PREDICTING PRIMARY ENERGY SAVINGS OF BUILDING RETROFIT MEASURES WITH DECISION-TREE-BASED ENSEMBLE METHODS

UDC 620.9:728.2]:005.322

Mirko M. Stojiljković, Marko G. Ignjatović, Goran D. Vučković

University of Niš, Faculty of Mechanical Engineering in Niš, Serbia

Abstract. *Primary energy is the quantity often used to express the total amount of consumed or saved energy. It is especially important for certification and standards compliance evaluation of buildings. In order to accurately assess primary energy consumption and savings of building retrofit measures, one needs adequate models that can be based on the knowledge of phenomena, on the collected data, or both. This paper analyzes the learning performance and accuracy of data-driven models based on decision trees that predict primary energy savings related to building retrofit measures. It uses supervised machine learning methods — regression decision trees and ensemble methods based on them — to train, validate, and test such models. Ensemble methods based on decision trees are powerful, accurate, but also convenient to prepare and fast to train. In addition, they calculate the relative importance of each feature. The research results with highly accurate data-driven models and consistent feature importance values.*

Key words: *buildings, data-driven models, ensemble methods, primary energy savings, supervised machine learning*

1. INTRODUCTION

Primary energy (PE) is a convenient quantity often used to express the total amount of consumed and saved energy when evaluating compliance to energy performance standards or certifying energy performance of buildings [1, 2]. It heavily depends on the properties of the thermal envelope of buildings, installed energy conversion and distribution systems, and occupants' behavior.

Accurate models for the prediction of energy performance are very important for planning and optimization related to buildings [3]. Foucquier et al. [4] divided these techniques into three categories: (1) white-box methods, (2) black-box methods, and (3) hybrid methods. White-box methods are based on physical models and can be further classified according to

the level of details they consider. Black-box methods use the existing data related to the building performance measure of interest and apply machine learning techniques. White-box methods are suitable when the physical phenomena are well understood and adequate input data exists. However, they can be overly complex and computationally-intensive, especially when high precision is required. Black-box methods are more efficient, do not rely on physical models, but their performance significantly depends on the volume and quality of used data. Hybrid methods combine white- and black-box approaches taking the advantages of both. Deb et al. [3] underlined that precise evaluation of energy consumption might be very hard when using white-box methods, due to complex physical models and that data-driven models can be an important alternative. Seyedzadeh et al. [5] consider machine learning methods promising for planning retrofit of complex buildings when intensive white-box calculations are an alternative.

Mosavi et al. [6] presented a wide range of machine learning applications related to energy systems and buildings including prediction of energy consumption, load, costs, solar radiation, wind speed, power quality, etc. Supervised machine learning methods are successfully used to predict heating and cooling demand of buildings [7–9], electricity demand [10, 11], indoor air temperature [12, 13], occupancy and users behavior [14, 15], etc. They can be applied to ensure thermal comfort [16–18], aid the optimization process, usually in combination with building simulation and genetic algorithm [19–23], or improve building control implementation [16, 24–29]. Unsupervised machine learning methods are also used for solving building-related problems [30]. For example, cluster analysis is often combined with supervised methods [31–34], often as a preprocessing technique. Reinforcement learning can be exploited for short-term prediction of energy consumption [35], especially in control applications [36].

Although many types of machine learning methods are present in buildings-related applications, regression analysis is dominant. It is usually performed with artificial neural networks (ANN) or support vector machines (SVM), while random forest (RF), gradient boosting (GB), linear regression, and other methods are used on several occasions.

Ahmad et al. [37] compared feedforward ANN and RF in terms of performance when predicting the energy consumption of a hotel building. Although ANN was slightly more accurate, they concluded that both methods are suitable to predict energy consumption and that the best method cannot be determined a priory, but by testing different options. They underlined that RF is faster to train, has fewer hyperparameters that need tuning, and handles categorical variables better and noticed that ensemble methods (including RF) are often ignored in buildings-related applications despite the results and attention they have in other fields. In the case considered by Touzani et al. [38], GB performed slightly better than RF when predicting energy consumption. Cui et al. [12] applied RF and GB (as well as other methods) as a part of a hybrid approach that predicts the indoor air temperature that combines the resistor-capacitor and black-box models. Seyedzadeh et al. [39] predicted heating and cooling loads with RF in the frame of a multi-objective optimization approach and Smarra [27] used if for data-driven model predictive control. Wang et al. [40] found extreme GB the best approach to predict cooling load one day ahead. RF was very well-performing when used for classification to learn human interactions [41] and predict occupancy [15].

