

# Automatisierte Bewertungsverfahren im Rahmen kreditwirtschaftlicher Wertermittlung

*Mehr Transparenz wagen*

Sebastian Hein, Lorenz Thomschke und Albert Gass

WORKING PAPER 2021-4-1

VALUE AG  
TEAM MARKTDATEN

Heidestraße 8  
10557 Berlin  
E-Mail: [info@value-marktdaten.de](mailto:info@value-marktdaten.de)  
Stand: 15. April 2021

## Hintergrund

Der digitale Wandel verändert auch die Immobilien- und Bewertungsbranche. BigData, SmartData oder künstliche Intelligenz haben längst Einzug in den Branchenwortschatz gehalten und werden gern als Heilsbringer für die Herausforderungen der Zukunft bemüht. Auch im Rahmen kreditwirtschaftlicher Bewertungsprozesse könnten daten- und modellbasierte Ansätze vielversprechende Lösungen bieten. Leider werden diese Ansätze kaum im Kontext empirischer Grundlagen diskutiert.

Eine differenzierte Betrachtung und Einordnung von verwendeter Datenbasis, herangezogener Modellansätze und damit einhergehender Schätzqualitäten fehlt im fachlichen wie auch im aufsichtsrechtlichen Diskurs. Deswegen ist der Standardisierungsgrad kreditwirtschaftlicher Bewertungsprozesse insbesondere im Kontext der Kreditvergabe vergleichsweise gering. Daneben ist auch die heterogene Immobiliendateninfrastruktur Grund für den geringen Digitalisierungsgrad in der kreditwirtschaftlichen Wertermittlung.

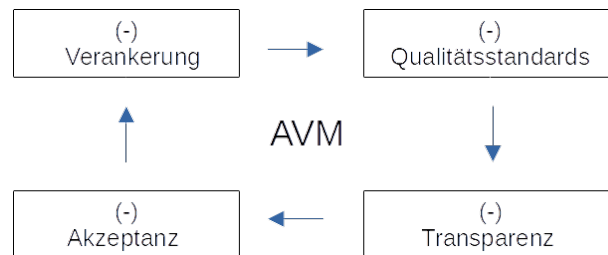
Gleichzeitig ist die Bankenlandschaft – gerade im Bereich der privaten Baufinanzierung – mit einem enormen Margendruck konfrontiert und gerät daher zunehmend unter Standardisierungs- und Digitalisierungszwang. Die Prozessoptimierung zielt dabei sowohl auf Kosteneinsparungen als auch auf die Entscheidungsgeschwindigkeit ab. Vielfach sind daher automatisierte Bewertungsverfahren (im Folgenden AVM) im Einsatz, die bereits seit geraumer Zeit von zahlreichen Anbietern zur Verfügung gestellt werden. Regulatorisch werden diese Verfahren aber lediglich als Entscheidungsunterstützungsmodelle eingeordnet und sind nicht direkt zur Bewertung bzw. zur Wertfestsetzung vorgesehen. Daher bestand bisher auch keine regulatorische Anforderung, allgemeine Qualitätsstandards für diese Modelle zu definieren.

Die Qualitätsbeurteilung der Verfahren bleibt damit den Anwendern überlassen. Wenn allerdings kaum Benchmarks zur Schätzgüte existieren und Modellannahmen, Aufbau sowie Backtesting Ergebnisse nicht offen gelegt werden, entsteht ein Black-Box-Charakter, der augenscheinlich nicht nur in Deutschland besteht:

„Whilst there are a large number of AVM vendors, the inner workings of the models and details of their specification are not released. Vendors do test their models regularly for accuracy, comparing individual property AVM valuations against achieved market prices, some claiming they have these figures independently assessed. However, the figures are not normally disclosed, especially by European vendors, and this non-disclosure puts a constraint on the analyses which can be objectively undertaken as regards an assessment of the reliability and accuracy of the models.“  
Matysiak (2017)

In der Folge rücken nicht Qualität, Schätzgüte und Datenbasis in den Fokus, sondern Innovationsversprechen oder Pricings. Anwender können zwischen den Bewertungslösungen qualitativ kaum noch unterscheiden. Der Markt wird intransparent noch bevor Marktreife erreicht ist. Außerdem leidet die Akzeptanz der Modelllösungen, wenn Schätzgenauigkeiten nicht offen thematisiert werden. Eine Emanzipation der AVMs vom Entscheidungsunterstützungsmodell hin zur umfänglichen Bewertungslösung bleibt somit verbaut.

