Energy for Sustainable Development 77 (2023) 101312

Contents lists available at ScienceDirect

Energy for Sustainable Development

journal homepage: www.journals.elsevier.com/energy-for-sustainable-development

Analysing the determinants of the Turkish household electricity consumption using gradient boosting regression tree

Denizhan Guven [a,*], M. Ozgur Kayalica [b]

[a] *Eurasia Institute of Earth Sciences, Istanbul Technical University, Istanbul, Turkey*
[b] *Energy Institute, Istanbul Technical University, Istanbul, Turkey*

ARTICLE INFO

Keywords:
Household electricity demand
Household Budget Survey
Social cost of Carbon
Gradient Boosting Regression tree
Machine Learning
Turkey

ABSTRACT

Residential buildings are the second largest electricity consumer in Turkey. Thus, the goal here is to detect the factors that determine the electricity consumption of the households in Turkey using the Household Budget Survey (HBS). This study applies Decision Tree (DT), Random Forest (RF) and Gradient Boosted Regression Tree (GBRT) methods. Since the GBRT method provides the lowest Root Mean Squared Error (RMSE), the impact of each variable on the electricity consumption is analysed with this method. The most critical determinant is found to be the household size, while income level and heating type are discovered as 2nd and 3rd most prominent determinants for household electricity demand. With the help of the Partial Dependence Plots (PDP) provided by the GBRT method, the impact of each categorical and continuous variable is presented. Using the results of PDPs, the monetary values of both electricity generation and the social cost of CO_2 emissions emitted into the atmosphere due to electricity generation are calculated for the most important determinants.

Introduction

Research background

The energy industry had concentrated only on the supply side of the system rather than the demand side until the first oil shock occurred in 1973. During this period, it is intended to match the energy demand by controlling the energy supply. However, after the first oil shock, the energy industry started to pay attention also to the demand side due to the abrupt price spikes (Bhattacharyya, 2011).

The electricity sector consists of different types of sub-sectors, namely residential, commercial, industrial, agricultural, community and transportation. The determiners and key factors of these sub-sectors are quite different. To investigate the impact of determiners, a sectoral approach to electricity demand analysis is very crucial.

In this context, as shown in Fig. 1, residential sector is responsible for about 22 % of overall energy consumption in 2019, whereas it accounts for approximately 17 % of global carbon dioxide (CO_2) emissions. Moreover, buildings account for more than 55 % of the global electricity demand (IEA (International Energy Agency), 2020, IEA (International Energy Agency), 2020b). On the supply side, coal-based power plants are the largest single emission releaser with a share of 30 % of global CO_2 emissions (IEA (International Energy Agency), 2019). Thus, arranging both the demand and supply side of the electricity sector takes a very important place for policymakers.

Following the 18th century, global CO_2 emissions have been increasing impetuously due to the first industrial revolution. Until this period, global CO_2 concentration was 280 ppm on average (Ritchie & Roser, 2019). However, in the period following the industrial revolution, CO_2 emissions have hit almost 420 ppm in 2020 spring as a record (SIO (Scripps Institution of Oceanography), 2020). Although the United Nations endeavours to keep CO_2 emission under control, CO_2 emissions continue to increase year by year. This necessitates to empower research in this area.

Research significance

It is a very well-known fact that CO_2 emissions are the number one key factor for climate change which is one of the most prominent phenomena of modern-day. Although CO_2 constitutes a very small part of the atmosphere, it is stated as the most significant *long-lived forcing* of climate change (Lacis, Hansen, Russell, Oinas, & Jonas, 2013). Nevertheless, country/state-based CO2 emission is not identical, and it has been varying over time. With the growing population, improving living

* Corresponding author at: Denizhan Guven, ITU Ayazaga Campus, Eurasia Institute of Earth Sciences, 34469, Maslak, Istanbul, Turkey.
E-mail address: guvende@itu.edu.tr (D. Guven).

https://doi.org/10.1016/j.esd.2023.101312
Received 30 January 2023; Received in revised form 23 August 2023; Accepted 29 September 2023
Available online 14 October 2023

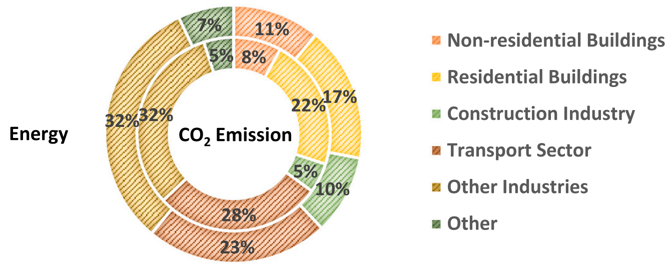

Fig. 1. Global shares of sectors in final energy (outer circle) and CO_2 emissions (inner circle) (Source: Authors' work using IEA (International Energy Agency), 2020, IEA (International Energy Agency), 2020b).

standards, and developing manufacturing sector, the share of Turkey in global CO_2 emissions has risen from 0.27 % to 1.15 % between the period 1970–2020 (BP (British Petroleum), 2021). Attempts to mitigate the impact of climate change by diminishing Greenhouse Gases (GHG) are crucial. Following COP21, Turkey has a challenging promise of reducing GHGs by almost one-fifth of its 2016 level (UNFCCC, 2015).

Turkey as an emerging country meets its electricity generation mostly from carbon-based fuels. Despite increasing investments in renewables, the portion of fossil-based electricity generation is about 67 % in total generation as shown in Fig. 2 (TEDC (Turkish Electricity Distribution Corporation), 2021a). Moreover, the self-sufficiency rate in electricity production has declined from 74 % in 1980 to 48.4 % in 2021 (author's calculation based on EXIST (Energy Exchange Istanbul), 2021 and TEDC (Turkish Electricity Distribution Corporation), 2019 data). A decrease in self-sufficiency rate causes a more fragile economy. This is the main engine of our starting point. Ascertaining determiners of electricity demand is very essential for constructing forthcoming government policies and designs that may also secure energy for future needs.

When it comes to the demand side, the electricity demand of Turkey has extended from 7.31 TWh in 1970 to 262.7 TWh in 2020, and the percentage of household electricity consumption (HEC) overall has jumped from 15.9 % to 23.1 % in this period (TEDC (Turkish Electricity Distribution Corporation), 2021b).

Research objectives

In light of this information, the objective here is to detect the factors that determine the HEC in Turkey using the Household Budget Survey (HBS). Many factors are affecting the residential electricity demand, such as household income, building properties, weather conditions, quantity and accessibility of household appliances, and consumer preferences. As a country with limited energy sources, Turkey should use these sources efficiently. Thus, evaluating the effects of these factors on HEC plays a fundamental role in determining future energy strategies, including capacity expansions, saving policies, and investment plans.

Adopted literature review

The following two subsections are dedicated for the literature review. While this subsection aims to provide an in-depth review of the literature on the determinants of HEC, the next one explicitly shows the gaps in the reviewed studies. A synopsis of the most recent works on this topic given in Table A.1 (in Appendices) will help to present the outcomes discussed in Adopted literature review and Research gap and motivation sections more efficiently. Eventually, this attempt eases the choice of the methodologies used for the analysis.

It is possible to handle the determinants of HEC under five titles, namely socio-economic structure of the household, appliances ownership, the impact of climate, location, and dwelling features.

A few studies are covering all types of determinants abovementioned. Kavousian, Rajagopal, and Fischer (2013) proposed a stepwise regression model to study the determinants in the USA. They find that climate, location and floor area are crucial determinants of HEC. In a similar study, based on the data of 500 Latvian households with a smart metering system, Laicāne, Blumberga, Rošā, and Blumberga (2014) analysed the influence of those determinants on HEC using regression analysis. More recently, Hernandez and Patiño-Echeverri (2019) used both Random Forest (RF) and Multiple Linear Regression (MLR) methods to analyse the most influential parameters on HEC in Mexico. This study revealed that location is much more important than socio-demographic characteristics and dwelling features.

Most of the studies in the literature consider socio-economic characteristics, dwelling features and household appliances ownership. While Bedir, Hasselaar, and Itard (2013) utilized regression for the Netherlands, Danlami (2017) applied the Ordinary Least Squares (OLS) method for Nigeria. Using a similar method to Danlami (2017) and quantile regression, Kim, Lee, Jin, Suh, and Song (2020) examined it for Korea. It was revealed that household size, household appliances ownership, age of household head, floor area, and refrigerator usage time are important for all quantiles. Zhang, Teng, and Zhou (2020) investigates it in China using the Seemingly Unrelated Regression (SUR) model. The results indicated that building features have a significant impact.

Dissimilar to the aforementioned studies, some studies also consider the influence of the location on HEC. Druckman and Jackson (2008)

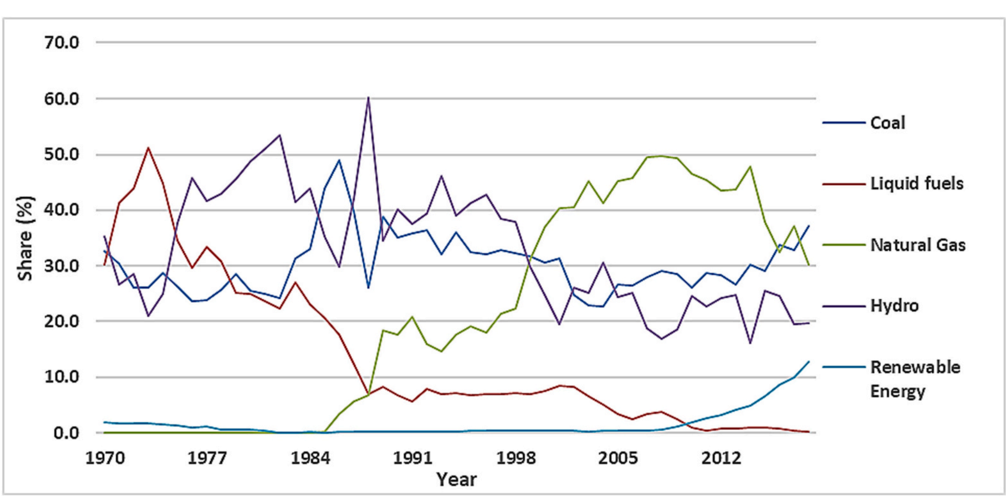

Fig. 2. Electricity generation shares by energy resources of Turkey (Authors' work using TEDC (Turkish Electricity Distribution Corporation) (2021a)).

used the Local Area Resource Analysis technique to determine the influence of socio-economic characteristics, dwelling features and location on HEC in the UK. It was stated that income is the most important factor, while dwelling type, household composition and location are also found to be exceedingly significant. By focusing on an Italian case study, Besagni and Borgarello (2018) utilized the OLS method to determine the effect of socio-economic characteristics, dwelling features, household appliances ownership and location on HEC. Socio-economic parameters were detected as the most influential factor on HEC. Moreover, Structural Equation Modelling was implemented by Đurišić, Rogić, Smolović, and Radonjić (2020) to analyse the factors which influence the HEC in Montenegro at most. Income, floor area, family composition and features of appliances became prominent as the most important factors.

Unlike previous studies, Jones, Fuertes, and Lomas (2015) presented a literature review on HEC. More than 62 factors were investigated under three main titles, including socio-economic features, dwelling properties and appliance ownership. Four socio-economic features, seven dwelling properties and nine appliance related factors were detected as having a positive impact.

