MACHINE LEARNING BASED OPTIMIZATION APPROACH FOR BUILDING ENERGY PERFORMANCE

Aslihan Senel Solmaz
Department of Architecture, Dokuz Eylul University, Izmir, Turkey

ABSTRACT

This paper proposes a machine learning based optimization approach to reduce building total electricity consumption for heating and cooling energy. The proposed approach includes three main steps. Firstly, sampling of base-case building design variables and corresponding simulation results were obtained in EnergyPlus to generate features/labels. As the next step, a Machine Learning (ML) Model based on Support Vector Regression algorithm to predict energy consumption was created. Thirdly, Bayesian optimization was performed on the ML model to determine the optimum values for design variables. A hypothetical office building was used as a case study in order to demonstrate the feasibility of the proposed approach. The results showed that the proposed machine learning based approach is very accurate in estimating the total energy consumption, and easily integrated to any black-box optimization techniques with much less computational cost.

INTRODUCTION

Improving energy efficiency in buildings is a complex process since buildings consist of numerous interrelated sub-systems (i.e. structural system and building materials, HVAC systems, building services) that influence the overall building performance (Hong et al., 2018). Applications of optimization methods with building simulation tools for handling these complex systems have emerged as a promising method, and therefore using optimization methods during this process play a key role for achieving high performance buildings (Evins, 2013).

In recent years, there are numerous studies that focus on optimizing building design and operations in order to ensure energy efficiency and thermal comfort in a cost effective manner. In most of these studies, the optimization process is based on connecting building performance simulation tools (e.g. EnergyPlus and TRNSYS) and algorithmic optimization engines (involving general optimization packages, custom programmed algorithms and special optimization tools) (e.g. GenOpt, BEopt and Matlab). For example, an optimization tool named MultiOpt based on a combination of genetic algorithm (GA) (NSGA-II) with TRNSYS energy simulation program was developed in order to determine the most effective building envelope and control system solutions while optimizing building energy consumption, cost, thermal comfort and life-cycle environmental impact (Chantrelle et al., 2011). A multi-objective optimization model involving direct coupling of the GenOpt optimization program, TRNSYS simulation program, and Matlab was introduced for identifying the most effective set of energy retrofit alternatives that maximize building energy savings and thermal comfort while minimizing energy retrofit costs simultaneously (Asadi et al., 2012). A coupling framework involving the EnergyPlus building energy simulation and the multi-dimensional numerical optimization employed by GenOpt optimization engine was used in order to find appropriate building envelope retrofitting applications that minimize building life cycle costs (Karaguzel et al., 2014). An optimization methodology based on coupling EnergyPlus with MATLAB was presented to optimize building envelope design parameters that ensure minimum energy demand while maximizing thermal comfort (Ascione et al., 2015). A simulation-based multi-objective optimization approach that couples a multi-objective particle swarm optimization algorithm embedded in jEPlus optimization engine with EnergyPlus simulation program was introduced in order to identify the most appropriate building solutions for minimizing building energy consumption (Delgarm et al., 2016). A decision-support framework including integration between a multi- objective optimization and sensitivity analysis performed by coupling GenOpt with

EnergyPlus was proposed in order to find the optimum set of energy retrofit solutions for an existing school building (Senel Solmaz et al., 2018). Similarly, A multi-objective optimization approach based on integration between EnergyPlus and GenOpt optimization program was used to optimize four objective criteria with: building heating and cooling energy savings, occupant thermal comfort and the financial metric, net present value (NPV) simultaneously. The "weighted-sum" approach was utilized to formulate optimization problem. As for the design variables, building envelope related parameters were chosen to find an optimum set of building energy efficiency solutions (Senel Solmaz, 2018).

