The potential of artificial neural networks to model daylight harvesting in buildings located in different climate zones

Raphaela Walger da Fonseca[1], Fernando O. R. Pereira[1], Konstantinos Papamichael[2]
[1]Environmental Comfort Laboratory, Federal University of Santa Catarina, Brazil
[2]California Lighting Technology Center, UC Davis, USA

Abstract

Tools able to allow fast construction performance feedback in early design stages are essential to the design process. Artificial neural networks (ANNs) have high potential to metamodeling building energy performance considering daylighting harvesting. The aim of the present study was to assess the potential of ANNs to predict the energy performance of daylit buildings located in different climate. The ANNs were trained using results from parametric computational energy simulations, by taking into consideration the key daylighting parameters. Accordingly, the ANN was able to predict the total energy consumption in different climates when cities from different hemispheres were taken into account.

Introduction

Numerous studies highlight the importance of taking into consideration the daylight integration in the early stages of the design process in order to understand its benefits (Pereira et al., 2005). The use of simplified simulation tools is an alternative to the lack of detailed building information inherent to the initial design phases, since computer simulations demand high detailing, as well as trained professionals, despite the fact that they consume more time. One of the main difficulties to integrate daylight to simplified tools used to predict the energy consumption in the building is the simultaneous influence of the electric lighting and of the air conditioning. Fonseca (2015) conducted a survey on simplified tools and found 72 tools available. These tools were analyzed according to their potential to assess the representation of light and thermal phenomena; building description,context and performance parameters; as well as their conception. The research concluded that, although there is a large number of daylighting tools, few of them allow assessing daylight impact on energy consumption, according to its implications in the electric lighting and HVAC use patterns. In addition, the existing tools have limitations inherent to their design (e.g., algorithms, methods, etc.) or application (e.g., methodological process such as the clustering of thermal zones, the definition of key variables or climate dynamism consideration).

The literature review evidenced the high potential of artificial neural networks (ANN) to predict energy consumption and even, to a lesser extent, to take daylighting into consideration (Datta and Tassou, 1998; Kalogirou, 2000; Kalogirou, 2001; Krarti, 2003; Neto and Fiorelli, 2008; Yezioro, Dong and Leite, 2008; Ekici and Aksoy, 2009; Magnier and Haghighat, 2010; Li, Su and Chu, 2011; Melo, 2012; Zhao and Magoulès, 2012; Aqlan et al., 2014; Buratti, Barbanera and Palladino, 2014; Coakley, Raftery and Keane, 2014; Fumo, 2014; Lepadatu, Judele and Rosu, 2014; Melo et al., 2014; Yang, Asadi and Geem, 2014; Kalogirou et al., 2015; Lu, Lü and Kibert, 2015; Zhang et al., 2015, Wong, Wan and Lam, 2010).

Artificial neural networks (ANNs) are nonlinear computational techniques that use artificial intelligence to represent or approximate systems (Haykin, 2001). Therefore, ANNs can be used as a simplified tool to predict energy consumption by taking daylighting into consideration, as well as few input information of instant results. However, gaps related to ANN accuracy, to the effects of building model orientation, or to aspects such as annual climate variation and different geographic locations were found.

The study addressed in the present paper is part of a more extensive research that investigated ANN potential to metamodel building description, context and performance variables focused on daylight-related energy performance in buildings. The ANNs were used as metamodel to model an energy simulation performed in the Energyplus software.

This research has deeply investigated the ANN aspects such as training and testing the suitability of data sets; networks' input (key-variables) and output parameters; and networks' architecture. The study focuses on building context key-variables in order to assess the ANNs potential to model the energy performance of buildings located in different climate zones.

Method

A systematic research approach was adopted in order to test ANN limitations concerning the key variables for daylight and energy performance modeling. The research was conducted in order to analyze: **i.** the influence of input parameter clusters from different building context categories; **ii.** the effect of grouping locations under significantly different geographic conditions, such as different hemispheres separated by the equatorial line, wherein the "orientation" parameter has opposite sunshine and radiation meanings; **iii.** the adequacy of different key–variable types to describe the thermal parameters of the climate; and **iv.** describing the sky conditions in different locations.

The ANNs were used as metamodels in the current investigation and their training was carried out based on results from energy simulations. The metamodel experiments were based on interactions (Table 2) and only used one luminous-thermal zone (Figure 1). It was done in order to compute the orientation effect, which is mitigated when the thermal zones are grouped. The study was based on a sequence of actions and the result of each action has motivated the next research stage. The predictive potential of the metamodels was assessed by comparing the results predicted through ANNs to the simulation results. If the error was lower than 5%, it was assumed that ANN was able to model the relations between key variables, whereas if it was higher than 5%, new alternatives were tried. This process was conducted by pushing the network's limits to force it to fail (e.g.: reducing training set, reducing input parameters, changing the ANN architecture, etc.). This approach aimed at checking the point in which the network would start failing. The tests started from the simplest ANN, which could learn and generalize solutions defined in the EasyNN software.

The research process was developed according to: a) variable selections; b) data sampling and parametrization; c) energy simulation; d) ANN training; e) result analyses and the set-up of the following research step.

The research steps

A) Variable selections:

The training, validation, and testing of ANNs required a set of data to be learned and generalized by the network. This set was established by combining key-variables in order to generate the parametric models.

The parameters of the building context were selected according to a metamodel based on few input key variables easy to be found, including: geographic location (Latitude, Longitude, Altitude); thermal climate features (Cooling Degree Days - CDD, Heating Degree Days - HDD); and daylight availability (Global Solar Radiation - Accumulated - GSR_a and Cloudiness - Accumulated - Cl_a). Eleven (11) cities with different climates were taken into account, nine cities from the Southern hemisphere and two from the Northern hemisphere. Table 1 shows the selected cities, as well as their respective building-context parameter values.

