Building and Environment 59 (2013) 107–115

Contents lists available at SciVerse ScienceDirect

Building and Environment

journal homepage: www.elsevier.com/locate/buildenv

Neural network and polynomial approximated thermal comfort models for HVAC systems

M. Castilla [a], J.D. Álvarez [b,*], M.G. Ortega [b], M.R. Arahal [b]

[a] University of Almería Agrifood Campus of International Excellence, ceiA3, Dpto. de Lenguajes y Computación, Área de Ingeniería de Sistemas y Automática, Ctra. Sacramento s/n, 04120 La Cañada (Almería), Spain
[b] University of Sevilla, Dpto. Ingeniería de Sistemas y Automática, Escuela Técnica Superior de Ingeniería, Camino de los Descubrimientos s/n, 41092 Sevilla, Spain

ARTICLE INFO

Article history:
Received 2 May 2012
Received in revised form
28 July 2012
Accepted 14 August 2012

Keywords:
Thermal comfort
Predictive Mean Vote
HVAC system
Estimated polynomial model
Artificial neural networks

ABSTRACT

Nowadays, the majority of people carry on their daily activities inside a building. This has motivated research directed to assure several comfort conditions. Thermal comfort is usually maintained by means of HVAC (Heating, Ventilation and Air Conditioning) systems. The most widely used thermal comfort index is the PMV (Predictive Mean Vote), which is computed considering measurements of several physical variables. The classical calculation of this index is expensive in computational terms, and the involved measurement requires a relatively extensive sensor network. This work proposes the use of two approximated models for the PMV index, one is based on an artificial neural network and the other makes use of polynomial expansions, aimed at using these approximated indices within model predictive control frameworks. In this context, the advantages of using approximated models are two-fold: the computational cost of the calculation of the index is reduced, allowing its use in real-time control of HVAC systems; and the network sensor size is decreased. These advantages entail economic benefits and promote the deployment of comfort controllers in larger structures. This paper illustrates the development of the above cited approximated models and includes experimental tests that rate the accuracy and benefits of the proposed models.

1. Introduction

Due to several facts, as the climate change, energy efficiency is becoming a very important topic at the moment. Recent studies show that approximately 40% of total world energy consumption is directly related to the energy consumption in buildings, more than half being used by HVAC (Heating, Ventilation and Air Conditioning) systems [1]. The main purpose of HVAC systems is to reach thermal comfort for building users. Furthermore, in developed countries, people usually spend 80% of their life indoors [2]. Thus, it is important to reach a tradeoff between energy saving and users welfare. This balance can be provided by means of suitable control strategies, such as, predictive [3–5], based on neural networks [6,7], adaptive and based on fuzzy logic [8–10]. The impact of appropriate HVAC control derives from the fact that either lack or poor indoor comfort has a direct effect on users' productivity and on the actual energy efficiency of the buildings [2]. This leads the research towards controlling HVACs to provide comfortable environments with the minimum possible energy consumption.

A useful tool to derive appropriate controllers is the development of models [11]. In the case of HVAC systems, the controlled variables should be thermal comfort and energy efficiency. This prompts to the use of thermal comfort indexes such as PMV (Predictive Mean Vote), a variable widely used [12]. However, the computational effort required by the classical procedure to obtain such an index is high, and also the involved measurements need a comprehensive sensor network [13]. Then, the use of approximated models can, on the one hand, reduce the computational cost required to compute the index, allowing its use in real-time control of HVAC systems; and, on the other hand, decrease the size of the network sensor.

Several works have dealt with approximated thermal models. Neural Networks are used in order to calculate approximated models for the PMV index in Refs. [14,15] or to predict indoor temperature instead of the PMV index, as in Ref. [16]. Finally, a fuzzy model of indoor thermal comfort is used in Ref. [17], with the aim to control the thermal comfort through an HVAC system.

This work proposes the use of two approximated models for the PMV index. The first one is obtained by means of Artificial Neural Networks (ANN). ANN have been used in the context of energy conversion and management [18–20] and, more specifically, to

* Corresponding author. Tel.: +34 950214155; fax: +34 950215129.
E-mail addresses: jalvarez19@us.es, jhervas@ual.es (J.D. Álvarez).

0360-1323/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.buildenv.2012.08.012

compute the thermal comfort index as previously has been mentioned. The second model is based on a polynomial approximation [21]. The main objective is to compare both approximations from different points of view, and afterwards, to choose the best option, which will be integrated in a control system to maintain thermal comfort with the possible minimum energy consumption. Moreover, the paper illustrates the development of the above cited approximated models and includes experimental tests, which have been carried out inside CDdI-CIESOL-ARFRISOL [22,23] building. These tests rate the accuracy and benefits of the proposed models.

2. Scope of the research

The CDdI-CIESOL-ARFRISOL (http://www.ciesol.es) research centre on solar energy has two plants with a total surface of 1072 m^2 and is located inside the Campus of the University of Almería, South-East Spain. It is one of the buildings (CDdI) that compose a singular strategic project called ARFRISOL together with others located in the Solar Platform of Almería, in CIEMAT-Madrid, in Barredo-Asturias Foundation and in CEDER-CIEMAT-Soria. ARFRISOL is a singular strategic project of the Spanish R&D plan 2004–2011 financed by FEDER funds and by the Spanish Ministry of Science and Innovation (MICINN). This project is leaded by the Energy Efficiency Unit of CIEMAT and relies on the participation of research centres like CIEMAT, the University of Almería or the University of Oviedo, the most important Spanish building companies, and some of the Spanish solar technological companies (http://www.arfrisol.es). Therefore, this research centre has been built following a bioclimatic architecture criteria, where the more representative rooms of the building have been monitored with the installation of the necessary sensors to estimate thermal comfort, see Fig. 1.

3. Thermal comfort

In accordance with the most part of international standards (such as in Refs. [12,24]) thermal comfort can be defined as: *"That condition of mind which expresses satisfaction with the thermal environment"* [25]. Thermal comfort depends on several circumstances. However, according to different researches in this area, although climates, living conditions and cultures differ around the world, the temperature that people choose for comfort under similar conditions of clothing, physical activity, humidity and air velocity is very similar [26].

Many authors have studied the problem of representing and/or computing the thermal comfort condition in a certain environment [27–30], as a result there are some indices and models in the bibliography of this area. The most extended is the PMV (Predicted Mean Vote) index, developed by Fanger [31] during the 70s to guarantee the thermal comfort of humans [28]. The PMV index predicts the mean response (in a statistical sense) about thermal sensation of a large group of people exposed to certain thermal conditions for a long time [32]. The value of PMV index is a seven-point thermal sensation scale that is shown in Table 1. To ensure a thermal comfort situation in a certain environment, different standards recommend to maintain the PMV index at level 0 with a tolerance of ±0.5 [33].