Although a simulation-based single or multi-objective optimization approaches have been studied extensively, one of main drawbacks of this process is the necessity of vast number of simulation evaluations, being computationally expensive. Simulation of detailed building energy models generally takes time due to its complexity, and during simulation-based optimization process, it is mostly necessary to make a large number of simulation evaluations. Hence these approaches may therefore become infeasible due to such high computational cost of building models. In order to overcome this, simplified models instead of detailed building models, namely “surrogate model”, have become popular. A surrogate model (meta-model) is a prediction model of the original simulation model, and it reliably mimics its behaviour. There are several methods have been used in order to construct the surrogate model based on the data set, e.g. artificial neural network (ANN), support vector machine (SVM) and Kriging (Nguyen et. al., 2014). Though a surrogate based optimization that indicates the use of developed prediction model instead of original simulation model in optimization process for the evaluation of individuals is outstanding in terms of computational efficiency, developing an accurate surrogate model is a big challenge of this process, and that still needs to be overcome (Li et al., 2017). When looking at the research effort in this area, for example, an optimization methodology based on integration of artificial neural network (ANN) with GA (NSGA-II) was proposed in order to minimise energy consumption and construction cost of residential buildings. First, the ANN based prediction model was developed, and then using that model in optimization, several envelope and HVAC related parameters were optimized to ensure building energy efficiency (Magnier and Haghighat, 2010). A prediction model based on ANN for thermal comfort index and annual energy consumption forecasting was developed, and then combined with multi-objective GA in order to identfy the optimum values of building envelope materials’ thermal properties that ensure the optimum energy consumption and thermal comfort concurrently (Gossard et al., 2013). Similarly, a multi-objective optimization model involving combination of ANN with GA was proposed for identifying the most effective building energy retrofit strategies that minimize building energy consumption, retrofit cost while maximize thermal comfort (Asadi et al., 2014). An optimization framework named as CASA based on integration of ANNs and multi-objective GA for assessment of cost-optimal energy retrofit strategies. ANNs was used to performance prediction of buildings, and by using that model GA was performed for optimal building energy retrofit solutions (Ascione et al., 2017). A machine learning based optimization approach involving sequential application of support vector machine (SVM) and GA was developed to optimize a set of occupant related design variables for minimizing cooling energy consumption and maximizing thermal comfort of building (Amasyali and Gohari, 2018). A robust surrogate ANN model for predicting building energy consumption was developed by using data from simulation-based multi-objective optimization model that was employed for determining an optimal set of renovation solutions in order to optimize building energy consumption, life-cycle cost (LCC), and life-cycle assessment (LCA) simultaneously (Sharif and Hammad, 2019). A multi objective optimization approach based on an ANN and best algorithm values derived from performance evaluation of four different multi-objective algorithms was proposed to design optimization of complex buildings for improving energy efficiency and indoor thermal comfort (Si et al., 2019). A surrogate model developed by using ANN was utilized to reduce the computing time. The results indicated that within the four multi-objective optimization algorithm, the NSGA-II performed the best.

The research efforts on building performance optimization summarized above are all significant and also indicate how important optimization of building design and operations per several performance criteria is. However, there is still room for investigating the integration of optimization algorithms with surrogate model approaches for getting model with good accuracy and less computational time. To this end, this study presents a machine learning (ML) based optimization approach using a machine learning method, Support Vector Regression (SVR), and Bayesian optimization in an integrative way to optimize building design parameters and minimize building total electricity for energy consumption.

RESEARCH METHODOLOGY

The general framework of the proposed approach is presented in Figure 1. A machine learning optimization approach consists of the application of three sequential steps: 1) Creating a building energy model (base-case model) in EnergyPlus, followed by input parameter sampling and energy simulations, 2) Feeding simulated input-output relations to ML algorithm as features/labels and model creation, 3) Bayesian black-box optimization to minimize the total electricity for building energy consumption.

In order to create a database for ML based prediction model training, design variables were defined with min-max value ranges, input sampling was done using uniform distrubution, and a number of EnergyPlus input files (.idf) from sampled values were generated. After that, the evalution/simulation of all these files were executed in EnergyPlus. At this point, in order to automate the whole process of both the generation of input (.idf) files and reading EnergyPlus output files (.eso) from associated simulation runs, custom Python scripts were used.