The luminous-thermal zone is 16m long, 8m deep and 3m tall, the window is in the larger dimension (Figure 2). The 4 cardinal orientations, 5 window-to-wall-ratio variations (WWR), 5 visible transmittance combinations (VT), and the solar heat gain coefficient (SHGC) were simulated (Table 2). The output parameters herein taken into consideration were Total Energy consumption, Cooling, Heating and Lighting consumption. The minimum and maximum values of each variable were chosen to avoid ANN extrapolation.

B) Sampling data and parametrization:

The key variables were punctually chosen according to the research goals; the sample was the permutation of the values presented in Table 2. The parametric modeling used an Excel spreadsheet macro adapted from Westphal (2012). This macro combines the variables, assembles the *.idf files of EnergyPlus (D.O.E-US, 2012), triggers the simulations and records their results.

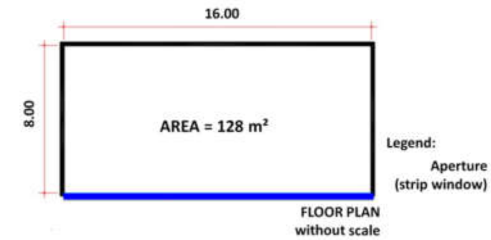


Figure 1: The luminous-thermal zone floor plan.

Table 2: Parametric variables of the building context

VARIABLE	VARIATIONS
Orientation	N (0°); E (90°); S (180°); W (270°)
WWR (%)	20; 35; 50; 65; 80
Glass (VT/SHGC)	0.25/0.467; 0.45/0.602; 0.57/0.62; 0.74/0,623; 0.91/ 0.898

C) Energy simulations:

The energy simulations were performed in the EnergyPlus software (D.O.E-US, 2012), the daylighting was taken into account and assessed through the split flux method (object:Daylighting:Controls).

D) ANN training:

Two software applications were used to generate the neural networks: MATLAB (The Mathworks Inc., 2011) and EasyNN-plus (Wolstenholme, 2013). The software selection aimed at facilitating rapid tests and programming flexibility. The MATLAB is a well-established software that uses ANN toolboxes to develop ANNs. The software has several file export formats to facilitate the interface with other software, as well as to make it possible programming several neural network types. In turn, the greater freedom of programming also requires greater dedication to the code writing. Alternatively, although the EasyNN-plus software is more limited, it has a user-friendly interface, as well as useful tools, for example, it indicates the number of neurons in the hidden layer in a particular network configuration. The software presents a programming option; however, as it was not designed for programming, this option is more limited. Thus, the EasyNN-plus was used along with MATLAB. The first one was used to define the simplest ANN the EasyNN-plus could learn, whereas MATLAB was used in all other tests in order to allow performing a larger number of tests.

The following steps were adopted to generate the neural networks: pattern definition, network initialization, training parameter definition, network training and network testing.

The pattern definition consisted of importing the data set, as well as of defining the input and output vectors, the variable presentation mode (for example, binary, real, etc.) and the data presentation mode to the network, either in order or randomly.

The network initialization was conducted through data set import and normalization.

Next, the training methods were defined, as well as the way of dividing the data set into training, validation and testing sets. The activation/transfer function, the network architecture type, the learning process, the learning algorithm and the learning rule were set.

The multiple training applied to the same set was adopted, and the set was randomly divided in training and testing rates. The multiple training was selected to allow observing the network-performance trends, since the aim of the study was to assess the limitations of the neural networks to model the daylighting effects on energy consumption, instead of proposing an efficient ANN. The sets were divided as follows: 80% of the data to training, and 20% to testing (literature test approach). Whenever the validation set was applicable, between 10% and 20% of the data belonging to the training set were selected to validate it. Thus, 5 to 10 neural networks were trained for each investigation in the current study.

The log-sigmoid function was adopted as activation/transfer function due to its great applicability (The Mathworks Inc., 2011). Such function requires normalizing the network input data, which should be limited between 0 and 1 (Equation 1).

$$\left(z_i^k\right)_n = \frac{z_i^k - z_{\min}^k}{z_{\max}^k - z_{\min}^k} \quad (1)$$

wherein: k = is the optimal maximization attribute 1;
i = is the alternative value z^k_i;
z^k_{min} = minimum set value; and
z^k_{max} = maximum set value.

The networks' architecture was non-recurrent and feedforward through multiple layers by the adoption of the multilayer perceptron – MLP.

The heuristic solutions used to set the ANN MLP size should take into consideration the network's convergence and generalization power (Silva, 2005). The EasyNN automatic function, and a heuristic commonly found in discussion forums about neural networks, were used to define the number of neurons in the hidden layers (Equation 2).

$$N_{HIDDEN} = 1.5 * N_V \quad (2)$$

Wherein:
N_{hidden}: is the number of neurons in the hidden layer; and
N_V: is the number of input variables.

The tests followed the recommendations by Silva (2005) in order to fix the network architecture and size, as well as to change the training set; or to fix the training set and vary the ANN architecture.

The supervised learning was the learning process applied to the neural networks; the networks' input and output sets were known. The herein adopted learning algorithm was the error backpropagation algorithm. Delta rule, which consists of using the Gradient Descent (GD), was originally applied to this algorithm. The use of regularization methods is one of the ways to improve network generalization and avoid overfitting. Thus, the Bayesian Regularization (BR) algorithm was also approached.

The following parameters were set to control the network training: learning, momentum and minimum gradient rates; validation rules; and training stop criteria. The values of these parameters were defined based on the literature and on experimentation, according to each test.

