Land Use Policy 99 (2020) 104889

Contents lists available at ScienceDirect

Land Use Policy

journal homepage: www.elsevier.com/locate/landusepol

A mass appraisal assessment study using machine learning based on multiple regression and random forest

Seckin Yilmazer[a,b], Sultan Kocaman[a,*]

[a] Hacettepe University, Department of Geomatics Engineering, 06800, Beytepe, Ankara, Turkey
[b] General Directorate of Land Registry and Cadastre 06100, Cankaya, Ankara, Turkey

ARTICLE INFO

Keywords:
Real estate valuation
Mass appraisal
Multiple regression analysis
Random Forest

ABSTRACT

Mass appraisal is a complex matter because it depends on several categorical and continuously changing or constant parameters. In addition, development of new assessment approaches for mass appraisal of real estate properties in highly complex urban environments is desirable. The advancements in geospatial technologies and machine learning algorithms open up new horizons. For this reason, the purpose of the present study is to compare one conventional stepwise linear multiple regression (MRA) and one more automated machine learning approach, random forest (RF), for mass appraisal in an urban residential area where commercial properties are also available. A part of Mamak District, Ankara, Turkey is selected as the study area since the property values are diverse and representative. Additionally, the district has a complex and developing urban structure. The data employed in the study were compiled under a cadastral modernization project of General Directorate of the Land Registry and Cadastre of Turkey (GDLRC) and were based on the reports of licensed experts (~50 %), court reports (~20 %), field surveys, or a combined analysis of all. Consequently, the data used in the study has a high level of confidence. The initial set of parameters used in both methods reflect the most frequently observed characteristics of the real estate properties in the study area that are also effective on the values. The stepwise MRA required manual adjustments of the final parameter set by the expert, whereas RF eliminated unusable parameters fully automatically. The method performance was assessed by using a subset of the training data as a random test. According to the accuracy assessment results, the RF (Adjusted R^2 0*.734;* the total variance explained from the model) slightly outperforms the MRA (Adjusted R^2 *0.696*) where the optimal parameters were set by the human expert. Finally, the results exhibited are promising for quick assessment of mass appraisal and a comprehensive discussion is presented in the study.

1. Introduction

Mass appraisal of real estate properties is gaining importance due to the large share of real estate values in economic measures, which became one of the development indicators of countries. The real estate market has important economic dimensions in the gross national products depending on the population growth and increase in urban development. The real estate appraisal efforts introduce in the activity areas of several public and private institutions throughout the world and should be performed by employing unbiased, objective and scientific methods for the purpose of determining the actual state, rights and obligations of property values.

The appraisal efforts can be categorized as individual and mass appraisal based on the amount of properties evaluated during the appraisal process, and both approaches aim at estimating the market value as precise as possible (Williamson et al., 2010). Individual methods are most commonly used for the appraisal of only one or a small number of properties, since valuation of a large number of properties by individual methods is time consuming and requires more exhaustive and also more expensive processes. With mass appraisal, the number of appraised properties changed from the single units into large numbers (Clapp, 2003). According to the International Association of Assessing Officers (IAAO, 2013) '*mass appraisal is the process of valuing a group of properties as of a given date using common data, standardized methods, and statistical testing*'(Eckert et al., 1990). Depending on the recent gigantic advancements in machine learning (ML) techniques, mass appraisal methods have taken their share and evolved into more automated approaches. Several mass appraisal approaches such as; ML-based approaches (Antipov and Pokryshevskaya, 2012; Masías et al., 2016), computer assisted mass appraisal methods (McCluskey et al.,

* Corresponding author.
E-mail addresses: seckinyilmazer@hacettepe.edu.tr (S. Yilmazer), sultankocaman@hacettepe.edu.tr (S. Kocaman).

https://doi.org/10.1016/j.landusepol.2020.104889
Received 3 September 2019; Received in revised form 22 April 2020; Accepted 29 June 2020

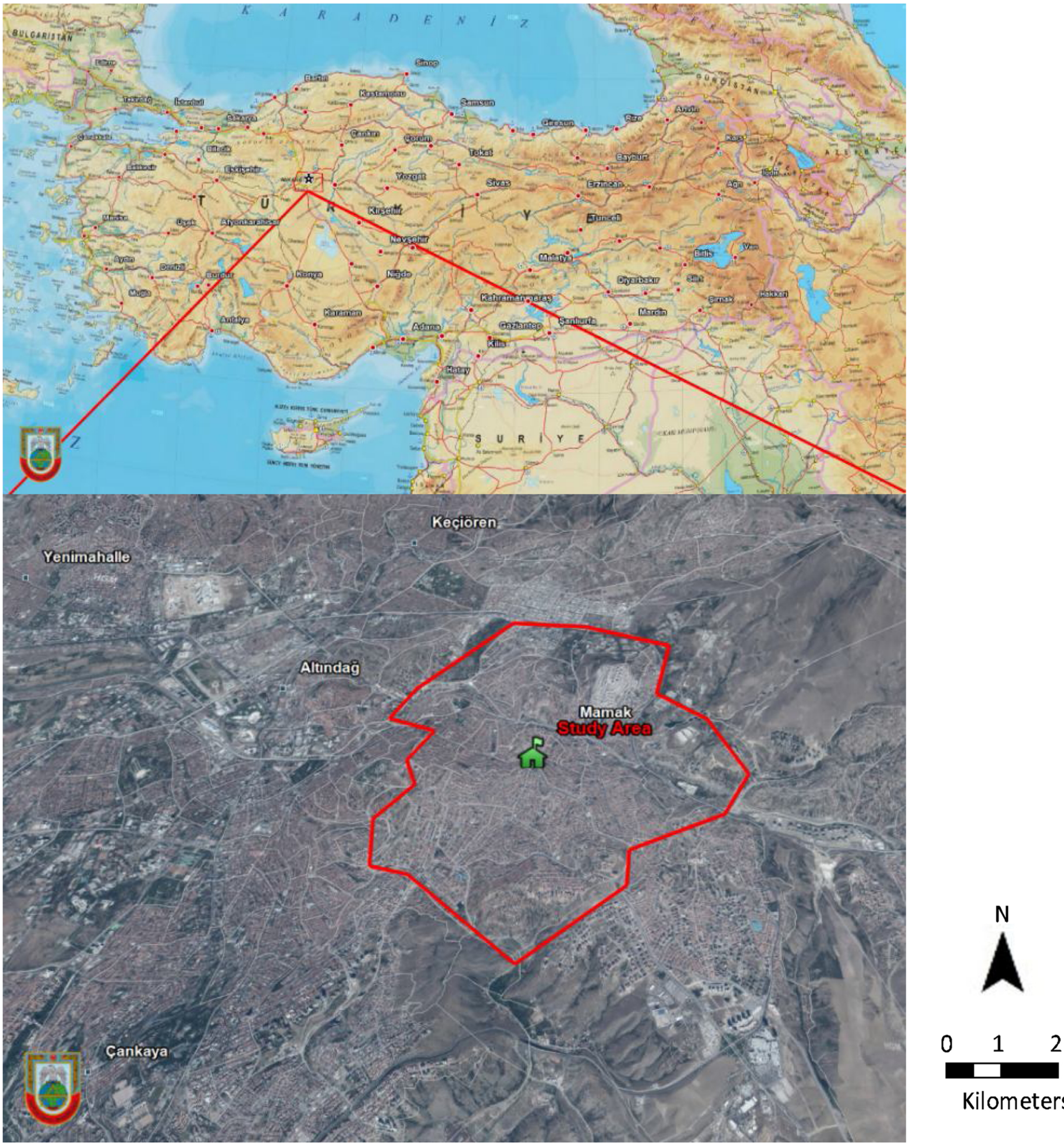

Fig. 1. The location and the boundary of the study area (red polygon shown below) with upper left coordinates: 39°59′4.75″N, 32°46′46.32″E; lower right coordinates: 39°52′45.33″N, 32°56′46.65″E (Image credit: GDM, 2018) (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

1997) and automated sales comparison approaches (Masías et al., 2016; Robbins, 2001) have been encountered in the literature. In recent years, statistics-based ML regression methods became a good alternative to the other more conventional appraisal approaches.