The PMV index is defined by the six variables that appear in Table 2 namely, metabolic rate, clothing insulation, air temperature, mean radiant temperature, air velocity and air relative humidity. The acquisition of the major part of these variables is carried out using a simple methodology [34]. However, clothing insulation and human activity are variables not easily measurable except in controlled experiments. The main reason is, that, they depend of the actual situation of the users at every time. The values of both variables can be found in manuals and standards such as in Refs. [12,31]. For example, clothing insulation for a typical office are 1.0 clo and 0.5 clo for Winter and Summer, respectively, whereas a typical value used for human activity is 1.0 met. Finally, mean radiant temperature ($\bar{t}_r$), that can be defined as the uniform temperature of an imaginary enclosure in which, radiant heat transfer from the human body equals the radiant heat transfer in the actual nonuniform enclosure [26], can be estimated using different methods:

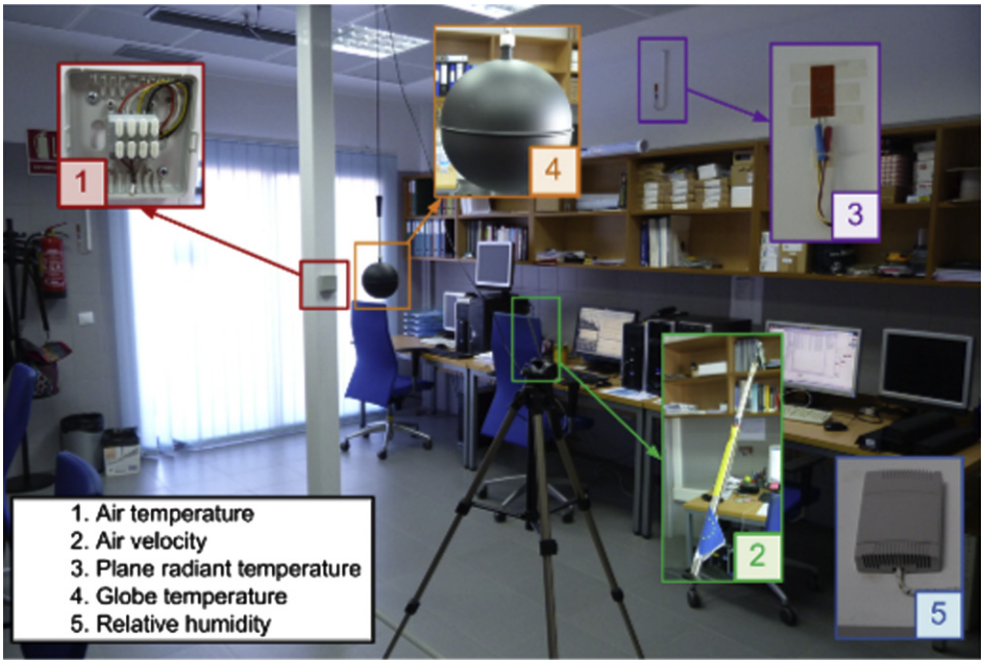

Fig. 1. Typical room inside the CDdI-CIESOL-ARFRISOL building.

Table 1
Thermal sensation scale.

PMV	Sensation
+3	Hot
+2	Warm
+1	Slightly warm
±0	Neutral
−1	Slightly cool
−2	Cool
−3	Cold

- From the plane radiant temperature in six opposite directions, weighted according to the projected area factors for a person, as Eq. (1) [24]. Using this methodology it is necessary a set of six sensors exclusively to estimate mean radiant temperature.

$$\bar{t}_r = (0.18[t_{pr}(up)+t_{pr}(down)]+0.22[t_{pr}(right)+t_{pr}(left)] +0.30[t_{pr}(front)+t_{pr}(back)])/(2[0.18+0.22+0.30]) \quad (1)$$

- From a black globe thermometer, which consists of a hollow sphere usually 150 mm of diameter, coated in flat black paint with a thermocouple build at its centre [26]. Thus, the mean radiant temperature can be calculated as a function of the globe temperature, t_g, using Eq. (2). More information about how to use the black globe thermometer and how to obtain the mean radiant temperature through the globe temperature can be found in Refs. [26,35,36]. With this strategy, only an extra sensor is necessary to estimate mean radiant temperature, so, it is a more efficient methodology, in economic terms.

$$\bar{t}_r = \left[(t_g+273)^4+\frac{1.10\cdot10^8\cdot V_a^{0.6}}{\epsilon\cdot D^{0.4}}(t_g-t_a)\right]^{1/4}-273 \quad (2)$$

where

- t_{pr}: plane radiant temperature [°C].
- $\bar{t}_r$: mean radiant temperature [°C].
- t_g: globe temperature [°C].
- V_a: air velocity [m/s].
- t_a: air temperature [°C].
- D: globe diameter [m].
- ϵ: emissivity (0.95 for black globe).

The PMV index can be estimated through the six variables that appear in Table 2 as shown in Eq. (3) [31]

$$PMV = \left(0.303e^{(-0.036\,M)}+0.028\right)L \quad (3)$$

In the previous equation, L is the thermal load in the human body (W/m^2), defined as the difference between the internal heat production and the heat loss which occurs when the person is in a thermal situation and M is the metabolic rate (W/m^2). Thermal load is estimated using Eq. (4) [31].

$$\begin{aligned} L = {} & (M-W)-0.0014\cdot M\cdot(34-t_a)-3.05\cdot10^{-3}\cdot[5733 \\ & -6.99\cdot(M-W)-p_a]-0.42\cdot(M-W-58.15) \\ & -1.72\cdot10^{-5}\cdot M\cdot(5867-p_a)-39.6\cdot10^{-9}\cdot f_{cl}\cdot\Big[\\ & \times(t_{cl}+273)^4-(\bar{t}_r+273)^4\Big]-f_{cl}\cdot h_c\cdot(t_{cl}-t_a) \end{aligned} \quad (4)$$

where

$$\begin{aligned} t_{cl} = {} & 35.7-0.028(M-W)-0.155\cdot I_{cl}\cdot\Big[39.6\cdot10^{-9}\cdot f_{cl}\cdot\Big[\\ & \times(t_{cl}+273)^4-(\bar{t}_r+273)^4\Big]+f_{cl}\cdot h_c\cdot(t_{cl}-t_a)\Big] \end{aligned} \quad (5)$$

$$h_c = \begin{cases} 2.38\cdot(t_{cl}-t_a)^{0.25}, & A>12.1\cdot\sqrt{v_a} \\ 12.1\cdot\sqrt{v_a}, & A\le 12.1\cdot\sqrt{v_a} \end{cases} \quad (6)$$

$$A = 2.38\cdot(t_{cl}-t_a)^{0.25} \quad (7)$$

$$f_{cl} = \begin{cases} 1.0+0.2\cdot I_{cl}, & I_{cl}\le 0.5\ \text{clo} \\ 1.05+0.1\cdot I_{cl}, & I_{cl}>0.5\ \text{clo} \end{cases} \quad (8)$$

In the previous equations the following variables and parameters have been used:

- W: effective mechanical power [W/m^2]. The effective mechanical power can be defined as the work realized by muscles to do a certain task.
- p_a: partial water vapour pressure in the air [Pa]. The relative humidity, y_{RH} is the ratio, in percent, between the partial water vapour pressure in the air and the saturation pressure of the water vapour at certain temperature.
- t_{cl}: clothing surface temperature [°C].
- h_c: convective heat transfer coefficient [W/(m^2 °C)].
- f_{cl}: clothing area factor [−].

Table 2
Variables which define comfort [32].

