

Transparenz und Güte der Ergebnisse von Wertermittlungen – Teil 1: Grundüberlegungen für eine moderne Wertermittlung

Transparency and Quality of Valuation Results – Part 1: Basic Thoughts on Modern Valuation

Peter Ache

Zusammenfassung

Die Immobilienwertermittlung befindet sich an einem Wendepunkt: Die zunehmende Verfügbarkeit digitaler Daten und der fortschreitende Einsatz Künstlicher Intelligenz eröffnen vielversprechende Möglichkeiten zur Entwicklung präziserer und transparenterer Wertermittlungsmodelle. Der vorliegende Beitrag diskutiert die Bedeutung evidenzbasierter Ansätze und rückt die Modellgüte als zentralen Qualitätsindikator in den Fokus der Diskussion. Dabei werden die Limitationen traditioneller Methoden, die häufig auf subjektiver Expertise basieren, angesprochen und datenbasierte Verfahren als zukunftsweisende Alternativen hervorgehoben.

Die aktuellen Veränderungen des Wertermittlungsrechts in Deutschland kann derzeit als eine eher langsame und vorsichtige, also konservative Anpassung der regulatorischen Rahmenbedingungen im Bereich der Immobilienwertermittlung wahrgenommen werden. Vor dem Hintergrund gesellschaftlicher Herausforderungen, wie dem Klimawandel, dem anhaltenden Wohnungsmangel und den veränderten geopolitischen Bedingungen, ist jedoch davon auszugehen, dass die Anforderungen an die Transparenz von Immobilienwerten und Bewertungsverfahren in naher Zukunft erheblich steigen werden. Diese Entwicklungen könnten zu tiefgreifenden und disruptiven Anpassungen der bestehenden Regelwerke führen, die die Branche nachhaltig transformieren werden.

Ein besonderer Schwerpunkt dieses Beitrages liegt auf der Notwendigkeit klar definierter und transparenter Prozesse, um den sich dynamisch verändernden Anforderungen an Immobilienmarktanalysen und -wertermittlungen gerecht zu werden. Dabei geht es insbesondere um die Definition des Begriffs der »Modellgüte« sowie um die Vorstellung von Ansätzen zur Beschreibung und Darstellung von Metriken, die die Leistungsfähigkeit und Qualität von Modellen und deren Ergebnisse beschreiben können.

Schlüsselwörter: Immobilienwertermittlung, Modellgüte, Künstliche Intelligenz, evidenzbasierte Ansätze, Transparenz, Immobilienmarkt

Summary

Real estate valuation is at a turning point: The increasing availability of digital data and the growing use of artificial intelligence offer promising opportunities for the development of more pre-

cise and transparent valuation models. This paper analyses the significance of evidence-based approaches and places model quality as a central indicator of quality at the forefront of the discussion. It highlights the limitations of traditional methods, which often rely on subjective expertise, and underscores data-driven procedures as forward-looking alternatives.

The current changes to valuation law in Germany are perceived as a rather slow and cautious, thus conservative, adjustment of the regulatory framework in the field of real estate valuation. Against the backdrop of societal challenges such as climate change, the persistent housing shortage, and shifting geopolitical conditions, it can be expected that the demands for transparency in property values and valuation procedures will increase significantly in the near future. These developments could lead to profound and disruptive changes to the existing regulatory frameworks, transforming the industry in a lasting way.

A particular focus of this paper lies on the necessity of clearly defined and transparent valuation processes to meet the dynamically changing requirements of market analyses and real estate valuations. This includes defining the concept of »Model reliability« and introducing approaches to describe and present metrics that can measure the performance and quality of models and their outcomes.

Keywords: *real estate valuation, model quality, artificial intelligence, evidence-based approaches, transparency, real estate market*

1 Einleitung

Der Begriff »Künstliche Intelligenz« (KI) steht derzeit im Zentrum zahlreicher Fachdiskussionen, Veranstaltungen und Initiativen. Insbesondere »Large Language Models (LLM)« wie »ChatGPT«, »Gemini« oder mittlerweile eine Vielzahl anderer Programme haben durch ihre breite und unkomplizierte Anwendung einen deutlichen Anstieg im Einsatz Künstlicher Intelligenz bewirkt.

Eine wesentliche Grundlage und zwingende Voraussetzung für den Erfolg solcher KI-Anwendungen sind jedoch verfügbare und qualitativ hochwertige Daten, mit denen die Systeme trainiert werden können. Diese können

dann auf Basis bestehender Informationen Vorhersagen treffen bzw. Vorschläge machen; sei es ein Code im Rahmen von Programmierarbeiten (z.B. GitHub Copilot), die Internetsuche auf der Grundlage von Fragen (perplexity.ai) oder seien es Anwendungen im medizinischen Bereich als unterstützende Systeme bei Operationen, in der Forschung oder bei der Identifizierung von Betrug und Straftaten.

Grundlegend hierfür ist die realitätsgetreue Modellierung, um auf Basis gut dokumentierter, repräsentativer und unverzerrter Daten zukünftige Phänomene oder auch nur einen nächsten Satz in einem Schriftstück vorhersagen zu können. Nur wenn ausreichend verlässliche Daten vorliegen, können Modelle eine hohe Leistungsfähigkeit und eine für den jeweiligen Zweck ausreichende Prognosicherheit gewährleisten.

Wird dieser Aspekt im Modellbildungsprozess nicht ausreichend berücksichtigt, können Verzerrungen, Ungenauigkeiten oder eine unzureichende Modellleistung schwerwiegende Folgen haben. Im Gesundheitssystem kann dies lebensbedrohliche Konsequenzen für Patienten haben, im Finanz- und Immobiliensektor zu erheblichen Verlusten oder gar Krisen führen und in organisatorischen Bereichen wie der Luftfahrt gravierende Störungen bis hin zu katastrophalen Unfällen verursachen.

