25th International Congress on Modelling and Simulation, Darwin, NT, Australia, 9 to 14 July 2023
mssanz.org.au/modsim2023

An application of machine learning in real estate economics: what extra benefits could machine learning techniques provide?

F. Chan [a], **R. Schulz** [b] **and** **Z. Zhang** [a]

[a] *School of Accounting, Economics and Finance, Curtin University, GPO BOX U1987, Perth, Western Australia*
[b] *Business School, University of Aberdeen, Dunbar Street, Aberdeen, United Kingdom, AB24 3QY*
Email: zhuoran.zhang@curtin.edu.au

Abstract: The applications of machine learning (ML) techniques in real estate practices has become popular recently. Specifically, ML techniques are often used to develop Automated valuation models (AVM), which purpose is to provide a price estimate of a particular property at a specified time. The main objective is to minimise human intervention in the process of price estimation. Apart from the users providing a set of inputs which, in the present context, would be a set of property features, the AVM would provide a price estimate without any human intervention in the estimating process. The recent literature show that such prediction process performed much better than the traditional approach. However, the presence of missing values in the data remains a major challenge in developing AVM using ML techniques. Therefore, this paper examines different approaches to handle missing values in the context of developing AVMs using various tree based algorithms.

This paper has two main objectives: (i) it examines the performance of Gradient Boosting Machine (GBM) in the presence of missing values. Recent literature suggested that GBM can still provide accurate prediction in the presence of missing values and this paper examines this claim in the context of AVMs; (ii) Using GBM, this paper compares some common strategies in managing missing values as well as a special strategy that only applicable for machine learning methods. This helps to identify any extra benefits in using machine learning techniques. It is worthwhile noting that data can be missing in both training and testing stages. Ideally, a model can be trained with missing values in the training set as well as having the ability to make predictions when some of the inputs (features) are missing in the test set.

The results show that the proposed GBMs can predict more accurately than the traditional hedonic model across different forecast criteria. The results are also unaffected by the choice of missing value strategies. This is consistent with the results in recent studies. In addition, the proposed implementation required minimum human intervention in both training and testing stages, even in the presence of missing values. Thus, machine learning methods would appear to be more efficient in addition to being more accurate in predicting house price.

Keywords: *Machine learning, Automated Valuation Model (AVM), Gradient Boosting Machine (GBM), missing values*

1 INTRODUCTION

The purpose of an Automated valuation model (AVM) is to provide a price estimate of a property at a particular time without, or with minimum, human intervention after the models are established (RICS 2021). In AVM literature, statistical models have been used to predict market prices of properties (Anglin & Gencay 1996, Parmeter et al. 2007, Haupt et al. 2010, for example). Recently, machine learning techniques have also been applied to predict prices of residential properties. For examples, Kagie & Van Wezel (2007), Kok et al. (2017), Mayer et al. (2019), and Schulz & Wersing (2021) apply suites of machine learning models, such as Random Forest, Gradient Boosting Machine (GBM) and Neural Network among others. The results generally show that machine learning methods could produce more accurate predictions than classical parametric models, but this benefit is relatively less compared with non-parametric models. Interestingly, GBM has always been ranked high, if not the highest, among the highest tier. However, discussions on their limitations are often lacking in these studies. Perhaps more importantly, the proposed AVMs in these studies still require substantial human intervention even after having been trained due to some reasons, such as missing values, which makes them not truly automated.