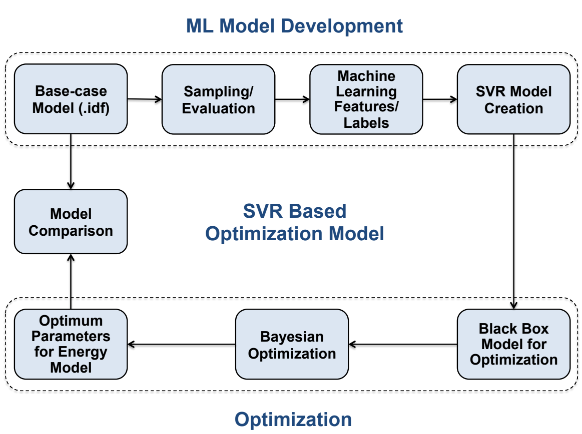


Figure 1. A machine learning based optimization framework

After obtaining the sampled inputs and associated output values (i.e. machine learning features and labels), the prediction model was generated. Lastly, once trained and validated, the prediction model was used as a black-box model as evaluation function for building energy consumption estimation in Bayesian optimization.

CASE STUDY

Building Energy Model (Base-case model)

In this study, a medium-sized office building in EnergyPlus (Figure 2), which is one of the 16 commercial reference building models developed by the U.S. Department of Energy (DOE), was used as a base-case model. Therefore, it is compatible with EnergyPlus (Deru et al., 2006). The base-case model is a three-story building oriented in south-north direction. There are a total of 3 plenums, 12 perimeter zones, and 3 core zones. Total floor area is 4982 m^2 (53625.8ft^2), while the total building height is 11.88 m (38.98ft). The glazing are uniformly distributed in the horizontal direction with a glazing ratio of 33 %.

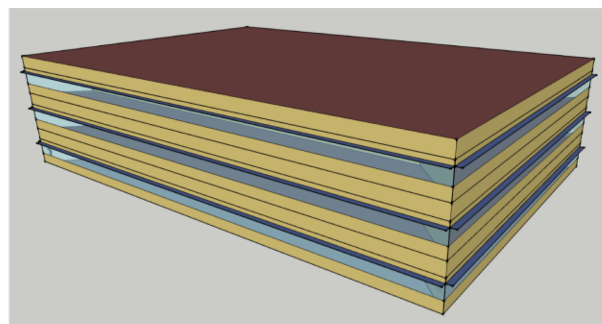


Figure 2. The base-case building's 3D energy model

Although the original geometry and HVAC system (the multi-zone variable air volume (MZ-VAV)) of the DOE reference building were kept, the envelope materials and their thermophysical properties were modified based on the construction standards in Turkey (Turkish Standard Institution, 2008). The building location is in Izmir-Turkey which has hot-humid climate features, and simulations were done using ASHRAE IWEC (International Weather for Energy Calculations) weather file. The cooling set point is 24°C and the heating set point is 22°C while the setback temperatures are 26°C and 16°C respectively.

The total energy consumption in this paper is the electricity used only for heating and cooling, not including facility, interior equipment and lights.

Machine Learning Model Creation and Validation

Input sampling and evaluation

In order to train the ML model, a dataset including input and associated output values is needed. The input is several envelope design variables highly influencial on building energy performance The min-max ranges were defined for each input as part of sampling process (Table 1). As seen on Table 1, some variables were defined with only one parameter while others had more.

For instance, wall, roof and floor have only thermal insulation material thickness as an input and the rest of the thermophysical properties of insulation material such as thermal conductivity, specific heat and density values were fixed. On the other hand, windows had U-value and Solar Heat Gain Coefficient (SHGC) at the same time.

Table 1 Selected design variables and their min-max value ranges for input sampling process.1 meter is equal to 3.28ft.

NO	VARIABLES	UNIT	MIN	MAX
1	Thermal Insulation thickness of exterior wall	m	0.001	0.12
2	Thermal Insulation thickness of roof	m	0.001	0.20
3	Thermal Insulation thickness of floor	m	0.001	0.10
4	Window U value	W/m^{2}K	1	6.0
5	Window SHGC value	-	0.5	1.0
8	Shading material depth (*South wind.)*	m	0.001	1.0
7	Shading material depth (*West wind*.)	m	0.001	1.0
8	Shading material angle (*South wind.*)	degree (°)	45	90
9	Shading material angle (*West wind*.)	degree (°)	45	90

By using random uniform distribution data sampling, a sample of 100 values within defined min-max ranges were simultaneously generated for each input. Then, model reproducer code processed base-case EnergyPlus input file (.idf), found the relevant variables and assigned values based on the generated input matrix, and saved as a new (.idf) file for each sample. A total of 100 (.idf) files were generated, and further simulated in EnergyPlus, and output values from output files (.eso) including building total energy consumption values were gathered automatically. The total simulation time of 100 cases took around 1.3h (46 sec for each simulation) using a computer equipped with Intel i7 Quad-Core CPU 2.2 GHz, 8 GB RAM. The distribution of total energy consumption is presented in Figure 3 where the lowest, mean and highest values in total energy consumption across all 100 cases are 163,433 kWh, 222,527 kWh, and 292,475 kWh respectively with 28,776 kWh standart deviation.

