten nur die Bandbreite von -46 bis 30 enthalten, was bedeutet, dass der Effekt durch diese Variable mit von -8,5 % bis 5,9 % zu beziffern ist und daraus eine Bandbreite von 14,4 % für diese Schätzung abgeleitet werden kann. Betrachtet man das 1 % und 99 % Quantil, liegen die Werte zwischen -26 und 16 (Bandbreite: 8 %) und beim 5 % und 95 % Quantil ist das Niveau bereits auf -11 und 7 gesunken (Bandbreite: 3,4 %). Dies bedeutet, dass bei durchschnittlichen Siedlungslagen der Effekt mit ca. $\pm 1,7\%$ bis $\pm 4\%$ eingestuft werden kann.

Erwähnenswert bei dieser Variable ist, dass eigentlich alle Radien zwischen 3 und 15 gut im Modell funktioniert haben. Die verwendete Variable mit dem größten Radius (750m) bildet neben der Muldenlage im Speziellen auch noch die Aussichts- lage der Immobilie besser als die anderen ab und liefert somit den stärksten Effekt im Modell.

4.3.3.1.3 Akademikeranteil

Obwohl diese Variable bereits durch die Verwendung der hybriden Variablen auf Zähl- sprengelbasis im Modell enthalten ist, kann durch die feine Aufgliederung der Raster- information das Modell noch weiter verbessert werden. Der Akademikeranteil hat hier nur einen sehr geringen Effekt von 1,5 %. Beim ETW Modell, wie in Kapitel 4.3.4 beschrieben, stellt sich dieser Effekt stärker dar.

KAPITEL 4. VALIDIERUNG DER MIKROLAGEVARIABLEN MIT EINEM GENERALISIERTEN ADDITIVEN MODELL (GAM)

Family: gaussian

Link function: identity

Formula:

```
lnp ~ dum_angebot + dum_gemeinnutz + dum_bewerter + s(area_weight) +
  s(area_gst, by = bundes_AUT) + s(area_gst, by = bundes_WIEN) +
  s(kp_year) + s(baujahr_berech) + s(zsp_iu_lnefh_2012) + s(zsp_am_lnefh_2012) +
  rh + antwortid_heiz_efh + antwortid_bad_qual + bad_size_kat +
  zusatzbad + terr_wohn_vorh + antwortid_zustand_haus + bundes +
  br_her_so_ant + ra_kk_med_15 + lp_ln_akademiker_ant + s(ra_ocl_nat_mean_all) +
  s(ra_laerm_all_log_sum_q)
```

Parametric coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	7.2356640	0.0416766	173.615	< 2e-16 ***
dum_angebot	0.0894971	0.0125221	7.147	1.06e-12 ***
dum_gemeinnutz	-0.1116342	0.0388380	-2.874	0.004071 **
dum_bewerter	0.0342163	0.0082519	4.146	3.45e-05 ***
rh	0.0449414	0.0138519	3.244	0.001187 **
antwortid_heiz_efh2	-0.0889612	0.0137311	-6.479	1.04e-10 ***
antwortid_bad_qual1	0.1097648	0.0144323	7.606	3.56e-14 ***
antwortid_bad_qual3	-0.0509113	0.0093017	-5.473	4.70e-08 ***
bad_size_kat1	-0.0482155	0.0121128	-3.981	7.01e-05 ***
zusatzbad	0.0264982	0.0080510	3.291	0.001006 **
terr_wohn_vorh	0.0487004	0.0094930	5.130	3.04e-07 ***
antwortid_zustand_haus1	0.3025157	0.0170078	17.787	< 2e-16 ***
antwortid_zustand_haus2	0.1857719	0.0151198	12.287	< 2e-16 ***
bundesBGL	-0.0556288	0.0198254	-2.806	0.005043 **
bundesKTN	0.0205595	0.0176165	1.167	0.243260 *
bundesNOE	0.0339695	0.0155811	2.180	0.029306 *
bundesOOE	0.0341669	0.0172432	1.981	0.047611 *
bundesSBG	0.0786539	0.0339050	2.320	0.020403 *
bundesT	0.1675098	0.0298752	5.607	2.21e-08 ***
bundesVBG	0.1338287	0.0324063	4.130	3.71e-05 ***
bundesW	2.4560042	0.0596371	41.182	< 2e-16 ***
br_her_so_ant	-0.1407210	0.0439743	-3.200	0.001385 **
ra_kk_med_15	0.0019207	0.0004981	3.856	0.000117 ***
lp_ln_akademiker_ant	0.0152972	0.0036149	4.232	2.37e-05 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Approximate significance of smooth terms:

	edf	Ref.df	F	p-value
s(area_weight)	7.655	8.554	194.161	< 2e-16 ***
s(area_gst):bundes_AUT	4.993	6.035	2710.162	< 2e-16 ***
s(area_gst):bundes_WIEN	2.364	2.883	304.947	< 2e-16 ***
s(kp_year)	2.103	2.653	15.423	4.40e-09 ***
s(baujahr_berech)	5.940	6.970	92.281	< 2e-16 ***
s(zsp_iu_lnefh_2012)	3.692	4.687	9.441	1.57e-08 ***
s(zsp_am_lnefh_2012)	2.614	3.399	15.888	2.56e-11 ***
s(ra_ocl_nat_mean_all)	2.248	2.893	8.834	1.07e-05 ***
s(ra_laerm_all_log_sum_q)	2.104	2.695	18.541	5.74e-11 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) = 0.846 Deviance explained = 84.8%
 GCV score = 0.044019 Scale est. = 0.043368 n = 3833
 AIC = -1092.25 BIC = -731.45

Abbildung 4.2: Finales Modell EFH

4.3.3.2 Nichtlineare Mikrolagevariablen (EFH)

4.3.3.2.1 Erreichbarkeit

In Kapitel 3.4 wurde eine Vielzahl von Erreichbarkeitsvariablen generiert. Eine Auflistung findet sich im Anhang in Tabelle B.2. Für Einfamilienhäuser hat sich der Mittelwert über die Klassen 1, 2 und 3 als beste Variable herausgestellt, da dieser als Zentralitätsindex (zu allen Zentren) zu sehen ist. Durch die Mittelwertbildung über alle drei Klassen, welche jeweils die Entfernung zum erst-, zweit- und drittnächsten definierten Zentrum enthalten, entsteht eine indirekte Gewichtung der Zielorte. Dies ist darauf zurückzuführen, dass, wie bereits in Kapitel 3.4.4 beschrieben, in den höheren Klassen (2 und 3) immer auch die Ziele der darunterliegenden Klassen enthalten sind und dadurch die Landeshauptstädte und die Städte der Klasse 2 höher als jene der Klasse 3 gewichtet sind.

Betrachtet man nun den Funktionsverlauf der Erreichbarkeitsvariable in Abbildung 4.4(h), kann man ablesen, dass mit abnehmender Erreichbarkeit, wie erwartet, ebenfalls der Wert der Immobilie abnimmt. Auffällig beim Funktionsverlauf ist, dass etwa bei einem Erreichbarkeitswert von ca. 80 die Kurve deutlich flacher wird und dann kaum mehr eine zusätzliche Wertminderung auftritt. Plakativ gesprochen macht es für den Immobilienpreis keinen Unterschied ob man 80 oder 120 Minuten zur nächsten Landeshauptstadt fährt. In Abbildung 4.3 ist der verwendete Erreichbarkeitskennwert für Österreich dargestellt.¹⁰⁶ Deutlich sichtbar (Abbildung 4.3) ist, dass nur sehr entlegene Regionen in den Zentralalpen oder im nördlichen Waldviertel überhaupt einen derart schlechten Erreichbarkeitswert aufweisen. Hier hat eine weitere Verschlechterung der Erreichbarkeit keine Relevanz mehr für den Marktwert einer Liegenschaft. An dieser Stelle ist noch anzumerken, dass die Erreichbarkeitsvariablen, wenn sie ausländische Zielpunkte enthalten (siehe Kapitel 3.4.6), keinen Einfluss haben. Neben der im Modell verwendeten Erreichbarkeitsvariable „ra_laerm_all_log_sum“ funktionieren auch einige andere wie beispielsweise „ra_ocl_min_3_nat_mean“, „ra_ocl_1_mean“ oder „ra_ocl_3_nat_mean“¹⁰⁷ gut im Modell.

¹⁰⁶ Weitere Informationen zur Konstruktion dieser Variable sind in Kapitel 3.4 zu finden.

¹⁰⁷ Die Definitionen dazu sind in Tabelle B.2 zu finden.

Die erstgenannte liefert allerdings ein etwas besseres Modellergebnis.

Zur Höhe des Effektes ist zu sagen, dass, obwohl die Erreichbarkeit bereits auf Gemeindeebene in den darunterliegenden Modellen berücksichtigt wird, sich mit der Rastererreichbarkeitsvariable noch eine relativ hohe Verbesserung im Modell erreichen lässt. Nachdem der höchste Wert bei 7,3 % und der niedrigste bei -9,5 % liegt, ergibt sich eine Bandbreite von rund 16,8 %, was ein äußerst wertrelevanter Effekt für die Preise bei Einfamilienhäusern ist.

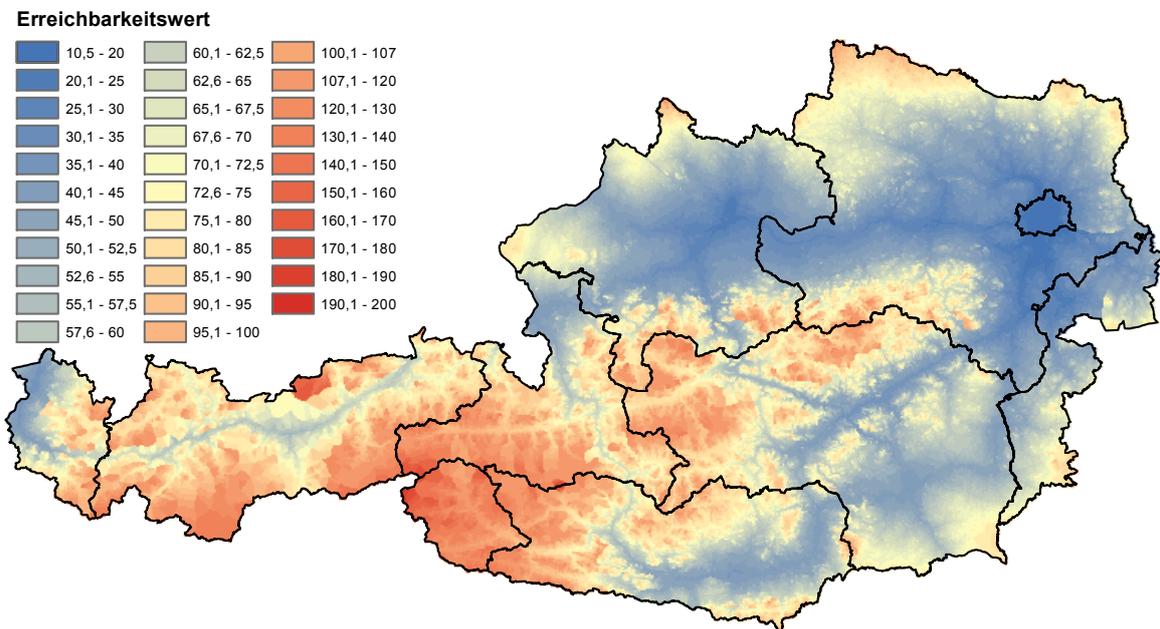


Abbildung 4.3: Erreichbarkeitsindex

4.3.3.2.2 Lärm

Jede im Kapitel 3.3 beschriebenen Variable aus dem Lärmmodell funktioniert grundsätzlich als erklärende Variable im Preismodell. Verwendet man anstelle der Summe über alle Lärmgruppen aus Formel 3.25 (Variable: „ra_laerm_all_log_sum“, Abbildung: 3.17(c)) die einzelnen Lärmgruppen (siehe Tabelle B.2), erhält man ähnliche Ergebnisse. Der Unterschied liegt lediglich darin, dass der in Abbildung 4.4(i) dargestellte Funktionsverlauf in mehrere Kurven aufgeteilt wird. Dadurch werden für die unterschiedlichen Lärmquellen jeweils eigene Funktionen geschätzt. Dies hat zur Folge, dass

die Lärmgruppe B (Bahnlärm)¹⁰⁸ im Vergleich mit allen anderen Lärmgruppen einen deutlich steileren Kurvenverlauf nach unten aufweist. Das bedeutet, dass Bahnlärm anscheinend stärker negativ wahrgenommen wird als alle anderen Lärmquellen und dementsprechend die Preisverminderung etwas stärker ausfällt. Da die Gesamtsumme des Lärms allerdings beim Modellvergleich die besten Kennzahlen liefert, wird im finalen Modell diese herangezogen, um die Wertverminderung durch Lärm darzustellen.

Betrachtet man nun den Kurvenverlauf in Abbildung 4.4(i), kann man ablesen, dass ab ca. 60 dB der Wertverlust für zusätzliche Lärmbeeinträchtigung deutlich stärker wird. Die Bandbreite bei der Lärmbelastung ist mit 13,4 % höher als erwartet und liegt zwischen 6,2 % bei minimalem Lärmpegel und bei -7,2 % mit sehr hohen dB-Werten. Speziell im leisen Bereich unter 50 dB erscheint die Werterhöhung etwas zu progressiv. Würde man bei diesem Wert abschneiden hätte man einen ähnlich guten Kurvenverlauf wie bei ETW in Abbildung 4.8(h) dargestellt. In diesem Fall wäre auch der Wertebereich mit ca. 9 % nahezu ident mit ETW.

¹⁰⁸ Siehe dazu auch in Tabelle 3.5.

KAPITEL 4. VALIDIERUNG DER MIKROLAGEVARIABLEN MIT EINEM GENERALISIERTEN ADDITIVEN MODELL (GAM)

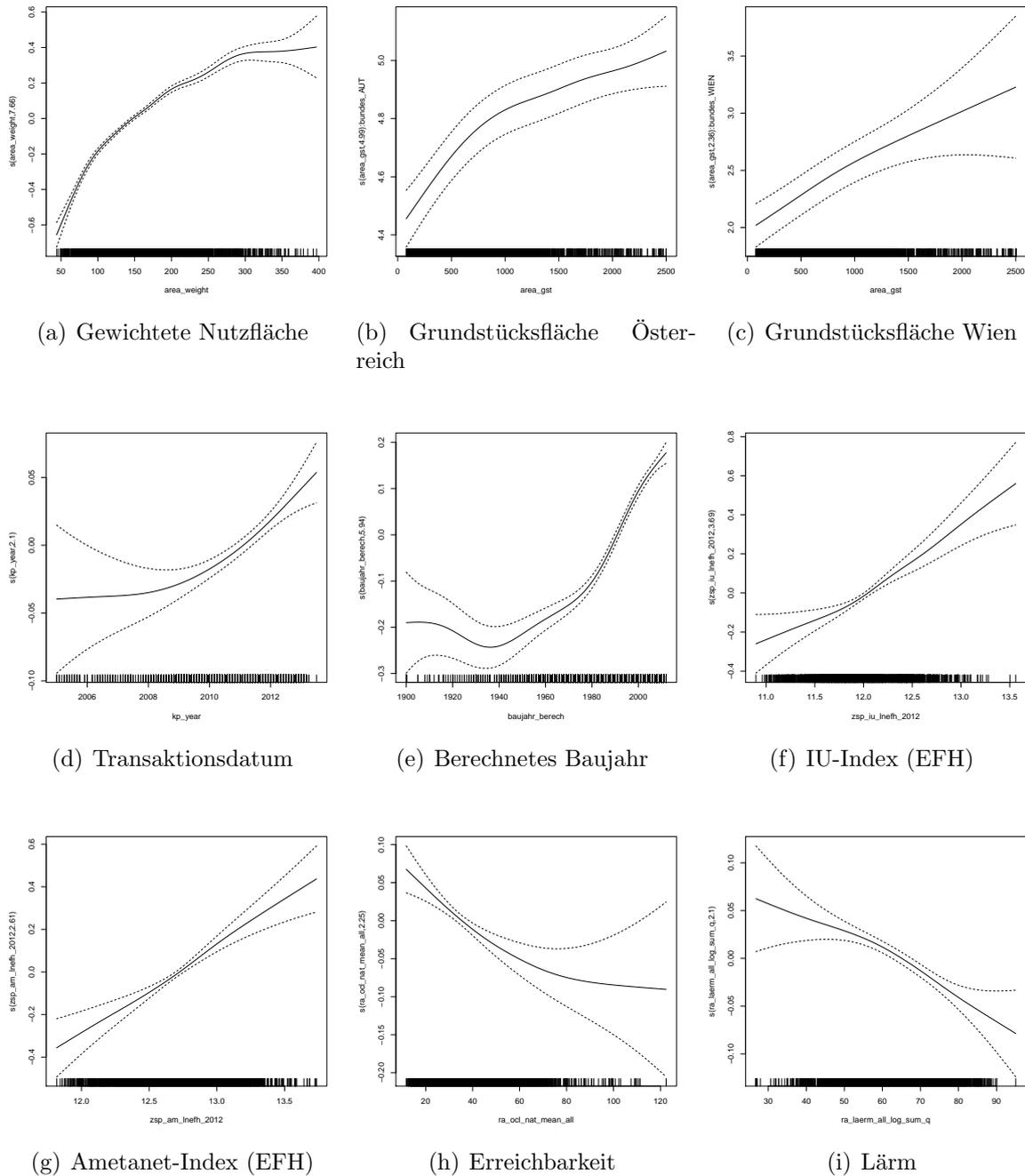


Abbildung 4.4: Nichtlineare Funktionen des finalen Modells (EFH)

4.3.3.3 Modellqualität EFH

In der nachstehenden Abbildung 4.5 wird die Schätzgenauigkeit des Modells graphisch dargestellt. In Abbildung 4.5(a) wird der Regressand mit den gefitteten Werten ver-

glichen. Man kann deutlich erkennen, dass die geschätzten Werte entlang der Winkelhalbierenden (45° Linie) angesiedelt sind. Beim Residuencheck in Abbildung 4.5(b) werden die Residuen anhand der Schätzung (Prediction) aus dem Modell ausgewertet. Die Verteilung der Datenpunkte ist hier zufällig um die Mitte zentriert, und es sind keine systematischen Abweichungen erkennbar, wie es bei einem guten Modell auch sein soll. Auch die Verteilung der Residuen (Abbildung: 4.5(c)) ist nahezu ident mit der in rot eingezeichneten Normalverteilung der Daten, was ebenfalls eine gute Modellqualität bestätigt.

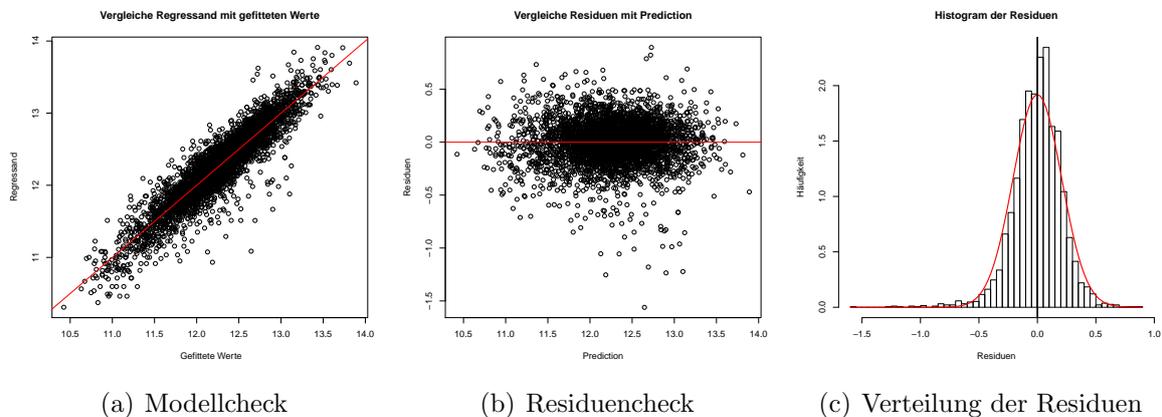


Abbildung 4.5: Modellqualität des finalen Modells (EFH)

Weitere Modellkennzahlen zu diesem Modell¹⁰⁹ und Vergleiche mit anderen Modellen sind im Kapitel 4.4 zu finden.

4.3.4 Mikrolagevariablen im Modell Eigentumswohnung (ETW)

Wie schon zuvor im Kapitel 4.3.3 für EFH, werden nun ebenfalls die Effekte der in dieser Arbeit generierten Mikrolagevariablen auf ETW genauer erläutert. Es werden hierbei allerdings hauptsächlich die Unterschiede zum EFH Modell (Kapitel: 4.3.3) beschrieben, gleiche oder ähnliche Effekte werden nicht wiederholt und sind im vorherigen Kapitel nachzulesen.

¹⁰⁹ Siehe Abbildung 4.2.

4.3.4.1 Lineare Mikrolagevariablen (ETW)

In diesem Kapitel werden die im ETW-Modell verwendeten linearen Mikrolagevariablen beschrieben und die daraus ablesbaren Effekte diskutiert.

4.3.4.1.1 Ausländeranteil

Wie bereits zuvor in Kapitel 4.3.3.1.1 für EFH beschrieben, funktioniert der Ausländeranteil auf Rasterebene im Modell sehr gut. Bei ETW fällt der Effekt mit -15,9 % im Vergleich mit EFH noch um -2,8 % stärker aus. Dies hängt damit zusammen, dass es in den städtischen Regionen mit vielen Eigentumswohnungen (z.B.: in Wien und den restlichen Landeshauptstädten) Gebiete gibt, die deutlich höhere Ausländeranteile aufweisen als die klassischen Einfamilienhaussiedlungsräume.