Aus diesem Grund sind die Validierung und Überprüfung von Modellen und Angaben zur Genauigkeit der Ergebnisse von großer Bedeutung und können nicht deutlich genug betont werden. Auch wenn diese Anforderungen in der Immobilienwertermittlung traditionell derzeit noch als weniger kritisch gesehen werden, ist es immer entscheidend, die Leistungsfähigkeit und Aussagekraft der Bewertungsmodelle im Blick zu behalten. Dabei geht es aktuell um die Daten, die für die Wertermittlung erforderlich sind und in Deutschland aktuell vor allem durch die Gutachterausschüsse für Grundstückswerte ermittelt und durch Sachverständige gemäß Baugesetzbuch (BauGB) und Immobilienwertermittlungsverordnung (ImmoWertV) genutzt werden. Diese Daten sind Aggregationen und basieren vorwiegend auf Kaufpreisen, teilweise aber auch auf der Expertise der Gutachterausschüsse.

Mit großer Wahrscheinlichkeit ist zukünftig davon auszugehen, dass Modelle entwickelt werden müssen, die direkt auf Kaufpreisen basieren und so unmittelbar zu einer präziseren Bewertung der jeweiligen Immobilie führen, als dies derzeit praktiziert wird. Allerdings kommt es auch hier darauf an, das Bewertungsergebnis so transparent zu gestalten, dass Nutzerinnen und Nutzer den Grad der Präzision einschätzen können.

Der Wandel in den Anforderungen an eine Immobilienwertermittlung hoher Güte in Deutschland ist damit klar umrissen. Dies führt aber zu einer kritischen Betrachtung der bisher zurückhaltend behandelten Genauigkeits- oder besser gesagt Gütekriterien in der traditionellen Immobilienwertermittlung. Aufgrund sogenannter begrenzter Datenmengen und – früher – nicht ausreichender Rechenkapazitäten liegt der Schwerpunkt traditionell auf der Ex-

pertise der Sachverständigen, während statistische Analysen und Genauigkeitsbetrachtungen eine untergeordnete Rolle spielen. Dies betrifft sowohl die Ermittlung der für die Wertermittlung erforderlichen Daten als auch die Ermittlung von Verkehrswerten, wie sie in § 194 BauGB definiert sind.

Gute Ansätze dazu sind in der Immobilienwertermittlungsverordnung 2021 (vom 14. Juli 2021, ImmoWertV 2021, BGBl. I S. 2805) und deren Muster-Anwendungshinweisen (ImmoWertA) vom 20. September 2023 erkennbar. Sie gehen jedoch vor dem Hintergrund der schnellen Entwicklungen nicht weit genug und es besteht die Gefahr, dass Anwendungen mit Künstlicher Intelligenz ungewollte Fakten schaffen, die im Nachhinein schwer rückgängig gemacht werden können.

In diesem Beitrag soll besonderes Augenmerk auf die Grundlagen transparenter Prozesse der Modellbildung und der letztendlich erreichten Ergebnisgüte gelegt werden. Dies kann zunächst nur ein Anfang sein und soll im Umfeld der traditionellen Immobilienwertermittlung zur Diskussion und Weiterentwicklung anregen. Der Fokus liegt dabei auf dem Prozess der Modellbildung bei der Wertermittlung von Immobilien und der Entwicklung von Kriterien zur Beurteilung der Güte der Modelle.

2 Künstliche Intelligenz und evidenzbasierte Wertermittlung

Die menschliche Wahrnehmung erfasst die Realität durch die bewusste oder unbewusste Beobachtung von Phänomenen und deren schnelle (mithilfe von Heuristiken) oder langsame (durchdachten) Interpretation (Graf 2015); dadurch entsteht eine subjektive Umsetzung der Wirklichkeit. Im Gegensatz dazu dienen Modelle realer Phänomene dazu, die Struktur der Wirklichkeit möglichst objektiv abzubilden oder zumindest wesentliche Grundzüge davon zu repräsentieren (Coombs et al. 1975). Die Grundlage der Modellbildung zur Analyse des Immobilienmarktes stellen beobachtbare Daten dar, etwa zu den Eigenschaften von Immobilien (nach ImmoWertV: Grundstücksmerkmale), zu den Umständen von Erwerbsvorgängen, zu erzielten Preisen, Mieten oder Pachten sowie zu Akteuren, die die Immobilien veräußern oder erwerben. Dabei muss berücksichtigt werden, dass die den Preis beeinflussenden Merkmale nicht immer die gleichen Effekte auf den Preis haben. Gleiche Eigenschaften einer Immobilie können je nach Kontext der Preisverhandlung unterschiedlich gewichtet werden, was auch für zeitlich variierende Marktphänomene oder geografische Unterschiede gilt. Dies unterscheidet die Preisbeobachtung auf dem Immobilienmarkt fundamental von der Messung physikalischer Größen, wie etwa Strecken, bei denen der Kontext – ob in München oder Hamburg gemessen – irrelevant ist.

Im Zusammenhang mit der zu erwartenden stärkeren Nutzung von Künstlicher Intelligenz rückt die Verwendung

von Machine-Learning-Modellen noch stärker in den Fokus, als es bislang der Fall war. So definiert das Europäische Parlament und der Rat in der Verordnung (EU) über Künstliche Intelligenz (KI-Verordnung, sog. AI Act) vom 13. Juni 2024 (ABl. DE L 2024/1689) den Begriff »KI-System« als

»[...] ein maschinengestütztes System, das für einen in unterschiedlichem Grade autonomen Betrieb ausgelegt ist und das nach seiner Betriebsaufnahme anpassungsfähig sein kann und das aus den erhaltenen Eingaben für explizite oder implizite Ziele ableitet, wie Ausgaben wie etwa Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen erstellt werden, die physische oder virtuelle Umgebungen beeinflussen können; [...]« (Art. 3 Nr. 1 AI Act)

Die zur Vorbereitung der KI-Verordnung der EU eingesetzte »High Level Expert Group« fasst den Begriff noch konkreter und definiert unter Nr. 4 (Updated definition of AI):