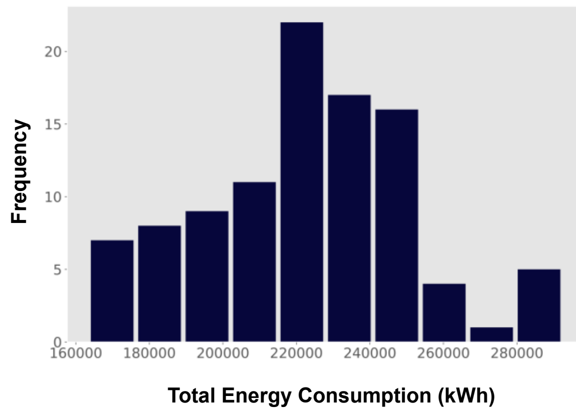
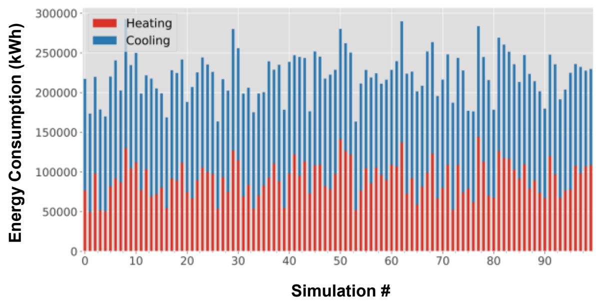


Figure 3 Distribution of energy consumption data obtained in EnergyPlus.

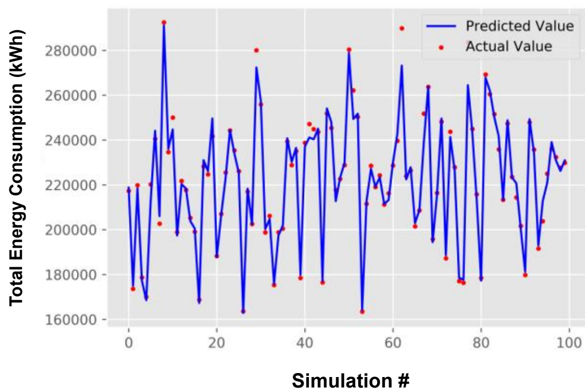
SVR-based ML model development and performance evaluation

The obtained input/output data was used as features/labels for SVR model generation. SVR algorithm was chosen because it works great for such non-linear models using kernel transformation. Scikit-Learn ML package was used for SVR model creation, training, validation and saving (Pedregosa et al., 2011). A sample of 100 cases was used for SVR training and Root Mean Square Error (RMSE) was used as loss function. During model generation, 5-fold cross validation was used to ensure good generatilization of the model, and hyper-parameter tuning was performed for better RMSE.

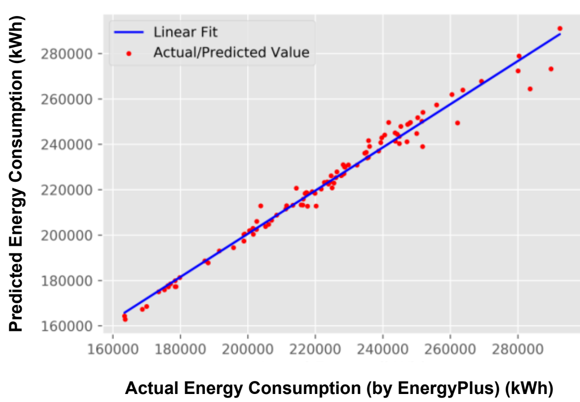
The actual heating and cooling energy values were plotted in Figure 4a, while the actual and predicted total energy consumption values for each simulation plotted in Figure 4b shows almost a perfect match, with RMSE of 0.1412 for normalized output values. Figure 4c shows great correlation between the actual and predicted values and linear-fit with Coefficient of Determination (R^2) value of 0.9813. Both plots show how well the SVR model performs compared to EnergyPlus simulation software. Hence, the SVR model created can instantly predict building energy consumption with great accuracy compared to 46sec simulation time in EnergyPlus.



