

Comparison of machine learning techniques for predicting energy loads in buildings

Uma comparação de técnicas de aprendizado de máquina para a previsão de cargas energéticas em edifícios

Grasiele Regina Duarte
Leonardo Goliatt da Fonseca
Priscila Vanessa Zabala Capriles Goliatt
Afonso Celso de Castro Lemonge

Abstract

Machine learning methods can be used to help design energy-efficient buildings reducing energy loads while maintaining the desired internal temperature. They work by estimating a response from a set of inputs such as building geometry, material properties, project costs, local weather conditions, as well as environmental impacts. These methods require a training phase which considers a dataset drawn from selected variables in the problem domain. This paper evaluates the performance of four machine learning methods to predict cooling and heating loads of residential buildings. The dataset consists of 768 samples with eight input variables and two output variables derived from building designs. The methods were selected based on exhaustive research with cross validation. Four statistical measures and one synthesis index were used for the performance assessment and comparison. The proposed framework resulted in accurate prediction models with optimized parameters that can potentially avoid modeling and testing various designs, helping to economize in the initial phase of the project.

Keywords: Energy efficiency. Heating and cooling loads. Machine learning.

Resumo

Métodos de aprendizagem de máquina podem ser usados para auxiliar o projeto de edifícios energeticamente eficientes, reduzindo cargas de energia enquanto se mantém a temperatura interna desejada. Eles operam estimando uma resposta a partir de um conjunto de entradas tais como a geometria do edifício, propriedades do material, custos do projeto, condições do tempo no local e impacto ambiental. Esses métodos requerem uma fase de treinamento que considera uma base de dados construída a partir de variáveis selecionadas no domínio do problema. Este trabalho avalia o desempenho de quatro métodos de aprendizado de máquina na predição de cargas de resfriamento e aquecimento de edifícios residenciais. A base de dados do treinamento consiste de oito variáveis de entrada e duas variáveis de saída, todas derivadas de projetos de edifícios. Os métodos foram selecionados de acordo com uma pesquisa exaustiva e ajustados por uma estratégia com validação cruzada. Para a avaliação foram usadas quatro medidas estatísticas de desempenho e um índice de sintetização e resultados. Essa estratégia resultou em algoritmos com parâmetros otimizados e permitiu obter resultados competitivos com os apresentados na literatura.

Palavras-chave: Eficiência energética. Cargas de aquecimento e resfriamento. Aprendizado de máquina.

Grasiele Regina Duarte
Universidade Federal de Juiz de Fora
Juiz de Fora - MG - Brasil

Leonardo Goliatt da Fonseca
Universidade Federal de Juiz de Fora
Juiz de Fora - MG - Brasil

Priscila Vanessa Zabala Capriles Goliatt
Universidade Federal de Juiz de Fora
Juiz de Fora - MG - Brasil

Afonso Celso de Castro Lemonge
Universidade Federal de Juiz de Fora
Juiz de Fora - MG - Brasil

Recebido em 28/11/16
Aceito em 23/03/17

Introduction

The basic principle of building energy efficiency is to use less energy for operations including heating, cooling, lighting and other appliances, without affecting the health and comfort of its occupants. Improving the energy efficiency of functional buildings brings many environmental and economic benefits such as reduced greenhouse gas emissions and operational cost savings. In many developed and developing countries, energy efficiency has become the main way to meet a rising energy demand (FRIESS; RAKHSHAN, 2017).

In order to reduce the energy demand growth and decrease the amount of energy used associated with buildings, it is critical to understand how energy is distributed throughout a building, and how building parameters contribute to energy consumption (MUSTAFARAJ *et al.*, 2014). Simulation tools can provide a reliable framework for assessing energy distribution in buildings and can help designers to understand the importance of building and weather parameters. However, when considering the decision-making process during the project cycle, carrying out a set of simulations can lead to complex scenarios and may be time-consuming. In order to avoid these drawbacks, machine learning methods can be used for energy demand prediction. These methods require the shortest amount of time in order to model the entire building and they are becoming commonly used for preliminary estimations (MELO *et al.*, 2016).

Although building orientation and layout have been shown to be highly important in reducing building energy consumption in cold and hot climates, the design can be often constrained by the specific characteristics of the building planned and the size, shape, and orientation of the building plot. Energy-efficient buildings with special designs such as orientation, insulation and windows are being appropriately adapted to withstand severe weather conditions (HOLOPAINEN, 2017). Natural ventilation (MARCONDES *et al.*, 2010) and natural light (FONSECA; DIDONE; PEREIRA, 2012) also play an important role in energy saving. Additionally, one can have buildings with walls composed by different materials (SPECHT *et al.*, 2010) and the consideration of daylight when evaluating buildings regarding energy performance (DIDONE; PEREIRA, 2010).

In a general context, climatic conditions in residential buildings may be determined by using technologies such as air conditioners and heaters. However, using this equipment constantly can generate high energy consumption. An alternative to reduce the use of cooling and heating equipment, maintaining the desired indoor climate conditions,

is to design energy-efficient buildings able to produce such conditions. In order to assess the energy efficiency of a building, its heating and cooling loads should be estimated and analysed based on physical characteristics defined during the design process. Moreover, information such as global location, the purpose of the building, occupation and activity level should be taken into consideration. Among the computational tools for this purpose are those that simulate scenarios which often produce accurate results. For instance, Mustafaraj *et al.* (2014) developed a 3D model related to building architecture, occupancy and heating, ventilation and air conditioning operations. Two calibration stages were considered and the final model identified monthly savings of energy between 20 and 27%. In the Brazilian context, simulation results indicated possible savings in electricity consumption of up to 26% for optimized designs (KRÜGER; MORI, 2012).

Although helpful and interesting in the design cycle, such tools may require advanced knowledge of the user due to the multidisciplinary aspect. In addition, simulations may consume considerable financial and computational costs and results may vary depending on the software used. It should be mentioned that accurate cooling load (CL) and heating load (HL) estimations and correctly identifying parameters that significantly affect building energy demand are necessary to determine appropriate equipment specifications, install systems properly and optimize building designs.

An alternative approach to tackle these drawbacks is to develop a predictive surrogate model that can accurately predict energy consumption based on a few common factors. If the predictive model accurately estimates the simulation model results, then this model could be used instead of the simulation software to estimate performance for different conditions while potentially requiring less information. Considering the context of energy performance in buildings, various efforts to build alternative surrogate predictive models can be identified in the literature.