4.3.4.1.2 Arbeitsstätten pro Einwohner

Diese Variable findet nur bei ETW Verwendung und kann als eine Art Zentrumsproxy gesehen werden. Eine hohe Anzahl an Arbeitsstätten ist zwar punktuell auch in Industrie- und Gewerbegebieten zu finden, allerdings befinden sich dort üblicherweise keine Wohnbauten. Es wird in der direkten Umgebung dieser der positive Effekt durch die Lärmvariable vermindert. Die meisten Arbeitsstätten sind bislang noch immer in den Stadt- und Ortszentren angesiedelt, was demnach eine zentrale Lage der Wohnung verspricht. Bei dieser Rastervariable kommt ebenfalls eine 250m Rastermaschenweite zum Einsatz, was einem 6,25 ha großem Areal entspricht, wodurch punktuelle Spitzenwerte abgefangen werden. Wertet man die Daten aus der Rasterschicht aus, kommt man auf positive Wertänderungen von bis zu 5,5% in Lagen mit einer hohen Anzahl von Arbeitsstätten, wie beispielsweise im 1. Bezirk in Wien. Das oberste Perzentil im Raster wurde auf den Wert des 99 Perzentil gesetzt, um die Extremwerte in dieser Datenschicht zu bereinigen. Der maximal vorkommende Wert im Raster ist somit 1,2 Arbeitsstätten pro Einwohner. Höhere Werte wurden den vorhandenen Immobilienpreisen bei ETW nicht zugeordnet und werden deshalb auch im Modell nicht geschätzt.

4.3.4.1.3 Anzahl der Wohngebäude

Der zweite etwas stärkere Zentrumsproxy ist die in Abbildung 4.6 dargestellte logarithmierte Anzahl der Wohngebäude aus dem Gebäuderaster. Dafür wurden die Variablen¹¹⁰ „gr_02_2_plus_etw_12“, „gr_03_wohngeb_gem_12“ mit dem Wert 1¹¹¹ addiert und logarithmiert.

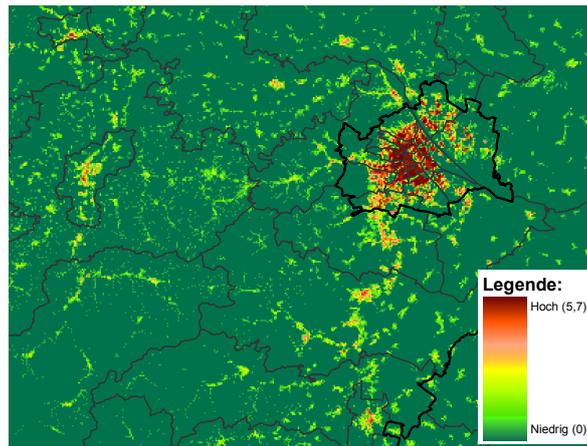


Abbildung 4.6: Anzahl der Wohngebäude

Legt man die höchsten Werte aus dem Raster um, kann es zu einer maximalen Wertsteigerung von 11,9 % kommen, wobei hohe Werte nur im Zentrum von Wien auftreten. In den restlichen Landeshauptstädten liegt das Maximum zwischen 8 % und 9,5 %. In den nicht so dicht bebauten Landeshauptstädten St. Pölten, Klagenfurt und Bregenz sind ein bis zwei Rasterzellen mit bis zu ca. 8 % zu finden. Die generelle Bebauungsdichte bei Wohngebäuden weist in diesen Städten in den meisten Zentrumslagen nur eine Wertsteigerung von ca. 7 % auf. Für Eisenstadt liegt die Wertsteigerung noch etwas niedriger, hier ist das Maximum in einer Rasterzelle bei 7,8 %, die restlichen Zentrumslagen sind durchwegs zwischen 6 % und 7 % angesiedelt.

Man kann also festhalten, dass ein mit Wohngebäuden dicht bebautes Zentrum, welches einem hohen Siedlungsdruck ausgesetzt ist, zu deutlich höheren Immobilienpreisen neigt als dünn besiedelte Gebiete. Es werden mit dieser Variable auch die Preisniveaunterschiede in den Zentren, die keine Landeshauptstädte sind, wie beispielsweise Wels,

¹¹⁰ Siehe dazu in Tabelle B.2.

¹¹¹ Der Logarithmus von 0 kann nicht berechnet werden.

Krems oder Villach etc. gut im Modell abgebildet, da natürlich auch dort dieser Zentrumseffekt auftritt. In kleineren Gemeinden wird die Zentrumslage auch abgebildet, dort ist allerdings der Effekt schwächer ausgeprägt, da in diesen üblicherweise keine so hohen Dichtewerte bei der Wohnbebauung erreicht werden.

4.3.4.1.4 Speckgürtel rund um Städte

Dies ist eine Dummyvariable, welche das direkte Umland von den Landeshauptstädten kennzeichnet und somit die gut erreichbaren und stadtnahen Wohngegenden, die innerhalb von 7 bis 15 Minuten erreichbar sind, ausweist. Erwähnenswert ist, dass sich dieser Speckgürtel-effekt in Wien anders verhält als in den restlichen Landeshauptstädten, da man sich mit der knappen Zeitschranke noch immer im Stadtgebiet befindet und somit die durch die Dummy gekennzeichneten gut erreichbaren Wohngegenden weitgehend innerhalb von Wien liegen. Im Vergleich mit allen anderen Wohngebieten weisen diese Speckgürtelwohnungen ein um 2,5 % höheres Preisniveau auf. Es wurde auch versucht, diese Variable für Wien separat zu schätzen, was allerdings zu einem nahezu identen Ergebnis geführt hat und diese Unterscheidung somit im Modell nicht vorkommt. Durch diese Variable wird bei optischer Überprüfung der geschätzten Werte auch sichtbar, dass der Übergang zwischen Stadt und Land (ausgenommen Wien) nun etwas weicher verläuft. Mehr dazu ist im Kapitel 5.1 zu finden.

4.3. FINALE MODELLE UND SCHLUSSFOLGERUNGEN

Family: gaussian

Link function: identity

Formula:

```
lnp ~ dum_angebot + dum_gemeinnutz + s(kp_year) + s(baujahr_bereich) +
  antwortid_bad_qual + bad_size_kat + zusatzbad + antwortid_zustand_haus +
  s(area_weight, by = bundes_AUT) + s(area_weight, by = bundes_WIEN) +
  s(zsp_am_lnetw_2012) + s(zsp_iu_lnetw_2012_q) + antwortid_zustand_wohnung +
  antwortid_heiz_etw + balc_vorh + lift_up + antwortid_garage_3 +
  bundes + br_her_so_ant + s(lp_ln_akademiker_ant) + br_arbst_01_pop_01 +
  ln_wohngeb + speckguertel + s(ra_laerm_all_log_sum) + s(ln_ra_ocl_1_min_1_q)
```

Parametric coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	7.862060	0.134024	58.662	< 2e-16 ***
dum_angebot	0.056399	0.012202	4.622	3.95e-06 ***
dum_gemeinnutz	-0.074376	0.020046	-3.710	0.000211 ***
antwortid_bad_qual1	0.079903	0.018663	4.281	1.91e-05 ***
antwortid_bad_qual3	-0.061455	0.010052	-6.114	1.09e-09 ***
bad_size_kat1	-0.021409	0.010135	-2.112	0.034731 *
zusatzbad	0.042128	0.016506	2.552	0.010749 *
antwortid_zustand_haus1	0.222877	0.044262	5.035	5.03e-07 ***
antwortid_zustand_haus2	0.160911	0.043876	3.667	0.000249 ***
antwortid_zustand_wohnung1	0.104120	0.011819	8.810	< 2e-16 ***
antwortid_zustand_wohnung3	-0.059177	0.012930	-4.577	4.90e-06 ***
antwortid_zustand_wohnung4	-0.169833	0.030705	-5.531	3.44e-08 ***
antwortid_heiz_etw2	-0.058186	0.015226	-3.821	0.000135 ***
balc_vorh	0.025805	0.008418	3.065	0.002192 **
lift_upTRUE	0.068286	0.009095	7.508	7.75e-14 ***
antwortid_garage_3	0.036542	0.009461	3.863	0.000114 ***
bundesBGL	0.397530	0.061627	6.451	1.28e-10 ***
bundesKTN	0.358565	0.033376	10.743	< 2e-16 ***
bundesNOE	0.455336	0.031826	14.307	< 2e-16 ***
bundesOOE	0.461525	0.033683	13.702	< 2e-16 ***
bundesSBG	0.483824	0.035523	13.620	< 2e-16 ***
bundesSTMK	0.400314	0.031663	12.643	< 2e-16 ***
bundesT	0.526170	0.032004	16.441	< 2e-16 ***
bundesVBG	0.505248	0.037601	13.437	< 2e-16 ***
br_her_so_ant	-0.173129	0.044278	-3.910	9.42e-05 ***
br_arbst_01_pop_01	0.044629	0.018099	2.466	0.013722 *
ln_wohngeb	0.019749	0.005055	3.906	9.56e-05 ***
speckguertel	0.024682	0.011001	2.244	0.024924 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Approximate significance of smooth terms:

	edf	Ref.df	F	p-value
s(kp_year)	4.022	4.953	86.736	< 2e-16 ***
s(baujahr_bereich)	4.748	5.763	95.685	< 2e-16 ***
s(area_weight):bundes_AUT	8.603	9.347	500.408	< 2e-16 ***
s(area_weight):bundes_WIEN	6.000	7.193	719.031	< 2e-16 ***
s(zsp_am_lnetw_2012)	8.496	8.922	17.592	< 2e-16 ***
s(zsp_iu_lnetw_2012_q)	6.450	7.688	3.234	0.001363 **
s(lp_ln_akademiker_ant)	3.064	3.831	6.427	5.19e-05 ***
s(ra_laerm_all_log_sum)	2.235	2.862	5.789	0.000758 ***
s(ln_ra_ocl_1_min_1_q)	2.176	2.770	4.533	0.004578 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) = 0.866 Deviance explained = 86.9%
GCV score = 0.043664 Scale est. = 0.0427 n = 3285
AIC = -963.31 BIC = -515.02

Abbildung 4.7: Finales Modell ETW

4.3.4.2 Nichtlineare Mikrolagevariablen (ETW)

Nachfolgend werden nun die als nichtlineare Terme im ETW-Modell verwendeten Mikrolagevariablen aufgelistet und erörtert.

4.3.4.2.1 Akademikeranteil

Im Gegensatz zum EFH-Modell wird bei ETW der logarithmierte Akademikeranteil als nichtlineare Variable verwendet, da hier der Verlauf der Funktion, wie man in Abbildung 4.8(g) erkennen kann, eindeutig nichtlinear ist. Bei Eigentumswohnungen ist gemäß dieser Grafik 4.8(g) bis zum Wert von 2, was 7,4 % Akademikeranteil bedeutet, ein minimaler Abwärtstrend erkennbar, welcher hier als Gerade interpretiert werden kann und somit bedeutet, dass keine Wertveränderung bei einem geringen Akademikeranteil ableitbar ist. Ausgehend von 2 bis zum Wert von 4,03, was mit 56,3 % das Maximum im Datensatz ist, steigt die Kurve stark an und zeigt, dass ein Akademikeranteil über 7,4 % eine deutliche Wertsteigerung für Eigentumswohnungen mit sich bringt. Geht man davon aus, dass die anfängliche Gerade bis zum Wert 2 auf der Y-Achse bei 0 liegt, so ist der höchste schätzbare Wert im Modell 11,6 % über diesem. Dies bedeutet, dass eine ETW um 11,6 % mehr Wert ist, wenn der Akademikeranteil in der Umgebung 56,3 % beträgt, was im Vergleich mit dem sehr geringen Effekt bei EFH von ca. 0,8 % bei diesem Anteil doch sehr hoch ist.

4.3.4.2.2 Lärm

Wie bereits im Kapitel 4.3.3.2.2 für EFH erklärt, kommt hier bei ETW die gleiche Lärmvariable zum Einsatz. Wie in Abbildung 4.8(h) ersichtlich, ist der Kurvenverlauf ähnlich wie schon zuvor bei EFH. Hier ist nun der gewünschte Effekt, dass bis zur störenden Dezibelhöhe von ca. 60 keine Wertverringerung bei den Immobilien nachweisbar ist, noch deutlicher erkennbar. Ab 60 dB beginnt sich der niedrigere Immobilienpreis abzuzeichnen und ab ca. 75 dB wird der Preisverfall noch deutlicher. Die Bandbreite zwischen leise (2 %) und laut (-7,2 %) ist mit 9,2 % geringer als beim EFH. Die Wertveränderung durch Lärmbeeinträchtigung kann allerdings, da im sehr leisen Bereich

unter 50 dB nur wenige Beobachtungen vorhanden sind, für beide Immobilientypen als ähnlich angesehen werden.

4.3.4.2.3 Erreichbarkeit

Für die Abbildung der Erreichbarkeit im Modell wird bei ETW im Gegensatz zum EFH nun die logarithmierte „raocl1_min_1“ Variable, welche die Entfernung zur nächsten Landeshauptstadt beschreibt, verwendet. Die Verwendung des bei EFH beschriebenen Zentralitätsindex funktioniert bei ETW nicht, da die Eigentumswohnungen wie in Abbildung 4.1 dargestellt, generell eher in zentralen Lagen angesiedelt sind.

Bei ETW sind gemäß der geschätzten Funktion der Erreichbarkeit bis zum Wert von 2, was umgerechnet 7 Fahrminuten sind, keine erwähnenswerten Auswirkungen auf den Immobilienpreis beobachtbar. Ab der 7 Minutengrenze – dort wo übrigens auch die Speckgürtelvariable¹¹² beginnt – fällt die Kurve nach unten ab und mit zunehmender Entfernung von den Landeshauptstädten werden die Eigentumswohnungen immer weniger wert. In diesem Datensatz ist die maximale Entfernung einer Wohnung zur Landeshauptstadt rund 76 Minuten, was dem Wert von 4,33 in der Grafik entspricht. Da gut erreichbare Wohnungen einen Wert von 1,5 % und die am schlechtesten erreichbare -6,3 % haben, ist der Wertebereich mit 7,8 % zwischen diesen zu beziffern. Der bei EFH auftretende Effekt, dass bei sehr schlecht erreichbaren Objekten der Wert nicht mehr weiter fällt, kann hier nicht beobachtet werden. Dies ist darauf zurückzuführen, dass in äußerst peripheren Gebieten mit sehr niedrigen Grundstückspreisen kaum Wohnungen errichtet werden und dementsprechend dort keine Beobachtungen für Wohnungen im Datensatz vorhanden sind.

¹¹² Siehe dazu in Kapitel 4.3.4.1.4.

KAPITEL 4. VALIDIERUNG DER MIKROLAGEVARIABLEN MIT EINEM GENERALISIERTEN ADDITIVEN MODELL (GAM)

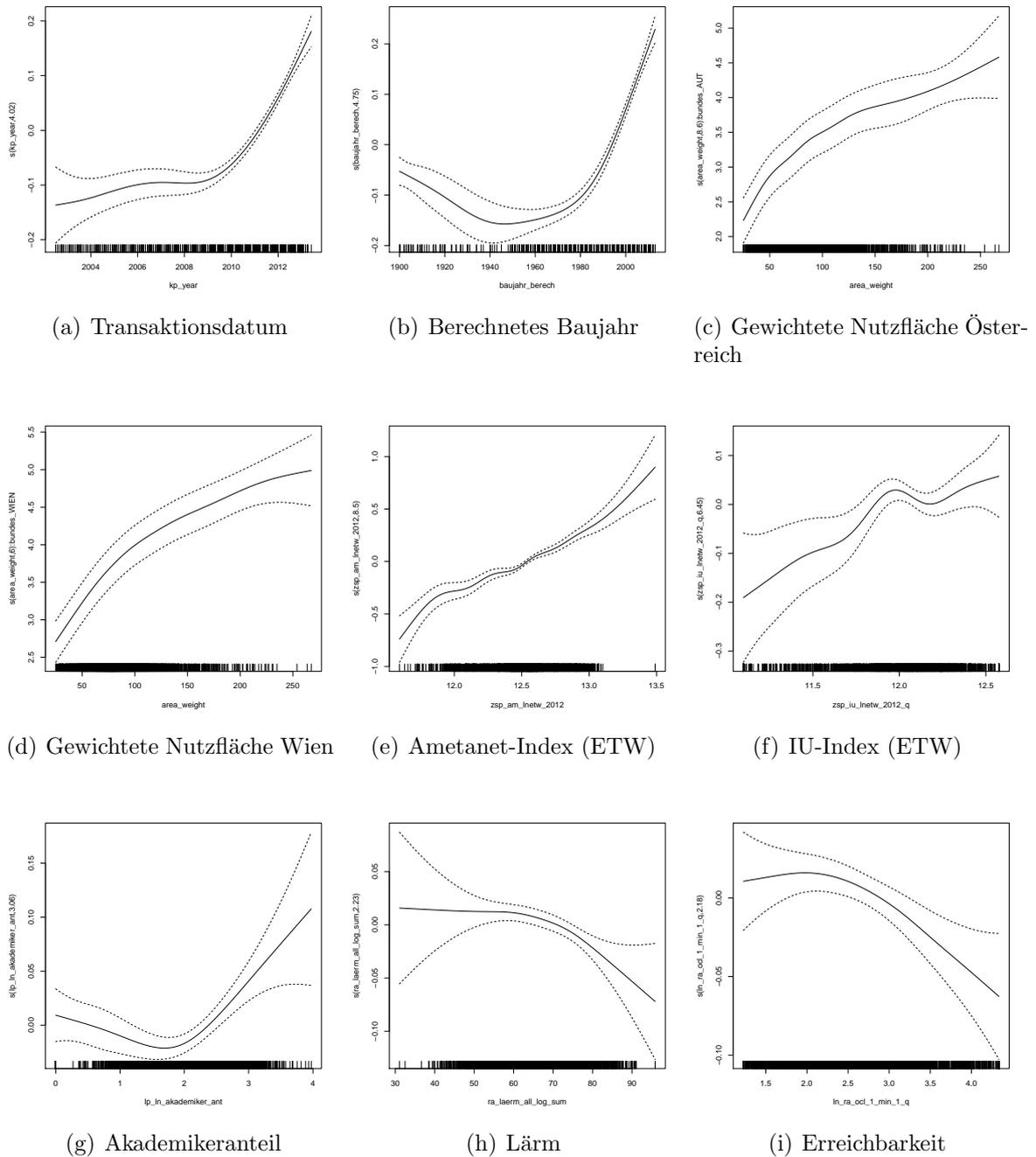


Abbildung 4.8: Nichtlineare Funktionen des finalen Modells (ETW)

4.3.4.3 Modellqualität ETW

Wie bereits zuvor für EFH im Kapitel 4.3.3.3 wird hier nun auch die Modellqualität für das ETW-Modell zur visuellen Überprüfung graphisch dargestellt. Vergleicht man die

se Modellcheckdiagramme (Abbildung 4.9(a)) mit jenem vom EFH (Abbildung 4.5(a)) kann man feststellen, dass die geschätzten Werte bei Wohnungen etwas besser mit den ursprünglichen korrelieren als bei Einfamilienhäusern, was bedeutet, dass die Schätzqualität des Wohnungsmodells etwas besser als jene des Hausmodells ist. Dieser Umstand ist darauf zurückzuführen, dass Wohnungen von der Konzeption grundsätzlich viel einheitlicher und auch einfacher aufgebaut sind als Häuser und dementsprechend auch genauer in einem derartigen Modell abgebildet werden können. Das gleiche Bild wiederholt sich natürlich auch beim Vergleich der Residuen mit der Schätzung (Prediction) in der zweiten Abbildung 4.9(b), auch hier streuen die geschätzten Werte bei ETW etwas weniger. Das Histogramm mit den Residuen (Abbildung 4.9(c)) ist auch in diesem Fall wieder sehr nahe an der Normalverteilung, was nochmals die Modellqualität unterstreicht.

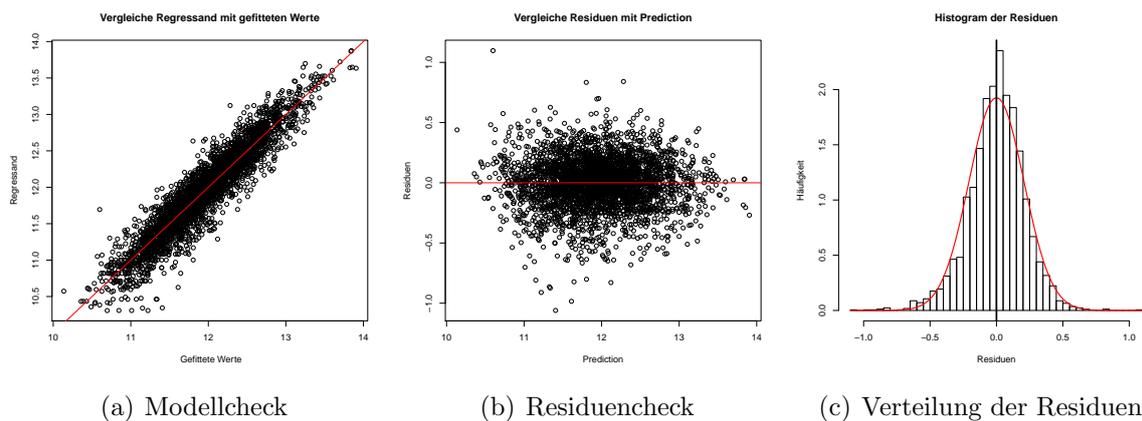


Abbildung 4.9: Modellqualität des finalen Modells (ETW)

4.4 Modellvergleiche

Abschließend werden die Vorteile der bereits in Kapitel 4.1 beschriebenen hybriden Modellierung in Kombination mit den in dieser Arbeit erstellten rasterbasierten Mikro-lagevariablen anhand von Modellvergleichen aufgezeigt. Alle in den nachstehenden Tabellen 4.3 und 4.4 angeführten Vergleichsmodelle sind im Anhang (Kapitel B.5) mit den entsprechenden Regressionsoutputs und Funktionskurven der nichtlinearen Terme zu finden.