### Das Akzeptanz-Transparenz-Dilemma



Letztlich ergibt sich aus der Gemengelage ein Dilemma: Ohne gesetzliche Verankerung entstehen keine Qualitätsstandards, ohne Qualitätsstandards entsteht keine Transparenz und ohne Transparenz keine Akzeptanz.<sup>1</sup> Das Akzeptanz-Transparenz-Dilemma droht langfristig als Innovationsbremse zu wirken.

„In the future, it is expected that progress in information technology and the development of a large property and transaction database will increase the precision of advanced statistical models. A strict restriction on the use of those models could hamper development in this market and the overall progress of the valuation market“ (EBA 2020, S.82)

In der *Novellierung der Leitlinien zur Kreditvergabe und Überwachung* befasst sich die Europäische Bankenaufsicht EBA (2020) nun eingehend mit der Rolle statistischer Modelle im Rahmen der kreditwirtschaftlichen Wertermittlung. Dabei wird befürchtet, dass eine strikte Beschränkung der Verwendung statistischer Modelle die Entwicklung und den Gesamtfortschritt des Bewertungsmarktes behindern könnte. Statt automatisierte Bewertungsverfahren strikt zu verbieten, empfiehlt sie vielmehr aufsichtsrechtliche Standards, um die Transparenz und Genauigkeit der Modellergebnisse zu gewährleisten, (EBA 2020, S.81-83).

### Neue Anforderungen und Chancen durch EBA Leitlinien

Das im Juni 2020 veröffentlichte Papier der EBA (2020) mit Leitlinien zur Kreditvergabe und -überwachung formuliert erstmals allgemeine Anforderungen für statistische Modelle im Rahmen von Bewertungen.<sup>2</sup> Demnach sollten statistische Modelle hinsichtlich ihrer Konfidenzniveaus und Annahmen für die Beteiligten nachvollziehbar sein. Vor allem die berücksichtigten objektspezifischen Variablen sowie die Unsicherheit der Modelle sollten offengelegt werden. Außerdem wird in der Leitlinie gefordert, dass die verwendeten statistischen Modelle für Bewertungen

1. objekt- und standortspezifisch in ausreichender Granularität vorliegen,
2. valide und genau sind sowie einem robusten und regelmäßigen Backtesting gegen die tatsächlich beobachteten Transaktionspreise unterliegen,
3. auf einer ausreichend großen und repräsentativen Stichprobe basieren, die auf beobachteten Transaktionspreisen beruht und
4. auf aktuellen Daten von hoher Qualität basieren. (EBA 2020, S.58)

<sup>1</sup>Die Reihenfolge ist hier beliebig und soll nicht die zugrundeliegende Henne-Ei-Frage beantworten. Es ist auch möglich, dass die fehlende Qualität und Transparenz der AVM initial zu einer fehlenden gesetzlichen Verankerung geführt haben.

<sup>2</sup>EBA Guidelines on loan origination and monitoring.

Laut EBA sind bei der Verwendung von Modellansätzen letztlich die Institutionen für die Angemessenheit und Leistungsfähigkeit der Modelle verantwortlich und sie müssen ebenso sicherstellen, dass die Dokumentation der Modelle auf dem neuesten Stand ist (EBA 2020, ebd.).

Daraus resultieren nicht nur neue Anforderungen, sondern ebenfalls Chancen für die Transparenz und damit letztlich auch für die Qualität und die Akzeptanz von AVM-Lösungen. Vor allem aber dürfte die neue Leitlinie eine verstärkte Auseinandersetzung mit statistischen Modellen, Stichproben und Bewertungsqualität auslösen. Dies ist im Hinblick auf einheitliche Standards zur Qualitätsbeurteilung nach über einer Dekade der AVM-Nutzung längst überfällig.