Using extensive parametric thermal simulations, Pessenlehner and Mahdavi (2003) examined the influence of morphological parameters that define residential building shapes for heating loads. Based on experiments carried out by Pessenlehner and Mahdavi (2003), Tsanas and Xifara (2012) provided a meticulous statistical analysis to gain important insight of the underlying properties of input and output variables. Using the same data collected by Pessenlehner and Mahdavi (2003), Cheng and Cao (2014) and Chou and Bui (2014)

implemented artificial intelligence techniques to predict the energy performance of buildings. Catalina, Virgone and Blanco (2008) developed a set of multiple regression models to predict the monthly heating demand for single-family residential sector in temperate climates. Jinhu *et al.* (2010) built a forecasting model combining Principal Component Analysis to extract the most important features and a weighted support vector regression model for cooling load prediction.

Kwok, Yuen and Lee (2011) used an artificial neural network model to simulate the total building cooling load of an office building in Hong Kong. Online building energy predictions with neural networks and genetic algorithms can also be used in some applications (YOKOYAMA; WAKUI; SATAKE, 2009). Other alternatives include data-driven models (CANDANEDO; FELDHEIM; DERAMAIX, 2017), agent-based modeling (AZAR; NIKOLOPOULOU; PAPADOPOULOS, 2016), graphical approaches (O'NEILL; O'NEILL, 2016) and bio-inspired techniques such as genetic algorithms (BRE *et al.*, 2016).

Chou and Bui (2014) suggested further studies focusing on the optimization of parameters of the model to achieve improvements in their accuracy in predicting heating and cooling loads in buildings. Following their suggestion, the objective of this paper is to use four predictive machine learning techniques which implement a model selection procedure that automatically searches for the best model in a set of user-defined parameters to assess and evaluate the performance of alternative building designs in the early stages of the design process. In addition, this optimized model can help architects to analyze the relative impact of significant parameters of interest while maintaining energy performance standard requirements.

The remainder of this paper is organized as follows: the second section describes the data set, the machine learning methods, the model selection procedure and the performance measures used in

this paper. The third section validates and analyses the performance of all models and compares the results of the simulation. In the same section, a discussion is conducted considering the performance of each method, their strengths and limitations. The last section presents the conclusions.

Method

Machine learning methods can be adopted to estimate \ response from a set of inputs. These methods require a training phase, called supervised training, which considers a dataset drawn from selected variables in the problem domain. The dataset used in the training phase should represent as much as possible the context of the problem in which the tool will be used. This choice may influence their accuracy considerably.

Dataset

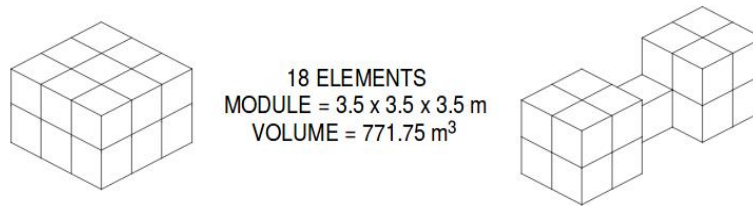
The dataset used in this study is available in Tsanas and Xifara (2012). The data were obtained by the simulation of a set of buildings using a software called Ecotect. Ecotect is an environmental analysis tool compatible with building information modeling software, such as Autodesk Revit Architecture, and is used to perform a comprehensive preliminary building energy performance analysis. It includes a wide range of analysis functions with a highly visual and interactive display enabling analytical results to be presented directly in the context of the building model (YANG; HE; YE, 2014). The dataset consists of eight input variables and two output variables, shown in Table 1. A modular geometry system was derived based on an elementary cube ($3.5 \times 3.5 \times 3.5\text{m}$). In order to generate different building shapes, eighteen such elements were used according to Figure 1. A subset of twelve shapes with distinct relative compactness values (see Table 1) was selected for the simulations, as shown in Figure 2.

Table 1 - Representation of the input and output variables

Description	Type of input/output	Min.	Max.	Mean
Relative Compactness (RC)	Set	0.62	0.98	0.76
Surface area	Set	514.5	808.5	671.71
Wall area	Set	245	416.5	318.50
Roof area	Set	110.25	220.5	176.60
Overall height	Set	3.5	7	5.25
Orientation	Set	2	5	3.50
Glazing area	Set	0	0.4	0.23
Glazing area distribution	Set	0	5	2.81
Heating Load (HL)	Range	6.01	43.1	22.31
Cooling Load (CL)	Range	10.9	48.03	24.59

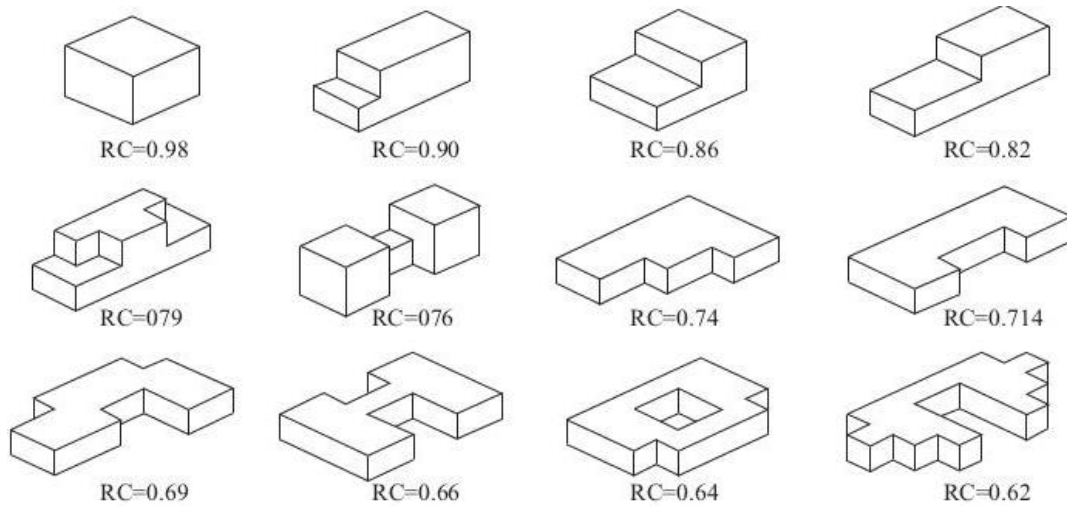
Source: Tsanas and Xifara (2012).

Figure 1 - Generation of shapes based on eighteen cubical elements



Source: Pessenlehner and Mahdavi (2003).

Figure 2 - Relative Compactness coefficient variation



Source: Chou and Bui (2014).