Grundsätzlich kann man bei EFH und ETW feststellen, dass sich beide Modelle durch die Verwendung der hybriden und der Rastervariablen qualitativ deutlich verbessern. Bei Betrachtung von AIC und BIC kann man erkennen, dass der modellverbessernde Effekt durch die hybriden Variablen etwas stärker als bei den Rastervariablen ist. Dies ist darauf zurückzuführen, dass durch die zwei hybriden Variablen¹¹³ die räumliche Variation des Preisniveaus beschrieben wird, und ohne diese Variablen nur der „breite“ LIEBE-Datensatz herangezogen wird und dadurch die Schätzung mit einer geringeren räumlichen Genauigkeit auskommen muss. Zieht man zum Vergleich zusätzlich noch R^2 und EDEV heran, wird nochmals bestätigt, dass sich die Modelle massiv verbessern.

4.4.1 EFH

Wie bereits zuvor im Kapitel 4.3.4.3 erwähnt, ist das ETW-Modell im Vergleich mit EFH bezüglich der Schätzqualität etwas besser. Dieser Umstand wird durch die Vergleiche der Modellergebnisse, in den nachstehenden Tabellen 4.3 und 4.4 bestätigt, wie man anhand von R^2 und EDEV erkennen kann.

Durch die (Raster-) Mikrolagevariablen verbessert sich das EFH-Modell allerdings deutlicher als ETW, da bei diesem die unterschiedlichen Lageeffekte weiter gestreut sind und dadurch ein größeres Verbesserungspotential vorhanden ist. Die maximale Verbesserung vom Basismodell ohne hybride Variablen und Rastervariablen bis zum finalen Modell beträgt bei EFH 2749 AIC bzw. 0,16 R^2 Punkte.

Modellvarianten	AIC	BIC	R^2	EDEV	GFG
nicht hybrid, ohne Raster	1656.7116	1898.2786	0.6829	0.6859	37.6421
nicht hybrid, mit Raster	113.6271	457.1957	0.7889	0.7918	53.9586
hybrid, ohne Raster	-977.8140	-655.1863	0.8410	0.8431	50.6089
hybrid, mit Raster (final)	-1092.2454	-731.4511	0.8460	0.8482	56.7141

Tabelle 4.3: Modellvergleich EFH

¹¹³ Siehe dazu Kapitel 4.1.1.

4.4.2 ETW

Obwohl das ETW-Modell bereits ohne hybride Variablen bzw. Rastervariablen ein hohes R^2 aufweist, kann mittels der neuen Variablen im Modell noch eine weitere Verbesserung erzielt werden. Im Vergleich zu EFH bringen die Rastervariablen hier eine nicht ganz so starke Verbesserung beim Modellvergleich. Wenn man allerdings die Regressionsoutputs der verschiedenen Varianten vergleicht, kann man schnell feststellen, dass die finale Modellspezifikation dennoch die beste ist. Wie auch schon bei EFH angeführt, beträgt die Differenz zwischen dem Basismodell ohne hybride Variablen und Rastervariablen bis zum finalen Modell beim AIC 1417 und beim R^2 0,074 Punkte.

Modellvarianten	AIC	BIC	R^2	EDEV	GFG
nicht hybrid, ohne Raster	454.0670	760.9987	0.7922	0.7952	49.3404
nicht hybrid, mit Raster	-322.5290	74.6690	0.8366	0.8398	64.1452
hybrid, ohne Raster	-884.3817	-505.5632	0.8622	0.8647	61.1307
hybrid, mit Raster (final)	-963.3109	-515.0190	0.8659	0.8688	72.5252

Tabelle 4.4: Modellvergleich ETW

4.4.3 Allgemeines zum Modellvergleich

Grundsätzlich kann man zu den unterschiedlichen Modellen also festhalten, dass die hybriden Variablen ein etwas höheres Verbesserungspotential als die Rastervariablen aufweisen, wenn man die Differenz bei der Modellqualität zwischen „nicht hybrid, mit Raster“ und „hybrid, ohne Raster“ bezogen auf das Ursprungsmodell („nicht hybrid, ohne Raster“) vergleicht. Die Mikrolagevariablen bringen dem Modell, mit den auf Zählsprenkel basierenden hybriden Variablen, allerdings noch eine weitere Detaillierungsstufe und schaffen es somit noch einige bis dahin unberücksichtigte Effekte wie beispielsweise Lärm oder Ausländeranteil durch das Modell erklären zu lassen. Insgesamt führt also die Rastertechnik in Kombination mit den hybriden Variablen zu einem sehr guten und konsistenten hedonischen Preismodell.

Kapitel 5

Schlussbetrachtungen

In diesem abschließenden Kapitel werden die Modellergebnisse visuell dargestellt und die Vorteile der Verwendung von Rastervariablen in hedonischen Preismodellen diskutiert (Kapitel 5.1). Des Weiteren werden noch einige Informationen über die softwaretechnische Verwendung dieses Modells in der Praxis angeführt (Kapitel 5.2.1), sowie mögliche weitere Entwicklungsschritte diskutiert (Kapitel 5.2).

5.1 Zusammenfassung der Ergebnisse

Für die graphische Darstellung der Modellergebnisse wird für die zwei Objektarten EFH und ETW ein vordefiniertes Objekt herangezogen und dann über alle Standorte in ganz Österreich simuliert. Die genaue Spezifikation dieser beider Simulationsobjekte ist in der nachstehenden Tabelle 5.1 angeführt.

Objekteigenschaft (Variablennamen) ¹¹⁴	EFH	ETW
Berechnetes Baujahr (baujahr_berech)	1993 (20 J.)	1993 (20 J.)
Gewichtete Nutzfläche (area_weight)	180m ²	85m ²
Heizungstyp (antwortid_heiz)	Zentralheizung	Zentralheizung
Ausstattung Bad (antwortid_bad_qual)	Mittel	Mittel
Kleines Badezimmer (bad_size_kat1)	Nein	Nein
Zusätzliche Badezimmer (zusatzbad)	Ja	Nein
Zustand Haus (antwortid_zustand_haus)	Mittel	Mittel
Zustand Wohnung (antwortid_zustand_wohnung)		Mittel
Balkon (balc_vorh)		Ja
Liftinteraktion (lift_upTRUE)		Nein
Garage (antwortid_garage_3)		Ja
Grundstücksfläche (area_gst)	650m ²	
Reihenhaus (rh)	Nein	
Terrasse (terr_wohn_vorh)	Ja	

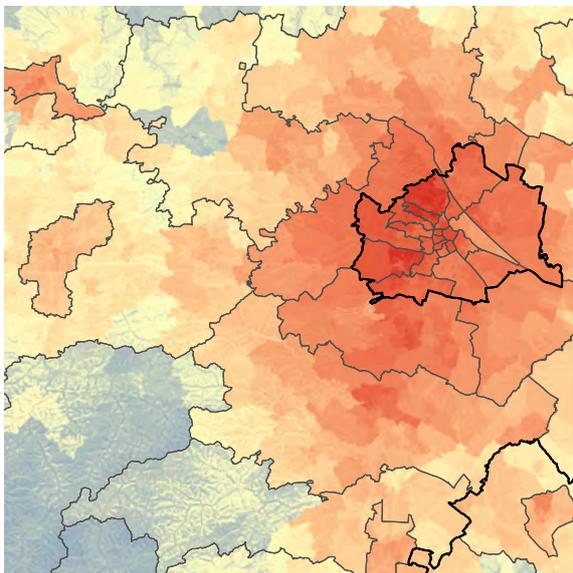
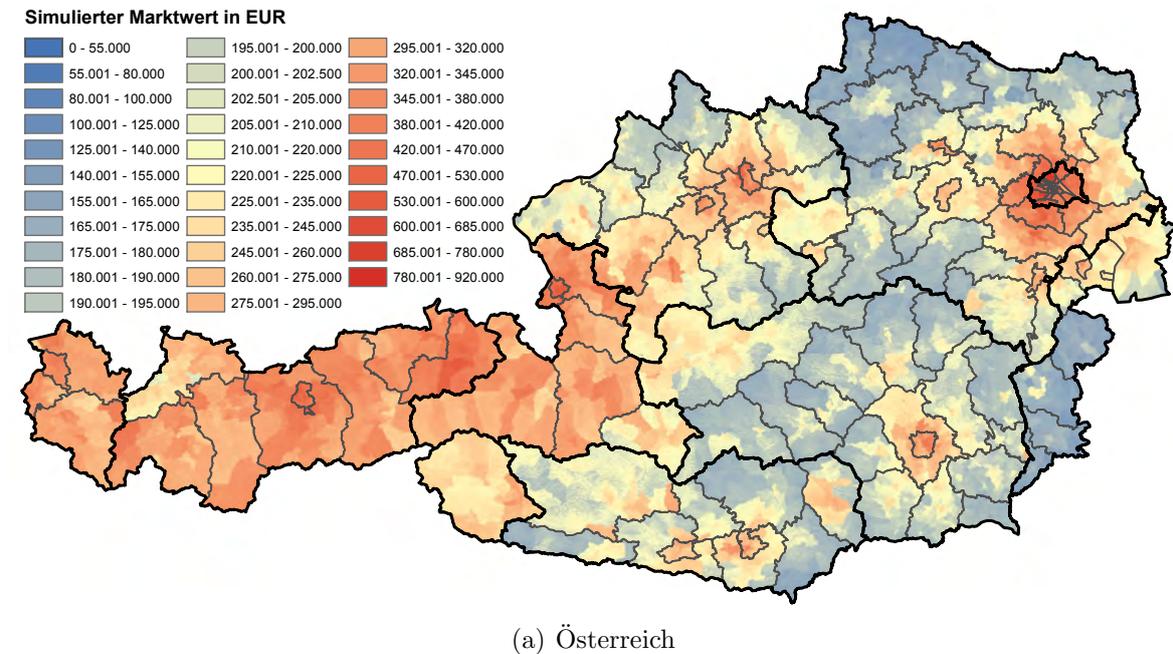
Tabelle 5.1: Simualtionsobjekte

Verwendet man für eine derartige Simulation keine Rastervariablen, kann man in der Statistiksoftware R die Schätzergebnisse mittels der eingebauten Funktion ausgeben lassen. Da in dieser Arbeit aber auch die Werte der Mikrolagevariablen im Rasterformat miteinbezogen werden müssen, kommt ein eigens in R und Python (ArcPy) programmiertes dreistufiges Verfahren zur Anwendung. Hierbei werden zuerst die Variablenwerte auf Zählsprengebasis summiert und dann die Rastervariablen dazugezählt. Im dritten Schritt werden dann die definierten Objekteigenschaften für die Simulation hinzuaddiert. Man erhält hierdurch die Gesamtschätzung basierend auf allen unterschiedlichen Datenbasisniveaus (Zählsprenge – Raster – Objekt). Da bei nichtlinearen Termen Werte vorkommen können, die sich außerhalb der Schätzung befinden, wurden für diese Werte die Funktionsverläufe anhand der Splinebasen extrapoliert. Alle Werte, die um mehr als 10 % über den geschätzten Wertebereich hinaus extrapoliert wurden, sind dann bei diesem 10 % Wert gedeckelt worden, um keine Verzerrungen im Modell zu erhalten. Betroffen davon ist nur die Variable Akademikeranteil bei ETW.

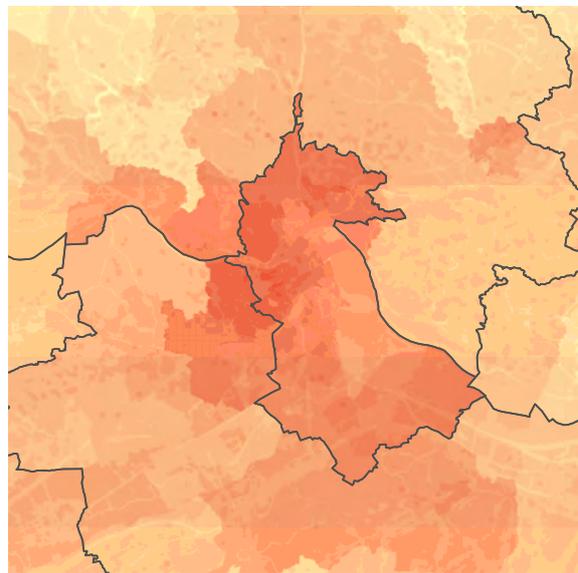
¹¹⁴ Weitere Informationen zu den Variablen sind in Tabelle 4.2 zu finden.

5.1.1 Modell EFH

Die Ergebnisse dieser Simulation für EFH sind in Abbildung 5.1 mit 3 verschiedenen Zoomstufen dargestellt.



(b) Wien und Umland



(c) Linz und Umland

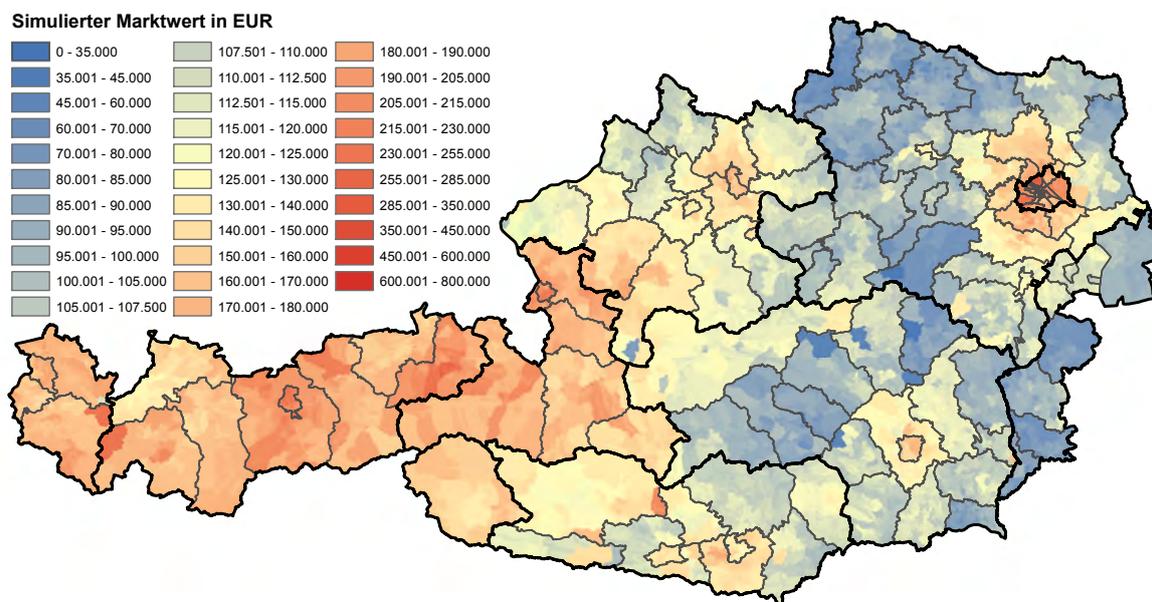
Abbildung 5.1: Simuliertes EFH

Die in der Österreichkarte (Abbildung 5.1(a)) dargestellte Legende gilt auch für die Zooms auf Wien (Abbildung 5.1(b)) und Linz (Abbildung 5.1(c)). Das bekannte Preisgefälle zwischen Ost- und Westösterreich ist in dieser Simulation deutlich erkennbar. Demnach sind neben den Landeshauptstädten und Wien die höchsten Immobilienpreise in Vorarlberg und Tirol bzw. Teilen von Salzburg zu finden.

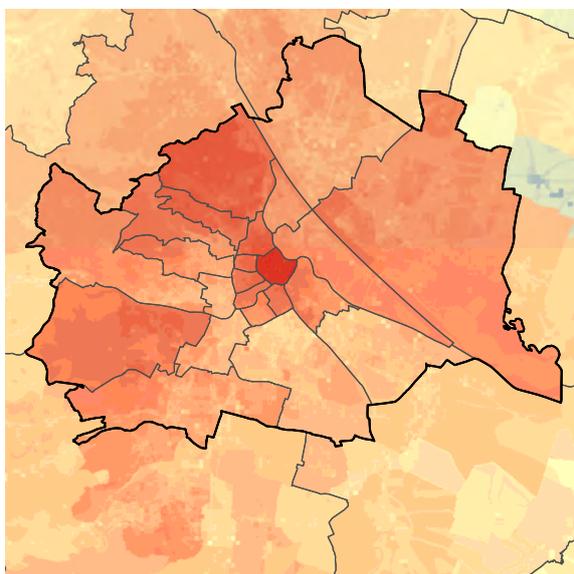
Betrachtet man die zwei Zooms auf Wien und Linz etwas genauer, kann man feststellen, dass in den Preisniveaus auch die Mikrolagevariablen deutlich erkennbar sind. Beispielsweise ist im südwestlichen Teil des Wienausschnittes deutlich die durch die Konvex-Konkav-Variable entstandene Abbildung der Geländestruktur zu erkennen, und es sind demnach in engen Tallagen niedrigere Immobilienpreise vorhanden. Teilweise sind auch hochrangige Straßen- und Bahntrassen erkennbar, welche indirekt durch die Lärmvariable abgebildet werden. Allgemein kann man beim Zoom auf Wien auch noch feststellen, dass die Zählsprengelgrenzen in der Preisstruktur kaum mehr erkennbar sind und die Übergänge zwischen diesen aufgrund der neuen Rastervariablen sehr fließend geworden sind. Lediglich einige wenige Gemeinde- und Bezirksgrenzen sind in den Preisniveaus aufgrund der Datenstruktur in den hybriden Variablen noch erkennbar, aber auch diese Übergänge wurden im Vergleich mit dem Modell ohne Rastervariablen deutlich fließender. Betrachtet man die Niveaus innerhalb von Wien, kann man feststellen, dass die hochpreisigen Immobilienlagen in den westlichen Außenbezirken in Wien (18., 19. und teilweise der 16. und 17.) sowie im 13. Bezirk deutlich anhand der dunkelroten Farbe (hohe Preise) erkennbar sind. Da der Ausschnitt von Linz einen kleineren Maßstab aufweist, können in diesem bereits deutlich die wichtigsten Straßenzüge und Bahnstrecken, welche Lärm emittieren, abgelesen werden. Durch den höheren Lärm des am östlichen Rand von Linz gelegenen Voestgeländes wird auch dieses Gebiet im Vergleich mit den umliegenden flächendeckend abgewertet. Generell gilt auch hier wieder die Aussage, dass die Übergänge durch die Rastertechnik fließender werden. Des Weiteren kann man hier auch Mikrolagen (z.B.: nordöstlich der Linzer Stadtgrenze), die kleiner als Zählsprengel sind beobachten, welche in sich geschlossen höhere oder niedrigere Werte aufweisen – sprich: es werden durch das Modell „gute“ und „schlechte“ kleinräumige Lagen erkannt.

5.1.2 Modell ETW

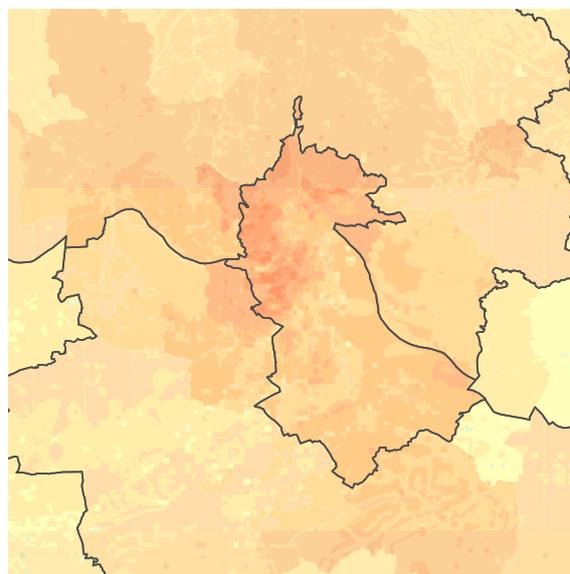
In der nachstehenden Abbildung 5.2 ist die Schätzung aus dem ETW-Modell ebenfalls für die drei gleichen Standorte dargestellt.



(a) Österreich



(b) Wien



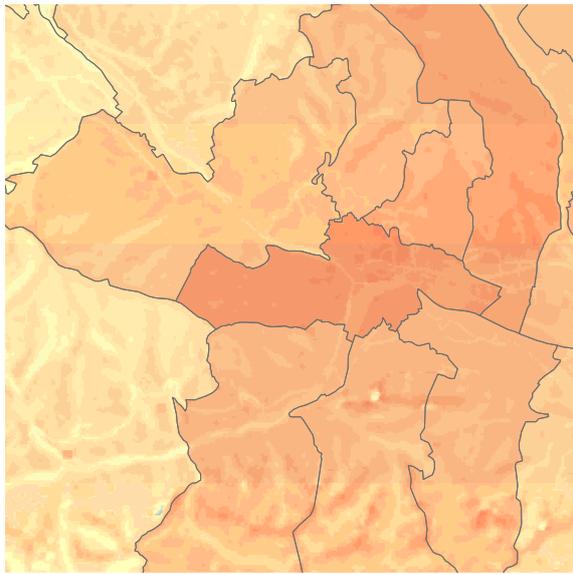
(c) Linz und Umland

Abbildung 5.2: Simuliertes ETW

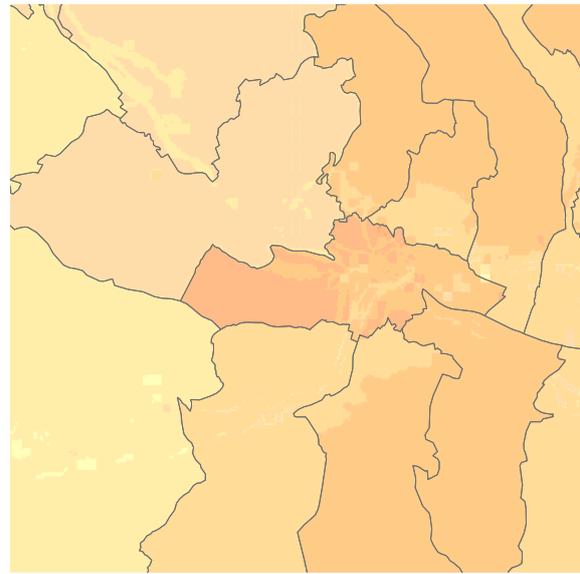
Wie schon zuvor bei Einfamilienhäusern ist beim Zoom auf ganz Österreich das Ost-West-Preisgefälle deutlich zu erkennen. Die niedrigsten Immobilienpreise sind hier wie auch beim EFH-Modell im nördlichen Waldviertel und im südlichen Burgenland sowie in peripheren Lagen in der Steiermark zu finden. Beim Zoom auf Wien (Abbildung 5.2(b)) sind neben den teureren Wohnlagen in den inneren Bezirken auch – wie schon bereits beim Einfamilienhaus – die Toplagen in den westlichen Außenbezirken gut erkennbar. Interessant an dieser Darstellung ist, dass hier nun durch die Verwendung des Ausländeranteils als Rastervariable auch die Zonen mit hohem Ausländeranteil (10., 11., 15. und Teile des 16. Bezirks) deutlich anhand des Preisniveaus identifiziert werden können. Auch beim Linzer Beispiel kann man wieder gut erkennen, dass die Schätzung auch in Hinsicht auf die räumliche Verteilung des Preisniveaus ein homogenes Bild mit fließenden Übergängen liefert.