(a)



(b)



(c)

Figure 4. ML model validation: a) Actual heating and cooling energy consumption values from 100 input files b) The actual (EnergyPlus) vs. predicted (SVR) total energy consumption values with RMSE of 0.1412. c) Almost perfect linear relationship between the actual and predicted energy consumption values and linear-fit with R^2 of 0.9813.

Bayesian Optimization

The SVR model created was used as the black-box for the optimization step and Bayesian Optimization was chosen as the algorithm in order to determine the optimal values for defined design variables and minimize the building energy consumption. Bayesian black-box Optimization is an appropriate selection due to the lack of any analytic equation to describe the energy model, which is complex and noisy in nature. It treats the energy model as a true black-box and tries to minimize *f(x)* by adjusting inputs in trial-error approach. The Bayesian optimization algorithm of the Scikit-Optimize package (GitHub, 2019) was used for single-objective optimization.

The optimization algorithm takes a total of 9 decision variables as inputs: thermal insulation thicknesses of exterior wall (x_{wall}), roof (x_{roof}) and floor (x_{floor}), window U ($x_{win\text{-}U}$) and SHGC ($x_{win\text{-}SHGC}$) values, shading material depth values for south ($x_{shadD\text{-}s}$) and west ($x_{shadD\text{-}w}$) windows, and shading material angle values for south ($x_{shadA\text{-}s}$) and west ($x_{shadA\text{-}w}$) windows, and a single output, total energy consumption *f(x)*. Hence the optimization problem can be formulated as follows:

$$
\begin{aligned}
& \min f(x) \\
\text{subject to} \quad & 0.01 \le x_{wall} \le 0.12 \\
& 0.01 \le x_{roof} \le 0.2 \\
& 0.01 \le x_{floor} \le 0.1 \\
& 1 \le x_{win-U} \le 6 \\
& 0.5 \le x_{win-SHGC} \le 1 \\
& 0 \le x_{shadD-s} \le 1 \\
& 0 \le x_{shadD-w} \le 1 \\
& 45 \le x_{shadA-s} \le 90 \\
& 45 \le x_{shadA-w} \le 90
\end{aligned}
\tag{1}
$$

The Bayesian optimization was set with random input variables to start with and limited to 50 calls. Figure 5 shows the objective function (total energy consumption) at each optimization step. It makes ~35% improvement within the first 10 steps and quickly converges around 30th step with close to 90% improvement. The optimum value was found to be 129,670 kWh, with the best-reported values for input parameters (Table 2). The total simulation time for 50 iterations was less than 1min.

Plugging the best parameters from Table 2 into EnergyPlus and re-running the simulation again, we see a difference of 2.81% with the actual simulation result.

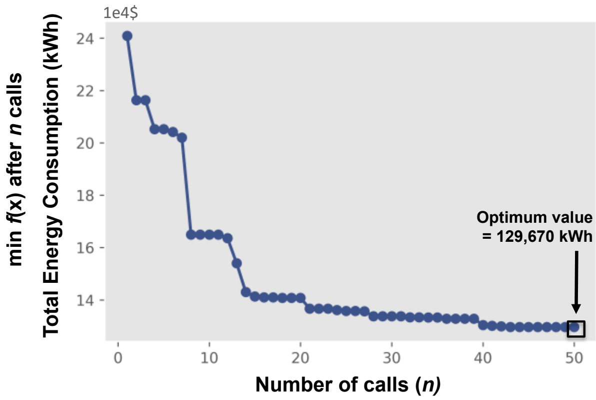


Figure 5 Optimization iterations and reduction in Total Energy Consumption.

