Acta Montanistica Slovaca
ISSN 1335-1788
actamont.tuke.sk

Investigation of the difference between the formal values of 2/B lands that have lost their forest status and the values of agricultural lands for sale in the market

Sultan CINAR BILGIS[1] *and Fatma BUNYAN UNEL*[2]*

Authors' affiliations and addresses:
[1] Department of Remote Sensing and Geographic Information Systems, Institute of Science, Mersin University, Mersin, Turkiye.
e-mail: sultancinar1998@gmail.com

[2] Department of Geomatics Engineering, Faculty of Engineering, Mersin University, Mersin, Turkiye.
e-mail: fatmabunel@mersin.edu.tr

***Correspondence:**
Fatma Bunyan Unel, Department of Geomatics Engineering, Mersin University, Mersin, Turkiye.
tel.: +90 324 361 0001 / 17357
e-mail: fatmabunel@mersin.edu.tr

Acknowledgement:
This study is adapted from Sultan CINAR (BILGIS)'s master's thesis conducted under the supervision of Asst. Prof. Dr. Fatma BUNYAN UNEL. The survey was reviewed and approved by the Mersin University Science Ethics Committee on 13 April 2021 (Approval No: 4).

How to cite this article:
Bilgis, S. C. and Unel, F. B. (2025). Investigation of difference between the formal values of 2/B lands that have lost their forest status and the values of agricultural lands for sale in the market. *Acta Montanistica Slovaca*, Volume 30 (4), 1079-1107

DOI:
https://doi.org/10.46544/AMS.v30i4.17

Abstract
The aim of this study is to compare two different value groups, namely 2/B and agricultural lands, in the Sarıkaya neighbourhood of Erdemli district, Mersin province. The first group comprises 414 formal values associated with 2/B agricultural lands, while the second group consists of 40 agricultural land prices compiled from market data. In order to determine the weighting of 31 criteria, a questionnaire was conducted with the participation of 40 experts and 210 citizens. The results of the frequency analysis of the questionnaire were taken into consideration, and the criteria were weighted with the Analytical Hierarchy Process (AHP) method. Criteria with weights below 0.050 and 0.010 were eliminated, and four different scenarios were generated using the formal value data set.
Scenarios were created for model validation using Multiple Regression Analysis (MRA) and Artificial Neural Networks (ANN) valuation models. Given this circumstance, it was determined to proceed with Scenario 3 in accordance with the results of performance analyses that included the mean absolute percentage error, mean absolute error, mean squared error, and R^2. In line with the final 6 and 17 criteria, value estimations were carried out with Ridge Regression Analysis (Ridge) and Random Forest (RF) valuation models. The value maps were produced with GIS software. The same process steps were performed with the final criteria by utilising the market values. A total of 380 sample points were selected from the formal and market value maps, and the permutation test was used to determine whether there was a difference between the two value groups. The test showed that there is an acceptable value difference between the two comparison groups.

Keywords
Agricultural lands, 2/B forest lands, land value, real estate valuation.

Introduction

Agricultural lands are the source of life for human beings, and they supply a large part of the world's food. Land used for agriculture cannot expand at the same rate as the population. As of March 2024, the world population exceeded 8 billion (Worldometer, 2024). It is forecasted that the population will be 9.7 billion in 2050 and 10.3 billion in 2100 (UN, 2022). Turkiye has a population of 85 million people (TUIK, 2024a). Population growth means that there will be a need for more food production and arable land. Although agricultural areas in Turkiye increased by 3.44% between 2020 and 2023, they decreased by 9.97% between 2004 and 2023 (TUIK, 2024b). Gross Domestic Product (GDP) increased by 4.5% in 2023, within which real estate activities increased by 2.7% and the agricultural sector decreased by 0.2% (TUIK, 2024c).

Sustainable development goals include ensuring that everyone benefits equally by making good use of soil and water resources (SDG, 2023). The resources are used for the economic, social, and environmental needs of present and future generations. The protection of sources can be achieved through sustainable land management (Enemark, 2004). Sustainable land management includes land administration that is land tenure, value, use, and development (Enemark, 2010; Enemark et al., 2014). Within the scope of sustainable development goals, the LADM Valuation Information Model diagrams were prepared and applied in Turkiye (Kara et al., 2020, 2021). The value is directly related to the characteristics and location of the real estate and is particularly utilised in taxation and valuation.

Sustainable processing and efficient use of the soil is only possible with planned and programmed land application activities. Since land management has an impact on every field, it concerns many disciplines and is subject to many legislations. One of the basic pieces of information used in relation to land is its value, particularly in sustainable management. Land value is determined by its surface area, geometric shape, soil fertility, topographic, geologic, and hydrographic features, and locational features such as distance to the road, market, the settlement centre, etc. (Alkan & Durduran, 2024; Çinar & Ünel, 2022). The Land Use Capability (LUC) classes of the lands were determined by examining the slope, drainage, soil depth, permeability, composition, stoniness, etc. LUC is divided into 8 classes according to the Technical Guideline (2017). In this case, Soils designated as first, second, third, and fourth class by the Ministry of Agriculture and Forestry are agricultural soils (Agricultural Reform Law, 1984). In addition, the world soil map was completed by FAO/UNESCO in 20 years (FAO, 2024). Eventually, it is important to have the same topographical and soil characteristics and a similar agricultural system (Colwell & Dilmore, 1999).

It was seen necessary to determine agricultural land values due to land wastage and underutilization in 1960. Agricultural land values depend on several criteria in a market economy. One of the major criteria is the land rent, which is directly related to the net income. While the net income and the rent are indicators of the demand for land, the supply and demand are direct determining criteria for land market prices (Burger, 1998; Hertel, 2011). The supply and demand are also influential for residential land values (Ma & Li, 2017; Zhang et al., 2023). Agricultural land values are calculated using the capitalisation rate, a traditional method that utilises net income (Fan et al., 2024; Perujo-Villanueva & Colombo, 2021; Stokes & Jansen, 2024).

According to the FAO (2008), land consolidation projects remarked that agricultural land markets could evolve in response to supply and demand at certain prices. FAO states that the projects can facilitate land mobility through bank transactions and funds. In this context, it is evident that mass land valuation in rural areas is particularly aimed at being applied in land consolidation studies (Asiama et al., 2018; Demetriou, 2016, 2017; Ertunç & Uyan, 2022). Automated valuation models (AVMs) were proposed due to the costliness and time-consuming nature of valuation processes. *The AVM is highly efficient compared to conventional land valuation methods since it may considerably reduce the time and resources used and provide transparency because the process has been converted from empirical to systematic, analytical, and standardized* (Demetriou, 2018, 21), Mass valuation studies were also carried out for urban area readjustment (Yomralioglu, 1993; Yomralioglu et al., 2007).

The mass valuation methods are used for real estate taxation purposes. In real estate appraisal, machine learning techniques are commonly used. Multiple Regression Analysis (MRA) (Abdulhafedh, 2022; Carbonara et al., 2021; Zhang, 2021; Uberti et al., 2018), Ridge Regression Analysis (Yazdani, & Raissi, 2023), Artificial Neural Networks (ANN) (Aydinoglu, & Sisman, 2024; Dewi et al., 2024; Przekop, 2022; Tabar et al., 2023), Random Forest (RF) (Antipov & Pokryshevskaya, 2012; Hong et al., 2020; Iban, 2022; Jafary et al., 2024; Moosavi, 2017; Pastukh, & Khomyshyn, 2025; Yazdani, 2021), and other techniques are frequently employed. The aforementioned studies serve purposes including determining criteria, extracting criteria, and developing methods for estimating real estate values. Especially ANN, one of the modern valuation methods, can make high-accuracy estimates with data containing multivariate and complex relationships (Alsahan & AlZaidan, 2024; Wu et al., 2022). RF is seen to be used effectively in real estate valuation studies by providing consistent estimates. It is also stated that it is stable, strong, and exhibits the best performance among the methods compared by the relevant articles in real estate valuation estimates (Antipov & Pokryshevskaya, 2017; Chen et al., 2022; Moreno-Foronda et al., 2025; Tochaiwat & Pultawee, 2024). An important source of income for countries is taxes and duties collected from real estate (Adil & Syafaruddin, 2024; Bahl & Martinez-Vazquez, 2007; Bird & Slack, 2004; OECD, 2024). The main issue for such real estate transactions is the real estate value (European Commission, 2014; Iara, 2015). The systematic and orderly processing of the resource is made possible by conducting valuation studies based on market values. For this reason, an accurate and up-to-date collection of market values is of great importance for the tax fee system. Entry into the value registration system based on citizen declaration paves the way for an informal economy in terms of real estate (IAAO, 2020; World Bank, 2020). In this case, the formal value recorded in the institutions and the market value circulating in the market are different. The question of this study is whether the difference between the formal and market value is statistically significant. Since it is difficult to find correctly recorded data, the valuation committees working within the institutions also experience problems in determining the value. Finding sample values in 2/B Agricultural Lands is significantly more challenging due to the unique circumstances present.

The aim of this study is to compare the formal and market values by evaluating the 2/B agricultural lands in the Sarıkaya neighbourhood of the Erdemli district of Mersin province with modern methods. The originality of the study is listed as the fact that the values of 2/B agricultural lands obtained under the name of the formal value are national property values, the property values are analysed with two different valuation models, value estimations are made by producing (4+4) different scenarios, and the predicted values are compared with each other.

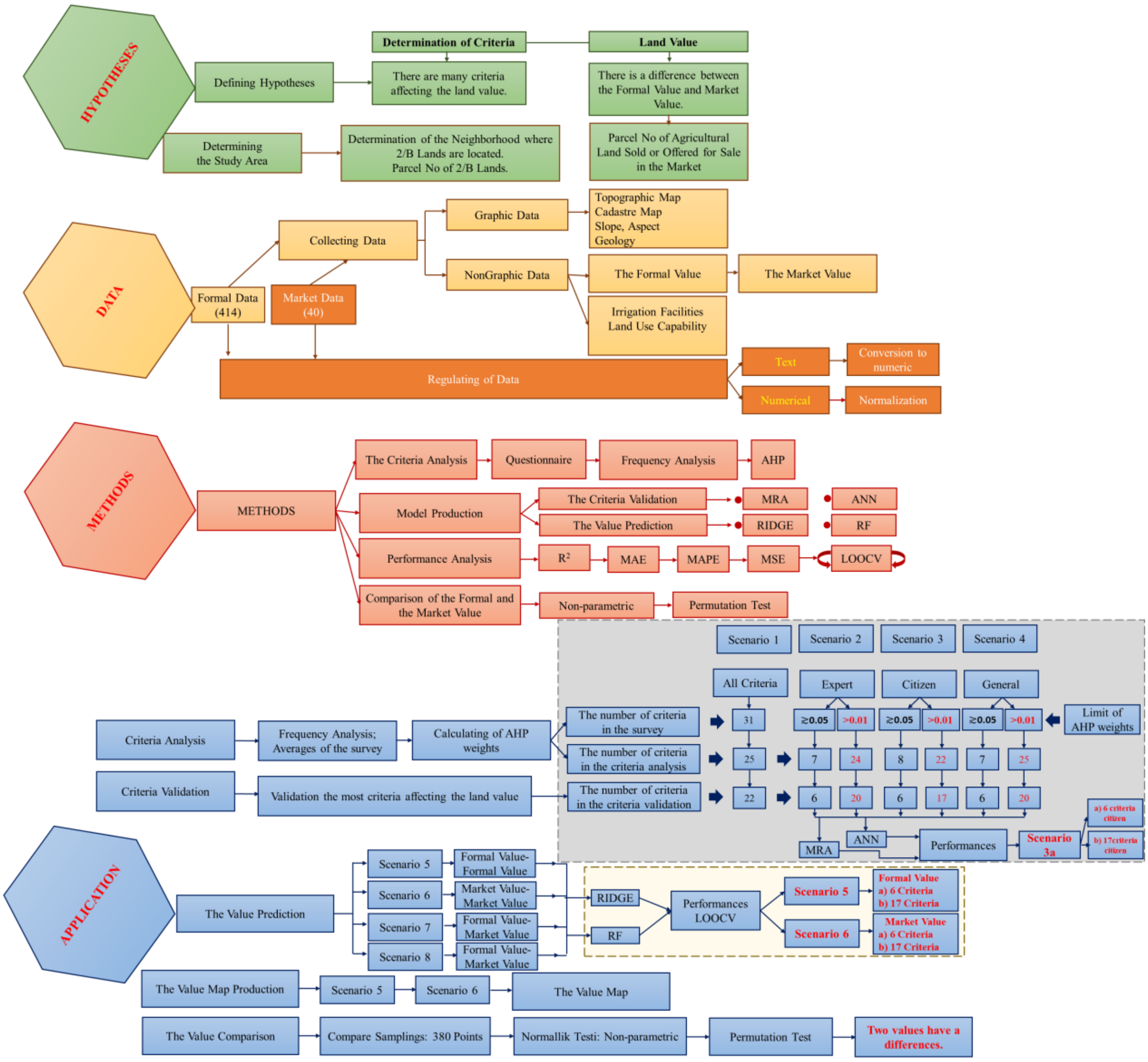

Fig. 1. The workflow diagram

The criteria affecting agricultural lands were investigated, and frequency analyses were performed using the questionnaire method and weighted by the AHP method according to the results of the analysis. The data set of agricultural lands comprises the formal values of 2/B agricultural lands (414) and the market values of agricultural lands traded through purchase and sale in the market (40), along with their characteristics. According to the criteria reduced by frequency and AHP methods, four scenarios were designed. The scenarios were validated using MLR and ANN, with the data split into 70% for training and 30% for testing. The performance analysis results of the best scenarios were then examined. According to the results, the number of criteria was reduced to 6 and 17. To compare the official and market values, four additional scenarios were developed based on these criteria and analyzed using Ridge and RF methods, resulting in the production of value maps. To analyze the difference between the formal and market values corresponding to 380 points on the value map, a permutation test, a nonparametric method, was employed.

Material and Methods

This section includes the workflow diagram, an overview of the study area, the characteristics of the data sets, and the methodologies utilized. The materials and methods used for this study are indicated, and the workflow is shown (Fig. 1). The workflow diagram comprises four stages: hypothesis, data, method, and application. This is a summary of the work steps performed in the study.

The Study Area

Mersin province, Erdemli district, Sarıkaya Neighbourhood was determined as the study area (Fig. 2). The reason why the study area is preferred is that the neighbourhood has more 2/B agricultural lands compared to other neighbourhoods.

Sarıkaya Neighbourhood is located 90 km from Mersin city centre and 31 km from the Erdemli district centre, with an elevation of 1300 meters. In 2012, the village transitioned to neighbourhood management under the scope of Law No. 6360 (Province and Districts Law, 2012), as mandated by Mersin Metropolitan Municipality. There is a primary school and a health centre in the neighbourhood. Transportation is provided by a minibus and a public bus. Since the residents of the district use it as a summer plateau, the population of the neighbourhood increases in the summer months and can reach up to 5000.

Forests and 2/B Agricultural Lands: According to the Forest Law No. 6831, the areas that are considered forest and those that are not forest are defined, and their borders are clearly stated. Trees and shrub communities that grow naturally or are cultivated by labour are referred to as forests, together with their locations. However, the following places are not considered forests (Forest Law, 1956):

- Reeds,
- Places covered with steppe plants,
- All kinds of thorns,
- Parks,
- Areas covered with trees and groves in cemeteries, etc.