»Künstliche Intelligenz (KI)« bezieht sich auf Systeme, die von Menschen entwickelt wurden und die angesichts eines komplexen Ziels in der physischen oder digitalen Welt agieren, indem sie ihre Umgebung durch Datenerfassung wahrnehmen, die gesammelten strukturierten oder unstrukturierten Daten interpretieren, um auf der Grundlage des aus diesen Daten gewonnenen Wissens Schlussfolgerungen zu ziehen und die besten Maßnahmen (gemäß vordefinierten Parametern) vorzuschlagen [...].« (Europäische Kommission, High Level Expert Group on Artificial Intelligence 2018)

Diese Definition impliziert, dass KI-Systeme aus den ihnen zur Verfügung gestellten Daten Modelle der Wirklichkeit erstellen, aus neuen Daten lernen und diese Modelle nutzen, um Vorhersagen zu treffen und Entscheidungen zu unterstützen. Für die Immobilienwertermittlung sind daher geeignete Kaufpreisdaten erforderlich, um Modelle zu trainieren, die einen zu erwartenden Kaufpreis zuverlässig und mit ausreichender Genauigkeit vorhersagen können.

3 Evidenzbasierte Wertermittlung

Dass in der heutigen Zeit enorme Mengen an Daten zu nahezu allen Aspekten unseres Lebens generiert werden, ist allgemein bekannt. Ebenso ist bekannt, dass Unternehmen wie z.B. Google, Amazon und Facebook diese Daten zu kommerziellen Zwecken analysieren und nutzen. Auch im Umfeld der Immobilienwirtschaft fallen eine Vielzahl von Daten an, die von Planungsdaten über Eigentümer- und Nutzerinformationen bis hin zu detaillierten Daten über

Bau- und Konstruktionsteile, Preise, Mieten und Verkaufsinformationen gehen.

Die traditionelle Immobilienwertermittlung hingegen basiert noch zunächst auf theoretischen, von ökonomischer Denkweise getriebenen Ansätzen; konkrete Immobilienmarktdaten zu Kaufpreisen spielen aus verschiedensten Gründen in der Regel lediglich in ihren aggregierten Formen eine Rolle und werden in gesetzlichen Normen für die Wertermittlung in Deutschland als »für die Wertermittlung erforderliche Daten« bezeichnet.

Dieser Ansatz ist im Verhältnis zu den heutigen Möglichkeiten und Anforderungen sehr konservativ und weist möglicherweise Unsicherheiten oder gar Verzerrungen auf, die nicht immer durch die sogenannte Expertise der Fachleute ausgeglichen werden können. Auch wird oft übersehen, dass Experten – wie etwa auch Richter, Lehrer, Polizisten und Führungskräfte – ebenfalls für psychologische Verzerrungen anfällig sind, insbesondere dann, wenn es um Entscheidungen mit hohen Unsicherheitsfaktoren geht (Kahneman et al. 1982). Die verfügbaren Daten werden derzeit unzureichend genutzt, um Verkehrswertermittlungen und die damit zusammenhängenden Entscheidungen valider zu gestalten.

Unter Berücksichtigung der Tatsache also, dass Märkte im Allgemeinen und Immobilienmärkte im Besonderen nicht nur von ökonomischen, sondern auch von psychologischen Faktoren beeinflusst werden, wird offensichtlich, dass herkömmliche ökonomische Methoden und die darauf fußenden Modelle den Immobilienmarkt möglicherweise nur unzureichend abbilden. Die Verzerrungen werden nach Erkenntnissen der Verhaltensökonomik umso größer, je höher die Risiken sind (Kahneman und Tversky 1979). Dies ist bei Investitionen in Immobilien aufgrund der vergleichsweise hohen Geldbeträge und der Langfristigkeit der Investitionen regelmäßig der Fall.

Es ist daher erforderlich, von theoretischen Bewertungsmodellen zu »evidenzbasierten Bewertungen (Evidence Based Valuation, EBV)« überzugehen. Dieser Ansatz der Bewertung basiert auf empirischen Daten, die direkt aus dem realen Markt stammen, darunter tatsächlich gezahlte Kaufpreise, Informationen über Verkäufer und Käufer sowie die spezifischen Umstände einzelner Transaktionen. Solche Bewertungen stützen sich nur marginal oder gar nicht auf theoretische ökonomische Annahmen. Allerdings ist es auch in diesem Kontext von entscheidender Bedeutung, die Güte der Vorhersagen transparent darzustellen, um deren Eignung für den vorgesehenen Zweck valide beurteilen zu können. Erst die Kenntnis über die Güte der Prognose ermöglicht es, die Zuverlässigkeit der Bewertung zu quantifizieren und sicherzustellen, dass die Ergebnisse den Anforderungen der jeweiligen Anwendung genügen.

4 Grundgesamtheit, Stichprobe und Repräsentativität

Im Kontext der Immobilienwertermittlung ist derzeit noch häufig unklar, welche Grundgesamtheit bei der Analyse von Kaufpreisen zugrunde gelegt wird. Für eine aussagekräftige Interpretation statistischer Analysen, die der Wertermittlung eines Grundstücks dienen, ist es jedoch unerlässlich, die Grundgesamtheit vor Beginn der Analyse präzise zu definieren. Bei der Untersuchung von Daten, wie beispielsweise Kaufpreisen für Eigentumswohnungen in einer Stadt, muss festgelegt werden, welcher Grundgesamtheit das betreffende Wertermittlungsobjekt zugeordnet werden soll. Dies erfordert eine klare und eindeutige Beschreibung der relevanten Merkmale, um die notwendige Transparenz sicherzustellen.

Die Definition der Grundgesamtheit im Rahmen der Wertermittlung lässt sich im Wesentlichen auf zwei zentrale Aspekte zurückführen:

- die Definition des räumlichen Teilmarktes (z. B. Land, Stadt, Stadtteil etc.)
- die Definition des sachlichen Teilmarktes (z. B. Einfamilienhaus, Büro, Hotel etc.)