The Relative Compactness (RC) indicator is used to show different building types and it is given by Equation 1:

$$RC = 6V^{2/3} / A \quad \text{Eq. 1}$$

Where:

V is the building volume; and

A is the surface area of the building.

The surface area was calculated as the total of the wall area, roof area and floor area. Figure 3 shows the details of the wall area, roof area, floor area and overall building height.

Four major orientations were considered in the experiments: north, east, west and south. Three percentages of the glazing area to floor area ratio were 10%, 25% and 40%. Moreover, five different glazing distributions were simulated:

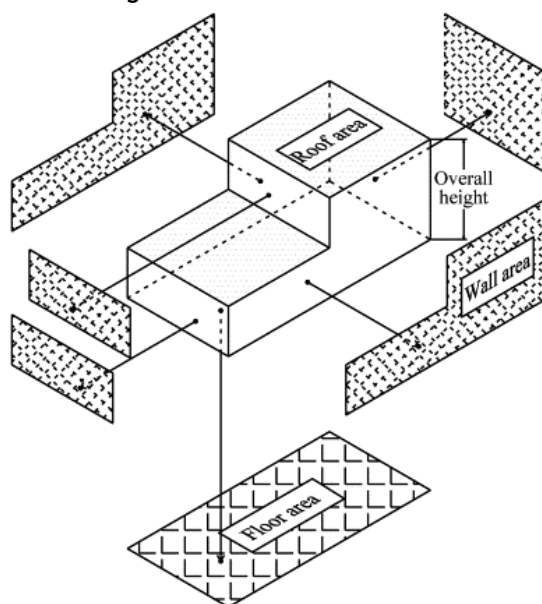
- (a) uniform: with 25% glazing for each face;
- (b) north: 55% for the north face and 15% for each of the other faces;
- (c) east: 55% for the east face and 15% for each of the other faces;
- (d) south: 55% for the south face and 15% for each of the other faces; and

- (e) west: 55% for the west face and 15% for each of the other faces.

Additionally, no glazing areas are simulated in the experiment. Finally, all the buildings were rotated to face the four cardinal directions. Based on this simulation setup, the dataset comprises $12 \times 3 \times 5 \times 4 + 12 \times 4 = 768$ samples of buildings. Table 1 provides the detailed input and output parameters in this study.

The simulation assumes the buildings are in Athens, Greece and each block is occupied by seven people doing sedentary activities, totaling a mean consumption of 70W. The indoor settings of the blocks were defined as: clothing: 0.6 clo, humidity: 60%, air speed: 0.30 m/s, lighting level: 300 lux (equivalent to five 9W LED lamps considering the lamp luminous efficacy as 80 lm/W and the given dimensions of the modular cube). The sensitive and latent internal heat gains were assumed as $5W/m^2$ and $2 W/m^2$, respectively. The air infiltration rate was 0.5 and the air change rate with wind sensitivity was 0.25 air charger per hour. Air change rate with wind sensitivity is an Ecotect parameter that modifies the air infiltration rate based on the current wind speed.

Figure 3 - Generic definition of building areas



Source: Chou and Bui (2014).

For the thermal properties, a mixed mode with 95% efficiency was used, a thermostat range of 19°-24° C, with 15-20 h of operation on weekdays and 10-20 h at weekends. It was considered that all buildings were constructed with the same material, all of which had the lowest U-value. The lower the U-value is, the better the material is as a heat insulator. The characteristics used (U-values between brackets) were: walls (1.780 W/m²K), floor (0.860 W/m²K), roofs (0.500 W/m²K) and windows (2.260 W/m²K). Additional details of the simulation experiments are provided by Tsanas and Xifara (2012).

Machine learning methods

In this study, the algorithms were programmed in Python 2.7 programming language using the *scipy* and *numpy* scientific computing libraries. The *pandas* package was used for data processing and analysis. The regression algorithms and cross validation approaches were implemented using the *Scikit-learn* machine learning library (PEDREGOSA *et al.*, 2011) and the *ffnet* package (WOJCIECHOWSKI, 2011). The following paragraphs describe the machine learning methods used in this paper.

Decision trees (DT) build classification or regression models in the form of a tree structure. They break down a dataset into smaller and smaller subsets while at the same time an associated decision tree is incrementally developed. The final result is a tree with decision and leaf nodes (HASTIE; TIBSHIRANI; FRIEDMAN, 2009). They take a set of attributes as input and return a

predicted value for the respective input. The decision, associated with the decision node, is made by running a test sequence (DUMONT, 2009): each internal node of the tree corresponds to a test of the value of properties and the branches of this node identify possible test values. Each leaf node specifies the return value if the leaf is reached. In this method, the estimated parameter is the maximum depth of the tree.

Support Vector Machines (SVM) (SHANMUGAMANI; SADIQUE; RAMAMOORTHY, 2015) are machine learning algorithms performing a linear combination of attributes by functions called kernel functions aimed to assign a class to a given sample. Different types of kernel functions can be used and different parameters can be varied according to the selected kernel. The SVM is commonly formulated as an optimization problem as follows (Equation 2):

$$\min_{w,b,\xi} \frac{1}{2} (w'w + C \sum \xi_i) \quad \text{Eq. 2}$$

$$\text{subject to } y_i(w\phi(z_i) + b) \geq 1 - \xi_i$$

$$\xi_i \geq 0, i = 1, 2, \dots, h$$

Where:

y_i are the outputs;

x_i are the input samples;

ϕ is used to transform the data to a high-dimensional space;

w represents the decision function coefficients;

the constant $C > 0$ is the error separating hyperplane;

h is the number of support vectors; and ξ is used to penalize the objective function.