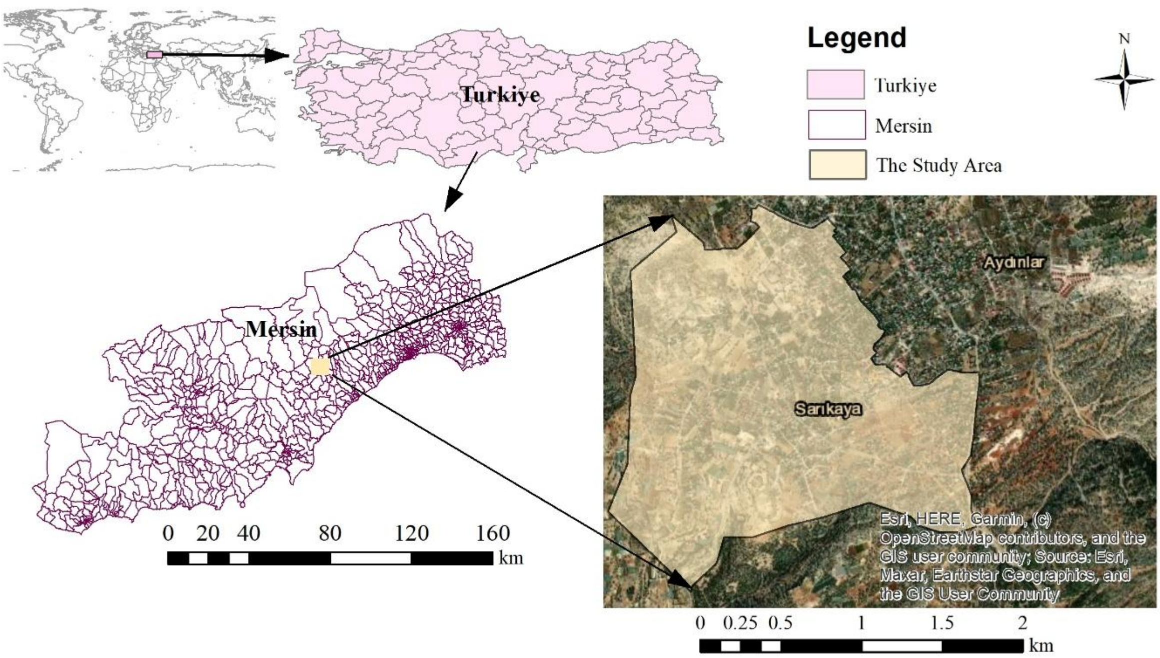

Fig. 2. The study area

The Forest Law No. 6831's 2 (B) article refers to lands taken out of forest boundaries and therefore, it is named as 2/B lands. Before 31 December 1981, these lands had completely lost their forest status in terms of science. If these places belong to the State, they are registered in the name of the Treasury. If these places belong to public organizations that are legal entities, they are registered or appropriated in the name of these organizations. If these places are private forests, they are registered in the name of their owners. In addition, the forest borders cannot be narrowed (Forest Law, 1956).

The study area comprises 981 2/B lands in total, with the distribution of these lands in the Sarıkaya neighbourhood presented spatially in Fig. 3. These were eliminated due to reasons such as being too small, creating extreme data, negatively affecting the analysis, and being unable to reach sales values. In the analysis, a total of 414 formal data points from 2/B agricultural lands, as valued by the Valuation Commission of the General Directorate of National Property, were used as samples.

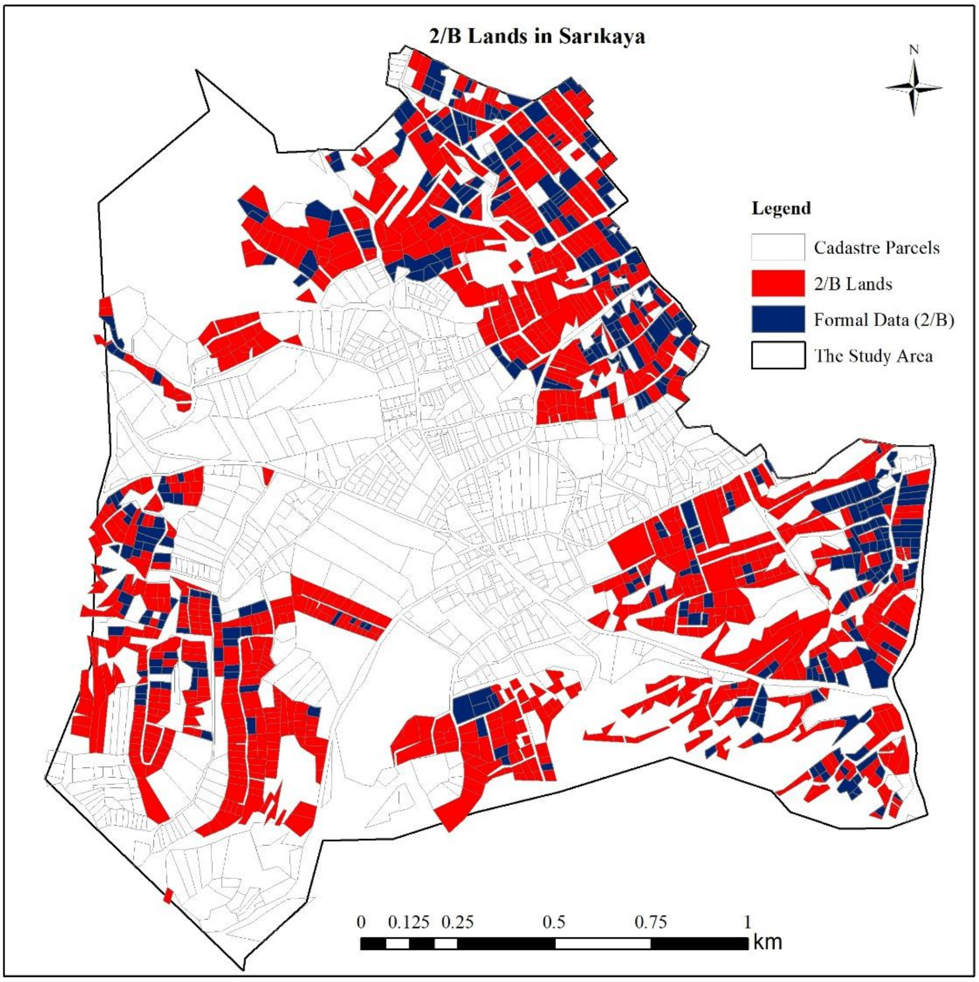

Fig. 3. Distribution of the 2/B lands in the Sarıkaya neighbourhood

The real estate information (real estate number, block/parcel) of the 2/B agricultural lands in the region and the unit m^2 price (formal values) specified in the report dated 12 February 2020, determined by the valuation commission, the area and spatial characteristics of the parcels (distances to district, road, nearest village, etc.) were collected.

The Agricultural Land (Farmland): Agricultural production must be done in a sustainable way if human life is to continue. To achieve this, agricultural lands with fertile soil are needed. These are protected under the Soil Protection and Land Use Law. In the law, Article 8, agricultural lands are classified according to their natural characteristics and importance in the country's agriculture. These are absolute agricultural lands, special crop lands, planted agricultural lands, and marginal agricultural lands (Land Use Law, 2005).

According to TUIK (2024b; 2024c) reports, it has been stated that in the change of agricultural lands in Turkiye between 2006 and 2012, the most frequent transformation of agricultural lands took place, followed by the transformation of forests and semi-natural areas (Bayar, 2018). Pasture-meadow areas have also seen significant alterations, and it has been determined that they have decreased by 61.5% over the past 70 years. Within the scope of the annulled law on land provision to farmers, many pasture-meadow areas have been appropriated by villagers. 7.5 million hectares of bushland were assigned to the Ministry of Forestry, and changes in pasture areas occurred due to land classification (Gökkuş, 2018). Before 31 December 1981, forest areas that lost their forest status were removed from the forest and converted into agricultural lands, known as 2/B lands.

To be used for analysis in the study, market data on 40 agricultural lands offered for purchase and sale were collected from an online announcement. Due to low trading, the minimum number of data points that would be statistically significant was provided. As a result of the analysis, the formal value and the market value were compared.

The Criteria Affecting Agricultural Land Value

There are many criteria that affect the value of real estate. Criterion name and number vary depending on the type of real estate, its location, and accessibility (Tempesta et al., 2021; Unel et al., 2023; Yalpır & Ünel, 2016). When agricultural lands are considered as a real estate type, they are grouped under 6 main headings: legal, social, physical, locational, agricultural, and land structure. There are many criteria under these headings, such as slope, soil characteristics, irrigation status, distance to the village, and distance to the road (Fig. 4).

The criteria were ranked as a result of the literature review (Dedeoğlu & Dengiz, 2018; Demetriou, 2016, 2018; Hüttel et al., 2016; Meyer & Früh-Müller, 2020; Öztürk et al., 2017; Sylla et al., 2019; Technical Guideline, 2017) and the existence of criteria data was investigated. In order for the data of the criteria to be analysed with statistical and modern valuation methods, the criteria must be compatible with software formats. In this context, the criteria affecting the value of agricultural land were listed (Fig. 4), and standards were determined. At the same time, these criteria are used in the survey application. While the main titles received scores out of 100, the criteria were assessed on a 5-point Likert scale.

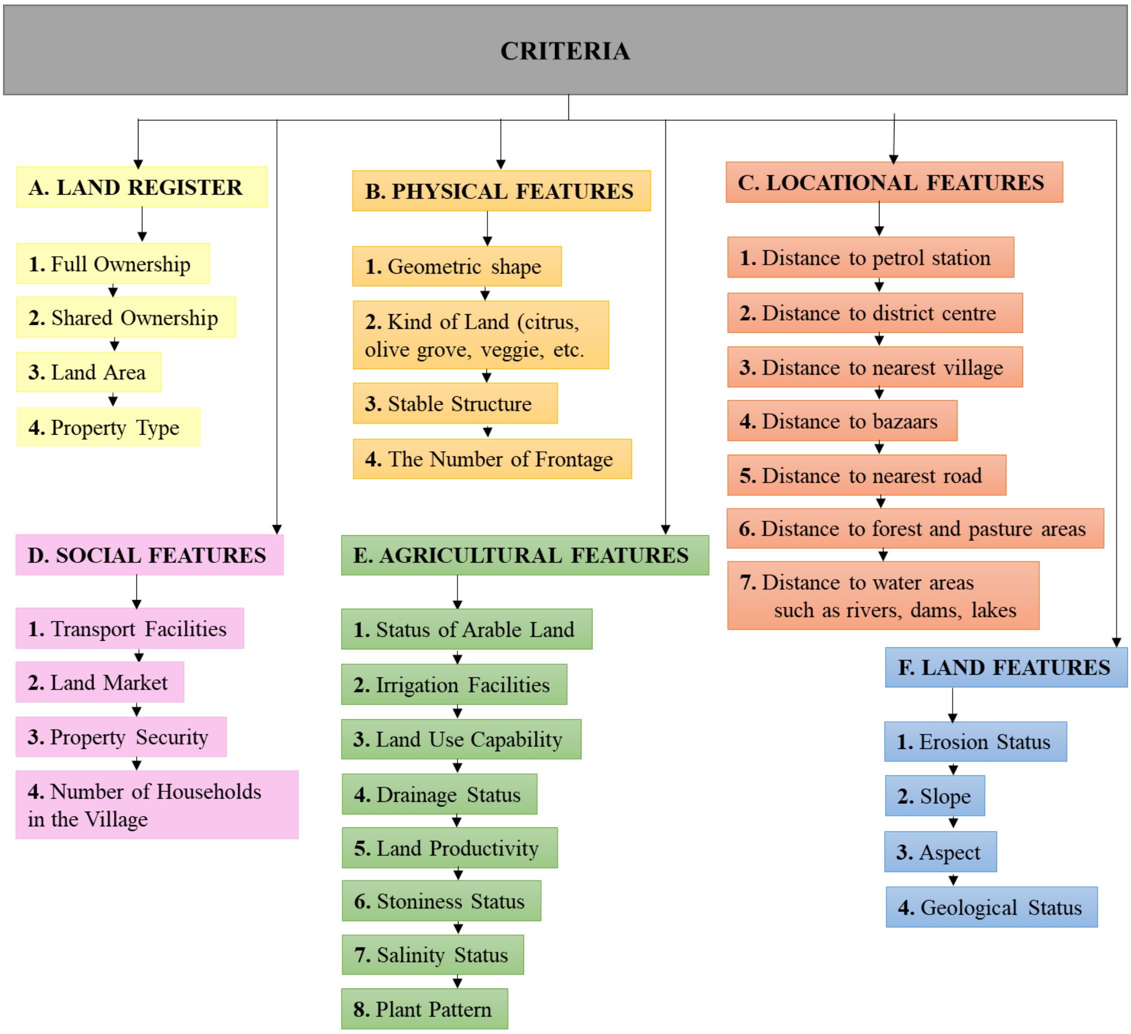

Fig. 4. The criteria of the survey application

The formal values were obtained from the public institution. The market values were gathered from a variety of sources, including citizens, real estate agents, and the internet. During the survey study, information regarding the market was also collected from members of the local community. In the Sarıkaya Neighbourhood, the formal value of 414 2/B agricultural lands and the market value of 40 agricultural lands traded in the market were collected. All data were normalized between 1 and 2 using the formula given in Eq. 1. After this process, all data were computed without units and between two predetermined numbers, producing a more accurate mathematical model.

$$z_i = \frac{x_i - min\ (x)}{max(x) - min\ (x)} + 1 \tag{1}$$

The Methods

The study consists of three stages: criteria extraction, model building, and comparing values. The methods applied are described in order, and the results are evaluated. The survey, frequency, and AHP methods were used to find the criteria to be subject to analysis. Multiple Regression Analysis (MRA), Ridge Regression (Ridge), Artificial Neural Networks (ANN), and Random Forest (RF) were used to estimate real estate value. The Permutation Test was used to find out whether there are differences between formal and market values.

A survey is a measurement-evaluation tool used to measure the general acceptance opinions of the people in a region about a subject. "Investigation of the criteria affecting the value of agricultural lands excluded from 2/B forests" was the focus of the survey in the study. It was carried out with the help of the data obtained with the questionnaire method:

- Whether there is a relationship between dependent and independent variables,
- determining the weights of independent variables,
- Formulating new assumptions (hypotheses) in the light of this information and planning new research.

The survey was conducted in April-May 2021. The participants consist of 210 citizens (farmers) living in the region and 40 experts who participated in the survey via e-mail. Experts are distributed as follows: 30% in geomatics, 14% in city and regional planning, 10% in agriculture, 7% in forestry, 3% in geological engineering, 3% in architecture, and the remainder in other fields. The criteria affecting agricultural land values were rated according to a 5-point Likert scale as "Very Bad (-5), Bad (-3), Ineffective (1), Good (3), Very Good (5)". The data from the survey were exported from Google Forms software in table format.

In this study, it was used LiCAD (2024) software for the processing of vector data, ArcGIS (ESRI, 2016) software for the production of value maps, MatLAB (MathWorks, 2010) software for the ANN model, SPSS (IBM, 2011) software for MRA and other statistical analyses, and Spyder (2023) open source software for Ridge and RF.

Criteria Analysis Methods: The analyses used for value estimation and criteria identification are different from each other. In this section, the criteria analyses, specifically the frequency and AHP methods, are investigated. After the answers of the survey participants were edited, the frequencies of the data were examined.

The frequency of an event is the number of times an experiment or study event occurs. These frequencies are usually plotted graphically and expressed by histograms (Kenney & Keeping, 1962). The survey answers were analyzed using SPSS software to identify the most frequently given answer and calculate participant averages.

The Analytic Hierarchy Process: "The Analytic Hierarchy Process (AHP) is a quantitative method for ranking and selecting decision alternatives according to multiple criteria." (Russell & Taylor, 2003). In other words, the Analytic Hierarchy Process is the process of developing numerical values to rank each decision alternative according to the degree to which it meets the decision maker's criteria. To provide numerical scores for the criteria, it is necessary to be an expert in real estate appraisal, considering the importance of each criterion in relation to the value of the real estate. In this study, experts and citizens were consulted to obtain general acceptance of the criteria weights. The results of the survey administered to experts and citizens have guided us on this issue. Table 1 shows the numerical **AHP scores** given to the criteria and their meanings.