## Qualitätsmaßstäbe im Kontext automatisierter Bewertung

Die EBA empfiehlt robuste und regelmäßige Backtestings gegen tatsächlich beobachtete Transaktionspreise (EBA 2020, ebd.).<sup>3</sup> Welcher Abweichungsgrad im Rahmen des Backtestings tolerierbar ist, bleibt allerdings offen. Das verwundert jedoch kaum, da selbst in der Fachliteratur Uneinigkeit über die Auswahl der Metriken zur Bestimmung der Modellabweichung herrscht (Steurer und Hill 2020).

$\text{MAE} = \frac{1}{N} \sum_{i=1}^N  e_i $	(Mean absolute error)
$\text{MDAE} = \text{MED}  e_i _{i=1}^N$	(Median absolute error)
$\text{RMSPE} = 100 \times \sqrt{\frac{1}{N} \sum_{i=1}^N (e_i^r)^2}$	(Root mean square percentage error)
$\text{MAPE} = \frac{100}{N} \sum_{i=1}^N  e_i^r $	(Mean absolute percentage error)
$\text{MDAPE} = 100 \times \text{MED}  e_i^r _{i=1}^N$	(Median absolute percentage error)
$\text{SMAPE} = \frac{100}{N} \sum_{i=1}^N \frac{ y_i - \hat{y}_i }{ y_i  +  \hat{y}_i }$	(Symmetric mean absolute percentage error)
$\text{FSD} = 100 \times \text{SD} (e_i^r)_{i=1}^N$	(Forecast Standard Deviation)
$Q_q = \frac{100}{N} \sum_{i=1}^N \mathbb{1} ( e_i^r  \leq q\%)$	( $Q_q$ )

Die hier dargestellten Qualitätsmaßstäbe dürften die gängigen Metriken sein, die zur Beurteilung von Schätzergebnissen herangezogen werden.<sup>4</sup> Bei welchem Maßstab ein Schätzverfahren am besten abschneidet, hängt neben der Modellspezifikation allerdings auch von der Verteilung der zu schätzenden Zielvariable ab. Im Rahmen automatisierter Wertermittlungen ist das der Transaktionspreis, der räumlich

<sup>3</sup>Alternative Verfahren wie z.B. der Abgleich mit sogenannten Appraisals, also gutachterlich festgestellte Werte, werden hier nicht thematisiert (vgl. dazu auch European AVM Alliance (2019)).

<sup>4</sup>Für alle Objekte  $i = 1, \dots, N$  aus einem Test- oder Validierungsdatensatz wird dabei auf Basis eines finalen Modells der Marktwert geschätzt. Die Abweichung zwischen Transaktionspreis ( $y_i$ ) und geschätztem Marktwert ( $\hat{y}_i$ ) gibt dann Auskunft über die Qualität des Schätzverfahrens. Allerdings heben sich positive und negative Abweichungen meist gegeneinander auf. Aus diesem Grund wird typischerweise die absolute Abweichung zur Berechnung des Key Performance Indicator (KPI) verwendet. Auf Basis der Abweichung  $e_i = y_i - \hat{y}_i$  bzw. der relativen Abweichung  $e_i^r = (y_i - \hat{y}_i)/y_i$  werden dann die angegebenen Metriken ausgewiesen. Eine ausführliche Diskussion der Metriken findet man beispielsweise bei Morley, Brito und Welling (2018).

sowie objekts- und segmentspezifisch mitunter starken Schwankungen unterliegt. Gleiches gilt – je nach Modellspezifikation – in der Regel auch für die Schätzergebnisse und deren Beurteilung. Letztlich können die unterschiedlichen Metriken verschiedene Stärken und Schwächen eines Schätzverfahrens aufdecken bzw. verschleiern. Optimalerweise sollte daher ein breites Spektrum möglicher Qualitätsmaßstäbe publiziert werden, um so eine möglichst hohe Transparenz in Anlehnung an wissenschaftliche Standards zu erzielen.