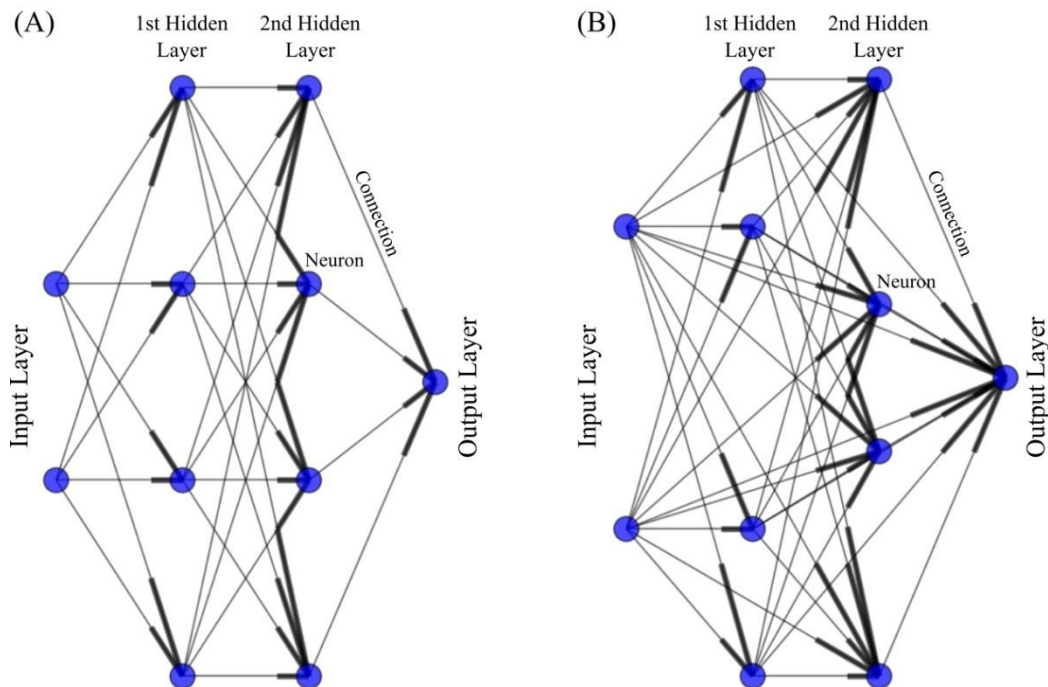
The dot product $\phi'(z_i) \cdot \phi(z_j)$ is replaced by kernel $K(z_i, z_j)$ that has some special properties. In this study, we used the linear kernel represented by $K(z_i, z_j) = z_i z_j$ and the radial basis kernel $K(z_i, z_j) = \exp(-\gamma \|z_i - z_j\|)$, where γ is a parameter of the radial basis function. The performance of the above methods depends on the appropriate choice of parameter C for the linear kernel and γ and C for the RBF kernel.

The Random Forest (RF) (HASTIE; TIBSHIRANI; FRIEDMAN, 2009) is an ensemble learning method for classification that operates by building k decision trees from the training set in k iterations. In each iteration, the training algorithm firstly randomly selects a set of samples from the training set. To reproduce a decision tree from this subset, the RF randomly chooses a subset of features as the candidate features for each node. Thus, each decision tree is built through the ensemble using random independent subsets of both features and samples. The prediction of a new sample class is performed as follows: each individual classifier votes and the most voted class is elected. The

minimum number of samples in newly created leaves is the parameter of this method.

The Multi-Layer Perceptron (MLP) (HAYKIN, 2008; NISSEN, 2005) was used in various areas, performing pattern recognition functions, control and signal processing. This architecture has one or more hidden layers, which comprise computational neurons, also called hidden neurons. The activity of hidden neurons is involved between the external and output layers of the network. Including one or more hidden layers, the network is able to capture non-linear relationships between inputs and outputs. This algorithm uses a number of hidden layers and neurons, the training algorithm, the connectivity and the normalization flag as parameters. If the renormalization flag is set to true, then the data are renormalized. The number of hidden layers is represented as a list of values. For instance, the configuration [5,5] indicates 2 hidden layers of 5 neurons each. The algorithm addresses the following methods to optimize the weights: l-bfgs refers to the Quasi-Newton approach, sgd refers to the Stochastic Gradient Descent Method, and tnc is the gradient information on the truncated Newton algorithm. The connectivity can be simply connected or fully connected. Figure 4 exemplifies the types of connectivity.

Figure 4 - Connectivities for a 2-[4-4]-1 neural network which has two inputs, two hidden layers (4 neurons in each) and one output



Note: the scheme of the simply connected scheme is shown in (A), where the neurons are connected only to the neurons of the previous layer, and in (B) there is the fully connected scheme, where a neuron is connected to all its predecessors.

Grid Search with Cross validation

In order to find the best predictive model and prevent overfitting, an approach based on the grid search and k-fold cross validation was implemented. It is well known that the suitable choice of parameter values of a machine learning method can cause a considerable impact on its accuracy. Furthermore, the optimal values for the parameters can vary according to the problem. Grid Search is a strategy for automatic and optimized parameter adjustments of the model. This technique builds a mesh from sets of predefined values for each parameter. For each possible combination of parameters, the predictive model is trained with some of the data, generating a set of outputs. The best parameter values are those that produced the best set of outputs. The number of configurations for the method is given by Equation 3:

$$\prod_{k=1}^P N_k \quad \text{Eq. 3}$$

Where:

P is the number of parameters; and

N_k is the number of values chosen for the k-th parameter (BERGSTRA; BENGIO, 2012).

In the training step, the strategy known as k-Fold cross validation was adopted, which divides the data set into k sets. The model is trained on k-1 sets and validated with the remaining part. Training and testing steps are repeated k times alternating the training and the testing sets. Figure 5 illustrates the application of k-Fold cross validation. In this study, k = 10 was adopted.

Performance evaluation

Multiple evaluating criteria were used to compare the performance of prediction models. Given a data set composed by N observations, the performance measures the Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) and Coefficient of Determination (R²), which are given by the following Equations (Equations 4-7):

$$MAE = \frac{1}{N} \sum |y_i - z_i| \quad \text{Eq. 4}$$

$$RMSE = \sqrt{\frac{1}{N} \sum (y_i - z_i)^2} \quad \text{Eq. 5}$$

$$MAPE = \frac{100}{N} \sum \frac{|y_i - z_i|}{y_i} \quad \text{Eq. 6}$$

$$R^2 = 1 - \frac{\sum (y_i - z_i)^2}{\sum (y_i - y_M)^2} \quad \text{Eq. 7}$$

Where:

y_i is the expected value for the output variable (HL or CL) with the input x_i ;

z_i is the predicted value for the same input x_i ; and

y_M is the average of the predicted values of the output variable y.

In order to obtain a comprehensive performance measure, the measures known as RMSE, MAE, MAPE and 1-R² are combined into a Synthesis Index (SI), as follows (CHOU; BUI, 2014) (Equation 8).

$$SI = \frac{1}{M} \sum_{i=1}^M \frac{P_i - P_{min,i}}{P_{max,i} - P_{min,i}} \quad \text{Eq. 8}$$

Where:

M is the number of performance measures; and

P_i is the performance measure.

The SI range is 0-1 and an SI value close to 0 indicates a highly accurate prediction model.