In developed countries, register databases or indices are used for the economic assessment of real estate markets, for land use planning and tax systems (Kara et al., 2017; Wilhelmsson, 2009). The property indices created using register databases are often updated with automated valuation models at certain intervals (Bagdonavicius and Deveikis, 2006) in order to serve to the real estate sector and the other related fields. In under developing countries, the real estate regulations and institutionalization of real estate markets are yet being established; and the time line indices, which are used to collect the property values at different times and essential for appropriate tax systems, have not been built yet. An appropriate tax system based on time line indices prevents from major tax losses. Mass appraisal studies have a crucial importance for the developing countries since they can stimulate the implementation and the establishment of the valuation processes and thus support the social and economic indicators. In addition, they may be used for the mass appraisal purposes in order to compute the past values with a confidence level with the help of data mining techniques and thus simulate the time line indices.

In a mass appraisal process, independent variables which affect the property values are determined and the model coefficients are computed with the help of a prediction method. Such a process requires accurate and reliable training (learning) data. Automatic appraisal systems perform the appraisal process at regular intervals with specified models or algorithms. The Multiple Regression Analysis (MRA) is one of the most common methods used in practice and for the mass appraisal studies in the literature and assessed with respect to other methods (i.e. Benjamin et al., 2004; Mark and Goldberg, 1988; Nghiep and Al, 2001; Zurada et al., 2011). The MRA method is based on close inspection of the model parameters by experts. Although ML-based techniques were often used for obtaining the house and commercial buildings prices separately (Chiarazzo et al., 2014), their use for the appraisal of all types of real estates has been increased (Antipov and Pokryshevskaya, 2012; Güneş and Yildiz, 2016; McCluskey et al., 2012; Trawiński et al., 2017). The decision tree (DT) and the random forest (RF) approaches are examples to ML-based methods that appear also in recent literature (Čeh et al., 2018; Park and Bae, 2015; Wager, 2014; Dimopoulos et al., 2018).

In fact, mass appraisal is a complex matter. The number of cases, the choice of independent variables, completeness and up-to-datedness of the input data, and selected prediction method affect the model outputs. Since the mass appraisal matters do not have any analytical solution, and it can be solved with a regression based empirical method or other more automated ML methods which are data driven. Such methodologies are open to development depending on the amount and the quality of data, and the methodological advancements. In addition, depending on the level of user intervention in variable selection and parameter optimization processes, the quality of the experts/employee may also have an influence. For this reason, increase in such studies may serve to understand the nature of the mass appraisal prediction problem. Considering this fact, the present study aims at evaluating two ML methods with different levels of automation for mass appraisal purpose in a part of Ankara, Turkey; and discussing their suitability, accuracy and reliability. This may be an important contribution to the literature because the data was compiled in a pilot appraisal project by GDLRC of Turkey with a team of experts via field work, and extensive analyses of court reports and other documents. It is well-known that the reliability of a statistical or ML prediction model depends on the quality and the number of the data employed during the model development. The selected ML methods are suitable to solve such problem. 37 features effective on the real estate values in the study area were selected out of a total of 96 variables identified by the GDLRC team to be used as initial list of independent variables. The number of independent variables is relatively large and is probably the first example in international literature.

2. Study area and data

Mamak is one of the major districts in Ankara, Turkey and has complex urban characteristics since it is a rapidly developing area. Mamak is surrounded by three other major districts (Altindag in the North, Elmadag in the East, Cankaya and Elmadag in the South, Cankaya and Altindag in the West) (Fig. 1). Cankaya district is an economically and socially developed area of Ankara, and the parts of Mamak which are close to Cankaya, are affected by this situation. The other parts of Mamak District are still developing. The land use land cover types in the area are agricultural fields (19,8 %), meadow and pasture (40,2 %), forest (6 %), and others (34 %) as of year 2019 (Mamak Municipality, 2019). The study area contains different socio-economical parts of Mamak and was selected to demonstrate the usability of the selected mathematical models for explaining such large differences.

According to the Land Registry and Cadastre Information System (TAKBIS) of GDLRC of Turkey, there are approximately 56.000 cadastral parcels and more than 11.000 individual units (e.g. condominiums, commercial properties, etc.) in Mamak. In this study, comprehensive information on the properties including the values were derived from the mass appraisal report obtained in a pilot project (GDLRC, 2014) and from TAKBIS. The pilot project team was composed of 6 field experts, 3 evaluation experts, and 8 assistance staff. 2 international experts also worked as external consultants to analyze the reliability and accuracy of the collected data. In addition to the data collected in the pilot project, TAKBIS real-time and historical data, licensed valuation company reports, court reports and municipal tender sales were also collected; and all of them were evaluated together in this study. On the other hand, the documents and digital archives stored in the geographical information systems (GIS) of Mamak Municipality were analyzed and further information was obtained from the interviews with municipality officers for verification of the dataset.

The study dataset comprises a total of 96 independent variables and 1200 cases and contains the residential transactions in the period of years 2013–2018. The 96 variables were defined in the pilot project via extensive field work to reflect the main characteristics of the real estate properties in all project districts and comprehend the most likely features that affect the price, such as proximities to the important cultural, recreational and health facilities, quality of the infrastructure and construction, etc. After the comprehensive data collection and analysis stage, the independent variables could be verified and the real property values, which are mostly based on the sale prices, could be separated from incompatible and corrupted data.

The property values were determined based on the reports of licensed experts (~50 %), court reports (~20 %), field surveys, or a combined analysis of all; since the declared property values, which are also used as base for tax calculations, may not represent the actual values. Although the minimum values to be declared are defined by municipalities; they are usually well below the real values. The common exceptions to this situation are the properties purchased using bank loans, where licensed experts assess the real value; and the property sales made by courts and municipalities, where the money transfers were made transparently. Therefore, the pilot project dataset is quite reliable and considered to reflect the correct values. The dataset was used to construct and validate the models and to compare their prediction performances. Since the values have temporal differences, i.e. include transfer prices between 2013–2018, they were updated using the Turkish Republic Central Bank (TCMB of Turkey) Residential Property Price Index (RPPI) for Ankara, and represent the values as of year 2018. The ratio of RPPI values of March 2013 and March 2018 (also shown in Fig. 2) is 1.57 for Ankara (data source: Electronic Data Delivery System, EVDS, service of Central Bank of Turkish Republic). Thus, the temporal differences (up to five years) in the dataset could be eliminated by updating the values with multiplying the RPPI ratio ($RPPI_{March\_2018}$ /$RPPI_{transfer\_date}$).