Parameter	Symbol	Range	Unit
Metabolic rate	M	0.8–4	met (W/m^2)[a]
Clothing insulation	I_{cl}	0–2	clo (m^2 °C/W)[b]
Air temperature	t_a	10–30	°C
Mean radiant temperature	$\bar{t}_r$	10–40	°C
Indoor air velocity	V_a	0–1	m/s
Air relative humidity	y_{RH}	30–70	%

[a] 1 met = 58.15 W/m^2.
[b] 1 clo = 0.155 m^2 °C/W.

4. Approximated PMV models

In this section, two approximated PMV models are presented: a model based on neural networks and a polynomial approximation. It is important to highlight that, as some PMV parameters have different values for Summer and Winter, e.g. clothing insulation, two models, one for Summer and other for Winter, must be calculated. To obtain these models some data-sets are necessary. The methodology applied to their construction is explained in Section 4.1. Whereas in Sections 4.2 and 4.3 the development to obtain the neural network and the polynomial models, respectively, is explained.

4.1. Data-sets construction

A methodology to estimate thermal comfort based on the PMV index is defined in the ISO 7730 standard [12], see Section 3. The ranges for the variables used to estimate the PMV index, i.e. the metabolic rate (M), the clothing insulation (I_{cl}), the air temperature (t_a), the relative humidity (y_{RH}), the air velocity (V_a) and the mean radiant temperature ($\bar{t}_r$), are shown in Table 2. Two of these variables, M and I_{cl}, are usually fixed to $M = 1$ met and $I_{cl} = 1$ clo or $I_{cl} = 0.5$ clo for Winter and Summer, respectively, that is, to typical values of an office environment. These values have been estimated by Fanger [31] and actually they are not considered inside the

developed approximations, since black-box models are used in this work. Therefore, the uncertainty associated with these variables can be associated with the other four variables (t_a, y_{RH}, V_a, $\bar{t}_r$). Thus, the PMV index is only depending on these four variables.

As was shown in Section 3, $\bar{t}_r$ cannot be obtained directly. On the contrary, it is estimated using either six plane radiant temperature sensors or a globe thermometer, t_g. For the current work, the last option has been considered. Even though temperature globe range is not specified in the standard, it can be calculated solving Eq. (9), which has been obtained from Eq. (2), and selecting the most logical roots, i.e. negative and irrational roots are rejected, and also positive roots out of range. Thus, it is possible to express the PMV index in terms of t_g instead of $\bar{t}_r$.

$$t_g^4 + (4\cdot 273)\cdot t_g^3 + \left(6\cdot 273^2\right)\cdot t_g^2 + \left[4\cdot 273^3 + \left(\frac{1.10\cdot 10^8\cdot V_a^{0.6}}{\epsilon\cdot D^{0.4}}\right)\right]\cdot t_g + \left[273^4 - \left(\frac{1.10\cdot 10^8\cdot V_a^{0.6}}{\epsilon\cdot D^{0.4}}\right)\cdot t_a - \left(\bar{t}_r + 273\right)^4\right] = 0 \tag{9}$$

Moreover, for these four variables (t_a, y_{RH}, V_a, t_g), subranges included into the ranges presented in Table 2 are defined depending on the place where the approximated models will be used, in this work the CDdI-CIESOL-ARFRISOL building located at the Campus of the University of Almería. Once the subranges are defined, see Table 3, two global data-sets (GS1 for Summer and GS2 for Winter) composed each one by 19,404 data points are obtained. The set of points covers the four-dimensional space of the independent variables (t_a, y_{RH}, V_a, t_g), where all possible combinations among each variable are considered. In addition, each variable changes inside a range delimited by appropriated values for the location of the research centre and with a fixed step size, see Table 3. Once these GS1 and GS2 have been defined, the PMV index can be estimated, using Eqs. (2)–(8), for all these combinations.

Following the standard methodology [37], each GSX, where $X = 1$ for Summer and $X = 2$ for Winter, has been split through random sampling without replacement in the following subsets, each one composed by 9702 data points:

- A training set (TRX) for obtaining both, the ANN parameters using a gradient descent algorithm and the polynomial model coefficients by means of a QR (orthogonal–triangular decomposition) factorization with pivoting.
- A testing set (TEX) for deciding both, the number of nodes in the hidden layer of the ANN and the order of the polynomial model.

Finally, the models are validated with real data-sets from the CDdI-CIESOL-ARFRISOL building. Each of these real data-sets corresponds to one day sampling with a sample time of $T_s = 60$ s, thus, their size is 1440 data points. Several real data-sets for each season have been chosen in order to cover the most usual cases since, as previously has been explained, different PMV models are needed for Summer and Winter.

Table 3
Ranges of variables considered to produce GS1 and GS2.

Variables	Symbol	Range	Size step	Unit
Air temperature	t_a	10–30	1.5	°C
Globe temperature	t_g	10–37.25	2	°C
Indoor air velocity	V_a	0.06–0.5	0.05	m/s
Air relative humidity	y_{RH}	30–70	4	%

For Summer, two data-sets are considered. The first one (VA1a) corresponds to a non-working day, where the real PMV index value is above the comfort band since the HVAC system was disconnected. The second data-set (VA1b) refers to a working day, where the usual occupants were inside the CDdI-CIESOL-ARFRISOL building and the HVAC system was working. Therefore, the real PMV index value for this data-set is inside the comfort band [−0.5-0.5]. In the same way, for Winter a working day and a non-working day have been selected as real validation data-sets. Moreover, the first data-set (VA2a) refers to a non-working day where the PMV index value is under the comfort band. For the second data-set (VA2b) the PMV index value is inside it.

Note that two more data-sets could be chosen for validation. One for Summer, where the real PMV is under the comfort band and another data-set for Winter, where the real PMV is above the comfort band. However, theses cases are unusual in the place where the CDdI-CIESOL-ARFRISOL building is located, Almería, and no real data have been found to be used as data-sets. A summary of the different data-sets used for training, testing and validation models is shown in Table 4.

Once trained, the goodness of fit of any of the approximations (P) (neural network or polynomial model) is reported using the root mean squared error RMS, computed over the samples of a particular set S and denoted as $e_S^{RMS}(P)$, thus:

$$e_S^{RMS}(P) = \sqrt{\frac{1}{\mathrm{card}(S)}\sum_{i=1}^{\mathrm{card}(S)}\left(y_i - \hat{y}_i(P)\right)^2} \tag{10}$$

where y_i stands for the correct value of the PMV for element $i \in S$ and $\hat{y}_i(P)$ is the approximation given by the model P for that particular element.

4.2. Neural network model for the PMV

Artificial Neural Networks (ANN) are universal approximators [38]. The most common type is a static feed-forward (FF) configuration that allows to approximate any nonlinear static mapping between input and output variables provided that certain conditions are met. These conditions, regarding the number of nodes of the ANN and the smoothness of the mapping, are often neglected, relying on a final validating step. This is due to the fact that these important factors are, in most of cases not known, leading to an iterative design procedure. Thus, ANN can be considered as black-box models where the model inputs are the number of neurons in the input layer, the model parameters are the number of neurons and the values of interconnection weights, which do not have any physical meaning, in the hidden layers and, at last, the outputs is the number of neurons in the output layer. A typical ANN scheme with two hidden layer can be observed in Fig. 2. The flexibility of ANN comes at a price, not only the number of nodes but also the selection of the weights must be decided by trial and error. In most applications, the iterative design procedure can be tackled using

Table 4
Data-sets description.