Results and discussion

Each machine learning method was trained and validated in 50 independent runs. Table 2 shows the set of parameters used as input for the grid search procedure, as well as the grid size. The machine learning methods appear in the first column: Decision Trees (DT), Multi-Layer Perceptron Neural Network (MLP), Random Forests (RF) and Support Vector Machines (SVM). The second column describes the parameter name for each method, while the third column shows the corresponding parameter settings. The last column shows the grid size, calculated as Equation (3). For example, the grid size for MLP is equal to 80: there are ten configurations for hidden layers, 4 distinct training algorithms and two connectivity schemes. Therefore, the grid size has $10 \times 4 \times 2 = 80$ possible arrangements of parameters. Other parameters involved in the methods, not defined for this step, were kept with the default values as set in the implementations in the scikit-learn package (PEDREGOSA *et al*, 2011).

Figure 6 illustrates the values of the four statistical measures averaged in 50 runs for the predicted heating and cooling loads. In each bar, the vertical black line indicates the standard deviation. For all machine learning methods implemented here, it can be observed that heating loads can be estimated more accurately than cooling loads. This conclusion is in agreement with other studies in the literature. Tsanas and Xifara (2012) conducted an extensive statistical analysis on the same dataset used in this paper. They found both heat and cooling loads are strongly positively correlated to Relative Compactness and overall height, and strongly negatively correlated with the surface area and roof area. The correlation coefficients and details of the statistical procedure can be found in Tsanas and

Xifara (2012). In their study, they concluded that heating loads are estimated with considerably greater accuracy than cooling loads because some variables interact more efficiently to provide an estimate of heating loads.

Taking into consideration the heating loading predictions, it can be observed that all methods produced similar results for all the statistical measures. However, random forests produced the best values for all statistical measures. The good performance of random forests can be explained by the internal optimization problem that is solved

during the training step, which internally accounts for redundant and interacting variables, leading to better prediction abilities. On the contrary, a similar behavior cannot be observed for cooling loading predictions. Clearly, multi-layer perceptron neural networks and support vector machines outperformed random forests and decision trees. The underlying relationships for cooling loads are quite complicated to be adequately captured by random forests and decision trees. In addition, as nonlinear estimators, MLP and SVM show more flexibility in their model parameters which lead to better predictions.

Figure 5 - Illustration of training (green) and testing (blue) sets for $k = 10$

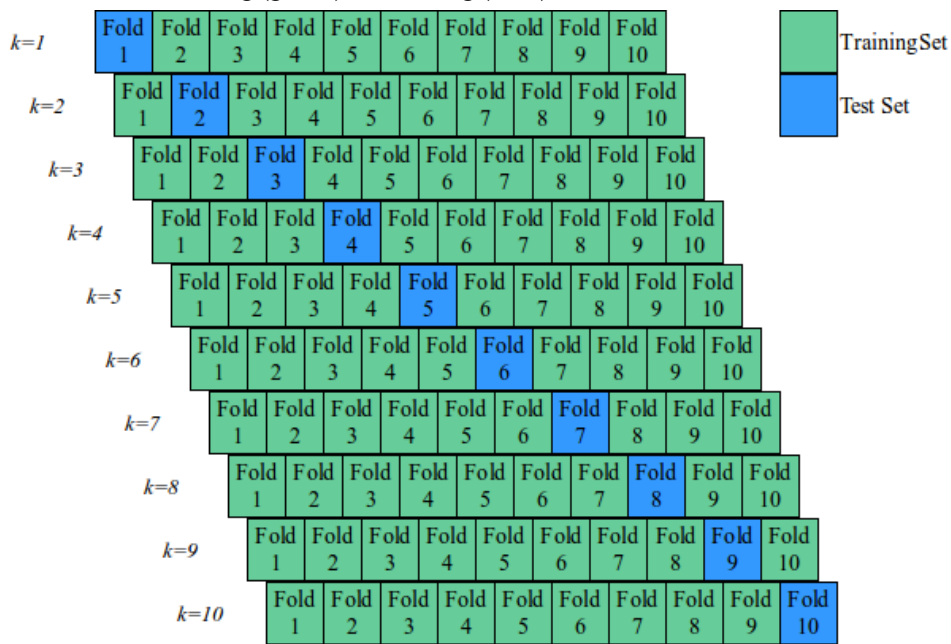
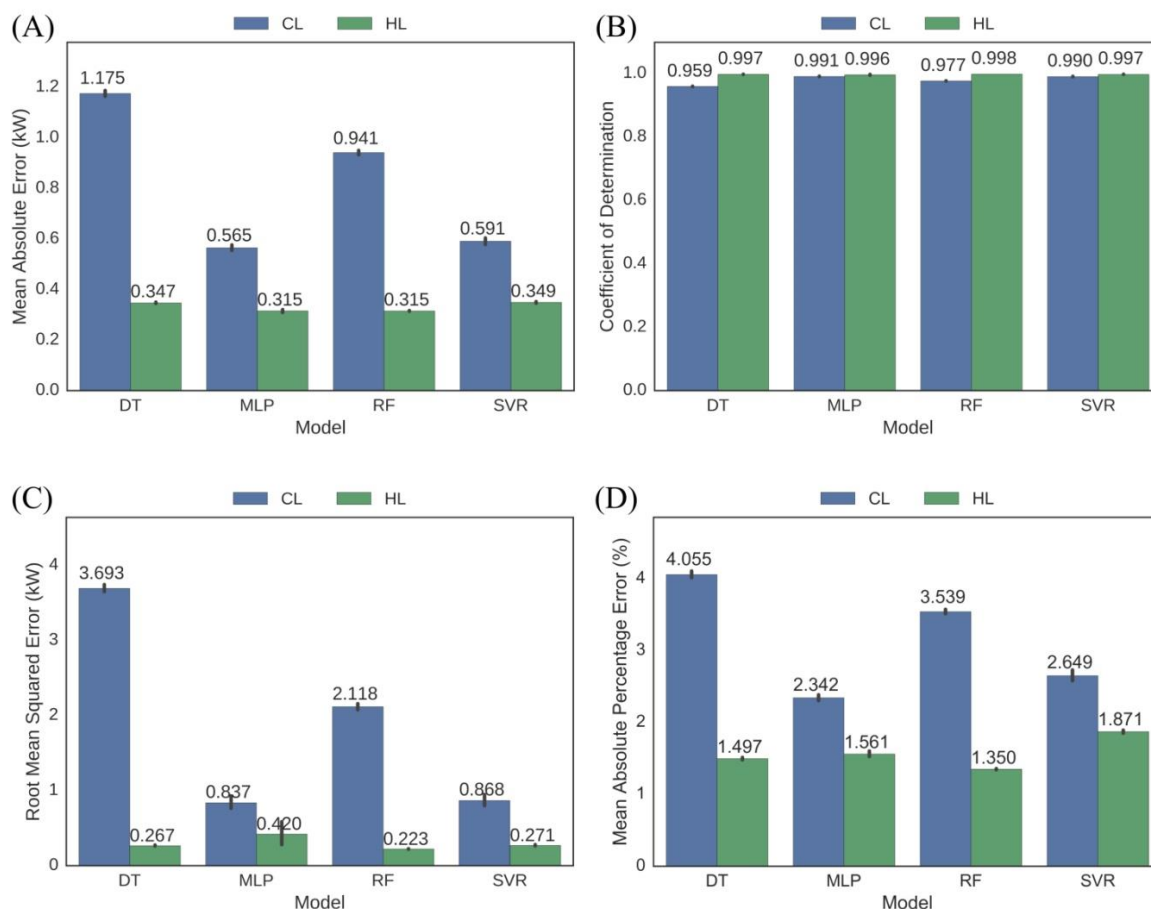