E) Result analyses and set up of the following research step:

The network performance analyses were based on the individual error analyses and on the mean error of the multiple training of ANNs. Thus, it was possible identifying the performance trends of the herein studied ANNs typologies. The performance analyses took into consideration the mean absolute error (MAE) (Equation 3), the mean absolute percentage error (MAPE) (Equation 4) and the root-mean-square error (RMSE) (Equation 5).

$$MAE = \frac{\sum_{t=1}^{n} \left|e_t\right|}{n} \quad (3)$$

wherein: MAE: is the mean absolute error;
e_t: is the error in period t; and
n: is the number of used records (periods).

$$MAPE = \frac{\sum_{t=1}^{n} \left|ep_t\right|}{n} \quad (4)$$

wherein: MAPE: is the mean absolute percentage error;
EP_t: is the percentage error in period t; and
n: is the number of used records (periods).

$$RMSE = \frac{\sum_{t=1}^{n} e_t^2}{n} \quad (5)$$

wherein: RMSE: is the root-mean-square error;
e_t: is the error in period t; and
n: is the number of used records (periods).

These errors were individually assessed in the training, validation and testing set, as well as in the total data set (the sum of the previous sets). It was done because most of the studies in the literature review just presented neural network errors in the total data set. Since the test set is generally smaller than the training set, the overall mean network error tends to the training set error, which, in most cases, is significantly lower than the test set error and masks the limitations of the metamodels. In addition to the error analyses performed in all ANN sets, each network output parameter was analyzed, individually and together, in order to find a global perspective about the networks' performance.

It is consensus that is necessary separating the data set percentage to be transferred to the network after its training is concluded in order to test the network's generalization power. However, it is believed that this method is not completely effective in the case of metamodels supported by ANNs based on results from parametric computational simulations. It is understood that these recommendations come from studies about ANNs based on real data, such as biological experiments based on sample collections, for example. In the case of metamodels based on parametric simulations, when a fraction of the set is separated for testing, the neural network previously "sees" these variables during the learning process. The network does not see exactly the same combination between variables in the test set, but rather the values adopted for these variables. Input combinations, herein referred to as "never seen" case sets, whose input variable values were not presented to the network during training, were proposed in order to generate a more demanding test method. Two locations were selected to be used as "never seen" cases, namely: São Paulo and Salvador (Table 1).

Testing the network by using the "never seen" cases is of fundamental importance to define the limitations of the metamodels, mainly when they are based on parametric metamodeling. As the number of cases to be simulated in order to train the networks is limited by time feasibility, the building solutions presented to the networks are also limited. Thus, the network must be able to generalize new combinations of the variable values presented to it (the herein named literature approach), as well as new values for these variables. Even within the limits of the values presented during training, these new values have not yet been presented to the network. Thus, the combination between the two testing methods is more effective, since the first allows checking the network generalization potential and the second one allows identifying network limitations.

In view of the aforementioned, the investigation followed the five action groups below:

Group 1 –Input key-variable grouping test

Action 1: Assessing the effect resulting from the clustering input parameters from different categories, according to the following sets:

Assessing the effect of grouping different input parameter categories according to the following sets:

- all input parameters (G+T+S): Geographic Location (G), Thermal climate features (T) and Sky conditions (S) **(base-case)**;
- G+T: Geographic Location (G) and Thermal Climate Features (T);
- G+S: Geographic Location (G) and Sky Conditions (S);
- T+S: Thermal climate features (T) and Sky conditions (S);
- G: Geographic Location (G);
- T: Thermal climate features (T); and
- S: Sky Conditions (S).

The initial ANN configuration was based on previous tests conducted in the global research the present study is part of (Fonseca, 2015). The Bayesian Regularization algorithm was adopted in this stage, and the hidden layer size was defined as four times the number of input variables.

Action 2: A new test set was conducted and it considered the number of hidden neurons. The hidden layer size was adjusted according to the input parameters of the layer size by following the size ratio of the base-case layers.

Group 2- Tests to group the locations by hemispheres

Action 1: Two neural network solutions based on the same network configuration found in the base-case solution used in the previous actions group were suggested: one comprising cities from both the Northern and the Southern hemispheres in the training set; and the other only comprising cities from the Southern hemisphere (Table 1). The "never seen" set was the same in both cases; this set comprised two cities from the Southern hemisphere - São Paulo and Salvador -, whose patterns were not presented to the network during training, but belonged to the same Brazilian bioclimatic zone - BBZ (ABNT, 2005) - of some other city that participated in such training. The network architecture adopted the all input parameters, G + T + S cluster, as the input parameter by using the Bayesian Regularization algorithm and 40 hidden neurons.

Action 2 - Defining the cities to be used to test the networks focused on avoiding extrapolation; the definition was based on the Latitude parameter. In addition, the city should be located in the same bioclimatic zone of some of the cities presented to the training set. Despite the adoption of several climate input parameters, they were not parametrically combined with each other, since together they represent each city. The range of every input parameter was jointly assessed in order to check whether the city selection criterion was appropriately chosen. This assessment allowed identifying the variables that could lead the network to extrapolation.

Action 3 - Based on this assessment, the networks were retested using the adjustments resulting from the previous action.

Group 3– Verifying the suitability of key variables to describe the thermal climate features

The applicability of another set based on primary thermal features, i.e., measurable features, was tested in order to check the suitability of the key variables selected to model the climate thermal features. The values found in January and July were selected to indicate the summer and winter conditions in both hemispheres, as well as to capture greater cold and heat variations. Thus, the selected variables were: Dry bulb temperature (summer) – mean (DBT_S); Dry bulb temperature (winter) – mean (DBT_W); Relative humidity (summer) – mean (RH_S); Relative humidity (winter) – mean (RH_W); Temperature variation (summer) – mean (ΔT_S); and temperature variation (winter) – mean (ΔT_W) (Table 1). The building description, geographic location and sky condition parameters were kept.

