WCCI 2012 IEEE World Congress on Computational Intelligence
June, 10-15, 2012 - Brisbane, Australia IJCNN

Neural Network PMV Estimation for Model-Based Predictive Control of HVAC Systems

Pedro M. Ferreira*†, Sérgio M. Silva‡, António E. Ruano‡†, Aldric T. Négrier‡ and Eusébio Z.E. Conceição‡
*Algarve Science & Technology Park, Universidade do Algarve - Pav. A5, Faro, Portugal
†Centre for Intelligent Systems, IDMEC, IST
‡Universidade do Algarve - FCT, Faro, Portugal
Email: {pfrazao, ssilva, aruano, a24232, econcei}@ualg.pt

***Abstract*—Heating, Ventilating and Air Conditioning (HVAC) systems are used to provide adequate comfort to occupants of spaces within buildings. One important aspect of comfort, the thermal sensation, is commonly assessed by computation of the Predicted Mean Vote (PMV) index. Model-based predictive control may be applied to HVAC systems in existing buildings in order to provide a desired degree of thermal comfort and simultaneously achieve significant energy savings. This control strategy may be formulated as a discrete optimisation problem and solved by means of structured search techniques. Finding the optimal solution depends on the ability of computing many PMV values in a small amount of time. As the PMV formulation involves iterative computations consuming variable time, it is crucial to have a method for fast, possibly constant execution time, computation of the PMV index. In this paper it is experimentally shown that an Artificial Neural Network (ANN) can estimate the PMV index with varying degrees of efficiency over the trade-off of accuracy versus computational speed-up.**

I. Introduction and Motivation

In European Union (EU) countries, primary energy consumption in buildings represents about 40% of the total energy consumption [1]–[3], and, with variations from country to country, half of this energy is spent for indoor climate conditioning. It is estimated that the use of efficient energy management systems in buildings can save up to 8% of the energy consumption in the entire EU [4]. Around 83% of the EU dwellings were constructed before 1990 and about 50% of them before 1970 [1]. Therefore it is of fundamental importance to control efficiently the existing HVAC systems, in order to decrease energy usage and increase compliance with the European Directive (2010/31/EU) on the energy performance of buildings [3].

A Model-Based Predictive Control (MBPC) methodology formulated with the purpose of efficiently controlling existing HVAC systems in public buildings has been one subject of research by the authors [5], [6]. The objective is the minimisation of the energy required to maintain a desired minimum comfort level for the occupants. Although the perception of comfort is related to several environmental factors such as lighting, temperature and air quality, in this work only the thermal comfort conditions aspect is addressed.

Most HVAC systems are composed of a number of external units located on the buildings roof, at least one internal unit in each independent room with the corresponding operating unit, and a PC management station to which all the units are connected, commonly via a LonWorks communication bus. This station is able to monitor and control many aspects of all the HVAC system. Those of major interest are: specifying a temperature set-point for a given room, switching the internal unit on or off, and disabling the operating unit so the occupants can not change an operating mode defined by a higher level control system. Typically, the HVAC systems are actuated manually by the occupants using the operating unit located in the room, which allows basic actions such as turning the system on or off, or setting a desired temperature set-point. The latter is commonly specified as an integer number within a restricted range, therefore defining a finite number of control alternatives. This is a characteristic that can be explored, as an approach to non-linear MBPC consists in discretising the control space into an appropriate finite set of control actions and performing a search for the optimal future control trajectory within the available set of control options. In this case the MBPC problem may be solved by means of discrete optimisation methods. Branch-and-Bound has been proposed [7] and applied successfully to this type of discrete (or discretised) non-linear MBPC problems [8]–[12], and therefore have been used in this context by the authors [5], [6].

A. Motivation

Figure 1 provides an illustration of the research experimental facilities available for the work summarised above. As the discrete MBPC problem is solved using the branch-and-bound search technique, finding the optimal solution depends on the ability of computing many PMV values in a small amount of time. As the PMV formulation involves iterative computations consuming variable time, it is crucial to have a method for fast, possibly constant execution time, computation of the PMV index. Also, one of the goals of the project is the development of a monolithic plug-in controller which will fuse the functionalities of the *control systems laboratory PC station*, the *wireless sensor network central node*, and part of the *HVAC management PC station* as depicted in Fig. 1. This should be a small, low cost device, with just enough computational capability to execute the required algorithms within a fixed small amount of time, therefore the PMV routine should be implemented by a fast constant execution time method.