Tab. 1. Degree of importance of criteria (R. W. Saaty, 1987; T. L. Saaty, 2008)

Degree of importance	Definition	Explanation
1	Equal importance	Two criteria or activities contribute equally to the objective
3	Moderate importance of one over another	Experience and judgment strongly favour one criterion or activity over another
5	Essential or strong importance	Experience and judgment strongly favour one activity over another
7	Very strong importance	A criterion or activity is strongly favoured, and its dominance is demonstrated in practice
9	Extreme importance	The evidence favouring one criterion or activity over another is of the highest possible order of affirmation
2, 4, 6, 8	Intermediate values	When compromise is needed

AHP equations are summarised. Firstly, Pairwise Comparison Matrix $\left[a_{ij}\right]$ constructs for each criterion group (Eq. 2) and normalised. AHP weights (w_i) of the criteria are found by averaging the rows (Eq. 3-4). The consistency is investigated with Eq. 5-9, and when Consistency Ratio CR ≤ 0.10, the Pairwise Comparison Matrices are reconstructed (R. W. Saaty, 1987; T. L. Saaty, 1990).

$$A=\left[a_{ij}\right]=\begin{bmatrix} 1 & a_{12} & a_{13} & \dots & a_{1n} \\ 1/a_{12} & 1 & a_{23} & \dots & a_{2n} \\ 1/a_{13} & 1/a_{23} & 1 & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ 1/a_{1n} & 1/a_{2n} & 1/a_{3n} & \dots & 1 \end{bmatrix} \quad (2)$$

$$a_{ij}^{*}=\frac{a_{ij}}{\sum_{i=1}^{n} a_{ij}} \quad (3)$$

$$w_i=\frac{\sum_{j=1}^{n} a_{ij}^{*}}{n} \quad (4)$$

$$D=\left[a_{ij}\right]_{n\times n}\times\left[w_i\right]_{n\times 1}=\left[d_i\right]_{n\times 1} \quad (5)$$

$$E=\frac{d_i}{w_i} \quad (6)$$

$$\lambda=\frac{\sum_{i=1}^{n} E_i}{n} \quad (7)$$

$$CI=\frac{\lambda-n}{n-1} \quad (8)$$

$$CR=\frac{CI}{RI} \quad (9)$$

a_{ij} **:** Pairwise Comparison Matrix elements
n **:** The number of criteria
$i, j = 1,2,3,.....,n$
CI : Consistency Index
RI : Random Index
CR : Consistency Ratio

Real Estate Valuation Methods: The method to be chosen for the valuation is determined according to the type and location of the real estate to be valued and the prevailing habits of the real estate market (Açlar & Çağdaş, 2008). In this study, the type of real estate is determined as agricultural land. ANN and RF are two modern valuation methods, while MRA and Ridge are two statistical valuation methods used for model verification.

Multiple regression analysis (MRA): The multiple regression analysis is generally used to investigate the existence of a relationship between multiple variables. If there is any relationship, it has been useful to explain how it happened. Variables consist of dependent and independent variables (Akış, 2013; Skiera et al., 2018; Sykes, 1993). In the valuation of agricultural lands in a region, the value is the dependent variable, while criteria such as land area, geometric shape, frontage length, etc., are considered as independent variables. With multiple regression analysis, the effect of independent variables on real estate value can be calculated.

The mathematical model of multiple regression analysis is calculated as given in Eq. 10 (Skiera et al., 2018; Szentesi et al., 2024).

$$y_i=\beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}+\cdots+\beta_k x_{ik}+u_i \quad (10)$$

x_{ik} : Independent variable (Land area, geometric shape, frontage length, etc.)
y_i : Dependent variable (The formal and market value of 2/B agricultural land)
β_0 : Constant coefficient
$\beta_1, \beta_2, \beta_3, \dots, \beta_k$: Coefficients of criteria
u_i : Error term
$i = 1,2,3 \dots, n$: Number of parcels with land value
$k = 1,2,3 \dots, m$: Number of criteria

Ridge regression analysis: Ridge regression is a method employed to mitigate the multicollinearity issue present in multiple linear regression (El-Dereny & Rashwan, 2011). In cases of high correlation among predictor variables, Ridge regression yields coefficients that enhance prediction and extrapolation compared to least squares and serves as a reliable method for variable selection (Marquardt & Snee, 1975). This approach reduces the model's error term while concurrently incorporating a penalty term (Eq. 11) to the sum of squares of the regression coefficients. Thus, overfitting of the model is mitigated by constraining the magnitude of the coefficients. Ridge regression yields more stable and dependable outcomes compared to classical linear regression, particularly in scenarios of strong variable correlation or when the dataset is limited while the variable count is elevated (Hoerl & Kennard, 1970; McDonald, 2009; Hoerl, 2020).

$$\widehat{\beta_R} = (X^T X + kI)^{-1} X^T y \tag{11}$$

Where $\widehat{\boldsymbol{\beta_R}}$ denotes the vector of regression coefficients estimated using the Ridge regression method, $\boldsymbol{X}$ represents the matrix of input (independent) variables, and $\boldsymbol{X^T}$ is the transpose of $\boldsymbol{X}$ matrix. $\boldsymbol{k}$ is the Ridge penalty (regularization) term that controls the degree of shrinkage applied to the coefficients. $\boldsymbol{I}$ refers to the identity matrix, which ensures that the regularization is applied uniformly across all coefficients. Finally, $\boldsymbol{y}$ is the vector of dependent variable values.

Artificial Neural Networks (ANN): Artificial neural networks are computer programs that mimic the biological neural networks of the human brain. Variable weights interconnect with each other, and each has its own information processing structures. The neural network method is a refinement of the human brain's iterative learning process. Training input and output data sets are introduced to the system. The system is first trained with this training set, and then the test data set is applied to the network learning. There are many learning algorithms used in artificial neural networks. The most commonly used learning algorithms are Hebb, Delta, Back Propagation, and Kohonen Function (Alpaslan, 2015; Bisi & Goyal, 2017; Rumelhart et al., 1985; Shanmuganathan & Samarasinghe, 2016).

The basic working steps of an artificial neural network can be listed as follows (Shanmuganathan & Samarasinghe, 2016).

- Selecting the appropriate data set from the working data set and introducing the input data set to the neural network
- Calculation of the output value of the neural network
- Calculating the error between the output value of the neural network and the desired value
- Calculating the neural network weight to minimize errors
- Continuing to teach the network until errors are acceptable.

The basic thing that is done in the learning process of Artificial Neural Networks is to change the weight values (Saraç, 2012). The most common aggregation function (Eq. 12) is the sum of weights (Güzel, 2018).

$$Net\ Input = \sum_{i=1}^{j} X_i W_i \tag{12}$$

ANNs have activation functions such as sigmoid, hyperbolic tangent (tanh), and rectified linear unit (ReLU). Sigmoid, which is widely used (Şengöz, 2018), was preferred for this study. In general, the data set is proportioned as 70% training and 30% testing (İlhan & Öz, 2020; Shobha Rani et al., 2018; Yalpir et al., 2014). The prediction process was performed with the non-linear sigmoid function of MATLAB software.

Random Forest Algorithm: Random forests (RF) were initially presented by Heath et al. (1996) for the aim of classification and decision making. The random forest algorithm, which was created by Ho (1995), is based on decision trees. (Murthy et al., 1994) investigated the benefits of randomization in the construction of oblique trees by using both real and artificial data.

Random forests are a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest (Breiman, 2001, 5). Random Forests can manage hundreds of variables without eliminating them or compromising accuracy. The RF process appears straightforward yet is challenging to study (Breiman, 2004).

Random Forest is a prediction technique with numerous iterations (Breiman, 1996). The Random Forest algorithm is a robust ensemble learning technique that can achieve high accuracy in regression tasks, even with a small number of observations (Biau & Scornet, 2016). This methodology employs iterative training of decision trees via bootstrapping, aggregating outcomes through voting or averaging. This framework enhances the model's overall efficacy by mitigating the likelihood of overfitting, particularly in small data sets (Han et al., 2021; Qi, 2012). It effectively captures non-linear correlations among variables and elucidates their significance. Therefore, it is a robust and dependable choice for researchers handling small data sets. It is rapid and simple to execute, yields very precise predictions, and accommodates several input variables without overfitting (Biau, 2012).

The model hyperparameters, which represented the standard implementation of Random Forest in software packages, are configured as follows: n_estimators=10, max_depth=2, and random_state=42, which is widely recognized for its efficacy in numerous scenarios. Nonetheless, modifying these hyperparameters in various manners can enhance the performance of RF (Probst et al., 2019). n_estimators=10 indicates that 10 decision trees are produced to achieve the benefit of reduced variance. max_depth=2 signifies the depth of each decision tree, hence mitigating the danger of overfitting by constraining the depth. The model was favored because of its ability to deliver high accuracy in datasets with a limited number of observations and its resilience to overfitting (El Hami, 2025; Han et al., 2021). The random_state was set to 42, guaranteeing the model's repeatability. Rectifying the random number generation enhances the reliability of statistical comparisons by guaranteeing that the model yields consistent findings. In the random forest methodology, predictions are produced using the widely recognized conceptual Eq. (13) (Biau, 2012; Breiman, 2001; Peng et al., 2022).

$$RF(x) = \frac{1}{N}\sum_{i=1}^{N} h(x; Z_{i1}, \dots Z_{is}; \omega) \tag{13}$$

where each $Z_{i1}, \dots Z_{is}$ denotes a subsample taken without replacement from the available training data, and ω denotes additional randomness injected into the base learner h.

Analysis of Differences: To determine whether there is any difference between the market value and the formal value, and if there is a difference, to analyse whether this difference is significant, a t-test was used. An independent sample t-test examines the significance of the difference between the arithmetic means obtained from two independent samples. The question is whether two measurements are significantly different from each other (Akdağ, 2011; Gerald, 2018). For the independent sample t-test, the process starts with the assumptions about the compared samples. These assumptions are as follows (Arslan, 2018).

- Two samples are independent of each other.
- The dependent variable is at the interval or ratio scale level.
- The distribution of raw scores of the population represented by each sample is normal.
- The variances of the universes represented by the samples are homogeneous.

A normality test was performed in SPSS software, and since the data were not normally distributed, they were analyzed with a permutation test.

This study employed a permutation test, which is independent of parametric assumptions, to ascertain whether a statistically significant difference exists between the two groups. The permutation test does not necessitate the assumption of independent and identically distributed (IID) data or a normal distribution, which are stringent prerequisites of traditional parametric tests (Bonnini et al., 2024). Rather, it depends on the more adaptable premise of exchangeability, which yields dependable outcomes, particularly in scenarios where the distribution is non-normal and the assumption of homogeneity of variance is unachievable.

Permutation tests provide a robust alternative, particularly in nonparametric contexts (Bertanha & Chung, 2023). Holt & Sullivan (2023) underscore that the permutation approach offers a versatile framework for statistical inference, allowing for significant conclusions to be derived from random permutations while preserving the integrity of the observed data. Moreover, this method is not only suitable for small samples, but also its asymptotic validity increases as the sample size increases. Kong et al. (2022) showed that permutation tests yield more robust findings than classical tests, particularly in high-dimensional data studies, and are not affected by parametric assumptions. Consequently, given the data structure and the limits of the assumptions in the study, the permutation test was deemed the most suitable procedure.

Performance Analysis: The mean absolute error (MAE) (Eq. 14) is a measure of the difference between two continuous variables. The mean absolute percentage error (MAPE) (Eq. 15) measures the average magnitude of the errors in a set of estimates (Aydın, 2018). The Mean Squared Error (MSE) (Eq. 16) provides the most accurate approximation of the degree of accuracy of the measurements. Since errors are squared in the calculation, larger errors within a measurement have a larger effect on the average, and thus the effect of large errors on the whole measurement can be determined. Coefficient of Determination (R^2) (Eq. 17) is a statistic loaded with systematic error. For a fixed number of independent variables, the level of systematic error decreases as the R^2 value increases and/or the sample size increases (Chen et al., 2024; Günel, 2003; Sharma et al., 2024). Leave One Out Cross Validation (LOOCV) is especially advantageous for small datasets, since it optimizes data utilization for training while delivering an impartial assessment of model efficacy (Shao & Er, 2016; Wong, 2015). Leave One Out Cross Validation mitigates the risk of overfitting in small sample sizes by evaluating the model on each observation independently. MAE, MAPE, and MSE with LOOCV are calculated repeatedly for each data row (Cha et al., 2020; Garreta & Moncecchi, 2013; Kuhn & Johnson, 2013).

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{14}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{y_i} \quad (15)$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (16)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (17)$$

n : Represents the number of samples;
y_i : The observed values;
$\hat{y}_i$: The predicted values;
$\bar{y}$: The mean of the observed values.

Results

The Criteria Extraction

The survey was used to determine which criteria affecting the value of 2/B agricultural lands in the study area would be considered and their respective weights. The (Cronbach) Alpha coefficient test was used to test the reliability of the questionnaire made on a 5-point Likert scale. As a result of the test, an alpha value of 0.844 was obtained. When the alpha value is interpreted according to the research by Taber (2018), it is concluded that the questionnaire is "Good".

To extract the criteria, frequency analysis was first carried out separately according to the survey results of experts, citizens, and the general public, which is a combined expert and citizen survey (Appendix A). In the frequency analysis, the total scores of the criteria were divided by the number of participants, and their averages were found. The main titles of the criteria, which are land register (Title Deed Information), physical, locational, agricultural, social, and land structure, were scored in such a way that their sum totalled 100 full points, and their averages were calculated. The criteria were scored on a scale of -5 to 5. These scores were summed in the same way, divided by the number of people participating in the survey, and average scores were calculated. When the values are negative, it means that that criterion has a negative effect on the average value of the people participating in the survey. After taking the absolute value of the negative values, the ranking process from largest to smallest was carried out.

The pairwise comparison matrices of the AHP analysis were created by considering the results of the frequency analysis. Pairwise comparisons of the criteria were made by determining the importance degrees of the criteria relative to each other according to the frequency analysis ranking. The Pairwise Comparison Matrix (Eq. 2) of AHP was obtained. Pairwise comparison matrices for the main headings of the citizen survey are provided (Tab. 2), along with the steps of the AHP methodology (Tabs. 2-4).

Tab. 2. Pairwise comparison matrix of the main headings

#		A Land Register	B Physical Features	C Locational Features	D Social Features	E Agricultural Features	F Land Features
A	**Land Register**	1	3	1/3	3	2	5
B	**Physical Features**	1/3	1	1/5	2	1/3	3
C	**Locational Features**	3	5	1	5	3	7
D	**Social Features**	1/3	1/2	1/5	1	1/5	2
E	**Agricultural Features**	1/2	3	1/3	5	1	5
F	**Land Features**	1/5	1/3	1/7	1/2	1/5	1
	Total	**5.367**	**12.833**	**2.210**	**16.500**	**6.733**	**23.000**

The total of each column is calculated in the pairwise comparison matrix (Tab. 2). The matrix is normalized by dividing each element by the sum of the columns (Eq. 3). The sum of the rows in the normalized matrix is calculated and divided by the number of main headings (Eq. 4) to determine the weights (W_1) associated with the main headings. The aggregate of the computed weights of the main headings totals 1 (Tab. 3).