Die Grundgesamtheit umfasst alle dem Markt zur Verfügung stehenden Objekte. Im Fall von z. B. Eigentumswohnungen in einer Stadt bedeutet dies, dass sämtliche vorhandenen Eigentumswohnungen einbezogen werden müssen, nicht nur die bereits veräußerten. Ein häufiges Missverständnis besteht darin, die Grundgesamtheit mit den in Kaufpreissammlungen der Gutachterausschüssen dokumentierten Kauffällen gleichzusetzen, die auf abgeschlossenen Grundstückskaufverträgen basieren. Diese Auffassung ist jedoch unzutreffend, da das zu bewertende oder zu veräußernde Objekt in der Regel noch keinen Preis aufweist – gerade dieser Preis, wegen seiner Allgemeingültigkeit für dieses Objekt als Wert bezeichnet – soll im Rahmen der Wertermittlung ermittelt werden.

Die Wertermittlung erfolgt im Idealfall – unabhängig vom gewählten Verfahren – auf Basis verfügbarer und geeigneter Kaufpreisdaten aus bestehenden Datenbanken. Diese Kaufpreise repräsentieren jedoch nur einen Teil der tatsächlichen Grundgesamtheit, sie müssen als Stichprobe aufgefasst werden.

Das Ziel einer statistischen Datenanalyse besteht in der Regel darin, Informationen zu einem interessierenden Parameter in einer unbekannten Grundgesamtheit zu gewinnen. Ein Parameter kann im Rahmen von Immobilienwertermittlungen beispielsweise ein Vergleichswert, ein Liegenschaftszinssatz, ein Sachwertfaktor oder ein anderer für die Wertermittlung relevanter Wert sein. Der grundlegende Gedanke dabei ist, dass die konkrete Wertermittlung auf ein Objekt abzielt, das nicht Teil der Stichprobe, jedoch der Grundgesamtheit ist.

Entscheidend ist daher, dass die Stichprobe die unbekannten Informationen der Grundgesamtheit mit der für die betreffende Wertermittlungsaufgabe erforder-

lichen Genauigkeit abbildet. Die Stichprobe muss repräsentativ sein, insbesondere in Bezug auf die zeitliche Nähe zum Wertermittlungstichtag sowie hinsichtlich der räumlichen und sachlichen Merkmale des relevanten Teilmarktes. Dies wird in § 9 ImmoWertV sehr klar definiert.

Dennoch wird der Begriff »Repräsentativität« in vielen Fällen zwar im Sinne der Anforderungen an gute statistische Analysen verwendet, der tatsächliche Inhalt bleibt jedoch regelmäßig im Ungefährten. Bezogen auf eine Grundgesamtheit von Immobilien dürfte damit gemeint sein, dass jedes Objekt dieser Grundgesamtheit dieselbe Wahrscheinlichkeit hat, veräußert oder bewertet zu werden. Dies wiederum bedeutet,

- dass Kaufpreise einer Stichprobe verzerrungsfrei sein müssen,
- die Kaufzeitpunkte gleichmäßig um den Wertermittlungstichtag (oder darauf anpassbar) streuen,
- der Kauffall um den geografischen Ort und
- um den sachlichen Teilmarkt der betreffenden Grundgesamtheit regelmäßig verteilt sein muss.

Oft ist allein schon die Beurteilung dieser Kriterien aufgrund viel zu weniger vorliegender Kaufpreise schlicht nicht möglich.

Allgemein kann die folgende Definition zur Beurteilung von Repräsentativität einer Stichprobe verwendet werden (Gabler und Quatember 2012):

»Eine Stichprobe heißt für eine Gesamtheit hinsichtlich einer interessierenden Verteilung oder eines interessierenden Parameters dieser Verteilung repräsentativ, wenn diese Verteilung oder der Parameter damit (zumindest annähernd) unverzerrt geschätzt werden können und bei dieser Schätzung auch eine vorgegebene Genauigkeitsanforderung eingehalten wird.«

Bezogen auf die Anforderungen aus der ImmoWertV kann dies so interpretiert werden, dass z. B. der geschätzte Erwartungswert für den Preis je Quadratmeter Wohnfläche für gebrauchte Eigentumswohnungen in der Stadt A unverzerrt ist und den Genauigkeitsanforderungen entspricht, wie sie die Wertermittlungsaufgabe erfordert. Dabei ist der Begriff »Genauigkeit« klarer zu fassen.

Wenn ein Wert für einen bestimmten Parameter für eine ganze Stadt gültig sein soll, jedoch die Stichprobe geografisch verzerrt, z. B. lediglich aus einem Stadtteil entnommen ist, und die Streuung der Kaufpreise so stark ist, dass eine dem Wertermittlungszweck entsprechende genaue Aussage zu dem Erwartungswert nicht möglich ist, liegt keine Repräsentativität der Stichprobe vor. Dies gilt ebenso, wenn beispielsweise immer nur sehr wenige Veräußerer (z. B. die Gemeinde oder ein Investor) für mehrere Objekte auftreten, die Kaufpreise insofern miteinander korreliert sind und sich dadurch zwar eine geringe Streuung ergibt, die jedoch eine Verzerrung in Bezug auf den interessierenden Parameter aufweist.

Liegt also Repräsentativität im o. g. Sinne nicht vor, was nicht selten der Fall ist, darf auch nicht von einer »repräsentativen Stichprobe« gesprochen werden, und es muss benannt werden, dass die Anforderungen an die Geeignetheit der für die Wertermittlung erforderlichen Daten nach § 9 Abs. 1 Satz 1 ImmoWertV nicht vorliegen. Die Stichprobe könnte allenfalls als »informative Stichprobe« bezeichnet werden, die Aufschluss über bestimmte Informationen, Trends oder Eindrücke zum Immobilienmarkt geben (Quatember 2015).