Out of the 96 variables, a total of 37 independent variables, which are significant for the study area were selected by the expert and utilized in the employed methods together with one dependent (target) variable, which is value. The initial list of independent variables is provided in Table A1 in Appendix. The collected data were analyzed to detect missing values and unusable ones. Out of 1200, a total of 38 cases were eliminated manually since they were incomplete. Descriptive statistics of all samples obtained from the employed dataset are provided in Table A2 in Appendix. The independent variables can

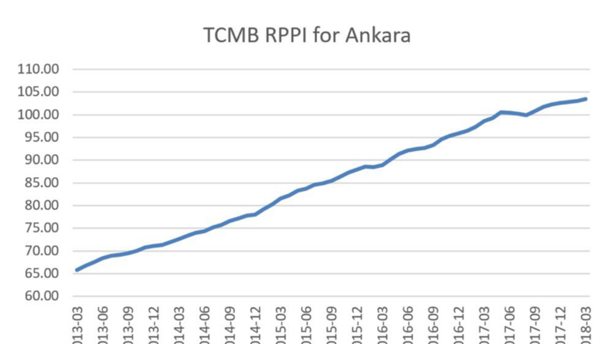

Fig. 2. TCMB RPPI for Ankara between March 2013-March 2018 (data source: evds2.tcmb.gov.tr).

Table 1
Numbers and percentages of training and test cases obtained from the SPSS Tool.

	Frequency	Percent
Testing Sample	230	20
Training Sample	932	80
Total	1162	100,0

be categorized as details on the ground, the zonal information, and information on the individual units. The target variable (*value*) reflects the property transfer price.

The training and test datasets were determined randomly using IBM SPSS statistics tool Released in 2015 (IBM Corp, A., 1968) for the MRA and with Matlab R2015 (MATLAB Statistics Toolbox, 2015) for the RF methods. The ratio of the training/test data cases is 80/20, which is selected empirically based on a literature check. Although this ratio can be found in some studies (e.g. Kouzani et al., 2007), comparable ratio values such as 75/25 or 70/30 are also available. The numbers and percentages of the training and test cases obtained from the SPSS Tool are given in Table 1.

3. Methodology

In order to obtain sufficient accuracy in a mass appraisal process, the determination of the right independent variables and selection of the appropriate prediction model should be performed properly. The method selection depends on several parameters such as the purpose of appraisal (expropriation, private sales, court enforcements, rentals, etc.), the number of properties to be valued, and the property types (residential, commercial, rental, etc.). There are mainly two stages such as model specification (design) and calibration in a ML-based mass appraisal process. The design of a model is based on the appraisal theory and the market analysis, and includes the selection of the variables considered for the model and defining the relationships among the variables and also with the property value. Model calibration (i.e. training) is the process of calculating the coefficients of the variables. Two ML methods, the stepwise MRA which requires more expert intervention; and the RF which is more automated, were employed for the mass appraisal purposes in this study. The methodological details are provided below.

3.1. Multiple regression analysis (stepwise)

The MRA, which is the most commonly used method in mass appraisal applications, was performed through the Stepwise Regression Method. MRA is the conventional option for mass housing price appraisal studies for less complex datasets (Chiarazzo et al., 2014). Many studies conducted to compare advanced ML methods with conventional MRA, calculate the error percentages between the models, and analyze them with a package of market samples where the sale price is a known factor (Mark and Goldberg, 1988). In the study of Mora-Esperanza (2004), the advanced ML methods showed an error rate of 5–10 % in real estate value prediction, where MRA results in a higher error rates of 10–15 %. Demetriou (2016) estimated the relationships between target and explanatory variables using the MRA. Whilst the standardized residuals indicate that estimation by the stepwise regression was pretty good in terms of fitting to a straight line, they do not inherently mean that the estimation model have enough capability to make a decision with the initial list of variables. Although the ease of use, the MRA method has some deficiencies (IAAO, 2013). In particular, the MRA method ignores spatial autocorrelation and spatial heterogeneity (Demetriou, 2016), although some models have options to give ranking for location quality (e.g. 1–5 from very bad to very good) as input. To overcome these shortcomings and to obtain good results, expert opinion is needed.

The MRA equation relates Y to a function of X and β, as follows:

$$Y = Y \Rightarrow f(x, \beta) \Rightarrow \Rightarrow \beta 0 + \beta Xij \tag{1}$$

where; Y is the target or dependent variable, X*i* is the independent variable(s) known or can obtained, and $\beta_{i\text{-}n}$ are constant coefficients obtained from regression.

The variables in the initial list were analyzed for multicollinearity. The following stages were performed during the mass appraisal with the MRA method:

- Variable determination for training (37 independent variables);
- Outlier and incomplete case elimination (38 cases were removed from the set);
- Multicollinearity check through correlation analysis between zonal parameters (e.g. distances to center, health, shopping and education facilities, etc.) and the removal of highly-correlated (i.e. correlation coefficient > 0.5) variables from the model;
- Random sample selection (80 % of the whole samples);
- Stepwise run of the linear model and elimination of the unstable or insignificant variables from the model after each step; and
- Model validation using the test samples (20 %).

3.2. Random forest

The RF is ensemble learning method that is the modified from the DT (Breiman, 2001; Ho, 1995) and requires supervised training. Therefore, it is necessary to have a training dataset consisting of response and explanatory variables. The DT is constructed by using the grown tree algorithm by recursively dividing the samples based on the importance and effectiveness of the variables in the training data set (Fan et al., 2006). DTs are constructed beginning with the root node of the tree and proceeding down to internal nodes and the last part of leaves (Quinlan, 1986). Although the DT is an alternative estimation method to linear MRA conventionally (Breiman et al., 1984), it is not preferred here since the growing up process usually results in exceeding number of trees, which can cause weak learning and overfitting (Liaw and Wiener, 2002). The basis of the method is the random selection of a sample selected from training variables and followed by training of the system with this subset data set to construct prediction model (Barandiaran, 1998). The resulted estimation of the ensemble is constructed from the discrete decisions by majority voting- classification or averaging-regression (Čeh et al., 2018). In this study, multiple tree estimators were created with the RF. These estimators have magnified according to the values of a vector formed by randomly selected samples from the training set (Breiman, 2001). As a result, using these predictors, the estimations appraised by taking the average of the predictions from individual trees. The RF model built for the study is demonstrated in Fig. 3.