5.1.3 Modell EFH vs. ETW

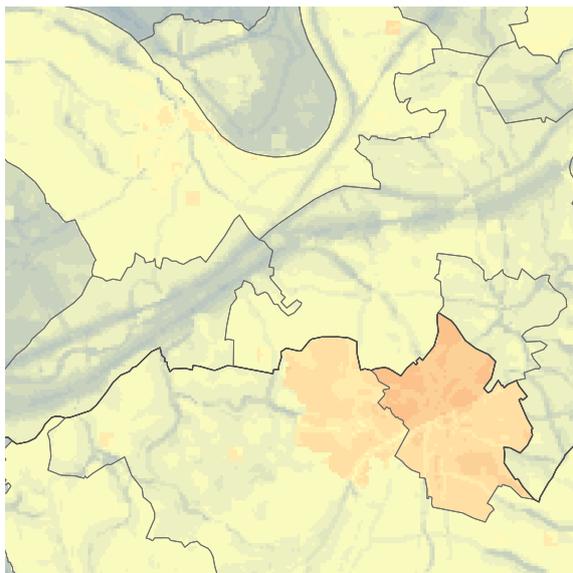
Abschließend sind in Abbildung 5.3 jeweils EFH und ETW nebeneinandergestellt, um auch die Unterschiede zwischen den zwei Modellen zu verdeutlichen. Es ist hier gut erkennbar, dass beispielsweise die Konvex-Konkav Variable bei ETW nicht zum Einsatz kommt, da die Struktur des Geländes nur beim Hausmodell erkennbar ist. Dass die Straßen bei ETW im Vergleich mit EFH nicht so gut erkennbar sind, liegt daran, dass die Erreichbarkeitsvariable bei ETW einen deutlich geringeren Einfluss hat und dadurch dieser Effekt nicht so stark ist wie bei EFH. Der Parameter der Lärmvariable hingegen ist bei beiden Modellen ähnlich, was zur Folge hat, dass die lärmemittierende höherrangige Verkehrsinfrastruktur auch im ETW-Modell abgebildet wird. Die größeren Rasterkacheln bei ETW in Abbildung 5.3(d) entstehen durch die Variablen Akademikeranteil, Ausländeranteil, Arbeitsstätten und Wohngebäude, welche als Raster mit 250m Auflösung im Modell verwendet werden. Vereinzelt kommt dies auch beim EFH-Modell vor. Dort wird allerdings nur der Akademiker- und Ausländeranteil mit dieser groben Auflösung verwendet, alle anderen Rastervariablen haben eine 50m Auflösung.



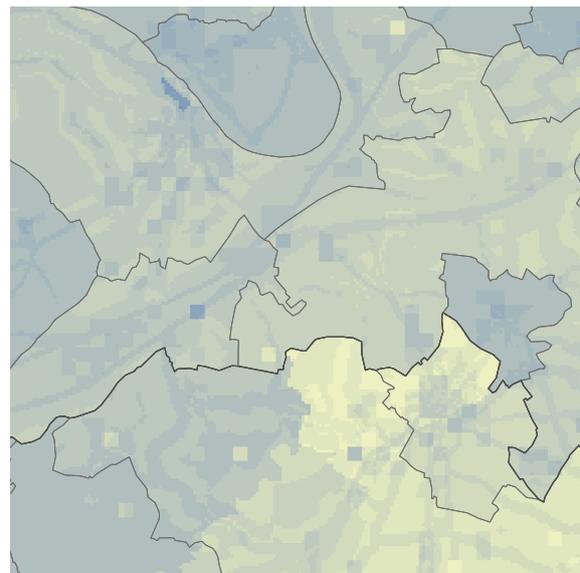
(a) Lienz EFH



(b) Lienz ETW



(c) Wieselburg EFH



(d) Wieselburg ETW

Abbildung 5.3: Vergleiche simuliertes EFH mit ETW

5.2 Diskussion und Ausblick

Wie diese Arbeit verdeutlicht, birgt die Verwendung von hochauflösenden Raster-Mikrolagevariablen als neue Modellierungsebene in statistisch-räumlichen Analysemodellen ein erhebliches Verbesserungspotential. Es können nun Lagewerte auch unter Einbeziehung modellierter Datenbasen wie etwa dem Lärm bis auf eine 50m Genauigkeit geschätzt werden. Die Verwendung von solch hohen Auflösungen bei Rasterlagevariablen wurde bislang in einer derart großen Analyseregion wie Österreich vermutlich aus EDV-Performancegründen noch nicht angestellt. Die heute zur Verfügung stehende Rechenleistung spielt mittlerweile selbst bei Berechnung von großen Analyseregionen, nur mehr eine untergeordnete Rolle. Da auch bereits die großen Datenanbieter wie z.B. Statistik Austria auf die Erfassung von Daten im Rasterformat setzten, wird es in Zukunft noch besser möglich sein, Mikrolagevariablen für diverse Modellrechnungen zu verwenden. Die in dieser Arbeit gezeigten Möglichkeiten der Modellierung von Mikrolagevariablen in einem GIS können als erster großer Schritt in die Richtung einer automatisierten mikrolagengenauen Modellschätzung gesehen werden. Bezogen auf Mikrolagefaktoren und deren Modellierung in der Immobilienbewertung ist das Potential aber noch lange nicht ausgeschöpft, und es wird auch zukünftig Weiterentwicklungen geben um die Mikrolageeffekte noch besser erklären zu können.

Einige im Zuge dieser Arbeit erstellten Mikrolagevariablen, wie beispielsweise die Wohnlage an einem Park oder die Lage an einem im Zentrum gelegenen Platz wurden bei der Variablenselektion, wie die finalen Modelle zeigen, bislang nicht berücksichtigt. Dies ist darauf zurückzuführen, dass diese Eigenschaften bei einer Schätzung zwar vom Niveau her gute Ergebnisse liefern, aber aufgrund der geringen Beobachtungszahl im breiten LIEBE-Datensatz die Variablensignifikanz noch nicht ausreicht, um diese im Modell zu verwenden. Da dieses Modell laufend mit neuen Daten befüllt wird, ist aber für die Zukunft zu erwarten, dass mit steigender Anzahl an Beobachtungen auch die Anzahl der verwendeten Mikrolagevariablen im Modell steigen wird.

5.2.1 Praktische Verwendung der Ergebnisse in einer Bewertungssoftware

Die in dieser Arbeit ermittelten Mikrolagefaktoren werden in der Praxis in ein hedonisches Preismodell eingegliedert, das noch über viele weitere Objektkriterien verfügt. Dieses Modell verwendet neben den in dieser Arbeit vorgestellten Methoden auch noch die in BRUNAUER 2010 S.43 ff und BRUNAUER, LANG und FEILMAYR 2013 beschriebene hierarchische Version von Strukturiert Additiven Regressionsmodellen (Strukturiert Additive Regressionsmodelle (STAR)) und ist – was die Schätzgenauigkeit betrifft – noch etwas besser.

In der UniCredit Group der Bank Austria AG wurde für die Sicherheitenbewertung bei Hypothekarkrediten im Zuge des LIEBE Projektes eine Software entwickelt, welche auf Basis eines hedonischen Preismodells funktioniert. Diese Software ermöglicht es, die in dieser Arbeit beschriebene Rastertechnik, sowie auch Interaktionseffekte, unterschiedliche Flächengewichte und differenzierte Zustandsnoten und deren jeweilige Parameter in einem Rechenmodell nachzubilden. Für Modellaktualisierungen oder Methodenumstellungen wird der Rechenkern der Software über diverse Steuertabellen neu befüllt, und es ist somit möglich, ohne weiteren Programmieraufwand laufende Modellupdates in die Software einzuspielen. Diese neue Software wird seit Mitte 2012 von speziell ausgebildeten Immobilienbewertern im Echtbetrieb in der Bank verwendet.

Anhang A

Verzeichnisse

A.1 Abkürzungsverzeichnis

AIC Akaike Informationskriterium

ASTER Advanced Spaceborne Thermal Emission and Reflection Radiometer

BEV Bundesamt für Eich- und Vermessungswesen

BEZ Bezirk

BIC Bayesianische Informationskriterium

DHM Digitales Höhenmodell

DM Data Management

DTV Durchschnittliche tägliche Verkehrsstärke

EDEV Erklärte Devianz

EFH Einfamilienhaus

ETW Eigentumswohnung

GFG Geschätzte Freiheitsgrade

FGDB File-Geodatabase

GAM Generalisierte Additive Modelle

GDEM Global Digital Elevation Map

GEM Gemeinde

GIS Geographisches Informationssystem

GNU General Public License

IU IMMOUnited
LAM Lambert Austria
LAT Latitude (Breitengrad)
LIEBE Liegenschaftsbewertung
LISA Land Information System Austria
LON Longitude (Meridian/Längengrad)
METI Ministry of Economy, Trade, and Industry
MIV Motorisierter Individualverkehr
NASA National Aeronautics and Space Administration
NK Nutzungskey
OG Obergruppe
ÖPNV Öffentlicher Personennahverkehr
OSM OpenStreetMap
SA Spatial Analyst
STAR Strukturiert Additive Regressionsmodelle
WGS84 World Geodetic System 1984
WKO Wirtschaftskammer Österreich
ZAMG Zentralanstalt für Meteorologie und Geodynamik
ZSP Zählsprengel

A.2 Literaturverzeichnis

- BAFU, BUNDESAMT FÜR UMWELT (2008): *Grundlagen Industrie- und Gewerbelärm*. Eidgenössisches Departement für Umwelt, Verkehr, Energie und Kommunikation (UVEK).
- BARTIK, T.J. und V.K. SMITH (1987): *Urban amenities and public policy. Handbook of Regional and Urban Economics*. Edited by E.S. Mills, Vol. 2, Chapter 31, S.271-318. Elsevier Science Publishers B.V.
- BRUNAUER, WOLFGANG A. (2009): *Spatially autoregressive hedonic regression models: an application for real estate in Vienna. Diplomarbeit*.
- (2010): *Modeling Real Estate Data with Structured Additive Regression Models. Dissertation*.
- BRUNAUER, WOLFGANG A., STEFAN LANG und WOLFGANG FEILMAYR (2013): *Hybrid multilevel STAR models for hedonic house prices. Jahrbuch für Regionalwissenschaft*. Springer-Verlag Berlin Heidelberg. DOI 10.1007/s10037-013-0074-9.
- BURROUGH, PETER A. und RACHAEL A. McDONNELL (1998): *Principles of Geographical Information Systems (Spatial Information Systems and Geostatistics)*. Oxford University Press. ISBN: 9780198233657.
- CHRISTALLER, WALTER (1980): *Die zentralen Orte in Süddeutschland: Eine ökonomisch-geographische Untersuchung über die Gesetzmäßigkeit der Verbreitung und Entwicklung der Siedlungen mit städtischen Funktionen. Ausgabe 3*. Wissenschaftliche Buchgesellschaft. ISBN: 3534044665.
- COURT, ANDREW T. (1939): *Hedonic price indexes - With automotive examples. The dynamics of automobile Demand*. General Motors Corporation (S.99-120). New York.
- EDLINGER, STAUDE und STOCK (2003): *Lawinenschutz. VU Nutzungskonflikte - Fachgeographische Übung*. Education Group.
- FAHRMEIR, LUDWIG, ALFRED HAMERLE und GERHARD TUTZ (1996): *Multivariate statistische Verfahren. 2. erweiterte Auflage*. Walter de Gruyter & Co., Berlin. ISBN: 3110138069.
- FAHRMEIR, LUDWIG, THOMAS KNEIB und STEFAN LANG (2007): *Regression. Modelle, Methoden und Anwendungen*. Springer-Verlag Berlin Heidelberg. ISBN: 9783540339328.
- FALLAST, KURT und GERD SAMMER (1994): *Grazer Lärmkataster. „Verkehrslärm“ für das Hauptstraßennetz*. Magistrat Graz - Amt für Umweltschutz - Referat für Lärmbekämpfung und Schallschutz.

- FREEMAN, A.MYRICK. III (2003): *The Measurement of Environmental and Resource Values. Theory and Methods*. Johns Hopkins University Press (RFF Press Book). Second Edition. Washington, DC: Resources for the Future. ISBN: 1891853635.
- GRILICHES, ZVI (1961): *Hedonic Price Indexes for Automobiles: An Econometric of Quality Change. Price Indexes and Quality Change*. The Price Statistics of the Federal Government. Report of the Price Statistics Review Committee. National Bureau of Economic Research und University of Chicago. Staff Paper 3. (S. 173 - 196). ISBN: 0870140728.
- HAMILTON, LAWRENCE C. (2009): *Statistics with STATA. Updatet for Version 10*. Brooks/Cole, Cengage Learning. ISBN: 9870459557869.
- HÄUPL, PETER (2008): *Bauphysik. Klima, Wärme, Feuchte, Schall ; Grundlagen, Anwendungen, Beispiele*. Ernst und Sohn Verlag für Architektur und technische Wissenschaften GmbH und Co. KG. ISBN: 9783433018422.
- KALIVODA, MANFRED T. und MONIKA BUKOVNIK (2010): *Verkehr und Umwelt. (Vorlesungsunterlagen)*. Department Raum, Landschaft und Infrastruktur (Universität für Bodenkultur).
- KOHLER, ULRICH und FRAUKE KREUTER (2008): *Datenanalyse mit Stata. Allgemeine Konzepte der Datenanalyse und ihre praktische Anwendung. 3.Auflage*. Oldenbourg Wissenschaftsverlag GmbH. ISBN: 9783486584561.
- KRAWIETZ, RHENA und WILFRIED HEIMKE (Dez. 2007): *Physik im Bauwesen. Grundwissen und Bauphysik*. München: Hanser Fachbuchverlag, S. 244. ISBN: 3446402764.
- LIGGES, UWE (2008): *Programmieren mit R. Dritte, überarbeitete und erweiterte Auflage*. Springer-Verlag Berlin Heidelberg. ISBN: 9783540799979.
- LUTZ, MARK (2006): *Programming Python. Powerful Object-Oriented Programming. 3rd Edition*. O'Reilly Media, Inc. ISBN: 0596009259.
- MITCHELL, ANDY (1999): *The ESRI Guide to GIS Analysis. Volume 1: Geographic Patterns & Relationships*. ESRI Press - Redlands, California. ISBN: 9781879102064.
- (2009): *The ESRI Guide to GIS Analysis. Volume 2: Spatial Measurements & Statistics*. ESRI Press - Redlands, California.
- MOORE, I.D., R.B. GRAYSON und A.R LADSON (1991): *Digital Terrain Modelling: A Review of Hydrological, Geomorphological and Biological Applications*. Hydrological Processes, Vol. 5, 3-30.
- MÜLLER, GERHARD und MICHAEL MÖSER (HRSG.) (2003): *Taschenbuch der Technischen Akustik. 3., erweiterte und überarbeitete Auflage*. Springer-Verlag Berlin Heidelberg. ISBN: 3540412425.

- PALMQUIST, RAYMOND B. (1991): *Hedonic Methods*. Edited by J. B. Braden und C.D. Kolstad. Measuring the Demand for Environmental Quality. Amsterdam: North-Holland, S. 77-120.
- (1999): *Hedonic Models. Handbook of Environmental and Resource Economics*. Edited by Jeroen C.J.M. van den Bergh. Edward Elgar Publishing (Chapter 53, S.765-776). ISBN: 9781858983752.
- RIEDL, LEOPOLD und ROBERT KALASEK (2008): *GIS - technische und methodische Grundlagen. Einheit 4, RasterGIS 1, Digitales Geländemodell, Map Algebra und RasterCalculator*. Fachbereich für Stadt- und Regionalforschung (Technische Universität Wien).
- ROSEN, SHERWIN (1974): *Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition*. Journal of Political Economy, Vol. 82, No. 1, Jan. - Feb., . S. 34-55. The University of Chicago Press.
- SEGER, MARTIN (1999): *Landinformationssystem Österreich - ein neuer Datensatz für regionale Raumordnungsfragen. Institut für Geographie (Universität Klagenfurt)*. 4. Symposium „Computergestützte Raumplanung“ CORP 99.
- WEBERNDORFER, RONALD (2009): *Auswirkungen von Naturgefahren auf die Immobilien- und Grundstückspreise. 1.Auflage*. IGEL Verlag GmbH. ISBN: 9783868151077.
- WEGSCHEIDER-PICHLER, ALEXANDRA (2009): *Umweltbedingungen, Umweltverhalten - Ergebnisse des Mikrozensus 2007*. Wien: Statistik Austria - Bundesanstalt Statistik Österreich. ISBN: 9783902587961.
- WILLEMS, WOLFGANG M., KAI SCHILD und SIMONE DINTER (2006): *Handbuch Bauphysik Teil 2. Schall- und Brandschutz, Fachwörterbuch deutsch-englisch, englisch-deutsch*. Vieweg und Sohn Verlag | GWV Fachverlage GmbH. ISBN: 3834801887.
- WOOD, SIMON N. (2006a): *Generalized Additive Models: An introduction with R*. Chapman und Hall/CRC Press. ISBN: 9781584884743.
- (2006b): *Low-Rank Scale-Invariant Tensor Product Smooths for Generalized Additive Mixed Models*. Biometrics 62, 1025-1036.
- (2006c): *On confidence intervals for GAMs based on penalized regression splines*. Australian und New Zealand Journal of Statistics 48, 445-464.
- ZEVEBERGEN, L.W. und C.R. THORNE (1987): *Quantitative Analysis of Land Surface Topography*. Earth Surface Processes und Landforms 12, 47-56.

A.3 Internet- und sonstige elektronische Quellen

APPCELERATOR, INC. (2010): *What is PyDev?* URL: <http://pydev.org/> (besucht am 01.02.2010).

BING MAPS (2012): *Microsoft Corporation.* URL: <http://www.bing.com/maps/?cc=at> (besucht am 25.10.2012).

BÖKEMANN, DIETER und WOLFGANG FEILMAYR (2005): *Kleinräumige Analyse von Immobilienpreisen. Immobilienindizes und Immobilienbewertung.* URL: <http://www.srf.tuwien.ac.at/feil/hedo.pdf> (besucht am 23.01.2008).

ESRI (2012): *ArcGIS Resource Center. Pfad: Professional Library / Geoprocessing / Geoprocessing tool reference / ...* URL: <http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html> (besucht am 06.06.2012).

FORUM SCHALL (2006): *Emissionsdatenkatalog. Modellierung und Dokumentation.* URL: http://www.umweltbundesamt.at/fileadmin/site/umwelthemen/laerm/forum_schall/downloads/Emissionsdatenkatalog_11_2006.pdf (besucht am 15.05.2012).

GNU R (2012): *Wikibooks, Die freie Bibliothek.* URL: http://de.wikibooks.org/wiki/GNU_R (besucht am 23.12.2012).

GOOGLE MAPS (2012): *Google Inc.* URL: <http://maps.google.at/> (besucht am 20.10.2012).

HEUTSCHI, K. und A. ROSENHECK (1997): *Bericht zum F+E-Projekt „Neues EMPA-Modell für Strassenlärm“ - Teil Quellenbeschreibung. Eidgenössische Materialprüfungs- und Forschungsanstalt (EMPA) - Abteilung Akustik und Lärmbekämpfung.* URL: http://www.empa.ch/plugin/template/empa/*/5577/ (besucht am 20.05.2012).

LISA (2012): *Land Information System Austria. Bundesministerium für Verkehr, Innovation und Technologie (bmvit), Austrian Space Applications Programme (ASAP 6 GMES), Österreichische Forschungsförderungsgesellschaft mbH (FFG).* URL: <http://www.landinformationssystem.at/> (besucht am 28.06.2012).

METI, NASA (2012): *ASTER Global Digital Elevation Model (ASTER GDEM). Easy-to-use topographic information of the global terrain. Japan Space Systems.* URL: <http://www.jspacesystems.or.jp/ersdac/GDEM/E/index.html> (besucht am 05.06.2012).

- PYTHON (2011): *Python v2.6.8 documentation. Tutorials, Library Reference, Language Reference, etc..Python Software Foundation*. URL: <http://docs.python.org/2.6/> (besucht am 30.10.2012).
- (2013): *Python Programming Language - Official Website. Python Software Foundation*. URL: <http://www.python.org/> (besucht am 06.02.2013).
- R (2012): *The R Project for Statistical Computing*. URL: <http://www.r-project.org/> (besucht am 10.03.2012).
- REED, SARAH E., JENNIFER L. BOGGS und JACOB P. MANN (2010): *SPreAD-GIS: an ArcGIS toolbox for modeling the propagation of engine noise in a wildland setting (Version 2.0). SPreAD-GIS User's Guide, v. 2.0 2 . The Wilderness Society, San Francisco, CA*. URL: <http://warnercnr.colostate.edu/~sereed/publications/ReedBoggsMann2010.pdf> (besucht am 20.05.2012).
- RIEDL, LEOPOLD (1999): *Leop's MapModels. User Manual Version 1.1 beta. Institut für Stadt- und Regionalforschung (Technische Universität Wien)*. (Besucht am 09.10.2012).
- SENGPIEL, EBERHARD (2012): *Forum für Mikrofonaufnahmetechnik und Tonstudioteknik. Dezibel-Tabelle - SPL - Lautstärkevergleichstabelle*. URL: <http://www.sengpielaudio.com/TabelleDerSchallpegel.htm> (besucht am 15.05.2012).
- SUNEARTHTOOLS (2012): *Sun Position. Tools für die Verbraucher und Designer des Sonnensystems*. URL: http://www.sunearthtools.com/dp/tools/pos_sun.php (besucht am 15.06.2012).
- TACHIKAWA, T. u. a. (2011): *Characteristics of ASTER GDEM Version 2. Earth Remote Sensing Data Analysis Center (ERSDAC), Mitsubishi Material Techno Corp., 3 University of Tokyo*. URL: http://www.jspacesystems.or.jp/ersdac/GDEM/ver2Validation/IGARSS2011_Proceedings_GDEM2.pdf (besucht am 05.06.2012).
- UMWELTBUNDESAMT (2012): *Flächennutzung in Österreich*. URL: <http://www.umweltbundesamt.at/umweltsituation/raumordnung/flchen-entw/> (besucht am 27.07.2012).
- VENABLES, W. N., D. M. SMITH und DAS R DEVELOPMENT CORE TEAM (2012): *An Introduction to R. Notes on R: A Programming Environment for Data Analysis and Graphics*. URL: <http://cran.r-project.org/doc/manuals/R-intro.html> (besucht am 01.11.2012).

