

# Weiterentwicklung der AKS – Implementierung neuer Auswertemethoden und Steigerung der Nutzer-Interaktion

Matthias Soot, Sebastian Zaddach, Alexandra Weitkamp, René Käker und Marcel Ziems

## Zusammenfassung

Die Kaufpreissammlung der Gutachterausschüsse ist die Basis für die Wertermittlung und Markttransparenz des deutschen Grundstücksmarkts. Eine ständige Weiterentwicklung des Aufbaus und der verwendeten Methodik zur Auswertung an die Marktgegebenheiten und Erkenntnisse der Forschung ist notwendig.

Diese Arbeit stellt ein gemeinsames Projekt des Landes Niedersachsens und der TU Dresden vor. Im Rahmen des Vorhabens wurde die Auswertekomponente der AKS Niedersachsen weiterentwickelt und insbesondere die Nutzerinteraktion, die Visualisierung, die Überprüfung von Modellannahmen und die Ausreißerdetektion verbessert. Dem Nutzer werden in einem Expertenmodus Metainformationen zu Kauffällen sowie Entscheidungshilfen angezeigt. In dem derzeit in der Entwicklung befindlichen vollautomatisierten Modus sind diese Entscheidungen vordefiniert und führen so zu einem einheitlichen Ergebnis. Mit diesem neuen Modul der AKS Niedersachsen steht zukünftig eine weitere Komponente zur Verfügung, die die Auswertung von Kauffällen optimiert und damit einen zusätzlichen Beitrag zur Markttransparenz leistet.

## Summary

*The purchase price collection is basis of real estate valuation and general market transparency of real estate market in Germany. A continuous development of the structure and the used methods of investigating purchase prices on the basis of changing market conditions and scientific findings is necessary.*

*In this paper, we present a joint project between the federal state of Lower Saxony and the TU Dresden. Within the project, the analyzing component of the AKS Niedersachsen has been developed. Especially, the user-interaction, the data visualization, the test of model assumptions and the detection of outliers was improved. Meta-information for the purchase prices and decision support is displayed in experts' mode. However, in a fully automated version, the decision borders are fixed to specific values. This leads to a better comparability of results. With the new module of the AKS Niedersachsen, a new component for market investigation is available. It has the potential to contribute to a higher market transparency.*

**Schlüsselwörter:** AKS, Automatisierte Kaufpreissammlung, Statistische Methoden, Nutzer-Interaktion

## 1 Einführung

Den vom Gesetzgeber vorgesehenen Beitrag zur Markttransparenz leisten die Gutachterausschüsse durch die Erfassung, Analyse und Veröffentlichung von Informationen zum Immobilienmarktgeschehen, u. a. Bodenrichtwerte und sonstige zur Wertermittlung erforderliche Daten. Die Transparenz auf dem Grundstücksmarkt wird letztlich durch die Qualität der Datenhaltung und der anknüpfenden Auswertemethoden wesentlich mitbestimmt.

Die Führung und Auswertung der Kaufpreissammlung ist Kernaufgabe der Gutachterausschüsse (§ 193 Abs. 5 Baugesetzbuch (BauGB)) – heute in digitaler Form. Neben der Datenbank mit Erfassungseinheit für Kauffälle bieten einige Kaufpreissammlungen auch die Möglichkeit der Kauffallanalyse.

Im Rahmen des gemeinsamen Projektes »Auswertemethoden der Automatisierten Kaufpreissammlung (AKS)« des Ministeriums für Inneres und Sport des Landes Niedersachsen und der Technischen Universität Dresden sollte daher eine Weiterentwicklung dieser Analysemethode untersucht und implementiert werden. Ziel war es, innovative Auswertemethoden für wertermittlungstechnische Fragestellungen zu untersuchen und zu implementieren. Darüber hinaus wurde die Nutzerinteraktion verbessert, um die Bedienbarkeit und Auswertung zu erleichtern. Die Implementierung erfolgt modular, um auch zukünftig eine Erweiterbarkeit zu gewährleisten. Im Folgenden wird nach Einführung in die Automatisierte Kaufpreissammlung Niedersachsen zunächst das Entwicklungskonzept vorgestellt. Anschließend werden wesentliche Erkenntnisse herausgestellt und darauffolgend anhand eines Beispiels verdeutlicht.

## 2 Die Automatisierte Kaufpreissammlung Niedersachsen

Die erste Version der AKS Niedersachsen hat ihren Ursprung Anfang der 1980er Jahre. Diese ab 1984 schrittweise in die Praxis eingeführte erste Generation der AKS beinhaltete bereits die Philosophie eines vollständigen, integrierten Programmsystems. Die AKS enthält als Kernkomponenten Module zur Informationsbeschaffung, Informationsverarbeitung sowie Abgabe der ausgewerteten Daten mit deren Bereitstellung als wesentliche Produkte der Gutachterausschüsse (Abb. 1).

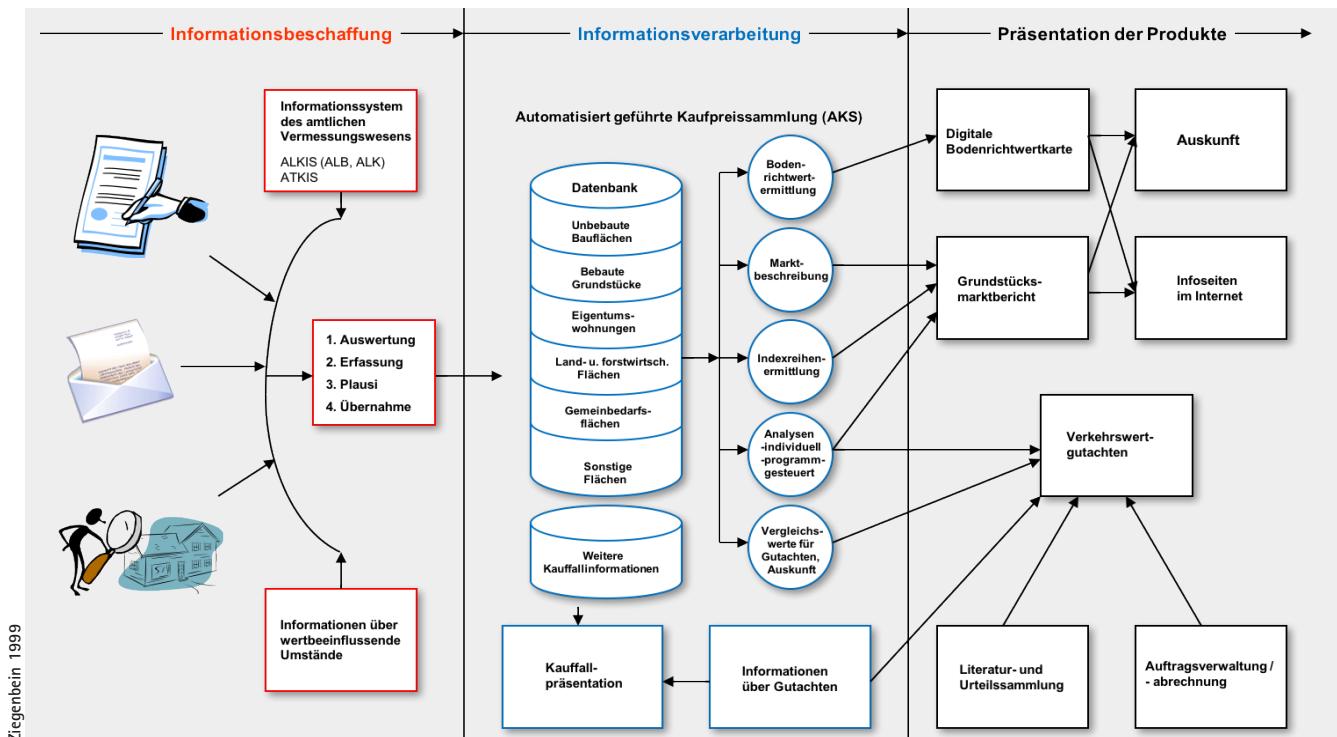


Abb. 1: Komponenten der AKS Niedersachsen

Von der Erhebung der originären Daten (Kaufpreise und weitere Informationen) über deren Prüfung und Plausibilisierung bis hin zur Auswertung und Ausgabe verschiedenster Statistikinformationen wurden die erforderlichen Programmkomponenten nach und nach in die AKS integriert:

- Voraussetzung für die Erfassung von Kauffallinformationen in einer Datenbank ist die **Informationsbeschaffung**. Die Definitionen der Grundstücksmerkmale, welche als potenziell wertrelevant standardmäßig zu erfassen sind, sind in den sogenannten Elementkatalogen zur Führung und Auswertung der AKS (vgl. LGLN 2015) zu finden. In Abhängigkeit der sachlichen Teilmärkte ist eine Mindestanzahl an Elementen vorgegeben, die zwingend in der AKS zu belegen sind, um standardisierte Auswertungen zu ermöglichen. Hierbei handelt es sich sowohl um Ordnungsmerkmale, die die räumliche Lage der Objekte beschreiben, als auch um Vertragsmerkmale, die nähere Angaben zu den spezifischen Eigenschaften eines einzelnen Kauffalls geben. Datengrundlage sind neben den Kaufverträgen die Geobasisdaten des amtlichen Vermessungswesens (bspw. Daten des Amtlichen Liegenschaftskataster-informationssystems (ALKIS) oder des Amtlichen Topographisch-Kartographischen Informationssystems (ATKIS)), örtliche Erhebungen oder Fragebögen, die an die Eigentümer verschickt werden. Neben den obligatorischen Elementen steht es den Gutachterausschüssen frei, ergänzende Informationen zu erheben, die als potenziell wertrelevant eingestuft werden.
- Darauf aufbauend erfolgt die **Informationsverarbeitung**. Getrennt nach sachlichen Teilmärkten können die erfassten Kaufverträge ausgewertet werden. Wich-

tigste Grundlage sind mathematisch-statistische Methoden, die von der deskriptiven Statistik für Zwecke der Stichprobenbeschreibung bis hin zur induktiven Statistik reichen – hier ist insbesondere der Fokus auf die multiple lineare Regressionsanalyse gelegt. Die AKS ermöglicht insofern sowohl die Auswertung einzelner Stichproben hinsichtlich statistischer Kennzahlen als auch die (halb-)automatisierte Ableitung von Indexreihen und Bodenrichtwerten sowie die für Vergleichswert- und Sachwertermittlung notwendigen Parameter. Speziell die multiple lineare Regressionsanalyse hat sich in den Jahrzehnten seit der Einführung der AKS als Standardauswertemethode etabliert. Der automationsgestützte, iterative Ablauf zur Ableitung einer optimalen Regressionsanalyse in der Wertermittlung ist Gegenstand zahlreicher Publikationen (u.a. Ziegenbein 1977, 1986, 1995, 1999). Einen Überblick über die implementierten mathematisch-statistische Methoden in der Verkehrswertermittlung geben insbesondere Brückner (1976) und Koch (1987).

- Im Sinne eines ganzheitlichen Produkts bietet die AKS Niedersachsen als dritte Kernkomponente die Bereitstellung von Analyseergebnissen für die **Präsentation der Produkte** der Gutachterausschüsse. Die Möglichkeiten reichen u.a. von der Erzeugung von Statistiken für die jährlichen Grundstücksmarktberichte über die Bereitstellung für Anwendungen im Internet (z.B. Immobilienpreiskalkulator) bis hin zu individuellen Auswertungen für die Erstellung von Verkehrswert-gutachten.

Dieser Artikel präsentiert vordergründig die Weiterentwicklung im Bereich der Informationsverarbeitung.

### 3 Das Projekt »Weiterentwicklung der AKS«

#### 3.1 Ausgangslage

Seit Einführung der AKS Niedersachsen wurden ca. 8 Millionen Kauffälle in den verschiedenen Teilmärkten erfasst. Allein in Niedersachsen werden in der Datenbank jährlich etwa 100.000 bis 150.000 Fälle neu erfasst. Diese stehen grundsätzlich für weitere Analysen und Kaufpreisauskünfte zur Verfügung. Da die Datenstruktur im

erfolgte eine Anbindung an das Open-Source GIS-System OpenJUMP. Zukünftig soll die Anbindung jedoch an die Open-Source-Software Q-GIS erfolgen, da diese neben einem größeren Funktionsumfang auch eine größere Verbreitung und Community zur Weiterentwicklung und Wartung hat. Ferner wurde die Schaffung einer Objektdatenbank zur Verwaltung von Kauffallfotos realisiert. Auch der Export von Daten und Auswertungen wurde in den letzten Jahren eingeführt.

Anpassungen im Bereich der Methodik der Auswerte-komponenten sind seit der Einrichtung der AKS nicht erfolgt: Hier besteht insbesondere Nachholbedarf im Bereich der Interaktion, der Visualisierung und der Verwendung von statistischen Tests – nicht zuletzt durch gestiegene Anforderungen an die Wertermittlung und einen erhöhten Bedarf an zeitnahen und zuverlässigen Markt-informationen. Im Fokus der Forschung standen daher in den letzten Jahren insbesondere alternative funktionale Modellierungen, bspw. durch die einfache Kollokation, deren erste Ansätze bereits durch Pelzer (1978) sowie Ziegenbein und Hawerk (1978) erprobt und durch Zaddach und Alkhatab (2013, 2014a, 2014b) aufgegriffen wurden. Des Weiteren wurde das Forschungsfeld der Bayes-Theorie für die Wertermittlung aufgegriffen; die Forschung umfasst bislang

- die Integration von Expertenwissen in das indirekte Vergleichswertverfahren (u.a. Alkhatab und Weitkamp 2012, Weitkamp und Alkhatab 2012),
- die Behandlung von Unsicherheiten in kaufpreisarmen Lagen über robuste Methoden (u.a. Weitkamp und Alkhatab 2014, Dorndorf et al. 2016, Soot et al. 2017),
- rekursiv anwendbare Schätzungen im indirekten Vergleichswertverfahren mit Auswertestrategie für eine erweiterte Unsicherheitsbetrachtung (u.a. Zaddach und Alkhatab 2013, 2014b, Zaddach 2016) sowie
- die Kombination mit der Fuzzy-Set-Theorie für die zusätzliche Berücksichtigung von Unschärfe im Bewertungsprozess (u.a. Alkhatab et al. 2016).

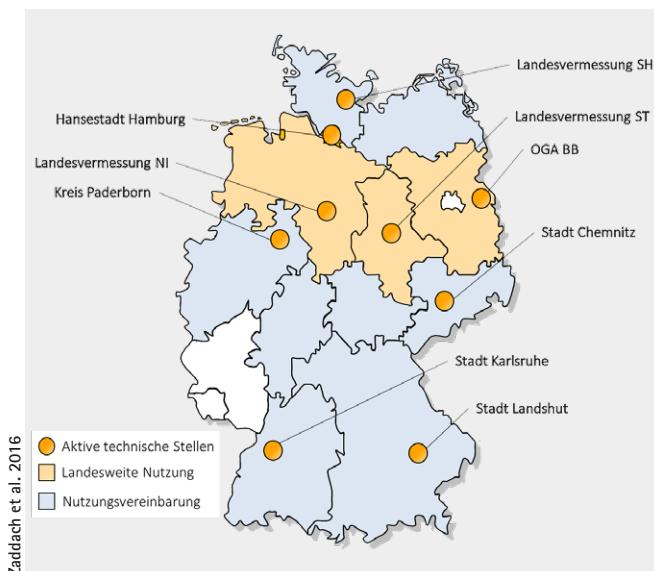


Abb. 2: Verwendung der AKS

Wesentlichen unverändert geblieben ist, konnte eine Überführung von alten Datenbanken als Vorgängerver-sionen i. d. R. verlustfrei stattfinden. In der Vergangenheit wurden sukzessive neue Auswertemodule eingeführt und nicht mehr benötigte entfernt. Neben dem Einsatz in Niedersachsen wird die AKS auch in anderen Bun-desländern teilweise oder flächendeckend verwendet (vgl. Abb. 2).

Zusätzlich zur eigentlichen Führung der Kaufpreissammlung (Erhebung, Bearbeitung, Prüfung und Ver-waltung der originären Kauffälle) enthält die AKS auch Möglichkeiten zur Ausgabe von Mengenstatistiken – bspw. für Grundstücksmarktberichte – oder statistische Auswertungs- und Analysemöglichkeiten. Diese Funktio-nen werden regelmäßig von den Gutachterausschüssen verwendet, um Analysen und Bewertungen durchzufüh-ren. Die Methoden beruhen weitestgehend auf den For-schungsergebnissen aus den 1970er Jahren (vgl. Ziegen-bein 1977).

Die seit ihrer Einführung vorgenommenen, fortlau-fenden Weiterentwicklungen der AKS beziehen sich vornehmlich auf rein technische Aspekte. Die letzte be-deutende Veränderung ist 2008 mit der Einführung der Version 4.0 vollzogen worden: Hier wurde eine grund-legend neue Benutzeroberfläche auf Java-Basis geschaf-fen und eine auf Geoobjekte spezialisierte Datenbank-managementsoftware (DBMS) eingeführt. Des Weiteren

In sämtlichen Forschungsarbeiten werden innovative sta-tistische Methoden für die Wertermittlung adaptiert und auf Praxisbeispiele in der Verkehrswertermittlung ange-wandt, wobei sich die Arbeiten bislang auf das indirekte Vergleichswertverfahren konzentrieren. Aufbauend auf diesen Ergebnissen bedarf es einer Prüfung der Übertrag-barkeit und Implementierung dieser Ansätze in die Aus-wertesoftware der AKS.