5 Methoden, Modelle und Modellbildung

Bei der Bestimmung des Immobilienwertes ist es das Ziel, eine fundierte Prognose des Wertes einer Immobilie zu erstellen. Dabei spielt die präzise Definition des Begriffs »Wert« eine zentrale Rolle. Im Allgemeinen wird dazu der in § 194 BauGB definierte »Verkehrswert« zugrunde gelegt. Ziel ist es, den Wert einer Immobilie zu ermitteln, für die kein Kaufpreis bekannt ist, wobei dieser Wert denjenigen Preis darstellen soll, der unter »normalen Marktbedingungen« mit hoher Wahrscheinlichkeit erzielt werden könnte.

Aus der Perspektive statistischer Anwendungen handelt es sich hierbei um die Prädiktion (Schätzung) eines Preises auf der Grundlage bestehender Beobachtungen von Objekten, die die Grundgesamtheit repräsentieren. Dieser prädiizierte Preis wird, um eine Abgrenzung zu einem »echten« Preis zu finden, als »Wert« bezeichnet, weil er allgemeingültiger ist als ein letztlich erzielter Preis. Die Grundlage der statistischen Schätzung bildet die Anwendung verschiedener statistischer Methoden. Die Anwendung dieser Methoden führt zur Entwicklung von Modellen, die die Realität abbilden. Diese Modelle sind naturgemäß mit Einschränkungen bezüglich der Schätzgenauigkeit zu Objekten der Grundgesamtheit verbunden, ermöglichen jedoch gleichzeitig ein tieferes Verständnis der Mechanismen in eben dieser Grundgesamtheit.

Dabei werden häufig die Begriffe »Methode« und »Modell« miteinander vermischt, was leicht zu Missverständnissen hinsichtlich der Beurteilung der Ergebnisse führen kann. Es ist daher von zentraler Bedeutung, diese Begriffe klar zu trennen und unabhängig voneinander zu betrachten. Die Wortwahl in der ImmoWertV weicht von dieser Sichtweise derzeit noch ab.

5.1 Methoden

Im Kontext der statistischen Datenanalyse wird der Begriff »Methode« als ein systematischer Ansatz verstanden, der zur Sammlung, Analyse, Interpretation und Ableitung von Schlussfolgerungen aus Daten verwendet wird. Im Wesentlichen handelt es sich dabei um Instrumente oder Techniken zur Erhebung von Daten, zur Analyse, zur Interpreta-

tion der Ergebnisse und zur Ableitung von Erkenntnissen aus der Grundgesamtheit von Beobachtungen über Phänomene.

Von besonderer Bedeutung ist, dass der Zweck der Analyse, die vorliegenden Daten sowie deren Güte maßgeblich die Wahl der Methode beeinflussen. Zu den statistischen Methoden zählen klassische Verfahren der Mittelwertbildung, Klassifikationsverfahren (wie etwa Entscheidungsbäume oder Random-Forest-Ansätze), Regressionsverfahren (beispielsweise die lineare Regression oder Regressionsmethoden mittels Entscheidungsbäumen), aber auch neuronale Netzwerke mit unterschiedlichen Tiefenstrukturen einschließlich Deep-Learning-Ansätzen. Die Methoden dienen dazu, die Komplexität der Daten zu modellieren und tiefergehende Einsichten in die Strukturen und Zusammenhänge der zugrunde liegenden Grundgesamtheit aus einer Stichprobe zu ermöglichen. Bei der Bezeichnung der Modelle wird oft die Bezeichnung der Methode vorangestellt (z. B. Median-Modell, Regressionsmodell, Random-Forest-Modell etc.).

5.2 Modelle

Im gleichen Kontext kann der Begriff »Modell« als eine vereinfachte, formale Repräsentation der Realität (Grundgesamtheit) verstanden werden. Ein Modell wird entwickelt, um spezifische Aspekte eines komplexen Phänomens oder Systems – wie beispielsweise die Bestimmung von Indikatoren für Immobilienpreise – besser zu verstehen. Dabei handelt es sich bei einem Modell um eine statistische Struktur, die die Beziehungen zwischen verschiedenen Variablen beschreibt und es ermöglicht, Vorhersagen zu treffen sowie Muster und Zusammenhänge in den Daten zu identifizieren und messbar zu machen. Diese Daten können sowohl aus quantitativen Erhebungen als auch aus qualitativen Einschätzungen von Experten oder anderen Quellen stammen. Eine wesentliche Voraussetzung ist dabei stets, dass die erhobenen Beobachtungen unabhängig voneinander sind und die Grundgesamtheit hinreichend repräsentieren. Die Notwendigkeit einer klaren und präzisen Definition der Grundgesamtheit versteht sich hierbei von selbst.

Es gibt verschiedene Typen von Modellen, die sich je nach Zielsetzung der Analyse und den angewandten Methoden unterscheiden. So beschreibt ein Regressionsmodell beispielsweise die Beziehung zwischen einer abhängigen quantitativen Variablen und einer oder mehreren unabhängigen Variablen. Klassifikationsmodelle hingegen dienen dazu, Beobachtungen vordefinierten Kategorien zuzuordnen. Darüber hinaus sind Random-Forest-Modelle, Neuronale Netzwerke und Deep-Learning-Modelle besonders geeignet, komplexe und nicht-lineare Zusammenhänge abzubilden.

Zusammenfassend ist festzustellen, dass ein Modell im Rahmen der statistischen Datenanalyse ein unverzichtbares Instrument darstellt, um das Verhalten von Systemen

zu erfassen und besser zu verstehen. Ein Modell ist somit das Ergebnis der Anwendung einer statistischen Methode auf eine spezifische und geeignete Stichprobe, die letztlich zur Veranschaulichung, Erklärung und Vorhersage von Phänomenen einzelner Objekte der Grundgesamtheit dient.

5.3 Modellbildung

Insgesamt ist die Methodologie der Modellbildung mit Immobilienmarktdaten als ein Prozess anzusehen, der von der Datensammlung und dem Input geeigneter »Rohdaten«, also den Kaufpreisen und ihren Entstehungsumständen, über deren Vorverarbeitung (Preprocessing), das Trainieren eines Modells sowie seiner Interpretation und Validierung mit Testdaten bis hin zu einem modellbasierten Ergebnis einer Immobilienwertermittlung reicht (Abb. 1).