WAHLBRINK, STEPHAN und WALWARE-TEAM (2012): *StatET for R*. URL: <http://www.walware.de/goto/statet> (besucht am 05.02.2012).

ZAMG (2011): *Energieoptimierung*. Zentralanstalt für Meteorologie und Geodynamik. URL: <http://www.zamg.ac.at/cms/de/produkte/klima/daten-und-statistiken/energieoptimierung> (besucht am 01.12.2012).

A.4 Rechtsquellen

BUNDES-LÄRMV (2006): *Bundes-Umgebungslärmschutzverordnung. Verordnung des Bundesministers für Land- und Forstwirtschaft, Umwelt und Wasserwirtschaft über die Methoden und technischen Spezifikationen für die Erhebung des Umgebungslärms.* BGBl. II Nr. 144/2006.

LBG (1992): *Liegenschaftsbewertungsgesetz. Bundesgesetz über die gerichtliche Bewertung von Liegenschaften.* BGBl. Nr. 150/1992.

SCHIV (2012): *Schienenverkehrslärm-Immissionsschutzverordnung. Verordnung des Bundesministers für öffentliche Wirtschaft und Verkehr über Lärmschutzmaßnahmen bei Haupt-, Neben- und Straßenbahnen.* BGBl. Nr. 415/1993.

A.5 Abbildungsverzeichnis

3.1	Flächeninformation aus Straßennamen (Beispiel: Wien, 1. Bezirk)	25
3.2	Vergleiche Nutzungsdaten mit Bing und Google (Beispiel: Linz)	28
3.3	Flächenanteile Österreich	30
3.4	Höhenmodell Österreich ¹¹⁵	32
3.5	Nachbarschaftsraster ¹¹⁶	33
3.6	Ausrichtung (Bezirke Innsbruck Stadt und Land)	34
3.7	Neigung [in°] (Bezirke Bludenz und Landeck)	36
3.8	Kreisnachbarschaft (Focal Statistics)	38
3.9	Geländekrümmung (Beispiel: Nördlicher Teil des Bezirks Innsbruck(Stadt))	40
3.10	Azimut α und Elevation(\sphericalangle) der Sonne in $AUT_{\bar{x}}$, am 01.01.2012, 12:00 Uhr	42
3.11	Hypothetische Beleuchtung (Innsbruck Stadt)	45
3.12	Verteilung der Lärmtypen	46
3.13	Dämpfungskonstante DK (Beispiel: Wien und Umgebung)	51
3.14	Pegelveränderung $\Delta P_{RB}[indB]$ (Beispiel: Wien)	53
3.15	Pegelveränderung $\Delta P_{Steigung}$ (Beispiel: Autobahn NK)11402	54
3.16	Flächendämpfung ΔP_{Fl} (Beispiel: Wien)	57
3.17	Simulierter Umgebungslärm (Beispiel Wien und Umgebung)	60
3.18	Reisegeschwindigkeiten (Beispiel: Ybbs an der Donau)	67
3.19	Cost Distance - Knotenreisekosten ¹¹⁷	70
3.20	Vertikaler Faktor ($VF_{Neigung}$)	72
3.21	Reisepfad, Reisedistanz und Reisekosten am Beispiel Linz	74
3.22	Vergleiche GoogleMaps-Routing mit Erreichbarkeitsmodell	75
3.23	Reisekosten zu den Landeshauptstädten (Kl_1)	78
3.24	Allokation der Reisekosten (Beispiel: nordwestliche Steiermark)	80
3.25	Entfernung zu den nächstgelegenen Zentren (Min_1)	81
3.26	Einzugsgebiete der definierten Zentren	82
3.27	Verknüpfung mit NK's (Beispiel: Graz Nord)	87

4.1	Verortung der LIEBE-Daten	92
4.2	Finales Modell EFH	108
4.3	Erreichbarkeitsindex	110
4.4	Nichtlineare Funktionen des finalen Modells (EFH)	112
4.5	Modellqualität des finalen Modells (EFH)	113
4.6	Anzahl der Wohngebäude	115
4.7	Finales Modell ETW	117
4.8	Nichtlineare Funktionen des finalen Modells (ETW)	120
4.9	Modellqualität des finalen Modells (ETW)	121
5.1	Simuliertes EFH	127
5.2	Simuliertes ETW	129
5.3	Vergleiche simuliertes EFH mit ETW	131
B.1	Basismodell EFH nicht hybrid, ohne Raster	160
B.2	Nichtlineare Funktionen des Basismodells EFH nicht hybrid, ohne Raster	161
B.3	Modellqualität des Basismodells EFH nicht hybrid, ohne Raster	161
B.4	Basismodell ETW nicht hybrid, ohne Raster	162
B.5	Nichtlineare Funktionen des Basismodells ETW nicht hybrid, ohne Raster	163
B.6	Modellqualität des Basismodells ETW nicht hybrid, ohne Raster	163
B.7	Basismodell EFH nicht hybrid, mit Raster	164
B.8	Nichtlineare Funktionen des Basismodells EFH nicht hybrid, mit Raster	165
B.9	Modellqualität des Basismodells EFH nicht hybrid, mit Raster	166
B.10	Basismodell ETW nicht hybrid, mit Raster	167
B.11	Nichtlineare Funktionen des Basismodells ETW nicht hybrid, mit Raster	168
B.12	Modellqualität des Basismodells ETW nicht hybrid, mit Raster	169
B.13	Basismodell hybrid, ohne Raster EFH	170
B.14	Nichtlineare Funktionen des Basismodells hybrid, ohne Raster (EFH) .	171
B.15	Modellqualität des Basismodells hybrid, ohne Raster (EFH)	172
B.16	Basismodell hybrid, ohne Raster ETW	173
B.17	Nichtlineare Funktionen des Basismodells hybrid, ohne Raster (ETW) .	174

B.18 Modellqualität des Basismodells hybrid, ohne Raster (ETW) 174

A.6 Tabellenverzeichnis

3.1	Datenquellen	18
3.2	Nutzungsdatenschicht	21
3.3	Sonnenposition am Standort $AUT_{\bar{x}}$ ¹¹⁸	42
3.4	Modellierbare Lärmquellen	47
3.5	Angenommene L_{eq} [in dB] pro Lärmquelle	48
3.6	Angenommenes $\frac{h}{b}$ Verhältnis pro Lärmquelle	52
3.7	Angenommene Dämpfungskonstante f_{DK} pro Lärmquelle	56
3.8	Verwendete Straßentypen	63
3.9	Merkmale bei Einwohnerzahlen im Regionalstatistischen Raster	84
4.1	Informationen zu den Immobilienpreisdaten	91
4.2	Variablen der Basismodelle	95
4.3	Modellvergleich EFH	122
4.4	Modellvergleich ETW	123
5.1	Simualtionsobjekte	126
A.1	Verwendete Software mit Kurzbeschreibung ¹¹⁹	148
B.1	Zielpunkte und Klassenzuordnung	152
B.2	Variablenbeschreibung	158
B.3	Transaktionen pro Bundesland und Objektart (2007-2012)	159
B.4	Transaktionen pro Jahr und Objektart	159

A.7 Verwendete Software

Software name	Beschreibung
Eclipse SDK V3.7.1	Software-Plattform und Integrierte Entwicklungsumgebung
Eclipse Pydev Plugin	Ermöglicht das Ausführen von Python in Eclipse
Eclipse StatET Plugin	Ermöglicht das Ausführen von R in Eclipse
ESRI ArcMAP 9.3 und 10 (inkl. Spatial Analyst)	GIS mit Pythonschnittstelle (ArcPy)
IrfanView 4.30	Bildbearbeitungsprogramm
JabRef 2.7.2	Literaturverwaltung
LateX, MiKTeX 2.9 im TeXnicCenter 1.0	Textverarbeitungsprogramm
Notepad++	Erweiterter Texteditor für viele Programmiersprachen
Microsoft Corporation, Office Excel 2010	Tabellenkalkulationsprogramm
R v2.14.0	Freie Programmiersprache für statistische Berechnungen (R Project for Statistical Computing) mit einer General Public License (GNU)
Stata 10 und 12.1 SE	Statistiksoftware
Python 2.6	Open Source Programmiersprache

Tabelle A.1: Verwendete Software mit Kurzbeschreibung¹²⁰

¹²⁰ Weiterführende Informationen sowie die Bezugsquellen sind für die methodisch relevanten Programme unter den obigen Verzeichnissen auf den Seiten 137 ff und 140 ff zu finden.

Anhang B

Tabellen und Grafiken

B.1 Erreichbarkeit

Zielpunkte

OP	Stadt	Einwohner	Land	Klasse (Kl_n)
1	Eisenstadt	11334	AUT	1
2	Bregenz	26752	AUT	1
3	Graz	226244	AUT	1
4	Innsbruck	113392	AUT	1
5	Klagenfurt	90141	AUT	1
6	Linz	183504	AUT	1
7	Salzburg	142662	AUT	1
8	Sankt Pölten	49121	AUT	1
9	Wien - (Innere Stadt)	1550123	AUT	1
10	Amstetten	22595	AUT	2
11	Baden	24502	AUT	2
12	Dornbirn	42301	AUT	2
13	Feldkirch	28607	AUT	2
14	Kapfenberg	22234	AUT	2
15	Klosterneuburg	24797	AUT	2
16	Krems an der Donau	23713	AUT	2
17	Leoben	25804	AUT	2
18	Leonding	22203	AUT	2
19	Mödling	20405	AUT	2
20	Steyr	39340	AUT	2
21	Traun	23470	AUT	2
22	Villach	57497	AUT	2
23	Wels	56478	AUT	2
24	Wien - Alsergrund		AUT	2
25	Wien - Brigittenau		AUT	2
26	Wien - Döbling		AUT	2
27	Wien - Donaustadt		AUT	2

ANHANG B. TABELLEN UND GRAFIKEN

OP	Stadt	Einwohner	Land	Klasse (Kl_n)
28	Wien - Favoriten		AUT	2
29	Wien - Floridsdorf		AUT	2
30	Wien - Hernals		AUT	2
31	Wien - Hietzing		AUT	2
32	Wien - Josefstadt		AUT	2
33	Wien - Landstraße		AUT	2
34	Wien - Leopoldstadt		AUT	2
35	Wien - Liesing		AUT	2
36	Wien - Margareten		AUT	2
37	Wien - Mariahilf		AUT	2
38	Wien - Meidling		AUT	2
39	Wien - Neubau		AUT	2
40	Wien - Ottakring		AUT	2
41	Wien - Penzing		AUT	2
42	Wien - Rudolfsheim-Fünfhaus		AUT	2
43	Wien - Simmering		AUT	2
44	Wien - Währing		AUT	2
45	Wien - Wieden		AUT	2
46	Wiener Neustadt	37627	AUT	2
47	Wolfsberg	25301	AUT	2
48	Ansfelden	14789	AUT	3
49	Bad Ischl	14081	AUT	3
50	Bad Radkersburg	1599	AUT	3
51	Bad Vöslau	10998	AUT	3
52	Bischofshofen	10087	AUT	3
53	Bludenz	13701	AUT	3
54	Braunau am Inn	16337	AUT	3
55	Bruck an der Leitha	7311	AUT	3
56	Bruck an der Mur	13439	AUT	3
57	Deutschlandsberg	7983	AUT	3
58	Eferding	3393	AUT	3
59	Enns	10611	AUT	3
60	Feldbach	4680	AUT	3
61	Feldkirchen in Kärnten	14030	AUT	3
62	Freistadt	7353	AUT	3
63	Fürstenfeld	5982	AUT	3
64	Gänserndorf	7928	AUT	3
65	Gmünd	5861	AUT	3
66	Gmunden	13184	AUT	3
67	Götzis	10097	AUT	3
68	Grieskirchen	4802	AUT	3
69	Güssing	3902	AUT	3
70	Hall in Tirol	11492	AUT	3
71	Hallein	18399	AUT	3
72	Hard	11471	AUT	3
73	Hartberg	6547	AUT	3
74	Hermagor	7232	AUT	3
75	Hohenems	13891	AUT	3
76	Hollabrunn	10685	AUT	3
77	Horn	6411	AUT	3

OP	Stadt	Einwohner	Land	Klasse (Kl_n)
78	Imst	8689	AUT	3
79	Jennersdorf	4236	AUT	3
80	Judenburg	10130	AUT	3
81	Kirchdorf an der Krems	4099	AUT	3
82	Kitzbühel	8574	AUT	3
83	Knittelfeld	12740	AUT	3
84	Köflach	10671	AUT	3
85	Korneuburg	11032	AUT	3
86	Kufstein	15358	AUT	3
87	Landeck	7336	AUT	3
88	Leibnitz	6892	AUT	3
89	Lienz	12079	AUT	3
90	Liezen	6908	AUT	3
91	Lilienfeld	3021	AUT	3
92	Lustenau	19709	AUT	3
93	Marchtrenk	11274	AUT	3
94	Mattersburg	6256	AUT	3
95	Melk	5222	AUT	3
96	Mistelbach	10644	AUT	3
97	Murau	2331	AUT	3
98	Mürzzuschlag	9569	AUT	3
99	Neunkirchen	11028	AUT	3
100	Neusiedl am See	5584	AUT	3
101	Oberpullendorf	2793	AUT	3
102	Oberwart	6696	AUT	3
103	Perchtoldsdorf	13998	AUT	3
104	Perg	7129	AUT	3
105	Rankweil	11171	AUT	3
106	Reutte	5719	AUT	3
107	Ried im Innkreis	11404	AUT	3
108	Rohrbach in Oberösterreich	2353	AUT	3
109	Rust	1714	AUT	3
110	Saalfelden am Steinernen Meer	15093	AUT	3
111	Sankt Andrä	10719	AUT	3
112	Sankt Johann im Pongau	10260	AUT	3
113	Sankt Veit an der Glan	12839	AUT	3
114	Schärding	5052	AUT	3
115	Scheibbs	4331	AUT	3
116	Schwaz	12212	AUT	3
117	Schwechat	15286	AUT	3
118	Spittal an der Drau	16045	AUT	3
119	Stockerau	14452	AUT	3
120	Tamsweg	5936	AUT	3
121	Telfs	12833	AUT	3
122	Ternitz	15232	AUT	3
123	Traiskirchen	15669	AUT	3
124	Tulln an der Donau	13591	AUT	3
125	Vöcklabruck	11697	AUT	3
126	Voitsberg	10074	AUT	3
127	Völkermarkt	11373	AUT	3

OP	Stadt	Einwohner	Land	Klasse (Kl_n)
128	Waidhofen an der Thaya	5750	AUT	3
129	Waidhofen an der Ybbs	11662	AUT	3
130	Wals-Siezenheim	11024	AUT	3
131	Weiz	8926	AUT	3
132	Wörgl	10885	AUT	3
133	Zell am See	9638	AUT	3
134	Zwettl-Niederösterreich	11630	AUT	3
135	Bolzano	102570	IT	2b
136	Bratislava	426927	SK	2b
137	Brno	368533	CZ	2b
138	Budapest	1702297	HU	2b
139	Ceske Budejovice	95071	CZ	2b
140	Friedrichshafen	58484	DE	2b
141	Győr	127594	HU	2b
142	Konstanz / Kreuzlingen	99763	DE	2b
143	Ljubjana	278638	SLO	2b
144	Maribor	116769	SLO	2b
145	München	1351747	DE	2b
146	Passau	50741	DE	2b
147	Praha	1223368	CZ	2b
148	Rosenheim	60674	DE	2b
149	Sankt Gallen	71126	CH	2b
150	Sopron	56394	HU	2b
151	Szombathely	79300	HU	2b
152	Trento	112793	IT	2b
153	UDINE	96750	IT	2b
154	Winterthur	96462	CH	2b
155	Zürich	358540	CH	2b
156	Vaduz	5109	CH	3b

Tabelle: B.1: Zielpunkte und Klassenzuordnung

B.2 Projektion im GIS

B.2.1 Zielprojektion

Im GIS wurden alle importierten Daten auf die nachstehende Lambert Austria Projektion für Österreich umgerechnet:

Lambert Austria

```
PROJCS["LAM_48_MGI",GEOGCS["GCS_Bessel_1841",
DATUM["D_Bessel_1841",
SPHEROID["Bessel_1841",6377397.155,299.1528128]],
```

```
PRIMEM["Greenwich",0.0],
UNIT["Degree",0.0174532925199433]],
PROJECTION["Lambert_Conformal_Conic"],
PARAMETER["False_Easting",400000.0],
PARAMETER["False_Northing",400000.0],
PARAMETER["Central_Meridian",13.3333333333333],
PARAMETER["Standard_Parallel_1",46.0],
PARAMETER["Standard_Parallel_2",49.0],
PARAMETER["Central_Parallel",48.0],
UNIT["Meter",1.0]]
```

B.2.2 Importprojektionen

Die GIS-Datensätze wurden mit folgenden Projektionen übergeben:

Lambert Austria Neu

```
PROJCS["MGI_Austria_Lambert",
GEOGCS["GCS_MGI",DATUM["D_MGI",
SPHEROID["Bessel_1841",6377397.155,299.1528128]],
PRIMEM["Greenwich",0.0],
UNIT["Degree",0.0174532925199433]],
PROJECTION["Lambert_Conformal_Conic"],
PARAMETER["False_Easting",400000.0],
PARAMETER["False_Northing",400000.0],
PARAMETER["Central_Meridian",13.3333333333333],
PARAMETER["Standard_Parallel_1",46.0],
PARAMETER["Standard_Parallel_2",49.0],
PARAMETER["Latitude_Of_Origin",47.5],
UNIT["Meter",1.0]]
```

ETRS 1989

```
PROJCS['ETRS_1989_LAEA',
GEOGCS['GCS_ETRS_1989',
DATUM['D_ETRS_1989',
SPHEROID['GRS_1980',6378137.0,298.257222101]],
PRIMEM['Greenwich',0.0],
UNIT['Degree',0.0174532925199433]],
PROJECTION['Lambert_Azimuthal_Equal_Area'],
PARAMETER['False_Easting',4321000.0],
PARAMETER['False_Northing',3210000.0],
PARAMETER['Central_Meridian',10.0],
PARAMETER['Latitude_Of_Origin',52.0],
UNIT['Meter',1.0]]"
```

WGS 1984

```
GEOGCS["GCS_WGS_1984",
DATUM["D_WGS_1984",
SPHEROID["WGS_1984",6378137,298.257223563]],
PRIMEM["Greenwich",0],
UNIT["Degree",0.017453292519943295]]
```

B.3 Modellvariablen

Variablenname	Beschreibung
Höhenmodell (Kapitel 3.2)	
ra_ausricht	Ausrichtung
ra_neigung	Hangneigung in Grad
ra_hoehe	Höhenlage in Meter
ra_hilshade	Schummerung für Karte
ra_hs_dez	Sonneneinstrahlung: hypothetische Beleuchtung Dezember
ra_hs_jun	Sonneneinstrahlung: hypothetische Beleuchtung Juni
ra_hs_mar	Sonneneinstrahlung: hypothetische Beleuchtung März
ra_hs_sep	Sonneneinstrahlung: hypothetische Beleuchtung September
ra_hs_mean	Sonneneinstrahlung: hypothetische Beleuchtung Durchschnitt
ra_kk_gis	Geländekrümmung (ArcGIS)
ra_kk_1	Geländekrümmung (Radius: 1 $St_{Typ} = \bar{x}$)
ra_kk_2	Geländekrümmung (Radius: 2 $St_{Typ} = \bar{x}$)
ra_kk_3	Geländekrümmung (Radius: 3 $St_{Typ} = \bar{x}$)
ra_kk_5	Geländekrümmung (Radius: 5 $St_{Typ} = \bar{x}$)
ra_kk_7	Geländekrümmung (Radius: 7 $St_{Typ} = \bar{x}$)
ra_kk_9	Geländekrümmung (Radius: 9 $St_{Typ} = \bar{x}$)
ra_kk_mean	Mittelwert der Geländekrümmung (Radius: Alle $St_{Typ} = \bar{x}$)
ra_kk_med_1	Geländekrümmung (Radius: 1 $St_{Typ} = \tilde{x}$)
ra_kk_med_2	Geländekrümmung (Radius: 2 $St_{Typ} = \tilde{x}$)
ra_kk_med_3	Geländekrümmung (Radius: 3 $St_{Typ} = \tilde{x}$)
ra_kk_med_5	Geländekrümmung (Radius: 5 $St_{Typ} = \tilde{x}$)
ra_kk_med_7	Geländekrümmung (Radius: 7 $St_{Typ} = \tilde{x}$)
ra_kk_med_9	Geländekrümmung (Radius: 9 $St_{Typ} = \tilde{x}$)
ra_kk_med_12	Geländekrümmung (Radius: 12 $St_{Typ} = \tilde{x}$)
ra_kk_med_15	Geländekrümmung (Radius: 15 $St_{Typ} = \tilde{x}$)
ra_kk_med_mean	Mittelwert Geländekrümmung (Radius: Alle $St_{Typ} = \tilde{x}$)
Lärmmodell (Kapitel 3.3)	
ra_laerm_max_10	Maximum über Lärmgruppe A
ra_laerm_max_17	Maximum über Lärmgruppe B
ra_laerm_max_pt"	Maximum über Lärmgruppe C+D
ra_laerm_all_max_sum	Maximum über alle Lärmgruppen
ra_laerm_log_sum_pt	Summe über Lärmgruppe C+D
ra_laerm_all_log_sum	Summe über alle Lärmgruppen
Erreichbarkeitsmodell (Kapitel 3.4)	
ra_ocl_1_min_1	Entfernung zum nächsten Ort der Klasse 1
ra_ocl_1_min_2	Entfernung zum zweitnächsten Ort der Klasse 1