The initial set of variables (37 independent and one dependent) and the same cases obtained after outlier and incomplete data elimination were employed in RF. The stages of mass appraisal with the RF regression model are:

- Random bootstrapping of samples from the whole data (at the ratio of 80 %)
- Running RF (Specified number of trees "n-trees", Number of samples calculate average values of estimators, "methods = regression")
- Validation of the results using the test samples (at the ratio of 20 %)

3.3. Performance check

Several statistical indicators and notations were computed and used for mass appraisal accuracy. Root mean square error (RMSE) is a measure of the average deviation of the estimates from the observed values (see Eq. (2)). R^2 is often used to measure the rate of variation of

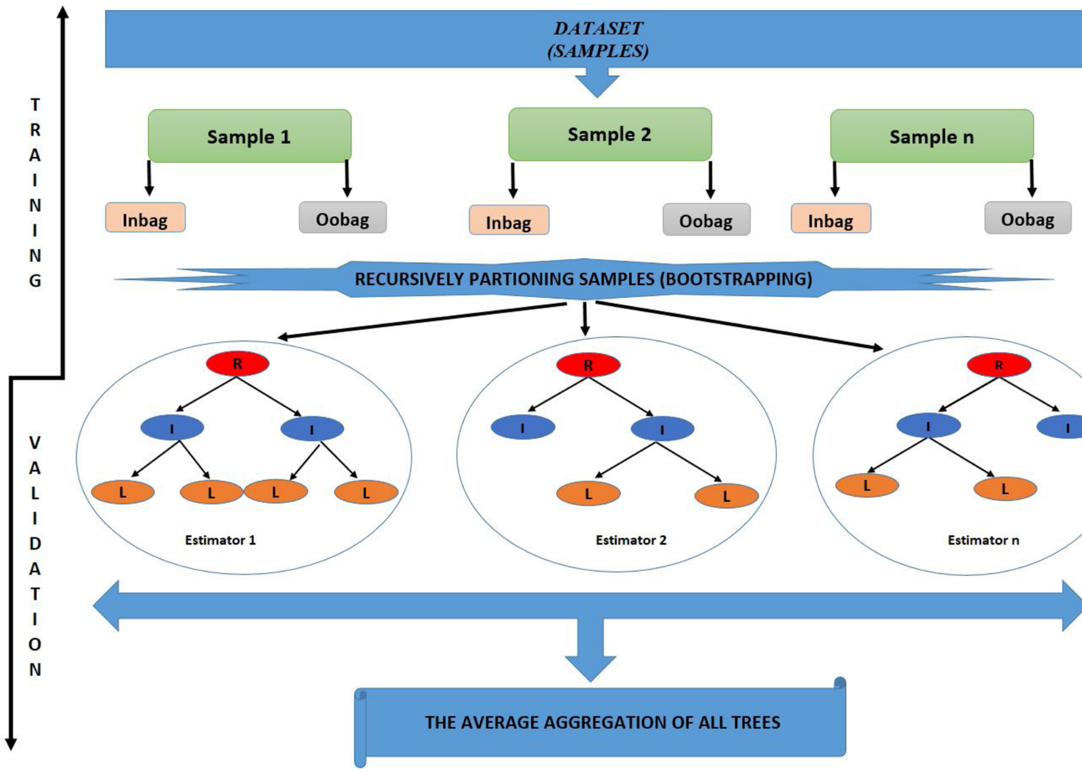

Fig. 3. The Random Forest Model employed in the study (modified after Li et al. (2017).

Table 2
Summary of the Stepwise MRA model results.

Model (step) number	R	R^2	Adjusted R^2	Std. Error of the Estimate
1	,777[a]	0.604	0.604	22339.34
2	,822[b]	0.676	0.675	20217.58
3	,835[c]	0.698	0.697	19538.61
4	,845[d]	0.713	0.712	19041.73
5	,849[e]	0.722	0.720	18776.02
6	,853[f]	0.728	0.726	18572.61
7	,856[g]	0.732	0.730	18426.66
8	,858[h]	0.737	0.735	18280.44
9	,860[i]	0.740	0.738	18168.75
10	,861[j]	0.742	0.739	18123.43
11	,862[k]	0.744	0.740	18077.52

dependent variables interpreted by the specified models (Fan et al., 2006). Furthermore, it is used to evaluate the model accuracy in regression methods. On the other hand, the Adjusted R^2 (Adj R^2) indicates that the percentage of variation explained by only significant variables that actually affect the dependent variable (given in Eq. (3)).

$$RMSE = \sqrt{n^{-1}\sum\nolimits_{1}^{n} (yi-\hat{y})^2} \quad (2)$$

$$R^2 = 1 - \left(\frac{MSE}{\sigma y^2}\right) \text{Adj } R^2 = 1 - \left(\frac{(1 - R^2)(p - 1)}{(p - N - 1)}\right) \quad (3)$$

where;

N is the number of cases, y_i is the i^{th} case, p is number of variables, $\hat{y}$ is the appraised mean, MSE is the mean square error.

RMSE, R^2 and Adj R^2 indicators are basic to assessing the fit of regression model's. On the other hand, Coefficient of Dispersion (COD), weighted mean (Wt.$\bar{R}$), and Price-Related Differential (PRD) are the ratio statistic indicators which can be used especially for mass appraisal accuracy assessments (IAAO, 2013). The COD is a widely used measure of appraisal uniformity. It expresses the average deviation of the ratios from the median as percentage. The Wt.$\bar{R}$ is the ratio of the average appraised value to the average sales prices and is also called aggregated ratio. The PRD measures the regressivity or progressivity of the

Table 3
Coefficients obtained from the Stepwise Regression Model.

Unstandardized Coefficients		Standardized Coefficients				95.0 % Confidence Interval for B	
Step11	B	Std. Error	Beta	T	Sig.	Lower Bound	Upper Bound
(Constant)	4955.90	3252.10		1.52	0.128	−1426.48	11338.28
SQ-METR	734.63	27.37	0.563	26.84	0	680.92	788.34
ONWH-FL	4207.75	363.00	0.234	11.59	0	3495.33	4920.16
NM-PARK	97.36	17.64	0.109	5.51	0	62.73	131.98
DIS-UNI	−1.65	0.32	−0.120	−5.22	0	−2.27	−1.02
ELVTOR	7607.48	1639.10	0.086	4.64	0	4390.68	10824.28
R-OPSQM	494.02	116.24	0.088	4.25	0	265.90	722.14
OCC-PER	6613.82	1594.45	0.070	4.15	0	3484.65	9742.98
STWIDTH	411.00	93.82	0.075	4.38	0	226.89	595.12
D-SLN	4997.35	1489.61	0.061	3.36	0.001	2073.92	7920.77
DIS-MRK	−1.32	0.56	−0.05	−2.37	0.018	−2.42	−0.23

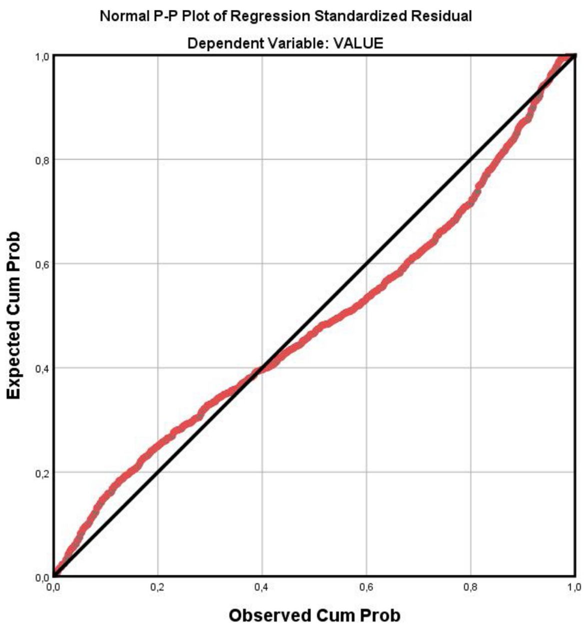

Fig. 4. The normal P-P Plot of standardized residuals obtained from the MRA method.