Only a few studies are examining the determinants for the Turkish household behaviour. Despite its importance, the literature on this topic for Turkey is very limited. Ozcan, Gulay, and Ucdogruk (2013) examined the effect of socio-economic, dwelling and location factors using a Multinomial Logit Model. They present that monthly income is the most significant component in energy choices. Guloglu and Akin (2014) used Partial Proportional Odds. More recently, Kayalica, Ozozen, Guven, Kayakutlu, and Bayar (2020) used both Artificial Neural Network (ANN) and Analytic Hierarchy Process (AHP) methods. They revealed that dwelling features have relatively higher influence on electricity consumption than household appliances ownership.

Research gap and motivation

As reviewed above, a large portion of the existing studies utilized only a certain method, especially an MLR or OLS model, due to their simplicity and high ability of interpretation. Nevertheless, linear models cannot offer a sufficient solution for non-linear problems. Hence, they hardly ascertain the hidden non-linear relationships between electricity consumption and household variables (Kim et al., 2020). Non-linear tree-based models (e.g., Gradient Boosted Regression Tree (GBRT) and RF) provide a better result, since they are capable of recognizing the nonlinear dynamics and relationships among variables.

The literature, particularly for Turkey, does not fully study tree-based machine learning techniques. More significantly, this is the first study employing the GBRT approach to examine the association between HEC and household characteristics. Furthermore, none of the studies abovementioned does thoroughly show how and how much each determinant impact HEC.

The novelty of this work is three-fold. First, unlike previous studies, this study analyses the determinants of household electricity consumption using three different methods. These are Random Forest (RF), Gradient Boosted Regression Tree (GBRT) and Decision Tree (DT) methods. This is because this method can distinguish the significance of variables and determine their partial-dependences (Wang, Hu, & Jiang, 2021). Secondly, this is the first study that utilized the GBRT method to detect the determiners of HEC. Last of all, none of the studies in the literature deeply present how and how much each determinant impact HEC. With this study, the non-linear relations between determinants and HEC will be presented using Partial Dependence Plots (PDPs) and Interaction Plots which are two of the outputs of abovementioned machine learning techniques. In addition, as a result of this analysis, the potential savings for electricity generation and CO_2 emissions are calculated. Hence, this work targets to fill these gaps in the literature.

Methodology and data

Methodology

To analyse the determinants of Turkish household electricity consumption, three different tree-based ML methods are employed in this study. Before establishing ML methods, the HBS data is divided into two parts, namely the training dataset and the test dataset. Following the training of models using the training dataset and parameter optimization, the established models are tested with the test dataset. The method with the lowest RMSE is selected to analyse Turkish household electricity consumption extensively. The framework of the proposed models is illustrated in Fig. 3.

Decision tree regression

Decision Tree (DT) operates within the realm of supervised machine learning, employing a hierarchy of rules. An illustrative instance featuring five regions is depicted in Fig. 4. In this case, the variables x_1 and x_2 undergo division based on predefined breakpoints (a_1- a_4).

The core principle behind the DT method involves breaking down a complex decision into a series of simpler decisions. During this process, the DT algorithm partitions the predictor space into J distinct and non-overlapping regions, such as R1, R2, …, RJ. The objective of the model is to determine the rectangles R1, R2, …, RJ that minimize the Residual Sum of Squares (RSS).

Random forest

The ensemble machine learning technique known as the Random Forest, first proposed by Breiman (2001), involves generating multiple randomized trees and making decisions based on non-parametric statistical regression principles (Xu et al., 2019). Once this collection of trees forms a forest, the Random Forest model combines their predictions through a voting mechanism. The final output of the Random Forest model (Eq. 1) is determined by averaging the predictions from all the individual trees.where c_{full} is the mean of the whole dataset, K is the total number of features, and J represents the number of trees in the forest (See Fig. 5.).

Gradient boosted regression tree

The Gradient Boosted Regression Tree method involves a sequence of shallow trees, where each tree builds upon the knowledge of its predecessor. The input for GBRT is a dataset $\left\{(x_i, y_i)\right\}_{i=1}^{n}$ and a differentiable loss function $L(y_i, F(x))$. GBRT operates as a gradient descent algorithm, and it is applicable with any loss function that is differentiable.

When configuring a GBRT algorithm, three key parameters play a pivotal role in influencing the model's performance. These factors encompass the learning rate, often referred to as shrinkage (R), the maximum depth of interactions (D), and the count of trees (K). Enhancing the model's performance hinges upon identifying the optimal combination of these parameters, a pursuit highlighted by Zhang and Haghani (2015). Lowering R and increasing K may better the generalizability capability of the established model and thus avoid over-fitting.

Tree-based techniques do not require a GPU (Graphics Processing Unit) for the training process, unlike Neural Networks (NN), despite the fact that NN is one of the most used methods in prediction studies. In addition, compared to NNs, tree-based approaches have a higher accuracy potential. While tree-based techniques outperform deep learning algorithms on tabular-style datasets, deep learning algorithms are thought to be more appropriate for applications involving image, audio, and language processing (Lundberg et al., 2020). Furthermore, some variables of the survey data are strongly correlated each. Since tree-based methods use decision trees, it is capable of handling multicollinearity issues (Yoon, 2021). For the purpose of this study, tree-based approaches are more beneficial in this regard.

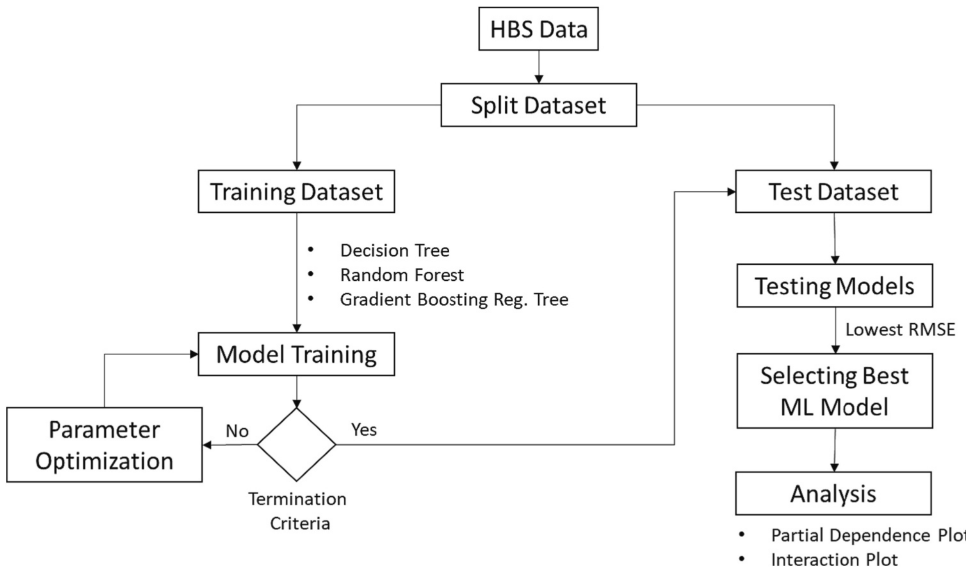

Fig. 3. Framework of the proposed structure.

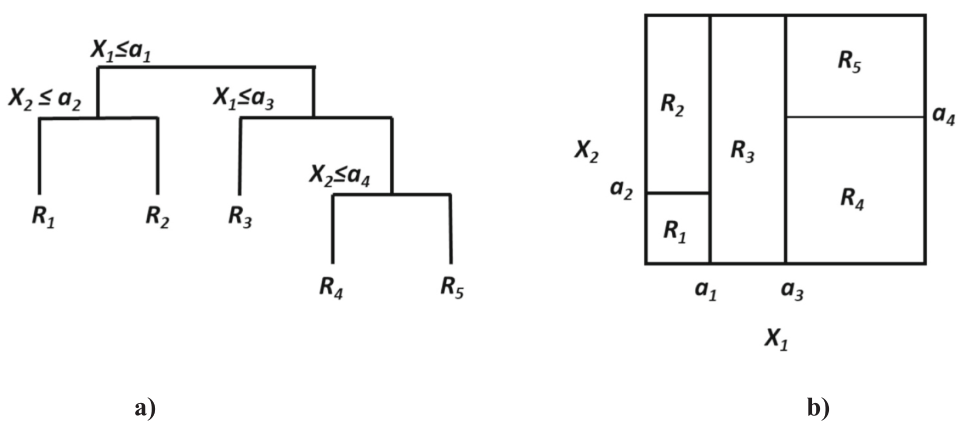

Fig. 4. a) The partition of a two-dimensional example b) The output of recursive binary splitting example.

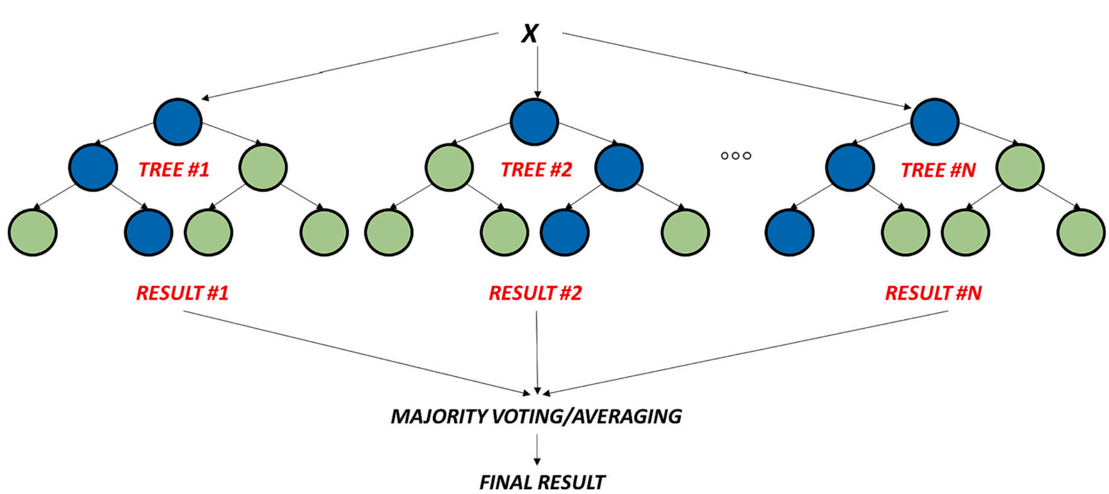

Fig. 5. Structure of Random Forest Algorithm

$$F(x) = \frac{1}{J}\sum_{j=1}^{J} c_{j_{full}} + \sum_{k=1}^{K}\left(\frac{1}{J}\sum_{j=1}^{J} contribution_j(x,k)\right) \tag{1}$$

Data

In this paper, the Household Budget Survey (HBS) is used to examine the determinants of HEC. HBS is one of the most important sources that provide information about the socio-economic structures, consumption habits of households, and test the validity of the applied socio-economic policies. HBS is the official survey managed by the Turkish Statistical Institute (TurkStat (Turkish Statistical Institute), 2019) and covers households' socio-demographic characteristics, income levels, dwelling properties, and appliances ownership.