This paper examines the possibility of using decision trees (DT) and the related ensemble methods — RF, GB, and extremely randomized trees (ET) — for predicting PE savings (PES) related to the energy retrofit measures of residential buildings connected to district heating systems.

2. PROBLEM FORMULATION

This paper applies DT and ensemble machine learning methods RF, ET, and GB to assess PES of retrofit measures for some of the typical Serbian residential buildings connected to district heating (DH) systems. Figure 1 displays the three-dimensional drawings of the buildings.
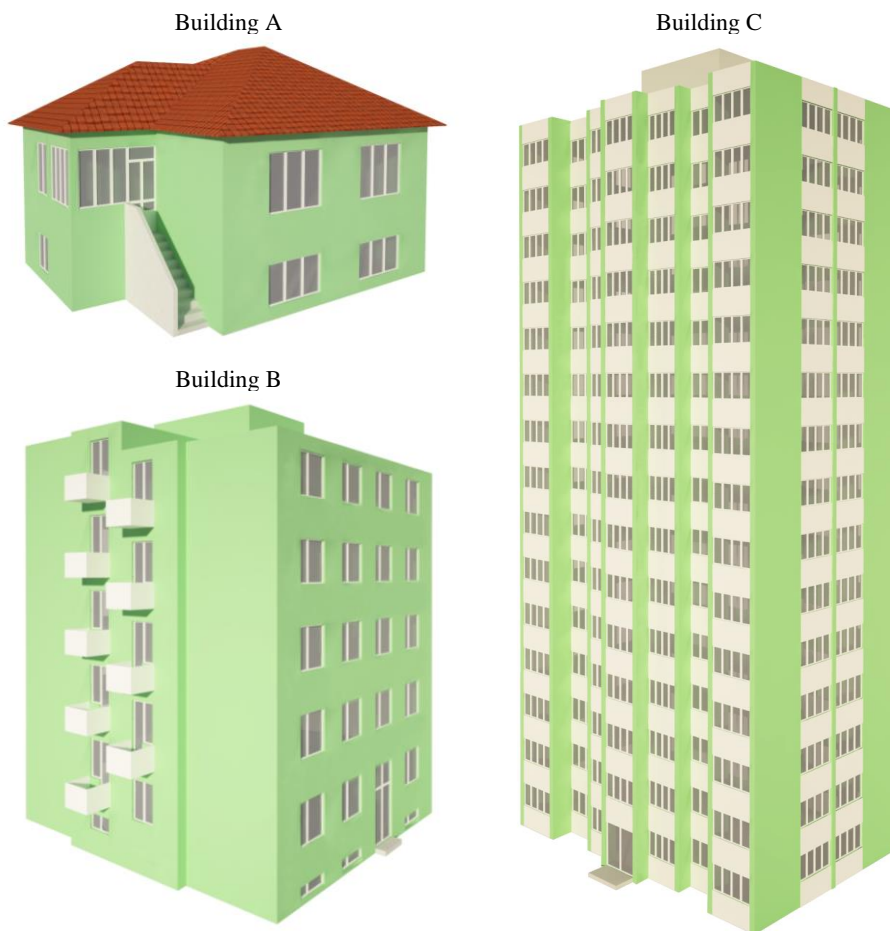
Building A

Building C

Building B

Fig. 1 Analyzed buildings

The main goal is to evaluate the performance of data-driven black-box surrogate models for predicting PES. The models should perform a regression analysis in which the output is PES and inputs are the quantities related to renovation measures. This paper focuses only on the improvement of the thermal envelope of a building, although the methodology might be extended to consider energy systems as well. There are six input variables (features):

1. The thermal conduction resistance of the insulation of exterior walls, in [m²K/W], calculated as $R_e=d/\lambda$, where d is the insulation thickness, in [m], and λ is the conductivity, in [W/(mK)];
2. The thermal conduction resistance of the insulation of interior walls, R_i, in [m²K/W];
3. The thermal conduction resistance of the insulation of floor, R_l, in [m²K/W];
4. The thermal conduction resistance of the insulation of ceiling or roof, R_u, in [m²K/W];
5. The U-value (thermal transmittance) of the windows, U_f, in [W/(m²K)];
6. The non-dimensional total solar energy transmittance of the windows, g.

The data for the regression analysis is obtained by calculating PES for a limited number of combinations of retrofit measures using the methodology described in [42]. PE conversion factors are also taken from [42]. Their values are $f_{DH}=1.8$ for DH and $f_E=2.5$ for electricity.