Name	Use	Origin	Size	Description
TR1	Training	GS1	9702	Summer training set
TE1	Testing	GS1	9702	Summer testing set
VA1a	Validation	Real data	1440	Summer validation set, non-working day
VA1b	Validation	Real data	1440	Summer validation set, working day
TR2	Training	GS2	9702	Winter training set
TE2	Testing	GS2	9702	Winter testing set
VA2a	Validation	Real data	1440	Winter validation set, non-working day
VA2b	Validation	Real data	1440	Winter validation set, working day

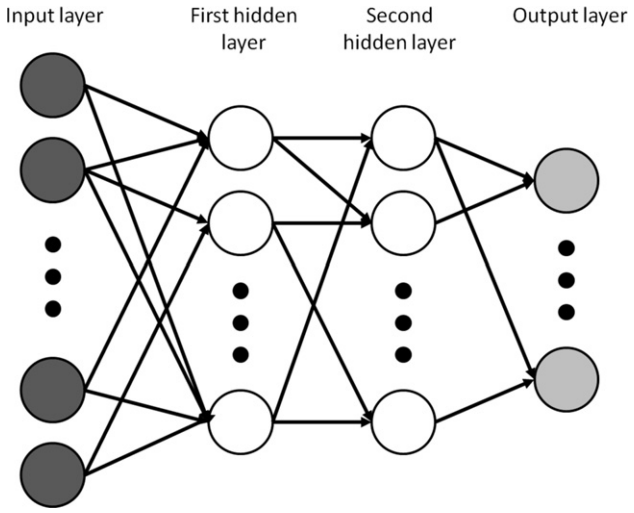

Fig. 2. Typical network scheme for a neural network with two hidden layers.

large amounts of data that are split in several sets, some of which are used for training/design and others to validate the solution.

In Ref. [14] an FF ANN is used to provide an estimation of Fanger's PMV model. The used explanatory variables are fixed and consist of air temperature, air wet bulb temperature, globe temperature, air velocity, clothing insulation and human activity. In this way, the authors avoid the need for hygrometers, which are complex and costly and provide a means for low-cost real-time control. Their training procedure is based on choosing some $2.3 \cdot 10^5$ data points covering the six-dimensional box originated by the Cartesian product of the intervals for each input variable. The intervals are of the form $[v^i_{\min}, v^i_{\max}]$ for each variable $i = 1,\cdots, 6$. The structure of the ANN is fixed a priori and consists of two hidden layers. The number of nodes per layer is also fixed to $6 \times 8 \times 4 \times 1$. The validation is done by comparing the ANN approximation with Fanger's model for an experiment in an office during a working day (08 h–17 h). This seminal work is expanded here in several directions. First, a study of the need for more data points has been performed. Then, the obtained results have been validated in a broader context by means of several experiments inside CDdI-CIESOL-ARFRISOL building. Finally, a comparison between this model and the polynomial approximation has been done.

4.2.1. Methodology

The first issue to be tackled is the ANN structure and, more particularly, the number of hidden layers. According to Ref. [39], just a hidden layer is needed to approximate any nonlinear smooth map. Using more hidden layers may reduce the number of nodes but that is irrelevant. The important issue is the network complexity, which is related to the number of parameters which, in turn, is related to the number of connections. The one-hidden layer structure allows for an easy comparison of different sizes with respect to estimated generalization error. The ANN to be used is thus formed of a hidden layer with n_h nodes with sigmoidal activation function and one output node with linear characteristic.

In the training process, input variables can be observed in Table 3. Training is performed using a variable-step gradient descent process, namely the MATLAB's implementation of the Levenberg–Marquardt algorithm. The trainlm is used for, at most, 30 iterations over the TR. The number of iterations is reduced to avoid overtraining. The usual procedure when using neural approximations is to normalize the components of the input vector. This is accomplished subtracting the mean value and dividing by the standard deviation for each variable. Once trained, the goodness of fit of a particular ANN N is reported, using the root mean squared error, see Eq. (10).

4.2.2. ANN size

An ANN with not enough nodes may be unable to reproduce the variations of PMV in the data-set. On the other hand, an ANN with more nodes than needed may lead to overtraining and degrade the generalization capabilities.

In order to select the most adequate network size several hundreds ANN are trained using data from the TR, with different random initial parameters and with different values of n_h. It is useful to show the relationship between the $e^{RMS}_{TE}(N)$ and network size n_h. Since the training of the networks relies on random elements, the error $e^{RMS}_{TE}(N)$ is a random variable, and hence, it is possible to compute its mean and standard deviation. In Fig. 3, the mean value for each n_h is plotted with solid line and (○) marks for testing sets for Summer, top, and Winter, bottom. The dashed lines above and below represent the mean plus and minus the standard deviation, respectively. For each value of n_h, 50 networks are trained and tested, the experimental mean and standard derivation are taken over these 50 trials for each n_h.

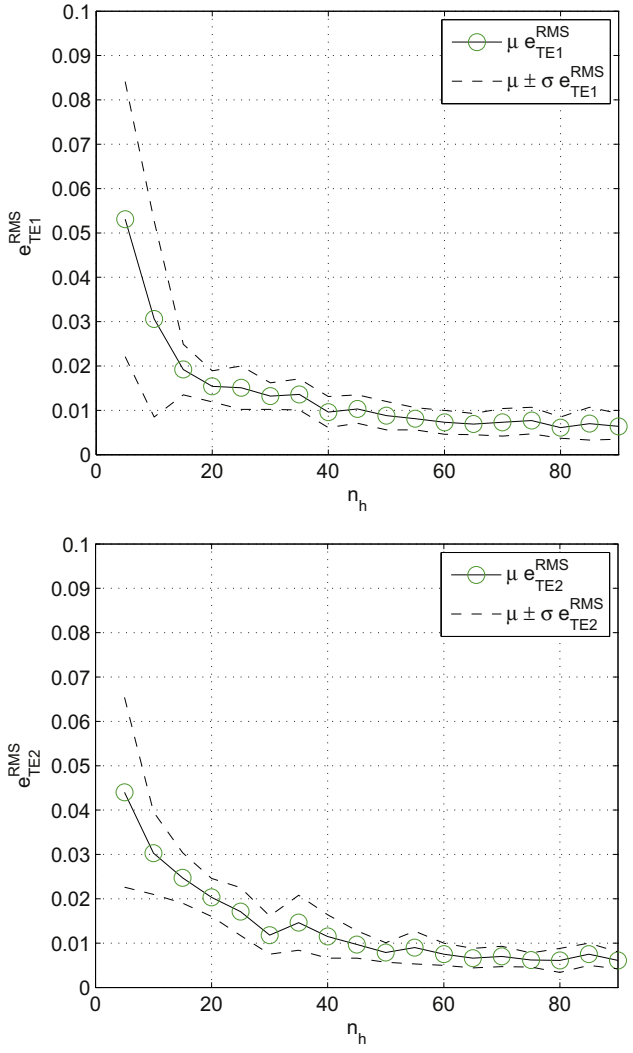

Fig. 3. Influence of network size n_h on the mean (μ) and standard error (σ) for the testing set for Summer (top) and Winter (bottom).