Der Mean Absolute Percentage Error (MAPE) sowie der Median Absolute Percentage Error (MDAPE) werden in zahlreichen Untersuchungen angegeben. Auch aufgrund der vergleichsweise einfachen Konstruktion und Verständlichkeit liefern sie gute Anhaltspunkte für die Beurteilung der Schätzgüte. Ein weiterer, häufig angegebener Maßstab ist die Absolute Percentage Error Range ( $Q_q$ ). Sie gibt an, wie viele Fälle des Backtestings sich innerhalb eines akzeptablen Fehlertoleranzkorridors ( $\pm q\%$ ) befinden. Als Fehlertoleranzen werden häufig Abweichungen von  $\pm 20\%$  akzeptiert, Fehlertoleranzen von  $\pm 10\%$  werden mitunter als unproblematisch eingestuft. Allerdings besteht auch hier kein Konsens: „What is an acceptable margin for error? This is a much-discussed concept.“ (Matysiak 2017, S.13).

## Auf der Suche nach einem Benchmark

Die European AVM Alliance (2019) schlägt als tolerierbares Maß  $\pm 20\%$  vom Transaktionspreis für die Fehlerstreuung vor. Matysiak (2017) evaluiert europäische und US-amerikanische Modelle und sieht mögliche Schätzgenauigkeiten und übliche Fehlertoleranzen im Bereich von  $\pm 10\%$  bis  $\pm 15\%$ . Die Evaluation der Schätzgenauigkeiten vollzieht er anhand der o.g. Metriken über die Intervalle  $\pm 5\%$ ,  $\pm 10\%$  und  $\pm 20\%$ .<sup>5</sup> Oftmals sei die Frage eines akzeptablen Fehlers aber eine Frage der Rechtsprechung.<sup>6</sup> Diesen Ansatz verfolgt auch Uter (2018) in ihrer Arbeit über die Approximationsgüte automatisierter Wertermittlungen innerhalb der Kleindarlehensgrenze der BelWertV. Sie zieht verschiedene, wenn auch ältere Urteile heran und leitet auf dieser Basis vier Stufen ab. Demnach wird eine übliche Abweichung dem Bereich  $\pm 10\%$  zugeordnet, tolerierte Abweichungen sieht sie bei  $\pm 20\%$  bzw. in Sonderfällen bei  $\pm 25\%$  und fehlerhafte Abweichungen im Bereich von  $> \pm 25\%$  um den tatsächlichen Preis. Auch Sanftenberg (2020) orientiert sich an der Rechtsprechung und leitet daraus Genauigkeitsanforderungen für Mehrfamilienhäuser von  $\pm 20\%$  bis  $\pm 30\%$  ab.

Da Matysiak (2017) im europäischen Raum keine Benchmarks findet, konzentriert er sich auf die Angaben von [Zillow](#) und der [HouseCanary Webseite](#). Zwischen den verschiedenen US-Staaten zeigen sich teilweise große Unterschiede. Die Median Absolutabweichung (MDAPE) schwankt zwischen 3,4 % und 15 %, der Anteil der Objekte innerhalb der  $\pm 10\%$  Spanne liegt zwischen 39,9 % und 81,5 %. Bei Zillow schwankt dieser Anteil je nach Staat zwischen 20 % und 92 %, die MDAPE liegt zwischen 3 % und 25 %. Laut aktuellen Angaben von Zillow liegt der US-weite Anteil der Schätzungen innerhalb der  $\pm 10\%$  Range für Off-Market Schätzer bei 59,2 %. Innerhalb der  $\pm 20\%$  Range liegen 80,5 %. Die US-weite Medianabweichung liegt bei  $\pm 7,7\%$ .

Matysiak (2017) hebt in seiner Studie hervor, dass die Schätzgenauigkeit von AVMs in Abhängigkeit von Objektgröße und -typ, Lage sowie Marktsegment variiert und dass es kaum zielführend sei, für alle Objekttypen einheitliche Maßstäbe anzusetzen.<sup>7</sup>

---

<sup>5</sup>Vgl. Matysiak (2017, S.6).