#### 3.2 Handlungsfelder und Projektphasen

Infolge der gewachsenen Ansprüche und Erwartungen an Wertermittlungsdaten – nicht zuletzt ausgelöst durch die Immobilien- und Finanzkrise – hinsichtlich zeitlicher Auflösung und Umfang sowie Unsicherheitsbetrachtun-gen und Genauigkeitsangaben wurden im Rahmen des hier präsentierten Projekts die Möglichkeiten der Kauf-

fallauswertung in der AKS Niedersachsen eruiert und Weiterentwicklungsmöglichkeiten aufgezeigt. Dabei wurde die Notwendigkeit deutlich, die Datenhaltung in ihrer starren Abgrenzung der sachlichen Teilmärkte neu zu konzipieren, die statistischen Auswertemethoden zu überarbeiten bzw. weitere zu implementieren und neue Standardschnittstellen für den Datenimport und -export zu definieren. Auch Überlegungen zu einem veränderten modularen Aufbau des Programmsystems der AKS erscheinen notwendig, um auf künftige Entwicklungen



Abb. 3: Darstellung der Arbeitspakete

flexibel eingehen zu können. Zusammenfassend lassen sich die Schwerpunkte des Projekts in Arbeitspakete untergliedern, die in Abb. 3 dargestellt sind und im Weiteren erläutert werden.

Im Bereich der *deskriptiven Statistik* soll insbesondere die Interpretierbarkeit von Größen gesteigert werden. Zu diesem Zweck kommen zukünftig moderne Visualisierungsmöglichkeiten zum Einsatz.

Der Bereich der *schließenden/induktiven Statistik* ist ebenfalls erweitert worden. Für die Auswertung der Kauffalldaten mit der Methodik der Regressionsanalyse bedarf es des Einhaltens – und insofern des Überprüfens – von Modellannahmen sowie einer aus statistischer Sicht optimalen Modellwahl: Im Programmsystem der AKS wurden vielfach pragmatische Ansätze implementiert, die durch heutige Erkenntnisse und Methoden ersetzt werden können. Dabei ist insbesondere der Umgang mit Ausreißern in der AKS weiterzuentwickeln: Gegebenenfalls liegen nur wenige Daten vor, sodass einerseits eine eindeutige Identifikation nicht möglich ist und andererseits durch Elimination von Ausreißern nicht auf Informationen verzichtet werden soll. Robuste Verfahren können hier Lösungswege bieten. Aufgrund der Wertermittlungsdaten ist ggf. für die Nutzung der multiplen linearen Regression eine Transformation von Daten notwendig – auch hier können neuere Erkenntnisse Eingang in die AKS finden.

## 4 Prozeduraler Ablauf

Das implementierte Auswertemodul verfügt über einen prozeduralen Ablauf, indem iterativ einzelne Schritte wiederholt werden können, bis das bestmögliche Modell identifiziert wird und alle außergewöhnlichen Kauffälle eliminiert wurden. Die neu entwickelten Module werden

anhand einer Auswertung einer realen Stichprobe mit der weiterentwickelten AKS Niedersachsen dargestellt.

### 4.1 Selektion der Stichprobe

In der AKS Niedersachsen kann nach allen Angaben, die zu einem Kauffall erfasst wurden, selektiert werden. Die Selektion erfolgt in erster Linie entsprechend der Erfahrung des Auswerts. Die Bildung von Teilmärkten stellt hierbei die größte Herausforderung dar (Liebig und Ache 2014). Für eine Auswertung mittels Regressionsanalyse muss im gesamten untersuchten Teilmarkt ein gleiches Marktverhalten vorherrschen. Typischerweise sind die Grenzen zwischen Teilmärkten nicht scharf abgrenzbar: Beispielsweise muss im Teilmarkt der Ein- und Zweifamilienhäuser i. d. R. eine Abgrenzung zum hochpreisigen Markt der Villen erfolgen. Eine vordefinierte Grenze existiert nicht. Daneben stellt auch die Abgrenzung räumlicher Teilmärkte insbesondere durch die Grenzen der Zuständigkeitsbereiche stets eine große Herausforderung dar.

In der AKS kann zunächst eine ausgedehnte Selektion erfolgen und untersucht werden, ob Kaufpreise in unterschiedlichen Lagen dem gleichen Teilmarkt zuzuordnen sind. Für dieses Beispiel wird der Teilmarkt der freistehenden Ein- und Zweifamilienhäuser in der Stadt Nienburg (Weser) selektiert. Hierbei wird der Teilmarkt der Nachkriegsgebäude (nach 1945) untersucht. Die Stichprobe umfasst 317 Kauffälle im Zeitraum 2013 bis 2015.

### 4.2 Selektion der Modellparameter

Die klassischen Regressionsmodelle im Bereich der Wertermittlung enthalten i. d. R. stets wiederkehrende Regressionsparameter. In den Marktberichten des Landes Niedersachsen finden sich für das Jahr 2017 bspw. nur 13 verschiedene Einflussgrößen zur Ableitung sämtlicher Marktdaten (Vergleichsfaktoren, Sachwertfaktoren und Liegenschaftszinssätze). Aus diesem Grund wurden zur Verbesserung der Bedienerfreundlichkeit eine Auswahl der regelmäßig auftretenden Parameter über Checkboxen und Dropdownfelder implementiert – unter Angabe der Anzahl der Belegung der Felder in der Datenbank (vgl. Abb. 4). Weiterhin ist eine eigene Zusammenstellung der Einfluss- und Zielgrößen (auch durch Kombinationen/Aggregationen verschiedener Größen) möglich.

### 4.3 Aufbereitung der Daten

Nach einer Selektion eines Datensatzes müssen zunächst die Daten aufbereitet werden. Schon hier können außergewöhnliche Kauffälle erkannt werden, die ggf. einem anderen Teilmarkt zuzuordnen sind.

### 4.3.1 Korrelogramm

Im neu entwickelten Modul der AKS kann die Stichprobe auf verschiedene Weise visualisiert werden. Zu der gewählten Stichprobe kann zunächst ein Korrelogramm dargestellt werden (vgl. Abb. 5). Hierzu wird die Korrelation zwischen allen Größen visualisiert. Des Weiteren wird der partielle Korrelationskoeffizient angegeben, der

Abb. 4: Darstellung der Auswahl der Parameter mittels Checkboxen und Drop-Down-Feldern

die Korrelation der beiden Größen ohne den Einfluss der anderen Größen darstellt. Im linken Teil der Abb. 5 ist der Fortgang des prozeduralen Ablaufs mit grünem Rahmen dargestellt.

Untereinander hoch korrelierte Größen sollten ausgeschlossen werden. In diesem Fall sind, wie zu erwarten, die Restnutzungsdauer und das Baujahr (partieller Korrelationskoeffizient von 0,83) sowie die Wohnfläche und die Bruttogrundfläche (partieller Korrelationskoeffizient von 0,7) hoch korreliert. Der verbleibende Parameter kann anhand des höheren partiellen Korrelationskoeffizienten mit der Zielgröße gewählt werden. In diesem Fall ist die Restnutzungsdauer mit dem Kaufpreis pro Wohnfläche höher korreliert (0,18) als das Baujahr (0,07) und die Wohnfläche (-0,55) höher als die Bruttogrundfläche (0,26).

### 4.3.2 Deskriptive Statistik

Im weiteren Schritt wird dem Nutzer die deskriptive Statistik aller Größen dargestellt. Hierzu werden neben den klassischen Größen wie Mittelwert, Median, Minimum, Maximum und Spannweite (vgl. Tab. 1) auch verschiedene weitere Visualisierungen angeboten.