Missing values are common in property transaction data, which creates an obstacle in developing an AVM. Missing values can appear in the training (or estimating) stage of the AVM, but also in the prediction stage. Note that while it is sometimes possible to fill in missing observations, this is usually done manually on an observation-by-observation basis. This is extremely time-consuming and becomes impossible, or at least, in practical, once a dataset is large. In other words, an ideal method should allow missing values in the training or estimation stage of an AVM, and it should also allow predicting when some of the inputs (features) are missing. Complete case strategy is by far the most common method of dealing with missing values in many studies, for example, Kok et al. (2017), Schulz et al. (2014) and Steurer et al. (2021). It only keeps those observations that have no missing values, while the others are dropped. Its advantage is simplicity, since a standard analysis can be applied without further modifications (Little & Rubin 2002). However, it often leads to a sizeable loss of observations, and it also means that the resulting AVM cannot make predictions with incomplete inputs. Actually, the observed part of incomplete observations that is not missing but dropped, could be valuable in estimation. Simply ignoring them is unwise and may increase the sample selection bias if missing values are not missing at random. The preferred strategies should preserve all relevant information from each observation and handle missing values during model training (or estimation), such as Knight et al. (1998), Kagie & Van Wezel (2007) and Hinrichs et al. (2021). These require appropriate data preparation before estimation or that the training (or estimation) procedure can handle missing values. In any case, the ability to handle missing values is a crucial requirement in developing a robust and practical AVM given their frequency in real estate property-level transaction data (Krause & Lipscomb 2016).

This paper examines the prediction performance of Gradient Boosting Machine (GBM) when developing an AVM given missing values in data. The choice of GBM are two-folded. First, it outperforms other machine learning models in a similar context (Kok et al. 2017, Mayer et al. 2019, Schulz & Wersing 2021). Second, GBM can incorporate different strategies to handle missing values, which will be comprehensively examined in this paper. Two of them (complete case strategy and multiple imputation strategy) are commonly used in a wide range of areas. The third strategy namely missing value node strategy, is designed for treelike machine learning models, which has the ability to accommodate missing values in both the training and prediction stages and does not require any additional computation to fill the missing values. In order to examine robustness of their performances, this paper compares these methods and strategies under two different loss (objective) functions with different forecast criteria. Models are trained using a rolling window procedure with the usual out-of-sample predictions to assess their predictive performance. The preference is that the model is able to provide the most accurate predictions and meanwhile requires no or less extra modifications for data. The results will shed light on three research questions, (i) is GBM a valuable method to develop AVM with excellent prediction performance similar to other prediction problems as shown in the recent literature; (ii) Does the prediction performance depend on the choice of loss functions and (iii) what other benefits do machine learning techniques provide other than accurate predictions? The rest of the paper is organized as follows. Section 2 explains the method we use in the study. Section 3 presents the dataset used briefly, and Section 4 shows the analysis procedure and its results. Section 5 concludes.

2 GRADIENT BOOSTING MACHINE

This section introduces the GBM. Consider a dataset $\mathcal{D}_N = \{\mathbf{x}_i; y_i\}_{i=1}^N$, where $\mathbf{x}_i = (x_{1i}, \ldots, x_{Mi})$ contains the i^{th} observation of the M predictors and define $X_m = (x_{m1}, \ldots, x_{mN})'$ as the m^{th} predictor that contains

N values. Thus, $\mathbf{X} = (\mathbf{x}'_1, \ldots, \mathbf{x}'_N)' = (X_1, \ldots, X_M)$ is the $N \times M$ data matrix that some elements could be missing.

Gradient boosting machine (Friedman 2001, 2002) is one of tree based machine learning methods, that the fundamental is decision tree[1] (Breiman et al. 1984). Initially, the basic concept of decision tree is recursive partitioning. This process could be roughly summarized by two steps (James et al. 2017):

1. The predictor space, the set of possible values for M predictors, is divided into J distinct and non-overlapping regions, $R_1, \ldots, R_J$.
2. The predictions of observations are calculated in their own region, R_j.