Table 2 - Parameters and their values for applications of grid searches

Method	Parameter Name	Parameter settings	Grid Size
DT	Max depth	[None, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 50]	14
MLP	Number of hidden layers and neurons	[5], [10], [20], [50], [100], [5, 5], [10, 10], [20, 20], [5, 5, 5], [10, 10, 10]	80
	Activation function	[logistic]	
	Training algorithm	[tnc, l-bfgs, sgd, rprop]	
	Connectivity	simply connected, fully connected	
RF	Number of trees	[10, 20, 30]	840
	Bootstrap	[True, False]	
	Max depth	[None, 1, 2, 4, 8, 16, 32]	
	Max features	[auto, 1.0, 0.3, 0.1]	
	Minimum sample leaf	[1, 3, 5, 9, 17]	
SVM	Max iterations	100000	294
	C	[1, 10, 100, 1000, 10^4 , 10^5 , 10^6]	
	σ	[1, 10^{-1} , 10^{-2} , 10^{-3} , 10^{-4} , 10^{-5} , 10^{-6}]	
	Base function	[rbf]	
	ϵ	[10^{-1} , 10^{-2} , 10^{-3} , 10^{-4} , 10^{-5} , 10^{-6}]	

Figure 6 - Barplots for the statistical measures for the heating load (HL in green) and cooling load (CL in blue)



Note: the performance metrics are (A) Mean Absolute Error (MAE), (B) Coefficient of Determination (R^2), (C) Root Mean Square Error (RMSE) and (D) Mean Absolute Percentage Error (MAPE). All the results are averaged on 50 runs.

To compare the performance of the developed models in this paper we used the Synthesis Index (SI). Table 3 lists the summary of averaged statistical measures for the cooling load (CL) and heating load (HL) for each model. Random forests had the best results based on the SI values for the heating load, while the multi-layer perceptron neural network model produced the best SI for cooling loads. Particularly, RF performs better for heating loads, as can be seen when comparing the SI values produced by RF and the remaining predictors. The conclusions obtained when analyzing the cooling loads are different: the support vector machine and multi-layer perceptron neural network show similar statistical measures. However, the neural network performed slightly better in all the measures. The previous analyses suggest two different machine learning models to predict heating and cooling loads. Interestingly, MLP and SVM, which produced the best statistical measures for cooling loads, presented the worst performance for heating loads.

The Synthesis Index (SI) values close to zero indicate a highly accurate prediction model. The performance metrics presented in the table are the Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE) and Coefficient of Determination (R^2). The machine learning models applied are Decision Trees (DT), Multi-Layer Perceptron Neural Network (MLP), Random Forests (RF) and Support Vector Machines (SVM).

Table 4 shows the average real time to perform the grid search and build the models with optimized parameters. The number of folds and the grid size are also shown. The computing time depends on the computational burden of the training algorithm of each model, the number of folds and the parameter grid size. Details of the implementation of this procedure can be found in Buitinck *et al.* (2013). Computer specifications are given as follows: CPU AMD Opteron Processor 6272 (64 cores of 2.1GHz and cache memory of 2MB), RAM of 250GB and operational system Linux Ubuntu 14.04.4 LTS. The

computation time data shows that the whole proposed framework can build optimized machine learning models within minutes. Once constructed, each optimized model performs the predictions quickly, promptly allowing for analysis and parameter testing in the design cycles.

Tables 5 and 6 present the statistical measures for the best models in this paper for both cooling and heating loads. In order to provide a comparison with other models in the literature, the tables also show the results obtained from other studies. Tsanas and Xifara (2012) implemented random forests, while Cheng and Cao (2014) used developed multivariate adaptive regression splines. Chou and Bui (2014) implemented an ensemble model, a linear combination of two or more models to enhance

performance. The results presented by Castelli *et al.* (2015) were obtained by genetic programming, an automated learning of computer programs using a process inspired by biological evolution. As can be seen in Table 4 for the heating load, the best model in this paper shows a better average performance for RMSE and obtained competitive results for the Mean Absolute Error and the Coefficient of Determination. For cooling loads, the Multi-Layer Perceptron model reaches the best average performance for all statistical measures. One of the most important features of neural networks is their flexibility and ability to learn highly nonlinear relationships based on the data. The search for optimized parameters can improve such features, as well as increasing the modeling flexibility.

Table 3 - Averaged statistical measures for cooling loads (CL) and heating loads (HL)

Output	Model	MAE	MAPE	RMSE	R ²	SI ^a
HL	DT	0.347	1.497	0.267	0.997	0.420
	MLP	0.315	1.561	0.420	0.996	0.602
	RF	0.315	1.350	0.223	0.998	0.000
	SVM	0.349	1.871	0.271	0.997	0.622
CL	DT	1.175	4.055	3.693	0.959	1.000
	MLP	0.565	2.342	0.837	0.991	0.000
	RF	0.941	3.539	2.118	0.977	0.553
	SVM	0.591	2.649	0.868	0.990	0.061

Table 4 - Average computing time to perform the grid search and build the models with optimized parameters

Model	Number of folds	Grid size	Average Time (s)	
			HL	Model
DT	10	14	1.6	DT
RF	10	840	70.0	RF
MLP	10	80	1448.5	MLP
SVM	10	294	524.3	SVM

Note: the machine learning models tested for heating (HL) and cooling loads (CL) are Decision Trees (DT), Multi-Layer Perceptron Neural Network (MLP), Random Forests (RF) and Support Vector Machines (SVM). The real time (in seconds) is averaged on 10 runs.