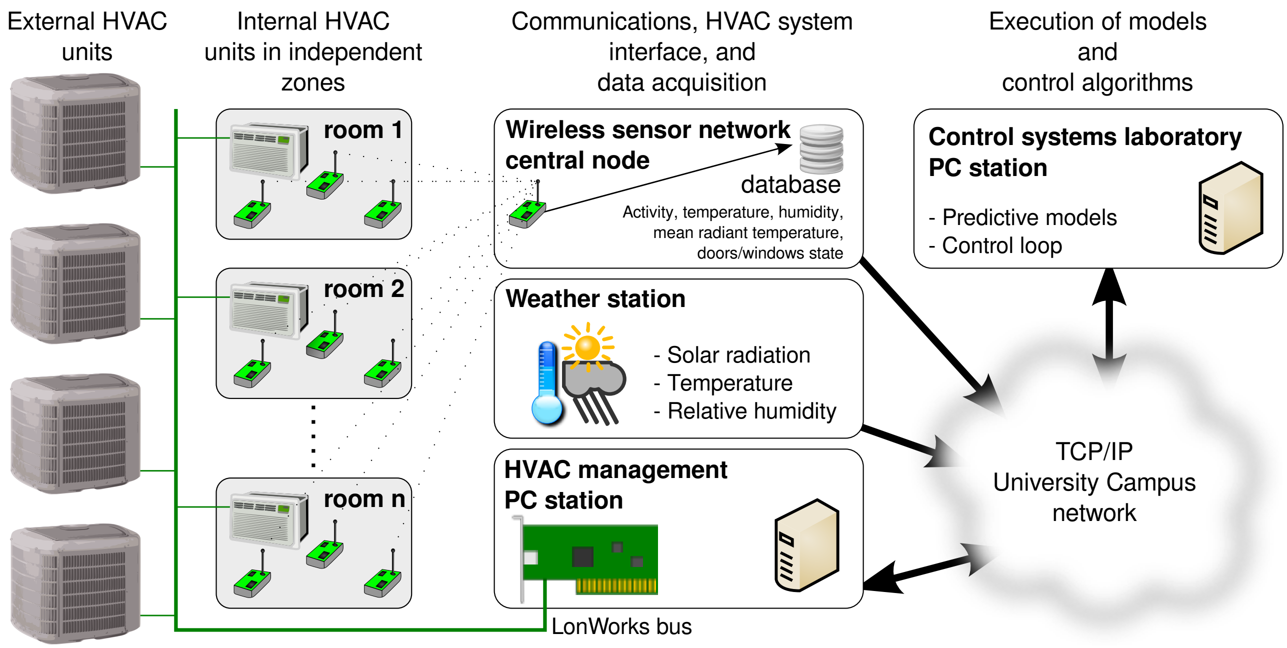

Fig. 1. Research facilities for HVAC control

Feed-forward ANNs are direct input-to-output connectionist computing structures capable of approximating a smooth function with arbitrary accuracy provided sufficient neurons are used. These features are the means to achieve the requirements stated above. The feed-forward ANN direct input-to-output structure provides the constant execution time, their ability to approximate non-linear functions provide the capability of approximating the PMV function. The accuracy of the approximation is related to the number of neurons used in the ANN hidden layer(s), which in turn is linearly related to the execution time. Consequently the problem consists in finding the appropriate trade-off between the PMV index approximation accuracy and the estimation execution time.

The application of an ANN to estimate the PMV index function has been studied before [13]–[15]. In all cases the approach taken was not the best for application in real-time control applications although this was the main motivation. In this work it is shown that by following a simpler and more appropriate approach for real-time control, it is possible to select a desired compromise between accuracy and execution time. More importantly it is shown that, when compared to the results in [13]–[15], an increased accuracy with shorter execution time may be achieved.

II. The PMV Index

The American Society of Heating Refrigerating and Air Conditioning Engineers (ASHRAE) proposed the thermal sensation scale with the purpose of quantifying the thermal sensation of people [16]. It uses a numerical coding, as described in table I, to express the qualitative thermal sensation. An index, designated PMV, was proposed by Fanger [17] in order to predict the average vote of a large group of persons on the thermal sensation scale. The index depends on six factors: metabolic rate, clothing insulation, air temperature and humidity, air velocity, and the mean radiant temperature. It is computed by means of a heat-balance equation [18], [19] given by,

TABLE I
The ASHRAE thermal sensation scale

cold	cool	slightly cool	neutral	slightly warm	warm	hot
-3	-2	-1	0	1	2	3

$$\begin{aligned} PMV = &\left(0.303e^{-0.036M} + 0.028\right)\left[(M-W) - \right.\\ &- 3.05\times10^{-3}\left[5733 - 6.99\left(M-W\right) - P_a\right] - \\ &- 0.42\left[(M-W) - 58.15\right] - \\ &- 1.7\times10^{-5}M\left(5867 - P_a\right) - \\ &- 0.0014M\left(34 - t_a\right) - \\ &- 3.96\times10^{-8} f_{cl}\left[\left(t_{cl}+273\right)^4 - \left(\bar{t}_r + 273\right)^4\right] - \\ &\left. - f_{cl}h_c\left(t_{cl} - t_a\right)\right] \ , \end{aligned} \tag{1}$$

where M and W are the metabolic rate and external work, both in W/m^2, P_a is the partial water vapour pressure in Pascal, and t_a and $\bar{t}_r$ are the air temperature and mean radiant temperature, in degrees Celsius. The surface temperature of clothing, t_{cl}, and the convective heat transfer coefficient, h_c, are given by,

$$\begin{aligned} t_{cl} = &35.7 - 0.028\left(M-W\right) - I_{cl}\left[3.96\times10^{-8} f_{cl}\times\right.\\ &\times\left[\left(t_{cl}+273\right)^4 - \left(\bar{t}_r+273\right)^4\right] + \\ &\left. + f_{cl}h_c\left(t_{cl}-t_a\right)\right] \ , \end{aligned} \tag{2}$$

and

$$h_c = \begin{cases} h_c^* & \text{if } h_c^* > 12.1\sqrt{V_a} \\ 12.1\sqrt{V_a} & \text{if } h_c^* < 12.1\sqrt{V_a} \end{cases} \tag{3}$$
$$\left(h_c^* = 2.38\left(t_{cl} - t_a\right)^{1/4}\right) \ ,$$

respectively. V_a is the air velocity in m/s and I_{cl} is the clothing thermal resistance in $m^2 \cdot {}^oC/W$. These two equations are solved iteratively until a prescribed degree of convergence is attained or a maximum number of iterations is reached. Finally, in (1) and (2), f_{cl}, which is the ratio of body surface area covered by clothes to the naked surface area, is defined by:

$$f_{cl} = \begin{cases} 1.00 + 1.290 I_{cl} & \text{if } I_{cl} \leq 0.078 \\ 1.05 + 0.645 I_{cl} & \text{if } I_{cl} > 0.078 \end{cases} . \tag{4}$$

The mean radiant temperature, $\bar{t}_r$, is a quantity which is hard to measure. The instrument most commonly employed in its determination is a black globe thermometer [20]. It consists of a black painted sphere with a temperature sensing device at its centre. Denoting the globe temperature by t_g, the mean radiant temperature may be determined as [21],

$$\bar{t}_r = \left[\left(t_g + 273\right)^4 + \right.$$
$$\left. + \ \frac{1.10 \times 10^8 V_a^{0.6}}{\epsilon D^{0.4}}\left(t_g - t_a\right)\right]^{1/4} - 273\,, \tag{5}$$

where D and ϵ are the globe diameter in meters and the globe emissivity coefficient, respectively.