This study uses the latest version (2019) of micro dataset of HBS. The data comprises the information gathered from 10,740 households in Turkey. The variables and their descriptions are given in Table 1. Furthermore, the distributions of variables are shown in Fig. 6.

Table 1
Description of survey data.

Variable	Description
Household Type	[1]: Single adult family, [21]: Couple without children, [22]: Elementary family with children, [23]: Single adult with children, [3]: Patriarchal or extended family, [4]: Students, workers, etc. living together
Dwelling Type	[1]: Detached residence, [121]: Twin or row house, [122]: Twin or row house (single entry), [131]: Apartments (<10 flats), [132]: Apartments (<10 flats) (single entry), [141]: Apartments (>10 flats), [142]: Apartments (>10 flats) (single entry), [2]: Other
Ownership	[1]: House owner [2]: Tenant [3]: Public housing [4]: Not a house owner but does not pay rent
Dwelling Age	[1]: 1945 and earlier, [2]: 1946–1960, [3]: 1961–1970, [4]: 1971–1980, [5]: 1981–1990, [6]: 1991–2000, [7]: 2001–2005, [8]: 2006 and later
Number of Rooms	1–99
Dwelling Area	10–999
Heating Type	[11]: Floor heating system (boiler etc.), [12]: Indoor central heating system, [13]: Remote heated central heating system, [2]: Natural gas stove, [32]: Air conditioner, [31]: Indoor central air conditioner, [90]: None, [98]: Other
Fuel Type	[111211]: Wood, [1111]: Coal, [13]: Turf, [121114]: Fuel oil, [1311]: Natural Gas, [1316]: LPG, [14]: Electricity, [1512]: Solar, [1513]: Thermal, [1514]: Wind, [98]: Other
Kitchen Fuel Type	[111211]: Wood, [1111]: Coal, [13]: Turf, [121114]: Fuel oil, [1311]: Natural Gas, [1316]: LPG, [14]: Electricity, [1512]: Solar, [1513]: Thermal, [1514]: Wind, [98]: Other
Hot Water Ownership	[1]: Yes [2]: No
Gas Ownership	[1]: Yes [2]: No
Cable TV Ownership	[1]: Yes [2]: No
Number of PC	[0]: or number of PC
Number of TV	[0]: or number of TV
Number of Refrigerator	[0]: or number of Refrigerator
Number of Freezer	[0]: or number of Freezer
Number of Dishwasher	[0]: or number of Dishwasher
Number of Microwave	[0]: or number of Microwave
Number of Washer	[0]: or number of Washer
Number of Dryer	[0]: or number of Dryer
Number of Air Conditioner	[0]: or number of Air Conditioner
Number of Game Console	[0]: or number of Game Console
Income	0–9,999,999.999 TL
Household Size	It is the household size calculated by taking into account the values of 1 for the first adult in the household, 0.5 for individuals aged 14 and over, 0.3 for individuals younger than 14 years old.
Average electricity consumption per month	0–9,999,999.999 kWh/month

In Turkey, the number of households in Turkey is approximately 25.3 million as of 2021 (TurkStat (Turkish Statistical Institute), 2022). Depending on the number of households in Turkey and sample size of HBS, the margin of error can be calculated as

$$\varepsilon = z \bullet \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \tag{2}$$

where $\hat{p}$ is the sample proportion (in this case, 0.5), n is the sample size (in this case, 10,740), and z is the z score (confidence interval, 99 %). As a result of Eq. 2, the margin of error is found to be 1.08. This means that there is a 99 % chance that the real value is within ±1.08 % of the surveyed value. Since 1.08 is less than the maximum margin of error (2.58) at the selected confidence level, the HBS is highly capable of representing the overall of Turkey (Wonnacott & Wonnacott, 1990).

Results and discussion

Estimation results

Measurement of error is one of the most critical phases to calculate the forecast accuracy and benchmark the forecast process. The accuracy can be calculated with various methods. In this study, the accuracies of models are examined with three different error measurements, namely Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE).

Mean Absolute Percent Error (MAPE) estimates the magnitude of the error in percentage terms and the formulation of it is given in Eq. 3.

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| \tag{3}$$

where A_t and F_t are actual value and forecasted value, respectively. The interpretation table of MAPE values is also shown in Table 2.

MAE is the arithmetic mean of the absolute errors between the actual value and predicted value. It is calculated as

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{4}$$

where y_i is the predicted value and $\hat{y}_i$the actual value. This calculation does not take into account the direction of errors. Besides, all individual differences have the same weight, so it does not penalise the larger differences. On contrary, RMSE is a quadratic scoring method. Thus, it penalises large errors with relatively higher weights. This denotes that RMSE is more functional in case large errors are particularly undesired. The equation of RMSE is given as

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{5}$$

All household electricity consumption models are run in *R Software* with the data aforementioned. The data is randomly separated into two parts, namely, the training dataset and the test dataset. Training and test datasets consist of values of 8055 and 2685 observations, respectively.

To establish the DT and RF models, the *rpart* and *randomForest* packages were used. Based on Eq. 8, MAPE of the DT model was calculated 15.9 % which can be labelled as a good forecast. RF model with 2250 trees and mtry = 3 provided the best MAPE with the value of 15.02 %. As mentioned in the previous section, the argument K (*mtry*) implies how many predictors is to be taken into consideration for each split of the tree. In addition to the *mtry* argument, the number of grown trees (*ntree*) is another argument that affects the performance of the model. Fig. 7 shows the RMSE comparison of different configurations of *ntree* and *mtry* arguments. The configurations with the lowest RMSE are chosen for determining the variance importance.

To establish the GBRT model, the *gbm* package of R software was

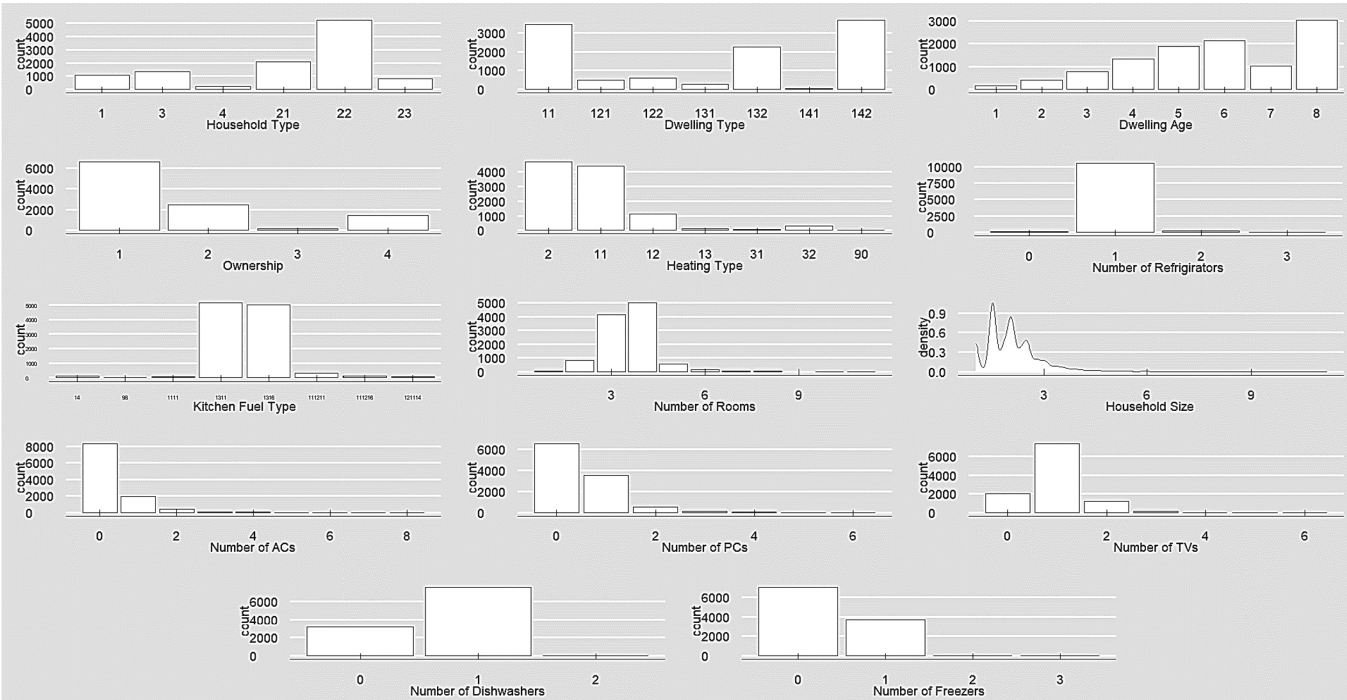

Fig. 6. Distributions of survey data.

Table 2
Interpretation of MAPE values (Source: Lewis, 1982).

Value	Interpretation
<10	Highly accurate forecasting
10–20	Good forecasting
20–50	Reasonable forecasting
>50	Inaccurate forecasting

used. As explained in Gradient boosted regression tree section, GBRT models have several tuning parameters. In this study, a grid search method is utilized to tune the parameters of GBRT model. Particularly, three learning rates (0.1, 0.01, and 0.001), and three tree complexities (1, 3, 5) are tested with the number of trees from 1 to 10,000.10-fold cross-validation is exhibited to determine performance robustness. After the designation of the hyper grid, the best tuning parameters were procured, and the MAPE value of fine-tuned GBRT model is calculated as 13.2 %. The best hyper grid structures used in the fine-tuned GBRT resulted from minimizing RMSE are given in Table 3.

In Table 4, MAPE, RMSE and MAE values of these models are given.

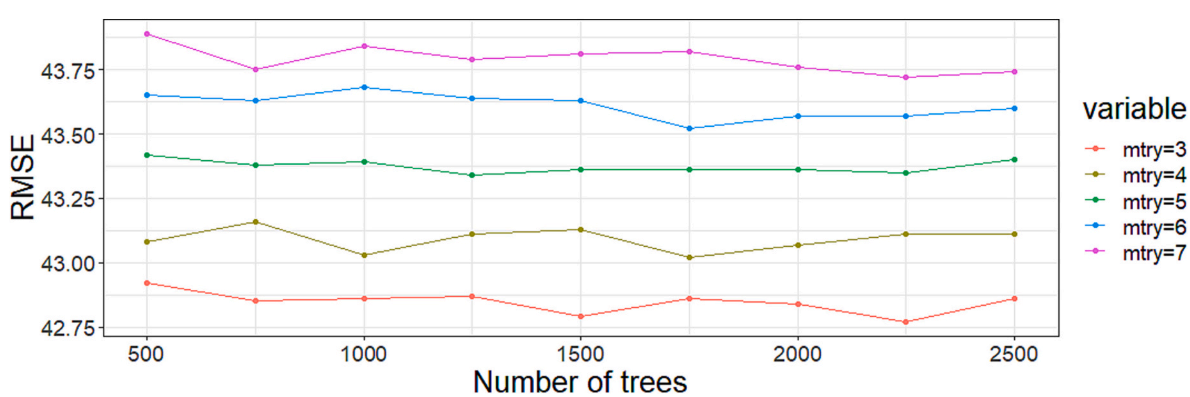

Fig. 7. RMSE values of RF models.

Table 3
Best hyper grid structures of training data for fine-tuning.