The regression analysis is performed for three existing residential buildings located in the City of Niš (Serbia), typical for Serbia. The buildings are heated from the DH system and cooled with electrical air-to-air split-systems. The main properties of the chosen buildings are shown in Table 1. It should be mentioned that the insulation of the interior walls has not been considered for building A.

Table 1 Main properties of considered buildings

Building	Floor area [m²]	No. of storeys	Type	Remark
A	185	2	Single-family	Masonry walls. Pitched roof of wood and tiles. No insulation. Old double wood windows.
B	755	5	Multi-family	Masonry walls without insulation. Insulated flat roof. Old double wood windows.
C	5718	15	Multi-family	Concrete construction. Old double wood windows.

Reference [43] also describes the buildings and considered retrofit measures.

3. METHODOLOGY

Regression models that map the input variables related to the building retrofit measures into PES are built and trained with DT and ensemble methods RF, ET, and GB.

DT is a supervised machine learning method that learns if-else rules using available observations and later applies these prediction rules. They are simple to understand and convenient to interpret and even visualize the results. They require little effort related to data preparation and work fine with categorical features. However, DT is prone to overfitting, i.e. creating excessively complex trees that sometimes fail to generalize well.

Ensemble methods based on DT use multiple trees (estimators) together to improve the prediction performance. Averaging methods use multiple independent estimators and average their results. RF and ET are ensembles of DT where each tree is trained with a sample of observations randomly chosen with replacement. The predicted value is calculated as the average of the predictions of all trees. Boosting methods like GB build estimators incrementally and sequentially.

RF, ET, and GB are very powerful and precise prediction methods. Like DT, they require little data preparation and handle categorical variables well. They have a small number of hyperparameters, which makes them easy to tune. They are usually much faster to tune and train compared to ANN and SVM and less computationally intensive. In addition, they can naturally estimate the relative importance of each feature, have the warm start option to add estimators without recalculations, and use the out-of-bag data instead of the validation set.

Regression models are trained using the mean square error (MSE) as a loss function to be minimized. Equation (1) defines MSE:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(o_i - p_i\right)^2 \tag{1}$$

where n is the number of observations, o is the observed value of the output and p is the predicted value of the output.

In this paper, the regression results are presented in terms of the root MSE (RMSE), the root square of MSE, which is more convenient for interpretation because it is expressed in the same physical units as the output variable. Equation (2) defines RMSE:

$$\text{RMSE} = \sqrt{\text{MSE}} \tag{2}$$

The data for training, validating, and testing DT, RF, ET, and GB models are provided by calculating annual heating and cooling demands, $Q_{H,nd}$ and $Q_{C,nd}$, as well as PE, all expressed in kWh/m², according to [42]. Once the heating and cooling demands are calculated, PE can be obtained with Eq. (3):

$$\text{PE} = f_{DH}\frac{Q_{H,nd}}{\eta_{HS}} + f_E\frac{Q_{C,nd}}{\text{SEER}} \tag{3}$$

where η_{HS} is the efficiency of the heating system and SEER is the seasonal energy efficiency ratio.

PES of a combination of retrofit measures is calculated as the difference between PE for the baseline (do-nothing) scenario, PE_b and PE for the observed combination of measures, PE_r, as shown in Eq. (4):

$$\text{PES} = \text{PE}_b - \text{PE}_r \tag{4}$$

Each combination of retrofit measures is a vector of six input variables (R_e, R_i, R_l, R_u, U_f, g), i.e. one observation. Each observation corresponds to one value of PES.

The entire dataset for each building is obtained in an exhaustive manner, i.e. by calculating PES for all combinations of retrofit measures. The training and validation data is randomly selected from the dataset. The models are tested with the remaining data, not used previously for training or validation. The learning curves that illustrate the dependence between the performance on the size of the training set are obtained with the k-fold cross-validation.

The optimization of hyperparameters has not been performed. Instead, fully grown trees were allowed. For each ensemble method, 100 estimators were used and the learning rate for GB was 0.1.

4. RESULTS AND DISCUSSION

For each building, one combination of retrofit measures represents one observation with the vector of input variables (R_e, R_i, R_l, R_u, U_f, g) and the output PES. PES values are calculated for each observation and each building. Building A has 19.800 observations (the insulation of the interior walls is not considered), while buildings B and C have 217.800 observations each. Maximal PES values are 238.44, 278.41, and 137.86 kWh/m², for buildings A, B, and C, respectively.