P_a, the water vapour pressure in Pascal, is easily related to the relative humidity of the air, h_a, by means of Antoine's equation [22]:

$$P_a = 10 h_a \mathrm{e}^{(16.6536 - 4030.183/(T_a + 235))} \tag{6}$$

By means of (1) to (6), the PMV appears conceptually as a function of six variables that can be measured or estimated:

$$PMV = f\left(t_a, t_g, h_a, V_a, I_{cl}, M\right) \tag{7}$$

Reference values for M and I_{cl} may be found in many handbooks related to HVAC systems and are also provided in the references presented in this section, which the interested reader should consult for an in-depth discussion of the PMV index model presented in (1) to (7).

III. Previous Work

Previous ANN models of the PMV index found in the literature employ Multi-Layer Perceptron (MLP) networks [13], [15] or Least-Squares Support Vector Machines (LSSVM) [14]. All cases share the input-output modelling approach presented in Fig. 2, except for [14] where, apparently, M is constant with a value of 58.2 W/m^2 (seated, relaxed condition), therefore not considered at the model input. Additionally, in [13] h_a is determined in relation to the wet-bulb temperature, t_{wb}, and in [14] $\bar{t}_r$ is made equal to t_a.

On what concerns the hidden layer activation functions, a sigmoid was used in [13], and a Gaussian kernel was considered in [14]. For [15] only a maximum error performance was available in addition to the information regarding the type of ANN and input-output structure employed.

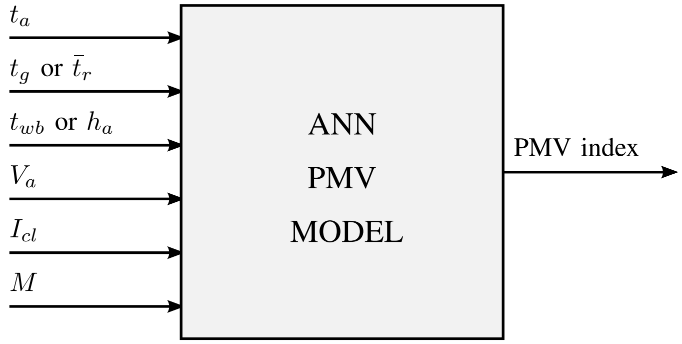

Fig. 2. Common PMV ANN modelling approach.

In Both [13] and [14], the training data sets were prepared firstly by specifying an operational range for each input variable, secondly by generating model input patterns with values from within these ranges. These are presented in Table II where the clothing insulation is shown in *clo*[1] (clothing insulation units). In [14] the values were taken at constant step sizes of 0.1, 7.5, 0.1, and 0.25, for t_a, h_a, V_a, and I_{cl}, respectively. This feature is not reported in [13], presumably as the data was selected randomly from within the ranges considered. Using these techniques 23040 data patterns were generated in [13] and 6035 in [14], for training the neural models. Regarding the last, the ranges and step sizes combine to 31570 points, therefore some subset selection was performed.

The results presented in [13] are the sum of the square of the errors obtained with training data (0.11) and a plot showing the *true* PMV line and the one obtained with the NN over 9 hours of consecutive data sampled at a rate of 10 minutes. The visible difference in the plot, constructed with 54 points, seems to account for more than 0.11 which does not favour the generalisation ability of the model. The results were obtained with a network having 6 inputs, two hidden layers with 8 and 4 neurons, and an output neuron. Collectively the model has 97 parameters to train.

In [14] the results reported are for 100 test points, different from the points in the training set, which cover a range of the PMV index from approximately -1.6 to 1.25. Clearly this range is not near the boundaries defining the thermal sensation scale, and is a reflection of the range selected for t_a. Within this test set the authors report a maximum absolute error of 0.0097 and a mean absolute error of 0.0022.

Considering the work in [15] the only data available was that the larger error was below 5%.

IV. Proposed Approach for Real-Time Control

For real-time control applications there are two important features that any PMV index approximation method should efficiently balance: accuracy and computing time. This means

[1] 1 *clo* = 0.155 $m^2 \times {}^oC/W$

TABLE II
INPUT RANGES ADOPTED IN PREVIOUS WORK MODELS [13], [14].

Ref.	t_a	t_g or $\bar{t}_r$	h_a or t_{wb}	V_a	I_{cl}	M
[13]	[16, 34]	$t_g = [14,\ 36]$	$t_{wb} = [8,\ 31]$	[0.1, 1.0]	[0.5, 1.0]	[58.2, 93.0]
[14]	[20, 28]	$\bar{t}_r = t_a$	$h_a = [20,\ 95]$	[0.1, 0.5]	[0.0, 1.5]	= 58.2
this work	[16, 32]	$t_g = [13,\ 35]$	$h_a = [20,\ 70]$	×	×	×

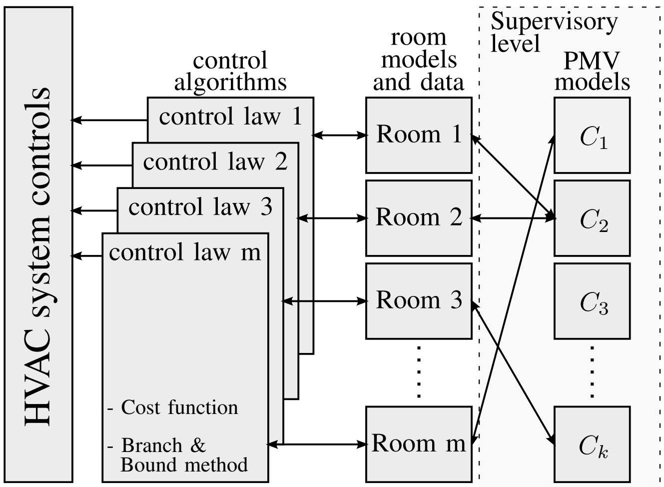

Fig. 3. Using a set of PMV models in an HVAC control system.

that the PMV models should simultaneously be as simple and as accurate as possible.

For most (if not all) HVAC real-time control applications, the environment is controlled in closed spaces where all occupants are assumed (and possibly advised) to be dressed similarly regarding the type of clothing they wear. Moreover it is likely that within each type of closed space they will be performing similar activities like attending a lecture, sitting writing a research paper, or having breakfast at the cafeteria. These two assumptions mean that for a given space it is possible to specify the values of the clothing insulation, I_{cl}, and the metabolic rate, M, therefore these may be removed from the PMV model input. If it is further assumed that the air velocity, V_a, varies little within the space and its value is determined by measurements, V_a may be considered constant and may also be removed from the PMV model input.