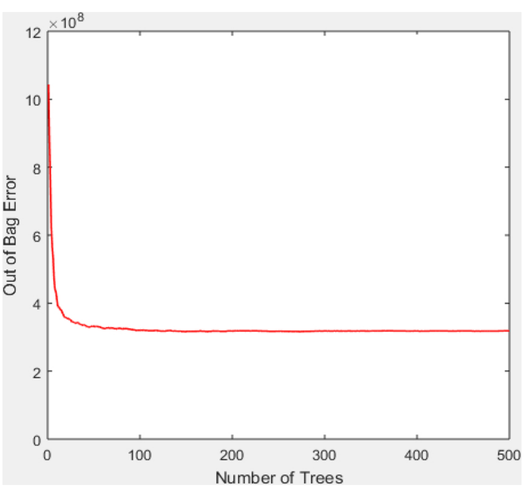

Fig. 5. Out of Bag Error Graphic obtained from the RF method.

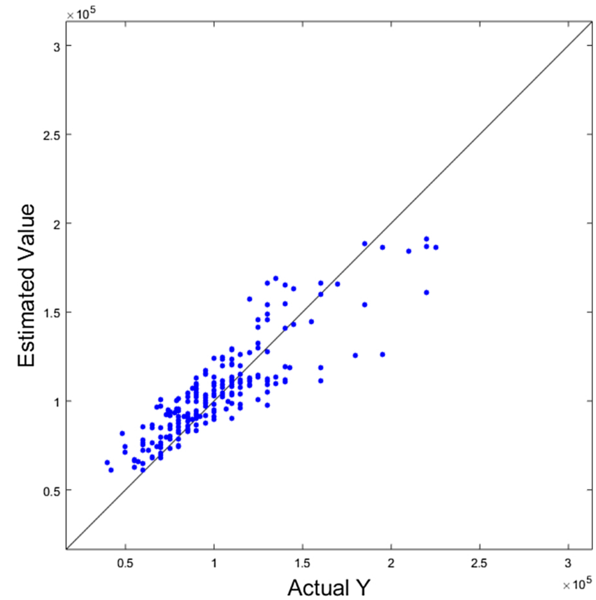

Fig. 6. Comparison of the given prices and estimated values (Y) in RF for the test cases.

assessments. Regressive appraisals occur when high-valued real estate properties are under-appraised relative to low-valued ones. Progressive appraisals occur when the opposite pattern occurs (IAAO, 2013). The formulations of COD, Wt.$\bar{R}$ and PRD are provided in Eqs. (4)–(7) and the confidence intervals of the indicators are provided in Table A3 in Appendix.

$$R_i = \frac{A_i}{S_i},\ (i = 1, \ldots, n). \tag{4}$$

$$COD = \%100\,x\,\frac{\frac{(\sum_1^n |R_i - \check{R}|)}{n}}{\check{R}} \tag{5}$$

$$Wt.\ \bar{R} = \frac{\sum_1^n A_i}{\sum_1^n S_i} \tag{6}$$

$$PRD = \frac{\bar{R}}{Wt.\bar{R}} \tag{7}$$

where;

n Number of observations

ANumerator of the *i-th* ratio (*i = 1, …, n*). (Appraisal Value)

S_i Denominator of the *i-th* ratio (*i = 1, …, n*). (Sale Price)

R_i The *i-th* ratio (i = 1, …, n). (Appraisal Ratio)

$\check{R}$Median

$\bar{R}$ Mean

Wt.$\bar{R}$Weighted Mean

4. Results and discussion

The results obtained from both methods and their comparisons are given in the following sub-headings. The discussions on the results are also provided in these sections.

4.1. Results of the stepwise MRA method

The correlation analysis results for the multicollinearity check are summarized in Table A4 in Appendix. Parameters having a correlation coefficient larger than 0.5 are considered as highly correlated. As can be seen in Table A4, the *distance to university* variable (DIS-UNI) is highly correlated with the *distance to center* (DIS-CEN), *distance to cultural activity area* (DIS-CUL) and *distance to hospital* (DIS-HOS) variables. On the other hand, the *distance to supermarket* (DIS-MAR) was highly correlated with the *distance to mall* (DIS-MAL), the *distance to recreational area* (DIS-REC), and the *distance to recycling area* (DIS-GAR). In order to avoid multicollinearity, these parameters (DIS-CEN, DIS-CUL, DIS-HOS, DIS-MAL, DIS-REC, DIS-GAR) were removed from the model.

The remaining 31 independent variables were employed in the initial MRA model. The variables that are unstable or insignificant were removed from the model in a Stepwise Regression Analysis. The resulting model output summaries and the final model coefficients are provided in Tables 2 and 3, respectively.

After analyzing the model summaries (Table 2) and coefficients (Table 3), although 11 steps were provided in the model summary table, it was observed that the DIS-BUS variable has very low effect and the results of Steps 10 and 11 are comparable. Therefore, parameters of the Step 10 were selected as the training model. When the appraisals are interpreted, the fixed value in the model is found ca. 4.956,00 TRY (Turkish Liras, I USD = 6 TRY by the end of 2019). The property value increase per gross square meter is 73,400 TRY. Being located on a higher floor for an individual unit has a positive contribution to the value. As the land share increases, the value of an individual unit also

Table 4
Accuracy assessment results for the MRA and the RF methods.

Method	Median	Mean	Wt.R̄	SD	COD	PRD	RMSE	R^2	Adj R^2
MRA	1.011	1.037	1.009	0.1653	12.742	1.027	17975	0.710	0.696
RF	1.048	1.061	1.029	0.1317	11.438	1.030	16486	0.749	0.734

Table 5
Statistical analysis results of the values (prices) obtained from both models.

Statistical Results	Residential Number	Total Value	Arithmetic Mean Values	Median Values
Stepwise Regression	1178	23.570.380,00	102.479,91	99.368,29
Random Forest	1178	24.056.571,00	104.593,79	102.198,50

Table 6
Annual real estate property tax rates inside and outside metropolitan areas in Turkey.

Property Type	Outside Metropolitan Areas	In Metropolitan Areas
Dwellings-Houses	0.1 %	0.2 %
Other Buildings	0.2 %	0.4 %
Agricultural Lands	0.1 %	0.2 %
Zoning Lands	0.3 %	0.6 %

increases. The residential values decrease when the distance to the university campuses (e.g. Hacettepe Medical School) increases. As mentioned before, the DIS-CEN, DIS-CUL, and DIS-HOS parameters were removed due to high correlation with the DIS-UNI parameter; so this parameter reflects their effects on the results to a certain extent as a matter of course. If the building has an elevator, the value increases in the order of 7.607,00 TRY. The validation results provide an RMSE 17.975 TRY; the Total Variance Explained Statistically (R^2) 0,710 and Adjusted R^2 0,696.

The normal probability plot (P-P) is a graphical technique to assess the distribution and to identify the divergence of the residuals from normal distribution. The P-P Plot not only assesses the normality, but it can also be used to compare any two distributions to determine how well they match each other. When the standardized residuals distribution is compared with the normal distribution here (Fig. 4), they seem to be compatible. In other words, as shown in the P-P Plot given in Fig. 4, a normal distribution of standardized residuals is indicated since they are located close to the diagonal line. A dataset that does not have normal distribution would not exhibit scattering along the diagonal line.