For each building, a part of all observations is used to train and validate data-driven models, while the rest is applied for testing. The testing and validation part varies between 0.01 and 0.6, i.e. between 1% and 60% of the entire dataset.

This paper first analyzes the learning performance of the data-driven models, i.e. the size of the training set required to obtain accurate models that capture the dependencies and generalize well. Figures 2–5 illustrate the learning curves of DT, RF, ET, and GB, respectively, according to the part of the dataset used for training and validation.

It can be seen that in most cases, when training and validation portions of the entire dataset are between 0.25 and 0.4 (25% and 40%), the training and validation RMSE averages come very close to each other, leaving a small gap in between. For buildings B and C, the values of RMSE are small all the time, while for building A, small datasets have a relatively large validation RMSE. This difference is the consequence of the fact that the dataset for building A is considerably smaller than for the other two buildings.

The standard deviation values for the validation sets are large for smaller datasets, especially in the case of building A, but drop significantly and remain below 1 kWh/ m² when the portion of the training and validation set exceeds 0.2.

Thus, the learning performance heavily depends on the size of the training set up to some point and then remains approximately constant. In the examples examined here, the threshold is around 0.3 (30% of data used for training and validation) or lower, depending more on the dataset and less on the method applied. This is confirmed with the prediction accuracy obtained with the test sets, as shown in Table 2. The results indicate that, in this case, it is possible to obtain accurate data-driven models for PES prediction that generalize well by calculating around 30% of the observations chosen randomly. For the rest of the retrofit combinations, PES can be calculated with the data-driven models.
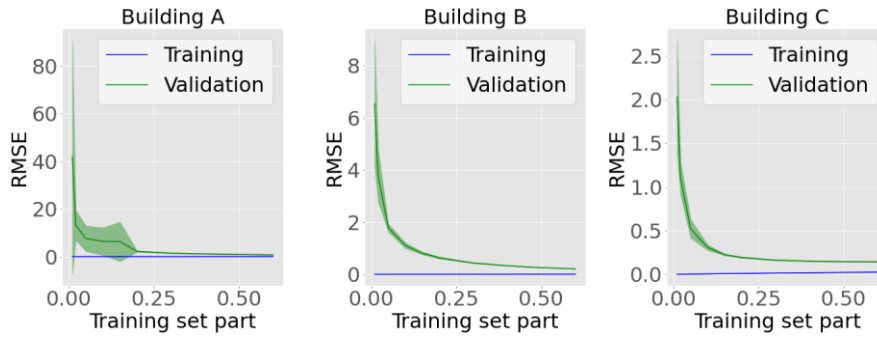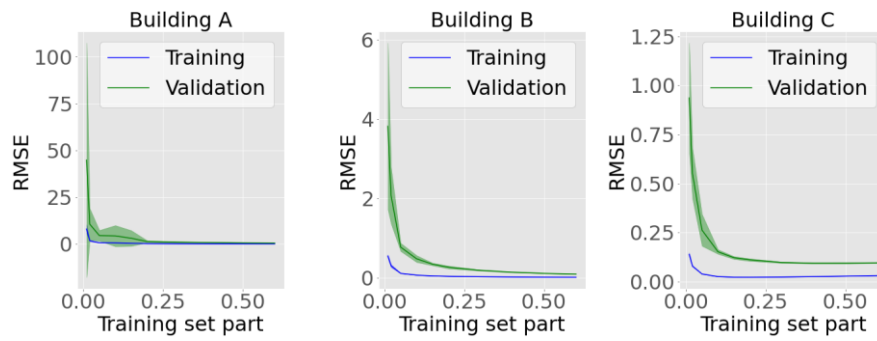
Fig. 2 Learning curve for decision trees
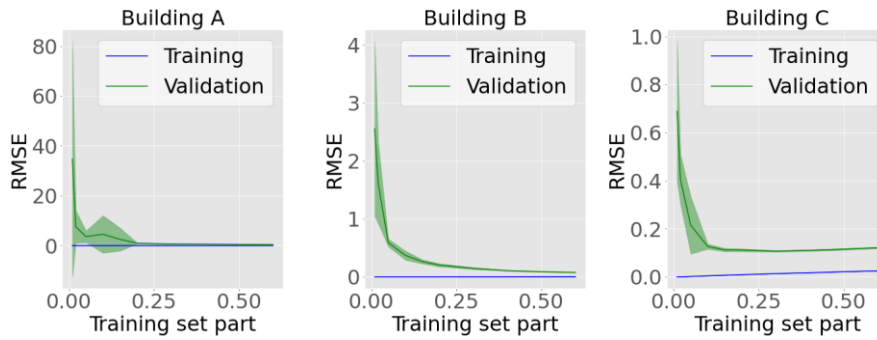