The same network configuration used in the previous action groups was herein adopted: G + T + S clustering using cities from both hemispheres by considering the adjustment of the cities in the training set, according to

Group 2 - Action 2. The networks comprising the six new thermal parameters were tested through the Bayesian Regularization and Gradient Descent algorithms, according to the stop criterion: *up to 30% of the validation set was within the error limit.*

Group 4 - Addition of Cumulative Horizontal Global Illuminance as a key variable to daylight availability

A single variable - Accumulated Global Horizontal illuminance (GHIa) - was added to complete the information about sky conditions in order to check the selected key variables suitability to model the Sky Conditions based on daylight availability (Table 1). The Gradient Descent algorithm test was also repeated for experimental purposes by using the stop criterion: *up to 30% of the validation set was within the error limit*. The original parameters concerning building description, geographic location and thermal climate features were kept.

Group 5 - Validating the metamodels test method based on parametric energy simulations: "literature" approach vs. "never seen" approach

A comparison between the "literature" and the "never seen" approach test methods was conducted in order to check the effectiveness of these test solutions in assuring model robustness. This information is especially important to check the applicability of ANNs to metamodel different climates, since the computational effort required to generate data for the metamodeling process should be taken into account in the metamodel feasibility verification. Thus, the current study kept the data set, as well as the same network configuration found in the base-case.

Discussion and result analysis

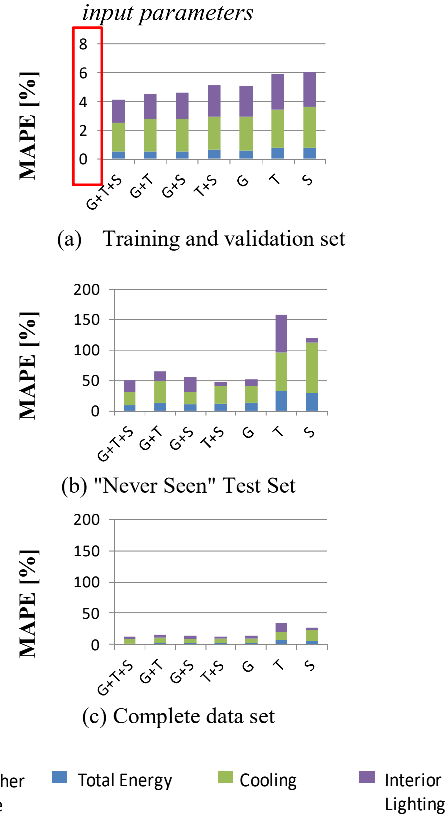
The error analysis presented in the figures found in the following section are based on the mean ANNs multiple training results.

Group 1 –Input key-variables grouping test

Action 1 - The network input parameter clustering test G + T + S overall presented the best performance, mainly when the training set was separately assessed. However, the difference between the performance of this cluster and that of the G + S, T + S and G clusters was lower than 2% (in absolute terms) when the entire network was assessed; and lower than 4 % in the "never seen" set. The clustering process applied to the three categories has generated a larger and, consequently, more complex network. Such outcome required greater computational effort and considerably increased execution time (training). Therefore, the balance between time and performance should be taken into consideration at the time to choose the model's input variables. In such cases, it is advisable to consider assessing the network performance per output parameter in order to prioritize the performance of the most relevant output parameter.

It was possible seeing that the training (Figure 2a) and "never seen" sets (Figure 2b) have presented the same error trend in the input parameters. However, they have shown different magnitudes; the "never seen" set has presented significantly higher errors, and the graph scales ranged from 8% to 200%. The G + T + S cluster was the best-performance solution, and it was followed by the two-group clustering solutions (for all sets). By observing the complete data set (Figure 2c), it was possible seeing that the Geographic Location (G) has presented better performance than G + T, wherein the MAPE was equal to 13% and 15%, respectively. The Geographic Location (G) features were more meaningful to ANN than the other categories, since the Geographic Location performance (G) alone was better than the performances of the Thermal Climate Features (T) and Sky Conditions (S). The relevance of the Sky Conditions appeared to be more substantial than the Thermal Climate Features. It happened because, when the Thermal Climate Features (G + T) were combined with the Geographic Location (G), they presented higher MAPE than the Sky Conditions (G + S), 15% and 14%, respectively. In addition, when the Thermal Climate Features (T) were separately observed, they presented 33% MAPE, whereas the Sky Conditions (S) presented 26%.

Figure 2- Comparison between the network groupings of input parameters



(a) Training and validation set

(b) "Never Seen" Test Set

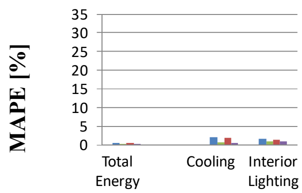
(c) Complete data set

Action 2 - The number of neuron adjustment tests applied to the seven network input parameter clusters has overall worsened the network performance in all the clusters, since the errors increased when the neurons were adjusted.

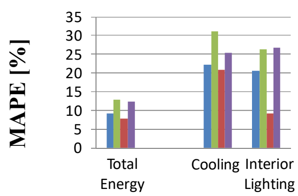
Group 2- Tests to group the locations by hemispheres

Action 1 – The global assessment of the network (Figure 3c) has shown that the use of cities from both hemispheres has positively affected the network performance. The same result was found in the "never seen" set (Figure 3b) in all network output parameters. The Heating parameter was the exception, which was benefited in both sets - training and "never seen" - when the hemispheres were separated. On the other hand, the training set errors (Figure 3a) were smaller when only the Southern hemisphere was taken into consideration.