Partitioning process starts from selecting the predictor (X_m) and the cut point (c), such that the predictor space is divided into two regions, $R_{Left}(m, c) = \{\mathbf{X}|x_{mi} < c\}_{i=1}^N$ and $R_{Right}(m, c) = \{\mathbf{X}|x_{mi} \geq c\}_{i=1}^N$. A more detailed description about decision tree can be found in Breiman et al. (1984) and Hastie et al. (2009). The algorithm of GBM is to grow a sequence of decision trees to improve predictions. Each new tree is grown using the information from previously grown trees, all the trees are in the same "family line" (James et al. 2017). The preliminary of GBM is to decide the hyperparameters[2], J terminal nodes (or the depth (or size) of each tree d), the number of trees (or iterations) B, and the learning rate λ. It starts of with an initial guess of response $F_0(\mathbf{X})$.

$$F_0(\mathbf{X}) = \gamma_0 = argmin_\gamma \sum_{i=1}^{N} L(y_i, \gamma). \tag{1}$$

where $L(\cdot)$ is the loss function. In each iteration b, the current "pseudo"-residuals are calculated using the previous information.

$$\tilde{y}_{ib} = -\Big[\frac{\partial L(y_i, F(\mathbf{x}_i))}{\partial F(\mathbf{x}_i)}\Big]_{F(\mathbf{X})=F_{b-1}(\mathbf{X})}, i = 1, \ldots, N. \tag{2}$$

Then, a J terminal nodes' tree is grown using the current "pseudo"-residuals $\tilde{y}_{ib}$ and all observations $\mathbf{x}_i$.

$$\{R_{jb}\}_1^J = J\ terminal\ nodes'\ tree(\{\tilde{y}_{ib}, \mathbf{x}_i\}_1^N). \tag{3}$$

The outputs of the tree in the iteration b are

$$\gamma_{jb} = argmin_\gamma \sum_{\mathbf{x}_i \in R_{jb}} L(y_i, F_{b-1}(\mathbf{x}_i) + \gamma). \tag{4}$$

Finally, the predictions are updated using the outputs.

$$F_b(\mathbf{X}) = F_{b-1}(\mathbf{X}) + \lambda \cdot \gamma_{jb} \cdot \mathbf{1}(\mathbf{x}_i \in R_{jb}). \tag{5}$$

After repeating B times, the final output predict values are

$$F_B(\mathbf{X}) = F_0(\mathbf{X}) + \sum_{b=1}^{B} \lambda\gamma_b. \tag{6}$$

The model details about GBM algorithm could be found in Friedman (2001, 2002).

Missing value node, a built-in technique, allows GBM to handle missing values in the fitting procedure. It treats missing values as a new category of observations when a node is split. Thus, GBM allows assigning missing values in the primary selected predictor of one split to a new node, rather than to the left or to the right as usual. The observations in the missing value node could be partitioned again as long as the estimation would be improved[3]. A schematic example is shown in Figure 1. If there are missing values in the selected predictor

[1] In this case, decision tree is for solving regression problem. The description introduces the characteristics of decision tree when it builds a regression model. For classification problem, the description is slightly different.

[2] These parameters could be called the tuning parameters also. They must be tuned for achieving the best performance. The depth (or size) of each tree d is commonly used in implementation packages rather than J terminal nodes.

[3] Missing value node is generated in each split as security mechanism, even the primary selected predictor of one split has no missing value. This is for prediction purpose, in case that missing values appear only in test data.

(X_1) of the top split, the observations with missing values are assigned to $R_{na,1} = \{\mathbf{X}|x_{1i} = NA\}_{i=1}^{N}$. The rest is distributed to the left if $x_{1i} < c_1$, and to the right if $x_{1i} \geq c_1$ $(i = 1, \ldots, N)$. If X_2 is unobserved in some observations, they are assigned to $R_{na,2}$ or $R_{na,3}$.