Shrinkage	Interaction depth	n.minobsinnode	bag.fraction	optimal_trees	min_RMSE
0.01	3	10	0.8	2287	38.99033
0.01	5	15	0.65	974	39.16137
0.01	3	10	0.65	2042	39.16348
0.01	5	10	0.80	1180	39.16434
0.01	3	15	1.0	2310	39.74524

Table 4
Error values of test data.

Model	MAPE	RMSE	MAE
Decision Tree	15.9	44.70	29.60
Random Forest	15.02	42.77	28.06
Gradient Boosted RT	13.2	40.55	26.19

As mentioned in the previous section, the most important feature of DT is the graphical display of diffraction points. The decision tree structure of household electricity consumption in Turkey is shown in Fig. 8.

As seen from Fig. 8, the most important factor is the income level of the household. The households with less income than 30,902 TL tends to use 63.1 kWh per month on average. On the contrary, electricity demand increases averagely to 87.1 kWh in the households with annual income more than 30,902 TL. Dishwasher ownership is a very important factor for the lower-income group, whereas air conditioner ownership becomes prominent in the higher-income group. The largest consumer group, which comprises 16.4 % of total households, is the higher-income group without air conditioner and less household size than 1.9. The households which have higher income level, heating type with air conditioner and natural gas stove are found to be the sub-group with the highest electricity demand. The households, which are smaller than 1.7, have less than 13,803 TL income and neither have air conditioner nor dishwasher are detected as the lowest energy demand group.

In contradistinction to the DT method, the RF method may not provide a certain tree plot since it creates various tree structures. However, it provides a diagram that shows the importance of variables on the dependent variable. The variable importance plot for household electricity demand is shown in Fig. 9.

As seen from Fig. 9, the most important variable on household electricity demand is the income level of the household. Dwelling area and household size also have a significant impact on household electricity demand. On the other hand, the number of game console, washer and refrigerator have very limited influence on determining the household electricity demand in Turkey. It can be revealed from Fig. 9 that following the income level, building properties are more important than household appliances to determine household electricity demand.

In Fig. 10, the permutation-based variable importance plot of GBRT model is shown. The *household size* is detected as the most significant variable, while *income level* and *heating type* are discovered as 2nd and 3rd most prominent determinants for household electricity demand forecasting. The least important factors are detected as *number of dryers, number of microwaves and hot water*.

The method with the lowest error rates is utilized to analyse the impact of each variable on the electricity consumption. Hence, based on the trained GBRT model, the partial influences of categorical factors and appliances ownership variables are shown in Fig. 11. For household type, it is seen that people living together, such as students and workers, and patriarchal or extended families tend to consume more electricity (83.5 kWh/month), whereas couples without children and single adult families demand less electricity (76.1 kWh/month). It can be said that there is almost no difference between dwelling types although apartments demand relatively higher amount of electricity. For the ownership category, people accommodated in public housing tend to consume less electricity compared to others. When it comes to heating type, buildings with remote heated central heating systems consume the least electricity (~75 kWh/month), while buildings heated with the air conditioner (~112.2 kWh/month) demand the most electricity. In Turkey, most central heating systems are powered with natural gas or fuel oil. Kitchens using natural gas consume less electricity (74 kWh/month) whereas kitchens using electricity causes more electricity consumption (~80 kWh/month). Moreover, dwelling age has a significant impact on HEC. The newer buildings tend to demand less electricity due to the improvement in insulation and electrical infrastructure (Lucon et al., 2014).

Analysis shows that gas ownership does not have a significant impact on HEC. On the other hand, as the number of PCs increases, an increase occurs in the electricity consumption as expected. However, it can be said that there is almost no difference between two and three PCs regarding electricity consumption.

Houses with one refrigerator consume 79 kWh/month on average whereas the electricity consumption with two refrigerators increases to 84 kWh/month. In addition, it can be revealed that the effect of freezers on HEC is almost the same as that of refrigerators. Houses with one

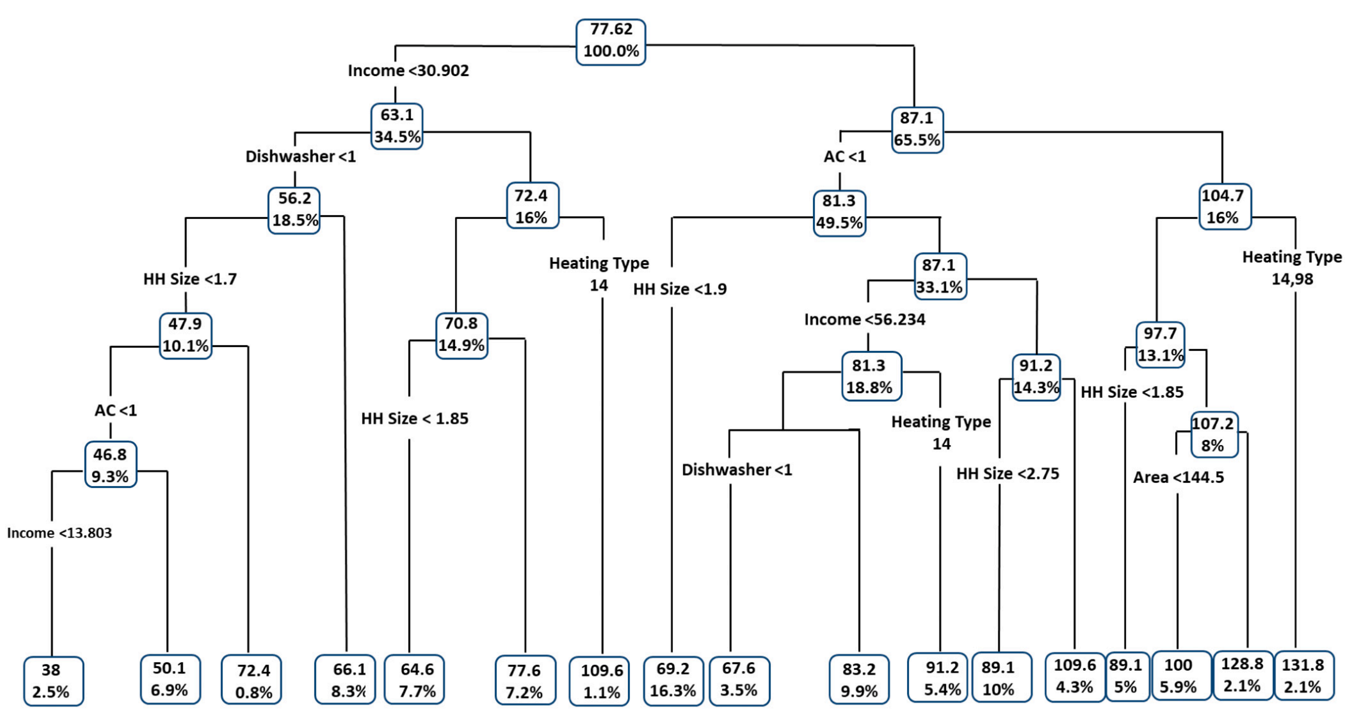

Fig. 8. Decision tree of electricity consumption.

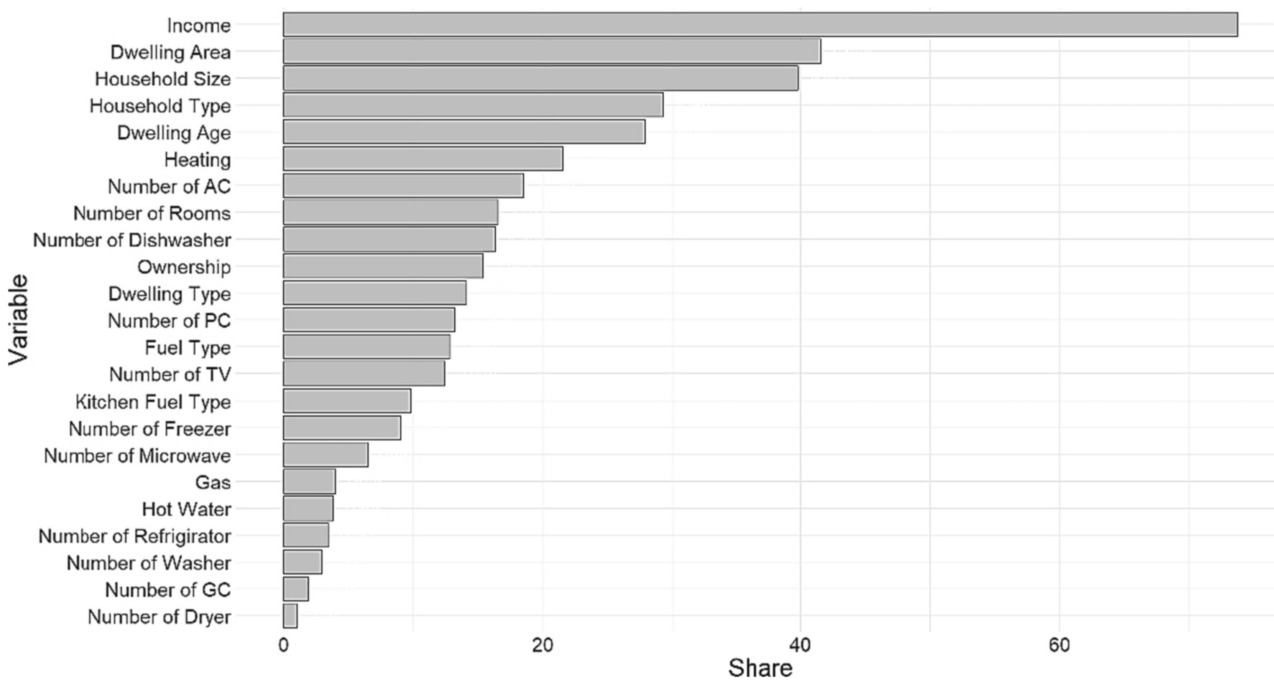

Fig. 9. Variable importance plot of random forest model.