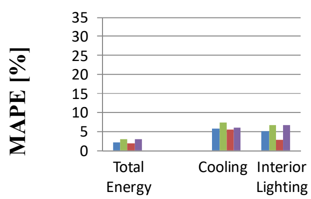
Figure 3 - Comparison between ANNs in cities from Both Hemispheres and from the Southern Hemisphere alone



(a) Training and validation set



(b) "Never Seen" Test Set



(c) Complete data set

Legend:
- G+T+S_B.H._HDD_CDD_BR
- G+T+S_S.H._HDD_CDD_BR
- G+T+S_B.H._RJ/SAL_HDD_CDD_BR
- G+T+S_S.H._RJ/SAL_HDD_CDD_BR

The reduction in the training set errors (Figure 3a) occurred when the cities were separated by hemisphere. It happened because this solution homogenized the training set, thus facilitating the network convergence. On the other hand, it harmed the generalization and reduced the performance of the "never seen" set (Figure 3b). Feeding the network with as many cases as possible may improve generalization. However, very extreme values, such as those found in Vancouver, whose climatic conditions are very different from those of the other cities, can become noisy data and undermine the forecasts, as it was recorded in the Heating parameter.

Action 2 – With respect to the assessment of the input parameter limits in the training and testing sets, it was necessary adjusting the network limits by transferring the data set belonging to Salvador City (SAL) to the training set, as well as by transferring the data set belonging to Rio de Janeiro City (RJ) to the "never seen" set. It happened because Salvador, which belongs to the "never seen" set, presented Latitude and Longitude values that forced the network to extrapolate when only cities from the Southern hemisphere were assessed. In addition, Salvador presented Altitude, HDD, CDD, GSR_a and Cl_a values close to the limits.

Action 3 - Firstly, the city switching effect was compared to the solution using Both Hemispheres, because this solution has presented the smallest errors in Action 1. The city switching has improved the performance of the four output parameters in the three sets (training, "never seen" and complete data set) (Figure 3). Lighting has presented the greatest error reduction due to city switching, and it reduced the MAPE by 11%, in absolute terms, in the “never seen” set. It indicates that the Lighting prediction is more affected by the input parameters extrapolating the training limits than the other outputs.

Then, the results were analyzed in the two training set solutions - Both Hemispheres and Southern Hemisphere only - and compared to the results in Action 1, before the cities were switched. The city switching did not change the pattern of errors found in Action 1, wherein the exclusion of cities from the Northern hemisphere has resulted in improved network training and in worsened network generalization. These results show that the Northern Hemisphere cities values in the input parameters were effective in increasing the ANN training data diversity, mainly because they were very different from the Southern Hemisphere ones. Therefore, they did not affect the error pattern observed before the exchange of the Southern Hemisphere cities .

The city switching was beneficial to the solution that considered both hemispheres, as well as to the solution that considered the Southern Hemisphere only, since it reduced the network forecast errors. However, this difference was more significant to Lighting in the case of Both Hemispheres; and to Cooling, in the case of the Southern Hemisphere, since it presented approximately 45% and 20% reduction, respectively, in comparison to the previous solution (Figure 3c).

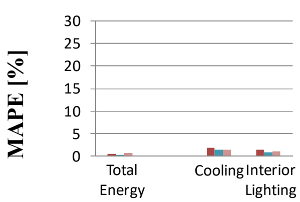
When the Southern Hemisphere was taken into consideration alone, the most affected input parameters by city switching were Latitude, Longitude and Accumulated Global Solar Radiation, since Salvador, which previously belonged to the “never seen” set, presented limiting values or values close to these limits. By confronting this observation with the MAPE reduction for Cooling, it was possible seeing that this parameter was more affected by the aforementioned input variables than the other outputs.

The adjustment in the limit values through city switching has improved the network performance, mainly in the solution that considered both hemispheres. When the Southern Hemisphere was assessed alone, the Lighting improvement was not significant, whereas there was no Heating improvement. Therefore, it was concluded that these output parameters were less affected by the input variables, whose limits have changed due the removal of cities belonging to the Northern Hemisphere.

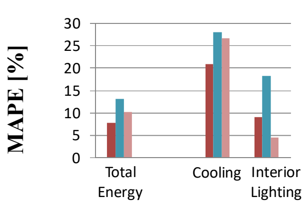
Group 3– Verifying the suitability of key variables to describe thermal climate features

It was possible seeing that the training (Figure 2a) and "never seen" sets (Figure 2b) have presented the same error trend in the input parameters. However, they have shown different magnitudes; the "never seen" set has presented significantly higher errors, and the graph scales ranged from 8% to 200%. The G + T + S cluster was the best-performance solution, and it was followed by the two-group clustering solutions (for all sets). By observing the total data set (Figure 2c), it was possible seeing that the Geographic Location (G) alone has presented better performance than G + T, wherein the MAPE was equal to 13% and 15%, respectively. The Geographic Location features were more representative than the others, since the individual Geographic Location performance (G) was better than the Thermal Climate Features (T) and Sky Conditions (S) performances. The relevance of the Sky Conditions appeared to be more representative than that of the Thermal Climate Features. It happened because, when the Thermal Climate Features (G + T) were combined with the Geographic Location (G), they presented higher MAPE than the Sky Conditions (G + S), 15% and 14%, respectively. In addition, when the Thermal Climate Features (T) were observed separately, they presented 33% MAPE, whereas the Sky Conditions (S) presented 26%.

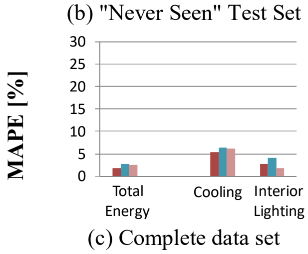
Figure 4 - Comparison between solutions with different thermal variables



(a) Training and validation set



(b) "Never Seen" Test Set



(c) Complete data set

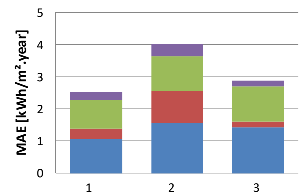
Legend:
G+T+S_B.H._RJ/SAL_HDD_CDD_BR
G+T+S_B.H._RJ/SAL_TBSV_TBSI_URV_URI_ΔTV_ΔTI_BR
G+T+S_B.H._RJ/SAL_TBSV_TBSI_URV_URI_ΔTV_ΔTI_GD

Action 2 - The number of neuron adjustment tests applied to the seven network input parameter clusters has overall worsened the network performance in all clusters, since the errors increased when the neurons were adjusted.