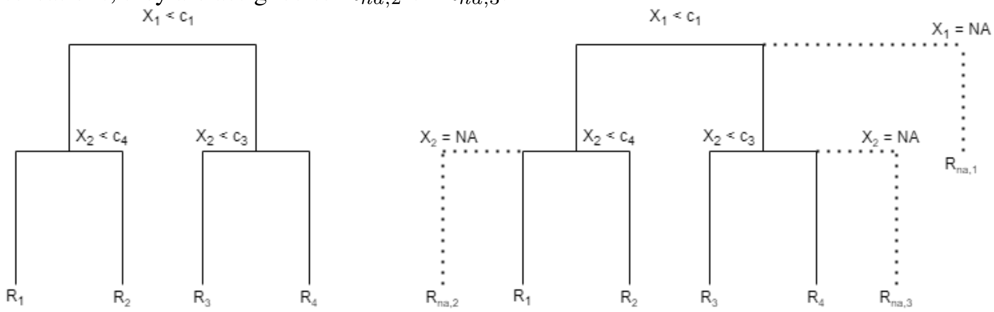

Figure 1. An example of missing value node in an individual tree.

3 A BRIEF DESCRIPTION OF DATA

The dataset is provided by the *Western Australian Land Information Authority*, that operates under the business name of *Landgate*. Only the market transactions of residential properties in the Greater Perth metropolitan area (excluding Mandurah) from 2015Q1 to 2020Q4 are included for the analysis. Each observation documents information about transaction information, parcel details, and dwelling details, such as sale price, date, location, property types[4] and the number of housing features. From 2015 to 2020, there are 21,231 (12% approx.) new established properties sold, 174,137 transactions (observations) in total. Missing values are preserved in the dataset.

The summary of variables is presented in Table 1. It summarizes the statistics for 174,137 observations. The price is deflated by the Residential Property Price Index (*RPPI*) from the ABS, the reference period is the financial year 2011–2012. In the table, some characteristics show low level of missingness, the rates are higher than 0.09% but lower than 1.6%. The missing rate of floor area, however, is around 35%, which is much heavier than the others. Compared with other datasets used in real estate literature, the missing rate of the dataset is mild[5].

Table 1. The summary statistics for the sold residential properties in the dataset, Jan 2015 - Dec 2020.

Variable	NA num	Mean	S.D.	Variable	NA num	Mean	S.D.
Price ($,000)	0	573.221	416.015	Dining	0	0.685	0.468
Land size (m^2)	0	927.504	5,054.839	Family	0	0.594	0.496
Floor area (m^2)	61,114	155.240	75.039	Game	0	0.203	0.407
Age (year)	1,243	27.504	21.937	Meal	0	0.308	0.464
Bedrooms	1,139	3.221	0.854	Study	0	0.197	0.407
Bathrooms	290	1.645	0.599	Car ports	0	1.499	0.770
Lounge	1,139	1.007	0.086	Tennis court	0	0.001	0.028
Kitchen	1,139	1.010	0.105	Pool	0	0.169	0.375
Tile-roof	2,707	0.798	0.401	latitude	0	-31.970	0.167
Brick-wall	169	0.935	0.247	longitude	0	115.856	0.096

4 PRICE PREDICTION FOR RESIDENTIAL PROPERTIES

To provide up-to-date price predictions, models are periodically maintained by adding new market transactions and removing old transactions in real estate industrial practice. In this study, rolling windows strategy is

[4]The most property type is houses (66.1%), followed by group houses (12.7%) and the rest types.

[5]Kagie & Van Wezel (2007) use the Dutch housing data that the missing rates of some features are from 0.2% to 70%. In Sydney (Australia) housing data, about 47% of full data have one or more missing values in the housing characteristics in Hill & Scholz (2018, Web appendix). Graz (Austria) residential transaction data used in Steurer et al. (2021) contain 11,250 incomplete cases, around 40% of total cases. In Zillow Prize data available in Kaggle, around 0.38% of observations don't have information about the number of bathrooms and bedrooms. 9.25% of lot size is unknown and 72.82% of air conditioner details are missing, etc.

applied to study the predictive performance of AVMs. The training datasets contain observations in two-year (eight-quarter) length period. We repeatedly shift this window by one quarter and re-apply the models. The testing datasets are constructed by observations transacted in the following quarter of training periods.