Fig. 3 Learning curve for random forest

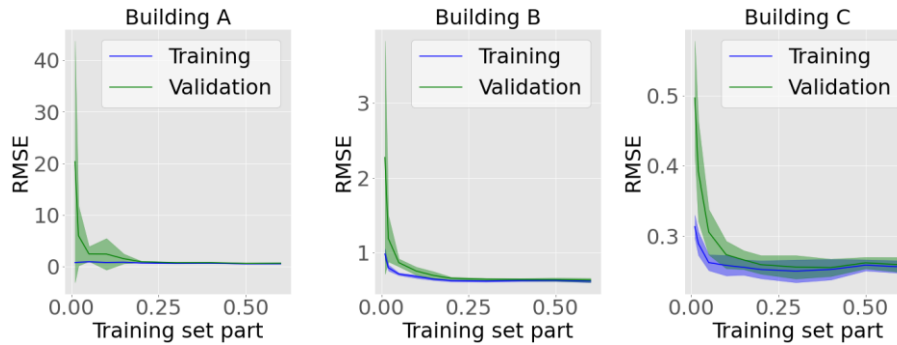**Fig. 4** Learning curve for extremely randomized trees

Fig. 5 Learning curve for gradient boosting

Table 2 presents RMSE values when 30% of all data is used for training and validation. RMSE for the training and validation sets is averaged over cross-validation folds. The results show that RMSE is far below 1 kWh/m² for the test sets, except when applying DT in the case for building A. They indicate very high accuracy of all methods. There is no single method that yields the best RMSE — RF, ET, and GB have the highest precision in some cases — and the differences are very small.

Table 2 Values of RMSE for the training set part of 0.3

Method	Building A			Building B			Building C		
	Train	Validation	Test	Train	Validation	Test	Train	Validation	Test
DT	0.0081	1.4634	1.4323	0.0017	0.4323	0.3944	0.0136	0.1616	0.1553
RF	0.1502	0.9093	0.7570	0.0285	0.1900	0.1694	0.0224	0.0969	0.0938
ET	0.0081	0.6690	0.6979	0.0017	0.1477	0.1303	0.0136	0.1071	0.1077
GB	0.6602	0.7940	0.6911	0.6202	0.6427	0.6551	0.2498	0.2555	0.2361

DT-based methods conveniently provide information on the relative importance, i.e. predictive power of each input variable. Figures 6–9 show the relative importance of the features for each of the used methods. For building A, very consistent results are obtained with all methods. The insulation of the exterior walls is the most important with 48–49%, followed by the floor insulation with about 25%, etc. For buildings B and C, two windows-related features have higher predictive power than the insulation. The relative contributions of the two vary, but their sum is approximately the same. This is valid for building A as well and might indicate the existence of the correlation among these two variables. The importance of other features are consistent across methods again.
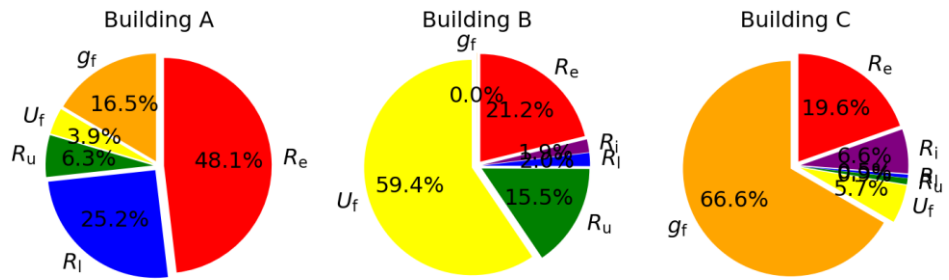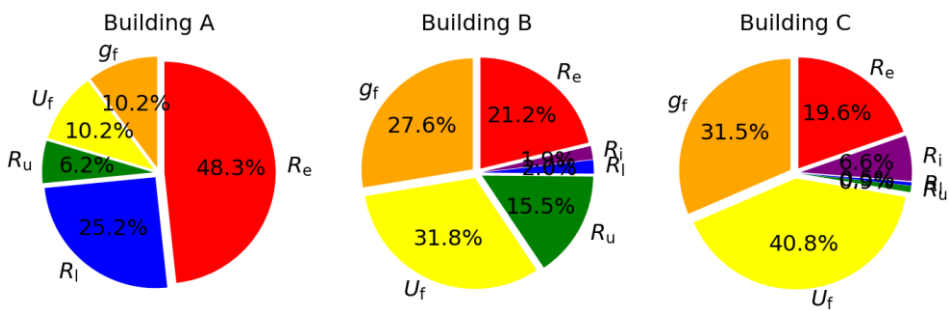
Fig. 6 Feature importance for decision trees