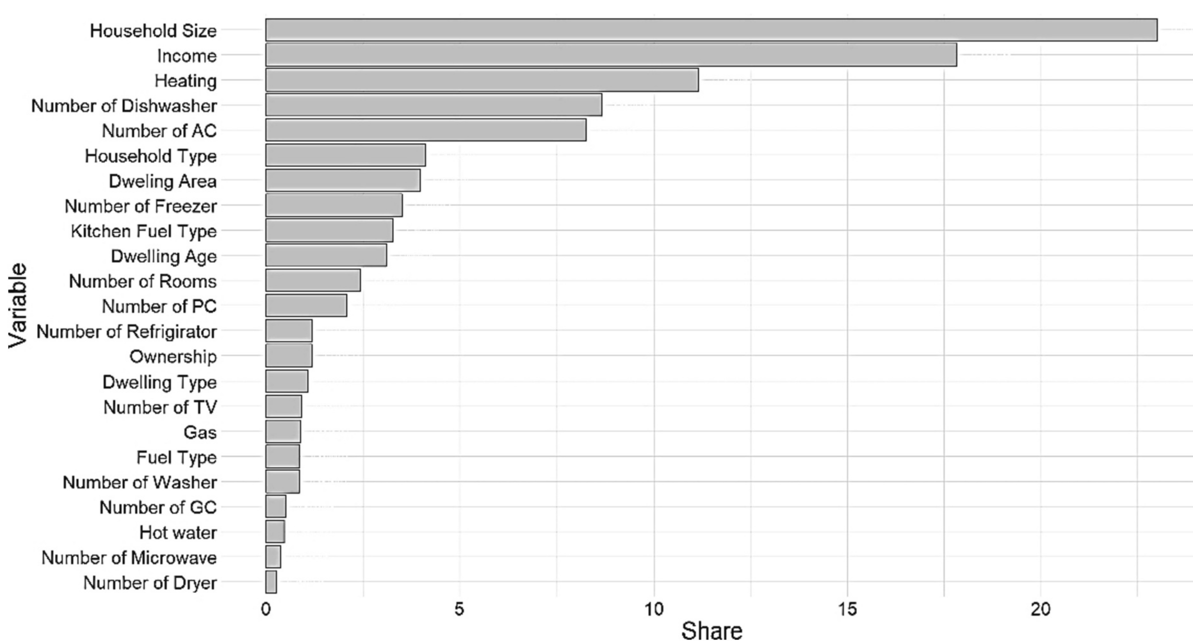

Fig. 10. Permutation based feature importance of variables.

dishwasher demand approximately 8 kWh/month more electricity. Microwave, washer and dryer have a very limited impact on household electricity consumption. Furthermore, a tumble dryer is not a widely-used household appliance due to the climate conditions, economic conditions and cultural preferences in Turkey.

Although air conditioners are not common in Turkey, it has a very powerful impact on household electricity demand. While one AC causes a 16.4 kWh/month increase in electricity consumption, two AC enhances the electricity consumption to 21 kWh/month.

When it comes to partial dependence plots of income and household size (See Fig. 12), it can be said that there is a positive correlation with HEC despite the fluctuations. Electricity consumption increases with the income of the household. That is, electricity is a normal good.

Regarding the household size, it has a positive influence on the electricity consumption as expected. The increasing number of people living in the apartment leads to consuming more electricity. However, the rate of increase in HEC is not the fold of the household size. The main reason for this is that most household appliances like refrigerators, dishwashers and air conditioners serve all household members at the same time. Thus, there is a decreasingly growing electricity demand for the increase in household size.

In the interaction plots, the predictors are investigated two at a time. An x–y grid is set up based on the possible combinations of predictors in the range of both variables. The rest of the predictors are held at either

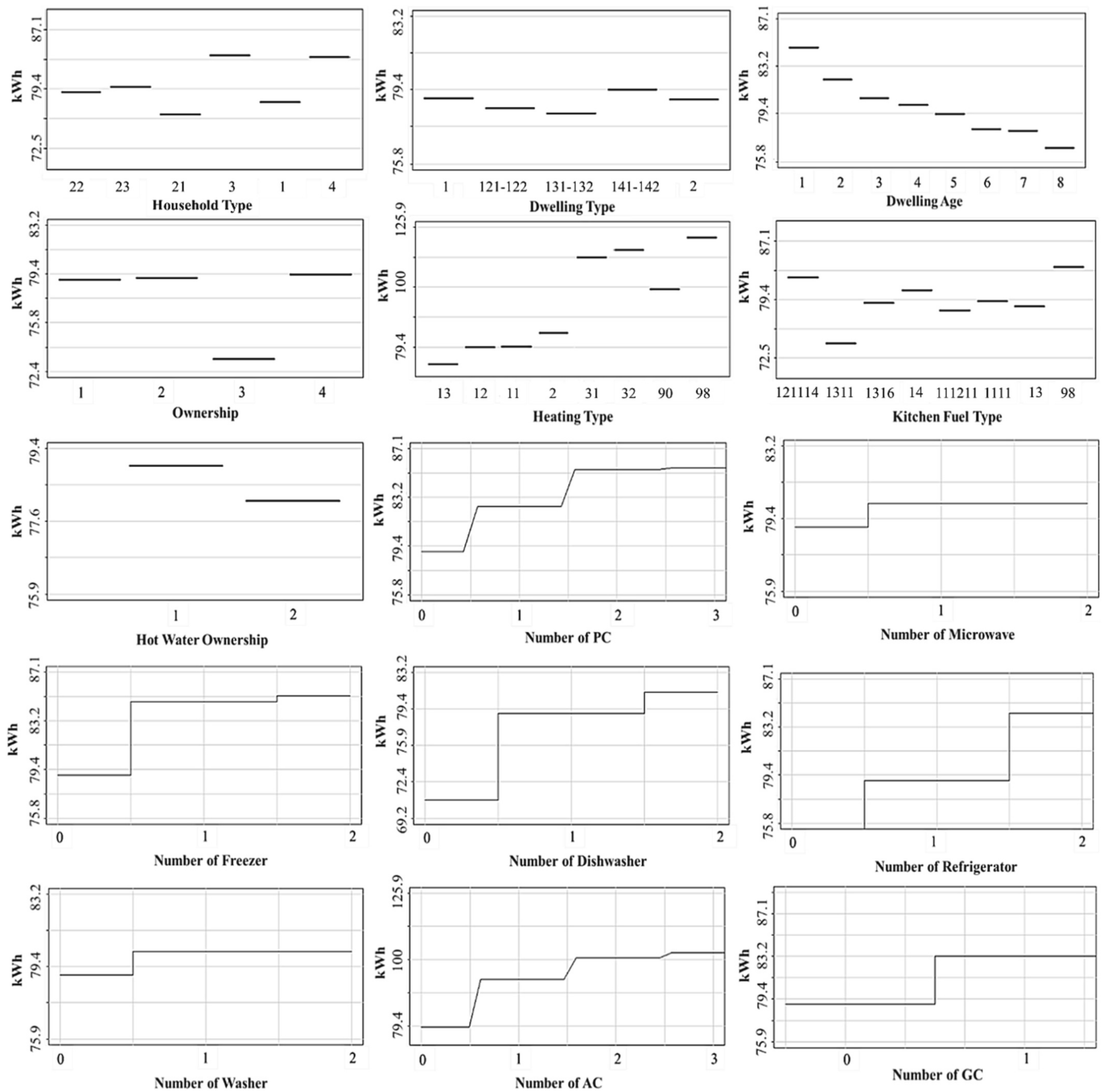

Fig. 11. Partial dependence plots of variables.

their means (for continuous values) or their most common value (for categorical variables). Model predictions are presented on the z-axis. The interaction plots of the most important variables are given in Fig. 13.

Discussion

Since this paper analyses the impact of dwelling properties, household characteristics and household appliances on residential electricity demands in Turkey, it principally focuses on and analogizes its outputs with previous studies specific to Turkey. However, none of the studies in the literature deeply present how and how much each determinant impacts household electricity consumption. Hence, in this section, only sorting of the impact of the variables is compared with previous studies.

As mentioned in the literature section, the number of studies specific to the Turkey case is very limited. In this context, Ozcan et al. (2013) stated that income is the most important component to determine HEC. This shows parallelism with our study in which income is found to be the second most important factor. Notwithstanding, Kayalica et al. (2020) presented that dwelling properties (e.g., floor area and number of rooms) are the most significant determinants for HEC dissimilar to this paper. Despite this difference, both this paper and Kayalica et al. (2020) revealed that household appliances have less impact on HEC compared to the dwelling features and household characteristics.

Furthermore, as the literature employs various methodologies for other countries/regions, it is worthwhile and stimulating to compare the results for Turkey to those for other countries/regions. To perform this, the most recent studies from the literature are gathered for comparison.

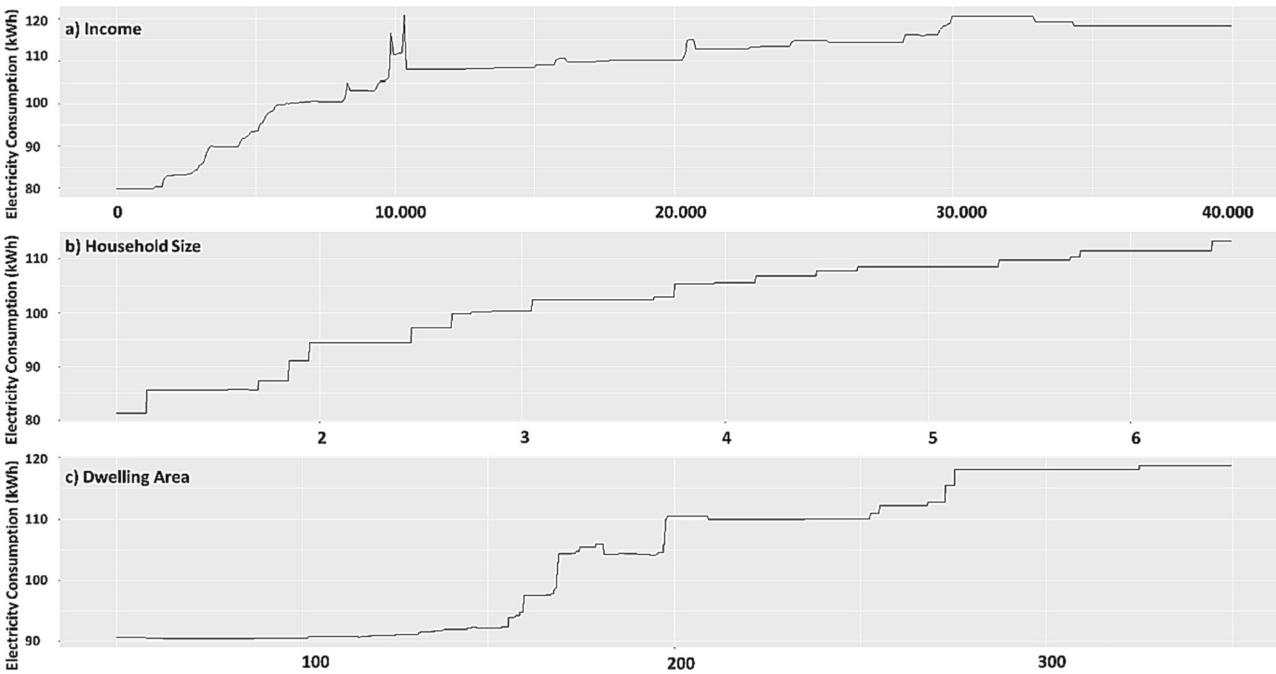

Fig. 12. Partial dependence plots of income, household size and floor area.

According to the results of both Kostakis (2020) and Kotsila and Polychronidou (2021), Greece, as a neighbour country of Turkey, shows parallelism to the findings of this paper. Both studies revealed that household size, income level and heating have a great influence on HEC. In another study, Besagni and Borgarello (2018) also found that socio-economic parameters, (e.g., household size and income) are the most influential factor in HEC in Italy. In light of this information, it can be said that our findings are in parallel with the literature.

In addition to providing the importance of variables in determining the residential electricity consumption, this paper also deeply presents how and how much each determinant impacts household electricity consumption as the results are given in the previous subsection. Using the generalisation capability of the HBS as mentioned in Data section, it is possible to implement the outputs of GBRT model to the total number of households in Turkey. Based on these results, it is possible to estimate potential electricity savings and possible to suggest policy implications for different determinants. In this way, the amount of potential CO_2 emission abatement and the social cost of this CO_2 may be calculated.