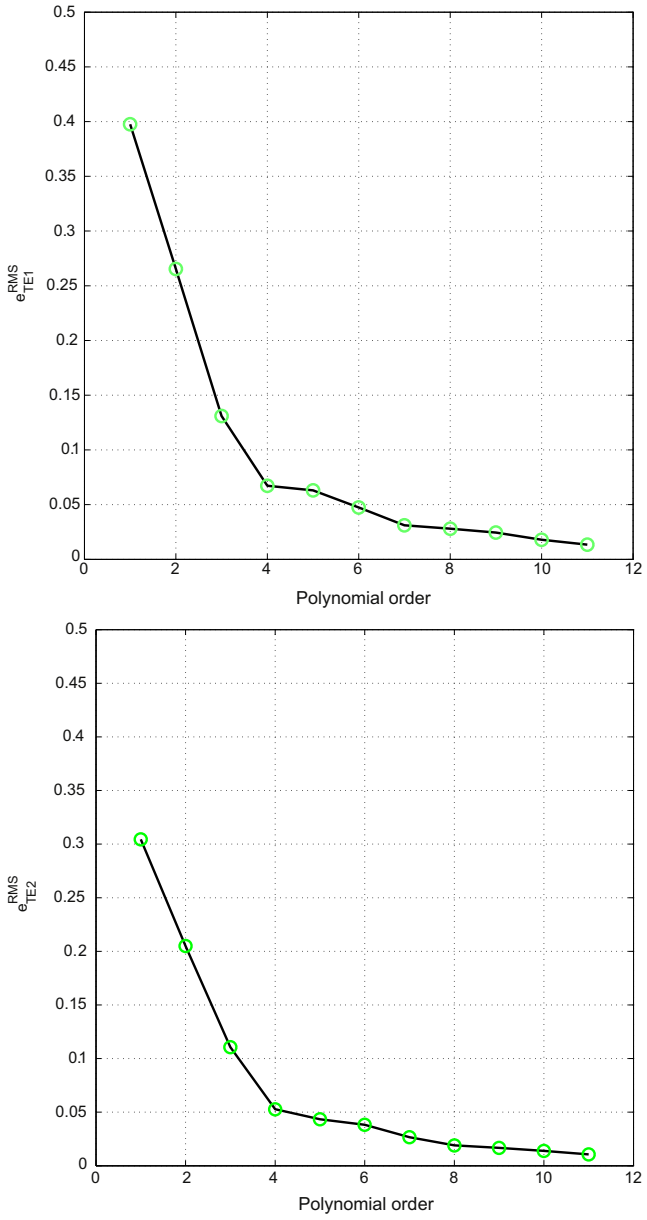

Fig. 4. Influence of polynomial degree n on the testing set errors for Summer (top) and Winter (bottom).

The mean value of $e_{TE}^{RMS}(N)$, for both Summer and Winter sets, does have a clear tendency to decrease with n_h. If one chooses $n_h = 50$, then the number of configurable parameters of the network is $n_p = 50\cdot(4+1)+50+1 = 301$ The graphs of Fig. 3 indicate that a mean RMS error about 0.01 is to be expected in both Summer and Winter cases. This would be tested later.

4.3. An estimated polynomial model for the PMV

The second model presented in this paper is based on a polynomial approximation for the PMV index. This model has been obtained using a polynomial regression modelling tool, more specifically, the MATLAB library *Polyfitn* [21]. *Polyfitn* is able to solve the coefficients of a polynomial regression model using classical linear least squares techniques. Several numerical methods have been used to implement *Polyfitn* library. However, to obtain a solution more stable is worth highlighting the use of QR factorization with pivoting for solving the system [21]. In addition, the components of the input vector are normalized before of obtaining the polynomial approximation.

As the same that in the ANN model case, two polynomial models have been obtained, one for each representative season of the year (Summer and Winter). Models that are ordinary generating functions of a four-dimensional array $a_{i,j,m,z}$, where the indexes i, j, m and z belong to set $C_l = 0,1,2,\ldots,n$ and n is the polynomial degree, are defined by Eq. (11).

$$\begin{aligned}\mathrm{PMV} &= f(t_a, y_{RH}, V_a, t_g)\\ &= \sum_{i,j,m,z\in C_l=\{0,\ldots,n\}}^{R(n+1)} a_{i,j,m,z}\cdot\left(t_a^i\cdot y_{RH}^j\cdot V_a^m\cdot t_g^z\right)\\ &= a_{7,0,0,0}\cdot t_a^7 + a_{6,1,0,0}\cdot t_a^6\cdot y_{RH} + a_{6,0,1,0}\cdot t_a^6\cdot V_a\\ &\quad + a_{6,0,0,1}\cdot t_a^6\cdot t_g + a_{6,0,0,0}\cdot t_a^6 + a_{5,2,0,0}\cdot t_a^5\cdot y_{RH}^2 + \cdots\\ &\quad + a_{0,0,0,5}\cdot t_g^5 + a_{0,0,0,4}\cdot t_g^4 + a_{0,0,0,3}\cdot t_g^3 + a_{0,0,0,2}\cdot t_g^2\\ &\quad + a_{0,0,0,1}\cdot t_g + a_{0,0,0,0}\end{aligned} \tag{11}$$

In the previous equation $a_{7,0,0,0}$, $a_{6,1,0,0}$, $a_{6,0,1,0}$,$\cdots$, $a_{0,0,0,2}$, $a_{0,0,0,1}$, $a_{0,0,0,0}$ are constant coefficients obtained from the TR1 for summer

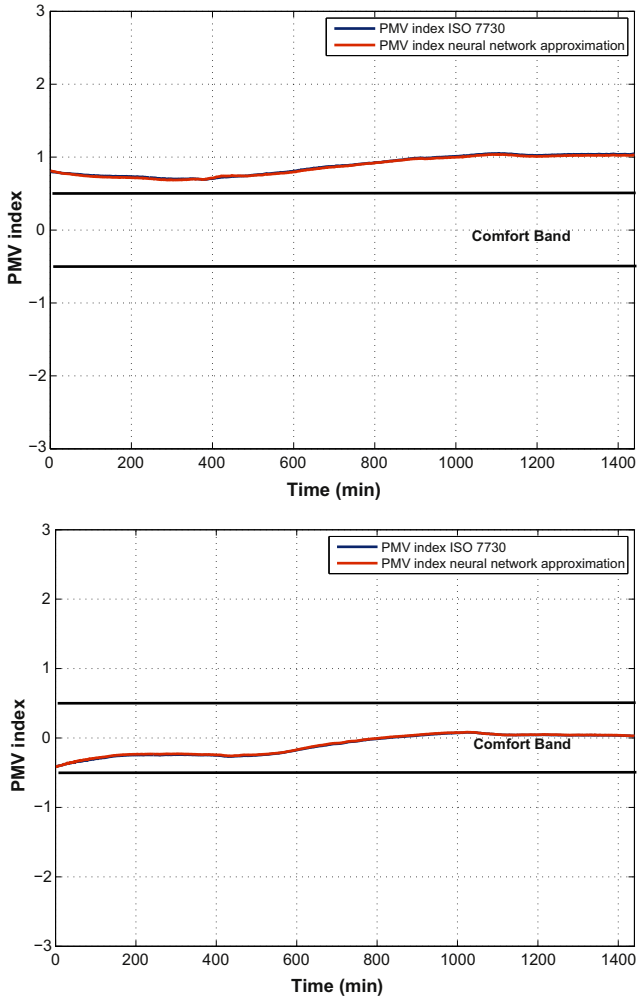

Fig. 5. Real PMV index and neural network approximation for Summer data: VA1a (top) and VA1b (bottom).