Tab. 3. Normalised matrix and weights

#		A	B	C	D	E	F	W_1
A	**Land Register**	0.186	0.234	0.151	0.182	0.297	0.217	**0.211**
B	**Physical Features**	0.062	0.078	0.091	0.121	0.050	0.130	**0.089**
C	**Locational Features**	0.559	0.390	0.453	0.303	0.446	0.304	**0.409**
D	**Social Features**	0.062	0.039	0.091	0.061	0.030	0.087	**0.061**
E	**Agricultural Features**	0.093	0.234	0.151	0.303	0.149	0.217	**0.191**
F	**Land Features**	0.037	0.026	0.065	0.030	0.030	0.043	**0.039**
	Total							**1.000**

To determine the consistency of the pairwise comparison matrix, the consistency ratio (CR) is computed. The pairwise comparison matrix is multiplied by the weight matrix (Eq. 5) to generate matrix **D**, which has dimensions of (nx1). The **D** matrix is divided by the criterion weight to derive the **E** matrix, as shown in Eq. 6. The arithmetic mean of the values in the **E** matrix is computed, followed by the calculation of the λ value as outlined in Eq. 7. The consistency index (CI) is determined using Eq. 8. The consistency ratio (CR) is obtained by dividing the result of Eq. 9 by the Random Index (RI), which is 1.240, as referenced in R. W. Saaty (1987). Since the CR is less than 0.10, the results demonstrate consistency (Tab. 4).

Tab. 4. Calculation of CI and CR

		D	E
A	**Land Register**	1.373	6.500
B	**Physical Features**	0.543	6.129
C	**Locational Features**	2.636	6.446
D	**Social Features**	0.373	6.073
E	**Agricultural Features**	1.199	6.274
F	**Land Features**	0.238	6.165
	Total		**37.587**

RI	**λmax**	**CI**	**CR**
1.240	6. 265	0.053	**0.042 ≤ 0.10**

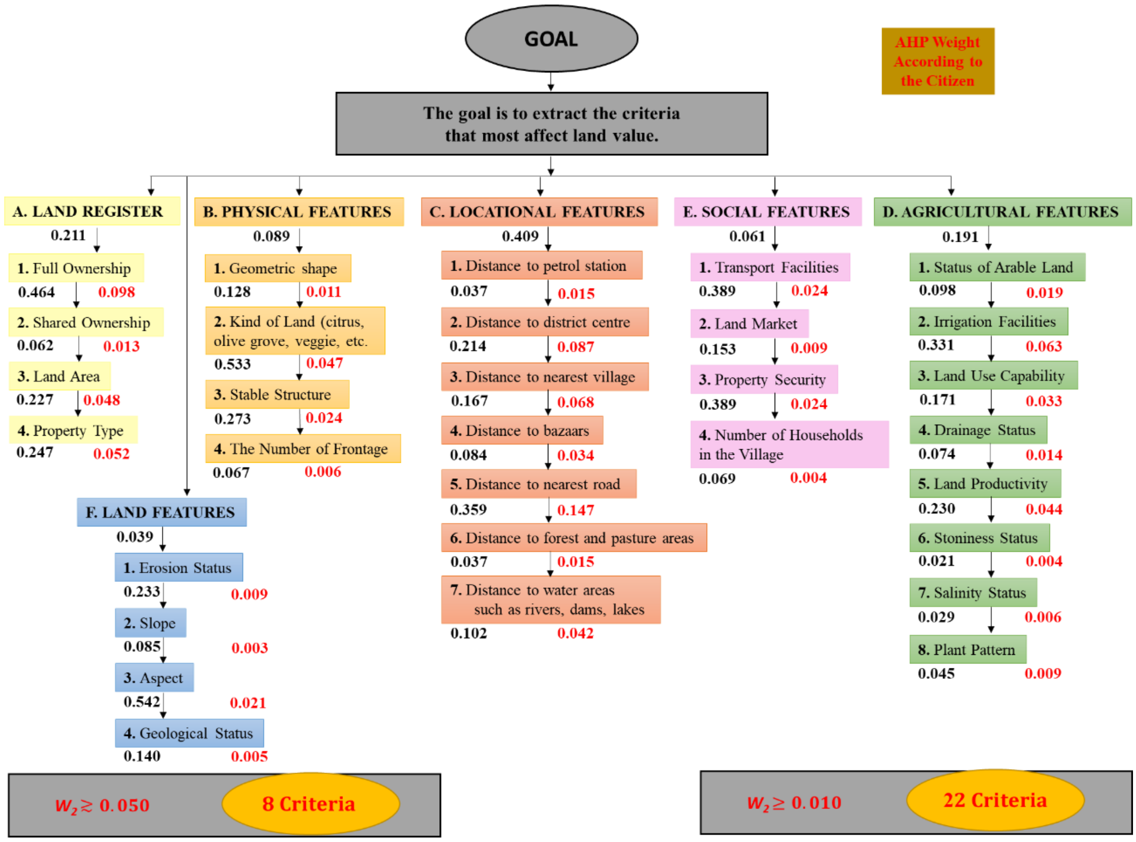

Fig. 5. AHP weight of all criteria according to the citizens

The weights of the main headings were determined, and AHP processes were similarly implemented for all the criteria under the main headings, following these process steps. Once the weights for each criterion were determined within their respective groups, the final weights of the criteria (w_2) were derived by multiplying these by the weights assigned to the main headings. Fig. 5 displays the final weights of the criteria derived from the citizen survey results, with w_2 highlighted in red. The operations for criteria extraction were systematically applied to both the expert and general public survey results, resulting in the calculation of the final weights of the criteria (Appendix A). The consistency of all criteria group operations was verified to ensure that CR is less than or equal to 0.10.

The criteria extraction process was conducted in two groups based on the outcomes of the final AHP weights (w_2 highlighted in red), adhering to the threshold limits of 0.050 and 0.010. I. Extraction: In the initial group [(a) in Tab. 5], the criteria with final AHP weights (w_2) exceeding approximately 0.050 ($w_2 \gtrsim 0.050$) include Full Ownership, Land Area, Property Type, Kind of Land, Distance to district centre, Distance to nearest village, Distance to nearest road, Irrigation Facilities, totalling 8 criteria identified as the most effective. II. Extraction: In the second group [(b) in Tab. 5], a total of 22 criteria exhibited final AHP weights (w_2) exceeding 0.010 ($w_2 \geq$ 0.010). The criteria identified as having a value less than 0.010 ($w_2 <$ 0.010) and demonstrating minimal impact on land value, based on the frequency analysis of the surveyed citizens, include the Number of Frontage, Land Market, Number of Households in the Village, Stoniness Status, Salinity Status, Plant Pattern, Erosion Status, Slope, and Geological Status. Similarly, the criteria extraction process was conducted based on the final AHP weights derived from both expert and general criteria.

Scenarios were developed to validate the criteria derived from the criterion extraction procedure. The scenarios include Scenario 1, encompassing all criteria, Scenario 2 for experts, Scenario 3 for citizens, and Scenario 4, which integrates both expert and citizen perspectives. According to the results of AHP analysis, Scenarios 2, 3, and 4 have two groups, represented as (a) ($w_2 \gtrsim$ 0.050) and (b) ($w_2 \geq$ 0.010). Every case possesses distinct criteria. The parameters for the extraction limit of w_2 weights are detailed in Appendix A. The number of criteria in the criteria analysis across all situations is also displayed in Table 5.

Tab. 5. The number of criteria in analyses and the scenarios

Scenario	Participants	The Number of Criteria in Criteria Analysis			The Number of Criteria in Model Validation Analysis		
Scenario 1	All criteria		31	25	MRA and ANN	22	
			(a) **I. Extraction**	(b) **II. Extraction**		(a) **I. Extraction**	(b) **II. Extraction**
Scenario 2	Expert	Frequency and AHP	7	24	MRA and ANN	6	20
Scenario 3	Citizen	Frequency and AHP	8	22	MRA and ANN	6	17
Scenario 4	General	Frequency and AHP	7	25	MRA and ANN	6	20

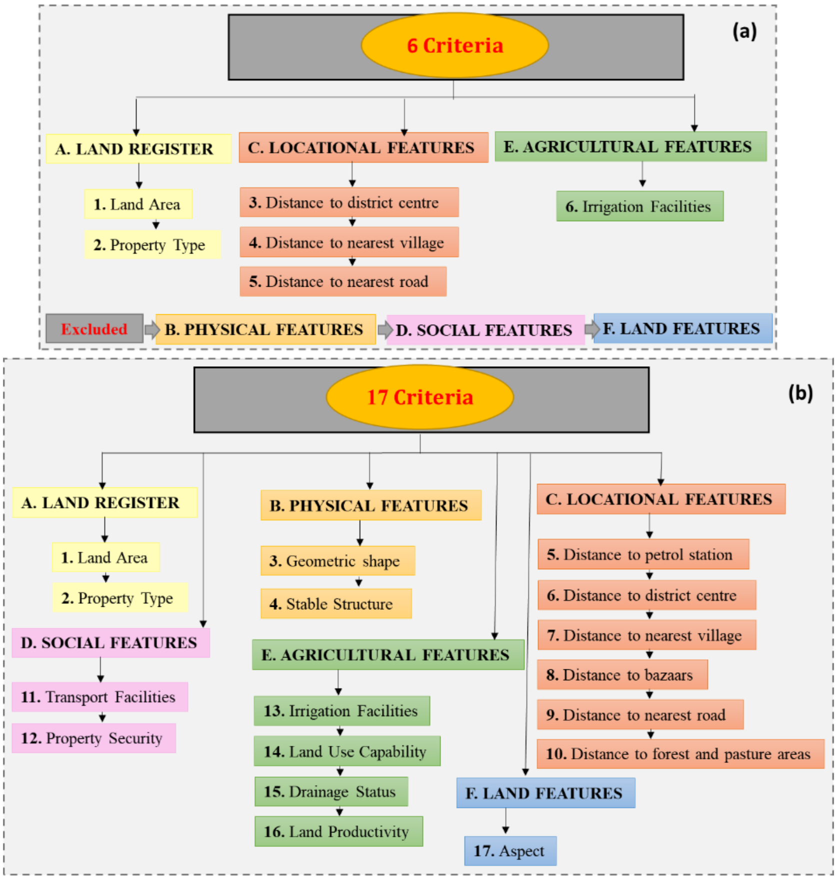

Fig. 6. The results of the criteria analyses

The Model Building

Since there were too many criteria, the criteria were eliminated after the AHP analysis. The criteria affecting the value were established in accordance with the survey opinions of the experts, citizens, and the general public. It is essential to determine which criteria to utilize and the number of criteria to apply in the process. To actualize it, it is essential to identify the criteria group that most accurately predicts land value and effectively elucidates the model. That is why these criteria were used for model validation using MRA and ANN. The results of the performance analysis of the models that were evaluated.

Data corresponding to the formal values were collected for model validation; however, not all data could be attained. Therefore, the criteria extraction process persisted, encompassing the validation phase. A total of 31 criteria were surveyed, with data collected for 25 of these criteria. Multiple Regression Analysis was performed with 25 criteria, but since 3 criteria were 'excluded criteria', model validation analysis was performed with 22 criteria. Consequently, model validation was conducted using 6 criteria and 17 criteria derived from the AHP criteria analysis based on the citizen survey data (Fig. 6). Similar reductions were observed in the other survey data (Tab. 5). Data were collected for each sample real estate, which included 414 2/B agricultural lands corresponding to 25 criteria. The data comprise both text and numerical values, with the numerical values measured in various units, including m, m^2, and %. Before performing value estimation analyses, all of the data should be arranged in a single format so that they are numerical and unitless. The text data were converted into numerical data by scoring, and the numerical data were normalised to the 1-2 range. Thus, the data were made ready for analysis.

After the data organisation process was completed, the data was divided into two groups: training and test data. Of the formal data, 290 training data points (70%) and 124 test data points (30%) were analyzed using valuation methods. Data organization and the separation of training and test data sets were conducted to predict the value of 40 agricultural lands based on 6-17 criteria identified after the model validation phase, 40 agricultural lands corresponding to 17 criteria. Of the market data, 30 training data points, which are approximately 70%, and 10 test data points, which are 30%, were analyzed with valuation methods (Fig. 7).

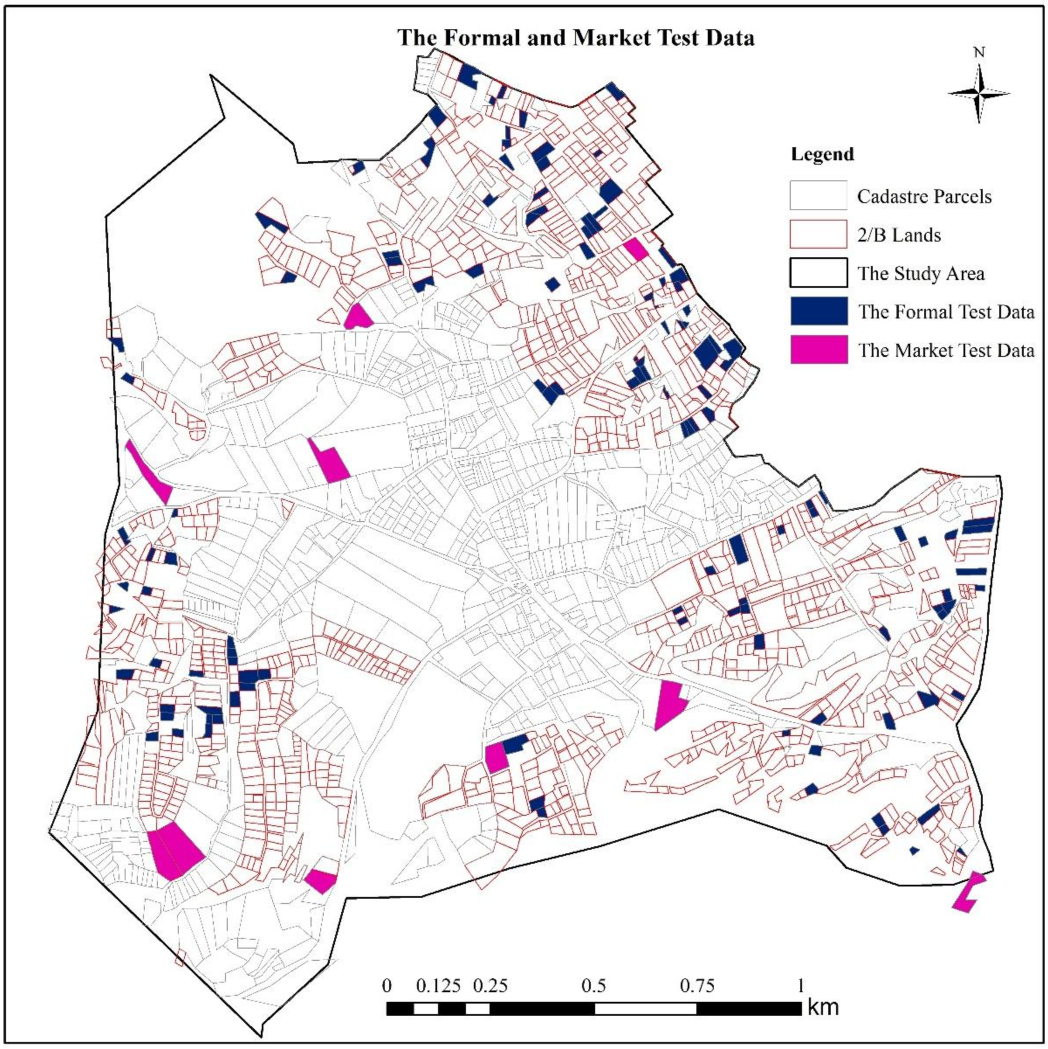

Fig. 7. Distribution of test data in the study area

Model Validation with Formal Values: MRA and ANN, widely used and generally accepted methods, were employed in the model validation process to clarify the extracted criteria. The characteristics of known-value 2/B lands constitute the formal data. The formal value is the dependent variable. The other criteria, which vary depending on the four scenarios, are the independent variables. These were arranged and prepared for the analyses. The number of criteria in Scenarios 1-4 varies according to the expert, citizen, and general public survey results. The models were produced by using the formal data belonging to the criteria in Scenarios 1-4. In the MRA, the mathematical model was produced with training data. The formal values were predicted using the model with both training and test data (Appendix B). Similarly, ANN was analysed using training data and predicted with the help of test data. For four scenarios, model validation was performed by examining performance analyses, which are R^2, MAE, MAPE, and MSE. Given that the training data were used to construct the prediction model, the R^2 value was utilized to determine the degree to which the model explained the data. To conduct an objective assessment of the test data's success, it was necessary to analyze the MAE, MAPE, and MSE values of the predictions (Tab. 6).