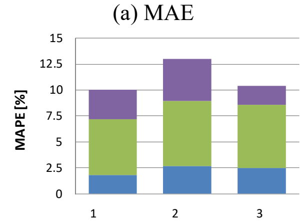
The replacement of the HDD and CDD parameters by the six new input thermal parameters did not improve the ANN performance (Figures 4 and 5). In addition, the larger number of input parameters increased the network complexity and, consequently, the execution time.

By comparing the results of the two algorithms in the new thermal parameter configuration, it was possible seeing that the Gradient Descent results were worse in the Total Energy Consumption and in the Cooling in the training set (Figure 4a), and better in all other output parameters in the "never seen" set (Figure 4b), as well as in the overall network performance (Figure 5b).

Figure 5 - Comparison between the networks and different thermal variables in the whole network set



(a) MAE



b) MAPE

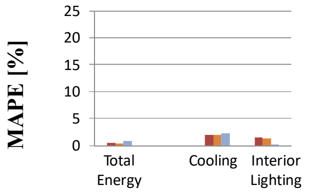
Legend:
(1) G+T+S_B.H._RJ/SAL_HDD_CDD_BR
(2) G+T+S_B.H._RJ/SAL_TBSV_TBSI_URV_URI_ΔTV_ΔTI_BR
(3) G+T+S_B.H._RJ/SAL_TBSV_TBSI_URV_URI_ΔTV_ΔTI_GD

Group 4 - Adding Cumulative Horizontal Global Illuminance as a daylight availability key variable

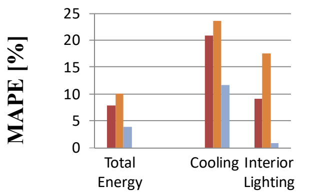
As for the new thermal parameters, the benefit from the inclusion of the Accumulated Global Horizontal Illuminance parameter was not conclusive due to conflicting results. It happened because the same original solution algorithm - the Bayesian Regularization algorithm (BR)- presented improved performance in the training set and worsened performance in the "never seen" set, as well as in the total data set (Figures 6a, b and c). On the other hand, the Gradient Descent algorithm presented lower performance in Total Energy Consumption and in Cooling in the training set (Figure 6a). However, all output parameters showed improved performance in the "never seen" set and in the total data set (Figures 6b and c).

Figures 7a and b show the overall network result, in which the addition of the Accumulated Global Horizontal Illuminance parameter increased the network errors in comparison to the same algorithm in the original configuration (BR). On the other hand, by comparing the two solutions that included the new input parameter, it was possible seeing that the Gradient Descent solution showed improved performance.

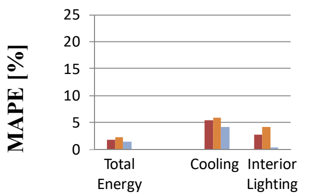
Figure 6 - Comparison between solutions and different sky condition variables



(a) Training and validation set



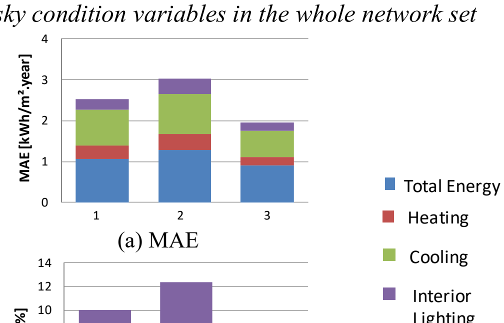
(b) "Never Seen" Test Set



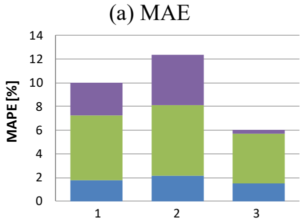
(c) Complete data set

Legend:
G+T+S_B.H._RJ/SAL_RSGa_Nba_BR
G+T+S_B.H._RJ/SAL_RSGa_Nba_IGHa_BR
G+T+S_B.H._RJ/SAL_RSGa_Nba_IGHa_GD

Figure 7 - Comparison between networks and different sky condition variables in the whole network set



(a) MAE



(b) MAPE

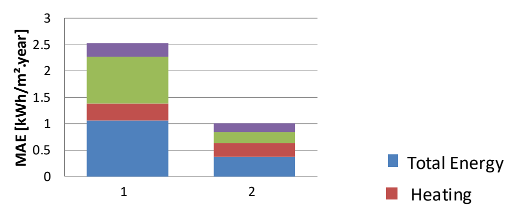
Legend:
(1) G+T+S_B.H._RJ/SAL_RSGa_Nba_BR
(2) G+T+S_B.H._RJ/SAL_RSGa_Nba_IGHa_BR
(3) G+T+S_B.H._RJ/SAL_RSGa_Nba_IGHa_GD

Group 5 - Validating the metamodels test method based on parametric energy simulations: assessment using the set-like test available in the literature

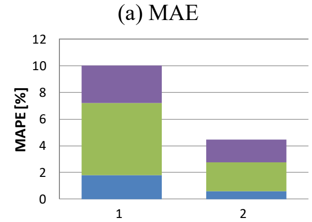
The comparison between the ANN test methods showed that the "never seen" set is essential for a rigorous approach used to test ANN as metamodel based on parametric studies, thus resulting in 50% difference in comparison to the test in the literature. Figure 8 depicts the comparison between the "never seen" set approach and the literature approach. It also shows the total data set (training + test) of errors in all the overlapping output parameters, thus resulting in 10% MAPE in the first approach and in less than 5% MAPE in the second approach.