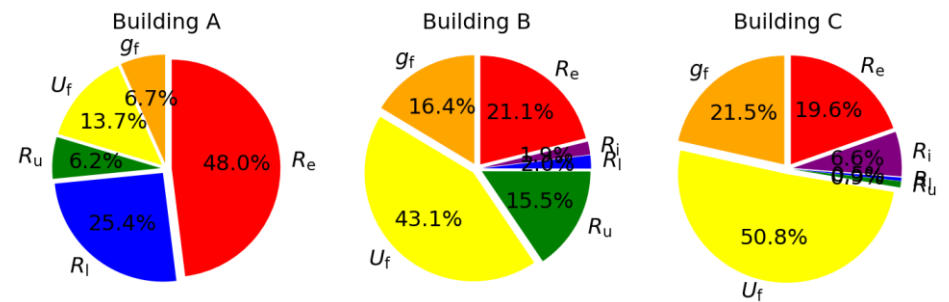**Fig. 7** Feature importance for random forest

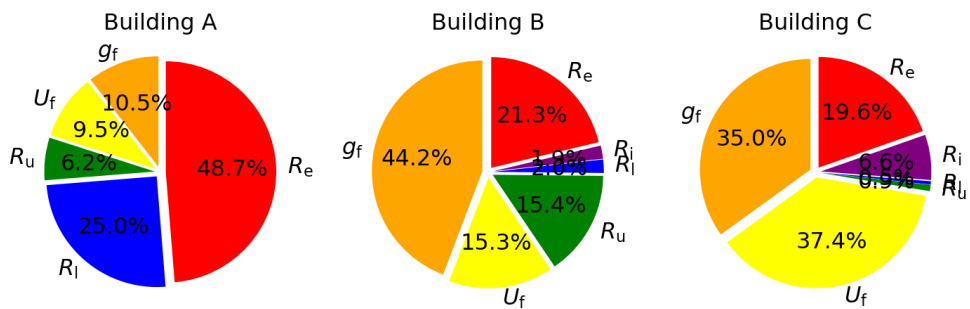**Fig. 8** Feature importance for extremely randomized trees

Fig. 9 Feature importance for gradient boosting

Thus, DT and DT-based ensembles provided highly accurate data-driven models that predict PES of the building retrofit measures, low variation across folds except for very small training sets, and consistent feature importance values.

5. CONCLUSION

This paper examined the applicability and accuracy of regression decision trees and related ensemble methods — random forest, gradient boosting, and extremely randomized trees — to create data-driven models that predict primary energy savings of retrofit measures for typical Serbian residential buildings connected to district heating systems. The considered measures are related to the improvements of the building thermal envelope.

In addition to being relatively fast to train, these methods yielded very accurate models, except in some cases with very small training sets. The k-fold cross-validation shows vary low variance across the folds, again except for some cases with small training sets. Ensemble methods gave slightly better results than decision trees. No particular ensemble method had the best performance in all cases. The applied methods have a convenient property to provide the relative importance of each feature. In most cases, all methods yield consistent results.

Future work might include: comparison of these methods against widely-used alternatives like artificial neural networks and support vector machines, both in terms of predictive accuracy and computational intensity; combine them with unsupervised learning methods; and test performance when considering the systems for energy supply and distribution in addition to the envelope-related measures.