Although the most common heating fuel in Turkey is natural gas, according to the TurkStat HBS, 6.67 % of households use air conditioner for space heating. While households with floor heating systems or central heating systems consume 79.4 kWh/month of electricity, this value rises to more than 112 kWh/month for households with air conditioners. Based on the output of the partial dependence analysis done in the previous section, it is possible to calculate the impact of this difference on the environment.

Based on the HBS, approximately 1.69 million households are heated using electricity. In this case, compared to heating with natural gas, an extra 550.8 GWh of electricity is consumed annually as a result of heating with electricity. The cost of this amount of electricity production in Turkey (~0.055 $/kWh based on PwC (PricewaterhouseCoopers), 2021) is about 30.3 M $. In addition to electricity production cost, during electricity production, an average of 578 g-CO_2/kWh carbon dioxide is emitted in Turkey (author's calculation based on the EXIST (Energy Exchange Istanbul), 2021 data). However, 370 g-CO_2/kWh will be emitted in case of heating with natural gas instead of electric heaters (IPCC (Intergovernmental Panel on Climate Change), 2014). Compared to the use of natural gas, an extra 114,566 tons of CO_2 will be emitted into the atmosphere annually.

Dwelling age is one of the dwelling characteristics which has a noteworthy impact on HEC. Although Turkey lies in an earthquake-prone region, the building stock of Turkey is quite old. Approximately 62.3 % of households live in buildings older than 20 years which were built before the 1999 Izmit Earthquake. These buildings are both vulnerable to earthquakes and consume more electricity. As mentioned before, the newer buildings tend to demand less electricity due to the improvement in insulation and electrical infrastructure. Utilizing the results of PDP analysis, it can be possible to calculate potential electricity and CO_2 emission savings regarding the dwelling age (See Table 5). Based on this calculation, in case this building stock is totally renewed, 715.1 GWh/year less electricity would be consumed and 1.3 million tons/year less CO_2 emissions would be emitted into the atmosphere. The total economic value of these savings is more than 41.2 M $ per year.

The penetration rate of the dishwasher among households is about 70.7 % in Turkey (authors' calculation based on HBS data). If even half of these households switch to a higher energy class dishwasher, 20 % energy savings can be achieved, and thus, 171.9 GWh of electricity could be saved per annum. In this way, it is possible to save 9.5 M $ on electricity production costs. In addition, 99,369 tons/year less CO_2 emission, whose SCC is more than 443 k $/year, would be emitted into the atmosphere.

The penetration of the refrigerator into the household in Turkey has reached 100 % today. It is known that approximately 20 % of refrigerators (approximately 4.9 million units) are in class B and lower (Pala & Esen, 2018). By converting those to Class A or higher, energy usage decreases by 20 %, and approximately 929 GWh of energy can be saved annually. While the electricity generation cost is 51.1 M$, there is also a 537 kton CO_2 emission reduction potential whose SCC is more than 2.4 M $ per year.

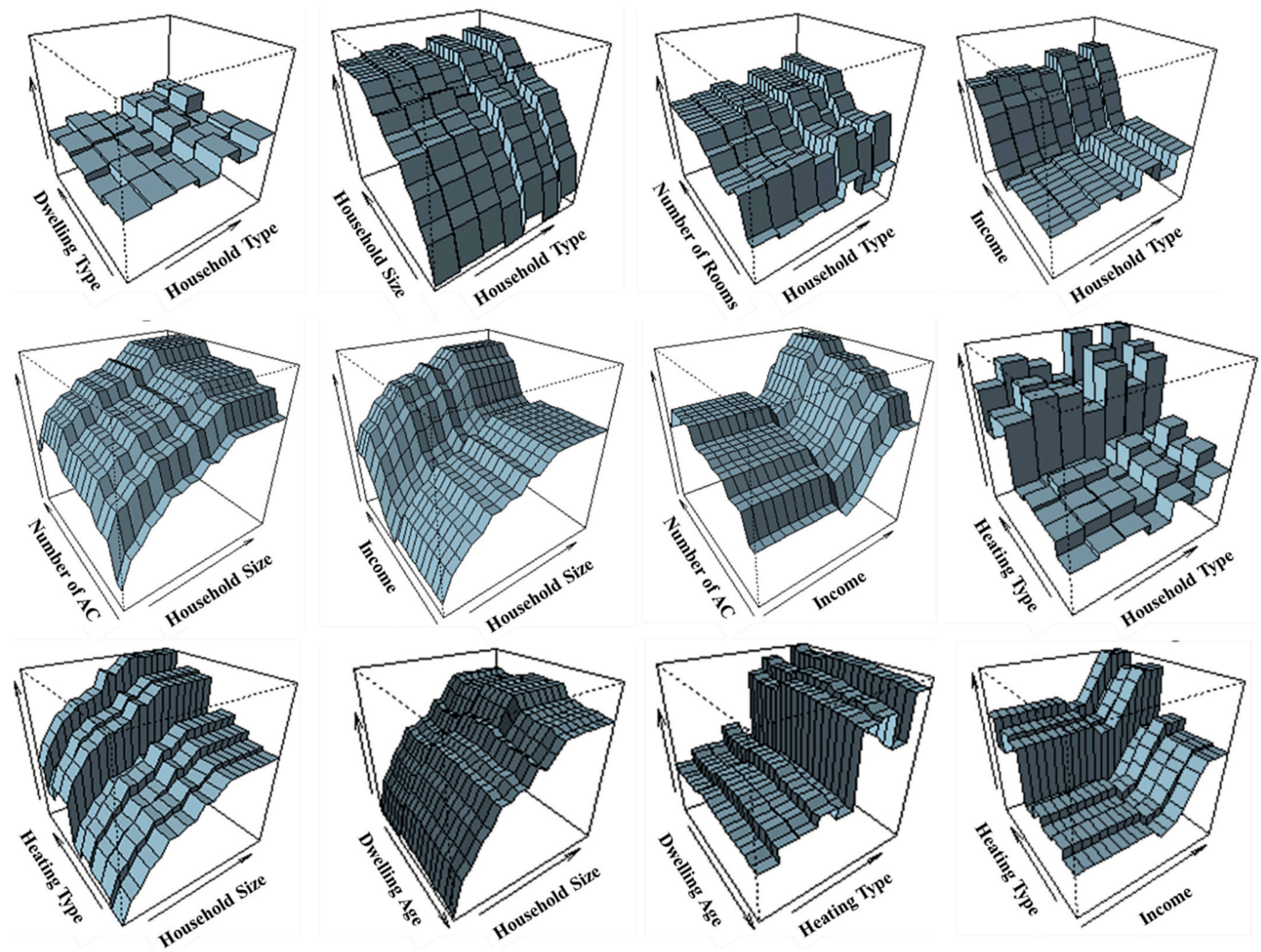

Fig. 13. Interaction plots of most important variables.

Table 5
Results of electricity consumption analysis for dwelling age.

Dwelling building year*	Consumption difference with new building** (kwh/year)	Percentage of Dwelling age (%)	Number of households	Potential electricity saving (GWh/year)	Potential electricity generation saving (M $/year)	CO_2 emission saving (ton/year)	SCC (k $/year)
1	96	1.4	363,202	34.87	1.92	20,153.38	89.9
2	72	3.8	962,250	69.28	3.81	40,045.02	178.6
3	60	7.1	1,806,578	108.39	5.96	62,652.1	279.4
4	54	12.5	3,160,333	170.66	9.39	986,640.3	439.9
5	50.4	17.5	4,438,617	223.71	12.30	129,302.2	576.7
6	21.4	19.9	5,056,533	108.21	5.95	62,545.3	278.9

As a Mediterranean country, Turkey is one of the countries which will be affected most by global warming (Lionello & Scarascia, 2018). Hence, air conditioning is getting more intention due to climate change, especially in the southern part of Turkey. As given in the previous section, AC causes a 16.4 kWh increase in electricity consumption. In Turkey, there are more than 4.75 million AC in residential buildings. With the integration of passive thermal management systems, such as TSW (thermochromic smart windows) and DPRC (daytime passive radiative coolers) which are two important technologies in this field, it is possible to save up to 17 % of electricity for air-conditioning (Lin et al., 2021). In this case, it can be said that there is a potential saving of 159.1 GWh of electricity and thus 91,970 tons of CO_2 emission per year. The monetary value of electricity generation and the SCC of CO_2 emission

Table 6
Incentive practices given to high efficiency white goods.

	Refrigerator			Washer	Dishwasher	Date
	A++	A+	A	AA	AA	
Italy	20 % discount					Since 2007
France	80€	30€				Since 2009
Belgium	150€	150€		150€		Since 2008
Spain	125€	105€	85€	105€	105€	Since 2008
Germany	150€					2009–2012
Denmark	130€	65€	65€			2006
Netherland	100€	100€	50€	100€	50€	2000–2003
USA	75-175			100$	100$	Since 2005

are calculated as 8.75 M $/year and 410.1 k $ per year, respectively.

Although the use of electric stoves in the kitchen consumes more electricity than natural gas and LPG, its use is not common (<0.8 %) in Turkey. Therefore, in general terms, it has little or no effect on electricity consumption.

Replacing inefficient old white goods with energy-efficient new ones provides great advantages to individuals, countries and the world. Therefore, modern world countries have been supporting this change as a state policy for years, and they accelerate this change with incentives. Some examples of incentives which are brought into force by modern world countries in different years are given in Table 6.

Policy implications

The social cost of CO_2 emission (SCC) for Turkey, which includes impacts on the environment and the human health, is estimated as 4.46 $/ton CO_2 for business-as-usual (BAU) scenario (Ricke, Drouet, Caldeira, & Tavoni, 2018). In the circumstances, the use of electricity instead of natural gas for heating results in a social cost of approximately 510.9 k $ per year in Turkey. Thus, the government needs to encourage a more environmentally friendly approach. That is, it should support the usage of natural gas instead of electricity in space heating, considering the high share of coal-based electricity generation in the electricity mix in Turkey. With the transition towards more cleaner electricity generation technologies, the average CO_2/kWh emitted from electricity generation (now 578 g-CO_2/kWh) would be less than natural gas (370 g-CO_2/kWh) in the future. After the threshold of 370 g-CO_2/kWh, electricity in space heating will be more environmentally friendly compared to natural gas. Furthermore, regulations should be changed to restrict the use of combi-type individual heating systems in buildings that have more than four independent areas. Conversion from the central system to the individual system should be made more difficult in existing buildings. Electricity generation should be provided while performing district heating where appropriate.

Renewing old buildings would create a more durable building stock to earthquakes and provide a great emission and energy savings. Since the lands in the big city centres are very valuable, efficient buildings can be built by demolishing and renovating the buildings with suitable floor conditions. In order to encourage this, some centres with very old buildings may be allowed to raise a floor, on the condition of making very good insulation and using efficient heating and hot water installations, beyond the regulations. Financial support should be provided for the rehabilitation of existing buildings (roof insulation, double glazing application) that ordinary citizens can benefit from, and allocations should be made from the funds created by the contributions of the energy service sector (natural gas and electricity sales companies).

It is also possible to monitor natural gas and electricity consumption on the basis of households or buildings, and to create coloured energy performance maps at the scale of neighbourhoods and settlements by obtaining indicators such as kWh/m^2, through the Geographic Information System (GIS), which is applied in developed countries and especially started to be used by metropolitan municipalities in Turkey. These maps concrete data in terms of determining priority places and encourage households in the selection of building rehabilitation areas.