<sup>6</sup>Vgl. Matysiak (2017, S.13).

<sup>7</sup>Im Gutachteralltag zeigt sich zudem, dass auch die Qualität der zur Verfügung gestellten Unterlagen, die Komplexität des Bewertungsgegenstandes sowie die Dichte verfügbarer Vergleichsdaten einen wesentlichen Einfluss auf die Schätzgenauigkeit hat.

„The message which emerges from the HouseCanary and Zillow figures is that accuracy levels vary across property size bands and geographically. Consequently, it is inappropriate to attribute a one-size fits all accuracy metric across all properties.“ (Matysiak 2017, S.12)

Für Deutschland liegen kaum Benchmarkwerte vor. Uter (2018) hat in einer geringen Anzahl an Vergleichen zu Verkehrswerten  $Q_{25}$ -Genauigkeitsraten für 3 verschiedene deutsche Programmlösungen zwischen 67 % und 73 % ermittelt. Sanftenberg (2020) berechnet auf Basis geringer Fallzahlen für Mietswohnhäuser in Berlin eine  $Q_{20}$ -Genauigkeitsrate von 90 %.<sup>8</sup> Ansonsten ist das Testfeld in Deutschland sehr unausgereift. Ein Methodenvergleich zu verschiedenen Verfahren der Immobilienpreismodellierung, der gerade im Kontext maschinellen Lernens und künstlicher Intelligenz ausgesprochen wichtig wäre, ist damit kaum möglich. Das liegt auch am Mangel einer umfangreichen Transaktionsdatenbasis zur Kreuzvalidierung.

Hinrichs, Kolbe und Werwatz (2020) bedienen sich beispielsweise der Kaufpreissammlung des Gutachterausschusses. Damit erzielen sie Genauigkeitsraten für  $Q_{20}$  ( $\pm 20\%$ ) zwischen 34 % und 43 %. Die durchschnittliche Abweichung anhand des MDAPE liegt zwischen 25 % und 30 %. Bemerkenswerterweise stellen Hinrichs, Kolbe und Werwatz (2020) zudem fest, dass auch die klassische OLS Schätzung eine sehr gute Performance im Hinblick auf die Qualität verspricht: „Strikingly, OLS, the simplest AVM procedure, has the lowest root mean squared prediction error.“ (Hinrichs, Kolbe und Werwatz 2020, S.30) Gleichzeitig geht aus der Studie auch hervor, dass die Attributierung der Kaufpreissammlungen zu gering ist, um die Potenziale alternativer Modellansätze auszuschöpfen.

## Transparenz und Nachvollziehbarkeit umfasst mehr als den Algorithmus

Die wenigen belastbaren Studien zum Thema Automatisierte Wertermittlung in Deutschland zeigen deutlich, dass es derzeit weder konkrete Minimalanforderungen an die Datengrundlage und das Schätzverfahren, noch einheitliche Standards zur Qualitätsmessung gibt. Das wäre aus unserer Sicht jedoch eine notwendige Voraussetzung, um das eingangs angedeutete Akzeptanz-Transparenz-Dilemma aufzulösen. Die Leitlinien der EBA sind ein wichtiger Schritt in diese Richtung. Die mannigfaltigen Evaluationsmöglichkeiten einer automatisierten Wertermittlung zeigen aber ebenso wie der wissenschaftliche Diskurs darüber, dass die Kriterien konkretisiert werden müssten. Auch Angaben zu statistischen Details sollten dabei nicht vergessen werden.<sup>9</sup>

Gleichwohl dürften auch bei einer transparenten und einheitlichen Kommunikation der Qualitätsmaßstäbe einige Fragen offen bleiben. So ist ein aussagekräftiger Vergleich zwischen verschiedenen Schätzverfahren nur dann möglich, wenn die herangezogenen Qualitätsmaßstäbe auf ein und derselben Datenbasis beruhen. Andernfalls bleibt stets unklar, ob unterschiedliche Ergebnisse am jeweiligen Schätzverfahren oder an der räumlichen bzw. segment- oder objektspezifischen Verteilung der gewählten Datengrundlage liegen. Ein einheitlicher, öffentlich verfügbarer Datensatz zur Evaluierung der vorhandenen AVMs dürfte daher ein vielversprechender Ansatz sein, die Qualität der Modelle adäquat einzustufen.