By defining a context vector $C = \{I_{cl},\, M,\, V_a\}$ and by using (1) to (6), a set of input-output data pairs may be generated in order to train an ANN model for the PMV index in the context C. This approach suggests that within an HVAC control system like the one in Fig 1, there are a set of distinct PMV models, PMV_i, each for a distinct context C_i. In this scheme, a supervisory system at an upper operational level will define the correspondence between each room controlled by HVAC systems and a pair $\{PMV_i, C_i\}$. The approach is illustrated in Fig. 3. The matching between the rooms and the assumed contexts can be done on the basis of the year season, of the purpose of the room, and maybe of any strong deviations of the outside weather from what is expected.

The consequence of using multiple PMV index models, is that for a specific context C_i, the model has increased accuracy and reduced computing time when compared to models considering multiple contexts. The increase in accuracy comes from the fact that there are less features to learn in the training data, whereas the decrease in computing time results from using fewer inputs and from the necessity to employ fewer neurons in the hidden layers in order to achieve a desired accuracy.

The models presented in previous work found in the literature are all multiple context models [13]–[15]. Although the motivation was the use in real-time control systems and the decrease in PMV index computing time, their input-output structure did not optimise properly the balance between accuracy and computing time for that purpose.

V. MATERIALS AND METHODS

In this work the models were obtained by means of Radial Basis Function (RBF) ANNs, a type of ANN commonly used as a function approximator. This section presents the experimental facilities, the RBF methodology employed, and a description of the modelling experiments.

A. Data Acquisition

The experiments were conducted in a classroom of the Faculty of Sciences and Technology of the University of Algarve, in the south of Portugal. This is one of 17 rooms equipped with HVAC systems and Wireless Sensor Networks (WSNs) that are being used in research for the efficient use of energy in public buildings. The WSNs are used to detect motion and the state of doors and windows (open or closed), and to measure the values of air temperature, globe temperature and relative humidity. All the measurements are done at 1 minute interval and stored in databases.

Temperature and relative humidity measurements where made using SHT11 sensors from Sensirion. The globe temperature was measured by one such sensor placed in the centre of a thin plastic sphere with a diameter of 125 mm, painted in matte black. Plastic was used instead of copper, which is more commonly found in this application, because it was easily available, but also because according to [21] it overcomes an undesirable high time constant that appears when copper is employed.

The air velocity in the room considered was measured by a BABUC probe in a 6×6 grid with a spacing of one meter, at a height of 1.2 meters above ground, and during a period of 3 minutes. Figure 4 illustrates the measurements in terms of maximum air velocity obtained. From these results an average

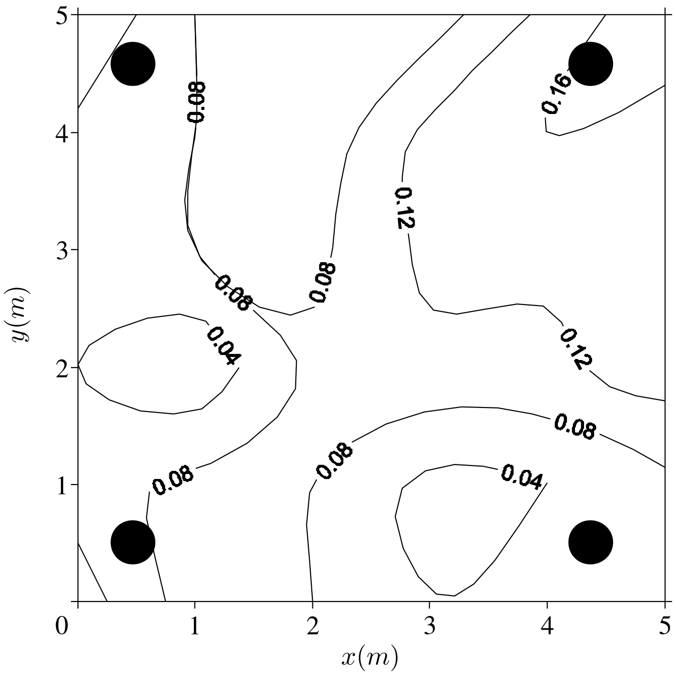

Fig. 4. Maximum air velocity (m/s) isolines in the classroom. Measurements were made at every (x, y) point. Black dots represent location of air ducts.

value of approximately 0.08 m/s was selected as a constant value to be used in the context vector of PMV index models. Regarding the two additional human factors in the context vector, a value of 69.78 W/m^2 was selected as the metabolic rate of a sedentary activity (see appendix A of [16]), and for the clothing insulation a value of 0.85 was used (see appendix B of [16]).

B. Radial Basis Function Neural Network

The RBF ANN was used in this work as a function approximator to the PMV index (1). The RBF models are trained using the Levenberg-Marquardt (LM) algorithm [23], [24] minimising a modified training criterion [25], [26].

RBF ANNs have the form,

$$\hat{y}\left(\mathbf{x}, \mathbf{w}, \mathbf{C}, \boldsymbol{\sigma}\right) = \sum_{i=0}^{n} w_i \varphi_i\left(\mathbf{x}, \mathbf{c}_i, \sigma_i\right), \tag{8}$$

where φ_i is the Gaussian function,

$$\varphi_i\left(\mathbf{x}, \mathbf{c}_i, \sigma_i\right) = e^{-\frac{1}{2\sigma_i^2}\|\mathbf{x}-\mathbf{c}_i\|^2},\ \varphi_0 = 1\,. \tag{9}$$

For a specified number of neurons, n, and for a determined set of inputs, $\mathbf{X}_t$, off-line training a RBF NN corresponds to determining the values of $\mathbf{w}$, $\mathbf{C}$, and $\boldsymbol{\sigma}$ such that (10) is minimised:

$$\Phi\left(\mathbf{X}_t, \mathbf{w}, \mathbf{C}, \boldsymbol{\sigma}\right) = \frac{1}{2}\left\|\mathbf{y} - \hat{\mathbf{y}}\left(\mathbf{X}_t, \mathbf{w}, \mathbf{C}, \boldsymbol{\sigma}\right)\right\|^2\,. \tag{10}$$