The model specification is similar to those of standard hedonic price models. The response is the logarithm of transaction price and the predictors are the housing characteristics, temporal and spatial factors ($\mathbf{x}_i$).

$$p_i = \log(P_i) = f(\mathbf{x}_i) + \epsilon_i. \tag{7}$$

where the function structure (f) is GBM[6] in this case. Three strategies for the missing value issue are applied, including missing value node strategy, complete case strategy and multiple imputation strategy[7]. In addition, two common loss functions (the least squares (LS) and the least absolute deviations (LAD)) are used to eliminate other interference. The linear models are the benchmark for the comparison. Thus, the players at the game table are introduced in Table 2, and all implementations could run automatically.

Table 2. The description of models.

Model	Model description
GBM (LS)	unmodified training sets (missing value node strategy), GBM and LS loss function.
GBM (LAD)	unmodified training sets (missing value node strategy), GBM and LAD loss function.
GBM (LS, MI)	multiple imputed training sets, GBM and LS loss function.
GBM (LAD, MI)	multiple imputed training sets, GBM and LAD loss function.
GBM (LS, C)	modified training sets with complete cases only, GBM and LS loss function.
GBM (LAD, C)	modified training sets with complete cases only, GBM and LAD loss function.
Linear (LS, MI)	multiple imputed training sets, linear function and LS loss function.
Linear (LAD, MI)	multiple imputed training sets, linear function and LAD loss function.

The log-scaled predictions are transformed back to the natural scale. We evaluate the performance of the candidate models based on the percentage prediction errors (e_k). The assessment measures we use are mean percentage error (MeanPE), median percentage error (MedianPE), mean absolute percentage error (MAPE), root mean squared percentage error (RMSPE), and percentage error range (PER) [8]. They cover bias, absolute difference, squared difference and error range, four major groups of metrics.

4.1 Results

The quarterly rolling windows are implemented on the dataset. Except models using multiple imputation strategy, the testing datasets used are the same, which may contain missing values. For models applying multiple imputation strategy, the missing values in the test sample are imputed by the imputation model for the training datasets, which follows the common process of multiple imputation strategy for forecasting. The overall results of the different implementations are shown in Table 3, which presents the overall error metrics through the 16 rolling windows.

4.2 Gradient boosting machine and linear hedonic model

Through results shown in the table, the GBMs clearly outperform the linear hedonic models with the same settings (GBMs (LS & LAD, MI) vs Linears (LS & LAD, MI)), that is in line with Mayer et al. (2019), Schulz & Wersing (2021). Averagely, the GBMs reduce mean absolute percentage error by around 8%, around 40% improvement to the linear hedonic models. They also present around 12% less for RMSPE, around 40% more accurate with respect to the benchmarks. For PER(10) (PER(20)), there are about 42% (15%) of observations that the absolute prediction errors are more than 10% (20%) of their valuations, there are

[6] The hyperparameters are tuned by grid searching and ten-fold cross-validation on the training sets. If two sets of hyperparameters have the same ranking, we choose the most parsimonious and conservative one. The idea is similar to the one standard error rule that is applied in Hastie et al. (2009).

[7] Missing value node strategy is directly applied on the unmodified training sets, the other two strategies need some adjustments on the data set.

[8] $PER(a) = \frac{1}{N_{test}} \sum_{k=1}^{N_{test}} \mathbf{1}(|e_k| \geq a)$

about 63% (36%) for the same metrics of the linear models. These improvements could be possibly due to modelling algorithm. GBM could measure more complex relationships among predictors rather than linearity. Additionally, partitioning process could distribute extreme values to some nodes to isolate them from the normal observations in the other nodes. These value might be well treated in their own nodes rather than with all observations. This may avoid huge accuracy drop in some extreme cases, especially compared with linear models.