Über die bisher schon mögliche Visualisierung als Histogramm hinaus kann die

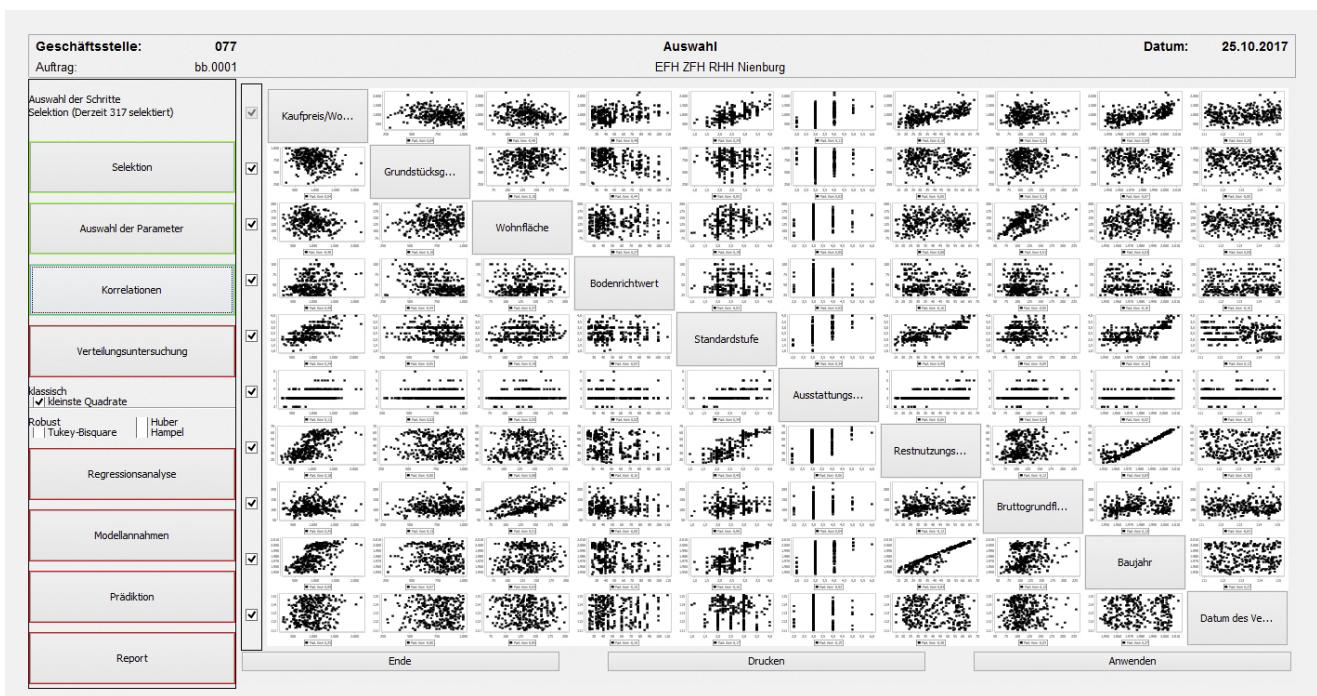


Abb. 5: Darstellung eines Korrelogramms

Tab. 1: Darstellung der deskriptiven Statistik der selektierten Stichprobe

Größe	Mittelwert	Median	Minimum	Maximum	Spanne
Kaufpreis/Wohnfläche	965,41	941,18	266,67	2164,95	1898,28
Wohnfläche	134,46	135,00	70	200	130
Grundstücksfläche	755,43	781,00	727,00	1000	728
Bodenrichtwert	50,17	42,00	25	110	85
Standardstufe	2,55	2,5	1	4	3
Restnutzungsdauer	39,73	36	15	68	53
Datum des Vertrages	2013	2013	2011	2015	4

Verteilung als Boxplot und QQ-Plot oder Tabelle dargestellt werden:

- Boxplots haben in diesem Zusammenhang den Vorteil, dass sie Kauffälle, die in den Verteilungen außergewöhnlich, d.h. weiter als die 1,5-fache Spanne des 50 % Quantils, sind, bereits als Ausreißer darstellen (Abb. 6 links). Diese sind mit einem Dreieck gekennzeichnet.
- In QQ-Plots werden die Quantile der untersuchten Verteilung gegen die Quantile einer vorgegebenen Verteilung (i. d. R. Normalverteilung) dargestellt. Durch den Verlauf der Geraden kann die Form der Verteilung festgestellt werden; schnell können Abweichungen der untersuchten Verteilung von der vorgegebenen visuell festgestellt werden (Abb. 6 rechts). Ausreißer sind durch ihr Abweichen der roten von der blauen Kurve erkennbar.

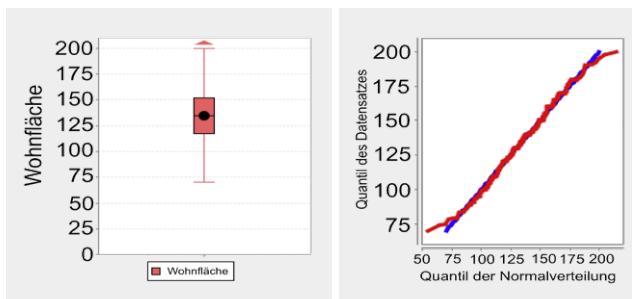


Abb. 6: Boxplot (links) und QQ-Plot (rechts)

Alle Darstellungen im neuen Modul der AKS Niedersachsen sind interaktiv gestaltet und können dazu genutzt werden, um Kauffälle im weiteren Verlauf der Auswertung als nicht zum Teilmarkt gehörend auszuschließen. Der Ausschluss von Kauffällen oder Spannen kann sowohl in den Grafiken als auch in den Tabellen geschehen. Die Kauffälle können auch zunächst in der weiteren Auswertung verbleiben und im Nachgang der Regressionsanalyse ggf. als außergewöhnlich identifiziert werden.

#### 4.4 Transformation der Daten

Viele Gutachter transformieren zu diesem Zeitpunkt die nicht normalverteilten Daten (Ziel- wie auch Einflussgrößen). In der Praxis der Wertermittlung hat sich diese eher pragmatische Vorgehensweise etabliert. Aus mathematisch-statistischer Sicht sei jedoch darauf hingewiesen, dass eine Transformation nur aufgrund nicht normalverteilter Daten nicht notwendig ist. Eine Transformation der Daten führt zu einem nicht linearen Zusammenhang zwischen Einfluss- und Zielgrößen. Daher ist eine Transformation nur zu empfehlen, wenn auch eine stichhaltige Begründung vorliegt – hier herauszustellen ist bspw. die fehlende Normalverteilung der Residuen.

Außerhalb der Wertermittlung erfolgen Transformationen vornehmlich über Exponenten oder den Logarithmus. Daneben erfolgt die Transformation häufig nur auf den Einflussgrößen. Da Wertermittlungsdaten i. d. R. einen positiven numerischen Wert aufweisen (alle Einflussgrößen  $>0$ ), muss die Transformation auch i. d. R. nur für diesen Bereich betrachtet werden. In Abb. 7 werden beispielhaft die Funktionen dargestellt, welche durch Nutzung einer exponentiellen oder logarithmischen Transformation linearisierbar sind. Erfolgt die Transformation auf der Zielgröße, wird der Zusammenhang zu allen Einflussgrößen nichtlinear. Sogenannte doppelt logarithmische Funktionen verwenden das statistische Bundesamt im Regressionsmodell des Häuserpreisindex; hier werden sowohl

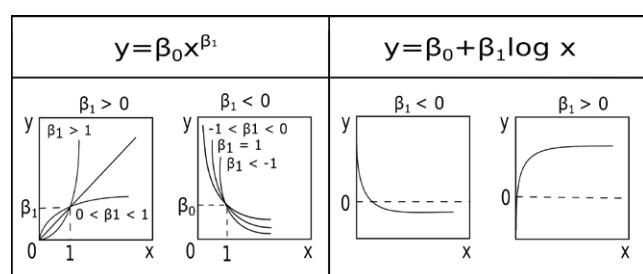


Abb. 7: Linearisierbare Funktionen. Links: Funktionsverläufe für eine Transformation der Zielgröße mit dem Exponenten. Rechts: Funktionsverläufe für eine Transformation mit dem Logarithmus.

die Ziel- als auch die Einflussgrößen transformiert ([www.destatis.de/DE/ZahlenFakten/GesamtwirtschaftUmwelt/Preise/Verbraucherpreisindizes/Methoden/Downloads/HedonischeMethodenHaeuserpreisindexWista05\\_2004.pdf](http://www.destatis.de/DE/ZahlenFakten/GesamtwirtschaftUmwelt/Preise/Verbraucherpreisindizes/Methoden/Downloads/HedonischeMethodenHaeuserpreisindexWista05_2004.pdf)).

Sofern die Residuenanalyse (Kap. 4.7, »Prüfung der Modellannahmen«) für eine Nichtlinearität Indizien liefert (basierend auf den Tests auf Normalverteilung), sollte in der Regel bei einer Auswertung zunächst mit einem linearen Modell gestartet und ggf. sollten zunächst nur einzelne Einflussgrößen transformiert werden. Nicht zu vernachlässigen ist, dass für die Transformation eine nachvollziehbare Begründung durch ein derartiges Marktverhalten existieren muss (Notwendigkeit der physikalischen Begründbarkeit). Für die Aufbereitung der Daten für den Marktbericht (Visualisierung in Grafiken) müssen die transformierten Funktionen zurücktransformiert werden.

## 4.5 Modellwahl

Ein Modell ist die Abstraktion der Wirklichkeit durch Beschreibung mit bekannten, quantifizierbaren Größen (Fahrmeir et al. 2007). Hier bedarf es der Nutzung von Gütemaßen, um zu entscheiden, ob eine Kombination von Einflussgrößen ein Modell optimal zu beschreiben vermag. Die Kombinationen werden strategisch auf das Optimum überprüft: Dies erfolgt i. d. R. durch Rückwärts-, Vorwärts- oder schrittweise Strategien (Chatterjee und Price 1995, S. 215).