The total energy consumption includes both heating and cooling energy consumption, so the result is expected to be a balanced decrease for both consumptions. According to optimization results, for example, the optimum thickness of thermal insulation material for wall, and floor were found 0.08 m (3.15in), and 0.01 m (0.39in) respectively. This value is 0.16 m (6.30in) for roof, which is much higher than wall and floor. This result is logical considering the positive impact of roof thermal insulation on both heating and cooling energy savings, so total energy savings on hot-humid climates, and also the size of roof area. In another words, application of insulation materials on roof surface is more effective than for wall and floor surface thermal insulation (Senel Solmaz et al., 2016). The thermal insulation material with almost lowest thickness (0.01 m) was assigned to the ground floor. This may be due to insulation material negatively affecting the buildings cooling energy consumption even if it has a positive impact on heating energy savings for hot-humid climates as cooling energy consumption is more dominated than heating energy consumption in base-case building (Senel Solmaz, 2018). Similarly, the optimum thermal insulation material's thickness for all exterior walls (0.08 m) is not the maximum amount.

Table 2 Optimum values for selected design variables. 1 meter is equal to 39.37in)

NO	VARIABLES	UNIT	OPTIMUM VALUES
1	Thermal Insulation thickness of exterior wall	m	0.08
2	Thermal Insulation thickness of roof	m	0.16
3	Thermal Insulation thickness of floor	m	0.01
4	Window U value	W/m^{2}K	1.0
5	Window SHGC value	-	0.50
8	Shading material depth (*South wind.)*	m	0.54
7	Shading material depth (*West wind*.)	m	1.0
8	Shading material angle (*South wind*.)	degree (°)	63.39
9	Shading material angle (*West wind*.)	degree (°)	90

As for the optimum values of window parameters, the optimization algorithm chose the window material having the lowest U (1.0 W/m^{2}K) value and mean SHGC value (0.5). When considering the positive impact of windows with lower U values on decreasing both heating and cooling energy consumption, it is logical to select the lowest U value for windows. For the SHGC, although it has positive effect on heating energy savings, it affects cooling energy savings negatively. Therefore, considering the total energy consumption, the optimization algorithm may try to make a balance on total energy consumption by choosing the almost medium value for window SHGC.

Lastly for the shading materials, the optimization algorithm found the optimum values for south window shading depth as 0.54 m, and 63.39 degrees for angle. As for the west window shading, the highest value was chosen for the shading depth (1.0 m), and shading angle was assigned as 90 degrees.

CONCLUSION

This paper proposed a machine learning based optimization approach in order to reduce building total energy consumption. The machine-learning model was built from simulation generated data and later used in optimization of building total energy consumption on nine building envelope related design variables (e.g. thermal insulation thicknesses of exterior wall, roof and floor, window U and SHGC values, shading material depth and angle values for south and west facades). The main findings made in the scope of this study are:

- The proposed machine learning based approach is very accurate in estimating the total energy consumption even with low number of features and easily integrated to any black-box optimization techniques.
- For the best iteration that ensures the minimum total energy consumption, the optimum values determined for selected variables on building envelope is: the optimum thickness of thermal insulation materials for wall, roof and floor were 0.08 m, 0.12 m and 0.01 m; the optimum U and SHGC values are 1.0 W/m^2K, and 0.5; the optimum values for shading depth and angle for south windows were 0.54 m and 63.39 degrees, and for west windows were 1.0 m and 90 degrees.
- The optimization algorithm could find the optimum result (with a deviation of 2.81% from EnergyPlus simulated value) in less then 1 min with only 50 iterations, as opposed to hours of optimization time using traditional approaches. Therefore, using ML surrogate model during optimization process for calculation of energy consumption provided a significant opportunity in term of computational cost.

As for the limitations of this research, the handled problem in this study was a single optimization problem and focused on only building total energy consumption period. Yet the modular nature of this framework can easily be extended to consider multiple objectives (e.g. thermal comfort, CO_2 emission, life-cycle cost), and optimize them simultaneously. This approach can also be extended to evaluate different type and/or size buildings in different weather conditions, and different energy systems (e.g. renewable energy technologies, HVAC systems). Another consideration is the number of design variables selected. The surrogate model can be built with more design variables (orders of magnitude more), but requires more simulation data beforehand in order to build a better model. Lastly, other ML algorithms can be explored with similar accuracy and robust prediction model.

REFERENCES

Asadi, E., Silva, M.G., Antunes, C.H., Dias, L. 2012. A Multi-Objective optimization model for building retrofit strategies using TRNSYS simulations, Genopt and MATLAB. Building and Environment, 56, 370-378.

