15th Triennial World Congress, Barcelona, Spain

CHOICE OF RBF MODEL STRUCTURE FOR PREDICTING GREENHOUSE INSIDE AIR TEMPERATURE

Ferreira, P.M. [*] **Ruano, A.E.** [**,*]

[*] *Universidade do Algarve, Faculdade de Ciências e Tecnologia*
Campus de Gambelas, 8000 Faro, Portugal
Email: pfrazao@ualg.pt, aruano@ualg.pt
[**] *Institute of Systems & Robotics, Portugal*

Abstract: The application of the radial basis function neural network to greenhouse inside air temperature modelling has been previously investigated by the authors. In those studies, the inside air temperature is modelled as a function of the inside relative humidity and of the outside temperature and solar radiation. Several training and learning methods were compared and the application of the Levenberg-Marquardt optimisation method was found to be the best way to determine the neural network parameters. A second-order model structure previously selected in the context of dynamic temperature models identification, was used. The model is intended to be incorporated in a real-time predictive greenhouse environmental control strategy. It is now relevant to question if the model structure used so far, selected in a different modelling framework, is the most correct in some sense. In this paper the usefulness of correlation-based model validity tests is addressed in order to answer the question mentioned above. *Copyright © 2000 IFAC*

Keywords: Neural Networks, Greenhouse Environmental Control, Model Validation, Radial Basis Functions, Temperature Prediction

1. INTRODUCTION

The main purpose of greenhouses is to improve the environmental conditions in which plants are grown. The aim of *greenhouse environmental control* (GEC) is to provide means to further improve these conditions in order to optimise the plant production process. The greenhouse climate is influenced by many factors, for example the outside weather, the actuators and the crop. Methods aimed at efficiently controlling the greenhouse climate environment must take these influences into account, and that is achieved by the use of models. Feed-forward layered neural networks (NNs) are widely applied in many fields of engineering to perform some type of non-linear data processing. In the fields of identification and modelling of non-linear systems their *universal approximator* property is exploited. In situations where the data generating function is a non-linear time-varying function, it is standard practice to train first the networks off-line, and subsequently to adapt the trained neural networks on-line. One type of feed-forward NN which in recent years has received growing interest is the *radial basis function* (RBF) NN. In the design process of a NN model several questions arise, which are of major importance to the performance of the designed NN: from a set of available variables in the form of data, which should be used as inputs to the NN? Having made this decision, are delayed input and output values important? What delays shall be employed? This paper addresses the usefulness of correlation-based model validity tests to help answer the last two questions. The tests will be applied to NN models obtained by off-line training. The quality of models is assessed by their prediction performance and the correlation tests.

[1] The authors would like to acknowledge the FCT (project MGS/33906/99-00 and grant SFRH/BD/1236/2000) for supporting this work.

The input-output structures considered will be tested for prediction in NNs adapted on-line. The following section presents previous work on the modelling of the greenhouse inside air temperature using RBF neural networks. In the next three sections the correlation-based validation tests, the RBF NN and the employed design methods will be briefly reviewed. In section 6 the experiments will be explained. Results from these experiments and their discussion are presented in section 7 and finally some conclusions are drawn in section 8.

2. PREVIOUS WORK

The design problem being considered is to model the inside air temperature of an hydroponic vegetable production greenhouse as a function of the outside solar radiation and temperature, and the inside relative humidity. Previous work on this NN modelling task relied in an input-output model structure selected from Cunha *et al.* (1996) in the context of dynamic temperature models identification. This structure was selected by means of the Akaike information criterion (Akaike, 1974) where several hypothesis were tested and the best one chosen. It is a second order model with one delay from the outside solar radiation to the inside air temperature. The application of the RBF NN to greenhouse inside air temperature modelling has been previously investigated (Ferreira *et al.*, 2000*a*). This type of feed-forward NN is structurally simple and may be characterised by a nonlinear-linear topology in the parameters. Existing hybrid training methods already reflect this structure, as found on RBFs, but fail to fully exploit it on the minimisation of a single explicit training criterion. An algorithm based on unconstrained deterministic optimisation using Levenberg-Marquardt (LM) methods, which exploits this feature on the minimisation of a new training criterion (Ferreira and Ruano, 2000) has been proposed and analysed, and a strategy for its on-line application is also suggested (Ferreira *et al.*, 2000*b*). A comparison study (Ferreira *et al.*, 2000*b*) where various off-line and on-line methods were considered revealed that the on-line LM method exploiting the separability of parameters achieved the best performance on this modelling problem. The basis for comparison was the one-time-step-ahead prediction error and the network size. This idea was reinforced by comparisons made (Ferreira and Ruano, 2001) over greater prediction horizons. This model is intended to be used in a greenhouse adaptive predictive hierarchical control scheme as shown in figure 1.