Auch wenn mit Vorlage der aktuellen EBA Leitlinien eine automatisierte Bewertung im Rahmen der Kreditvergabe weiterhin nur als Entscheidungsunterstützung vorgesehen ist, geht die Leitlinie explizit auf das Wechselspiel zwischen Mensch (Gutachter/Bewerter) und Maschine (statistisches Modell) ein.

<sup>8</sup>Allerdings erfolgt der Vergleich hier gegen gutachterlich ausgewiesene Marktwerte und nicht gegen Transaktionspreise.

<sup>9</sup>Die relative Abweichung  $e_i^r = (y_i - \hat{y}_i)/y_i$  kann beispielsweise auch als  $e_i^r = (y_i - \hat{y}_i)/\hat{y}_i$  definiert werden (Schulz, Wersing und Werwatz 2014, S.15). Beide Definitionen sind aus statistischer Sicht weder falsch noch richtig, sie führen i.d.R. aber zu unterschiedlichen Testergebnissen.

Die EBA unterstreicht dabei, dass die Wertvorschläge aus Modellen übernommen werden dürfen, wenn die Konfidenzmaße, die berücksichtigten Eingaben und die Annahmen hinreichend nachvollziehbar sind.

Für die Anwender ergeben sich daraus neue Herausforderungen bei der Bewertung teils komplexer statistischer Verfahren. Konsistente (Kreuz-)Validierungsdaten könnten hier viele Vorteile bieten und die Evaluation der unterschiedlichen Schätzverfahren auf ein neues Niveau heben. Das wiederum dürfte die Skepsis gegenüber Automatisierten Wertermittlungsverfahren enorm abmildern. Insgesamt plädieren wir daher dafür, den fach- und aufsichtsrechtlichen Diskurs zum Thema der Modelltransparenz und Nachvollziehbarkeit ganzheitlicher zu führen:

„Das Generieren transparenter und nachvollziehbarer Modelle ist aktuell eines der wichtigsten Forschungsfelder des Maschinellen Lernens, um die Entscheidungsfindung eines intelligenten Systems jederzeit verstehen und bewerten zu können. Im Gegensatz zu „Black-Box-Modellen“ – rein statistischen datengetriebenen Lernmodellen – beziehen transparente Modelle auch Fachwissen zur Erklärung ein, so dass die Logik oder die einzelnen Ausgaben des Systems besser nachvollziehbar sind.“  
(BaFin 2018)

Transparenz setzt allerdings nicht erst am Algorithmus an, sie beginnt bereits bei der Datenerhebung. „Es ist leicht, Vorschriften über die Theorie des Beweises aufzustellen, aber der Beweis selbst ist schwer zu führen,“ schreibt Giordano Bruno.<sup>10</sup> Im Bereich automatisierter Wertermittlungen dürfte es umgekehrt sein.

---

<sup>10</sup>Giordano Bruno, italienischer Philosoph (1548 - 1600). *Von der Ursache, dem Prinzip und dem Einen*, 2. Dialog.