Table 3. The predictive performance of models.

Model	MeanPE	MedianPE	MAPE	RMSPE	PER(10)	PER(20)
All test samples ($N = 113,104$)						
$GBM\ (LS)$	0.77	0.31	11.86	17.82	42.99	15.57
$GBM\ (LAD)$	0.84	0.18	11.77	18.52	41.60	15.14
$GBM\ (LS,\ MI)$	0.77	0.33	11.81	17.68	42.85	15.50
$GBM\ (LAD,\ MI)$	0.83	0.19	11.76	18.46	41.70	15.11
$GBM\ (LS,\ C)$	4.26	0.41	19.05	33.01	54.01	29.43
$GBM\ (LAD,\ C)$	3.84	0.12	18.95	32.44	53.28	29.52
$Linear\ (LS,\ MI)$	0.30	-3.17	19.84	30.19	64.03	36.42
$Linear\ (LAD,\ MI)$	3.58	0.14	20.04	30.58	62.69	35.98

4.3 Least squares (LS) and least absolute deviations (LAD)

Overall, LAD benefits AVMs more. On average, models using LAD reduce PER(10) by approximately 1%, Per(20) by roughly 0.5%. This indicates that more observations are accurately predicted when LAD is applied. LAD is less sensitive to outliers and can prevent their influence, because all errors are given the same weights when the loss is calculated. This could make estimates more accurate for most observations. In addition, LAD provides a special quantile of predictions' distribution, and these predictions come with a 50% confidence. This advantage and its extensions could benefit some special purposes in practice, such as risk management for mortgage portfolio.

4.4 Missing value strategies

Three missing value strategies are comprehensively compared within the GBM's family. Overall, the GBMs using missing value node strategy are the recommended choice regardless of loss functions. The complete case strategy is the simplest – the observations with missing information are discarded carelessly. However, the drawback is significant – the accuracy of prediction would be apparently decreased[9] when missing value occurs in testing sets. The GBMs ($LS,\ C$ and $LAD,\ C$) show roughly 7% more MAPE, 14% more RMSPE and 14% more PERs than the GBMs (LS and LAD). The multiple imputation strategy is to use complete cases and observed characteristics (omitting the target variable) to impute missing information multiple times in the training and testing sets. Compared with the GBMs (LS and LAD), the performance is at the same level. The differences in MAPE, RMSPE and PERs are negligible, less than around 0.1%. However, this strategy requires imputation steps before and after modelling, which makes the procedure much more complex than the other two. The above two have their own shortcomings, one has a simple process but low accuracy, the other is accurate enough but complex. The GBMs applying missing value node strategy become the game changer with a simple procedure and providing accurate predictions. This strategy allows missing values being simultaneously dealt with estimation in one step. It gives a clue or a chance to partition the incomplete cases, then, forecast using the observed information, regardless of missingness types.

5 CONCLUSIONS

This paper established several implementations of automated valuation models with different missing value strategies for residential property price prediction in Perth, Western Australia. There are three important insights proposed. Firstly, the results suggest that GBM is a competent model that can provide better prediction accuracy than the linear hedonic models completely, which is in line with the previous research in the literature. Then, loss function is a worthwhile aspect to investigate, especially when the least squares loss is not always

[9]The worse situation is that the prediction can't be provided, for example when complete case strategy is applied on linear model.

the best. Especially, in practice, there are diverse purposes of constructing an AVM with machine learning methods, not only for predicting the prices. Most importantly, GBM with missing node strategy and the least absolute deviation loss may be the best alternative for classic valuation models. The missing node strategy is designed for (tree-like) machine learning methods, it provides an unsophisticated implementation procedure and incredible prediction accuracy. This proves that machine learning methods indeed provide extra benefits, a better path for dealing missing values, other than precious forecasts. Additionally, missing values may not have to be arbitrarily or experientially dropped, but could be automatically and simultaneously handled within estimation stage, while no preparation, such as imputation, is required.