Tab. 6. Performance analyses of scenarios according to models

Scenario	Group	The Number of Criteria	Model	Training Data	Performance Analyses of Testing Data		
				R^2	MAE	MAPE	MSE
Scenario 1	all	22	MRA	0.604	0.043	0.038	0.006
			ANN	0.624	0.064	0.057	0.013
Scenario 2	a	6	MRA	0.106	0.059	0.051	0.007
			ANN	0.212	0.059	0.051	0.007
	b	20	MRA	0.572	0.043	0.038	0.006
			ANN	0.609	0.082	0.074	0.019
Scenario 3	***a***	***6***	***MRA***	***0.548***	***0.044***	***0.039***	***0.006***
			ANN	***0.609***	***0.049***	***0.043***	***0.010***
	b	**17**	**MRA**	**0.576**	**0.044**	**0.039**	**0.006**
			ANN	**0.690**	**0.078**	**0.070**	**0.018**
Scenario 4	a	6	MRA	0.037	0.058	0.051	0.007
			ANN	0.128	0.056	0.047	0.008
	b	20	MRA	0.596	0.043	0.038	0.006
			ANN	0.578	0.079	0.071	0.019

The model performance of Scenario 3, which has the least number of criteria, was scrutinised. Among these models, the performance analysis results of the ANN method of Scenario 3b are better. Given that the R^2 value (0.690) is close to 1, and the ratios of MAE (0.078), MAPE (0.070), and MSE (0.018) indicate that the model has a low ratio. As a final result, the criteria in Scenario 3, which is the best model, were extracted and verified as the criteria affecting the land value.

Production Model for Comparison of the Formal and Market Values: The final criteria affecting the land value are those belonging to Scenarios 3a and 3b, which are the most successful in the model validation. The formal data and market data that belong to the 6 (a) and 17 (b) criteria to be used in mass land valuation were prepared for analysis. Since formal data consists of 2/B agricultural lands and market values consist of privately owned agricultural lands, these two groups have different real properties. Therefore, a one-to-one comparison cannot be made. To make a comparison, Scenario 5, 6, 7, and 8 models were produced with 6 criteria (a) and 17 criteria (b). In Scenario 5, the formal training and test data set from the criteria decided in line with the citizen survey as a result of the evaluations made; in Scenario 6, market training and test data set collected from the market; in Scenario 7 and 8, the formal training data set and market test data set were used (Tab. 7).

Tab. 7. Comparison scenarios

Scenario	The Number of Criteria		Training Data		Test Data		Model Analysis
	(a)	(b)					
Scenario 5	6	17	Formal Values	290	Formal Values	124	RIDGE/RF
Scenario 6	6	17	Market Values	30	Market Values	10	RIDGE/RF
Scenario 7	6	17	Formal Values	290	Market Values	40	RIDGE/RF
Scenario 8	6	17	Formal Values	414	Market Values	40	RIDGE/RF

The comparison models are generated using Ridge and RF methods, which are capable of conducting robust analyses with small data input. Comparison models with Scenarios 5, 6, 7, and 8 were generated using training data with the final criteria. With these models, land values were estimated using test data. Performance analyses utilized R^2 for model explanation and LOOCV MAE, MAPE, and MSE for error assessment. The closest model training data' R^2 (0.742 in Ridge and 0.739 in RF) to 1 is **Scenario 6b** using market data. On the other hand, the test results of the Ridge model in **Scenario 5b**, where the formal values are used and the error rates are the lowest in the form of LOOCV MAE (0.044), MAPE (0.039), and MSE (0.005) (Tab. 8).

Tab. 8. The results of performance analysis

Scenario		**Model**	Train	**Performance Analyses of Testing Data**		
			R^2	**LOOCV MAE**	**LOOCV MAPE**	**LOOCV MSE**
Scenario 5	a	RIDGE	0.533	0.044	0.039	0.006
		RF	0.591	0.047	0.041	0.005
	b	RIDGE	0.559	0.044	0.039	0.005
		RF	0.624	0.046	0.041	0.005
Scenario 6	a	RIDGE	0.644	0.190	0.141	0.070
		RF	0.696	0.236	0.176	0.092
	b	RIDGE	0.742	0.211	0.156	0.071
		RF	0.739	0.235	0.173	0.089
Scenario 7	a	RIDGE	0.531	0.212	0.150	0.075
		RF	0.591	0.210	0.147	0.076
	b	RIDGE	0.560	0.220	0.156	0.079
		RF	0.624	0.212	0.150	0.077
Scenario 8	a	RIDGE	0.454	0.212	0.149	0.077
		RF	0.521	0.210	0.146	0.078
	b	RIDGE	0.487	0.221	0.156	0.082
		RF	0.546	0.212	0.148	0.079

Maps of Predicted Values with the Formal and Market Data: Since the formal and market values belong to different real properties, it is necessary to utilise the data under the same conditions for the comparison in the study area. For this reason, the processes were continued over the predicted values, where the formal data belonging to Scenario 5 and the market data belonging to Scenario 6 were tested. Predicted values were denormalised and converted into Turkish Liras. The visualization was achieved by producing value maps via Inverse Distance Weighting (IDW) in ArcGIS.

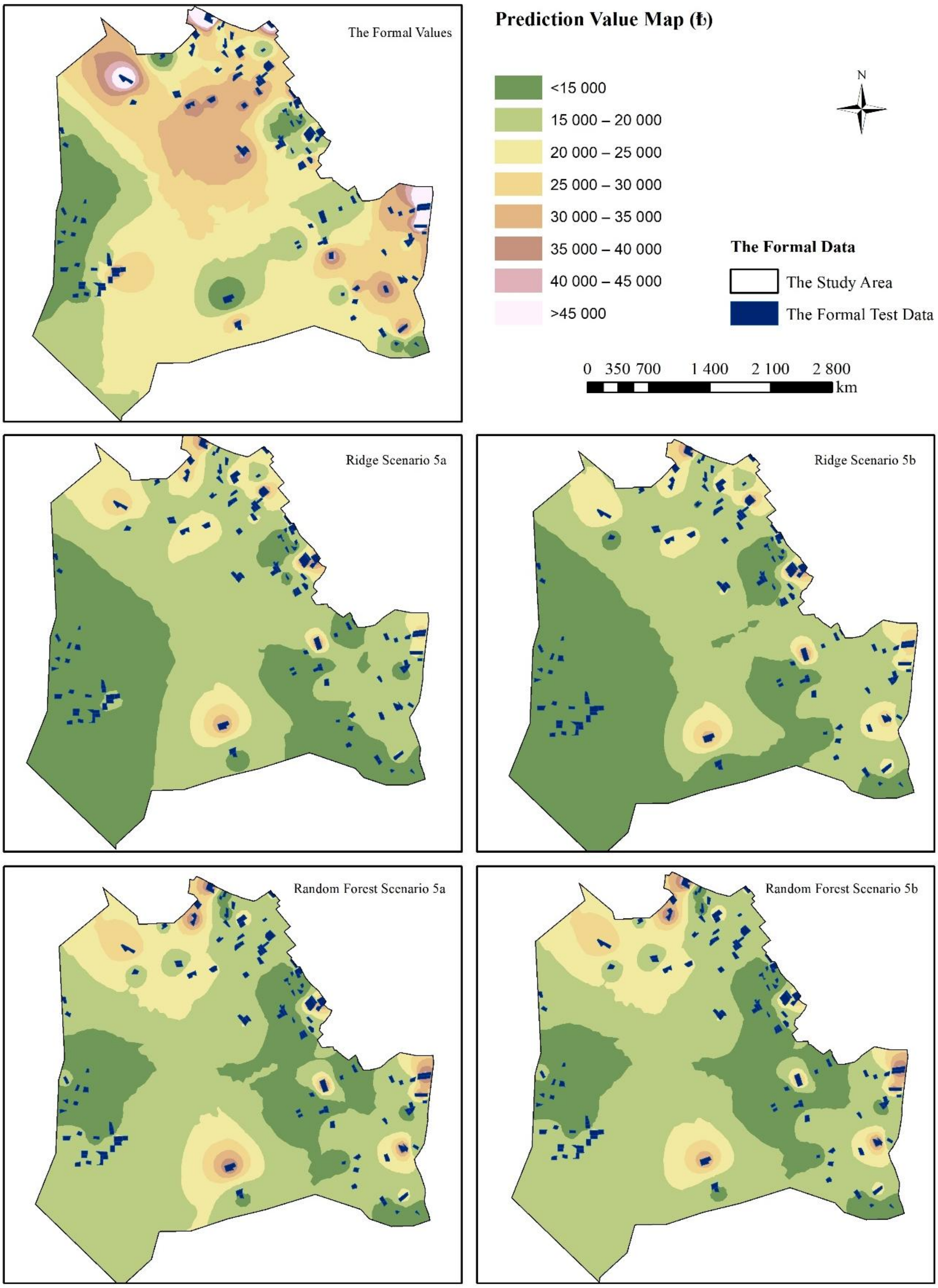

Fig. 8. The formal value and predicted values maps (Scenario 5)

With the help of the formal values in Scenario 5a and 5b, the statistical distribution of the values in the whole study area was visualised. Both the similarities of the maps were analysed, and the formal values corresponding to the samples to be taken for comparison were found. It is observed that the similarities between the formal value maps and the value maps estimated by the Ridge and RF methods are low. However, the RF method map shows a high similarity to the formal value map in the value ranges of less than 15,000 and 15,000 - 20,000 (Fig. 8).

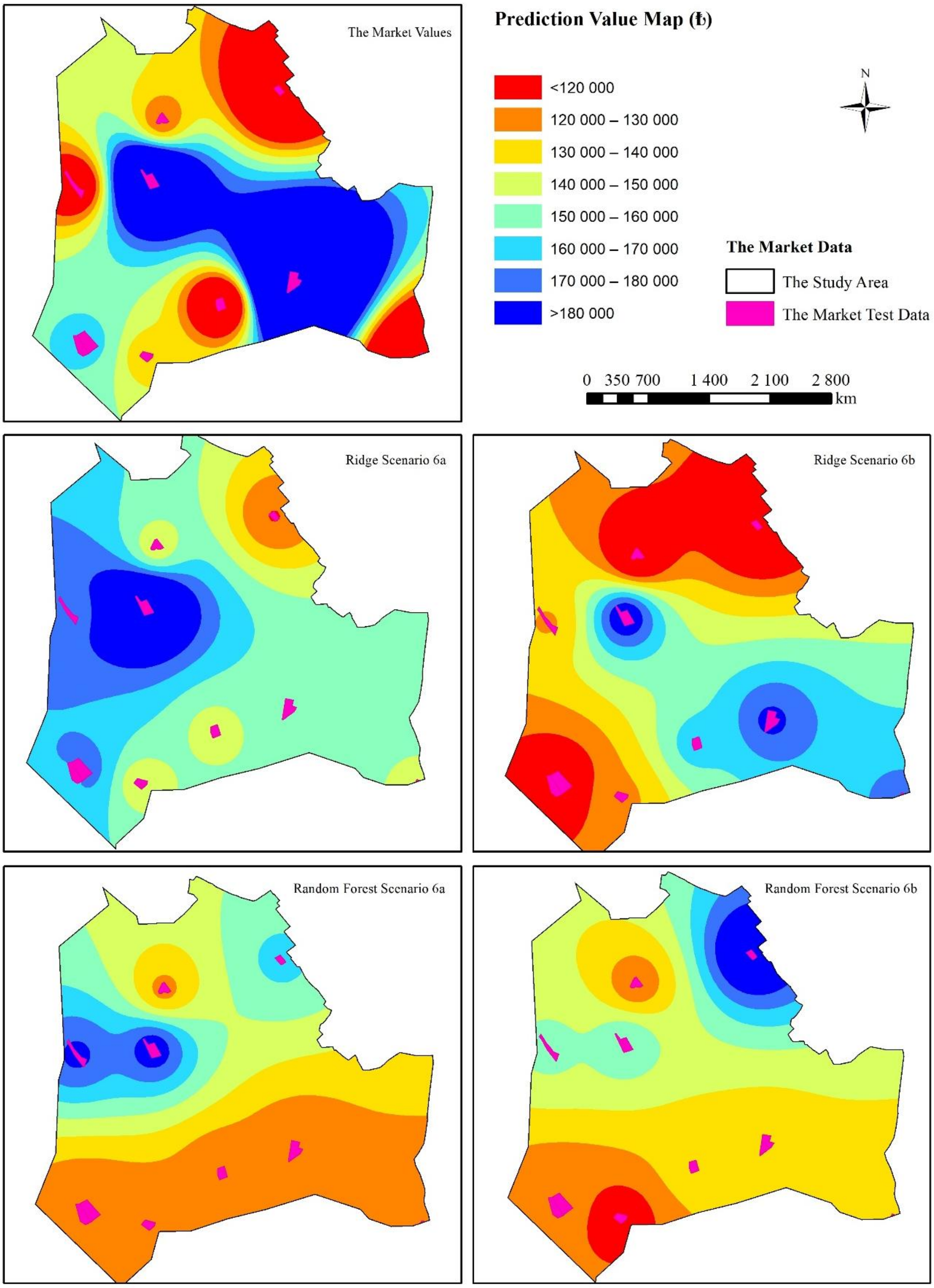

Fig. 9. The market value and predicted values maps (Scenario 6)

With the help of the market values in Scenario 6a and 6b, the statistical distribution of the values in the whole study area was visualised. Both the similarities of the maps were analysed, and the market values corresponding to the samples to be taken for comparison were found. It is observed that the similarities between the market value maps and the value maps estimated by the Ridge and RF methods are close. Most likely, the market value map is Ridge Scenario 6b (Fig. 9).

The Formal and Market Values Comparing

It is not possible to find both the formal and market value of the same parcel in the study area. For this reason, the estimated value maps were converted into raster format. In the study area, 380 sample points were randomly assigned to ensure a homogeneous distribution to compare values. The point data were matched with the pixel value of the value map in raster format. The value equivalents of all value maps were printed in the attribute table of the point (Fig. 10).

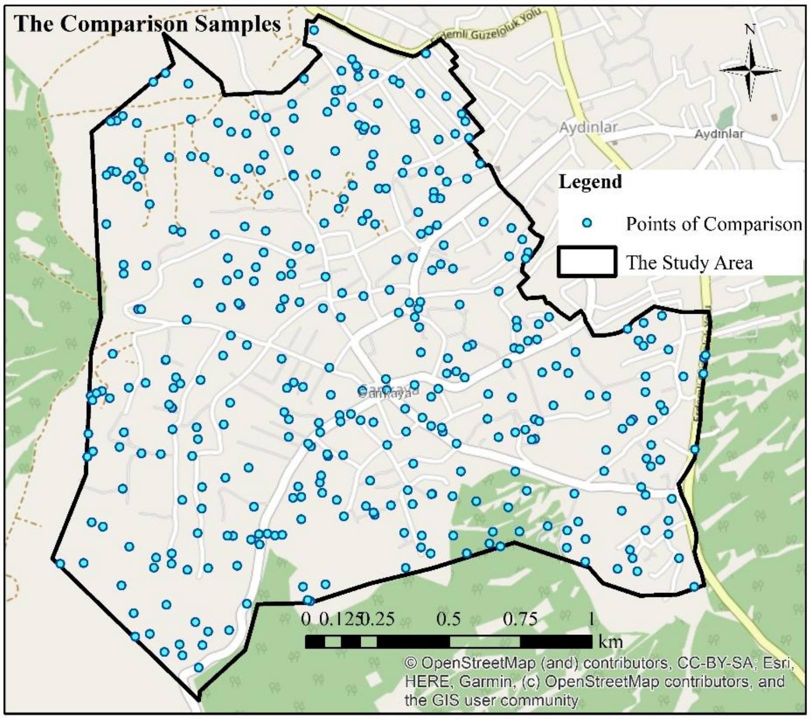

Fig. 10. Comparison samples

Normality test: Data groups were subjected to a normality test for the analysis of differences. The data group corresponding to each point was converted into Excel format, and the test was performed. Kolmogorov-Smirnov and Shapiro-Wilk tests were used since the number of data sets was 30 or more (Bayrak, 2021).