Table 5 - Heating load - comparison between the results of this study and those in the literature used as a reference

Reference	Model	MAE (kW)	RMSE (kW)	MAPE (%)	R ²
Tsanas and Xifara (2012)	Random forests	0.510	–	2.180	–
Cheng and Cao (2014)	Ensemble model	0.340	0.460	–	0.998
Chou and Bui (2014)	Ensemble model	0.236	0.346	1.132	0.999
Castelli <i>et al.</i> (2015)	Genetic programming	0.380	–	0.430	–
This paper	Random forests	0.315	0.223	1.350	0.998

Note: the best results are highlighted in bold. The performance metrics presented in the table are the Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE) and Coefficient of Determination (R²).

Table 6 - Cooling load - comparison between the results of our study and those in the literature used as a reference

Reference	Model	MAE (kW)	RMSE (kW)	MAPE (%)	R ²
Tsanas and Xifara (2012)	Random forests	1.420	–	4.620	–
Cheng and Cao (2014)	Ensemble model	0.680	0.970	–	0.990
Chou and Bui (2014)	Ensemble model	0.890	1.566	3.455	0.986
Castelli <i>et al.</i> (2015)	Genetic programming	0.970	–	3.400	–
This paper	Neural network	0.565	0.837	2.342	0.991

Note: the best results are highlighted in bold. The performance metrics presented in the table are the Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE) and Coefficient of Determination (R²).

Although the predictive models proposed here and those found in the literature produced accurate results, they should be used carefully. It should be mentioned that they are only applicable to the twelve specified building types considering the simulated experiment setup. Besides, even in computer simulations, uncertainties in thermal and physical properties of materials can influence thermal performance (SILVA; ALMEIDA; GHISI, 2017) and may be considered. Comprehensive tests using real data are necessary to assess the performance of the methods in real world situations, leading to the development of new and improved models. Some authors using data measured from a wireless sensor network have identified that atmospheric pressure, exterior air temperature and wind speed are important parameters to predict energy loads (CANDANEDO; FELDHEIM; DERAMAIX, 2017).

Conclusions

This paper evaluated the application of four machine learning methods to predict energy efficiency in residential buildings: decision trees, random forests, multi-layer perceptron neural networks and support vector machines. Their parameters were adjusted through the grid search and trained with cross validation. The dataset consists of a data set of 768 simulated buildings.

From the results obtained, random forests proved to be the best option for predicting heating loads while multi-layer perceptron neural networks produced the most accurate results for cooling loads. Support vector machines obtained accurate predictions for cooling loads, but with a slightly lower performance. After comparing them with the machine learning methods found in the literature, the results obtained in this paper show that the search in the parameters can generate accurate models, and are an alternative for early prediction of building cooling and heating loads. However, the machine learning methods developed here, even though accurate, are only applicable to the twelve specified building types in the simulated dataset.

The models with optimized parameters developed in this study are able to evaluate different sets of parameters, resulting in simulation settings that can potentially avoid modeling and testing various prototypes, helping to save resources in the initial phase of the design. Expecting to improve the results presented here, other machine learning methods can be implemented in further research. In addition, the grid search strategy can be replaced by an optimization evolutionary algorithm to set the parameters of the machine learning methods.

References

- AZAR, E.; NIKOLOPOULOU, C.; PAPADOPOULOS, S. Integrating and Optimizing Metrics of Sustainable Building Performance Using Human-Focused Agent-Based Modeling. *Applied Energy*, v. 183, p. 926–937, 2016.
- BERGSTRA, J.; BENGIO, Y. Random Search for Hyper-parameter Optimization. *Journal of Machine Learning Research*, v. 13, p. 281–305, fev. 2012.
- BRE, F. *et al.* Residential Building Design Optimisation Using Sensitivity Analysis and Genetic Algorithm. *Energy and Buildings*, v. 133, p. 853–866, 2016.
- BUITINCK, L. *et al.* API Design for Machine Learning Software: experiences from the scikit-learn project. In: EUROPEAN CONFERENCE ON MACHINE LEARNING AND PRINCIPLES AND PRACTICES OF KNOWLEDGE DISCOVERY IN DATABASES, Prague, 2013. *Proceedings...* Prague, 2013.
- CANDANEDO, L. M.; FELDHEIM, V.; DERAMAIX, D. Data Driven Prediction Models of Energy Use of Appliances in a Low-Energy House. *Energy and Buildings*, v. 140, p. 81–97, 2017.
- CASTELLI, M. *et al.* Prediction of Energy Performance of Residential Buildings: a genetic programming approach. *Energy and Buildings*, v. 102, p. 67–74, 2015.