ACKNOWLEDGEMENT

The authors would like to thank the financial supports from the School of Accounting, Economics and Finance, Curtin University and the Business School, University of Aberdeen.

REFERENCES

Anglin, P. M. & Gencay, R. (1996), 'Semiparametric estimation of a hedonic price function', *Journal of Applied Econometrics* **11**, 633–648.

Breiman, L., Friedman, J., Stone, C. J. & Olshen, R. A. (1984), *Classification and Regression Trees*, Belmont, Calif : Wadsworth International Group.

Friedman, J. H. (2001), 'Greedy function approximation: a gradient boosting machine', *The Annals of Statistics* **29**(5), 1189–1232.

Friedman, J. H. (2002), 'Stochastic gradient boosting', *Computational Statistics & Data Analysis* **38**(4), 367–378.

Hastie, T., Tibshirani, R. & Friedman., J. (2009), *The Elements of Statistical Learning : Data Mining, Inference, and Prediction*, second edition edn, New York, NY : Springer New York.

Haupt, H., Schnurbus, J. & Tschernig, R. (2010), 'On nonparametric estimation of a hedonic price function', *Journal of Applied Econometrics* **25**, 894–901.

Hill, R. J. & Scholz, M. (2018), 'Can geospatial data improve house price indexes? A hedonic imputation approach with splines', *Review of Income and Wealth* **64**, 737–756.

Hinrichs, N., Kolbe, J. & Werwatz, A. (2021), 'Using shrinkage for data-driven automated valuation model specification – a case study from Berlin', *Journal of Property Research* **38**, 130–153.

James, G., Witten, D., Hastie, T. & Tibshirani, R. (2017), *An Introduction to Statistical Learning: With Applications in R*, New York, NY: Springer New York.

Kagie, M. & Van Wezel, M. (2007), 'Hedonic price models and indices based on boosting applied to the Dutch housing market', *Intelligent Systems in Accounting, Finance and Management* **15**(3/4), 85–106.

Knight, J. R., Sirmans, C. F., Gelfand, A. E. & Ghosh, S. K. (1998), 'Analyzing real estate data problems using the gibbs sampler', *Real Estate Economics* **26**(3), 469–492.

Kok, N., Koponen, E.-L. & Martínez-Barbosa, C. (2017), 'Big data in real estate? From manual appraisal to automated valuation', *Journal of Portfolio Management* **43**(6), 202–211.

Krause, A. & Lipscomb, C. A. (2016), 'The data preparation process in real estate: Guidance and review', *Journal of Real Estate Practice and Education* **19**, 15–42.

Little, R. J. A. & Rubin, D. B. (2002), *Statistical Analysis with Missing Data*, Wiley Series in Probability and Statistics, second edn, John Wiley & Sons, Hoboken NJ.

Mayer, M., Bourassa, S. C., Hoesli, M. & Scognamiglio, D. (2019), 'Estimation and updating methods for hedonic valuation', *Journal of European Real Estate Research* **12**, 134–150.

Parmeter, C. F., Henderson, D. J. & Kumbhakar, S. C. (2007), 'Nonparametric estimation of a hedonic price function', *Journal of Applied Econometrics* **22**, 695–699.

RICS (2021), Automated valuation models, Insight paper, Royal Institution of Chartered Surveyors, London.

Schulz, R. & Wersing, M. (2021), 'Automated valuation services: A case study for Aberdeen in Scotland', *Journal of Property Research* **38**(2), 154–172.

Schulz, R., Wersing, M. & Werwatz, A. (2014), 'Automated valuation modelling: A specification exercise', *Journal of Property Research* **31**, 131–153.

Steurer, M., Hill, R. J. & Pfeifer, N. (2021), 'Metrics for evaluating the performance of machine learning based automated valuation models', *Journal of Property Research* **38**, 99–129.