The success in energy-saving efforts in all developed countries has been achieved by the financial support of the government. A fund should be established to provide funds for incentives and especially household savings investments. The savings from electricity generation could be transferred into this fund system. As a very low-interest credit system, it should be managed with a transparent mechanism including the banking sector, and the fund should be designed to terminate itself after a while following meeting the targets set from the beginning.

Conclusion

The objective of this paper is to detect the determinants of household electricity consumption in Turkey based on the Household Budget Survey using three different tree-based machine learning methods, namely Decision Tree, Random Forest and Gradient Boosted Regression Tree. In this study, the accuracies of established models are examined with three different error measurements; Mean Absolute Error, Mean Absolute Percentage Error and Root Mean Square Error. Since the GBRT method provides the lowest RMSE, the impact of each variable on the electricity consumption is analysed with this method. Household size is found to be the most important variable, while income level and heating type are discovered as 2nd and 3rd most prominent determinants for household electricity demand forecasting. With the help of the Partial Dependence Plots provided by machine learning techniques, the impact of each categorical and continuous variable is presented. Based on the results of partial dependence plots, the monetary values of both electricity generation and the social cost of CO_2 emissions emitted into the atmosphere due to electricity generation are calculated for most important determinants. As a result, it can be revealed that there is a huge potential to save on energy and decrease CO_2 emissions by changing existing inefficient household equipment with efficient ones, renewing or rebuilding old residential buildings, and applying state-of-the-art energy management systems and technologies.

There are limits of this study. Turkish HBS does not provide any information of the location of data points. Thus, it is not possible to include a location variable in the analysis. Moreover, during the Covid-19 pandemics, this survey could not be conducted for 2020 and 2021. The next HBS is expected to be published in mid-2024.

Since this study focuses on three tree-based ML methods, future studies should investigate the determinants of HEC using different ML algorithms such as Extreme Gradient Boosting Trees etc. Furthermore, future studies could delve into non-household level elements. These might encompass aspects like policy and regulatory dynamics, macro-economic influences, electricity pricing mechanisms, and environmental considerations. These factors, which extend beyond the scope of the present review, could be subjects of exploration in subsequent research.

Availability of data and materials

Data of this study were collected from the TurkStat. The research data cannot be shared publicly, because the individual privacy could be compromised.

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Table A.1
A brief summary of the literature.

Reference	Region	Method	Survey Size	Factors*				
				SE	DW	APP	CL	LOC
Druckman and Jackson (2008)	UK	Local Area Resource Analysis	~7000	X	X			X
McLoughlin, Duffy, and Conlon (2012)	Ireland	Multiple Linear Regression	4200	X	X			
Santamouris et al. (2013)	Greece	Cluster Analysis	598	X	X	X	X	
Bedir et al. (2013)	Netherlands	Regression Analysis	323	X	X	X		
Ozcan et al. (2013)	Turkey	Multinomial Logit Model	181,548	X	X			X
Kavousian et al. (2013)	USA	Stepwise Regression	1628	X	X	X	X	X
Wijaya and Tezuka (2013)	Indonesia	Multivariate data analysis	100	X	X	X		X
Guloglu and Akin (2014)	Turkey	Partial Proportional Odds Model	8640	X	X	X		
Laicāne et al. (2014)	Latvia	Regression Analysis	500	X	X	X	X	X
Romero-Jordán, Peñasco, and del Río (2014)	Spain	Partial Adjustment Model		X				
Longhi (2015)	UK	Ordinary Least Squares	21,393	X	X		X	
Huang (2015)	Taiwan	Quantile Regression Analysis	~15,000	X	X	X		X
Jones et al. (2015)	–	Review						
Esmaeilimoakher, Urmee, Pryor, and Baverstock (2016)	Australia	Survey Analysis	9	X	X		X	
Danlami (2017)	Nigeria	Ordinary Least Squares	548	X	X	X		
Yalcintas and Kaya (2017)	Hawaii	Linear Regression Analysis		X				
Besagni and Borgarello (2018)	Italy	Ordinary Least Squares	15,013	X	X	X		X
Lévy and Belaïd (2018)	France	Hierarchical Cluster Analysis	29,249	X	X			
Hernandez and Patiño-Echeverri (2019)	Mexico	Random Forest and Multinomial Logistic Regression	32,047	X	X	X	X	X
Zhang et al. (2020)	China	Seemingly Unrelated Regression (SUR)		X	X	X		
Trotta (2020)	Denmark	K-means Clustering and Multinomial Probit Regression	19,734	X	X			X
Kayalica et al. (2020)	Turkey	ANN and AHP	10,427	X	X	X		
Đurišić et al. (2020)	Montenegro	Structural Equation Modelling	964	X	X	X		X
Kim et al. (2020)	South Korea	Multiple Linear Regression and Decision Tree	71	X	X	X		
Huang (2020)	Taiwan	Logarithmic Mean Divisia Index	7677			X	X	
Kostakis (2020)	Greece	Quantile Regression Analysis	6176	X				X
Kim (2020)	South Korea	Ordinary Least Squares and Quantile Regression	2520	X	X	X		
Son and Yoon (2020)	Vietnam	Ordinary Least Squares and Quantile Regression Analysis		X	X			
Ali et al. (2021)	Malaysia	Multiple linear regressions	620	X	X			
Kotsila and Polychronidou (2021)	Greece	Ordinary Least Squares and log-linear regression	1801	X	X	X	X	
Ofetotse, Essah, and Yao (2021)	Botswana	K-means cluster analysis	310	X	X	X		
Li and Zhang (2022)	China	Tobit model	3232	X	X	X		
Park and Yun (2022)	Korea	Ordinary Least Squares and panel	225	X	X	X	X	
Nsangou et al. (2022)	Cameroon	ANN, Ordinary Least Squares and Decision Tree	1013	X	X	X	X	
Yarbaşı and Çelik (2023)	Turkey	Heckman Sample Selection	10,740	X	X	X		
Tete, Soro, Sidibé, and Jones (2023)	Burkina Faso	Survey Analysis	387	X	X	X		
Debebe, Senbeta, Diriba, Teferi, and Teketay (2023)	Ethiopia	Multivariate probit regression model	420	X	X	X		X
Karaaslan and Algül (2023)	Turkey	Quantile Regression Analysis	10,740	X	X	X		

* SE: Socio-economic, DW: Dwelling properties, APP: Household appliances, CL: Climate, LOC: Location.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Ali, S. S. S., Razman, M. R., Awang, A., Asyraf, M. R. M., Ishak, M. R., Ilyas, R. A., & Lawrence, R. J. (2021). Critical Determinants of Household Electricity Consumption in a Rapidly growing City. *Sustainability, 13*(8), 4441. https://doi.org/10.3390/su13084441

Bedir, M., Hasselaar, E., & Itard, L. (2013). Determinants of electricity consumption in Dutch dwellings. *Energy and Buildings, 58*, 194–207. https://doi.org/10.1016/j.enbuild.2012.10.016

Besagni, G. M., & Borgarello, M. (2018). The determinants of residential energy expenditure in Italy. *Energy, 165*, 369–386. https://doi.org/10.1016/j.energy.2018.09.108

Bhattacharyya, S. C. (2011). *Energy economics: Concepts, issues, markets and governance* (Springer Science & Business Media).

BP (British Petroleum). (2021). *Statistical Review of World Energy*, 2021.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32.

Danlami, A. H. (2017). Determinants of household electricity consumption in Bauchi State, Nigeria. *Hyperion Economic Journal, 5*(1), 16–28.

Debebe, B., Senbeta, F., Diriba, D., Teferi, E., & Teketay, D. (2023). Determinants of household energy choice for domestic chores: evidence from the Semien Mountains National Park and Adjacent Districts, Northwest Ethiopia. *Cleaner Energy Systems, 4*, 100063. https://doi.org/10.1016/j.cles.2023.100063

Druckman, A., & Jackson, T. (2008). Household energy consumption in the UK: a highly geographically and socio-economically disaggregated model. *Energy Policy, 36*(8), 3177–3192. https://doi.org/10.1016/j.enpol.2008.03.021

Đurišić, V., Rogić, S., Smolović, J. C., & Radonjić, M. (2020). Determinants of household electrical energy consumption: Evidences and suggestions with application to Montenegro. *Energy Reports, 6*, 209–217. https://doi.org/10.1016/j.egyr.2019.10.039

Esmaeilimoakher, P., Urmee, T., Pryor, T., & Baverstock, G. (2016). Identifying the determinants of residential electricity consumption for social housing in Perth, Western Australia. *Energy and Buildings, 133*, 403–413. https://doi.org/10.1016/j.enbuild.2016.09.063

EXIST (Energy Exchange Istanbul). 2021. Real Time Generation: 01.01.2021–31.12.2021 [dataset]. Accessed June 13, 2022. https://seffaflik.epias.com.tr/transparency/uretim/gerceklesen-uretim/gercek-zamanli-uretim.xhtml.

Guloglu, B., & Akin, E. (2014). Türkiye'de Hane Halkları Elektrik Talebininin Belirleyicileri: Sıralı Logit Yaklaşımı. *Siyaset, Ekonomi ve Yönetim Araştırmaları Dergisi, 2*(3), 1–20.

Hernandez, M. H., & Patiño-Echeverri, D. (2019). Classifying Residential Electricity demand in Mexico using Random Forest and Multinomial Logistic Regression. In *2019 FISE-IEEE/CIGRE Conference-Living the Energy Transition (FISE/CIGRE)* (pp. 1–6).

Huang, W. H. (2015). The determinants of household electricity consumption in Taiwan: evidence from quantile regression. *Energy, 87*, 120–133. https://doi.org/10.1016/j.energy.2015.04.101

Huang, Y. H. (2020). Examining impact factors of residential electricity consumption in Taiwan using index decomposition analysis based on end-use level data. *Energy, 213*, 119067. https://doi.org/10.1016/j.energy.2020.119067

IEA (International Energy Agency). (2019). *Global Energy & CO2 Status Report 2019*. Paris: IEA. https://www.iea.org/reports/global-energy-co2-status-report-2019/emissions. (Accessed 4 July 2022).

IEA (International Energy Agency). (2020). *Energy Technology Perspectives 2020*. Paris: IEA. https://www.iea.org/reports/energy-technologyperspectives-2020. (Accessed 12 July 2022).

IEA (International Energy Agency). 2020b. World Energy Balances 2020, IEA, Paris. Accessed July 13, 2022. https://www.iea.org/data-andstatistics?country=WORLD&fuel=Energy%20supply&indicator=TPESbySource.

IPCC (Intergovernmental Panel on Climate Change). (2014). Annex II: Metrics & Methodology. In *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge*. Cambridge University Press.