REFERENCES

1. ****: Directive (EU) 2018/2002 of the European Parliament And of the Council of 11 December 2018 amending Directive 2012/27/EU on energy efficiency (Text with EEA relevance), Official Journal of the European Union, L 328, pp. 210–230, 2018.
2. ****: Directive (EU) 2018/844 of the European Parliament and of the Council of 30 May 2018 amending Directive 2010/31/EU on the energy performance of buildings and Directive 2012/27/EU on energy efficiency (Text with EEA relevance), Official Journal of the European Union, L 156, pp. 75–91, 2018.
3. Deb, C. et al.: A review on time series forecasting techniques for building energy consumption, Renewable and Sustainable Energy Reviews, Vol 74, pp. 902–924, 2017.
4. Foucquier, A. et al.: State of the art in building modelling and energy performances prediction: A review, Renewable and Sustainable Energy Reviews, Vol. 23, pp. 272–288, 2013.
5. Seyedzadeh, S. et al.: Tuning machine learning models for prediction of building energy loads, Sustainable Cities and Society Vol. 47, 101484, pp. 1–18, 2019.
6. Mosavi, A. et al.: State of the Art of Machine Learning Models in Energy Systems, a Systematic Review, Energies, Vol. 12, No. 7, 1301, pp. 1–42, 2019.
7. Rahman, A. and Smith, A.: Predicting heating demand and sizing a stratified thermal storage tank using deep learning algorithms, Applied Energy, Vol. 228, pp. 108–121, 2018.
8. Sretenović, A. A. et al.: Support Vector Machine for the Prediction of Heating Energy Use, Thermal Science, Vol. 22, Suppl. 4, pp. S1171–S1181, 2018.

9. Bui, D.-K. et al.: An artificial neural network (ANN) expert system enhanced with the electromagnetism-based firefly algorithm (EFA) for predicting the energy consumption in buildings, Energy, Vol. 190, 116370, pp. 1–12, 2020.

10. Rahman, A. et al.: Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks, Applied Energy, Vol. 212, pp. 372–385, 2018.

11. Zeng, A. et al.: Prediction of building electricity usage using Gaussian Process Regression, Journal of Building Engineering, Vol. 28, 101054, pp. 1–8, 2020.

12. Cui, B. et al.: A hybrid building thermal modeling approach for predicting temperatures in typical, detached, two-story houses, Applied Energy, Vol. 236, pp. 101–116, 2019.

13. Mawson, V. J. and Hughes, B. R.: Deep learning techniques for energy forecasting and condition monitoring in the manufacturing sector, Energy & Buildings, Vol. 217, 109966, pp. 1–10, 2020.

14. Dai, X. et al.: A review of studies applying machine learning models to predict occupancy and window-opening behaviours in smart buildings, Energy & Buildings, Vol. 223, 110159, pp. 1–15, 2020.

15. Huchuk, B. et al.: Comparison of machine learning models for occupancy prediction in residential buildings using connected thermostat data, Building and Environment, Vol. 160, 106177, pp. 1–12, 2019.

16. Jin, X. et al.: Foresee: A user-centric home energy management system for energy efficiency and demand response, Applied Energy, Vol. 205, pp. 1583–1595, 2017.

17. Ngarambe, J. et al.: The use of artificial intelligence (AI) methods in the prediction of thermal comfort in buildings: energy implications of AI-based thermal comfort controls, Energy & Buildings, Vol. 211, 109807, pp. 1–15, 2020.

18. Chaudhuri, T. et al.: A feedforward neural network based indoor-climate control framework for thermal comfort and energy saving in buildings, Applied Energy, Vol. 248, pp. 44–53, 2019.

19. Asadi, E. et al.: Multi-objective optimization for building retrofit: A model using genetic algorithm and artificial neural network and an application, Energy and Buildings, Vol. 81, pp. 444–456, 2014.

20. Ascione, F. et al.: CASA, cost-optimal analysis by multi-objective optimisation and artificial neural networks: A new framework for the robust assessment of cost-optimal energy retrofit, feasible for any building, Energy and Buildings, Vol. 146, pp. 200–219, 2017.

21. Chen, X., and Yang, H.: A multi-stage optimization of passively designed high-rise residential buildings in multiple building operation scenarios, Applied Energy, Vol. 206, pp. 541–557, 2017.

22. Gossard, D. et al.: Multi-objective optimization of a building envelope for thermal performance using genetic algorithms and artificial neural network, Energy and Buildings, Vol. 67, pp. 253–260, 2013.

23. Sharif, S. A. and Hammad, A.: Developing surrogate ANN for selecting near-optimal building energy renovation methods considering energy consumption, LCC and LCA, Journal of Building Engineering, Vol. 25, 100790, pp. 1–15, 2019.

24. Domahidi, A. et al.: Learning decision rules for energy efficient building control, Journal of Process Control, Vol. 24, pp. 763–772, 2014.

25. Drgoňa, J. et al.: Approximate model predictive building control via machine learning, Applied Energy, Vol. 218, pp. 199–216, 2018.

26. Pallonetto, F. et al.: Demand response algorithms for smart-grid ready residential buildings using machine learning models, Applied Energy, Vol. 239, pp. 1265–1282, 2019.

27. Smarra, F.: Data-driven model predictive control using random forests for building energy optimization and climate control, Applied Energy, Vol. 226, pp. 1252–1272, 2018.