### 4.5.1 Gütemaße

In der Wertermittlung, wie auch in anderen Fachdisziplinen (Diskussion in Quinino und Reis 2012), wird sehr häufig das Bestimmtheitsmaß (auch  $R^2$ ) als Gütemaß für eine qualitative Aussage über eine Regressionsanalyse genutzt. Die Praxis nutzt das Bestimmtheitsmaß irrtümlich auch, um eine Aussage zum Vergleich verschiedener Modelle (unterschiedliche Ziel- und Einflussgrößen) zu verwenden. Goldberg (1991, S. 177 ff.) zeigt auf, dass das Bestimmtheitsmaß ungeeignet ist, um eine Aussage zum Modellvergleich zu tätigen. Auch hinsichtlich der Entscheidung über die optimale Kombination der Einflussgrößen ist das Bestimmtheitsmaß nur sehr beschränkt verwendbar. Mit Zunahme weiterer Parameter in das Regressionsmodell verteilt sich das Residuum auf alle Parameter. Somit kann durch sogenannte Überparametrisierung – rein mathematisch – ein deutlich höheres  $R^2$  erzielt werden, welches allerdings nicht tatsächlich einen besseren Erklärungsgehalt haben muss. Ein Vergleich von Bestimmtheitsmaßen zur Prüfung der Stärke einer Hypothese ist ebenfalls nicht geeignet. Selbst das oftmals verwendete angepasste Bestimmtheitsmaß sollte i. d. R. kritisch gewürdigt werden. Hier wird bei Hinzunahme von weiteren Parametern in das Regressionsmodell ein Strafterm eingeführt. Dieser Strafterm ist abhängig von

der Anzahl der Einflussgrößen (sowie auch der Kauffälle und der Höhe des Bestimmtheitsmaßes). Er reduziert mathematisch das Bestimmtheitsmaß bei Hinzunehmen weiterer Einflussgrößen. Der Strafterm ist jedoch vielfach zu zahm: Er reduziert das  $R^2$  oftmals nicht in der Größenordnung, in der das Hinzunehmen einer weiteren Einflussgröße es erhöht. Insbesondere bei großen Datensätzen ( $n > 300$ ) und hohen Bestimmtheitsmaßen verliert der Strafterm seine Wirkung. Daher sollte auch das angepasste  $R^2$  mit Vorsicht verwendet werden (Cameron 1993). Der Vorteil des Bestimmtheitsmaßes – speziell für die Praxis – ist lediglich in seiner Interpretierbarkeit zu sehen: Das Bestimmtheitsmaß misst, welcher Anteil der Streuung durch das Modell erklärt wird.

### 4.5.2 Informationskriterien AIC und BIC

Eine geeignete Alternative zur Auswahl des besten Modells (in Bezug auf die verwendeten Daten) liefern die so genannten Informationskriterien. Bei einer explorativen Untersuchung zur Ableitung eines Regressionsmodells in einer Vorwärts-, Rückwärts- oder schrittweisen Strategie liefern diese bessere Aussagen zur Vergleichbarkeit von Modellen. Etabliert haben sich die Parameter AIC (Aikakesches Informationskriterium) und BIC (Bayesisches Informationskriterium).

Während das AIC eher genutzt wird, um allgemeine Aussagen abzuleiten, wird das BIC genutzt, um möglichst genaue Theorieprüfungen durchzuführen. Für das AIC wird i. d. R. eine größere Stichprobe benötigt ( $n/k > 40$  mit  $n$  = Anzahl (Kauf-)Fälle,  $k$  = Anzahl Einflussgrößen). Bei kleinen Stichproben wird die Verwendung eines korrigierten AIC vorgeschlagen (Burnham 2004).

Zur Modellwahl werden dem Nutzer im neu entwickelten Modul neben dem klassischen Bestimmtheitsmaß auch das angepasste Bestimmtheitsmaß sowie auch die Modellwahlkriterien AIC und BIC angegeben. Anhand dieser kann bei der vorliegenden Stichprobe das optimale Modell mit 5 Parametern bestimmt werden (vgl. Tab. 2). Anders als bei dem Bestimmtheitsmaß sollten die Informationskriterien möglichst kleine Größen aufweisen. Das Modell mit den Einflussgrößen Wohnfläche, BRW, Standardstufe, Restnutzungsdauer und (Vertrags-)Datum (Angabe im Dezimaljahr) beinhaltet einerseits signifikante Parameter und weist sowohl das kleinste BIC als auch AIC auf.

In der Expertenversion der AKS wird diese Rückwärtsstrategie durch eine iterative Deaktivierung des schlechtesten Parameters durchgeführt. In einer weiteren, in Entwicklung befindlichen vollautomatischen Version wird das Modell anhand der genannten Kriterien durch das Programm gewählt (vgl. Chatterjee und Price 1995, S. 251).

Das Modell ergibt sich zu:

$$\begin{aligned} \text{KP/WF} = & -4835 - 2,698 \cdot \text{WF} + 6,160 \cdot \text{BRW} + 123,2 \\ & \cdot \text{STST} + 12,5 \cdot \text{RND} + 44,36 \cdot \text{Datum} \end{aligned}$$

Tab. 2: Darstellung der Modellwahlkriterien für verschiedene Modelle in einer Rückwärtsstrategie

Modell	$R^2$	Angepasstes $R^2$	AIC	BIC	Bemerkungen
Grundstücksgröße (F)	0,67	0,66	4538	4508	Grundstücksgröße nicht signifikant bestimmt
Wohnfläche (WF)					
Bodenrichtwert (BRW)					
Standardstufe (StSt)					
Restnutzungsdauer (RND)					
Datum					
Wohnfläche	0,67	0,66	4534	4507	Alle Parameter signifikant bestimmt, schlechterer $ t $ Value für Standardstufe
BRW					
Standardstufe					
RND					
Datum					
Wohnfläche	0,64	0,64	4589	4566	Alle Parameter signifikant bestimmt, schlechterer $ t $ Value für Datum des Kaufvertrages
BRW					
Datum					
RND					
Wohnfläche	0,61	0,60	4616	4596	Alle Parameter signifikant bestimmt, schlechterer $ t $ Value für Wohnfläche
BRW					
RND					
BRW	0,54	0,54	4664	4649	Alle Parameter signifikant bestimmt
RND					

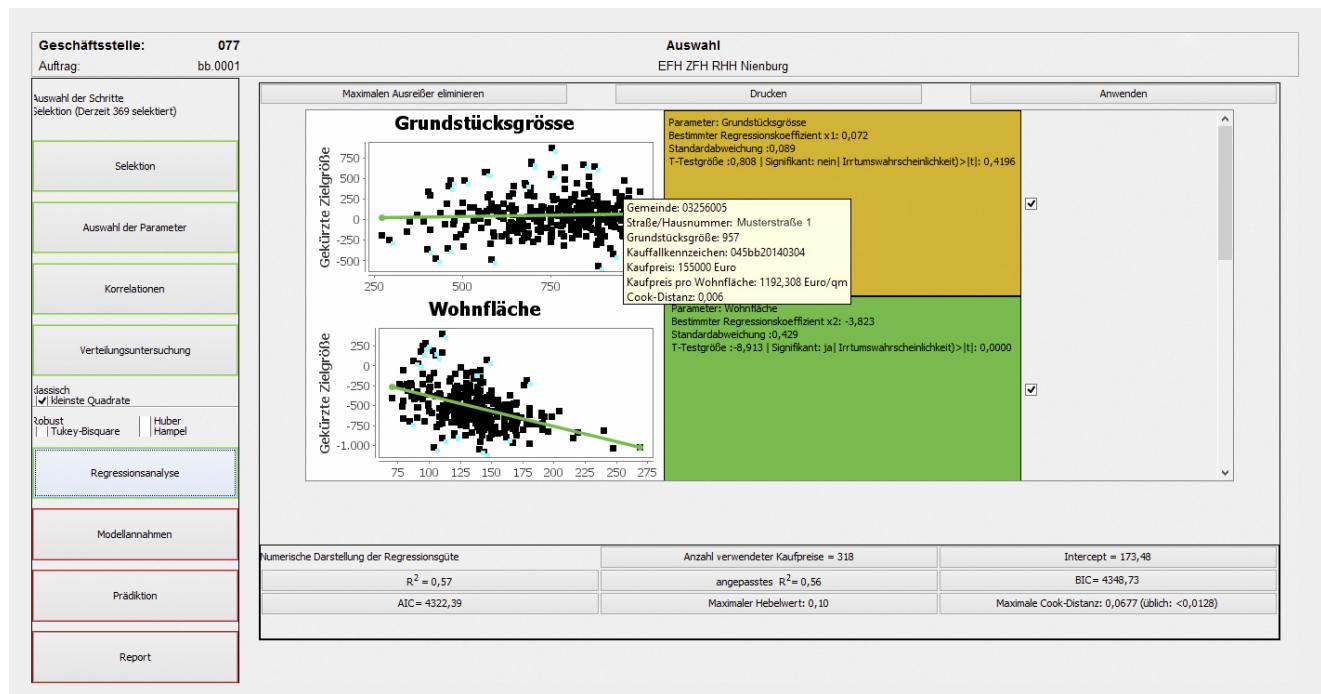


Abb. 8: Darstellung der Regressionsanalyse. Von der Analyse ausgeschlossene Kaufpreise werden als rote Punkte im Streudiagramm dargestellt. Punkte, die eine Cook-Distanz über dem Grenzwert haben, werden mit einem blauen Symbol versehen. Ist der Parameter signifikant bestimmt, wird die deskriptive Statistik grün hinterlegt. Die Gesamtstatistik wird dem Nutzer im unteren Fenster angezeigt.

## 4.6 Ausreißeridentifikation und robuste Methoden

Die Regression darf nicht durch Ausreißer beeinflusst sein. Daher müssen diese – ggf. iterativ – aus der Stichprobe entfernt werden. Ausreißer entstehen u.a. durch ungewöhnliche oder persönliche Verhältnisse. Nach § 7 ImmoWertV wird ein solcher Einfluss angenommen, wenn Kaufpreise erheblich von den Kaufpreisen und anderen Daten in vergleichbaren Fällen abweichen (Ausreißer).