Jones, R. V., Fuertes, A., & Lomas, K. J. (2015). The socio-economic, dwelling and appliance related factors affecting electricity consumption in domestic buildings. *Renewable and Sustainable Energy Reviews, 43*, 901–917. https://doi.org/10.1016/j.rser.2014.11.084

Karaaslan, K.Ç., & Algül, Y. (2023). Determinants of energy expenditures for Turkish households using quantile regression and data from an original survey in Turkey. *Environmental Science and Pollution Research, 30*(13), 38939–38954. https://doi.org/10.1007/s11356-022-24323-8

Kavousian, A., Rajagopal, R., & Fischer, M. (2013). Determinants of residential electricity consumption: using smart meter data to examine the effect of climate, building characteristics, appliance stock, and occupants' behavior. *Energy, 55*, 184–194. https://doi.org/10.1016/j.energy.2013.03.086

Kayalica, M. O., Ozozen, A., Guven, D., Kayakutlu, G., & Bayar, A. A. (2020). Electricity consumption analysis based on Turkish Household Budget surveys. *Energy, Ecology and Environment, 5*(6), 444–455. https://doi.org/10.1007/s40974-020-00193-z

Kim, M. J. (2020). Understanding the determinants on household electricity consumption in Korea: OLS regression and quantile regression. *The Electricity Journal, 33*(7), 106802. https://doi.org/10.1016/j.tej.2020.106802

Kim, Y. J., Lee, S. J., Jin, H. S., Suh, I. A., & Song, S. Y. (2020). Comparison of Linear and Nonlinear Statistical Models for Analyzing Determinants of Residential Energy Consumption. *Energy and Buildings, 223*, 110226. https://doi.org/10.1016/j.enbuild.2020.110226

Kostakis, I. (2020). Socio-demographic determinants of household electricity consumption: evidence from Greece using quantile regression analysis. *Current Research in Environmental Sustainability, 1*, 23–30. https://doi.org/10.1016/j.crsust.2020.04.001

Kotsila, D., & Polychronidou, P. (2021). Determinants of household electricity consumption in Greece: a statistical analysis. *Journal of Innovation and Entrepreneurship, 10*(1), 1–20. https://doi.org/10.1186/s13731-021-00161-9

Lacis, A. A., Hansen, J. E., Russell, G. L., Oinas, V., & Jonas, J. (2013). The role of long-lived greenhouse gases as principal LW control knob that governs the global surface temperature for past and future climate change. *Tellus B: Chemical and Physical Meteorology, 65*(1), 19734. https://doi.org/10.3402/tellusb.v65i0.19734

Laicāne, I., Blumberga, A., Rošā, M., & Blumberga, D. (2014). Determinants of household electricity consumption savings: a Latvian case study. *Agronomy Research, 12*(2), 527–542.

Lévy, J. P., & Belaïd, F. (2018). The determinants of domestic energy consumption in France: Energy modes, habitat, households and life cycles. *Renewable and Sustainable Energy Reviews, 81*, 2104–2114. https://doi.org/10.1016/j.rser.2017.06.022

Lewis, C. D. (1982). *Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting*. Butterworth-Heinemann.

Li, Z., L., Lv, and Z. Zhang. 2022. Research on the characteristics and influencing factors of chinese urban households' electricity consumption efficiency. Energies 15(20): 7748. doi:https://doi.org/10.3390/en15207748.

Lin, K., Chao, L., Lee, H. H., Xin, R., Liu, S., Ho, T. C., et al. (2021). Potential building energy savings by passive strategies combining daytime radiative coolers and thermochromic smart windows. *Case Studies in Thermal Engineering, 28*, 101517. https://doi.org/10.1016/j.csite.2021.101517

Lionello, P., & Scarascia, L. (2018). The relation between climate change in the Mediterranean region and global warming. *Regional Environmental Change, 18*(5), 1481–1493. https://doi.org/10.1007/s10113-018-1290-1

Longhi, S. (2015). Residential energy expenditures and the relevance of changes in household circumstances. *Energy Economics, 49*, 440–450. https://doi.org/10.1016/j.eneco.2015.03.018

Lucon, O., Ürge-Vorsatz, D., Ahmed, A. Z., Akbari, H., Bertoldi, P., Cabeza, L. F., et al. (2014). Buildings. In O. Edenhofer, R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, … J. C. Minx (Eds.), *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., et al. (2020). From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence, 2*(1), 56–67. https://doi.org/10.1038/s42256-019-0138-9

McLoughlin, F., Duffy, A., & Conlon, M. (2012). Characterising domestic electricity consumption patterns by dwelling and occupant socio-economic variables: an Irish case study. *Energy and Buildings, 48*, 240–248. https://doi.org/10.1016/j.enbuild.2012.01.037

Nsangou, J. C., Kenfack, J., Nzotcha, U., Ekam, P. S. N., Voufo, J., & Tamo, T. T. (2022). Explaining household electricity consumption using quantile regression, decision tree and artificial neural network. *Energy, 250*, 123856. https://doi.org/10.1016/j.energy.2022.123856

Ofetotse, E. L., Essah, E. A., & Yao, R. (2021). Evaluating the determinants of household electricity consumption using cluster analysis. *Journal of Building Engineering, 43*, 102487. https://doi.org/10.1016/j.jobe.2021.102487

Ozcan, K. M., Gulay, E., & Ucdogruk, S. (2013). Economic and demographic determinants of household energy use in Turkey. *Energy Policy, 60*, 550–557. https://doi.org/10.1016/j.enpol.2013.05.046

Pala, B. D., & Esen, G. K. (2018). Türkiye'de Ev Tipi Elektrikli Cihazlarda Enerji Etiketi Ve Enerji Verimliliği Farkindaliği Analizi. *Verimlilik Dergisi, 2*, 81–97.

Park, J., & Yun, S. J. (2022). Social determinants of residential electricity consumption in Korea: Findings from a spatial panel model. *Energy, 239*, 122272. https://doi.org/10.1016/j.energy.2021.122272

PwC (PricewaterhouseCoopers). (2021). *Overview of the Turkish Electricity Market Report*, 2021. https://www.invest.gov.tr/en/library/publications/lists/investpublications/overview-of-turkish-electricity-market.pdf. (Accessed 16 July 2022).

Ricke, K., Drouet, L., Caldeira, K., & Tavoni, M. (2018). Country-level social cost of carbon. *Nature and Climate Change, 8*, 895–900. https://doi.org/10.1038/s41558-018-0282-y

Ritchie, H., M. Roser, M. 2019. CO_2 and Greenhouse Gas Emissions. Accessed July 15, 2022. https://ourworldindata.org/co2-and-other-greenhouse-gas-emissions.

Romero-Jordán, D., Peñasco, C., & del Río, P. (2014). Analysing the determinants of household electricity demand in Spain. An econometric study. *International Journal of Electrical Power & Energy Systems, 63*, 950–961. https://doi.org/10.1016/j.ijepes.2014.06.005

Santamouris, M., Paravantis, J. A., Founda, D., Kolokotsa, D., Michalakakou, P., Papadopoulos, A. M., et al. (2013). Financial crisis and energy consumption: a household survey in Greece. *Energy and Buildings, 65*, 477–487. https://doi.org/10.1016/j.enbuild.2013.06.024

SIO (Scripps Institution of Oceanography). 2020. Scripps CO2 Program: Carbon dioxide concentration at Manua Loa Observatory. Accessed June 11, 2022. https://scripps.ucsd.edu/programs/keelingcurve/.

Son, H., & Yoon, S. (2020). Reducing energy poverty: Characteristics of household electricity use in Vietnam. *Energy for Sustainable Development, 59*, 62–70. https://doi.org/10.1016/j.esd.2020.08.007

TEDC (Turkish Electricity Distribution Corporation). 2019. Electricity Generation and Transmission Statistics [dataset]. Accessed July 18, 2022. https://www.teias.gov.tr/tr-TR/turkiye-elektrik-uretim-iletim-istatistikleri.

TEDC (Turkish Electricity Distribution Corporation). 2021a. Electricity Generation and Shares by Energy Resources of Turkey [dataset].

TEDC (Turkish Electricity Distribution Corporation). 2021b. Electricity Distribution and Consumption Statistics of Turkey Report [dataset].

Tete, K. H., Soro, Y. M., Sidibé, S. S., & Jones, R. V. (2023). Urban domestic electricity consumption in relation to households' lifestyles and energy behaviours in Burkina Faso: Findings from a large-scale, city-wide household survey. *Energy and Buildings, 285*, 112914. https://doi.org/10.1016/j.enbuild.2023.112914

Trotta, G. (2020). An empirical analysis of domestic electricity load profiles: who consumes how much and when? *Applied Energy, 275*, 115399. https://doi.org/10.1016/j.apenergy.2020.115399

TurkStat (Turkish Statistical Institute). 2019. Household Budget Survey Consumption Expenditures Combined Micro Data Set [dataset].

TurkStat (Turkish Statistical Institute). 2022. Number of households by size and type, 2015–2021. Accessed July 27, 2022. https://data.tuik.gov.tr/Bulten/DownloadIstatistikselTablo?p=4ChPevhNXoD1Z4Evl5HUrgYtwAlOmbR25a8rEXavtQYuzQT0TCghIu6t7r2ZyTS5.

UNFCCC. 2015. Republic of Turkey Intended Nationally Determined Contribution Report. https://www4.unfccc.int/sites/submissions/INDC/Published%20Documents/Turkey/1/The_INDC_of_TURKEY_v.15.19.30.pdf.

Wang, T., Hu, S., & Jiang, Y. (2021). Predicting shared-car use and examining nonlinear effects using gradient boosting regression trees. *International Journal of Sustainable Transportation, 15*(12), 893–907. https://doi.org/10.1080/15568318.2020.1827316

Wijaya, M. E., & Tezuka, T. (2013). A comparative study of households' electricity consumption characteristics in Indonesia: a techno-socioeconomic analysis. *Energy for Sustainable Development, 17*(6), 596–604. https://doi.org/10.1016/j.esd.2013.09.004

Wonnacott, T. H., & Wonnacott, R. J. (1990). *Introductory Statistics* (5th ed.). Wiley.

Xu, L., Chen, N., Zhang, X., Chen, Z., Hu, C., & Wang, C. (2019). Improving the north American multi-model ensemble (NMME) precipitation forecasts at local areas using wavelet and machine learning. *Climate Dynamics, 53*(1), 601–615. https://doi.org/10.1007/s00382-018-04605-z

Yalcintas, M., & Kaya, A. (2017). Roles of income, price and household size on residential electricity consumption: Comparison of Hawaii with similar climate zone states. *Energy Reports, 3*, 109–118. https://doi.org/10.1016/j.egyr.2017.07.002

Yarbaşı, İ. Y., & Çelik, A. K. (2023). The determinants of household electricity demand in Turkey: an implementation of the Heckman Sample selection model. *Energy, 283*, 128431. https://doi.org/10.1016/j.energy.2023.128431

Yoon, J. (2021). Forecasting of real GDP growth using machine learning models: Gradient boosting and random forest approach. *Computational Economics, 57*(1), 247–265. https://doi.org/10.1007/s10614-020-10054-w

Zhang, J., Teng, F., & Zhou, S. (2020). The structural changes and determinants of household energy choices and energy consumption in urban China: addressing the role of building type. *Energy Policy, 139*, 111314. https://doi.org/10.1016/j.enpol.2020.111314

Zhang, Y., & Haghani, A. (2015). A gradient boosting method to improve travel time prediction. *Transportation Research Part C: Emerging Technologies, 58*, 308–324. https://doi.org/10.1016/j.trc.2015.02.019