The comparison samples consist of a total of 380 points. Formal and market data are available corresponding to these points. The dataset comprises 380 formal data points and 380 market data points, resulting in a total of 760 data points. If the level of significance (Sig.) in the Kolmogorov-Smirnov column and Shapiro-Wilk column given in Table 9 is at $p < 0.05$, it is significant. The null hypothesis is rejected. For this reason, it is concluded that the distribution of the data is not normal. In other words, the fact that this value is significant in the test means that the data is not normally distributed. A permutation test, a non-parametric test, was performed to analyze data groups that were not normally distributed.

Tab. 9. Normality test

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Test Statistic	The Sample	Sig.	Test Statistic	The Sample	Sig.
The Formal-Market Values	0.248	760	0.000	0.854	760	0.000
The Formal-Market Ridge (5a-6a)	0.294	760	0.000	0.754	760	0.000
The Formal-Market Ridge (5b-6b)	0.281	760	0.000	0.789	760	0.000
The Formal-Market RF (5a-6a)	0.280	760	0.000	0.769	760	0.000
The Formal-Market RF (5b-6b)	0.284	760	0.000	0.765	760	0.000

Permutation Test Application: The comparison of the formal value, market value, and predicted values will be carried out with 380 sample points of data taken from the value maps.

Groups and hypotheses to be tested;

$\boldsymbol{M1}$**:** Median parameter of the first group method

$\boldsymbol{M2}$**:** Median parameter of the second group method

$\boldsymbol{H_0}$: There is no difference between the formal and market values.

$\boldsymbol{H_1}$**:** There is a significant difference between the formal and Market values.

$$H_0 : M1 = M2$$
$$H_1 : M1 < M2$$

The decision rule is determined according to the hypothesis H_1 with $\propto$ being the significance level.

- If the hypothesis H_1 is one-way, p = ExactSig.(1-tailed), H_0 is rejected if $p < \propto$, H_0 cannot be rejected if $p \geq \propto$.
- Comments are made in accordance with the decision.

As a result of the Permutation Test, group means and medians were found (Tab. 10):

Analysis of the Formal and Market Values reveals that the negative average indicates that the formal values in Group 1 are roughly 138,000 times less than the market value in Group 2. The mean and median differences are quite close to each other. Cohen's d (-3.683) and Rank-biserial correlation (1.000) indicate significant differences between the groups. The final decision: Since $\alpha = 0.05$ and p-value (2-tailed), $p = 0.000$ and $p < \propto$, H_0 is rejected. This means that there is a difference between the two values (the formal value and market value). As a result of the test, it is concluded that there is a statistically significant difference between the two data groups (Tab. 10).

Tab. 10. The results of the Permutation Test

Group1: Formal Group2: Market	The Formal-Market Values	Predicted Values of the Formal and Market			
		Ridge (5a-6a)	Ridge (5b-6b)	RF (5a-6a)	RF (5b-6b)
Mean difference	-138952.590	-142506.497	-124773.609	-121931.057	-125052.627
Median difference	-130114.050	-138847.650	-124500.750	-119445.950	-121999.000
Cohen's d (effect size)	-3.683	-11.747	-8.059	-10.531	-10.636
Rank-biserial correlation	1.000	1.000	1.000	1.000	1.000
Permutation statistic	-138952.590	-142506.497	-124773.609	-121931.057	-125052.627
p-value (2-tailed)		0.000	0.000	0.000	0.000

Discussion

This study seeks to evaluate the efficacy of current valuation methodologies by examining the disparity between the formal values of 2/B lands that have forfeited their forest designation and the market sales values of agricultural lands. The permutation test findings indicate a statistically significant difference between the two groupings (Tab. 10). This indicates that the officially assessed 2/B land prices, including the disparity between mortgage and real values (Seifert & Hüttel, 2020) and the juxtaposition of sales value and tax value (Ma & Swinton, 2012), fail to represent market reality accurately. Consequently, it can be inferred that formal values are not applicable to the value in the tax base. In this context, TKGM has continued its valuation activities by establishing a department within its body (TKGM, 2019) and performed the Land Registry Cadastre Modernisation Project with the support of the World Bank Sustainable Development Department (TKMP, 2014).

The inclusion of the special situation surrounding 2/B agricultural land complicates the study and transaction process due to significant restrictions. The special situation surrounding 2/B agricultural land causes the difficulty of finding healthy and up-to-date data from the market and the complexity of criteria linked to value data (Choumert & Phélinas, 2015; Unel & Yalpir, 2019; Yalpir & Unel, 2022) Consequently, the sample size was constrained, and it was noted that the results were significantly influenced by this limitation. The statistical power of the methodologies employed in the analysis enhanced the reliability of the findings.

There are many criteria affecting the land value. In one of them (Demetriou, 2016), analysis results clearly demonstrated that eight out of fourteen land valuation criteria related to location characteristics, legal, physical, and economic conditions are the most important. As irrigation Facilities, Sampson et al. (2019) found that the agricultural land values are 53% higher for irrigated parcels than non-irrigated parcels on average in Kansas. For real properties that are not registered in the land registry, up to 100 acres of dry land and 40 acres of irrigated land shall be determined in the name of the owner who proves with documents that he has been in possession as owner for at least twenty years without dispute and without interruption (Cadastre Law, 1987, 14). In other words, irrigated land seems to be 2.5 times more valuable than dry land according to the Cadastre Law. These highlight the significance of irrigation facilities in agricultural features.

This research employed a systematic questionnaire survey to derive valuation criteria from stakeholder feedback, distinguishing it from other investigations. The established criteria were then assessed and weighted by AHP (Fig. 5), facilitating a systematic prioritization. To assess the prediction efficacy of the constructed models, MRA and ANN were employed. This study further distinguishes itself by comparing formal and market-based land values. To examine any disparities between these two value sets, the normality of the data distribution was initially evaluated, and owing to the non-normal characteristics of the data, a Permutation Test was performed.

Due to the constraints of the special situation surrounding 2/B land and the paucity of data in rural regions, some difficulties were experienced in the model validation and analysis phase for predictive values. Notwithstanding this, Scenario 3 had superior performance in model validation. The primary reason identified was the valuation commission's determination of area-based fair value (Erdönmez, 2013) with traditional valuation methods (Çinar & Ünel, 2022). Due to the limited availability of market data, Ridge and RF analysis methods, which are statistically suitable for the number of observations according to the number of variables—the ratio of observations to variables —were utilized. The performance study revealed that the LOOCV error results were within an acceptable range (Tab. 8). The same group training and test data in Scenarios 5 and 6 were utilized to compare the formal value with the market value. Assuming the formal values were utilized as the tax base value, the prediction market values from the models of the formal values were conducted in Scenarios 7 and 8. The performance analysis results indicate a need for improvement. Therefore, the value maps in Figs 8 and 9 were produced according to the results of Scenarios 5 and 6 (Demetriou, 2018; Uberti et al., 2018). The value maps enable the juxtaposition of formal and market values and assess the value of every land within the study area via IDW. In addition, the pixel values associated with the formal and market values served as the foundation for statistical comparison.

The permutation test indicates a statistically significant disparity between the Formal and Market Values and the expected values. The proximity of the mean and median differences indicates the absence of extreme values in the distribution, suggesting a balanced distribution. Cohen's d (-3.683) signifies a very significant effect size, as per the threshold value established by Cohen (1988). The rank-biserial correlation of 1.000 indicates an almost total separation between the two value distributions. The results underscore the necessity for a reevaluation of valuation procedures in policy contexts pertaining to 2/B land or analogous state-managed ownership regimes, as the formal values established by the public agency do not consistently align with market-based values.

This study utilizes region-specific data and legislative frameworks; however, the statistical procedures of criterion analysis, model validation, value estimation, and benchmark testing are universally applicable and adaptable to diverse geographical and institutional situations. This study offers a reproducible methodology for comparing government and market-based land values to formulate strategies and policies for land planning in various nations experiencing land reform. Especially, in agricultural land valuation studies, aggregate valuation may be conducted utilizing reduced criteria. Consequently, with local calibration, it can facilitate universal applications in land value policy and practice.

Conclusion

Valuation of agricultural lands is an important issue in transactions such as inheritance sharing, taxation, insurance, and the determination of irrigation fees. Values of agricultural lands vary according to criteria such as topographic structure, productivity, irrigability, proximity to the city, etc. With these characteristics of agricultural lands, incorporating precise and current land values into decision-making processes would facilitate more rational planning, thereby enhancing efficient and effective land utilization. Agricultural land values are considered a basis for planning issues to prevent land wastage and underutilisation. Thus, the study's results endorse the formulation of more sustainable land management plans grounded in objective and data-driven valuation methodologies.

In this study, the criteria determination, data collection, data arrangement, normalisation, method application, and performance analyses were carried out with considerable effort. Since the criteria affecting agricultural land values are not clear and explicit, they were determined in the light of literature research and expert experience. These criteria are grouped under certain main headings and will serve as a basis for other studies. The questionnaire survey was conducted with citizens and experts who are important actors and users of real estate. According to the questionnaire results, frequency analyses of the criteria were conducted, and their averages were determined. Based on these findings, the criteria were weighted using the AHP method. The AHP weights of the criteria were examined, and those below 0.050 and 0.010 units were eliminated. Two different elimination processes were performed. The criteria whose value was affected as expert, citizen, and general public were obtained.

The model validation analyses were performed with the help of 8 models in total by using 4 scenarios as all criteria (Scenario 1), expert (Scenario 2), citizen (Scenario 3), and general public (Scenario 4), and by MLR and ANN valuation models. Performance analyses were realised with the estimation results of the formal and market values. According to the performance analysis results, both the MLR and ANN models in Scenario 3 yielded better results than the others. To compare formal and market values, the models must have the same number of criteria and data standards. Therefore, a total of 17 criteria were established. Subsequently, Scenario 5 and Scenario 6 were created using the data that fell within these criteria. The market data were incorporated into the model derived from the formal values, yielding Scenarios 7 and 8 as a hybrid model. The scenarios were analyzed using Ridge and RF methods, considering the specific nature of the real estate and the constraints of the available data. Value maps were produced with the results of the scenarios. Data from 380 points with homogeneous distribution in the value maps were obtained and compared to analyze whether the difference between the values was significant. Non-normally distributed data were analysed with the non-parametric permutation test, and it was concluded that there was a difference between the two data groups.

Clear and transparent value maps are essential for good governance and sustainable land management. Fair and equitable distribution of the rapid change in land use, driven by population growth, will be possible through the value-based implementation of urban transformation, land and land arrangements, land consolidation, etc. The value is the numerical expression of the corresponding to many criteria, and the fact that it reflects a whole should also be taken into consideration. Data sharing between institutions should be ensured with protocols, and more models should be developed through projects involving universities. For these processes, legal legislation explaining the technical application for mass valuation should be written for Turkiye. In this context, the criteria analyses should be renewed. It is suggested to produce better models by repeating the analyses of the data with different modern mass valuation methods, such as fuzzy logic, genetic algorithm, and fuzzy multi-criteria decision analysis. In addition, it can be tested whether there is a difference between the market value and the formal values, such as tax value, expropriation value, insurance value, etc., that are currently used. The increase in the difference in value ranges reveals the need for more research and model production on valuation. Analyses should be continued until this difference decreases. It should be planned to perform land classifications on a geometric basis with the criteria, which are the basic building blocks of agricultural lands, and to make groupings according to marginal, absolute, special crop, planted, and under-cover agricultural land types. Final criteria can be determined through the valuation of agricultural lands, enabling value-based pilot applications of land consolidation projects.

Appendix

A. Average (Avg.) of frequency analysis and W_2 AHP weights

The Criteria	EXPERTS		CITIZEN		GENERAL PUBLIC	
	Avg. of Frequency Analysis	W_2 AHP Weights	Avg. of Frequency Analysis	W_2 AHP Weights	Avg. of Frequency Analysis	W_2 AHP Weights
Full Ownership	3.52	0.024	3.50	0.098	3.50	0.105
Shared Ownership	-1.06	0.004	-2.21	0.013	-2.07	0.010
Land Area	2.61	0.010	2.60	0.048	2.60	0.030
Property Type	3.00	0.016	2.73	0.052	2.76	0.047
Geometric Shape	2.74	0.037	2.63	0.011	2.64	0.015
Kind of Land	2.48	0.013	2.99	0.047	2.93	0.039
Stable Structure	2.68	0.021	2.86	0.024	2.84	0.022
The Number of Frontage	2.81	0.062	2.55	0.006	2.59	0.010
Distance to Petrol Station	2.29	0.014	1.45	0.015	1.55	0.016
Distance to District Centre	2.55	0.029	2.85	0.087	2.82	0.108
Distance to Nearest Village	3.06	0.069	2.71	0.068	2.75	0.069
Distance to Bazaars	2.42	0.023	2.20	0.034	2.23	0.031
Distance to Nearest Road	3.45	0.099	3.88	0.147	3.83	0.134
Distance to Forest and Pasture Areas	2.23	0.007	1.45	0.015	1.54	0.016
Distance to Water Areas	2.29	0.011	2.55	0.042	2.51	0.044
Transport Facilities	3.39	0.008	3.20	0.024	3.22	0.025
Land Market	1.84	0.005	1.65	0.009	2.87	0.008
Property Security	3.58	0.014	3.15	0.024	3.21	0.018
Number of Households in the Village	2.16	0.002	1.77	0.004	1.82	0.004
Status of Arable Land	2.94	0.050	2.79	0.019	2.81	0.024
Irrigation Facilities	3.9	0.148	4.00	0.063	3.99	0.073
Land Use Capability	3.32	0.077	3.09	0.033	3.12	0.035
Drainage Status	1.9	0.031	2.19	0.014	2.16	0.015
Land Productivity	3.71	0.106	3.15	0.044	3.22	0.048
Stoniness Status	-0.23	0.010	-0.77	0.004	-0.71	0.004
Salinity Status	-0.16	0.012	-0.99	0.006	-0.89	0.006
Plant Pattern	1.84	0.021	1.65	0.009	1.68	0.010
Erosion Status	-1.65	0.004	-1.32	0.009	-1.36	0.010
Slope	-0.94	0.009	-0.91	0.003	-0.91	0.003
Aspect	2.48	0.039	2.89	0.021	2.84	0.019
Geological Status	1.26	0.022	1.02	0.005	1.05	0.004

B. MRA coefficients of four scenarios

No	The Criteria	Scenario 1	Scenario 2		Scenario 3		Scenario 4	
			a	b	a	b	a	b
	The Formal Value	The Dependent Variable						
					β			
	Constant	0.149	1.104	0.660	0.556	0.865	1.089	0.340
1	**Land Area**	0.560	-	0.547	0.549	0.551	-	0.559
2	**Geometric Shape**	-0.004	-	0.002	-	0.001	-	-0.002
3	**Property Type**	0.012	-	0.009	0.013	0.008	0.023	0.007
4	**Stable Structure**	-0.009	-	-0.007	-	-0.008	-	-0.008
5	**The Number of Frontage**	-0.011	0.238	-0.025	-	-	-	-0.007
6	**Distance to Petrol Station**	0.073	-	-0.034	-	-0.069	-	0.145
7	**Distance to District Centre**	0.048	-	0.068	0.010	0.058	0.066	0.031
8	**Distance to Nearest Village**	0.018	-0.040	-0.047	-0.080	-0.033	-0.052	-0.001
9	**Distance to Bazaars**	0.031	-	-0.085	-	-0.095	-	0.093
10	**Distance to Nearest Road**	0.000	-0.007	0.002	-0.014	-0.008	0.011	-0.004
11	**Distance to Forest and Pasture Areas**	-0.045	-	-	-	-0.020	-	-0.041
12	**Transport Facilities**	0.047	-	-	-	0.035	-	0.043
13	**Land Market**	0.043	-	-	-	-	-	-
14	**Property Security**	-0.032	-	0.009	-	0.005	-	-0.011
15	**Irrigation Facilities**	-0.011	-0.020	-0.003	0.010	-0.017	-0.030	-0.013
16	**Land Use Capability**	-0.146	-0.074	-0.030	-	-0.072	-	-0.111
17	**Drainage Status**	-0.019	-	-0.019	-	-0.020	-	-0.054
18	**Land Productivity**	-	-0.031	-	-	-0.036	-	-
19	**Stoniness Status**	-	-	-	-	-	-	-
20	**Salinity Status**	0.108	-	-0.003	-	-	-	-
21	**Plant Pattern**	-	-	-	-	-	-	-0.088
22	**Erosion Status**	0.085	-	-	-	-	-	0.082
23	**Slope**	0.001	-	-	-	-	-	-
24	**Aspect**	-0.005	-	-0.008	-	-0.005	-	-0.002
25	**Geological Status**	0.013	-	0.025	-	-	-	-
	R	0.777	0.326	0.756	0.740	0.759	0.192	0.772
	R^2	0.604	0.106	0.572	0.548	0.576	0.037	0.596
	Adjusted R^2	0.571	0.087	0.545	0.538	0.550	0.016	0.568
	Std. Error of the Estimate	0.066	0.096	0.068	0.068	0.068	0.100	0.066
	F-test	18.502	5.595	21.397	57.077	21.741	1.797	20.984
	F-test' Sig.	0.000	0.000	0.000	0.000	0.000	0.100	0.000

References

Abdulhafedh, A. (2022). Incorporating multiple linear regression in predicting the house prices using a big real estate dataset with 80 independent variables. Open Access Library Journal, 9(1), 1-21.