### 4.6.1 Ausreißeridentifikation

Aus mathematisch-statistischer Sicht gibt es verschiedene Methoden, um mit ungewöhnlichen Kaufpreisen umzugehen. Zur Identifikation können bspw. Methoden wie die Cook-Distanz oder ein Data-Snooping nach Baarda (1968) und Pobe (1976) durchgeführt werden:

- Die Cook-Distanz gibt an, wie stark sich das Ergebnis verändert, wenn der jeweilige Kauffall aus der Regressionsanalyse ausgeschlossen würde (Hedderich und Sachs 2012). Hohe Cook-Distanzen weisen darauf hin, dass es sich um einen außergewöhnlichen Kauffall handelt: Sind sie über einem Wert von  $4/n$  (im Beispiel insofern  $>0,013$ ), so liegt ein Indiz vor, dass der Kauffall einen hohen Einfluss auf die Regressionsfunktion hat (Fox und Long 1990).
- Das Data-Snooping nach Baarda und Pobe ist im geodätischen Kontext weit verbreitet. Hierbei werden die normierten Verbesserungen der einzelnen Kauffälle untersucht. Überschreiten diese einen Grenzwert, kann davon ausgegangen werden, dass es sich um einen ungewöhnlichen Kauffall handelt.

Jeder Kauffall, der eine Cook-Distanz über dem Grenzwert aufweist, wird in den Streudiagrammen gekennzeichnet (vgl. Abb. 8). Dieser kann in der Expertenversion durch den Nutzer deaktiviert werden. Hierzu kann iterativ der jeweils maximale Hebelwert deaktiviert und eine neue Regression gerechnet werden, bis alle Cook-Distanzen kleiner als der Grenzwert sind. Alternativ kann die Deaktivierung auch hier in der Grafik erfolgen. Neben der Cook-Distanz werden dem Nutzer in der Mouse-Over-

Funktion weitere Informationen zum jeweiligen Kauffall angezeigt. Alternativ zur Cook-Distanz kann die normierte Verbesserung zur Ausreißeridentifikation genutzt werden (Data-Snooping).

Nach dem Ausschluss der Kauffälle mit auffälligen Cook-Distanzen verbleiben im Anwendungsbeispiel 266 Kauffälle im Datensatz.

### 4.6.2 Robuste Methoden

Neben den Methoden zur Identifikation von Ausreißern – und der Folge, dass diese aus der Analyse ausgeschlossen werden – können auch robuste Methoden genutzt werden, bei denen die außergewöhnlichen Kauffälle an Gewicht verlieren. In Dorndorf et al. (2017) wurden verschiedene robuste Methoden auf ihre Eignung für die Wertermittlung untersucht. Neben der klassischen Regressionsanalyse wurden ein Data-Snooping nach Baarda, der Random Sample Consensus (RANSAC) und eine robuste bayessche Regression an einem simulierten Kauffalldatensatz getestet. Hierbei wurde festgestellt, dass bei großen Datensätzen ( $n > 100$ ) zwar das Data-Snooping die besten Resultate liefert, bei kleineren Stichproben (kaufpreisarme Lagen) sich jedoch die klassischen robusten Verfahren eignen. Sie weisen gegenüber dem Data-Snooping den Vorteil auf, dass die Kauffallinformationen nicht vollständig verloren gehen, sondern vielmehr erhalten bleiben. Lediglich das Gewicht der Einflüsse wird minimiert.

In der AKS sind nunmehr die klassischen robusten Verfahren in Form von M-Schätzern implementiert: Tukey-Bisquare, Huber und Hampel. Die Methoden unterscheiden sich durch eine unterschiedliche Gewichtungsfunktion (vgl. Abb. 9). Während der Huber-Schätzer auch Werten mit extremen Abweichungen ( $\epsilon$ ) noch ein Gewicht ( $w$ ) gibt, geben sowohl der Hampel-Schätzer als auch der Tukey-Bisquare-Schätzer Werten über einem definierten Bruchpunkt kein Gewicht mehr ( $w(\epsilon) = 0$ ). Diese sollten also insbesondere zum Einsatz kommen, wenn mit extremen Ausreißern im Datensatz gerechnet werden kann. Durch Auswahl des entsprechenden Schätzers vor der Regressionsanalyse wird die Berechnung mit der jeweilig

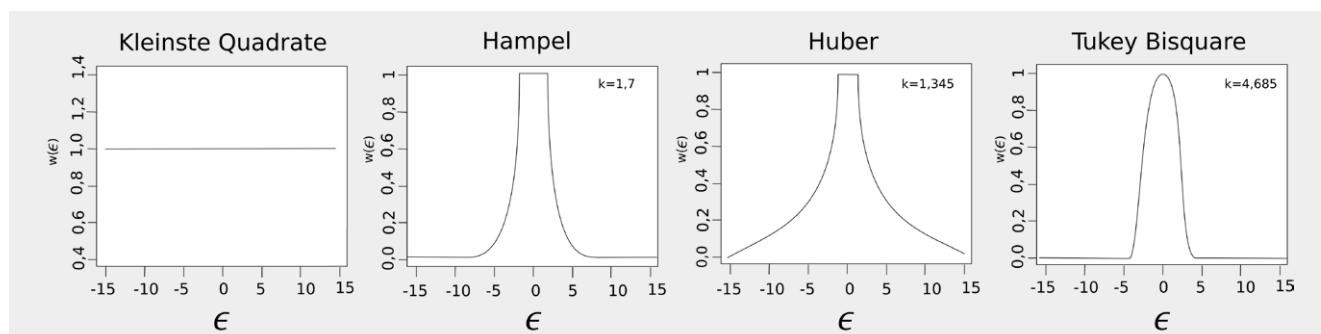


Abb. 9: Darstellung der verschiedenen Gewichtsfunktionen, für die im neuen AKS Modul implementierten M-Schätzer. Links ist die Gewichtsfunktion der klassischen Kleinste-Quadrat-Regression dargestellt. In den drei rechten Grafiken die jeweilige Gewichtsfunktion der robusten M-Schätzer.

aktivierten Methode durchgeführt. Im Streudiagramm werden dann bei der Verwendung der Robusten Methoden anstelle der Hebelwerte der einzelnen Kauffälle die Gewichte im Regressionsmodell angegeben.

#### 4.7 Prüfung der Modelannahmen

Eine Regressionsanalyse unterliegt verschiedenen Modellannahmen, die für eine zielführende Verwendung der Regressionsfunktion spätestens im Nachgang der Analyse zu überprüfen sind. Es sind nach Urban und Mayerl (2011, S. 120 ff.) folgende Bedingungen zu erfüllen:

- Normalverteilung der Residuen ( $\epsilon \sim N(0, \sigma)$ ),
- Homoskedastizität (Gleichverteilung der Varianzen) ( $Var(\epsilon_i) = \sigma^2 = const.$ ) und
- Multikollinearität (Korrelationen der Einflussgrößen untereinander) ( $x_j \neq c_0 + c_1 \cdot x_i, i \neq j$ ).

Mit dem Ziel, Modellverstöße aufzudecken, werden im Berechnungsbeispiel nun die einzelnen Modellannahmen geprüft. Neben den Größen selber wird auch der sogenannte *p*-Wert angegeben. Dieser gibt eine Aussage über die Wahrscheinlichkeit, mit der die Nullhypothese abgelehnt wird. Je kleiner die Testgröße ist, desto höher ist die Wahrscheinlichkeit, dass die Nullhypothese abgelehnt wird.

##### 4.7.1 Normalverteilung der Residuen

Die Normalverteilung der Daten selbst ist – anders als in Kleiber et al. (2017, S. 1455) beschrieben – nicht Voraussetzung für die Regressionsanalyse. Lediglich die Normalverteilung der Residuen ist eine Voraussetzung (vgl. Montgomery et al. 2006, S. 3). Des Weiteren sollten die Residuen insgesamt gleichmäßig verteilt sein und keine Systematiken aufweisen. Systematiken in den Residuen weisen auf Modellfehler hin: Eine tiefer gehende Untersuchung der Residuen ist daher stets zu empfehlen, um Modellverstöße aufzudecken.

Als Prüftest für die Annahme der Normalverteilung der Residuen hat sich der Shapiro-Wilk-Test etabliert. Dieser ist in seiner Testgüte dem ursprünglich in der AKS Niedersachsen verwendeten Chi<sup>2</sup>-Test deutlich überlegen (vgl. Hedderich und Sachs 2012, S. 421 u. 426), da dieser nur Schiefe und Exzess betrachtet, während andere Tests Rangfolgen berücksichtigen und eine empirische Verteilung gegenüberstellen. Der Shapiro-Wilk-Test eignet sich für kleine Stichproben und ist daher für die in der Wertermittlung oftmals kleinen Datensätze gut geeignet. Die Teststatistik wird hierbei durch eine Verhältniszahl der bei Normalverteilung erwarteten Varianz und der tatsächlichen Varianz gebildet. Für größere Stichproben (Anzahl

der Kauffälle  $n > 50$ ) weist der Kolmogorov-Smirnov-Test eine bessere Testgüte auf (vgl. Yazici und Yolacan 2007). In beiden Tests werden die Werte der Reihe nach geordnet (Ordnungsstatistik) und die Abweichungen zum erwarteten Wert untersucht.