Variablenname	Beschreibung
ra_ocl_1_min_3	Entfernung zum drittnächsten Ort der Klasse 1
ra_ocl_2_nAT_min_1	Entfernung zum nächsten Ort der Klasse 2
ra_ocl_2_nAT_min_2	Entfernung zum zweitnächsten Ort der Klasse 2
ra_ocl_2_nAT_min_3	Entfernung zum drittnächsten Ort der Klasse 2
ra_ocl_2_min_1	Entfernung zum nächsten Ort der Klasse 2b
ra_ocl_2_min_2	Entfernung zum zweitnächsten Ort der Klasse 2b
ra_ocl_2_min_3	Entfernung zum drittnächsten Ort der Klasse 2b
ra_ocl_3_nAT_min_1	Entfernung zum nächsten Ort der Klasse 3
ra_ocl_3_nAT_min_2	Entfernung zum zweitnächsten Ort der Klasse 3
ra_ocl_3_nAT_min_3	Entfernung zum drittnächsten Ort der Klasse 3
ra_ocl_3_min_1	Entfernung zum nächsten Ort der Klasse 3b
ra_ocl_3_min_2	Entfernung zum zweitnächsten Ort der Klasse 3b
ra_ocl_3_min_3	Entfernung zum drittnächsten Ort der Klasse 3b
ra_ocl_1_mean	Mittelwert Klasse 1
ra_ocl_2_nAT_mean	Mittelwert Klasse 2
ra_ocl_2_mean	Mittelwert Klasse 2b
ra_ocl_3_nAT_mean	Mittelwert Klasse 3
ra_ocl_3_mean	Mittelwert Klasse 3b
ra_ocl_min_1_mean	Mittelwert zum nächsten Ort der Klassen 1, 2, 3
ra_ocl_min_2_nAT_mean	Mittelwert zum zweitnächsten Ort der Klassen 1, 2, 3
ra_ocl_min_2_mean	Mittelwert zum zweitnächsten Ort der Klassen 1, 2b, 3b
ra_ocl_min_3_nAT_mean	Mittelwert zum drittnächsten Ort der Klassen 1, 2, 3
ra_ocl_min_3_mean	Mittelwert zum drittnächsten Ort der Klassen 1, 2b, 3b
ra_ocl_nAT_mean_all	Mittelwert über Klasse 1, 2, 3
ra_ocl_mean_all	Mittelwert über Klasse 1, 2b, 3b
Regionalstatistischer Raster (Kapitel 3.5.1)	
br_pop_01	Anzahl der Einwohner mit Hauptwohnsitz (2001)
br_pop_06	Anzahl der Einwohner mit Hauptwohnsitz (2006)
br_pop_11	Anzahl der Einwohner mit Hauptwohnsitz (2011)
br_pop_12	Anzahl der Einwohner mit Hauptwohnsitz (2012)
br_pop_11_12_ant	Veränderung der Hauptwohnsitze zwischen 2011 und 2012 in Prozent
br_pop_01_12_ant	Veränderung der Hauptwohnsitze zwischen 2001 und 2012 in Prozent
br_nws_01	Anzahl der Einwohner mit Nebenwohnsitz (2001)
br_nws_06	Anzahl der Einwohner mit Nebenwohnsitz (2006)
br_nws_01_ant	Anteil der Nebenwohnsitze (2001)
br_nws_06_ant	Anteil der Nebenwohnsitze (2006)
br_hh_01	Anzahl der Haushalte (2001)
br_all_01_12_ant	Anteil der 0 bis 2 jährigen Einwohner (2012)
br_all_02_12_ant	Anteil der 3 bis 5 jährigen Einwohner (2012)
br_all_03_12_ant	Anteil der 6 bis 9 jährigen Einwohner (2012)
br_all_04_12_ant	Anteil der 10 bis 14 jährigen Einwohner (2012)
br_all_05_12_ant	Anteil der 15 bis 19 jährigen Einwohner (2012)
br_all_06_12_ant	Anteil der 20 bis 24 jährigen Einwohner (2012)
br_all_07_12_ant	Anteil der 25 bis 34 jährigen Einwohner (2012)
br_all_08_12_ant	Anteil der 35 bis 44 jährigen Einwohner (2012)
br_all_09_12_ant	Anteil der 45 bis 64 jährigen Einwohner (2012)
br_all_10_12_ant	Anteil der 65 bis 84 jährigen Einwohner (2012)
br_all_11_12_ant	Anteil der über 84 jährigen Einwohner (2012)
br_her_aut_12_ant	Anteil der Einwohner mit Herkunft aus Österreich (2012)
br_her_eu_12_ant	Anteil der Einwohner mit Herkunft aus der EU (2012)

ANHANG B. TABELLEN UND GRAFIKEN

Variablenname	Beschreibung
br_her_afn_12_ant	Anteil der Einwohner mit Herkunft aus Nordafrika (2012)
br_her_afs_12_ant	Anteil der Einwohner mit Herkunft aus Südafrika (2012)
br_her_ams_12_ant	Anteil der Einwohner mit Herkunft aus Mittel- und Südamerika (2012)
br_her_asw_12_ant	Anteil der Einwohner mit Herkunft aus Westasien (2012)
br_her_asz_12_ant	Anteil der Einwohner mit Herkunft aus Zentral- Südasiem (2012)
br_her_XXX_12_ant	Anteil der Einwohner mit unbekannter/ungeklärter Herkunft (2012)
br_her_amn_12_ant	Anteil der Einwohner mit Herkunft aus Nordamerika (2012)
br_her_aso_12_ant	Anteil der Einwohner mit Herkunft aus Ostasien (2012)
br_her_aus_12_ant	Anteil der Einwohner mit Herkunft aus Australien (2012)
br_her_deu_12_ant	Anteil der Einwohner mit Herkunft aus Deutschland (2012)
br_her_eun_12_ant	Anteil der Einwohner mit Herkunft aus Nordeuropa (2012)
br_her_euw_12_ant	Anteil der Einwohner mit Herkunft aus Westeuropa (2012)
br_her_euo_12_ant	Anteil der Einwohner mit Herkunft aus Osteuropa (2012)
br_her_eus_12_ant	Anteil der Einwohner mit Herkunft aus Südeuropa (2012)
br_her_exj_12_ant	Anteil der Einwohner mit Herkunft aus Ex-Jugoslawien (2012)
br_her_tur_12_ant	Anteil der Einwohner mit Herkunft aus der Türkei (2012)
br_arbst_01	Anzahl der Arbeitsstätten (2001)
br_besch_01	Anzahl der Beschäftigten (2001)
br_gcb_01	Anzahl der Gebäude (2001)
br_gcb_06	Anzahl der Gebäude (2006)
br_whg_01	Anzahl der Wohnungen (2001)
br_whg_06	Anzahl der Wohnungen (2006)
br_wohngeb_01	Anzahl der Wohngebäude (2001)
br_wohngeb_06	Anzahl der Wohngebäude (2006)
br_wohngeb_01_ant	Anteile der Wohngebäude (2001)
br_wohngeb_06_ant	Anteile der Wohngebäude (2006)
br_arbst_01_pop_01	Arbeitsstätten pro Hauptwohnsitz (2001)
br_besch_01_pop_01	Beschäftigte pro Hauptwohnsitz (2001)
br_whg_01_pop_01	Wohnungen pro Hauptwohnsitz (2001)
br_whg_06_pop_06	Wohnungen pro Hauptwohnsitz (2006)
br_all_bis24_12_ant	Anteil der 0 bis 24 jährigen Einwohner (2012)
br_all_25bis64_12_ant	Anteil der 25 bis 64 jährigen Einwohner (2012)
br_all_65bis100_12_ant	Anteil der 65 bis über 84 jährigen Einwohner (2012)
br_her_nw_ant	Anteil der Einwohner mit Herkunft Nord- und Westeuropa, Nordamerika und Australien, Ostasien
br_her_so_ant	Anteil der Einwohner mit Herkunft Süd- und Osteuropa (inkl. Ex-Jugoslawien und Türkei)
br_her_aa_wz_xx_ant	Anteil der Einwohner mit Herkunft Afrika, Mittel- und Südamerika, West- und Zentral-Südostasien und ungeklärter Herkunft
br_her_aa_wz_ant	Anteil der Einwohner mit Herkunft Afrika, Mittel- und Südamerika, West- und Zentral-Südostasien
br_her_aut_ant	Anteil der Einwohner mit Herkunft Österreich
br_her_so_aa_wz_ant	Anteil der Einwohner mit Herkunft Süd- und Osteuropa, Afrika, Mittel- und Südamerika, West- und Zentral-Südostasien und ungeklärter Herkunft
br_her_so_xx_ant	Anteil der Einwohner mit Herkunft Süd- und Osteuropa und ungeklärter Herkunft
Raster Lebensphasen (Kapitel 3.5.2)	
lp_personen_im_hh_1_ant	Anteil der Personen in einem 1 Personen Haushalt lebend (2012)
lp_personen_im_hh_2_ant	Anteil der Personen in einem 2 Personen Haushalt lebend (2012)
lp_personen_im_hh_3_ant	Anteil der Personen in einem 3 Personen Haushalt lebend (2012)

Variablenname	Beschreibung
lp_personen_im_hh_4_ant	Anteil der Personen in einem 4 Personen Haushalt lebend (2012)
lp_personen_im_hh_5_ant	Anteil der Personen in einem 5 Personen Haushalt lebend (2012)
lp_personen_im_hh_6_ant	Anteil der Personen in einem 6 Personen Haushalt lebend (2012)
lp_personen_im_hh_7_ant	Anteil der Personen in einem 7 Personen Haushalt lebend (2012)
lp_lebensphase_a_ant	Anteil der Personen in einer Wohngemeinschaft lebend (2012)
lp_lebensphase_b_ant	Anteil der Personen in einer ersten Single Wohnung lebend (2012)
lp_lebensphase_c_ant	Anteil der Personen in einem 2 Personen Haushalt ohne Kind lebend und unter 40 Jahre alt (2012)
lp_lebensphase_d_ant	Anteil der Personen in einem mehrpersonen Haushalt mit 2 Erwachsenen und mit einem oder mehr Kindern mit höchstens 2 Jahren (2012)
lp_lebensphase_e_ant	Anteil der Personen in einem mehrpersonen Haushalt mit 2 Erwachsenen und mit einem oder mehr Kindern mit höchstens 4 Jahren (2012)
lp_lebensphase_f_ant	Anteil der Personen in einem mehrpersonen Haushalt mit 2 Erwachsenen und mit einem oder mehr Kindern mit höchstens 6 Jahren (2012)
lp_lebensphase_g_ant	Anteil der Personen in einem mehrpersonen Haushalt mit 2 Erwachsenen und mit einem oder mehr Kindern mit höchstens 13 Jahren (2012)
lp_lebensphase_h_ant	Anteil der Personen in einem mehrpersonen Haushalt mit 2 Erwachsenen und mit einem oder mehr Kindern mit höchstens 20 Jahren (2012)
lp_lebensphase_i_ant	Anteil der Personen in einem 2 Personen Haushalt ohne Kind lebend und unter 40 Jahre alt (2012)
lp_lebensphase_j_ant	Anteil der Personen in einem Singlehaushalt lebend (2012)
lp_akademiker_ant	Anteil der Akademiker (2012)
lp_entscheider_ant	Anteil der Entscheider. Als Entscheider werden Führungskräfte aus der Wirtschaft bezeichnet. Dies beinhaltet: Entscheidungsträger in Großunternehmen, Beteiligte an Unternehmen, Gewerbetreibende bzw. Selbständige, Vorstände und Vorstandsmitglieder und sonstige Führungskräfte der Wirtschaft (2012)
Raster Gebäudedaten (Kapitel 3.5.3)	
gr_l000250v3	ID der Rasterzelle aus dem 250m Raster der Statistik Austria (2012)
gr_gesamtanzahl_geb_12	Gesamtanzahl der Gebäude (Summe aus GEB_01 bis GEB_09) (2012)
gr_01_1etw_12	Anzahl der Gebäude mit einer Wohnung (2012)
gr_02_2_plus_etw_12	Anzahl der Gebäude mit zwei oder mehr Wohnungen (2012)
gr_03_wohngeb_gem_12	Anzahl der Wohngebäude für Gemeinschaften (2012)
gr_04_hotels_12	Anzahl der Hotels und ähnlicher Gebäude (2012)
gr_05_buero_12	Anzahl der Bürogebäude (2012)
gr_06_handel_12	Anzahl der Groß- und Einzelhandelsgebäude (2012)
gr_07_verkehr_12	Anzahl der Gebäude des Verkehrs- und Nachrichtenwesens (2012)
gr_08_industrie_lager_12	Anzahl der Industrie- und Lagergebäude (2012)
gr_09_kultur_freizeit_12	Anzahl der Gebäude für Kultur- und Freizeitzwecke (2012)
gr_gesamtanzahl_geb_12_ant	Anteil der Gebäude mit einer Wohnung (2012)
gr_01_1etw_12_ant	Anteil der Gebäude mit zwei oder mehr Wohnungen (2012)
gr_02_2_plus_etw_12_ant	Anteil der Wohngebäude für Gemeinschaften (2012)
gr_03_wohngeb_gem_12_ant	Anteil der Hotels und ähnlicher Gebäude (2012)
gr_04_hotels_12_ant	Anteil der Bürogebäude (2012)
gr_05_buero_12_ant	Anteil der Groß- und Einzelhandelsgebäude (2012)
gr_06_handel_12_ant	Anteil der Gebäude des Verkehrs- und Nachrichtenwesens (2012)

Variablenname	Beschreibung
gr_07_verkehr_12_ant	Anteil der Industrie- und Lagergebäude (2012)
gr_08_industrie_lager_12_ant	Anteil der Gebäude für Kultur- und Freizeitzwecke (2012)
Raster Klimadaten (Kapitel 3.5.4)	
ra_klima_hgt	Mittlere Jahressumme der Heizgradtage: „Heizgradtage errechnen sich aus der Summe der täglichen Differenzen zwischen der Raumtemperatur und der mittleren Außentemperatur während der gesamten Heizperiode. Heizgradtage sind meist bezogen auf eine Raumtemperatur von 20° C und eine Heizgrenze von 12° C.“ ¹²¹
ra_klima_neuschnee_sum	Schnee: Mittlere jährliche Summe der Neuschneehöhen.
ra_klima_rrsum_sommer	Niederschlag: Sommersumme (April- September)
ra_klima_rrsum_winter	Niederschlag: Wintersumme (Oktober- März)
ra_klima_sommert	Lufttemperatur: Sommertage > 25°
ra_klima_sonne_abs	Absolute Sonnenscheindauer in Stunden: Die mittlere Summe der absoluten Sonnenscheindauer des Monats Jänner.
ra_klima_sonne_sum”	Relative Sonnenscheindauer in %: Der mittlere monatliche Anteil der tatsächlichen an der maximal möglichen Sonnenscheindauer im Jänner.

Tabelle: B.2: Variablenbeschreibung

B.4 Anzahl im Grundbuch erfasster Transaktionen

Die nachstehend angeführten Tabellen B.3 und B.4 wurden auf Basis der IU-Kaufpreisdaten erstellt. Diese Datensammlung enthält ab Jänner 2008 alle im Grundbuch eingetragenen Transaktionen der Landeshauptstädte und seit Anfang 2009 eine österreichweite Abdeckung.

Da der Verkaufszeitpunkt einer Immobilie nicht mit dem Datum der Grundbucheintragung ident ist, sind in diesem Datensatz vereinzelt sehr alte Transaktionen (bis 1948) enthalten. Aus diesem Grund wurden für die Auswertungen alle Transaktionen, welche vor dem Jahr 2007 getätigt wurden, auf das Jahr 2007 gesetzt.

¹²¹ ZAMG 2011

B.4.1 Transaktion pro Bundesland

Bundesland	EFH	ETW	Summe
Burgenland	2.748	1.208	3.956
Kärnten	3.810	7.684	11.494
Niederösterreich	12.727	17.616	30.343
Oberösterreich	8.013	15.409	23.422
Salzburg	2.954	13.086	16.040
Steiermark	8.787	19.827	28.614
Tirol	2.673	16.630	19.303
Vorarlberg	2.121	9.913	12.034
Wien	3.670	50.488	54.158
Summe	47.503	151.861	199.364

Tabelle B.3: Transaktionen pro Bundesland und Objektart (2007-2012)

B.4.2 Transaktion pro Jahr

Transaktionsjahr	EFH	ETW	Summe
2007	1.374	9.507	10.881
2008	4.866	21.350	26.216
2009	10.366	31.816	42.182
2010	11.028	33.987	45.015
2011	10.974	32.982	43.956
2012	8.895	22.219	31.114
Summe	47.503	151.861	199.364

Tabelle B.4: Transaktionen pro Jahr und Objektart

B.5 Vergleichsmodelle

B.5.1 Das Basismodell nicht hybrid, ohne Raster

B.5.1.1 EFH

Family: gaussian
Link function: identity

Formula:

```
lnp ~ dum_angebot + dum_gemeinnutz + dum_bewerter + s(area_weight) +
s(area_gst, by = bundes_AUT) + s(area_gst, by = bundes_WIEN) +
s(kp_year) + s(baujahr_berech) + rh + antwortid_heiz_efh +
antwortid_bad_qual + bad_size_kat + zusatzbad + terr_wohn_vorh +
antwortid_zustand_haus + bundes
```

Parametric coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	7.29641	0.05878	124.128	< 2e-16 ***
dum_angebot	0.12471	0.01775	7.025	2.53e-12 ***
dum_gemeinnutz	-0.02941	0.05558	-0.529	0.59680
dum_bewerter	0.07179	0.01156	6.211	5.84e-10 ***
rh	0.15348	0.01942	7.902	3.57e-15 ***
antwortid_heiz_efh2	-0.13012	0.01958	-6.644	3.49e-11 ***
antwortid_bad_qual1	0.19090	0.02051	9.309	< 2e-16 ***
antwortid_bad_qual3	-0.10952	0.01314	-8.336	< 2e-16 ***
bad_size_kat1	-0.04167	0.01730	-2.409	0.01606 *
zusatzbad	0.05484	0.01150	4.769	1.92e-06 ***
terr_wohn_vorh	0.05947	0.01357	4.383	1.20e-05 ***
antwortid_zustand_haus1	0.26818	0.02420	11.080	< 2e-16 ***
antwortid_zustand_haus2	0.15522	0.02146	7.233	5.70e-13 ***
bundesBGL	-0.10616	0.02458	-4.318	1.61e-05 ***
bundesKTN	0.01024	0.02452	0.418	0.67624
bundesNOE	0.12033	0.01726	6.972	3.66e-12 ***
bundesOOE	0.05977	0.02262	2.642	0.00827 **
bundesSBG	0.37629	0.04128	9.115	< 2e-16 ***
bundesT	0.35337	0.03264	10.828	< 2e-16 ***
bundesVBG	0.35110	0.03974	8.835	< 2e-16 ***
bundesW	2.68297	0.09469	28.333	< 2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Approximate significance of smooth terms:

	edf	Ref.df	F	p-value
s(area_weight)	6.853	7.978	137.160	< 2e-16 ***
s(area_gst):bundes_AUT	4.385	5.324	1289.853	< 2e-16 ***
s(area_gst):bundes_WIEN	2.503	3.059	147.673	< 2e-16 ***
s(kp_year)	1.000	1.000	9.219	0.00241 **
s(baujahr_berech)	2.901	3.645	63.098	< 2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) = 0.683 Deviance explained = 68.6%
GCV score = 0.090168 Scale est. = 0.089283 n = 3833
AIC = 1656.71 BIC = 1898.28

Abbildung B.1: Basismodell EFH nicht hybrid, ohne Raster

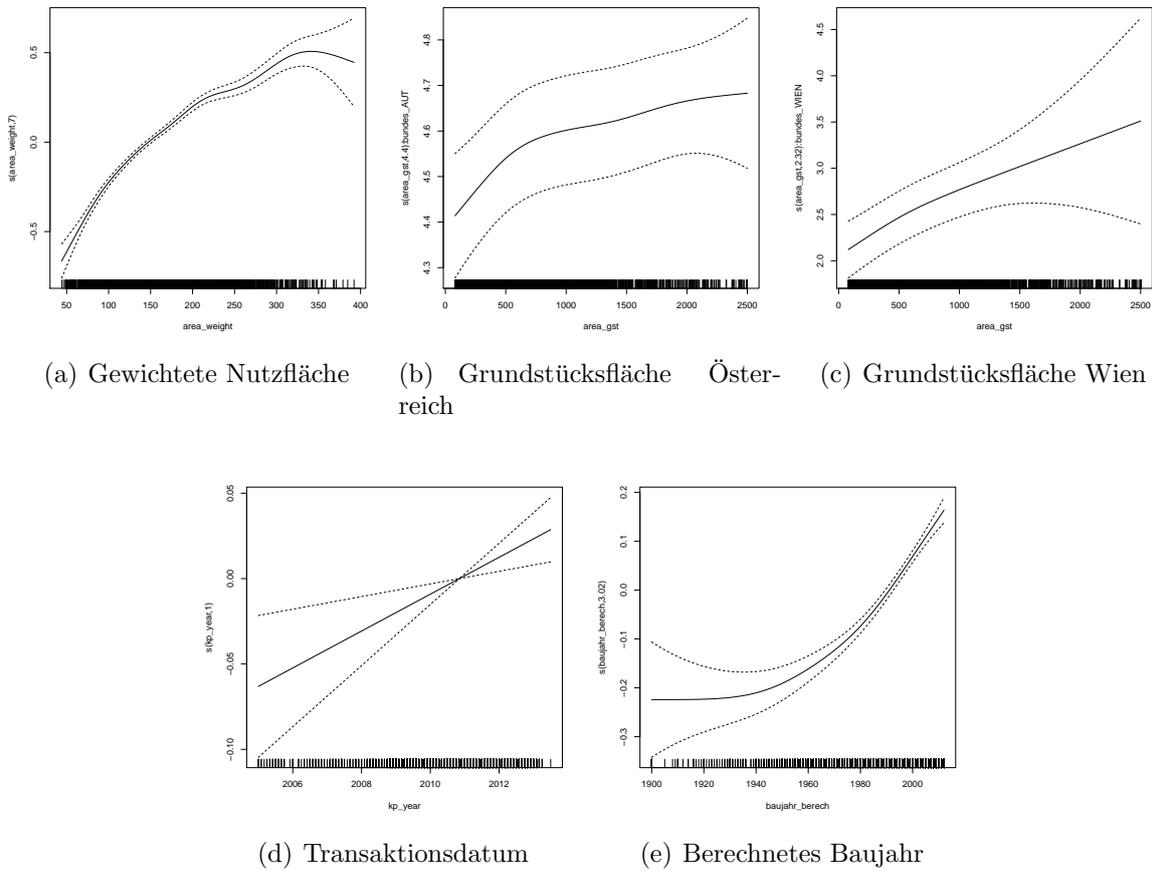


Abbildung B.2: Nichtlineare Funktionen des Basismodells EFH nicht hybrid, ohne Raster