Immobilienwertermittlung mit modellierten Daten

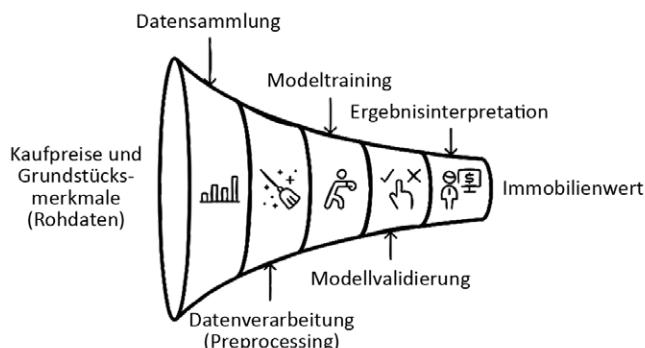


Abb. 1: Immobilienwertermittlung und Datenmodellierung

Dabei können bei der Modellbildung unterschiedliche Methoden im o.g. Sinne eingesetzt werden, bei denen die Güte der ermittelten Modelle verglichen wird und dann das »beste Modell« ausgewählt und ggf. weiter evaluiert wird. In der Regel ist die Modellbildung ein iterativer Prozess.

Statistisch ausgelegte Modellbildungsansätze haben in Bezug auf die in Deutschland normierten Wertermittlungsverfahren eher Ähnlichkeit mit dem traditionellen Vergleichswertverfahren, während die Sachwert- und Ertragswertverfahren weiter abweichen, da sie auf theoretischen Annahmen beruhen und erst durch Faktoren wie den Sachwertfaktor oder den Liegenschaftszinssatz an den Markt angepasst werden. Verhaltensökonomische Verzerrungen im Markt oder in der Wertermittlung selbst werden dabei in der Regel nur unzureichend berücksichtigt und können stark unterschätzt werden. So zeigte sich etwa um das Jahr 2020, dass Liegenschaftszinssätze mancherorts negativ waren oder nahe Null lagen, was den klassischen Annahmen einer angemessenen Verzinsung des eingesetzten Kapitals widersprach. Erwerber von Mehrfamilienhäusern haben anscheinend weitere Aspekte in ihre Überlegungen einbezogen, die das klassische Ertragswertverfahren nicht abbilden konnte. Dies könnten beispielsweise auch spekulative Überlegungen zu einem späteren gewinnbringenden

Verkauf sein, die in der Ableitung des Liegenschaftszinssatzes dadurch ihren Niederschlag finden, dass der Kaufpreis im Verhältnis zu den Erträgen zu hoch erscheint.

Die Güte der Ergebnisse klassischer Verfahren hängt daher stark von der Verfügbarkeit und der Genauigkeit von für die Wertermittlung erforderlichen Daten ab, was die Entwicklung von statistischen Modellen notwendig macht, die Marktphänomene realitätsnah abbilden – besonders in unsicheren sozialen, ökonomischen und politischen Zeiten, wie wir sie derzeit erleben.

Mit verbessertem Zugang zu Immobilienmarktdaten und der steigenden Akzeptanz gut und fortwährend trainierter Wertermittlungsmodelle müssen traditionelle Methoden zunehmend in den Hintergrund treten. Für die Immobilienwertermittlung ist es daher entscheidend, den Modellbegriff so zu verstehen, dass ein Modell auf Basis geeigneter Daten trainiert und anschließend mit anderen Daten (Testdaten) der gleichen Grundgesamtheit evaluiert wird, um eine fundierte Aussage über seine Güte zu treffen. Dazu ist das »permanente Lernen« des Modells aus immer wieder neu hinzugefügten Daten im Blick zu behalten, da die Beurteilung der Marktteilnehmer über die Bedeutung bestimmter Einflüsse sich dynamisch mit der Zeit verändert.

Erst auf einer solchen gedanklichen Grundlage ist ein Modell geeignet, realitätsnahe Verkehrswerte oder für die Wertermittlung erforderliche Daten im Sinne der ImmoWertV zu liefern. Dabei ist anzumerken, dass die Bezeichnung »Modell« sowohl in der ImmoWertV als auch in der ImmoWertA nicht dem korrekten statistischen Verständnis des Modellbegriffs entspricht und hier ein klärender Nachbesserungsbedarf ins Auge gefasst werden sollte.

6 Güte der Ergebnisse

Für Einschätzungen zur Güte eines Modells bildet die Abweichung des Modells von der Grundgesamtheit den Kern. Da aber die Grundgesamtheit in der Regel nicht bekannt ist, werden diese Abweichungen auf der Grundlage der genutzten Stichprobe geschätzt. Basis dabei ist die Abweichung des tatsächlichen (wahren) Werts in der Stichprobe von seinem prädizierten Wert. Diese Abweichungen werden als »error (Fehler)« oder »residuals (Residuen)« bezeichnet. Der Begriff stammt aus dem lateinischen Wort »erratum« für »Fehler«. Als Abkürzung wird der griechische Buchstabe »Epsilon (ϵ)« verwendet. Auf Basis der Residuen wird mit unterschiedlichen Metriken und Intervallen um die Schätzwerte herum auf die Güte des Modells in Bezug auf die Grundgesamtheit geschlossen.

Im Sinne einer transparenten Immobilienwertermittlung gewinnen die Forderungen nach der Darstellung und detaillierten Beschreibung der verwendeten statistischen Methoden sowie der auf dieser Grundlage entwickelten Modelle zunehmend an Bedeutung. Diese Ansätze sind sowohl in der ImmoWertV als auch in den zugehörigen

Anwendungshinweisen (ImmoWertA) bereits erkennbar. So wird in § 6 Abs. 4 ImmoWertV formuliert, dass der Verkehrswert auf Grundlage des Verfahrenswertes des angewandten Wertermittlungsverfahrens »unter Würdigung seiner oder ihrer Aussagefähigkeit« zu ermitteln ist. Die ImmoWertA konkretisieren diese Forderung in Ziff. 6.(4) dahingehend, dass die Aussagefähigkeit »insbesondere in Hinblick auf die Qualität der zugrunde gelegten Daten« zu würdigen ist. Allerdings bleibt unklar, was genau unter »Qualität« in diesem Zusammenhang zu verstehen ist.