Please note, that in contrast with (8) and (9), (10) is now applied to a set of training input patterns, $\mathbf{X}_t$, and not to a single input pattern, $\mathbf{x}$. As the model output is a linear combination of the neuron activation functions output (8), (10) can be given as,

$$\Phi\left(\mathbf{X}_t, \mathbf{w}, \mathbf{C}, \boldsymbol{\sigma}\right) = \frac{1}{2}\left\|\mathbf{y} - \boldsymbol{\phi}\left(\mathbf{X}_t, \mathbf{C}, \boldsymbol{\sigma}\right)\mathbf{w}\right\|^2\,, \tag{11}$$

where omitting the dependence of $\boldsymbol{\varphi}$ on $\mathbf{C}$ and $\boldsymbol{\sigma}$,

$$\boldsymbol{\phi}\left(\mathbf{X}_t, \mathbf{C}, \boldsymbol{\sigma}\right) = \left[\boldsymbol{\varphi}\left(\mathbf{x}\left(1\right)\right)\ \boldsymbol{\varphi}\left(\mathbf{x}\left(2\right)\right) \cdots \boldsymbol{\varphi}\left(\mathbf{x}\left(N\right)\right)\right]^T.$$

By computing the global optimum value ($\mathbf{w}^*$) of the linear parameters $\mathbf{w}$, with respect to the nonlinear parameters $\mathbf{C}$ and $\boldsymbol{\sigma}$, as a least-squares solution,

$$\mathbf{w}^* = \boldsymbol{\phi}^+\left(\mathbf{X}_t, \mathbf{C}, \boldsymbol{\sigma}\right)\mathbf{y}\,, \tag{12}$$

where "+" denotes a pseudo-inverse operation, and replacing (12) in (11), the training criterion to determine the nonlinear parameters $\mathbf{C}$ and $\boldsymbol{\sigma}$ is:

$$\Psi\left(\mathbf{x}_{t,\mathbf{C},\boldsymbol{\sigma}}\right) = \frac{1}{2}\left\|\mathbf{y} - \boldsymbol{\phi}\left(\mathbf{X}_t, \mathbf{C}, \boldsymbol{\sigma}\right)\boldsymbol{\phi}^+\left(\mathbf{X}_t, \mathbf{C}, \boldsymbol{\sigma}\right)\mathbf{y}\right\|^2. \tag{13}$$

This criterion is independent of the linear parameters $\mathbf{w}$, and reflects the nonlinear-linear parameters structure of the RBF NN, by separating their computation, each type being determined by an adequate method.

The initial values for the neuron centre positions are randomly selected from the training data, and the spreads of the neuron activation functions are initialised using the simple rule in [27, p. 299]. The training procedure progresses iteratively using the LM algorithm minimising criterion (13), until a pre-specified number of iterations is reached. For more details about the training algorithm and the training criterion the reading of [26], [28], [29], or [30] is recommended.

C. Data Sets

Although the models were tested using data measured in a room as described in section V-A, the input vector for training the RBF ANNs, defined as $\mathbf{X}^t = [\mathbf{t}_a\ \bar{\mathbf{t}}_r\ \mathbf{h}_a]$, was built using randomly generated data. $\mathbf{t}_a$, $\bar{\mathbf{t}}_r$ and $\mathbf{h}_a$ are vectors of N values taken from the ranges presented in table II, on the row labelled as *this work*. They were constructed as follows:

1) $\mathbf{t}_a$ and $\mathbf{h}_a$ were selected randomly from a uniform distribution of values in the ranges specified;
2) For each value $t_{a\,k}$ in $\mathbf{t}_a$, a corresponding value t_{g_k} in $\mathbf{t}_g$ was generated using,

$$t_{g_k} = t_{a\,k} + \rho\left(-3.0, 3.0\right)\,,$$

where $\rho\left(a, b\right)$ is a random number from the uniform distribution in the range $[a,\, b]$. Implicitly it is assumed that $t_a - 3 < t_g < t_a + 3$;
3) $\bar{\mathbf{t}}_r$ was obtained by means of (5), considering $\mathbf{t}_a$ and $\mathbf{t}_g$ just described, $V_a = 0.08$, $D = 0.125$, and $\epsilon = 0.95$;

In order to determine the corresponding output PMV index values, the context vector was defined as described in subsection V-A, as $C = \{0.85,\, 69.78,\, 0.08\}$. Then, $\mathbf{Y}^t$ was constructed by means of (1), using each triplet X_k^t in $\mathbf{X}^t$ along with the values in C. This procedure gives rise to the training data set, $\mathbf{D}^t = \{\mathbf{X}^t, \mathbf{Y}^t\}$.

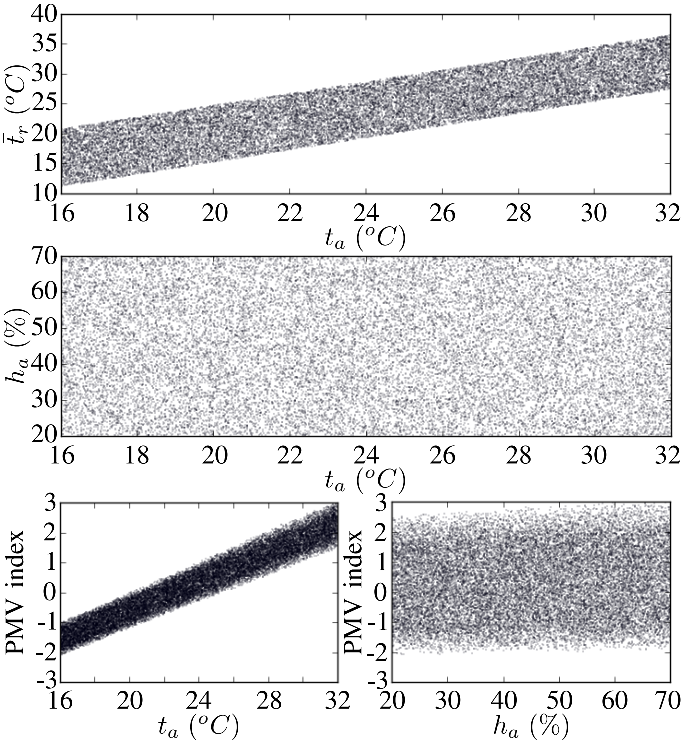

Fig. 5. Input-output variables distribution within the ranges of values considered. The plots correspond to data set $\mathbf{D}^v$.