4.2. The RF results

The initial list of variables given in Table A1 was employed in this model. 1162 cases were split as for training (80 %) and validation (20 %) randomly by the RF method. The out of bag error (OOB) which is an indicator and a major tool of measuring the estimation error of RF using bagging (Breiman, 1996) to sub-sample data cases used for training, is calculated to determine the number of trees to be used in the model. According to Oshiro et al. (2012) sometimes the number of trees have no significance on performance. Fig. 5 shows the OOB error graph obtained after the RF was performed. The tree value where the graph becomes parallel to the y axis appears to be approximately 100. For this reason, 100 was selected as the number of trees since from this number on, the minimum OOB error is obtained. If the number of trees is chosen greater than 100, the OOB score does not change significantly, but the operational expense will increase.

The appraisal value (estimated value) and the actual prices (sale prices) for the test data are shown in Fig. 6. It can be seen that the given and estimated values are gathered around the diagonal in general, but divergence occur for the properties with higher prices. The validation results (test samples) provide an RMSE 16.486, Total Variance Explained Statistically (R^2) 0,749, and Adjusted R^2 0,734.

4.3. Comparison and discussion

The comparison summary of the statistical assessment results obtained from both methods are provided in Table 4. From the results, it can be observed that the RF model explaining rate and the average deviation are slightly better than the MRA results. Both methods achieved very close PRD scores, suggesting both methods yielded consistent indications of vertical inequity. On the other hand, both of two methods achieved accepted levels of COD and PRD scores according to the IAOO standards (see Table A3). The RF model can determine the suitable variables fully automatically with a comparable level of accuracy of the human expert. The main reasons for the RF method to obtain the better results are: i) the method can solve nonlinear interactions unlike MRA; ii) it obtains the best tree structure by the boosting method; and iii) it can eliminate over-fitting errors. A statistical analysis and comparison of the outputs from both models is provided in Table 5.

The annual real estate property tax rates in Turkey vary depending on the type of property and the city zoning plan (Table 6). In addition, the property ownership transfer tax rate is 2% both for the seller and the buyer (4 % in total) and the taxes are charged separately. Based on the appraisal results, a simple analysis was performed to find out the ratio of tax losses caused after property transfers and provided in Table 7. From the Table, it can be seen that the ratio of the paid taxes and the taxes losses was almost 1:1. For the transfer tax calculations, the values declared jointly by the seller and the buyer are used, although this value also cannot be less than the municipality defined minimum as mentioned before. The minimum declaration value of the property is updated by municipality in case of a higher sale price upon the buyer's declaration, which is obligatory after the ownership transfer. Thus, the municipal tax authority can use the new declaration

Table 7
Comparison of the declared property values for the property transfers and the estimated values for the test samples in the study area (values are given in TRY).

Number of cases (test)	Sum of the declared values for sale price	Sum of the paid taxes[a]	Sum of the estimated values with RF	Taxes based on the RF results	Difference btw. the paid and actual taxes	Ratio btw. the paid and actual taxes
230		496.960,00		962.262,84	4657.302,84	1,94

[a] 4% of the declared values.

as the municipal property value for annual tax calculations. Based on the outcomes of this study, it can be seen that neither the sale prices declared to the land registry offices nor the municipal registered property values reflect the actual sale prices apart from few exceptions (e.g. bank loans as mentioned previously). Since the tax losses are currently in the focus of the Turkish State, such analyses may be useful for policy updates.

5. Conclusions

The study presented here is an interesting example of the comparison of the ML methods for mass appraisal of real estate properties and to understand the details of ML model specification and model calibration for this purpose. The selected area, Mamak District of Ankara, Turkey, is a rapidly developing urban area. Two ML-based methods, the MRA and RF, with different levels of expert interactions were evaluated for the mass appraisal purposes. The initial set of independent variables were selected carefully out of a larger set based on the expected influence on the values in the district.

The RF algorithm can find both linear and non-linear interactions between the independent variables and the target variable. Furthermore, partitioning the tree can also help to find minimized interactions of variables as can be seen in the model summaries table. The RF algorithm is capable of explaining the dependent variable accurately and produces slightly better results than the MRA method for mass appraisal.

Although the initial set of parameters is the same for both models, the MRA required parameter set optimization, because linear regression method could not overcome the multicollinearity problem and the nonlinear interactions between explanatory variables; whereas the RF omitted the unusable variables fully-automatically and is capable of working with complex datasets. The performances of the methods were analyzed by using a subset of the training data as test.

The results obtained with the residential appraisal approach can be used for the correction of declared tax values. In the actual real estate tax system of Turkey, municipalities calculate the property tax values based on taxpayers' declarations. The case study shows large differences between the declared and the estimated values and the State has lost ca. 50 % of the tax income for the test data.

It is evident that the generalization capacity and prediction accuracy of the results obtained from MRA and RF approaches depend on not only quality of the dataset but also number of cases. Due to this reason, the models constructed in the study will open to development depending on increase in cases. In addition, depending on the region, the choice of independent variables may also differ. Due to lack of analytical solution of mass appraisal matter, construction of data driven models for such problem is mandatory and increase in the number of similar studies can be expected in the future.

CRediT authorship contribution statement

Seckin Yilmazer: Conceptualization, Formal analysis, Methodology, Data curation, Investigation, Software, Writing - original draft. **Sultan Kocaman:** Conceptualization, Formal analysis, Methodology, Supervision, Validation, Writing - review & editing.

Declaration of Competing Interest

The authors declare no competing interests.

Acknowledgements

The authors sincerely thank to Pilot Appraisal Project Team of GDLRC for collecting and providing the data.

Appendix A

Table A1
The initial list of variables prepared by the expert.

Nmb	ID	Value Type or Measurement Unit	Variable Descriptive	Variable Type In/Out
1	IU-ID	Integer Number	'IndividualUnit_TitleID'	Desc. - Out
2	BALC	Square Meter	'Balcony'	Indep. - In
3	BL	Boolean	'BuildingLicence'	Indep. - In
4	D-SLN	Boolean	'DependentSaloonOrNot'	Indep. - In
5	D-STTS	Boolean	'DevelopingStatus'	Indep. - In
6	DIS-BAZ	Meter	'DistanceToBazaar'	Indep. - In
7	DIS-BUS	Meter	'DistanceToBusStop'	Indep. - In
8	DIS-CEN	Meter	'DistanceToCenter'	Indep. - In
9	DIS-COL	Meter	'DistanceToCollege'	Indep. - In
10	DIS-CUL	Meter	'DistanceToCulturalArea'	Indep. - In
11	DIS-GAR	Meter	DistanceToGarbageColl'	Indep. - In
12	DIS-HOS	Meter	'DistanceToHospital'	Indep. - In
13	DIS-ROD	Meter	'DistanceToMainRoad'	Indep. - In
14	DIS-MAL	Meter	'DistanceToMall'	Indep. - In
15	DIS-MRK	Meter	'DistanceToMarket'	Indep. - In
16	DIS-MOS	Meter	'DistanceToMosque'	Indep. - In
17	DIS-SCHL	Meter	'DistanceToPimarySchool'	Indep. - In
18	DIS-REC	Meter	'DistanceToRecreationalArea'	Indep. - In
19	DIS-UNI	Meter	'DistanceToUniversity'	Indep. - In
20	ELVTOR	Boolean	'Elevator'	Indep. - In

(*continued on next page*)

Table A1 (*continued*)