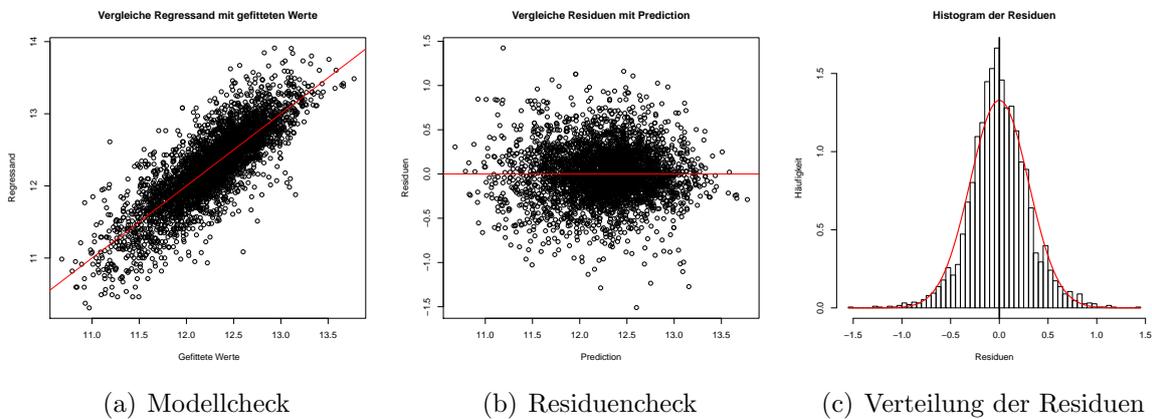


Abbildung B.3: Modellqualität des Basismodells EFH nicht hybrid, ohne Raster

B.5.1.2 ETW

Family: gaussian

Link function: identity

Formula:

```
lnp ~ dum_angebot + dum_gemeinnutz + s(kp_year) + s(baujahr_bereich) +
  antwortid_bad_qual + bad_size_kat + zusatzbad + antwortid_zustand_haus +
  s(area_weight, by = bundes_AUT) + s(area_weight, by = bundes_WIEN) +
  antwortid_zustand_wohnung + antwortid_heiz_etw + balc_vorh +
  lift_up + antwortid_garage_3 + bundes
```

Parametric coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	8.20981	0.19958	41.135	< 2e-16 ***
dum_angebot	0.03594	0.01510	2.380	0.017357 *
dum_gemeinnutz	-0.07575	0.02477	-3.058	0.002249 **
antwortid_bad_qual1	0.12623	0.02310	5.465	4.97e-08 ***
antwortid_bad_qual3	-0.09938	0.01242	-8.005	1.65e-15 ***
bad_size_kat1	-0.01858	0.01257	-1.478	0.139514
zusatzbad	0.07093	0.02045	3.468	0.000531 ***
antwortid_zustand_haus1	0.25868	0.05489	4.712	2.55e-06 ***
antwortid_zustand_haus2	0.16733	0.05443	3.074	0.002127 **
antwortid_zustand_wohnung1	0.10386	0.01517	6.846	9.04e-12 ***
antwortid_zustand_wohnung3	-0.05084	0.01604	-3.169	0.001545 **
antwortid_zustand_wohnung4	-0.12784	0.03800	-3.364	0.000776 ***
antwortid_heiz_etw2	-0.09512	0.01881	-5.056	4.53e-07 ***
balc_vorh	0.01895	0.01038	1.827	0.067821 .
lift_upTRUE	0.06235	0.01088	5.733	1.08e-08 ***
antwortid_garage_3	0.02931	0.01158	2.530	0.011466 *
bundesBGL	0.04978	0.05807	0.857	0.391364
bundesKTN	0.32888	0.03867	8.505	< 2e-16 ***
bundesNOE	0.39581	0.03671	10.782	< 2e-16 ***
bundesOOE	0.37597	0.03945	9.531	< 2e-16 ***
bundesSBG	0.64915	0.04102	15.827	< 2e-16 ***
bundesSTMK	0.32497	0.03693	8.799	< 2e-16 ***
bundesT	0.65748	0.03721	17.667	< 2e-16 ***
bundesVBG	0.52500	0.04102	12.800	< 2e-16 ***

 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

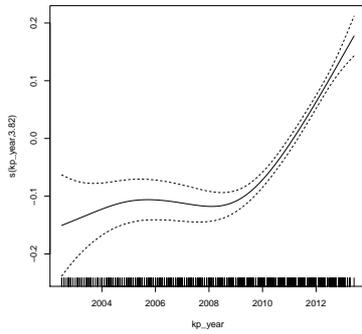
Approximate significance of smooth terms:

	edf	Ref.df	F	p-value
s(kp_year)	3.797	4.693	52.66	<2e-16 ***
s(baujahr_bereich)	6.160	7.219	54.98	<2e-16 ***
s(area_weight):bundes_AUT	7.706	8.778	279.85	<2e-16 ***
s(area_weight):bundes_WIEN	8.947	9.521	377.61	<2e-16 ***

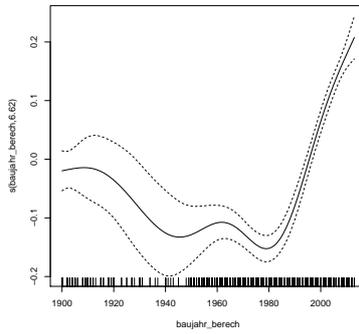
 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) = 0.792 Deviance explained = 79.5%
 GCV score = 0.067203 Scale est. = 0.066194 n = 3285
 AIC = 454.07 BIC = 761

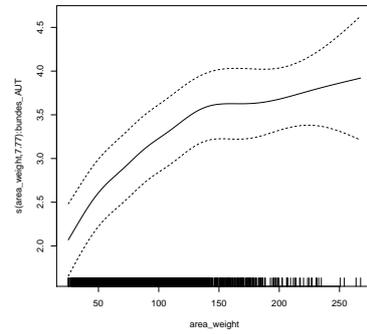
Abbildung B.4: Basismodell ETW nicht hybrid, ohne Raster



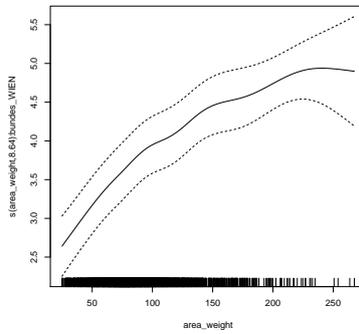
(a) Transaktionsdatum



(b) Berechnetes Baujahr

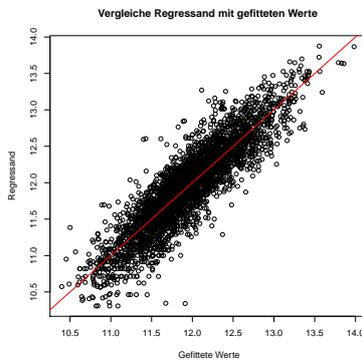


(c) Gewichtete Nutzfläche Österreich

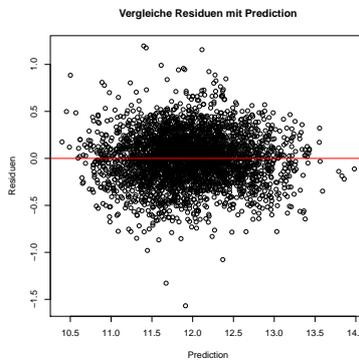


(d) Gewichtete Nutzfläche Wien

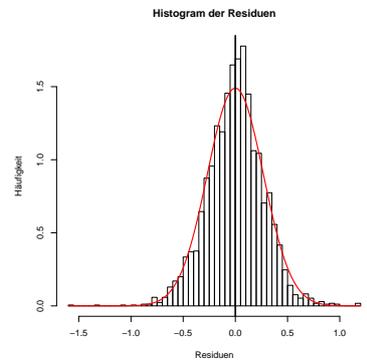
Abbildung B.5: Nichtlineare Funktionen des Basismodells ETW nicht hybrid, ohne Raster



(a) Modellcheck



(b) Residuencheck



(c) Verteilung der Residuen

Abbildung B.6: Modellqualität des Basismodells ETW nicht hybrid, ohne Raster

B.5.2 Das Basismodell nicht hybrid, mit Raster

B.5.2.1 EFH

Family: gaussian

Link function: identity

Formula:

```
lnp ~ dum_angebot + dum_gemeinnutz + dum_bewerter + s(area_weight) +
s(area_gst, by = bundes_AUT) + s(area_gst, by = bundes_WIEN) +
s(kp_year) + s(baujahr_berech) + rh + antwortid_heiz_efh +
antwortid_bad_qual + bad_size_kat + zusatzbad + terr_wohn_vorh +
antwortid_zustand_haus + bundes + br_her_so_ant + ra_kk_med_15 +
lp_ln_akademiker_ant + s(ra_ocl_nat_mean_all) + s(ra_laerm_all_log_sum_q)
```

Parametric coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	7.2427972	0.0501893	144.310	< 2e-16 ***
dum_angebot	0.0865251	0.0148105	5.842	5.59e-09 ***
dum_gemeinnutz	-0.0970749	0.0454977	-2.134	0.032938 *
dum_bewerter	0.0557201	0.0097854	5.694	1.33e-08 ***
rh	0.0812591	0.0160161	5.074	4.09e-07 ***
antwortid_heiz_efh2	-0.0953425	0.0160440	-5.943	3.06e-09 ***
antwortid_bad_qual1	0.1364261	0.0168561	8.094	7.74e-16 ***
antwortid_bad_qual3	-0.0699877	0.0108897	-6.427	1.46e-10 ***
bad_size_kat1	-0.0403489	0.0141780	-2.846	0.004453 **
zusatzbad	0.0391417	0.0094186	4.156	3.31e-05 ***
terr_wohn_vorh	0.0424916	0.0111036	3.827	0.000132 ***
antwortid_zustand_haus1	0.3060197	0.0198959	15.381	< 2e-16 ***
antwortid_zustand_haus2	0.1832561	0.0176665	10.373	< 2e-16 ***
bundesBGL	-0.2395045	0.0211784	-11.309	< 2e-16 ***
bundesKTN	0.0209922	0.0200812	1.045	0.295921
bundesNOE	-0.1057275	0.0161457	-6.548	6.60e-11 ***
bundesOOE	-0.0149424	0.0192840	-0.775	0.438472
bundesSBG	0.4393834	0.0341817	12.854	< 2e-16 ***
bundesT	0.5696954	0.0295018	19.311	< 2e-16 ***
bundesVBG	0.3695925	0.0328357	11.256	< 2e-16 ***
bundesW	2.4892981	0.0815794	30.514	< 2e-16 ***
br_her_so_ant	-0.1161896	0.0509820	-2.279	0.022721 *
ra_kk_med_15	0.0006431	0.0005795	1.110	0.267109
lp_ln_akademiker_ant	0.0586467	0.0039969	14.673	< 2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

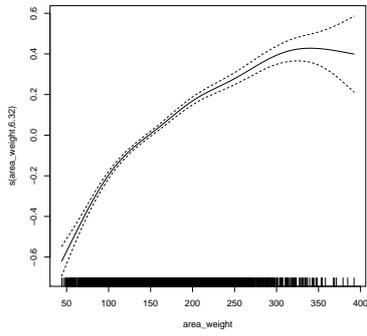
Approximate significance of smooth terms:

	edf	Ref.df	F	p-value
s(area_weight)	6.595	7.758	168.901	< 2e-16 ***
s(area_gst):bundes_AUT	4.433	5.382	1890.080	< 2e-16 ***
s(area_gst):bundes_WIEN	2.564	3.133	177.129	< 2e-16 ***
s(kp_year)	5.531	6.585	3.749	0.000645 ***
s(baujahr_berech)	5.024	6.055	62.772	< 2e-16 ***
s(ra_ocl_nat_mean_all)	5.676	6.870	178.264	< 2e-16 ***
s(ra_laerm_all_log_sum_q)	1.135	1.259	75.328	< 2e-16 ***

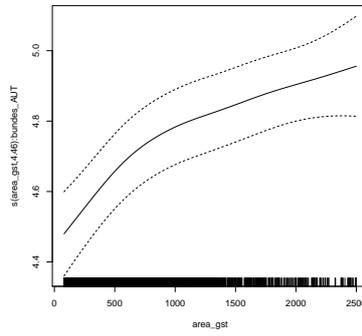
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) = 0.789 Deviance explained = 79.2%
GCV score = 0.060292 Scale est. = 0.059443 n = 3833
AIC = 113.63 BIC = 457.2

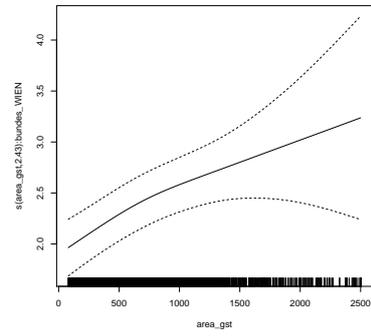
Abbildung B.7: Basismodell EFH nicht hybrid, mit Raster



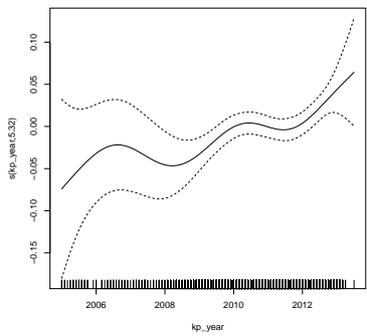
(a) Gewichtete Nutzfläche



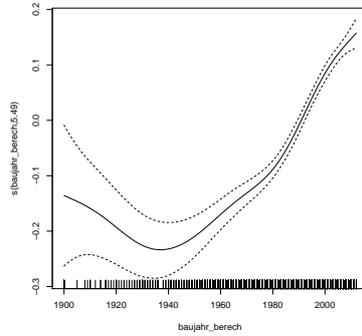
(b) Grundstücksfläche Österreich



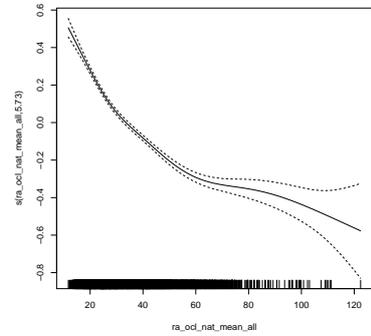
(c) Grundstücksfläche Wien



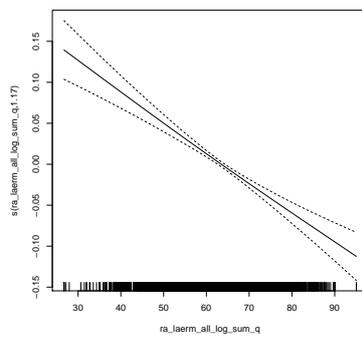
(d) Transaktionsdatum



(e) Berechnetes Baujahr



(f) Erreichbarkeit



(g) Lärm

Abbildung B.8: Nichtlineare Funktionen des Basismodells EFH nicht hybrid, mit Raster

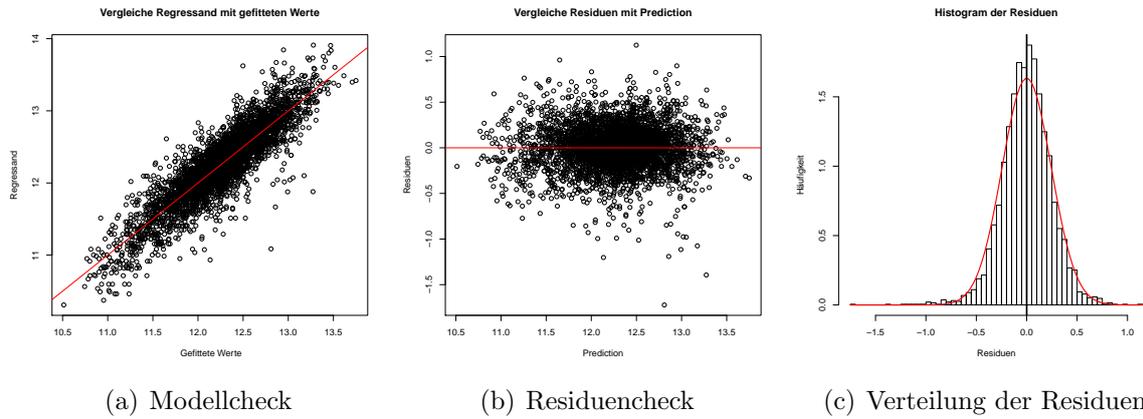


Abbildung B.9: Modellqualität des Basismodells EFH nicht hybrid, mit Raster

B.5.2.2 ETW

Family: gaussian

Link function: identity

Formula:

```
lnp ~ dum_angebot + dum_gemeinnutz + s(kp_year) + s(baujahr_berech) +
  antwortid_bad_qual + bad_size_kat + zusatzbad + antwortid_zustand_haus +
  s(area_weight, by = bundes_AUT) + s(area_weight, by = bundes_WIEN) +
  antwortid_zustand_wohnung + antwortid_heiz_etw + balc_vorh +
  lift_up + antwortid_garage_3 + bundes + br_her_so_ant + s(lp_ln_akademiker_ant) +
  br_arbst_01_pop_01 + ln_wohngeb + speckguertel + s(ra_laerm_all_log_sum) +
  s(ln_ra_ocl_1_min_1_q)
```

Parametric coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	8.114220	0.182477	44.467	< 2e-16 ***
dum_angebot	0.048717	0.013408	3.633	0.000284 ***
dum_gemeinnutz	-0.065909	0.022017	-2.993	0.002779 **
antwortid_bad_qual1	0.100765	0.020557	4.902	9.97e-07 ***
antwortid_bad_qual3	-0.073115	0.011047	-6.618	4.24e-11 ***
bad_size_kat1	-0.014787	0.011164	-1.325	0.185419
zusatzbad	0.060982	0.018204	3.350	0.000817 ***
antwortid_zustand_haus1	0.238672	0.048815	4.889	1.06e-06 ***
antwortid_zustand_haus2	0.164761	0.048371	3.406	0.000667 ***
antwortid_zustand_wohnung1	0.100484	0.013163	7.634	2.98e-14 ***
antwortid_zustand_wohnung3	-0.058225	0.014269	-4.081	4.60e-05 ***
antwortid_zustand_wohnung4	-0.142712	0.033741	-4.230	2.41e-05 ***
antwortid_heiz_etw2	-0.067837	0.016774	-4.044	5.37e-05 ***
balc_vorh	0.027230	0.009263	2.940	0.003309 **
lift_upTRUE	0.063953	0.010046	6.366	2.21e-10 ***
antwortid_garage_3	0.022197	0.010400	2.134	0.032887 *
bundesBGL	0.054555	0.054073	1.009	0.313090
bundesKTN	0.296316	0.038151	7.767	1.07e-14 ***
bundesNOE	0.399592	0.037060	10.782	< 2e-16 ***
bundesOOE	0.427838	0.038670	11.064	< 2e-16 ***
bundesSBG	0.628062	0.039845	15.763	< 2e-16 ***
bundesSTMK	0.310653	0.036558	8.498	< 2e-16 ***
bundesT	0.676690	0.036630	18.473	< 2e-16 ***
bundesVBG	0.539719	0.039861	13.540	< 2e-16 ***
br_her_so_ant	-0.329142	0.047374	-6.948	4.47e-12 ***
br_arbst_01_pop_01	0.075217	0.019904	3.779	0.000160 ***
ln_wohngeb	0.031404	0.005506	5.703	1.28e-08 ***
speckguertel	-0.001873	0.019212	-0.097	0.922353

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

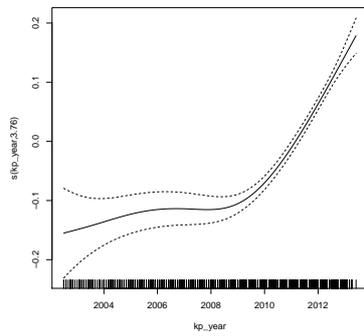
Approximate significance of smooth terms:

	edf	Ref.df	F	p-value
s(kp_year)	3.534	4.383	77.314	< 2e-16 ***
s(baujahr_berech)	5.077	6.119	79.301	< 2e-16 ***
s(area_weight):bundes_AUT	8.807	9.440	349.269	< 2e-16 ***
s(area_weight):bundes_WIEN	8.411	9.272	443.045	< 2e-16 ***
s(lp_ln_akademiker_ant)	3.560	4.410	60.470	< 2e-16 ***
s(ra_laerm_all_log_sum)	1.415	1.735	8.206	0.00057 ***
s(ln_ra_ocl_1_min_1_q)	6.611	7.817	32.020	< 2e-16 ***