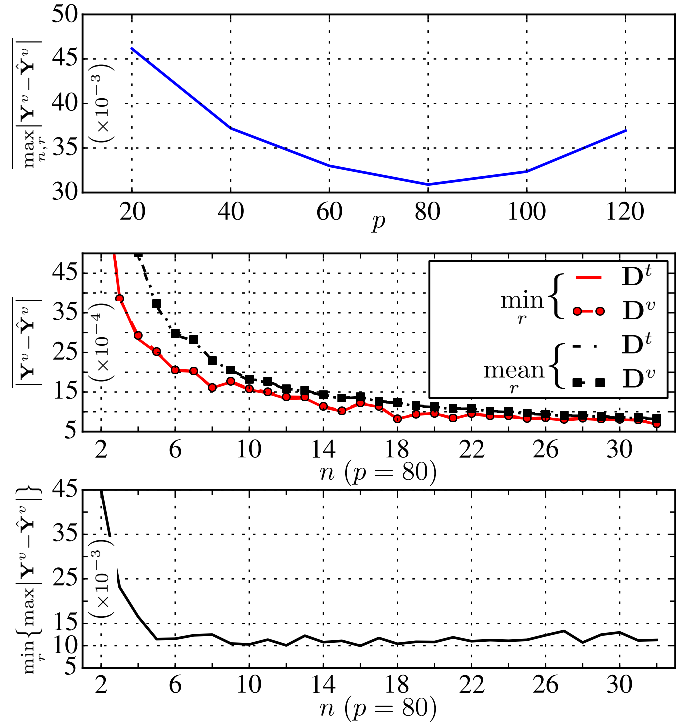

Fig. 6. Results regarding the selection of the best training set size, the generalisation capacity of the models, and the accuracy achieved.

Using the approach just described, an additional data set, $\mathbf{D}^v$, was prepared in order to validate the models with unseen data, after the training stage. $\mathbf{D}^v$ has 23100 training pairs ($N = 23100$).

Figure 5 shows how the input-output patterns are distributed within the ranges specified for the input variables, considering the data in $\mathbf{D}^v$ for the context C. It may be seen on the top plot how $\bar{t}_r$ distributes densely around the values of t_a for the whole range considered. The same is visible in the middle plot for the relative humidity and air temperature. The bottom plots, from left to right, illustrate the scattering of PMV values around the air temperature and relative humidities, respectively. This data set allows good model evaluation as it efficiently covers the variability of combinations that occurs between input variables.

D. Modelling experiments

Although the input-output structure of the model has been specified, there are still two design parameters that need to be determined: the number of neurons and the number of training patterns. For the first, an exhaustive search was conducted over the range $[2, 32]$. For the second, a search was also conducted as described in the following. It is known that the *ideal* number of training patterns is to some extent related to the number of parameters of the model being fitted [31], [32]. Considering the RBF in (8) and (9) for the PMV index model presented, each neuron accounts for 5 parameters, therefore the total number of parameters is given by $n \times 5$, n being the number of neurons employed. By specifying a number of patterns (p) per model parameter it becomes possible to determine the value of N for the training data set $\mathbf{D}^t$, as $N = n \times 5 \times p$. In this case a search was made for p in $\{20, 40, 60, 80, 100, 120\}$. For each of these values, n was varied in the range $[2, 32]$ as already mentioned. Considering that the RBF parameters are initialised randomly, 20 trials were executed for each (n, p) pair. Each trial consisted on the application of the modified training criterion LM algorithm for 200 iterations.

VI. Results and Discussion

Having determined all the models spanned by $(n, p)_r \big|_{r=1}^{20}$, where r is the trial number, the first result sought was a decision on the number of training patterns. The decision was made on the basis of the models maximum absolute error obtained on the validation data set, $\mathbf{D}^v$. For each value of p, the average of that error was computed over all the models (for all n and r). The top plot of Fig. 6 illustrates the result, where it is clearly seen that, on average, it is best to train the models with 80 training patterns for each model parameter. The middle plot illustrates the best relation between the results obtained with the training set and those obtained with the validation set, for $p = 80$. The plot presents the minimum and mean of the average absolute error obtained over the 20 trials for each number of neurons n. The difference is negligible, which allows to conclude that the models provide excellent generalisation capability. To this respect, it should be noted that the validation data set has 23100 points, a value comparable to the training set in [13], and that with only 1600 training patterns (4 neurons case), near 25% of the number

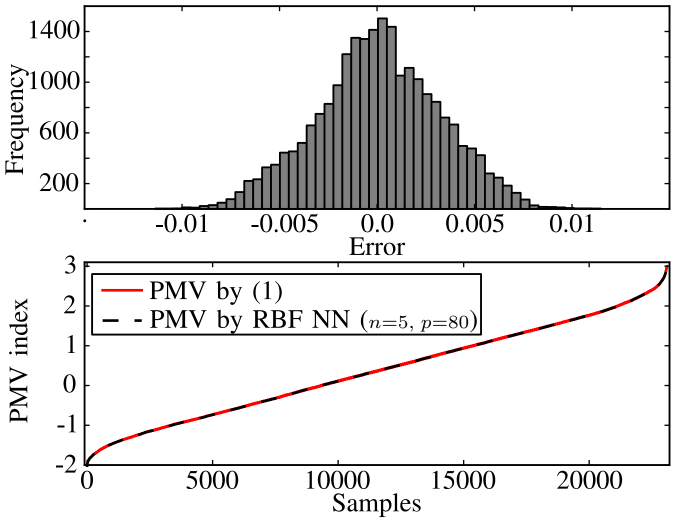

Fig. 7. PMV index fitting on validation data set, obtained with a 5 neuron RBF model. Histogram of error at the top plot.

used in [14], a maximum absolute error of approximately 0.015 is obtained both in training and validation data sets. Still for $p = 80$, as it showed the best generalisation, the lower plot in Fig 6 shows for each number of neurons, the minimum value, obtained over the 20 trials, of the maximum absolute error in the validation data set. It may be concluded that using more than 10 neurons is not necessary as no significant improvements are obtained.