Açlar, A., & Çağdaş, V. (2008). Taşınmaz (gayrimenkul) değerlemesi (2. Baskı). TMMOB, Harita ve Kadastro Mühendisleri Odası (HKMO).

Adil, M., & Syafaruddin, S. (2024). Property Taxes and Their Implications on the Real Estate Market: A Literature Review. Advances in Taxation Research, 2(2), 65-77. https://doi.org/10.60079/atr.v2i2.281

Agricultural Reform Law, (1984). Pub. L. No. 3083, Sulama Alanlarında Arazi Düzenlenmesine Dair Tarım Reformu Kanunu (Agricultural Reform Law on Land Regulation in Irrigation Areas,), 18592.

Akdağ, M. (2011). SPSS'de İstatistiksel Analizler. https://avesis.inonu.edu.tr/mustafa.akdag/dokumanlar

Akış, B. (2013). İstatistiki Yöntemlerle Değer Belirleme ve Değer Haritası Üretimi Selçuklu Örneği. Selçuk Üniveristesi, Fen Bilimleri Enstitüsü.

Alkan, T., & Durduran, S. S. (2024). Türkiye'de Tarım Arazilerinin Değerlemesine İlişkin Genel Bir Değerlendirme. Osmaniye Korkut Ata Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 7(2), 953–972. https://doi.org/10.47495/okufbed.1332617

Alpaslan, H. İ. (2015). TMS/TFRS Kapsamında Gayrimenkul Değerleme ve Gerçeğe Uygun Değerin Tespitinde Emsal Karşilaştırma ve Gelir İndirgeme Yöntemleri Üzerine bir Uygulama. Nevşehir Hacı Bektaş Veli Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 4, 17–35.

Alsahan, I. M., & AlZaidan, Z. I. (2024). Unleashing the Power of Artificial Intelligence in Real Estate Valuation: Opportunities and Challenges Ahead. Journal of Artificial Intelligence General science (JAIGS), ISSN: 3006-4023, 3(1), 89-96.

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 systems with applications, 39(2), 1772-1778.

Antipov, E. A., & Pokryshevskaya, E. B. (2017). Are box office revenues equally unpredictable for all movies? Evidence from a random forest-based model. Journal of Revenue and Pricing Management, 16, 295-307.

Arslan, K. (2018). SPSS'de Bağımsız Örneklem T-Testi (Independent Sample T-Test). https://www.galloglu.com/blog/bagimsiz-orneklem-t-test-SPSS-independent-sample-t-test

Asiama, K. O., Bennett, R., Zevenbergen, J., & Asiama, S. O. (2018). Land valuation in support of responsible land consolidation on Ghana's rural customary lands. Survey Review, 50(361), 288–300. https://doi.org/10.1080/00396265.2018.1467672

Aydın, C. (2018). Makine Öğrenmesi Algoritmaları Kullanılarak İtfaiye İstasyonu İhtiyacının Sınıflandırılması. European Journal of Science and Technology, 14, 169–175. https://doi.org/10.31590/ejosat.458613

Aydinoglu, A. C., & Sisman, S. (2024). Comparing modelling performance and evaluating differences of feature importance on defined geographical appraisal zones for mass real estate appraisal. Spatial Economic Analysis, 19(2), 225-249.

Bahl, R. & Martinez-Vazquez, J. (2007). The Property Tax in Developing Countries: Current Practice and Prospects, Lincoln Institute of Land Policy, WP07RB1.

Bayar, R. (2018). Arazi kullanımı açısından Türkiye'de tarım alanlarının değişimi. Coğrafi Bilimler Dergisi, 16(2), 187–200. https://doi.org/10.1501/cogbil_0000000197

Bayrak, H. (2021). SPSS ile Veri Analizi—Normal Dağılım Testleri. https://dijilopedi.com/spss-ile-veri-analizi-normal-dagilim-testleri/

Bertanha, M. & Chung, E. (2023). Permutation Tests at Nonparametric Rates. Journal of the American Statistical Association, 118:544, 2833-2846, DOI:10.1080/01621459.2022.2087660

Biau, G. (2012). Analysis of a Random Forests Model. Journal of Machine Learning Research, 13, 1063–1095.

Biau, G., & Scornet, E. (2016). A random forest guided tour. Test, 25(2), 197-227.

Bird, R. M., & Slack, E. (2004). International Handbook of Land and Property Taxation. Edward Elgar Publishing Limited, UK. https://doi.org/10.4337/9781845421434

Bisi, M., & Goyal, N. K. (2017). Artificial Neural Network for Software Reliability Prediction. 1st ed., Wiley. https://doi.org/10.1002/9781119223931

Bonnini, S., Assegie, G.M., & Trzcinska, K. (2024). Review about the Permutation Approach in Hypothesis Testing. Mathematics, 12, 2617. https://doi.org/10.3390/math12172617

Breiman, L. (1996). Bagging Predictors, Machine Learning. Kluwer Academic Publishers, 24, 123–140.

Breiman, L. (2001). Random Forests. Kluwer Academic Publishers, Machine Learning, 45, 5–32.

Breiman, L. (2004). Consistency for A Simple Model of Random Forests. Statistics Department University of California at Berkeley, Technical Report 670.

Burger, A. (1998). Land valuation and land rents in Hungary. Land Use Policy, 15(3), 191–201. https://doi.org/10.1016/S0264-8377(98)00013-1

Cadastre Law, (1987). Pub. L. No. 3402, Kadastro Kanunu (Cadastre Law).

Carbonara, S., Faustoferri, M., & Stefano, D. (2021). Real Estate Values and Urban Quality: A Multiple Linear Regression Model for Defining an Urban Quality Index. Sustainability, 13, 13635. https://doi.org/10.3390/su132413635

Cha, G. W., Moon, H. J., Kim, Y. M., Hong, W. H., Hwang, J. H., Park, W. J., & Kim, Y. C. (2020). Development of a prediction model for demolition waste generation using a random forest algorithm based on small datasets. International Journal of Environmental Research and Public Health, 17(19), 6997. doi:10.3390/ijerph17196997

Chen, C., Ma, X., & Zhang, X. (2024). Empirical Study on Real Estate Mass Appraisal Based on Dynamic Neural Networks. Buildings, 14(7), 2199. https://doi.org/10.3390/buildings14072199

Chen, M., Liu, Y., Arribas-Bel, D., & Singleton, A. (2022). Assessing the value of user-generated images of urban surroundings for house price estimation. Landscape and Urban Planning, 226, 104486.

Choumert, J., & Phélinas, P. (2015). Determinants of agricultural land values in Argentina. Ecological Economics, 110, 134–140. https://doi.org/10.1016/j.ecolecon.2014.12.024

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences. Second Edition, New York: Lawrence Erlbaum Associates.

Colwell, P. F., & Dilmore, G. (1999). Who Was First? An Examination of an Early Hedonic Study. Land Economics, 75(4), 620. https://doi.org/10.2307/3147070

Çinar, S., & Ünel, F. B. (2022). 2/B orman vasfını yitirmiş araziden tarım arazisine dönüşen taşınmazların toplu değerlemesi. Geomatik, 7(2), 112–127. https://doi.org/10.29128/geomatik.900457

Dedeoğlu, M., & Dengiz, O. (2018). Coğrafi Bilgi Sistemleri İle Entegre Edilen Çok Kriterli Karar Destek Analiz Yaklaşımı Kullanılarak Arazi Uygunluk Sınıflarının Belirlenmesi. Süleyman Demirel Üniversitesi Ziraat Fakültesi Dergisi, 13(2), 60–72.

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. https://doi.org/10.1016/j.landusepol.2016.03.008

Demetriou, D. (2017). A spatially based artificial neural network mass valuation model for land consolidation. Environment and Planning B: Urban Analytics and City Science, 44(5), 864–883. https://doi.org/10.1177/0265813516652115

Demetriou, D. (2018). Automating The Land Valuation Process Carried Out In Land Consolidation Schemes. Land Use Policy, 75, 21–32. https://doi.org/10.1016/j.landusepol.2018.02.049

Dewi, M. R., Ulama, B. S. S., & Susilaningrum, D. (2024). The Comparison of Hedonic Regressiıon and Artificial Neural Network in the Development of Mass Appraisal Model: Case Study: Residential Property in Surabaya City. Jurnal Riset dan Aplikasi Matematika (JRAM), 8(2), 94-104.

El-Dereny M., & Rashwan, N. I. (2011). Solving Multicollinearity Problem Using Ridge Regression Models. Int. J. Contemp. Math. Sciences, 6(12), 585 – 600.

El Hami, A. (2025). Methods and Applications of Artificial Intelligence, Dynamic Response, Learning, Random Forest, Linear Regression, Interoperability, Additive Manufacturing and Mechatronics. ISTE Ltd and John Wiley & Sons, Inc. 77-100.

Enemark, S. (2004). Building Land Information Policies. UN, FIG, PC IDEA Inter-Regional Special Forum on The Building of Land Information Policies in the Americas. FIG, Aguascalientes, Mexico.

Enemark, S. (2010). Land Management: A Global Perspective. Journal on Geoinformatics, Nepal, 21–25. https://doi.org/10.3126/njg.v9i1.39686

Enemark, S., Bell, K. C., Lemmen, C., & McLaren, R. (2014). Fit-For-Purpose Land Administration: Joint FIG / World bank publication. International Federation of Surveyors (FIG).

Erdönmez, C. (2013). Research on Selling of 2B Areasfrom Point of National Forestry Programme of Turkey (2B alanlarının satışının Türkiye Ulusal Ormancılık Programı açısından irdelenmesi). Kastamonu University Journal of Forestry Faculty, 13(2), 307-324.

Ertunç, E., & Uyan, M. (2022). Land valuation with Best Worst Method in land consolidation projects. Land Use Policy, 122(106360). https://doi.org/10.1016/j.landusepol.2022.106360

ESRI, (2016). ArcGIS Desktop: Release 10.5. Environmental Systems Research Institute, California, USA.

European Commission, (2014). Cross country review of taxes on wealth and transfers of wealth, Specific Contract No8 TAXUD/2013/DE/335, Based on Framework Contract No TAXUD/2012/CC/117.

Fan, W., Zhang, Y., Chen, N., & Nie, W. (2024). A Review of Rural Land Capitalization: Current Status and Further Research. Land, 13(3), 401. https://doi.org/10.3390/land13030401

FAO. (2008). Opportunities to mainstream land consolidation in rural development programmes of the European Union. Food and Agriculture Organization of the United Nations. https://www.fao.org/4/i0091e/i0091e00.pdf

FAO, U. (2024). FAO/UNESCO Soil Map of the World. Food and Agriculture Organization of the United Nations, FAO Soil Portal. https://www.fao.org/soils-portal/data-hub/soil-maps-and-databases/faounesco-soil-map-of-the-world/en/

Forest Law, (1956). Pub. L. No. 6831, Orman Kanunu (Forest Law), 9402.

Garreta, R., & Moncecchi, G. (2013). Learning scikit-learn: Machine Learning in Python. UK: Packt Publishing. Birmingham B3 2PB, UK. ISBN 978-1-78328-193-0

Gerald, B. (2018). A Brief Review of Independent, Dependent and One Sample t-test. International Journal of Applied Mathematics and Theoretical Physics, 4(2), 50–54. https://doi.org/10.11648/j.ijamtp.20180402.13

Gökkuş, A. (2018). Meralarımız ile ilgili bir değerlendirme. TÜRKTOB Dergisi, 25, 6–8.

Günel, A. (2003). Regresyon Denkleminin Başarısını Ölçmede Kullanılan Belirleme Katsayısı ve Kritiği. Doğuş Üniversitesi Dergisi, 4(2), 133–140.

Güzel, Y. (2018). Biyolojik Sinir Sistemi ve Yapay Sinir Ağı Hücresi. https://medium.com/@yasinguzel/yapay-zeka-ders-notlar%C4%B1-03-biyolojik-sinir-sistemi-ve-yapay-sinir-a%C4%9F%C4%B1-h%C3%BCcresi-6555add68d80

Han, S., Williamson, B. D., & Fong, Y. (2021). Improving random forest predictions in small datasets from two-phase sampling designs. BMC Medical Informatics and Decision Making, 21, 322. https://doi.org/10.1186/s12911-021-01688-3

Heath, D. Kasif, S. & Salzberg, S. (1996). Committees of Decision Trees Cognitive Technology: In Search of a Humane Interface B. Gorayska and J.L. Mey (Editors), Elsevier Science B.V., Chapter 18, 305-317.

Hertel, T. W. (2011). The Global Supply and Demand for Agricultural Land in 2050: A Perfect Storm in the Making? American Journal of Agricultural Economics, 93(2), 259–275. https://doi.org/10.1093/ajae/aaq189

Ho, T. K. (1995). Random Decision Foserst, AT & T Bell Laboratories, USA.

Hoerl, A. E., & Kennard, R. W. (1970). *Ridge regression: Biased estimation for nonorthogonal problems*. Technometrics, 12(1), 55–67.

Hoerl, R. W. (2020). Ridge Regression: A Historical Context. Technometrıcs, 62(4), 420–425.