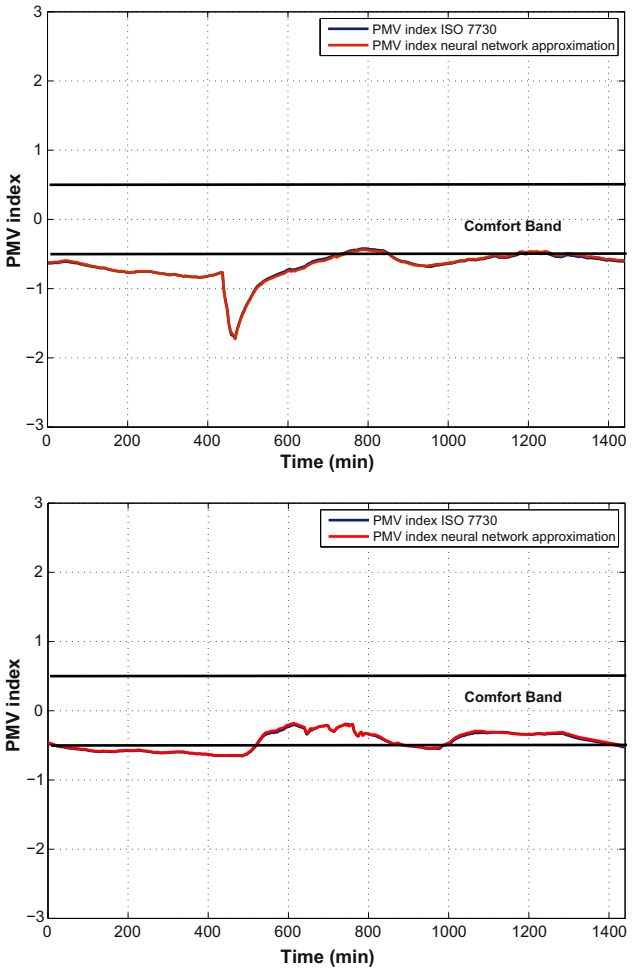

Fig. 6. Real PMV index and neural network approximation for Winter data: VA2a (top) and VA2b (bottom).

polynomial model, and TR2 for winter polynomial model. At last, $R(n+1)$ is a distribution defined by Eq. (12).

$$R(k) = \sum_{i=0}^{k} Q_k \tag{12}$$

where

$$Q_k = Q_{k-1} + \sum_{i=0}^{k} i \quad \text{and} \quad Q_0 = 0 \tag{13}$$

In this work, n has been chosen based on having similar configurable parameters, or freedom degrees, than the ANN models. Thus, for a seventh degree polynomial, i.e. $n = 7$, the number of terms is equal to 330, which is similar than the one obtained in the ANN models. However, the e_S^{RMS} index for testing sets S = TE1 (Summer) and S = TE2 (Winter) is almost tree times bigger than the one for ANN models, see Fig. 4. Thus, in the polynomial case, a mean RMS error about 0.03 is to be expected in both Summer and Winter cases. At last, as the number of coefficients is to high, 330, for lack of space, the polynomial structure is not expanded at all in Eq. (11).

5. Results and discussion

5.1. Neural network models results

As previously was commented, the goodness of fit of the polynomial models presented in Section 4.2 is characterized by the root mean square error (e_S^{RMS}), defined in Eq. (10). In Figs. 5 and 6, the real and estimated PMV is shown for the four validation data-sets. With respect to Summer results, the neural network model obtains $e_{\mathrm{VA1a}}^{\mathrm{RMS}} = 0.0117$ and $e_{\mathrm{VA1b}}^{\mathrm{RMS}} = 0.0079$, for validation data-sets VA1a, VA1b, respectively, see Fig. 5. These results are inside the expected range obtained with the testing data-set, and can be considered good enough in order to use the neural network model when computing the PMV index.

On the other hand, with respect to Winter data-sets, the neural network model obtains a root mean square error value of $e_{\mathrm{VA2a}}^{\mathrm{RMS}} = 0.0145$ and $e_{\mathrm{VA2b}}^{\mathrm{RMS}} = 0.0123$, for validation data-sets VA2a, VA2b, respectively, see Fig. 6. As for the Summer case, these results are inside the expected range and can be considered pretty good. Moreover, they probe the efficiency of the neural network model at time to calculate the PMV index with real data from both, non-working and working days.

5.2. Polynomial models results

As in the previous experiments, the polynomial models are tested with the validation data-sets VA1a, VA1b, VA2a and VA2b. The polynomial degree has been selected as $n = 7$, therefore, the number of configurable parameters is similar to the one used in the neural network model. From VA1a data-set, a value of $e_{\mathrm{VA1a}}^{\mathrm{RMS}} = 0.0065$ is obtained, this result is better than expected, see

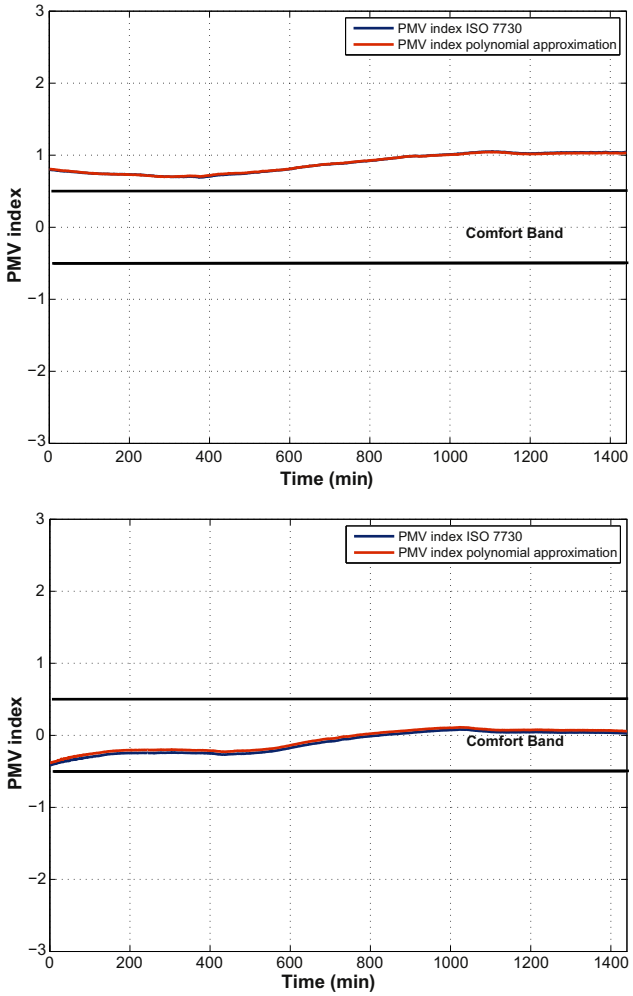

Fig. 7. Real PMV index and polynomial approximation for Summer data: VA1a (top) and VA1b (bottom).