The model with 5 neurons in the lower plot of Fig. 6, corresponding to 26 parameters, achieves an average and maximum absolute error of 0.0025 and 0.011, respectively. These values are comparable to the results in [14], 0.0022 and 0.0097 (obtained with 100 testing points), although the RBF model is incomparably smaller. Only two results are presented in [13]: a Sum of the Square of the Errors (SSE) on the training set of 0.11, and a figure showing how the model with 97 parameters behaved on a 9 hours experiment within a room. Regarding the first it is a bit less than half the value of 0.23 obtained by the 5 neuron model. Regarding the figure we may only comment that the well visible errors clearly show that the model does not generalise properly, probably due to a large training set and also, possibly, due to overtraining, considering the small value of SSE.

Figure 7 shows the fitting of the validation data set by the 5 neurons RBF model, as well as the histogram of the error obtained. In order to provide a more realistic evaluation of the PMV index model, it was applied to a set of data acquired in the room described previously. The data acquisition took place during a system identification experiment where pseudo-random binary sequences were being applied to the air conditioning set-point, hence there was significant variability in the room environment. The result is shown in Fig. 8. The estimates provided by the model are extremely accurate: the average and maximum absolute errors were 0.0014 and 0.0075.

A final note is due on the important trade-off between

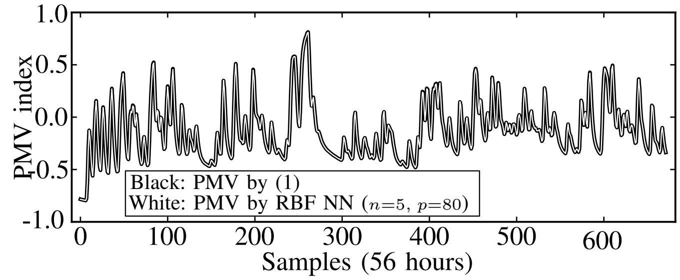

Fig. 8. PMV given by (1) (thick black line) and by 5 neuron RBF ANN (thin white line) during a system identification experiment. Please note that the lines are coincident, hence the white line is within the black line.

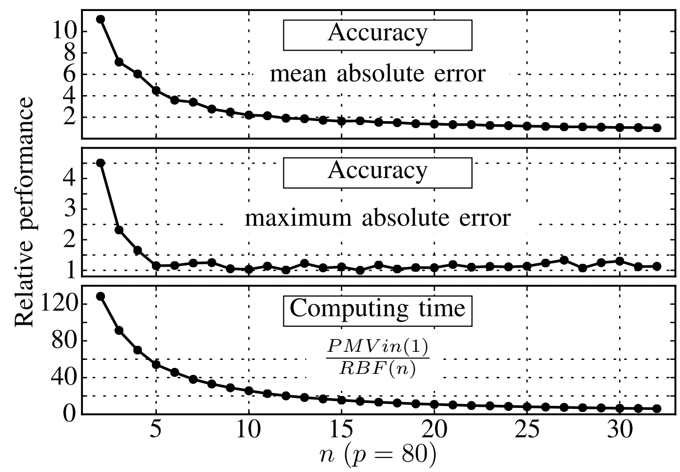

Fig. 9. Relative performance regarding computing time and accuracy.

computing time and estimation accuracy. Figure 9 highlights the relative performance of RBF models in terms of computing time and accuracy. The lower plot shows the ratio between the PMV given by (1) and the one computed by an RBF ANN with n neurons. It may be seen that a 13 neurons model is about 20 times faster than (1) and a 5 neurons model is approximately 55 times faster. The limit case of interest is 4 neurons corresponding to a speed-up near 70. The upper plot in the same figure shows the relative accuracy in terms of average absolute error. The limiting case of interest is maybe 12 neurons, about 20 times faster, with the double of the smaller average error achieved with 32 neurons. Although this plot may suggest that less than 12 neurons achieves a bad performance, the middle plot clears this impression. It shows the relative performance in terms of the maximum absolute error. It may be seen that with only 5 neurons the best performance is almost achieved, although in terms of average error the model is a bit worse than the best one. Using these curves one may select a specific model with the desired balance between accuracy and speed-up.

When compared to the models in [13] or [14], those here presented show better estimation accuracy, specifically on unseen data in real application, provide a wider coverage of the PMV input variables and of the thermal sensation scale, and achieve speed-up improvements. In this last case the gain is significant, being very large when compared to [14] and estimated to be about 3.5 times faster than the model in [13].

VII. Conclusions

A methodology has been presented in order to obtain a set of radial basis function neural network models to estimate the PMV index. As the models are necessary for real-time control, the input-output structure was simplified by separating dynamical environment variables from human factors that may be considered constant. The models show good estimation accuracy over wide ranges of the input variables and provide good coverage of the thermal sensation scale. This has been shown with data not used during the training stage, including data collected in a classroom used for HVAC control research. When compared to previous work, the design of the models was planned in detail with the purpose of using them in real-time control applications, the generalisation of the models was tested thoroughly, and a procedure was shown in order to select a model on the basis of a desired compromise between speed-up and estimation accuracy as opposed to trial and error.

Acknowledgement

The authors thank the Portuguese National Science Foundation for funding this work with project PTDC/ENR/73345/2006. The first author thanks the European Commission for the grant PERG-GA-2008-239451. This work was supported by FCT, through IDMEC, under LAETA.

References

[1] B. Poel, G. van Cruchten, and C. A. Balaras, "Energy performance assessment of existing dwellings," *Energy and Buildings*, vol. 39, pp. 393–403, 2007.

[2] C. Balaras, E. Dascalaki, A. Gaglia, K. Droutsa, and S. Kontoyiannidis, "Energy performance of european buildings," in *Proceedings of the Energy Sustainability Conference*, Long beach, USA, Jun. 2007, pp. 387–396.

[3] E. U. Parliament and E. U. Council, "On the energy performance of buildings," pp. 13–34, Jun. 2010.