Holt, C. A., & Sullivan, S. P. (2023). Permutation tests for experimental data. Experimental Economics, 26, 775–812. https://doi.org/10.1007/s10683-023-09799-6

Hong, J., Choi, H., & Kim, W.S. (2020). A house price valuation based on the random forest approach: the mass appraisal of residential property in South Korea. International Journal of Strategic Property Management, 24(3), 140-152. https://doi.org/10.3846/ijspm.2020.11544

Hüttel, S., Wildermann, L., & Croonenbroeck, C. (2016). How do institutional market players matter in farmland pricing? Land Use Policy, 59, 154–167. https://doi.org/10.1016/j.landusepol.2016.08.021
IAAO, (2020). Standard on Property Tax Policy. The International Association of Assessing Officers, Kansas City.
Iara A. (2015). Wealth distribution and taxation in EU Members. WORKING PAPER N. 60, Luxembourg.
Iban, M. C. (2022). An explainable model for the mass appraisal of residences: The application of tree-based Machine Learning algorithms and interpretation of value determinants. Habitat international, 128, 102660.
IBM, (2011). IBM SPSS Statistics 20 Version, International Business Machines Corporation, New York, USA.
İlhan, A. T., & Öz, S. N. (2020). Yapay Sinir Ağlarının Gayrimenkullerin Toplu Değerlemesinde Uygulanabilirliği: Gölbaşı İlçesi Örneği. Hacettepe Üniversitesi Sosyal Bilimler Dergisi, 2(2), 160–188.
Jafary, P., Shojaei, D., Rajabifard, A., & Ngo, T. (2024). Automated land valuation models: A comparative study of four machine learning and deep learning methods based on a comprehensive range of influential factors. Cities, 151, 105115.
Kara, A., Çağdaş, V., Isikdag, U., Van Oosterom, P., Lemmen, C., & Stubkjaer, E. (2021). The LADM Valuation Information Model and its application to the Turkey case. Land Use Policy, 104, 105307. https://doi.org/10.1016/j.landusepol.2021.105307
Kara, A., Van Oosterom, P., Çağdaş, V., Işıkdağ, Ü., & Lemmen, C. (2020). 3 Dimensional data research for property valuation in the context of the LADM Valuation Information Model. Land Use Policy, 98, 104179. https://doi.org/10.1016/j.landusepol.2019.104179
Kenney, J. F., & Keeping, E. S. (1962). Frequency, Mathematics of Statistics, Part 1. https://en.wikipedia.org/wiki/Frequency_(statistics)
Kong, E., Wang, L., Xia, Y., & Liu, J. (2022). A permutation test for two-sample means and signal identification of high-dimensional data. Statistica Sinica, 32(4), 89–108. https://doi.org/10.5705/ss.202019.0425
Kuhn, M., & Johnson, K. (2013). *Applied Predictive Modeling*. Springer, New York.
Land Use Law, (2005). Pub. L. No. 5403, Toprak Koruma ve Arazi Kullanımı Kanunu (Soil Protection and Land Use Law), 25880.
LiCAD, (2024). Lider Yazılım (Lider Software), Lider Bil. Tek. Tur. G. San. ve Tic. Ltd. Şti., Denizli, Turkiye.
Ma, H., & Li, J. (2017). The Impacts of Supply and Demand Analysis on the Price of the Real Estate Market. Proceedings of the 7th International Conference on Education, Management, Information and Mechanical Engineering (EMIM 2017). 7th International Conference on Education, Management, Information and Mechanical Engineering (EMIM 2017), Shenyang, China. https://doi.org/10.2991/emim-17.2017.384
Ma, S., & Swinton, S. M. (2012). Hedonic Valuation of Farmland Using Sale Prices versus Appraised Values. Land Economics, 88(1), 1–15. https://doi.org/10.3368/le.88.1.1
Marquardt, D. W., & Snee, R. D. (1975). Ridge Regression in Practice. The American Statistician, 29(1), pp. 3-20. https://www.jstor.org/stable/2683673
MathWorks, (2010). MATLAB (Version R2010a). The MathWorks Inc., Massachusetts, USA.
McDonald, G. C. (2009). Ridge regression. Wiley Interdisciplinary Reviews: Computational Statistics, 1(1), 93-100.
Meyer, M. A., & Früh-Müller, A. (2020). Patterns and drivers of recent agricultural land-use change in Southern Germany. Land Use Policy, 99, 104959. https://doi.org/10.1016/j.landusepol.2020.104959
Moosavi, V. (2017). Urban data streams and machine learning: a case of swiss real estate market. arXiv preprint arXiv:1704.04979.
Moreno-Foronda, I., Sánchez-Martínez, M.-T. & Pareja-Eastaway, M. (2025). Comparative Analysis of Advanced Models for Predicting Housing Prices: A Review. Urban Sci., 9, 32. https://doi.org/10.3390/urbansci9020032
Murthy, S. K., Kasif S., & Salzberg, S. (1994). A System for Induction of Oblique Decision Trees. Journal of Articial Intelligence Research, 2, 1-32.
OECD, (2024). Revenue Statistics 2024: Health Taxes in OECD Countries, OECD Publishing, Paris, https://doi.org/10.1787/c87a3da5-en
Öztürk, G., Engindeniz, S., & Bayraktar, Ö. V. (2017). İzmir deki Sulanabilir Tarım Arazilerinin Değerini Etkileyen Faktörlerin Analizi. Selcuk Journal of Agricultural and Food Sciences, 31(3), 75–87. https://doi.org/10.15316/SJAFS.2017.38
Pastukh, O., & Khomyshyn, V. (2025). Using ensemble methods of machine learning to predict real estate prices. arXiv preprint arXiv:2504.04303.
Peng, W., Coleman, T., & Mentch, L. (2022). Rates of convergence for random forests via generalized U-statistics. Electronic Journal of Statistics, 16(1), 232-292.
Perujo-Villanueva, M., & Colombo, S. (2021). Impact of parcel fragmentation on the calculation of the real estate value of land belonging to farms. New Medit, 20(1). https://doi.org/10.30682/nm2101g

Probst, P., Wright, M. N., & Boulesteix, A. L. (2019). Hyperparameters and Tuning Strategies for Random Forest, Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 9(3), e1301. https://doi.org/10.1002/widm.1301

Province and Districts Law, (2012). Pub. L. No. 6360, On dört ilde büyükşehir belediyesi ve yirmi yedi ilçe kurulması ile bazı kanun ve kanun hükmünde kararnamelerde değişiklik yapılmasına dair kanun (Regarding the Amendment of Some Laws and Decree-Laws with Metropolitan Mu-nicipality in Fourteen Province and the Establishment of Twenty-Seven Districts).

Przekop, D. (2022). Artificial neural networks vs spatial regression approach in property valuation. Central European Journal of Economic Modelling and Econometrics, 199-223.

Qi, Y. (2012). Random Forest for Bioinformatics. In: Zhang, C. and Ma, Y.Q. Ed., Ensemble Machine Learning, Springer, US, 307-323. http://dx.doi.org/10.1007/978-1-4419-9326-7_11

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1985). Learning Internal Representations by Error Propagation. University of California, Institute for Cognitive Science.

Russell, R. S., & Taylor, B. W. (2003). Operations Management. Prentice Hall.

Saaty, R. W. (1987). The Analytic Hierarchy Process-What it is and how it is Used. Mathl Modelling, 9(3–5), 161–176.

Saaty, T. L. (1990). How to make a decision: The Analytic Hierarchy Process. European Journal of Operational Research, 48, 9–26.

Saaty, T. L. (2008). Decision Making with the Analytic Hierarchy Process. International Journal of Services Sciences, 1(1), 83. https://doi.org/10.1504/IJSSCI.2008.017590

Sampson, G. S., Hendricks, N. P., & Taylor, M. R. (2019). Land market valuation of groundwater. Resource and Energy Economics, 58, 101120. https://doi.org/10.1016/j.reseneeco.2019.101120

Saraç, E. (2012). Yapay Sinir Ağları Metodu ile Gayrimenkul Değerleme. Yüksek Lisans Tezi. İstanbul Kültür Üniversitesi.

SDG. (2023). The Sustainable Development Goals Report 2023: Special Edition. United Nations Department of Economic and Social Affairs. https://doi.org/10.18356/9789210024914

Seifert, S., & Hüttel, S. (2020). Common values and unobserved heterogeneity in farmland auctions in Germany. Agricultural Land Markets Efficiency and Regulation, DFG Research Unit 2569 FORLand, Humboldt-Universität Zu Berlin, 6(21).

Shanmuganathan, S., & Samarasinghe, S. (Eds.). (2016). Artificial Neural Network Modelling (Vol. 628). Springer International Publishing. https://doi.org/10.1007/978-3-319-28495-8

Shao, Z., & Er, M. J. (2016). Efficient leave-one-out cross-validation-based regularized extreme learning machine. Neurocomputing, 194, 260-270.

Sharma, H., Harsora, H. & Ogunleye, B. (2024). An Optimal House Price Prediction Algorithm: XGBoost. Analytics. 3, 30–45. https:// doi.org/10.3390/analytics3010003

Shobha Rani, N., Chandan, N., Sajan Jain, A., & R. Kiran, H. (2018). Deformed character recognition using convolutional neural networks. International Journal of Engineering & Technology, 7(3), 1599. https://doi.org/10.14419/ijet.v7i3.14053

Skiera, B., Reiner, J., & Albers, S. (2018). Regression Analysis. In C. Homburg, M. Klarmann, & A. Vomberg (Eds.), Handbook of Market Research (pp. 1–29). Springer International Publishing. https://doi.org/10.1007/978-3-319-05542-8_17-1

Spyder, (2023). Spyder: The Scientific Python Development Environment (Version 5.x) software, Spyder Development Team. https://github.com/spyder-ide/spyder

Stokes, J., & Jansen, J. (2024). Farmland Valuation: Understanding Income Capitalization and Cap Rates. N Agricultural Economics, Cornhusker Economics, the University of Nebraska–Lincoln. https://agecon.unl.edu/cornhusker-economics

Sykes, A. O. (1993). An Introduction to Regression Analysis. Coase-Sandor Institute for Law & Economics Working Paper No. 20.

Sylla, M., Lasota, T., & Szewrański, S. (2019). Valuing Environmental Amenities in Peri-Urban Areas: Evidence from Poland. Sustainability, 11(3), 570. https://doi.org/10.3390/su11030570

Szentesi, S. G., Pantea, M. F., Trifan, V. A., Mazuru, L. I., & Szentesi, N. F. G. (2024). Standardization of Regression Equation Parameters in the Case of Multiple Linear Regression for an Econometric Model Development to Determine the Price of Apartments. Proceedings of the International Conference on Business Excellence, 18(1), 2344–2352. https://doi.org/10.2478/picbe-2024-0198

Şengöz, N. (2018). Yapay Sinir Ağları: Yapay Zeka ve Muhteşem Beyin. https://tr.newworldai.com/yapay-sinir-aglari-nilgun-sengoz/

Tabar, M. E., Sisman, A., & Sisman, Y. (2023). A Real Estate Appraisal Model with Artificial Neural Networks and Fuzzy Logic: A Local Case Study of Samsun City. International Real Estate Review, 26(4).

Taber, K. S. (2018). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. Research in Science Education, 48(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Technical Guideline, (2017). Pub. L. No. 30265, Technical Instruction on Soil and Land Classification Standards (Toprak ve Arazi Sınıflaması Standartları Teknik Talimatı).

Tempesta, T., Foscolo, I., Nardin, N., & Trentin, G. (2021). Farmland value in the "Conegliano Valdobbiadene Prosecco Superiore PGDO" area. An application of the Hedonic Pricing method. Aestimum, 78, 5–33. https://doi.org/10.36253/aestim-10826

TKMP, (2014). Gayrimenkul değerleme bileÇeni pilot uygulama taslak tamamlanma raporu, Tapu Kadastro Modernizasyon Projesi-TKMP, ÇŞB, TKGM, Kadastro Dairesi Başkanlığı, Ankara.

TKGM, (2019). Tapu ve Kadastro Genel Müdürlüğü, Taşınmaz Değerleme Dairesi Başkanlığı, Ankra. https://www.tkgm.gov.tr/tasinmaz

Tochaiwat, K., & Pultawee, P. (2024). House type specification for housing development project using machine learning techniques: A study from bangkok metropolitan region, thailand. Nakhara: Journal of Environmental Design and Planning, 23(1), 403-403.

TUIK. (2024a). Statistic Data Portal. Turkish Statistical Institute (Türkiye İstatistik Kurumu-TUIK), Statistic Data Portal. https://data.tuik.gov.tr/Kategori/GetKategori?p=Nufus-ve-Demografi-109

TUIK. (2024b). TUIK Data: Biruni. Turkish Statistical Institute, The centralised distribution system. https://biruni.tuik.gov.tr/medas/?locale=tr

TUIK. (2024c). Dönemsel Gayrisafi Yurt İç Hasıla (53756). Turkish Statistical Institute, Data Portal. https://data.tuik.gov.tr/Bulten/Index?p=Donemsel-Gayrisafi-Yurt-Ici-Hasila-IV.-Ceyrek:-Ekim-Aralik-ve-Yillik,-2023-53756#:~:text=%C3%9Cretim%20y%C3%B6ntemiyle%20Gayrisafi%20Yurt%20%C4%B0%C3%A7i,milyar%20402%20milyon%20olarak%20ger%C3%A7ekle%C5%9Fti.

Uberti, M. S., Antunes, M. A. H., Debiasi, P., & Tassinari, W. (2018). Mass appraisal of farmland using classical econometrics and spatial modeling. Land Use Policy, 72, 161–170. https://doi.org/10.1016/j.landusepol.2017.12.044

UN. (2022). World Population Prospects 2022: Summary of Results [United Nations]. https://doi.org/10.18356/9789210014380

Unel F.B., Kusak L. & Yakar M. (2023). GeoValueIndex map of public property assets generating via Analytic Hierarchy Process and Geographic Information System for Mass Appraisal. Aestimum, 82, 51-69. doi: 10.36253/aestim-14110

Unel, F. B., & Yalpir, S. (2019). Reduction of mass appraisal criteria with principal component analysis and integration to GIS. International Journal of Engineering and Geosciences, 4(3), 94–105. https://doi.org/10.26833/ijeg.458430

Wong, T. T. (2015). Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. Pattern recognition, 48(9), 2839-2846.

World Bank, (2020). The Property Tax Diagnostic Manual. Authors: Kelly, R., White, R. & Anand, A. Property Tax Diagnostic Manual. Washington, D.C., USA: World Bank Group.

Worldometer. (2024). Worldometer—Real time world statistics. World Population. https://www.worldometers.info/

Wu, L., Zhang, Y., Wei, Y., & Chen, F. (2022). A BP Neural Network-Based GIS-Data-Driven Automated Valuation Framework for Benchmark Land Price. Complexity, 2022(1), 1695265.

Yalpir, S., Durduran, S. S., Unel, F. B., & Yolcu, M. (2014). Creating A Valuation Map In GIS Through Artificial Neural Network Methodology: A Case Study. Acta Montanistica Slovaca, 19(2), 79–89.

Yalpir, S., & Unel, F. B. (2022). Multivariate statistical analysis application to determine factors affecting the parcel value to be used mass real estate valuation approaches. International Journal of Engineering and Geosciences, 7(1), 32–42. https://doi.org/10.26833/ijeg.862563

Yalpır, Ş., & Ünel, F. B. (2016). Investigation and Reduction of Criteria Affecting The Value of Land Plot in Turkey and International Standards by Factor Analysis. Afyon Kocatepe University Journal of Sciences and Engineering, 16(025502), 303–322. https://doi.org/10.5578/fmbd.28134

Yazdani, M. (2021). Machine learning, deep learning, and hedonic methods for real estate price prediction. arXiv preprint arXiv:2110.07151.

Yazdani, M., & Raissi, M. (2023). Real Estate Property Valuation using Self-Supervised Vision Transformers. arXiv preprint arXiv:2302.00117.

Yomralioglu, T. (1993). A Nominal Asset Value-Based Approach for Land Readjustment and its Implementation Using Geographical Information Systems. University of Newcastle upon Tyne, PhD.

Yomralioglu, T., Nisanci, R., & Yildirim, V. (2007). An Implementation of Nominal Asset Based Land Readjustment. FIG Working Week 2007, Hong Kong SAR, China.

Zhang, Q. (2021). Housing price prediction based on multiple linear regression. Scientific Programming, 2021(1), 7678931.

Zhang, Y., Zhou, R., & Chen, N. (2023). Identification and correction of ratchet effect of residential land price: Empirical study on urban agglomeration in China. Applied Economics, 55(50), 5847–5863. https://doi.org/10.1080/00036846.2022.2140768