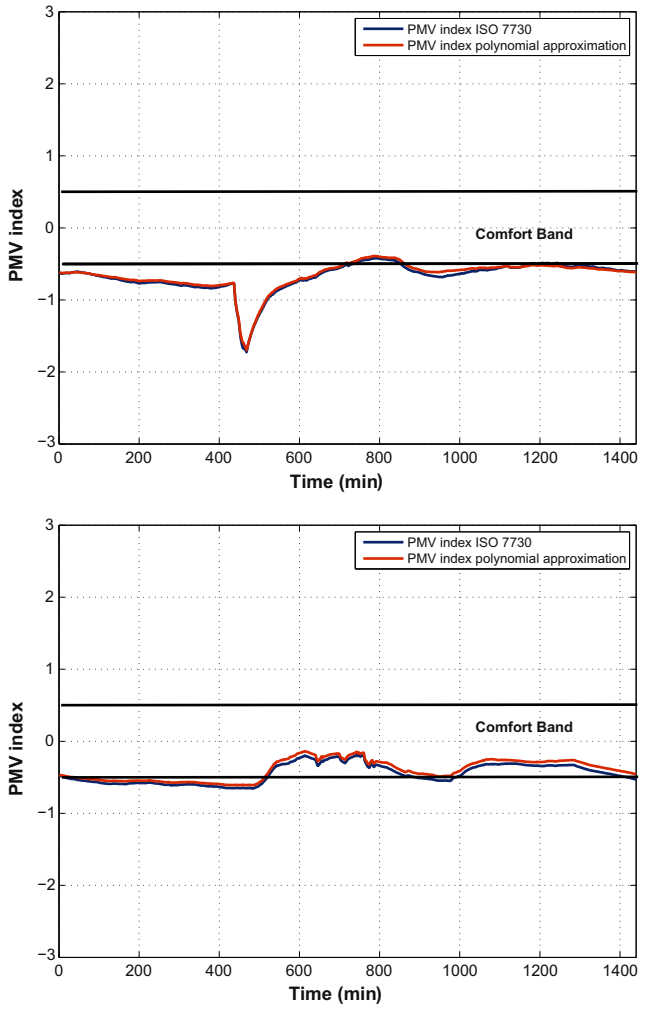

Fig. 8. Real PMV index and polynomial approximation for Winter data: VA2a (top) and VA2b (bottom).

top graph at Fig. 7, where a root mean square error about 0.03 was predicted, and even better than the one obtained by the neural network model. On the other hand, with the validation data-set for a Summer working day, see bottom graph at Fig. 7, the polynomial model obtains $e^{\mathrm{RMS}}_{\mathrm{VA1b}} = 0.0357$, which is worse than the one from the neural network model, although it is close to the expected value.

With respect to Winter case, the results obtained by the polynomial model for the non-working and working days validation data-sets are $e^{\mathrm{RMS}}_{\mathrm{VA2a}} = 0.0296$ and $e^{\mathrm{RMS}}_{\mathrm{VA2b}} = 0.0528$, respectively. The first one, i.e., the results for the non-working day data-set VA2a, are inside the expected range, see top graph at Fig. 8. Nevertheless, the results for the working day data-set VA2b, are enough worse than expected, see bottom graph at Fig. 8.

Even so, it is important to highlight that, although the polynomial model results are, in general, worse than the ones obtained from the neural network model, these results are good enough in order to use the polynomial model at time to compute the PMV index.

Table 5
Comparison for Summer models using e^{RMS}_{S} for sets $S\in$ {TR1,TE1,VA1a,VA1b}.

Model	TR1	TE1	VA1a	VA1b
ANN model with $n_{\mathrm{h}} = 50$	0.0052	0.0051	0.0117	0.0079
Polynomial model order $n = 7$	0.0182	0.0194	0.0065	0.0357

Table 6
Comparison for Winter models using e^{RMS}_{S} for sets $S\in$ {TR2,TE2,VA2a,VA2b}.

Model	TR2	TE2	VA2a	VA2b
ANN model with $n_{\mathrm{h}} = 50$	0.0175	0.0172	0.0145	0.0123
Polynomial model order $n = 7$	0.0255	0.0266	0.0296	0.0528

5.3. Comparison between models

This section is devoted to carry out a comparison between both models based on the root mean square error e^{RMS}_{S}. First, in Table 5, the error associated with the data-sets for Summer is shown. As can be observed in that table, the differences between the two Summer models are very small. However, it may be said that the quality of the results obtained with ANN model are better than the ones obtained with the polynomial model. More specifically, as previously has been commented, this model obtains an excellent result with VA1b. However, with the validation data-set for a Summer non-working day, VA1a, the results obtained, though good, are worse than the ones obtained with the polynomial approximation.

As in the previous case, for Winter case, in each model the errors are very similar for all data-sets: the training set (TR2), the testing set (TE2), the first real validation set for a non-working day (VA2a) and the second real validation set for a working day (VA2b), see Table 6.

It is important to highlight that, sometimes, the obtained results are not close to the expected ones, as in the case of the polynomial model for the validation data-set VA2b, or the neural network model for the validation data-set VA1b. But, in general, all the obtained results are close to the expected ones, i.e. a root mean square error about 0.01 and 0.03 is reached for neural network and polynomial model, respectively. As a conclusion, the neural network model is preferable to the polynomial model since it obtains better results with a similar number of configurable parameters. Even so, the polynomial model can be used as a good approximation at time to calculate the PMV index. In addition, although in general the results obtained with the ANN model are better, the polynomial model can be derived in a more easier way, hence, it can be linearized around an operation point easily, allowing the implementation of linear controllers.

6. Conclusions

The use of automatic controllers for HVAC systems benefits from the existence of appropriate models of thermal comfort. Approximate models for comfort index have the advantage of reducing computational cost, allowing its use in real-time control. Additionally, the needed sensor network has a reduced size.

In this paper, two approximated models for the PMV index have been examined. The first one is an ANN structure where different number of hidden nodes have been considered. The second one is a polynomial series truncated at power $n = 7$. A high enough number of configurable parameters has been chosen in both models in order to obtain a good accuracy at time to approximate the PMV index.

The performance of these two approximated models have been compared using different data-sets, which cover all the possibilities. From the models results it is important to highlight that: (i) the PMV can be modelled with both kind of models and (ii) as expected, there is a tradeoff between model accuracy and model complexity.

As a conclusion, with similar configurable parameters, the neural network model has better results than the ones obtained from the polynomial model. However, both models can be used at time to approximate the PMV index. As future work, these models

will be used together with predictive controllers at time to control the users' thermal comfort inside a building.

Acknowledgements

The authors are very grateful to Andalusia Regional Government (Consejería de Economía, Innovación y Ciencia), Spain, for financing this work through the Programme "Formación de personal docente e investigador predoctoral en las Universidades Andaluzas, en áreas de conocimiento deficitarias por necesidades docentes (FPDU 2009)". This is a programme co-financed by the European Union through the European Regional Development Fund (ERDF).

This work has been partially funded by the following projects: PSE-ARFRISOL PS-120000-2005-1 and DPI2010-21589-C05-04 (financed by the Spanish Ministry of Science and Innovation and EU-ERDF funds); PHB2009-0008 (financed by the Spanish Ministry of Education; CNPq-BRASIL; CAPES-DGU 220/2010). The authors would like to thank all companies and institutions included in the PSE-ARFRISOL project.