Nmb	ID	Value Type or Measurement Unit	Variable Descriptive	Variable Type In/Out
21	FR-GARD	Meter	'FrontGarden'	Indep. - In
22	ISCOPLX	Boolean	'IsIntheComplexStyleOrNot'	Indep. - In
23	ROADFR	Boolean	'MainRoadFrontageOrNot'	Indep. - In
24	H-MAX	Meter	'MaxConstHeightOnAPlot'	Indep. - In
25	NM-BATH	Meter	'NumberOfBaths'	Indep. - In
26	NM-FACA	Integer Number	'NumberOfFacade'	Indep. - In
27	NM-FLOR	Integer Number	'NumberOfFloors'	Indep. - In
28	NM-PARK	Integer Number	'NumberOfParkingArea'	Indep. - In
29	NM-ROOM	Integer Number	'NumberOfRooms'	Indep. - In
30	OCC-PER	Boolean	'OccupancyPermit'	Indep. - In
31	ONWH-FL	Integer Number	'OnWhichFloor'	Indep. - In
32	CL-PARK	Boolean	'ParkingOrNot'	Indep. - In
33	R-CLSQM	Square Meter	'ResExtraClosedArea'	Indep. - In
34	R-SQM	Square Meter	'ResidenceGrossArea'	Indep. - In
35	R-OPSQM	Square Meter	'ResidenceGrossOpenSpace'	Indep. - In
36	SGARDN	Meter	'SideGardenWidth'	Indep. - In
37	STWIDTH	Meter	'StreetWidth'	Indep. - In
38	SPOOL	Boolean	'SwimmingPool'	Indep. - In
39	VALUE	TRY; USD 6 TRY ≈ 1 USD 7 TRY ≈ 1 EURO	'Price' - Value	Dep. - In

Table A2
Descriptive statistics of all dependent and independent variables given in Table A.1.

Variables	Valid	Miss	Mean	Median	Std. Deviation	Min	Max
BALC	1162	0	0.91	1	0.28	0	1
BL	1162	0	0.96	1	0.19	0	1
D-SLN	1162	0	0.75	1	0.43	0	1
D-STTS	1162	0	0.49	0	0.5	0	1
DIS-BAZ	1162	0	905.76	810.5	552.91	1	2743
DIS-BUS	1162	0	255.62	243	141.25	1	776
DIS-CEN	1162	0	3229.08	2902.5	1767.02	1	8949
DIS-COL	1162	0	844.83	736	632.37	1	4488
DIS-CUL	1162	0	4574.05	4204	2508.25	1	12128
DIS-GAR	1162	0	5992.99	5957.5	2578.88	1	13263
DIS-HOS	1162	0	3935.12	3349	2551.93	1	11619
DIS-ROD	1162	0	160.39	141.5	149.6	1	1118
DIS-MAL	1162	0	2741.71	2195.5	1915.38	1	8754
DIS-MRK	1162	0	997.7	581	1440.02	1	6999
DIS-MOS	1162	0	286.89	273	162.54	1	1091
DIS-SCHL	1162	0	481.85	433.5	309.03	1	1864
DIS-REC	1162	0	7412.79	7837	2409.08	1	11441
DIS-UNI	1162	0	3806.93	3176	2569.38	1	11949
ELVTOR	1162	0	0.21	0	0.41	0	1
FR-GARD	1162	0	5.69	5	1.89	0	20
COMPLX	1162	0	0.05	0	0.22	0	1
ROADFR	1162	0	0.24	0	0.43	0	1
H-MAX	1162	0	11.64	12	2.22	9	33
NM-BATH	1162	0	1.9	2	0.58	1	6
NM-FACA	1162	0	2.1	2	0.56	1	3
NM-FLOR	1162	0	6.43	6	3.17	1	36
NM-PARK	1162	0	16.86	8	39.91	0	321
NM-ROOM	1162	0	3.92	4	0.67	2	8
OCC-PER	1162	0	0.83	1	0.37	0	1
ONWH-FL	1162	0	1.36	1	2.02	−5	14
CL-PARK	1162	0	0.94	1	0.24	0	1
R-CLSQM	1162	0	49.14	48	14.73	14	130
SQ-METR	1162	0	106	102	26.78	43	264
R-OPSQM	1162	0	8.57	8	6.08	0	86
SGARDN	1162	0	3.17	3	1.23	1	15
STWIDTH	1162	0	13.37	12	6.4	0	50
SPOOL	1162	0	−0.58	−1	0.81	−1	1
VALUE	1162	0	102.928,27	100.000	35.087,39	30.000	340.000

Table A3
Ratio Study Uniformity Standards Indicating Acceptable General Quality (IAOO, 2013).

General Property Class	Jurisdiction Size/Profile/Market Activity	COD Range
Residential improved (single family dwellings, condominiums, manuf. housing, 2 − 4 family units)	Large to mid-sized jurisdictions/older & newer properties/less active markets	5.0–15.0

CODs lower than 5.0 may indicate sales chasing or non-representative samples.
PRD's for each type of property should be between 0.98 and 1.03 to demonstrate vertical equity.
Definitions of indicators;

- The Standard Deviation (SD) is the average distance of the ratios from the ratio mean.
- The Coefficient of Dispersion (COD) is a widely used measure of appraisal uniformity. It expresses as a percentage of the average deviation of the ratios from the median.
- The price-related differential (PRD) measures the regressivity or progressivity of the assessments. Regressive appraisals occur when high-valued real estates are under-appraised relative to low-valued ones. Progressive appraisals occur when the opposite pattern occurs.

Table A4
Correlation matrix of the initial list of variables obtained from MRA.

Correlations Coefficients														
Meter	DIS-BAZ	DIS-BUS	DIS-CEN	DIS-COL	DIS-CUL	DIS-GAR	DIS-HOS	DIS-ROAD	DIS-MAL	DIS-MRK	DIS-MOS	DIS-SCHL	DIS-REC	DIS-UNI
DIS-BAZ	1.00	0.27	0.07	0.33	−0.08	0.38	−0.02	0.32	0.32	0.02	0.20	0.04	0.25	0.06
DIS-BUS	0.27	1.00	0.13	0.13	0.01	0.34	0.02	0.29	0.28	0.16	0.22	0.05	0.35	0.05
DIS-CEN	0.07	0.13	1.00	0.47	**0.84**	0.40	**0.86**	0.23	0.32	**0.64**	0.35	0.39	0.02	**0.78**
DIS-COL	0.33	0.13	0.47	1.00	0.41	0.16	0.42	0.08	0.04	0.11	0.25	0.18	−0.16	0.43
DIS-CUL	−0.08	0.01	**0.84**	0.41	1.00	0.17	**0.98**	0.17	0.03	0.52	0.35	0.44	−0.24	**0.94**
DIS-GAR	0.38	0.34	0.40	0.16	0.17	1.00	0.29	0.13	**0.91**	**0.73**	0.14	−0.07	0.56	**0.36**
DIS-HOS	−0.02	0.02	**0.86**	0.42	**0.98**	0.29	1.00	0.17	0.16	**0.63**	0.34	0.40	−0.23	**0.97**
DIS-ROD	0.32	0.29	0.23	0.08	0.17	0.13	0.17	1.00	0.03	0.03	0.14	0.03	0.01	0.17
DIS-MAL	0.32	0.28	0.32	0.04	0.03	**0.91**	0.16	0.03	1.00	**0.75**	0.11	0.00	**0.66**	0.20
DIS-MRK	0.02	0.16	**0.64**	0.11	0.52	**0.73**	0.63	0.03	**0.75**	1.00	0.19	0.14	0.25	**0.64**
DIS-MOS	0.20	0.22	0.35	0.25	0.35	0.14	0.34	0.138	0.11	0.19	1.00	0.35	0.13	0.34
DIS-SCHL	0.04	0.05	0.39	0.18	0.44	−0.07	0.40	0.03	0.00	0.14	0.35	1.00	0.09	0.33
DIS-REC	0.25	0.35	0.02	−0.16	−0.24	**0.56**	−0.23	0.01	**0.66**	0.25	0.13	0.09	1.00	−0.26
DIS-UNI	0.06	0.05	**0.78**	0.43	**0.94**	0.36	**0.97**	0.17	0.20	**0.64**	0.34	0.33	−0.26	1.00

References

Antipov, E.A., Pokryshevskaya, E.B., 2012. Mass appraisal of residential apartments: an application of random forest for valuation and a CART-based approach for model diagnostics. Expert Syst. Appl. 39, 1772–1778.