28. Wang, J. et al.: Data-driven model predictive control for building climate control: Three case studies on different buildings, Building and Environment, Vol. 160, 106204, pp. 1–12, 2019.

29. Yang, S. et al.: Model predictive control with adaptive machine-learning-based model for building energy efficiency and comfort optimization, Applied Energy, Vol. 271, 115147, pp. 1–16, 2020.

30. Miller, C. et al.: A review of unsupervised statistical learning and visual analytics techniques applied to performance analysis of non-residential buildings, Renewable and Sustainable Energy Reviews, Vol. 81, pp. 1365–1377, 2018

31. Ding, Y. et al.: Effect of input variables on cooling load prediction accuracy of an office building, Applied Thermal Engineering, Vol. 128, pp. 225–234, 2018.

32. Naganathan, H. et al.: Building energy modeling (BEM) using clustering algorithms and semi-supervised machine learning approaches, Automation in Construction, Vol. 72, pp. 187–194, 2016.

33. Shang, C., and You, F.: A data-driven robust optimization approach to scenario-based stochastic model predictive control, Journal of Process Control, Vol. 75, pp. 24–39, 2019.

34. Westermann, P. et al.: Unsupervised learning of energy signatures to identify the heating system and building type using smart meter data, Applied Energy, Vol. 264, 114715, pp. 1–14, 2020.

35. Liu, T. et al.: A novel deep reinforcement learning based methodology for short-term HVAC system energy consumption prediction, International Journal of Refrigeration, Vol. 107, pp. 39–51, 2019.

36. Wang, Z. and Hong, T.: Reinforcement learning for building controls: The opportunities and challenges, Applied Energy, Vol. 269, 115036, pp. 1–18, 2020.
37. Ahmad, M. W. et al.: Trees vs Neurons: Comparison between random forest and ANN for high-resolution prediction of building energy consumption, Energy and Buildings, Vol. 147, pp. 77–89, 2017.
38. Touzani, S. et al.: Gradient boosting machine for modeling the energy consumption of commercial buildings, Energy and Buildings, Vol. 158, pp. 1533–1543, 2018.
39. Seyedzadeh, S. et al.: Data driven model improved by multi-objective optimisation for prediction of building energy loads, Automation in Construction, Vol. 116, 103188, pp. 1–12, 2020.
40. Wang, Z. et al.: Building thermal load prediction through shallow machine learning and deep learning, Applied Energy, Vol. 263, 114683, pp. 1–14, 2020.
41. Ghahramani, A. et al.: Learning occupants' workplace interactions from wearable and stationary ambient sensing systems, Applied Energy, Vol. 230, pp. 42–51, 2018.
42. ****: Rulebook on energy efficiency of buildings, Official Gazette of the Republic of Serbia, Vol. 61, 2011. (In Serbian)
43. Stojiljković, M. M. et al.: Cost-optimal energy retrofit for Serbian residential buildings connected to district heating systems, Thermal Science, Vol. 23, Suppl. 5, pp. 1707–1717, 2019.

PREDVIĐANJE UŠTEDE PRIMARNE ENERGIJE MERA ZA RENOVIRANJE ZGRADA ANSAMBL METODAMA NA BAZI STABALA ODLUČIVANJA

Primarna energija je veličina kojom se često izražava ukupna količina potrošene ili sačuvane energije. Posebno je važna za sertifikaciju zgrada i proveru usklađenosti sa standardima. Da bi se precizno procenila potrošnja ili ušteda primarne energije, potrebni su adekvatni modeli koji mogu biti zasnovani na znanju o fenomenima, na prikupljenim podacima ili oba. U ovom radu se analiziraju performanse učenja i preciznost modela zasnovanih na podacima i kreiranim na bazi stabala odlučivanja koji predviđaju uštedu primarne energije mera renoviranja zgrada. Rad koristi metode nadgledanog mašinskog učenja — regresiona stabla odlučivanja i ansambl metode zasnovane na njima — za treniranje, validaciju i testiranje ovih modela. Ansambl metode zasnovane na stablima odlučivanja su moćne, precizne, a uz to pogodne za pripremu i brzo treniranje. Pored toga, one računaju relativni značaj svake ulazne promenljive ponaosob. Istraživanje je rezultiralo veoma preciznim modelima zasnovanim na podacima i konzistentnim vrednostima značaja promenljivih.

Ključne reči: *zgrade, modeli zasnovani na podacima, ansambl metode, ušteda primarne energije, nadgledano mašinsko učenje*