Appendix A. Nomenclature

ϵ	emissivity (0.95 for black globe)
e_S^{RMS}	root mean square error for a particular data-set S
D	globe diameter [m]
f_{cl}	clothing area factor [−]
h_{c}	convective heat transfer coefficient [W/(m^2 °C)]
I_{cl}	clothing insulation [(m^2 °C)/W]
L	thermal load in the human body [W/m^2]
M	metabolic rate [W/m^2]
N	particular artificial neural network
n	polynomial degree
n_{h}	network size
n_{p}	configurable parameters of the network
p_{a}	partial water vapour pressure in the air [Pa]
S	particular data-set
t_{a}	air temperature [°C]
t_{cl}	clothing surface temperature [°C]
t_{g}	globe temperature [°C]
t_{pr}	plane radiant temperature [°C]
$\bar{t}_{\mathrm{r}}$	mean radiant temperature [°C]
V_{a}	air velocity [m/s]
W	effective mechanical power [W/m^2]
y_{RH}	air relative humidity [%]

References

[1] Yang IH, Yeo MS, Kim KW. Application of artificial neural network to predict the optimal start time for heating system in building. Energy Convers Manag 2003;44(17):2791–809.
[2] Kolokotsa D, Tsiavos D, Stavrakakis G, Kalaitzakis K, Antonidakis E. Advanced fuzzy logic controllers design and evaluation for buildings' occupant thermal-visual comfort and indoor air quality satisfaction. Energy Build 2001;33:531–43.
[3] Donaisky E, Oliveira G, Freire Z, Mendes N. PMV-based predictive algorithms for controlling thermal comfort in building plants. In: Proceedings of the 16th IEEE international conference on control applications, Singapore; 2007, p. 182–87.
[4] Freire R, Oliveira G, Mendes N. Thermal comfort based predictive controllers for building heating systems. IFAC workshop on energy saving control in plants and buildings, Bansko, Bulgaria; 2006.
[5] Dumur D, Boucher P, Murphy K, Déqué F. Comfort control in residential housing using predictive controllers. In: Proceedings of the 1997 IEEE international conference on control applications, Hartford, Connecticut, U.S.A; 1997, p. 265–70.
[6] Congradac V, Kulic F. HVAC system optimization with CO_2 concentration control using genetic algorithms. Energy Build 2009;41:571–7.
[7] Ahmed O, Mitchel J, Klein S. Application of general regression neural networks (GRNN) in HVAC process identification and control. ASHRAE Trans 1996;102: 1147–56.
[8] Nicol JF, Humphreys MA. Adaptive thermal comfort and sustainable thermal standards for buildings. Energy Build 2002;34:563–72.
[9] Federspiel C, Asada H. User – adaptable comfort control for HVAC systems. J Dyn Syst Meas Control 1994;116:474–86.
[10] Dounis A, Santamouris MJ, Lefas C. Building visual comfort control with fuzzy reasoning. Energy Convers Manag 1993;34(1):17–28.
[11] Camacho EF, Bordons C. Model predictive control. London: Springer-Verlag Ltd.; 2004.
[12] ISO7730. Moderate thermal environments – determination of the PMV and PPD indices and specification of the conditions for thermal comfort. International Organisation for Standardisation; 1994.
[13] Braun JE. Intelligent building systems – past, present and future. In: Proceedings of the 2007 American control conference, New York, U.S.A; 2007, p. 4374–81.
[14] Atthajariyakul S, Leephakpreeda T. Neural computing thermal comfort index for HVAC systems. Energy Convers Manag 2005;46:2553–65.
[15] Liu W, Lian Z, Zhao B. A neural network evaluation model for individual thermal comfort. Energy Build 2007;39:1115–22.
[16] Ruano A, Crispim E, Conceiçao E, Lúcio M. Prediction of building's temperature using neural networks models. Energy Build 2006;38:682–94.
[17] Homoda R, Saharia K, Almuribb H, Nagi F. RLF and TS fuzzy model identification of indoor thermal comfort based on PMV-PPD. Build Environ 2012;49:141–53.
[18] Yokoyama R, Wakui T, Satake R. Prediction of energy demands using neural network with model identification by global optimization. Energy Convers Manag 2009;50:319–27.
[19] Lee YW, Chang TL. Application of NARX neural networks in thermal dynamics identification of a pulsating heat pipe. Energy Convers Manag 2009;50:1069–78.
[20] Liang J, Du R. Design of intelligent comfort control system with human learning and minimum power control strategies. Energy Convers Manag 2008;49:517–28.
[21] D'Errico J. Polyfitn. http://www.mathworks.com/matlabcentral/fileexchange/10065. [last checked 11 February 2011].
[22] Castilla M, Álvarez JD, Berenguel M, Pérez M, Guzmán JL, Rodríguez F. Comfort optimization in a solar energy research center. In: IFAC conference on control methodologies and technology for energy efficiency, Vilamoura, Portugal; 2010.
[23] Castilla M, Álvarez JD, Berenguel M, Rodríguez F, Guzmán JL, Pérez M. A comparison of thermal comfort predictive control strategies. Energy Build 2011;43:2737–46.
[24] ASHRAE55. Thermal environment conditions for human occupancy. American Society of Heating Ventilating and Air-conditioning Engineers; 1992.
[25] Fanger P. Assessment of man's thermal comfort in practice. Br J Ind Med 1973; 30:313–24.
[26] ASHRAE. ASHRAE handbook – fundamentals. Refrigerating American Society of Heating and Air-Conditioning Engineers; 2005.
[27] Hoof JV. Forty years of Fanger's model of thermal comfort: comfort for all? Indoor Air 2008;18(3):182–201.
[28] Orosa JA. Research on general thermal comfort models. Eur J Sci Res 2009; 27(2):217–27.
[29] Sherman M. A simplified model of thermal comfort. Energy Build 1985;8(1): 37–50.
[30] Wan JW, Yang K, Zhang WJ, Zhang JL. A new method of determination of indoor temperature and relative humidity with consideration of human thermal comfort. Build Environ 2009;44:411–7.
[31] Fanger P. Thermal comfort analysis and applications in environment engineering. McGraw Hill; 1972.
[32] IDAE. Reglamento de instalaciones térmicas en los edificios. Tech. Rep. Ministerio de Industria, Turismo y Comercio; 2007.
[33] Liang J, Ruxu D. Thermal comfort control based on neural network for HVAC application. In: Proceedings of the 2005 IEEE conference on control applications, Toronto, Canada; 2005, p. 819–24.
[34] Tse WL, Chan WL. A distributed sensor network for measurement of human thermal comfort feelings. Sens Actuators A Phys 2008;144:394–402.
[35] Bedford T, Warmer C. The globe thermometer in studies of heating and ventilation. J Hyg 1934;34:458–73.
[36] Vernon H. The globe thermometer. Proc Inst Heat Vent Eng 1932;39.
[37] Reed RD, Marks RJ. Neural smithing. Supervised learning in feedforward artificial neural networks. The MIT Press; 1999.
[38] Cybenko G. Approximation by superpositions of a sigmoidal function. Math Control Signals Syst 1989;2:303–14.
[39] Huang GB, Chen YQ, Babri H. Classification ability of single hidden layer feedforward neural networks. IEEE Trans Neural Netw 2000; 11(3):799–801.