Asadi, E., Silva, M. G. da, Antunes, C. H., Dias, L., Glicksman, L. 2014. Multi-objective optimization for building retrofit: A model using genetic algorithm and artificial neural network and an application. Energy and Buildings, 81, 444-456.

Ascione, F., Bianco, N., De Masi, R., Mauro, G., Vanoli, G. 2015. Design of the building envelope: A novel multi-objective approach for the optimization of energy performance and thermal comfort. Sustainability, 7(8), 10809- 10836.

Ascione, F., Bianco, N., De Stasio, C., Mauro, G.M., Vanoli, G.P. 2017. 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, 146, 200-219.

Amasyali, K., El-Gohary, N. 2018. Machine Learning-based occupant energy use behavior optimization, Construction Research Congress 2018, New Orleans, Louisiana, USA.

Chantrelle, F.P., Lahmidi, H., Keilholz, W., Mankibi, M.E., Michel, P. 2011. Development of a multicriteria tool for optimizing the renovation of buildings. Applied Energy, 88, 1386-1394.

Delgarm, N., Sajadi, B., Kowsary, F., Delgarm, S. 2016. Multi-objective optimization of the building energy performance: A simulation-based approach by means of particle swarm optimization (PSO). Applied Energy, 170, 293-303.

Deru, M., Griffith, B., Torcellini, P. 2006. Establishing Benchmarks for DOE Commercial Building R&D and Program Evaluation. California, National Renewable Energy Laboratory (NREL).

Evins, R. 2013. A review of computational optimisation methods applied to sustainable building design. Renewable and Sustainable Energy Reviews, 22 (2013), 230–245.

Gossard, D., Lartigue, B., Thellier, F. 2013. Multi-objective optimization of a building envelope for thermal performance using genetic algorithms and artificial neural network. Energy and Buildings, 67, 253-260.

GitHub. 2019. Skopt module. available at: https://scikit-optimize.github.io/ (accessed 09.28.2019).

Hong, T., Langevin, J., Sun, K. 2018. Building simulation: Ten challenges. Building Simulation, 11, 871–898.

Karaguzel, O., Zhang, R., Lam, K. 2014. Coupling of whole-building energy simulation and multi-dimensional numerical optimization for minimizing the life cycle costs of office buildings. Building Simulation, 7, 111-121.

Li, K., Pan, L., Xue, W., Jiang, H., Mao, H. 2017. Multi-objective optimization for energy performance improvement of residential buildings: A comparative study. Energies, 10(2), 245.

Magnier, L., Haghighat, F. 2010. Multiobjective optimization of building design using TRNSYS simulations, genetic algorithm, and Artificial Neural Network. Building and Environment, 45(3), 739-746.

Nguyen, AT., Reiter, S., Rigo, P. 2014. A review on simulation-based optimization methods applied to building performance analysis. Applied Energy, 113, 1043-58.

Pedregosa, F., Ga, I., Varoquaux, A., et. al. 2011. Scikit-learn: Machine Learning in Python. J. Mach. Learn. Res., 12, 2825-2830.

Senel Solmaz, A., Halicioglu, F.H., Gunhan, S. 2016. A simulation based optimization approach for determining the optimum energy saving solutions for buildings, Third IBPSA-England Building Simulation & Optimization Conference 2016, Newcastle upon Tyne, UK.

Senel Solmaz, A., Halicioglu, F.H., Gunhan, S. 2018. An approach for making optimal decisions in building energy efficiency retrofit projects. Indoor and Built Environment, 27 (3), 348-368.

Senel Solmaz, A. 2018. Optimisation of energy performance and thermal comfort of an office building. GRADEVINAR, 70 (7), 581-592.

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

Si, B., Wang, J., Yao, X., Shi, X., Jin, X, Zhou, X. 2019. Multi-objective optimization design of a complex building based on an artificial neural network and performance evaluation of algorithms. Advanced Engineering Informatics, 40 (2019), 93-109.

Turkish Standard Institution. 2008. TS 825—The Thermal Insulation Requirements for Buildings (revision) in Turkish Context. Ankara.

US Department of Energy (DOE) & National Renewable Energy Laboratory (NREL). 2019. EnergyPlus, available at: https://energyplus.net/ (accessed 10.12.2019).