[4] A. Dexter, "Intelligent buildings: fact or fiction?" *HVAC&R Research*, vol. 2, no. 2, pp. 105–106, 1996.

[5] P. M. Ferreira, S. Silva, and A. E. Ruano, "Model based predictive control of hvac systems for human thermal comfort and energy consumption minimisation," in *CESCIT'12 - 1^{st} IFAC Conference on Embedded Systems, Computational Intelligence and Telematics in Control*. Würzburg, Germany: International Federation of Automatic Control, Apr. 2012, to appear.

[6] P. M. Ferreira, S. M. Silva, and A. E. Ruano, "Energy savings in hvac systems using discrete model-based predictive control," Dec. 2012, submitted.

[7] J. M. Sousa, R. Babuška, and H. B. Verbruggen, "Fuzzy predictive control applied to an air-conditioning system," *Control Engineering Parctice*, vol. 5, no. 10, pp. 1395–1406, 1997.

[8] J. A. Roubos, S. Mollov, R. Babuška, and H. B. Verbruggen, "Fuzzy model-based predictive control using takagi-sugeno models," *International Journal of Approximate Reasoning*, vol. 22, pp. 3–30, 1999.

[9] M. Berenguel, F. Rodríguez, F. G. Acién, and J. L. García, "Model predictive control of ph in tubular photobioreactors," *Journal of Process Control*, vol. 14, pp. 377–387, 2004.

[10] L. F. Mendonça, J. M. Sousa, and J. M. G. Sá da Costa, "Optimization problems in multivariable fuzzy predictive control," *International Journal of Approximate Reasoning*, vol. 36, pp. 199–221, 2004.

[11] P. M. Ferreira and A. E. Ruano, "Discrete model based greenhouse environmental control using the branch & bound algorithm," in *Proceedings of the 17^{th} IFAC World Congress*, vol. 17, Seoul, Korea, Jul. 2008.

[12] P. M. Ferreira, "Application of computational intelligence methods to greenhouse environmental control," Ph.D. dissertation, Faculty of Sciences and Technology, University of Algarve, Faro, Portugal, Jan. 2008.

[13] S. Atthajariyakul and T. Leephakpreeda, "Neural computing thermal comfort index for hvac systems," *Energy Conversion and Management*, vol. 46, no. 15-16, pp. 2553–2565, 2005.

[14] M. Kumar and I. Kar, "Non-linear hvac computations using least square support vector machines," *Energy Conversion and Management*, no. 50, pp. 1411–1418, 2009.

[15] Y. Jian and X. Jin, "Research on the bpnn in the prediction of pmv," *Applied Mechanics and Materials*, vol. 29-32, pp. 2804–2808, 2010.

[16] ANSI and ASHRAE, "Thermal environmental conditions for human occupancy," 2004.

[17] P. O. Fanger, *Thermal comfort: analysis and applications in environmental engineering*. New York: McGraw-Hill, 1972.

[18] J. Kang, Y. Kim, H. Kim, J. Jeong, and S. Park, "Comfort sensing system for indoor environment," in *International Conference on Solid State Sensors and Actuators*, vol. 1, Jun. 1997, pp. 311–314.

[19] I. O. for Standardization (ISO), "Iso 7730: Moderate thermal environments – determination of the pmv and ppd indices and specification of the conditions for thermal comfort," Genéve, Switzerland, 1994.

[20] T. Bedford and C. Warner, "The globe thermometer in studies of heating and ventilation," *J. Hyg. (Lond.)*, vol. 34, no. 4, pp. 458–473, Dec. 1934.

[21] M. S. Owen, Ed., *2009 ASHRAE Handbook - Fundamentals*, si ed. ASHRAE, 2009.

[22] K. Cena and J. Clark, Eds., *Bioengineering, Thermal Physiology and Comfort*, ser. Studies in Environmental Science. Elsevier B.V, 1981, vol. 10.

[23] K. Levenberg, "A method for the solution of certain non-linear problems in least squares," *The Quarterly of Applied Mathematics*, no. 2, pp. 164–168, 1944.

[24] D. W. Marquardt, "An algorithm for least-squares estimation of non-linear parameters," *Journal of the Society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, Jun. 1963.

[25] A. Ruano, D. Jones, and P. Fleming, "A new formulation of the learning problem of a neural network controller," in *Proceedings of the 30^{th} IEEE Conference on Decision and Control*, vol. 1, 1991, pp. 865–866.

[26] P. M. Ferreira and A. E. Ruano, "Exploiting the separability of linear and non-linear parameters in radial basis function neural networks," in *IEEE Symposium 2000: Adaptive Systems for Signal Processing, Communications, and Control*, Canada, 2000, pp. 321–326.

[27] S. Haykin, *Neural Networks: a Comprehensive Foundation*, 2^{nd} ed. Prentice Hall, Inc., 1999.

[28] P. M. Ferreira, E. Faria, and A. E. Ruano, "Neural network models in greenhouse air temperature prediction," *Neurocomputing*, vol. 43, no. 1-4, pp. 51–75, 2002.

[29] A. E. Ruano, P. M. Ferreira, and C. Fonseca, "An overview of non-linear identification and control with neural networks," in *Intelligent Control Using Soft-Computing Methodologies*, ser. Control Series, A. E. Ruano, Ed. IEE Publishing, 2005, vol. 70, pp. 37–87.

[30] P. M. Ferreira and A. E. Ruano, "Evolutionary multiobjective neural network models identification: Evolving task-optimised models," in *New Advances in Intelligent Signal Processing*, ser. Studies in Computational Intelligence. Springer, 2011, vol. 372/2011, pp. 21–53.

[31] J. Sjöberg and L. Ljung, "Overtraining, regularization, and searching for minimum with application to neural networks," in *In Preprint IFAC Symposium on Adaptive Systems in Control and Signal Processing*, 1994, pp. 669–674.

[32] S. Amari, N. Murata, K.-R. Müller, M. Finke, and H. Yang, "Asymptotic statistical theory of overtraining and cross-validation," *IEEE Transactions on Neural Networks*, vol. 8, no. 5, pp. 985–996, Sep. 1997.