Geht es um die Ergebnisse der Anwendung von Wertermittlungsmethoden (Verfahrenswerte), was unzweifelhaft intendiert ist, so muss die **Güte des Ergebnisses** im Fokus stehen. Dabei kann dieser Begriff mit mehreren Dimensionen beschrieben werden:

1. Die **Genauigkeit** des Ergebnisses beschreibt, wie nah das Modell an dem »wahren« Verhalten der Merkmale in der Grundgesamtheit liegt.
2. Die **Robustheit** des Ergebnisses beschreibt, wie stabil das Ergebnis gegenüber Änderungen der Eingangsdaten, also z.B. der Verwendung einer anderen Stichprobe aus der gleichen Grundgesamtheit, ist (Modell-Performance).
3. Die **Validität** der Methoden und Annahmen beschreibt, ob die richtigen Annahmen zu den Eingangsgrößen für das Modell und die zu der Wertermittlungsaufgabe passenden Methoden verwendet wurden.
4. Die **Relevanz** des Ergebnisses beschreibt, ob das Ergebnis für die Wertermittlungsaufgabe inhaltlich und kontextuell geeignet ist; z.B., ob der Verkehrswert nach § 194 BauGB ermittelt werden soll oder ein überschlägiger Schätzwert auf der Grundlage von Angebotsdaten.
5. Die **Transparenz** beschreibt, ob das Ergebnis reproduzierbar und die zugrunde liegenden Daten, Annahmen und Methoden sowie das Modell nachvollziehbar dokumentiert sind.

Zusammengefasst kann die Güte eines Modells durch seine **Ergebnisgenauigkeit** (durch das Konfidenzintervall) und seine **Leistungsfähigkeit** (Modell-Performance) beurteilt werden.

Im Kontext der Anwendung der Wertermittlungsverfahren und hier insbesondere im Rahmen der Würdigung der Aussagefähigkeit der Verfahrenswerte nach § 6 Abs. 4 ImmoWertV i. V. m. Ziff. 6.(4) ImmoWertA kommt es also nicht nur auf den Verfahrenswert selbst an, sondern auch auf seine Güte im o. g. Sinne.

Bei der Ermittlung von für die Wertermittlung erforderlichen Daten gilt, dass es neben dem Wert selbst ebenfalls auf seine Güte ankommt. Wird etwa der Liegenschaftszinssatz abgeleitet, so ist es erforderlich, die Güte des veröffentlichten Wertes zu dokumentieren (vgl. Anhang A.II ImmoWertA) und in den ImmoWertA wird angeregt, wie die Güte eines Wertes dargestellt werden sollte. Dies erfolgt durch die Darstellung eines Konfidenzintervalls (auch als Vertrauensintervall oder Vertrauensbereich bezeichnet). Nur durch die Angabe des Erwartungswertes in Kombina-

tion mit einer quantitativen Darstellung seiner Genauigkeit ist es möglich, die Auswirkungen der zugrunde liegenden Daten auf den Liegenschaftszinssatz, den Sachwertfaktor oder den Vergleichswertfaktor sowie damit auch auf den Verfahrenswert insgesamt zu bewerten.

Der Begriff »Güte der Ergebnisse« lässt sich im Kontext der Immobilienwertermittlung also dahingehend präzisieren, dass der jeweilige Verfahrenswert – also der aus der Stichprobe ermittelte statistische Erwartungswert – zusammen mit seinem **Konfidenz- bzw. Vertrauensintervall** betrachtet werden sollte. Gleichermaßen gilt für die sonstigen für die Wertermittlung erforderlichen Daten, ermittelt durch die Gutachterausschüsse oder ggf. auch andere Einrichtungen. Das Konfidenzintervall ermöglicht eine fundierte Aussage über die Genauigkeit des ermittelten Wertes und somit auch über die Aussagefähigkeit der zugrunde liegenden Daten sowie anderer relevanter Eingangsinformationen. Dies bezieht sich ausdrücklich nicht auf den nach Abwägung der Güte der Verfahrenswerte durch den Sachverständigen festzustellenden Verkehrswert, der sowohl im nationalen als auch im internationalen Kontext der Wertermittlung als »Punktwert« definiert ist. Ebenso wenig bezieht es sich auf den zu veröffentlichten Bodenrichtwert; dies wird in § 13 Abs. 2 Satz 3 ImmoWertV eindeutig und klar definiert, für den Verkehrswert liegt es wohl auf der Hand und ist nicht ausdrücklich definiert.

Insgesamt ist es von entscheidender Bedeutung, den Begriff »Ergebnisgenauigkeit« von dem Begriff »Leistungsfähigkeit (Modell-Performance)« abzugrenzen. Während die »Ergebnisgenauigkeit« die Frage klärt, wie nah das Ergebnis eines spezifischen Verfahrens an dem »wahren Wert« liegt, fokussiert die »Modell-Performance« auf die Fähigkeit des Modells, Ergebnisse zu liefern, die auch bei wechselnden Stichproben aus der gleichen Grundgesamtheit in ihrer Genauigkeit (oder auch Ungenauigkeit) vergleichbar bleiben.

Beide Konzepte sind für die transparente Immobilienwertermittlung unerlässlich, jedoch adressieren sie unterschiedliche Aspekte des Bewertungsprozesses. Die Ergebnisgenauigkeit ist direkt mit der Vertrauenswürdigkeit des spezifischen Ergebnisses verknüpft, während die Modell-Performance den gesamten Prozess der Modellentwicklung und -anpassung bewertet, um zu beurteilen, ob das Modell sowohl theoretisch fundiert und praktisch anwendbar ist, also die Grundgesamtheit ausreichend widerspiegelt. Eine sorgfältige Differenzierung dieser beiden Konzepte trägt erheblich zur Verbesserung der Transparenz und Nachvollziehbarkeit in der Immobilienwertermittlung bei.