Bagdonavicius, A., Deveikis, S., 2006. Automated valuation models in Lithuania. A Paper Presented at the FIG Workshop on Automated Valuation Models of Real Estate Purposes 19–20.

Barandiaran, I., 1998. The random subspace method for constructing decision forests. IEEE Trans. Pattern Anal. Mach. Intell. 20 (8), 832–844 20, 1-22.

Benjamin, J., Guttery, R., Sirmans, C., 2004. Mass appraisal: an introduction to multiple regression analysis for real estate valuation. J. Real Estate Pract. Educ. 7, 65–77.

Breiman, L., 1996. Bagging predictors. Mach. Learn. 24, 123–140.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Breiman, L., Friedman, J., Stone, C.J., Olshen, R., 1984. Classification and Regression Trees. CRC Press, United States.

Čeh, M., Kilibarda, M., Lisec, A., Bajat, B., 2018. Estimating the performance of random forest versus multiple regression for predicting prices of the apartments. ISPRS Int. J. Geo-Inf. 7, 168.

Chiarazzo, V., Caggiani, L., Marinelli, M., Ottomanelli, M., 2014. A Neural Network based model for real estate price estimation considering environmental quality of property location. Transp. Res. Procedia 3, 810–817.

Clapp, J.M., 2003. A semiparametric method for valuing residential locations: application to automated valuation. J. Real Estate Fin. Econom. 27, 303–320.

Demetriou, D., 2016. The assessment of land valuation in land consolidation schemes: the need for a new land valuation framework. Land Use Policy 54, 487–498.

Dimopoulos, T., Tyralis, H., Bakas, N.P., Hadjimitsis, D., 2018. Accuracy measurement of Random Forests and Linear Regression for mass appraisal models that estimate the prices of residential apartments in Nicosia, Cyprus. Adv. Geosci. 45, 377–382.

Eckert, J.K., Gloudemans, R.J., Almy, R.R., 1990. Property Appraisal and Assessment Administration. Intl Assn of Assessing Off.

Fan, G.-Z., Ong, S.E., Koh, H.C., 2006. Determinants of house price: a decision tree approach. Urban Stud. 43, 2301–2315.

GDLRC, 2014. Pilot Mass Appraisal Report of Mamak. General Directorate of Land Registry and Cadastre, Ankara.

GDM, 2018. General Directorate of Mapping, Sphere V4 2.98. GDM, Ankara.

Güneş, T., Yildiz, Ü., 2016. Property valuation and taxation for improving local governance in Turkey. Land Tenure J. 15, 141–160.

Ho, T.K., 1995. Random decision forests. In: Proceedings of 3rd International Conference on Document Analysis and Recognition. IEEE. pp. 278–282.

IAAO, 2013. Standard on Ratio Studies. International Association of Assessing Officers (IAAO), Kansas City, Missouri, USA.

IBM Corp, A, 1968. IBM SPSS Statistics for Windows, 2015 ed. IBM Corp.

Kara, A., Çağdaş, V., Isikdag, U., van Oosterom, P.J.M., Lemmen, C., Stubkjær, E., 2017. Towards an International Data Standard for Immovable Property Valuation. FIG Working Week 2017, 29 May-2 June, Helsinki, Finland.

Kouzani, A.Z., Nahavandi, S., Khoshmanesh, K., 2007. Face classification by a random Forest. TENCON - 2007 IEEE Region 10 Conference Paper.

Li, K., Yu, N., Li, P., Song, S., Wu, Y., Li, Y., Liu, M., 2017. Multi-label spacecraft electrical signal classification method based on DBN and random forest. Plos One 12, e0176614.

Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. R. News 2, 18–22.

Mamak Municipality, M., 2019. Zoning Plan. Mamak Municipality.

Mark, J., Goldberg, M.A., 1988. Multiple regression analysis and mass assessment: a review. Appraisal J. 56, 89.

Masías, V., Valle, M., Crespo, F., Crespo, R., Vargas, A., Laengle, S., 2016. Property valuation using machine learning algorithms: a study in a Metropolitan-Area of Chile. Selection at the AMSE Conferences-2016 97.

MATLAB Statistics Toolbox, 2015. The MathWorks, Inc., Natick, Massachusetts, United States.

McCluskey, W., Deddis, W., Mannis, A., McBurney, D., Borst, R., 1997. Interactive application of computer assisted mass appraisal and geographic information systems. JPVI 15, 448–465.

McCluskey, W., Davis, P., Haran, M., McCord, M., McIlhatton, D., 2012. The potential of artificial neural networks in mass appraisal: the case revisited. J. Fin. Manage. Prop. Construct. 17, 274–292.

Mora-Esperanza, J., 2004. La inteligencia artificial aplicada a la valoración de inmuebles: Un ejemplo para valorar Madrid. CT Catastro 50, 51–67.

Nghiep, N., Al, C., 2001. Predicting housing value: a comparison of multiple regression analysis and artificial neural networks. J. Real Estate Res. 22, 313–336.

Oshiro, T.M., Perez, P.S., Baranauskas, J.A., 2012. How many trees in a random Forest? Lecture Notes in Computer Science, vol 7376 Springer.

Park, B., Bae, J.K., 2015. Using machine learning algorithms for housing price prediction:

the case of Fairfax County, Virginia housing data. Expert Syst. Appl. 42, 2928–2934.
Quinlan, J.R., 1986. Induction of decision trees. Mach. Learn. 1, 81–106.
Robbins, M., 2001. Process for automated real estate valuation. Google Patents.
Trawiński, B., Telec, Z., Krasnoborski, J., Piwowarczyk, M., Talaga, M., Lasota, T., Sawiłow, E., 2017. Comparison of expert algorithms with machine learning models for real estate appraisal. In: 2017 IEEE International Conference on Innovations in Intelligent Systems and Applications (INISTA). IEEE. pp. 51–54.
Wager, S., 2014. Asymptotic Theory for Random Forests. arXiv Preprint. .
Wilhelmsson, M., 2009. Construction and updating of property price index series: the case of segmented markets in Stockholm. Prop. Manage. 27, 119–137.
Williamson, I., Enemark, S., Wallace, J., Rajabifard, A., 2010. Land administration for sustainable development. Citeseer.
Zurada, J., Levitan, A., Guan, J., 2011. A comparison of regression and artificial intelligence methods in a mass appraisal context. J. Real Estate Res. 33, 349–387.