This assessment was considered important, since the use of the literature approach alone could mask the model performance in cases never seen before, but which are within the limits of the examples presented to the network. If the 50% error difference between the two methods referred to 5% absolute MAPE value in the total data set, it was much higher in the testing set (e.g. 18% Cooling), see Figure 8. Thus, the combination of the two methods allowed saying that the network has the potential to predict results with errors lower than 5%. It results from the assessment of the network convergence (training) and generalization (literature approach test). However, there is up to 20% error limitation, depending on the never seen case and on the output parameter per final use ("never seen" set test).

Figure 8 - Comparison between ANN performances tested according to the literature method and to the "never seen" set



(a) MAE



(b) MAPE

Legend:
(1) G+T+S_B.H._RJ/SAL_NSS
(2) G+T+S_B.H._RJ/SAL_LS

Conclusion

As general conclusion, ANNs were able to predict the Total Energy Consumption in different climates by taking into consideration the cities from different hemispheres. The other outputs required more trainings if the "never seen" test set was taken into account. If just the literature test set was taken into consideration, the mean percentage error would be lower than 3%, but when the "never seen" test was taken into account, the errors for Cooling, even for the best solution, reached 20% (see Figure 3b, G+T+S_B.H_RS/SAL_HDD_CDD_BR). However, the result analysis pointed out that improvements in ANN, such as input selection, as attested in the key-variables grouping or in the suitability tests, as well as in data set (cities replacement), could enable ANN to metamodel the other outputs presenting lower errors.

With respect to the four topics addressed in the current investigation (underlined), it can be concluded that:

i. The influence of input parameter clusters from different building–context categories – the cluster categories were extremely significant, since different groups had different influences on the network performance. The Geographic Location parameters were the most important ones. The combination between different clusters has also led to different error patterns. The effects of each cluster were differently perceived by the ANNs output parameters in the "never seen" set test.

ii. The effect of clustering significantly different geographic locations, such as different hemispheres separated by the equatorial line, where the "orientation" parameter has opposite sunshine and radiation meanings – the clustering of locations from both hemispheres was, in general, more efficient than the use of locations from the Southern Hemisphere alone, since it resulted in networks with greater generalization power. However, variables presenting spurious values, such as HDD in Vancouver, may affect the network convergence regarding the output patterns most influenced by this variable, in the present case, the Heating. The effects of changing the clusters according to the hemisphere were not perceived by the output parameters in the same way.

iii. The adequacy of different types of key variables to describe the thermal parameters of climate - the replacement of the derived parameters by the primary ones in order to describe the thermal features of the building was inadequate. According to the tests performed in the current study, HDD and CDD are the thermal climate features most adequate to model the energy consumption parameters. They were selected because the studied models are conditioned. The so-called primary parameters may be adequate to other studies or ANN approaches.

iv. Describing the sky conditions in different locations – the information about accumulated radiation and cloudiness was used to describe the Sky Conditions. It was not possible stating that the addition of information about illuminance was useful to the network, since the results were divergent with respect to the model improving the assessed algorithms, sets and outputs. Therefore, the increase in the number of description parameters does not necessarily improve the model.

Acknowledgement

We would like to acknowledge the Research Funding Institutions CAPES and CNPq for providing nancial support to the present study.

References

Aqlan, F., et al. (2014) Integrating Artificial Neural Networks and Cluster Analysis to Assess Energy Efficiency of Buildings. In: Proceedings of the 2014 Industrial and Systems Engineering Research Conference. 8 p.

Brazilian Association for Technical Standards (2005) National Standard NBR 15220-3: Building thermal performance - Brazilian bioclimatic zoning and constructive guidelines for single family dwellings of social interest. (in Portuguese).

Buratti, C.; Barbanera, M.; Palladino, D. (2014) An original tool for checking energy performance and certification of buildings by means of Artificial Neural Networks. Applied Energy, v. 120, p. 125-132.

Coakley, D.; Raftery, P.; Keane, M. (2014) A review of methods to match building energy simulation models to measured data. Renewable and Sustainable Energy Reviews, v. 37, p. 123-141.

Datta, D.; Tassou, S. A. (1998) Artificial neural network based electrical load prediction for food retail stores. Applied Thermal Engineering, v. 18, n. 11, p. 1121-1128.

D.O.E -US. (2012) EnergyPlus program .v. 6. Available at https://energyplus.net/.

Ekici, B. B.; Aksoy, U. T. (2009) Prediction of building energy consumption by using artificial neural networks. Advances in Engineering Software, v. 40, n. 5, p. 356-362.

Fonseca, R. W. (2015) Daylighting harvesting for non-residential buildings: the possibilities and limitations of the artificial neural networks. (thesis) Federal University of Santa Catarina, 457 p. (in Portuguese).

Fumo, N. (2014) A review on the basics of building energy estimation. Renewable and Sustainable Energy Reviews, v. 31, p. 53-60.

Haykin, S. (2001) Artificial Neural Networks: Principles and Practices. ARTMED Publisher 900 p. (in Portuguese).

Kalogirou, S. A. (2000) Applications of artificial neural-networks for energy systems. Applied Energy, v.6 7. p. 17-35.

________. (2001) Artificial neural networks in renewable energy systems applications: a review. Renewable and Sustainable Energy Reviews, v.5, n.4, p.373-401. 2001.

Kalogirou, S. A.; et al. (2015) Artificial neural networks for the generation of a conductivity map of the ground. Renewable Energy, v.77, p.400-407.

Krarti, M. (2003) An Overview of Artificial Intelligence-Based Methods for Building Energy Systems. Journal of Solar Energy Engineering, v. 125, n. 3, Aug. 2003, p. 331-342.

Lepadatu, D.; Judele, L.; Rosu, (2014) A. Applying Artificial Neural Networks to Estimate the Energy Performance of Buildings In: 2014 Proceedings of Second Intl. Conf. on Advances In Civil, Structural and Environmental Engineering-. 5 p.