7 Fazit

Dass die Bundesrepublik Deutschland vor enormen Herausforderungen steht, wird allgemein nicht bezweifelt und ist Inhalt täglicher Nachrichten. Dass dies auch die Wertermittlung von Immobilien betrifft, mag zunächst nicht so

wichtig erscheinen und das Gefühl der Betroffenheit hält sich insgesamt in Grenzen. Wird jedoch die Bedeutung der Werte von Immobilien und deren Entwicklung für jede Privatperson, für Investorinnen und Investoren und nicht zuletzt auch für den Staat in den Blick genommen, verändert sich die Perspektive. Die Einordnung von Immobilienwerten beispielsweise in das Konzept von Klimaschutzmaßnahmen, den Ausbau von Wind- und Sonnenenergieanlagen, in die Finanzierung von staatlichen Eingriffen in das Privateigentum beim Ausbau des ÖPNV, die Entwicklung von Förderkulissen im Zusammenhang mit der Schaffung von bezahlbarem Wohnraum etc. erweitert den Blick auf die Bedeutung von Transparenz und Güte von Immobilienbewertungen und -marktanalysen erheblich. Es bleibt der Schluss, dass die Immobilienwertermittlung vor einem Paradigmenwechsel steht, der sich durch die Immobilienwertermittlungsverordnung (ImmoWertV 2021) zwar andeutet, jedoch keinen disruptiven Charakter zeigt.

Wegen der zunehmenden und schneller werdenden Veränderungen auf dem Immobilienmarkt kann davon ausgegangen werden, dass es zu einer erheblichen Veränderung der Anforderungen an die Immobilienwertermittlung kommen kann. Eine Rolle spielt dabei, dass die zunehmende Verfügbarkeit von Daten und der Einsatz moderner Technologien, wie der Einsatz Künstlicher Intelligenz, neue Möglichkeiten für präzisere und objektivere Bewertungen und Analysen eröffnet und gleichzeitig der Einsparung von Personal und anderen Ressourcen entgegenkommt. Die traditionellen Ansätze, die auf der individuellen Expertise von Sachverständigen beruhen, reichen in absehbarer Zeit sicher nicht mehr aus, um den steigenden Anforderungen gerecht zu werden. Stattdessen müssen evidenzbasierte Methoden in den Fokus rücken, um Verzerrungen zu minimieren, die Güte der Ergebnisse von Wertermittlungen zu verbessern und die Immobilienmarktanalysen einem breiten wissenschaftlichen und kritischen Diskurs zu Methoden, Modellen und Ergebnissen zugänglich zu machen.

Ein zentraler Punkt dabei ist die Bedeutung der Definition und der Bewertung der »Ergebnisgüte«. Diese beschreibt nicht nur die Genauigkeit der Ergebnisse, sondern auch die Robustheit, Validität und Relevanz eines Modells im spezifischen Anwendungskontext. Dieses wiederum ist entscheidend, um das Vertrauen in die Ergebnisse und ihre Eignung für verschiedene Wertermittlungszwecke zu gewährleisten. Die Ergebnisgüte stellt sicher, dass ein Modell nicht nur theoretisch fundiert und in sich logisch ist, sondern auch in der Praxis zuverlässige Ergebnisse liefert und nachvollziehbar ist.

Die präzise Definition von Ergebnisgüte schafft einen Rahmen, innerhalb dessen die objektive Bewertung der Leistungsfähigkeit von Modellen ermöglicht wird und beurteilt werden kann und wie genau die Ergebnisse einer Schätzung mit diesem Modell sind. Transparent kommunizierte Informationen über die Methoden, die Datenbasis und die Güte der Ergebnisse sind unerlässlich, um Nutzerinnen und Nutzern die erforderliche Grundlage für fundierte Entscheidungen zu geben. Nur durch die konsequente Anwendung und Weiterentwicklung solcher Ansätze kann die Immobilienwertermittlung zukunftssicher gestaltet werden und den komplexen Anforderungen des modernen Marktes gerecht werden.

Literatur

- Coombs, C. H., Dawes, R. M., Tversky, A., Wendt, D. (1975): Mathematische Psychologie: Eine Einführung. Beltz, Weinheim.
- Europäische Kommission, High-Level Expert Group on Artificial Intelligence (2018): A definition of AI: Main capabilities and scientific disciplines. Brüssel.
- Gabler, S., Quatember, A. (2012): Das Problem mit der Repräsentativität von Stichprobenerhebungen. In: Jahrbuch 2012, hrsg. vom Verband Schweizer Markt- und Sozialforschung, 17–19.
- Graf, R. (2015): »Heuristics and Biases« als Quelle der Vorstellung – Verhaltensökonomische Forschung in der Zeitgeschichte. In: Zeithistorische Forschungen/Studies in Contemporary History, Heft 3/2015, 12. Jg., 511–519. DOI: 10.14765/zff.dok-1432.
- Kahneman, D., Slovic, P., Tversky, A. (eds.) (1982): Judgment under uncertainty: Heuristics and biases. Cambridge University Press.
- Kahneman, D., Tversky, A. (1979): Prospect Theory: An Analysis of Decision under Risk. In: Econometrica, Vol. 47, No. 2, 263–292. DOI: 10.2307/1914185.
- Quatember, A. (2015): Statistischer Unsinn: Wenn Medien an der Prozenthürde scheitern. Springer Spektrum, Berlin und Heidelberg. DOI: 10.1007/978-3-662-45335-3.

Kontakt

Dipl.-Ing. Peter Ache
FIG – International Federation of Surveyors
Vorsitzender der FIG-Commission 9 »Valuation and the Management of Real Estate«
peter.ache.fig@achemail.de

Dieser Beitrag ist auch digital verfügbar unter www.geodaeie.info.