Die Prüfgröße für die Normalverteilung nach Kolmogorov-Smirnov liegt im Berechnungsbeispiel bei 0,05, also unter dem kritischen Wert (bei 5 % Irrtumswahrscheinlichkeit) mit  $1,358/\sqrt{n}$  (vgl. Hedderich und Sachs 2012, S. 422): Die Nullhypothese (Residuen sind normalverteilt) wird angenommen. Neben der Angabe der Prüfgröße werden dem Nutzer die studentisierten Residuen in Form von Histogrammen, QQ-Plots und Boxplots dargestellt.

##### 4.7.2 Homoskedastizität

In der linearen Regression wird unterstellt, dass die Residuen gleichmäßig streuen (Homoskedastizität). Die Residuen dürfen daher nicht von unabhängigen Variablen oder der Reihenfolge der Kauffälle abhängig sein (Heteroskedastizität). Liegt eine ungleichmäßige Verteilung der Varianzen vor, werden zwar die Schätzwerte der Parameter (Regressionskoeffizienten) unverzerrt ermittelt, die Varianzen sind jedoch verzerrt. Dies hat zur Folge, dass das Schätzverfahren nicht mehr effektiv erfolgt (insbesondere Tests versagen, da die Formeln für die Standardfehler keine Gültigkeit mehr haben). Zur Prüfung auf Homoskedastizität werden die Residuen hinsichtlich einer Homogenität in den Varianzen untersucht. Hierzu kann bspw. der Breusch-Pagan-Test zum Einsatz kommen (es wird die Beziehung zwischen standardisierten Residuen und den unabhängigen Variablen getestet). Hierbei wird untersucht, ob die Fehlervarianz  $\sigma_i^2$  konstant ist. Dazu wird eine Hilfsregression mit den Residuen durchgeführt und es wird überprüft ob alle Residuen konstant sind – also alle Parameter dieser Regression gleich Null. Die Teststatistik errechnet sich dann aus dem Vergleich der Residuen der Hilfsregression mit den Residuen der Ausgangsregression. Die Prüfgröße ist hierbei  $\chi^2$  verteilt.

Für das Berechnungsbeispiel wird die Hypothese der Multikollinearität angenommen: Die Prüfgröße des Breusch-Pagan-Tests liegt unter dem Grenzwert.

##### 4.7.3 Multikollinearität

Wird eine Größe bereits durch eine andere im Modell enthaltene Größe beschrieben (ggf. auch nur zu Teilen), liegt Multikollinearität vor. Diese kann durch die Untersuchung der Varianzen aufgedeckt werden: Hierzu ist der Varianzinflationsfaktor (VIF) als Indikator zu empfehlen. Je größer der VIF ist, desto wahrscheinlicher liegen Multikollinearitäten vor. Der Variationsinflationsfaktor  $VIF_j$  wird aus der reziproken Toleranz  $Tol_j = 1 - R_j^2$  (unabhängige Variabilität) berechnet (mit  $R_j$ : Bestimmtheitsmaß der Regressionsanalyse der verbleibenden Einflussgrößen auf die Einflussgröße  $j$ ).

Urban und Mayerl (2011, S. 232) nennen einen üblichen Grenzwert für den VIF von >10, um Multikollinearität aufzudecken. Sie empfehlen jedoch einen niedrigeren Grenzwert im Bereich von 5. Bei ersten Analysen mit dem Datensatz in Nienburg (Weser) hat sich der Grenzwert von 10 als zu hoch erwiesen. Bei den Analysen und im Vergleich zu herkömmlichen Vorgehensweisen wurde regelmäßig ein Grenzwert von 8 als geeignet festgestellt: Hier müssen jedoch weitere Untersuchungen – auch in anderen Teilmärkten – erfolgen, um diesen Wert zu bestätigen.

Im vorliegenden Berechnungsbeispiel liegen alle Größen unter dem Grenzwert von 8. Die maximalen Werte liegen im Bereich von 3,8 für die Standardstufe und 3,9 für die Restnutzungsdauer. Die Variablen enthalten insofern einen eigenständigen Erklärungsgehalt. Wären die anderen (korrelierten) Einflussgrößen nicht schon bei der Korrelationsuntersuchung ausgeschlossen worden, würde spätestens hier der Hinweis erfolgen, dass diese Größen bereits durch die anderen Größen beschrieben werden.

Neben den genannten Größen wird als weitere Größe auch die Autokorrelation angegeben. Diese gibt eine Information, ob das Signal (zur Interpretation von Wertermittlungsdaten als Signal sei z.B. auf Zaddach (2016) verwiesen) mit sich selbst korreliert ist, d.h. ob bspw. zyklische Phänomene auftreten. Die Autokorrelation ist eher informativ und keine direkte Voraussetzung zur Annahme des Regressionsmodells.

#### 4.8 Prädiktion

Nach dem Abschluss der Regressionsanalyse hat der Nutzer nun die Möglichkeit, konkrete Bewertungen durchzuführen. Die Funktionsgleichung der Regression ergibt sich nach der Ausreißerelimination und Bestätigung aller Modellannahmen für das gezeigte Beispiel mit:

$$\text{KP/WF} = -4835 - 2,698 \cdot \text{WF} + 6,160 \cdot \text{BRW} + 123,2 \cdot \text{STST} + 12,5 \cdot \text{RND} + 44,36 \cdot \text{Datum [dezimal]}$$

Diese Funktionsgleichung steht in dem neu entwickelten Modul der AKS Niedersachsen für konkrete Bewertungsfälle zur Verfügung: Es können die Einflussgrößen des Bewertungsobjektes eingegeben und das Ergebnis angezeigt werden. Als weitere Information wird nochmals die deskriptive Statistik (Minimum und Maximum der einzelnen Einflussgrößen) angezeigt, um einzuordnen, ob die Datengrundlage für die Bewertung (Bewertung im Rahmen der Stichprobe) geeignet ist.

#### 4.9 Bericht

Abschließend hat der Nutzer die Möglichkeit, sich einen Bericht mit allen wesentlichen Größen der Auswertung

ausgeben zu lassen. Dieser kann als Dokumentation dienen. Neben deskriptiven Größen können hier auch Visualisierungen ausgegeben werden.

### 5 Fazit und zukünftige Entwicklungen

Mit dem neu implementierten Modul werden dem Nutzer der AKS mehr Möglichkeiten der Interaktion mit den Daten geliefert. Ihm werden nun Entscheidungsgrößen nach dem Stand der Technik zur Verfügung gestellt. Neben der Erhöhung des Visualisierungsgrades ist ein übersichtlicher, prozeduraler Ablauf der Untersuchung von Daten möglich. Um dem in vielen Teilmärkten vorherrschenden Marktdruck gerecht zu werden, wird auch ein höherer Automatisierungsgrad bei der Auswertung benötigt. In der Abschlussphase des Projektes wird neben der Version mit hoher Nutzerinteraktion auch eine Version mit hohem Automatisierungsgrad entwickelt. Mit dem neu erstellten Modul wurde gezeigt, dass die AKS Niedersachsen modular erweiterbar und so für zukünftige Entwicklungen und Ergänzungen geeignet ist.

Um eine Ableitung von Marktdaten auch in kaufpreisarmen Lagen zu ermöglichen, ist es notwendig, auch weitere Quellen für Marktinformationen zu nutzen. Hier wären statistische Informationen, z.B. der Landesämter für Statistik, aber auch eigene erhobene Daten (z.B. Passantenfrequenzen) denkbar. In der AKS soll zukünftig die Möglichkeit eröffnet werden, diese Informationen einzupflegen und für Auswertungen zu nutzen. Des Weiteren soll die Einbindung von Expertenwissen in Form von Gutachten und Befragungen ermöglicht werden. Insbesondere die umfangreiche qualitative Beschreibung von Immobilien auf Vermarktungsplattformen wie Immobilienscout24 und ImmoWelt liefern viele Informationen, die i.d.R. umfangreicher sind als die Informationen der Eigentümer aus Fragebögen. Bei einer geeigneten Würdigung der Angebotspreise könnten diese gemeinsam mit Kaufpreisen ausgewertet werden. Erste Untersuchungen wurden bereits in Soot et al. (2016) durchgeführt.

Mit der Kaufpreissammlung der Gutachterausschüsse hat Deutschland im internationalen Vergleich ein einzigartiges Instrument zur Marktbeobachtung. Um weiterhin eine unabhängige Institution zur Marktbeobachtung zu haben, ist es dringend notwendig, die in den Datenbanken abgelegten Daten mit einer ausreichenden Anzahl an Metainformationen zu ergänzen. Damit ist es möglich, im Nachgang entsprechende Auswertungen und Marktberichte zu erstellen.

### Literatur

- Alkhatab, H., Weitkamp, A. (2012): Bayesischer Ansatz zur Integration von Expertenwissen in die Immobilienbewertung, Teil 1. In: zfv – Zeitschrift für Geodäsie, Geoinformation und Landmanagement, Jg. 137 (2), S. 93–102.