## Literatur

- BaFin, Bundesanstalt für Finanzdienstleistungsaufsicht (2018). „Big Data und künstliche Intelligenz: Maschinen dürfen auch bei automatisierten Prozessen nicht die Verantwortung tragen“. In: *BaFinJournal*. URL: [https://www.bafin.de/SharedDocs/Veroeffentlichungen/DE/Fachartikel/2018/fa\\_bj\\_1806\\_BDAI\\_Interview.html](https://www.bafin.de/SharedDocs/Veroeffentlichungen/DE/Fachartikel/2018/fa_bj_1806_BDAI_Interview.html).
- EBA, European Banking Authority (2020). „Guidelines on loan origination and monitoring“. In: *Final Report*. URL: [https://www.eba.europa.eu/sites/default/documents/files/document\\_library/Publications/Guidelines/2020/Guidelines%20on%20loan%20origination%20and%20monitoring/884283/EBA%20GL%202020%2006%20Final%20Report%20on%20GL%20on%20loan%20origination%20and%20monitoring.pdf](https://www.eba.europa.eu/sites/default/documents/files/document_library/Publications/Guidelines/2020/Guidelines%20on%20loan%20origination%20and%20monitoring/884283/EBA%20GL%202020%2006%20Final%20Report%20on%20GL%20on%20loan%20origination%20and%20monitoring.pdf).
- European AVM Alliance (2019). „European Standards for Statistical Valuation Methods for Residential Properties“. In: *EAA Standards*. URL: [https://www.europeanavmalliance.org/publications.html?file=files/ea/Downloads/EAA\\_Standards\\_2nd\\_Edition.pdf](https://www.europeanavmalliance.org/publications.html?file=files/ea/Downloads/EAA_Standards_2nd_Edition.pdf).
- Hinrichs, Nils, Jens Kolbe und Axel Werwatz (2020). *AVM and high dimensional data: Do ridge, the lasso or the elastic net provide an „automated“ solution?* Techn. Ber. FORLand-Working Paper. URL: <https://www.econstor.eu/bitstream/10419/227605/1/FORLand-2020-22.pdf>.
- Matysiak, George Andrew (2017). „Automated Valuation Models (AVMs): a brave new world?“ In: *Conference Paper*. URL: [https://www.researchgate.net/publication/319355261\\_Automated\\_Valuation\\_Models\\_AVMs\\_A\\_brave\\_new\\_world](https://www.researchgate.net/publication/319355261_Automated_Valuation_Models_AVMs_A_brave_new_world).
- Morley, Steven Karl, Thiago Vasconcelos Brito und Daniel T. Welling (2018). „Measures of model performance based on the log accuracy ratio“. In: *Space Weather* 16.1, S. 69–88.
- Sanftenberg, Anne (2020). „Value Drivers and Valuation Algorithms of Residential Investments“. In: *International Journal of Sciences: Basic and Applied Research* 54 (2). URL: <https://core.ac.uk/download/pdf/337387768.pdf>.
- Schulz, Rainer, Martin Wersing und Axel Werwatz (2014). „Automated valuation modelling: a specification exercise“. In: *Journal of Property Research* 31.2, S. 131–153. URL: <https://edoc.hu-berlin.de/bitstream/handle/18452/5138/46.pdf?sequence=1>.
- Steurer, Miriam und Robert J Hill (2020). „Metrics for Measuring the Performance of Machine Learning Prediction Models: An Application to the Housing Market“. In: *Graz Economics Papers – GEP*. URL: [https://www.researchgate.net/profile/Miriam-Steurer/publication/338992074\\_Metrics\\_for\\_Measuring\\_the\\_Performance\\_of\\_Machine\\_Learning\\_Prediction\\_Models\\_An\\_Application\\_to\\_the\\_Housing\\_Market/links/5e37fbdf92851c7f7f1822b7/Metrics-for-Measuring-the-Performance-of-Machine-Learning-Prediction-Models-An-Application-to-the-Housing-Market.pdf](https://www.researchgate.net/profile/Miriam-Steurer/publication/338992074_Metrics_for_Measuring_the_Performance_of_Machine_Learning_Prediction_Models_An_Application_to_the_Housing_Market/links/5e37fbdf92851c7f7f1822b7/Metrics-for-Measuring-the-Performance-of-Machine-Learning-Prediction-Models-An-Application-to-the-Housing-Market.pdf).
- Uter, Aline (2018). „Zur Approximationsgüte automatisierter Wertermittlungen innerhalb der Kleindarlehensgrenze der BelWertV am Beispiel ausgewählter Referenzobjekte“. *Magisterarb. Hochschule Anhalt – Bernburg*. URL: <https://opendata.uni-halle.de/bitstream/1981185920/12720/1/Masterarbeit%20Aline%20Uter.pdf>.