- CATALINA, T.; VIRGONE, J.; BLANCO, E. Development and Validation of Regression Models to Predict Monthly Heating Demand for Residential Buildings. **Energy and Buildings**, v. 40, n. 10, p. 1825–1832, 2008.
- CHENG, M.-Y.; CAO, M.-T. Accurately Predicting Building Energy Performance Using Evolutionary Multivariate Adaptive Regression Splines. **Applied Soft Computing**, v. 22, p. 178–188, 2014.
- CHOU, J.-S.; BUI, D.-K. Modeling Heating and Cooling Loads by Artificial Intelligence for Energy-Efficient Building Design. **Energy and Buildings**, v. 82, p. 437–446, 2014.
- DIDONE, E. L.; PEREIRA, F. O. R. Simulação Computacional Integrada Para a Consideração da Luz Natural na Avaliação do Desempenho Energético de Edificações. **Ambiente Construído**, Porto Alegre, v. 10, n. 4, p. 139–154, out./dez. 2010.
- DUMONT, M. Fast Multi-Class Image Annotacion With Random Subwindows and Multiple Output Randomized Trees. In: INTERNATIONAL CONFERENCE ON COMPUTER VISION THEORY AND APPLICATIONS, 6., Lisboa, 2009. **Proceedings...** Lisboa, 2009.
- FONSECA, R. W. de; DIDONE, E. L.; PEREIRA, F. O. R. Modelos de Predição da Redução do Consumo Energético em Edifícios que Utilizam a Iluminação Natural Através de Regressão Linear Multivariada e Redes Neurais Artificiais. **Ambiente Construído**, Porto Alegre, v. 12, n. 1, p. 163–175, jan./mar. 2012.
- FRIESS, W. A.; RAKHSHAN, K. A Review of Passive Envelope Measures for Improved Building Energy Efficiency in the {UAE}. **Renewable and Sustainable Energy Reviews**, v. 72, p. 485–496, 2017.
- HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The Elements of Statistical Learning – Data Mining, Inference, and Prediction**. 2nd. ed. New York: Springer, 2009.
- HAYKIN, S. O. **Neural Networks and Learning Machines**. 3th. ed. New Jersey: Prentice Hall, 2008.
- HOLOPAINEN, R. Cost-Efficient Solutions for Finnish Buildings. In: PACHECO-TORGAL, F. *et al.* (Eds.). **Cost-Effective Energy Efficient Building Retrofitting**. [S.l.]: Woodhead Publishing, 2017.
- JINHU, L. *et al.* Applying Principal Component Analysis and Weighted Support Vector Machine in Building Cooling Load Forecasting. In: COMPUTER AND COMMUNICATION TECHNOLOGIES IN AGRICULTURE ENGINEERING, 2010. **Proceedings...** 2010.
- KRUGER, E. L.; MORI, F. Análise da Eficiência Energética da Envoltória de Um Projeto Padrão de Uma Agência Bancária em Diferentes Zonas Bioclimáticas Brasileiras. **Ambiente Construído**, Porto Alegre, v. 12, n. 3, p. 89–106, jul./set. 2012.
- KWOK, S. S.; YUEN, R. K.; LEE, E. W. An Intelligent Approach to Assessing the Effect of Building Occupancy on building Cooling Load Prediction. **Building and Environment**, v. 46, n. 8, p. 1681–1690, 2011.
- MARCONDES, M. P. *et al.* Conforto e Desempenho Térmico nas Edificações do Novo Centro de Pesquisas da Petrobras no Rio de Janeiro. **Ambiente Construído**, Porto Alegre, v. 10, n. 1, p. 7–29, jan./mar. 2010.
- MELO, A. *et al.* A Novel Surrogate Model to Support Building Energy Labelling System: a new approach to assess cooling energy demand in commercial buildings. **Energy and Buildings**, v. 131, p. 233–247, 2016.
- MUSTAFARAJ, G. *et al.* Model Calibration for Building Energy Efficiency Simulation. **Applied Energy**, v. 130, p. 72–85, 2014.
- NISSEN, S. **Neural Networks Made Simple**. Software 2.0, v. 2, p. 14–19, 2005.
- O’NEILL, Z.; O’NEILL, C. Development of a Probabilistic Graphical Model for Predicting Building Energy Performance. **Applied Energy**, v. 164, p. 650–658, 2016.
- PEDREGOSA, F. *et al.* Scikit-Learn: machine learning in Python. **Journal of Machine Learning Research**, v. 12, p. 2825–2830, 2011.
- PESSENLEHNER, W.; MAHDAVI, A. Building Morphology, Transparency, and Energy Performance. In: EIGHTH INTERNATIONAL IBPSA CONFERENCE, Eindhoven, 2003. **Proceedings...** Eindhoven, 2003.
- SHANMUGAMANI, R.; SADIQUE, M.; RAMAMOORTHY, B. Detection and Classification of Surface Defects of Gun Barrels Using Computer Vision and Machine Learning. **Measurement**, v. 60, p. 222–230, 2015.
- SILVA, A. S.; ALMEIDA, L. S. S.; GHISI, E. Análise de Incertezas Físicas em Simulação Computacional de Edificações Residenciais. **Ambiente Construído**, Porto Alegre, v. 17, n. 1, p. 289–303, jan./mar. 2017.

SPECHT, L. P. *et al.* Análise da Transferência de Calor em Paredes Compostas por Diferentes Materiais. **Ambiente Construído**, Porto Alegre, v. 10, n. 4, p. 7–18, out./dez. 2010.

TSANAS, A.; XIFARA, A. Accurate Quantitative Estimation of Energy Performance of Residential Buildings Using Statistical Machine Learning Tools. **Energy and Buildings**, v. 49, p. 560–567, 2012.

WOJCIECHOWSKI, M. Application of Artificial Neural Network in Soil Parameter Identification for Deep Excavation Numerical Model. **Computer Assisted Mechanics and Engineering Science**, v. 18, n. 4, p. 303–311, 2011.

YANG, L.; HE, B.-J.; YE, M. Application Research of Ecotect in Residential Estate Planning. **Energy and Buildings**, v. 72, p. 195–202, 2014.

YOKOYAMA, R.; WAKUI, T.; SATAKE, R. Prediction of Energy Demands Using Neural Network With Model Identification by Global Optimization. **Energy Conversion and Management**, v. 50, n. 2, p. 319–327, 2009.

Acknowledgments

This research was supported by the following Brazilian agencies CNPq (grant 305099/2014-0), CAPES and FAPEMIG (grants TEC APQ 01606/15 and TEC PPM 388/14).

Grasiele Regina Duarte

Departamento de Ciência da Computação, Instituto de Ciências Exatas | Universidade Federal de Juiz de Fora | Rua José Lourenço Kelmer, s/n, São Pedro | Juiz de Fora - MG - Brasil | CEP 36036-900 | Tel.: (32) 2102-3327 | E-mail: grasird@yahoo.com.br

Leonardo Goliatt da Fonseca

Departamento de Mecânica Aplicada e Computacional, Faculdade de Engenharia | Universidade Federal de Juiz de Fora | Tel.: (32) 2102-3469 | E-mail: goliatt@gmail.com

Priscila Vanessa Zabala Capriles Goliatt

Departamento de Ciência da Computação, Instituto de Ciências Exatas | Universidade Federal de Juiz de Fora | Tel.: (32) 2102-3481 Ramal 31 | E-mail: priscilacapriles@gmail.com

Afonso Celso de Castro Lemonge

Departamento de Mecânica Aplicada e Computacional, Faculdade de Engenharia | Universidade Federal de Juiz de Fora | Tel.: (32) 3229-3412 | E-mail: afonso.lemonge@ufjf.edu.br

Revista Ambiente Construído

Associação Nacional de Tecnologia do Ambiente Construído
Av. Osvaldo Aranha, 99 - 3º andar, Centro
Porto Alegre - RS - Brasil
CEP 90035-190
Telefone: +55 (51) 3308-4084
Fax: +55 (51) 3308-4054
www.seer.ufrgs.br/ambienteconstruido
E-mail: ambienteconstruido@ufrgs.br