- Akhatab, H., Weitkamp, A., Zaddach, S., Neumann, I. (2016): Unsicherheitsschätzung des Verkehrswertes durch ein Fuzzy-Bayes Vergleichswertverfahren. In: *zfv – Zeitschrift für Geodäsie, Geoinformation und Landmanagement*, Jg. 141 (3), S. 169–178.
- Baarda, W. (1968): A testing procedure for use in geodetic networks. PhD thesis, Delft University of Technology, Publications on Geodesy, Jg. 2 (5), Netherlands Geodetic Commission, Delft, Niederlande.
- Brückner, R. (1976): Mathematische Statistik bei der Ermittlung von Grundstückswerten. Kontaktstudium. Wissenschaftliche Arbeiten der Lehrstühle für Geodäsie, Photogrammetrie und Kartographie an der Technischen Universität Hannover, Band 65.
- Cameron, S. (1993): Why is the R squared adjusted reported? In: *Journal of Quantitative Economics*. Jg. 9, S. 183–186.
- Chatterjee, S., Price, B. (1995): Praxis der Regressionsanalyse. 2. Aufl., Lehr- und Handbücher der Statistik, München: Oldenbourg.
- Daniel, C., Wood, F.S. (1980): Fitting equations to data: computer analysis of multifactor data. 2. Aufl., Wiley series in probability and mathematical statistics, New York: Wiley.
- Dorndorf, A., Soot, M., Weitkamp, A., Alkhatab, H. (2016) Development of a Robust Bayesian Approach for Real Estate Valuation in Areas with Few Transactions Development of a Robust Bayesian Approach for Real Estate Valuation in Areas with Few Transactions. In: FIG Working Week 2016, Christchurch (New Zealand), 2–6 May 2016.
- Dorndorf, A., Soot, M., Weitkamp, A., Alkhatab, H. (2017): A heuristic robust approach for real estate valuation in areas with few transactions. In: FIG Working Week 2017, Helsinki (Finland), 29 May–2 June 2017.
- Fahrmeir, L., Kneib, T., Lang, S. (2007): Regression: Modelle, Methoden und Anwendungen. Berlin: Springer.
- Fox, J., Long, J.S. (1990): Modern methods of data analysis. Newbury Park, Calif.: Sage Publications.
- Goldberg, A. (1991): A Course in Econometrics. Cambridge: Harvard University Press.
- Hedderich, J., Sachs, L. (2012): Angewandte Statistik. Heidelberg: Springer Gabler.
- Kenneth, P.B., David, R.A. (2004): Multimodel Inference: Understanding AIC and BIC in Model Selection. In: *Sociological Methods & Research*, Jg. 33 (2), S. 261–304.
- Kleiber, W., Fischer, R., Werling, U. (2017): Verkehrswertermittlung von Grundstücken. Köln: Bundesanzeiger Verlag.
- Koch, K.-R. (1987): Statistische Methoden zur Analyse von Grundstückspreisen. In: *ZfV – Zeitschrift für Vermessungswesen*, Jg. 112 (12), S. 617–621.
- LGLN (2015): Handbuch für die Führung und Auswertung der Kaufpreissammlung (Handbuch-AKS), Stand: 04.03.2015. Landesamt für Geoinformation und Landesvermessung Niedersachsen, Hannover.
- Liebig, E., Ache, P. (2014): Immobilienmärkte in Deutschland – Strukturen und Trends. In: Kummer, K., Frankenberger, J., Kötter, T. (Hrsg.): Das deutsche Vermessungs- und Geoinformationswesen. Berlin: Wichmann, S. 553–580.
- Montgomery, D.C., Peck, E.A., Vining, G.G. (2006): Introduction to linear regression analysis. 4. ed., Wiley series in probability and statistics, Hoboken, NJ: Wiley-Interscience.
- Pelzer, H. (1978): Ein indirektes Vergleichswertverfahren unter Anwendung statistischer Methoden. In: *ZfV – Zeitschrift für Vermessungswesen*, Jg. 103 (6), S. 245–254.
- Pope, A.J. (1976): The Statistics of Residuals and the Detection of Outliers. NOAA Technical Report NOS65 NGS1, US Department of Commerce, National Geodetic Survey, Rockville, Maryland.
- Quinino, R. C., Reis, E. A. (2012): Using the coefficient of determination R<sup>2</sup> to test the significance of multiple linear regression. In: *Teaching Statistics*, Jg. 35 (2), S. 84–88.
- Soot, M., Weitkamp, A., Alkhatab, H., Dorndorf, A., Jeschke, A. (2016): Analysis on Different Market Data for Real Estate Valuation – Investigations on German Real Estate Market. In: FIG Working Week 2016, Christchurch (New Zealand), 2–6 May 2016.
- Soot, M., Weitkamp, A., Dorndorf, A., Alkhatab, H., Jeschke, A. (2017): Different Regions with Few Transactions – An Approach of Systematization. In: FIG Working Week 2017, Helsinki (Finland), 29 May–2 June 2017.
- Urban, D., Mayerl, J. (2011): Regressionsanalyse: Theorie, Technik und Anwendung. Wiesbaden: VS Verlag.
- Weitkamp, A., Alkhatab, H. (2012): Bayesischer Ansatz zur Integration von Expertenwissen in die Immobilienbewertung, Teil 2. In: *zfv – Zeitschrift für Geodäsie, Geoinformation und Landmanagement*, Jg. 137 (2), S. 103–114.
- Weitkamp, A., Alkhatab, H. (2014): Die Bewertung kaufpreisarmer Lagen mit multivariaten statistischen Verfahren – Möglichkeiten und Grenzen robuster Methoden bei der Auswertung weniger Kauffälle. In: *AVN – Allgemeine Vermessungs-Nachrichten*, Jg. 121 (1), S. 2–11.
- Yazici, B., Yolacan, S. (2007): A comparison of various tests of normality. *Journal of Statistical Computation and Simulation*, Jg. 77 (2), S. 175–183.
- Zaddach, S. (2016): Zum Beitrag Bayesscher Schätzverfahren in der Verkehrswertermittlung. Dissertation, Wissenschaftliche Arbeiten der Fachrichtung Geodäsie und Geoinformatik der Leibniz Universität Hannover, Band 321.
- Zaddach, S., Alkhatab, H. (2013): Quantifying the impact of uncertainty to the market value by introducing a bayesian sales comparison approach. In: Proceedings of the RICS Cobra Research Conference, New Delhi (India), 10.–12.09.2013.
- Zaddach, S., Alkhatab, H. (2014a): Least Squares Collocation as an Enhancement to Multiple Regression Analysis in Mass Appraisal Applications. In: *Journal of Property Tax Assessment & Administration*, Jg. 11 (1), S. 47–66.
- Zaddach, S., Alkhatab, H. (2014b): Propagating the uncertainty of the market value by the use of a Bayesian regression approach. In: Proceedings of the XXV FIG Congress, Kuala Lumpur (Malaysia), 16–21 June 2014.
- Zaddach, S., Soot, M., Käker, R., Weitkamp, A., Ziems, M. (2016): Entwicklung neuer Auswertemethoden in der AKS. In: *Nachrichten der Niedersächsischen Vermessungs- und Katasterverwaltung*, Jg. 66 (1), S. 3–11.
- Ziegenbein, W. (1977): Zur Anwendung multivariater Verfahren der mathematischen Statistik in der Grundstückwertermittlung. Dissertation, Wissenschaftliche Arbeiten der Lehrstühle für Geodäsie, Photogrammetrie und Kartographie an der Technischen Universität Hannover, Band 77.
- Ziegenbein, W. (1986): Zur Analyse der automatisiert geführten Kaufpreissammlungen für Grundstücke in Niedersachsen. In: *Nachrichten der Niedersächsischen Vermessungs- und Katasterverwaltung*, Jg. 36 (3), S. 195–218.
- Ziegenbein, W. (1995): Programmgesteuerte Regressionsanalyse und Vergleichswertermittlung im Programmsystem AKS. In: *Nachrichten der Niedersächsischen Vermessungs- und Katasterverwaltung*, Jg. 45 (4), S. 321–335.
- Ziegenbein, W. (1999): Wertermittlungsinformationssystem Niedersachsen. In: *Nachrichten der Niedersächsischen Vermessungs- und Katasterverwaltung*, Jg. 49 (3), S. 121–126.
- Ziegenbein, W., Hawerk, W. (1978): Erfahrungen bei der Prädikation. In: *ZfV – Zeitschrift für Vermessungswesen*, Jg. 103 (6), S. 254–261.

## Kontakt

Matthias Soot (M.Sc.) | Prof. Dr.-Ing. Alexandra Weitkamp  
TU Dresden, Geodätisches Institut, Professur für Landmanagement  
Helmholzstraße 10, 01069 Dresden

Dr.-Ing. Sebastian Zaddach | Dipl.-Ing. René Käker  
Niedersächsisches Ministerium für Inneres und Sport  
Referat 47 – Vermessung und Geoinformation, Kampfmittelbeseitigung  
Lavesallee 6, 30169 Hannover

Dr.-Ing. Marcel Ziems  
Landesamt für Geoinformation und Landesvermessung Niedersachsen (LGLN), Fachgebiet 223 – Wertermittlungsinformation, Geschäftsnachweise, KLR  
Podbielskistraße 331, 30659 Hannover

Dieser Beitrag ist auch digital verfügbar unter [www.geodaezie.info](http://www.geodaezie.info).