Li, K.; Su, H.; Chu, J. (2011) Forecasting building energy consumption using neural networks and hybrid neuro-fuzzy system: A comparative study. Energy and Buildings, v. 43, n. 10, p. 2893-2899.

Lu, T.; Lü, X.; Kibert, C. (2015) A hybrid numerical-neural-network model for building simulation: A case study for the simulation of unheated and uncooled indoor temperature. Energy and Buildings, v. 86, p. 723-734.

Melo, A. P.; et al. (2014) Development of surrogate models using artificial neural network for building shell energy labeling. Energy Policy, v. 69, 9 p.

Magnier, L.; Haghighat, F. (2010) Multiobjective optimization of building design using TRNSYS simulations, genetic algorithm, and artificial neural network. Building and Environment, v. 45, n. 3, p. 739-746.

Neto, A. H.; Fiorelli, F. A. S. (2008) Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption. Energy and Buildings, v. 40, n.1 2, p. 2169-2176.

Pereira, F. O. R.; et al. (2005) An Investigation about the Consideration of Daylighting along the Design Stages. In: Proceedings of 2005 Passive and Low Beirut, 2005. p. 1025-1030.

Silva, R. M. (2005) Artificial Neural Networks applied to Intrusion Detection in TCP / IP Networks. Pontifical Catholic University of Rio de Janeiro. Rio de Janeiro. 144 p. (in Portuguese).

The Mathworks INC. MatLab 2011. (2011) MATLAB Documentation. Natick: The Mathworks Inc.

Westphal, F. S. (2012) Macro Excel Visual Basic para parametrização de modelos para o EnergyPlus.

Wolstenholme, S. (2013) EasyNNplus software v.1.6.

Wong, S. L.; Wan, K. K. W.; Lam, T. N. T. (2010) Artificial neural networks for energy analysis of office buildings with daylighting. Applied Energy, v. 87, n. 2, p. 551-557.

Yezioro, A.; Dong, B.; Leite, F. (2008) An applied artificial intelligence approach towards assessing building performance simulation tools. Energy and Buildings, v. 40, n. 4, p. 612-620.

Zhao, H.-X.; Magoulès, F. (2012). A review on the prediction of building energy consumption. Renewable and Sustainable Energy Reviews, v. 16, n. 6, p. 3586-3592.

Yang, X.-S.; Asadi, S.; Geem, Z. (2014) Sustainable Building Design: A Review on Recent Metaheuristic Methods. In: (Ed.). Recent Advances in Swarm Intelligence and Evolutionary Computation: Springer International Publishing, v. 585, p.203-223.

Zhang, Y.; et al. (2015) Comparisons of inverse modeling approaches for predicting building energy performance. Building and Environment, v. 86, p. 177-190.

Table 1 –Data from the cities chosen for the training and testing of artificial neural networks.

	Geographic location			Thermal climate features		Daylight availability		Step 3 - Primary thermal climate features						Step 4 - Daylight availability
LOCATIONS	**Latitude (°)**	**Longitude (°)**	**Altitude (m)**	**[2]HDD (°C)**	**[3]CDD (°C)**	**[4]GSR_a (Wh/m^2 n)**	**[5]Cl_a (tenths n)**	**[6]DBT_S (°C)**	**[7]DBT_W (°C)**	**[8]RH_S (%)**	**[9]RH_W (%)**	**[10]ΔT_S (°C)**	**[11]ΔT_W (°C)**	**[12]I GHIa (lux n)**
[1]BBZ 1 - Curitiba	-25.5	-49.2	908	21	427	1520920	64912	20.9	13.6	83.9	83.8	8.7	9.7	28400
BBZ 2 - BH	-19.9	-44	785	0	1487	1915844	41260	24.0	18.9	73.4	68.6	7.9	12.4	14700
BBZ 3 - Fpolis	-27.7	-48.5	7	2	1135	1647171	56937	24.4	17.1	84.7	85.1	7.1	6.8	16600
BBZ 3 - SP (“never seen” test set)	-23.6	-46.6	803	1	854	1678614	58717	22.8	16.0	78.1	76.9	8.4	9.6	47600
BBZ 4 - Brasília	-15.9	-47.9	1061	0	1219	1962601	49949	22.1	19.0	78.1	63.1	9.2	14.7	42000
BBZ 5/8 - RJ	-22.9	-43.2	-3	0	2184	1843298	47841	26.6	20.8	77.0	79.1	6.4	6.4	17700
BBZ 6 - CGR	-20.5	-54.7	556	2	2236	1927772	44176	25.8	21.0	77.3	65.9	10.0	11.0	21900
BBZ 7 - Cuiabá	-15.7	-56.1	182	0	3206	1948652	50085	27.5	23.4	75.8	63.5	8.1	13.0	25100
BBZ 8 - Salvador (“never seen” test set)	-13	-38.5	51	0	2901	1926311	48568	27.2	24.3	76.7	78.6	3.8	2.9	12700
BBZ - Vancouver	49.2	-123.2	2	901	5	1229518	59161	17.0	3.2	86.6	76.2	5.0	9.2	800
BBZ - Phoenix	33.5	-112	337	28	2661	2094203	27787	35.6	13.0	36.9	29.9	11.5	12.1	15800

Source: Adapted from Weather Data (US-DOE, 2011).

[1]Brazilian Bioclimatic Zone - BBZ; [2]Heating Degree Days – HDD; [3]Cooling Degree Days - CDD, [4]Global Solar Radiation - Accumulated - GSR_a; and [5]Cloudiness - Accumulated - Cl_a; [6]Dry bulb temperature (summer) – average (DBT_S); [7]Dry bulb temperature (winter) – average (DBT_W); [8]Relative humidity (summer) – average (RH_S); [9]Relative humidity (winter) – average (RH_W); [10]Temperature variation (summer) – average (ΔT_S); and [11]Temperature variation (winter) – average (ΔT_W); [12]Global Horizontal illuminance - Accumulated (GHIa)