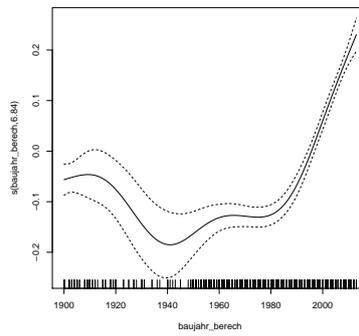
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) = 0.837 Deviance explained = 84%
 GCV score = 0.053063 Scale est. = 0.052027 n = 3285
 AIC = -322.53 BIC = 74.67

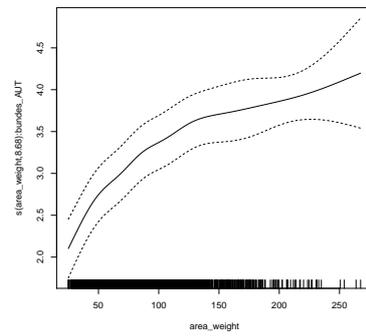
Abbildung B.10: Basismodell ETW nicht hybrid, mit Raster



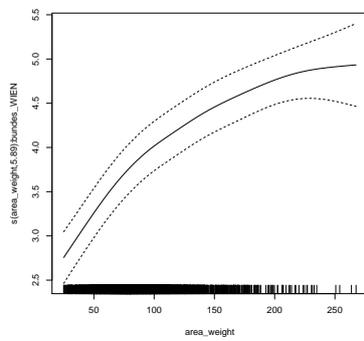
(a) Transaktionsdatum



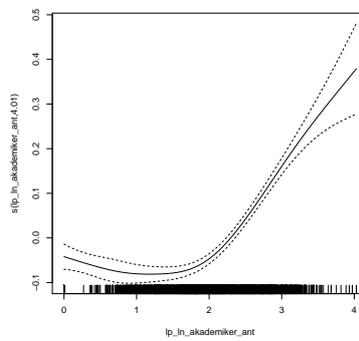
(b) Berechnetes Baujahr



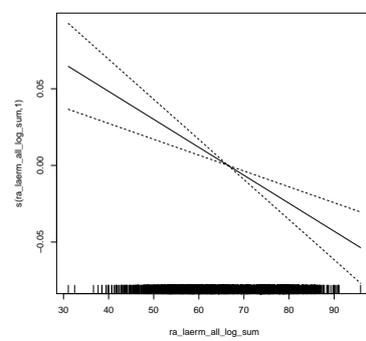
(c) Gewichtete Nutzfläche Österreich



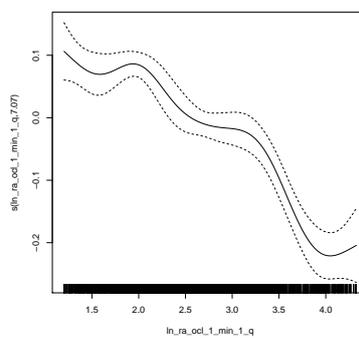
(d) Gewichtete Nutzfläche Wien



(e) Akademikeranteil



(f) Lärm



(g) Erreichbarkeit

Abbildung B.11: Nichtlineare Funktionen des Basismodells ETW nicht hybrid, mit Raster

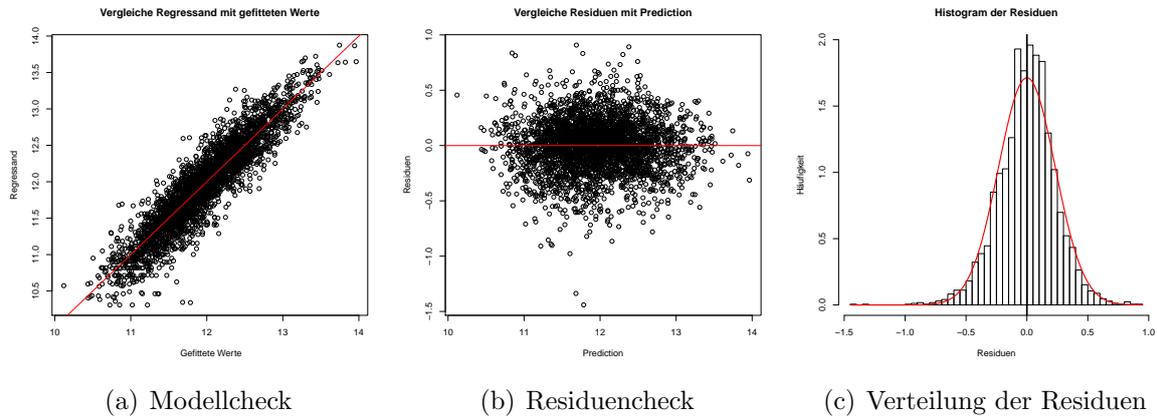


Abbildung B.12: Modellqualität des Basismodells ETW nicht hybrid, mit Raster

B.5.3 Das Basismodell hybrid, ohne Raster

B.5.3.1 EFH

Family: gaussian

Link function: identity

Formula:

```
lnp ~ +dum_angebot + dum_gemeinnutz + dum_bewerter + s(area_weight) +
s(area_gst, by = bundes_AUT) + s(area_gst, by = bundes_WIEN) +
s(kp_year) + s(baujahr_berech) + s(zsp_iu_lnefh_2012) + s(zsp_am_lnefh_2012) +
rh + antwortid_heiz_efh + antwortid_bad_qual + bad_size_kat +
zusatzbad + terr_wohn_vorh + antwortid_zustand_haus + bundes
```

Parametric coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	7.233330	0.043872	164.872	< 2e-16	***
dum_angebot	0.096382	0.012713	7.581	4.28e-14	***
dum_gemeinnutz	-0.100171	0.039419	-2.541	0.011088	*
dum_bewerter	0.035108	0.008387	4.186	2.90e-05	***
rh	0.054621	0.014102	3.873	0.000109	***
antwortid_heiz_efh2	-0.091254	0.013928	-6.552	6.45e-11	***
antwortid_bad_qual1	0.117856	0.014637	8.052	1.08e-15	***
antwortid_bad_qual3	-0.051049	0.009453	-5.400	7.07e-08	***
bad_size_kat1	-0.053939	0.012274	-4.394	1.14e-05	***
zusatzbad	0.027677	0.008171	3.387	0.000713	***
terr_wohn_vorh	0.052105	0.009632	5.410	6.71e-08	***
antwortid_zustand_haus1	0.303065	0.017254	17.565	< 2e-16	***
antwortid_zustand_haus2	0.184537	0.015345	12.026	< 2e-16	***
bundesBGL	-0.014930	0.018821	-0.793	0.427678	
bundesKTN	0.020154	0.017863	1.128	0.259285	
bundesNOE	0.068671	0.013587	5.054	4.53e-07	***
bundesOOE	0.041806	0.017047	2.452	0.014234	*
bundesSBG	0.030546	0.033402	0.914	0.360514	
bundesT	0.095734	0.027250	3.513	0.000448	***
bundesVBG	0.098406	0.032466	3.031	0.002454	**
bundesW	2.465429	0.060711	40.609	< 2e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

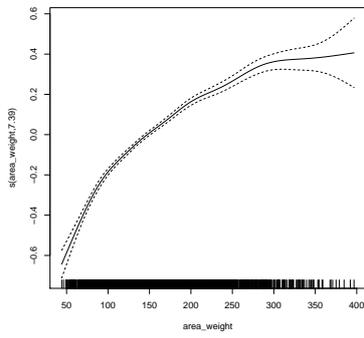
Approximate significance of smooth terms:

	edf	Ref.df	F	p-value	
s(area_weight)	7.392	8.385	192.35	< 2e-16	***
s(area_gst):bundes_AUT	5.306	6.392	2229.83	< 2e-16	***
s(area_gst):bundes_WIEN	2.363	2.882	297.33	< 2e-16	***
s(kp_year)	2.306	2.902	15.66	7.20e-10	***
s(baujahr_berech)	6.088	7.111	93.06	< 2e-16	***
s(zsp_iu_lnefh_2012)	6.154	7.363	11.87	1.06e-15	***
s(zsp_am_lnefh_2012)	1.000	1.000	66.09	5.79e-16	***

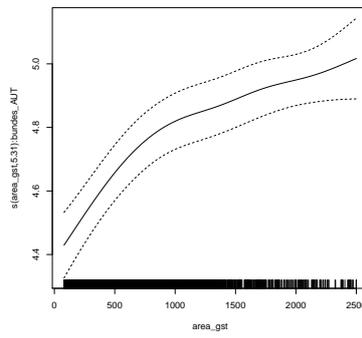
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) = 0.841 Deviance explained = 84.3%
GCV score = 0.045351 Scale est. = 0.044752 n = 3833
AIC = -977.81 BIC = -655.19

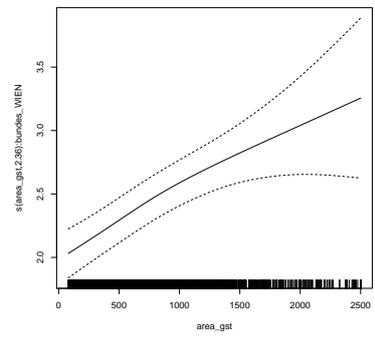
Abbildung B.13: Basismodell hybrid, ohne Raster EFH



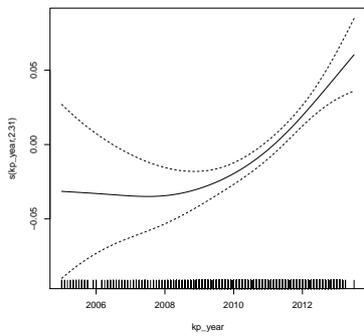
(a) Gewichtete Nutzfläche



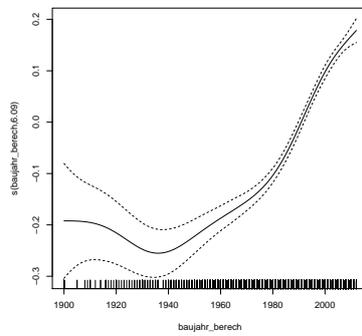
(b) Grundstücksfläche Österreich



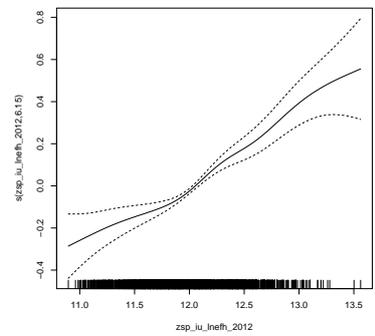
(c) Grundstücksfläche Wien



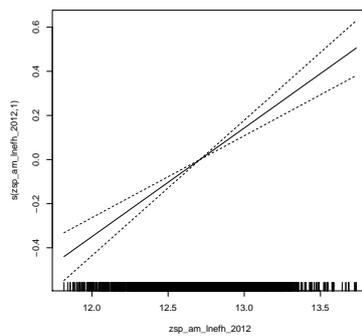
(d) Transaktionsdatum



(e) Berechnetes Baujahr



(f) IU-Index (EFH)



(g) Ametanet-Index (EFH)

Abbildung B.14: Nichtlineare Funktionen des Basismodells hybrid, ohne Raster (EFH)

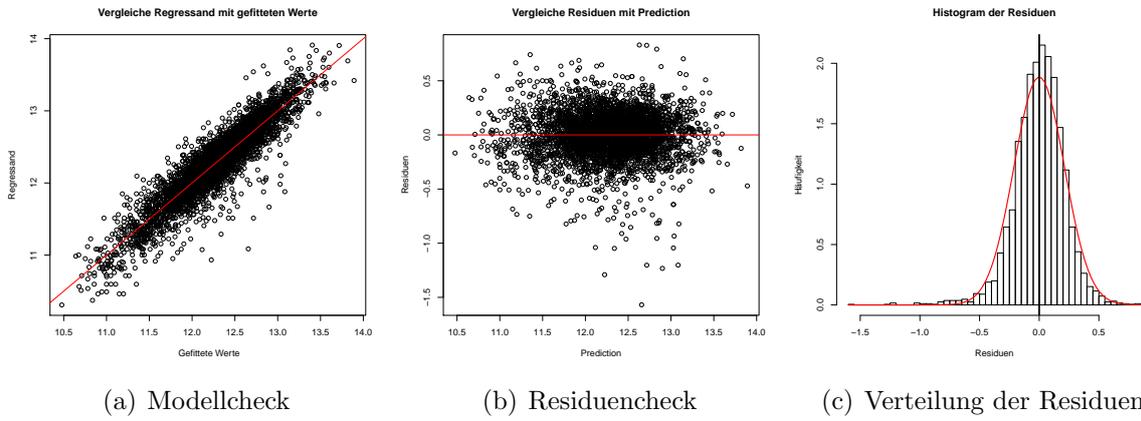


Abbildung B.15: Modellqualität des Basismodells hybrid, ohne Raster (EFH)

B.5.3.2 ETW

Family: gaussian
Link function: identity

Formula:

```
lnp ~ dum_angebot + dum_gemeinnutz + s(kp_year) + s(baujahr_berech) +
  antwortid_bad_qual + bad_size_kat + zusatzbad + antwortid_zustand_haus +
  s(area_weight, by = bundes_AUT) + s(area_weight, by = bundes_WIEN) +
  s(zsp_am_lnetw_2012) + s(zsp_iu_lnetw_2012_q) + antwortid_zustand_wohnung +
  antwortid_heiz_etw + balc_vorh + lift_up + antwortid_garage_3 +
  bundes
```

Parametric coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	7.850250	0.127167	61.732	< 2e-16 ***
dum_angebot	0.057100	0.012354	4.622	3.95e-06 ***
dum_gemeinnutz	-0.077008	0.020274	-3.798	0.000148 ***
antwortid_bad_qual1	0.078636	0.018868	4.168	3.16e-05 ***
antwortid_bad_qual3	-0.064207	0.010172	-6.312	3.12e-10 ***
bad_size_kat1	-0.021918	0.010252	-2.138	0.032589 *
zusatzbad	0.042865	0.016683	2.569	0.010233 *
antwortid_zustand_haus1	0.230545	0.044760	5.151	2.75e-07 ***
antwortid_zustand_haus2	0.166514	0.044381	3.752	0.000179 ***
antwortid_zustand_wohnung1	0.104170	0.012036	8.655	< 2e-16 ***
antwortid_zustand_wohnung3	-0.059189	0.013081	-4.525	6.26e-06 ***
antwortid_zustand_wohnung4	-0.169070	0.031107	-5.435	5.88e-08 ***
antwortid_heiz_etw2	-0.061416	0.015375	-3.994	6.63e-05 ***
balc_vorh	0.023887	0.008493	2.812	0.004947 **
lift_upTRUE	0.067362	0.008920	7.552	5.54e-14 ***
antwortid_garage_3	0.043964	0.009480	4.638	3.66e-06 ***
bundesBGL	0.458483	0.060924	7.525	6.77e-14 ***
bundesKTN	0.376144	0.032481	11.581	< 2e-16 ***
bundesNOE	0.449092	0.030574	14.689	< 2e-16 ***
bundesOOE	0.458933	0.032961	13.924	< 2e-16 ***
bundesSBG	0.457077	0.034724	13.163	< 2e-16 ***
bundesSTMK	0.418210	0.030910	13.530	< 2e-16 ***
bundesT	0.500877	0.030941	16.188	< 2e-16 ***
bundesVBG	0.500683	0.036984	13.538	< 2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

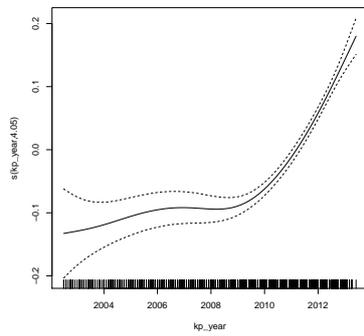
Approximate significance of smooth terms:

	edf	Ref.df	F	p-value
s(kp_year)	4.045	4.981	81.288	< 2e-16 ***
s(baujahr_berech)	4.969	6.002	93.728	< 2e-16 ***
s(area_weight):bundes_AUT	8.405	9.244	515.553	< 2e-16 ***
s(area_weight):bundes_WIEN	5.471	6.615	783.481	< 2e-16 ***
s(zsp_am_lnetw_2012)	8.681	8.966	24.783	< 2e-16 ***
s(zsp_iu_lnetw_2012_q)	6.828	8.001	4.835	6.11e-06 ***

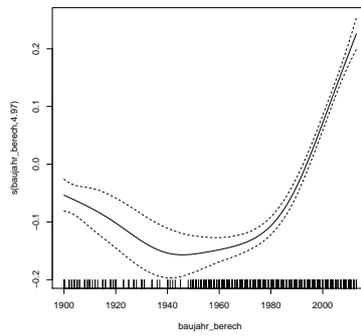
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) = 0.862 Deviance explained = 86.5%
GCV score = 0.044719 Scale est. = 0.043887 n = 3285
AIC = -884.38 BIC = -505.56

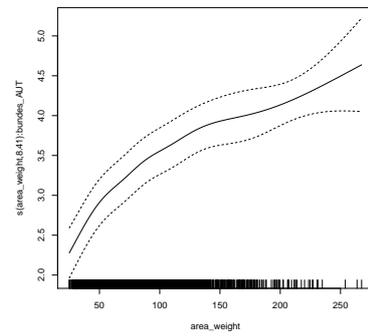
Abbildung B.16: Basismodell hybrid, ohne Raster ETW



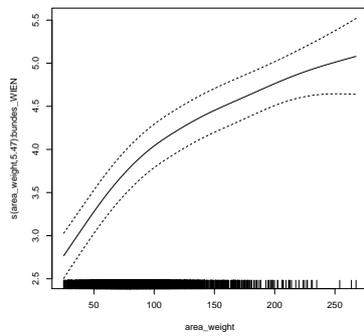
(a) Transaktionsdatum



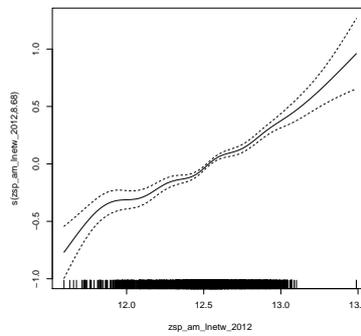
(b) Berechnetes Baujahr



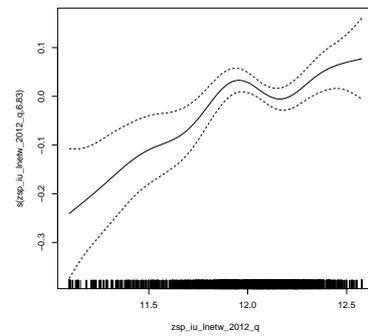
(c) Gewichtete Nutzfläche Österreich



(d) Gewichtete Nutzfläche Wien

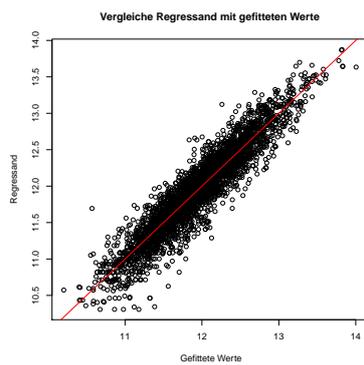


(e) Ametanet-Index (ETW)

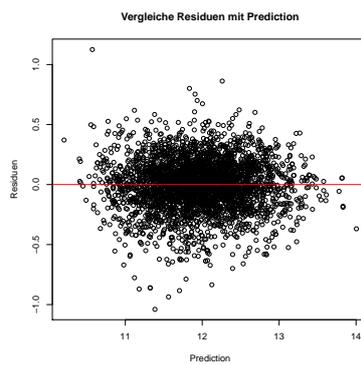


(f) IU-Index (ETW)

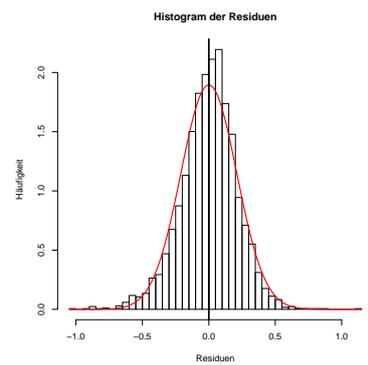
Abbildung B.17: Nichtlineare Funktionen des Basismodells hybrid, ohne Raster (ETW)



(a) Modellcheck



(b) Residuencheck



(c) Verteilung der Residuen

Abbildung B.18: Modellqualität des Basismodells hybrid, ohne Raster (ETW)

Curriculum Vitae

Anschrift/Kontakt

Dipl.-Ing. Ronald Simon Weberndorfer

1160 Wien, Austria

Email: weberndorferronald@gmx.at

Ausbildung

1988 bis 1996:

Volksschule und Hauptschule Ottensheim

1996 bis 2001:

Höhere technische und gewerbliche Bundeslehranstalt für Forstwirtschaft (HBLAF) in Gainfarn, Bad Vöslau (Försterschule)

2001 bis 2008:

Studium Raumplanung und Raumordnung an der TU Wien (Dipl.-Ing.)

Diplomarbeit: Auswirkungen von Naturgefahren auf Immobilien- und Grundstückspreise

2008:

Prämierung der Diplomarbeit durch den Ausbildungspreis 2008 der FMA und IFMA Austria

2009:

Verleihung eines Sonderpreises (GEFMA Förderpreis 2009) in der Fachkategorie Risikoanalyse für die Diplomarbeit

2010 bis 2013:

Dr.-Studium der techn. Wissenschaften Raumplanung und Raumordnung an der TU
Wien

Jobs

Bis 2005:

Diverse Schulpraktika und sonstige Studentenjobs.

2005 bis 2008:

Selbstständiger Gewerbeinhaber (Dienstleistungen in der automatischen Datenverarbeitung und Informationstechnik).

Projektbasierende Zusammenarbeit mit:

- ALPINE ENERGIE GmbH & Co.KG
- VALOR Financial Services GmbH
- Sonstige Kleinaufträge

Seit 2008:

Techniker für statistische Analysen und Geoinformatik.

IRG – Immobilien Rating GmbH (Ein Unternehmen der Bank Austria UniCredit),
Wien