3. THE VALIDATION TESTS

Correlation-based model validity tests are methods that permit the detection of inadequacy of fitted models. In the case of NN modelling these inadequacies

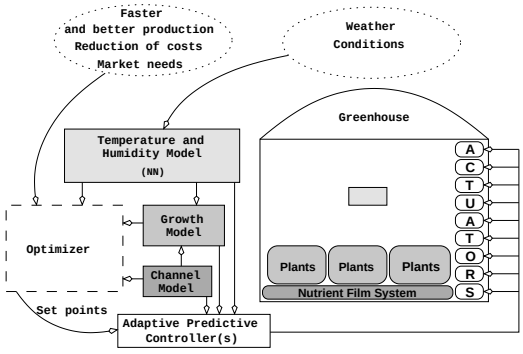

Fig. 1. Hierarchical greenhouse environmental control strategy

can be due to various reasons: insufficient number of neurons, under-training or over-training, incorrect inputs selection, inappropriate training algorithm, noisy data, among others. Independently of the cause, validation tests should give indication that the model is not correct in some way. Validation tests for linear models are well documented in the literature (Ljung, 1987). These involve the computation of the cross-correlation function between the model error and the inputs, and the auto-correlation function of the error. It is known that these tests by themselves are insufficient as they are unable to cope with non-linear effects in the data. Billings and Voon (1986) and Billings and Zhu (1993) have shown that additional tests based on higher-order correlation functions using errors, inputs and outputs are necessary to detect non-linear effects in the prediction error. If the fitted model is adequate the following conditions should hold:

$$\begin{aligned} R_{ee}(\tau) &= \delta(\tau) \\ R_{ue}(\tau) &= 0 \quad ,\forall\tau \\ R_{u^2{}'e}(\tau) &= 0 \quad ,\forall\tau \\ R_{u^2{}'e^2}(\tau) &= 0 \quad ,\forall\tau \\ R_{e(eu)}(\tau) &= 0 \,, \tau \geq 0 \\ R_{e^2e^2}(\tau) &= \delta(\tau) \\ R_{(ye)e^2} &= k\delta(\tau) \\ R_{(ye)u^2} &= 0 \quad ,\forall\tau \end{aligned} \tag{1}$$

If the normalized correlation functions (1) lie within the 95% confidence bands at $1.96/\sqrt{N}$, N being the size of the data vectors, the model is considered adequate.

4. OVERVIEW OF RBFNS

A RBF NN consists of three fully connected layers. The first is the input layer connecting the source nodes to the hidden layer of the network, which is composed of a certain number of units, called neurons. The outputs of the hidden layer are then linearly combined by a set of parameters to produce the overall network response in the output layer. This way the network

performs a mapping, f, from an input space, $\mathcal{X}^d$, to an output space, $\mathcal{Y}^m$. The hidden layer applies a non-linear transformation to the inputs generating an hidden space which in general has a higher dimension than $\mathcal{X}$. Broomhead and Lowe (1988) proposed the RBF network which is described by the following equation:

$$f(x_j) = \sum_{i=1}^{n} \alpha_i \varphi\left(\left\|x_j - c_i\right\|\right) \tag{2}$$

where the $\{c_i\}_{i=1}^n$ are a set of points called centers which, together with the set of weights, $\{\alpha_i\}_{i=1}^n$, have to be chosen in order to minimise the distance, from the approximation f to the target y, stated as:

$$\mathcal{E}(f) = \sum_{j=1}^{N} (y_j - f(x_j))^2 \tag{3}$$

Defining $\Lambda = [\alpha_i, \ldots, \alpha_n]^T$ as the linear weight vector, eq. (2) can be written in order to the weights in the following compact form,

$$\Lambda = \Phi^+ \mathbf{y} \tag{4}$$

where $\mathbf{y}$ is an N-by-1 vector of the desired target values and Φ is an N-by-n matrix whose elements $\varphi_{j,i}$ are the values of the radial basis functions centred at $\{c_i\}_{i=1}^n$ and evaluated at the points $\{x_j\}_{j=1}^N$. Φ^+ denotes the pseudo-inverse of Φ. The most used function in RBF NNs is a Gaussian function of the form:

$$\varphi_i(x_j) = e^{-\frac{1}{2\sigma_i^2}\left\|x_j - c_i\right\|^2}$$

5. TRAINING METHODS

5.1 *Off-line training*

In this approach the centre locations, the spreads of centers and the output linear weights are all determined under a *supervised learning procedure* based on unconstrained deterministic optimisation. Basically new parameter values are calculated in an iterated manner in order to minimise the cost function,

$$\mathcal{E} = \frac{1}{2} \sum_{i=1}^{N} (t(i) - y(i)) \tag{5}$$

where $\mathbf{t}$ is the vector of target values and vector $\mathbf{y}$ is defined from eq. (2) as $\mathbf{y} = [f(x_1), \ldots, f(x_N)]^T$. Eq. (5) can also be rewritten as:

$$\mathcal{E} = \frac{1}{2} \|\mathbf{t} - \mathbf{y}\|_2^2 \tag{6}$$

Let $\mathbf{u} = [\alpha_1, \ldots, \alpha_n]^T$, $\mathbf{v} = \left[c_1^T, \ldots, c_n^T, \sigma_1, \ldots, \sigma_n\right]^T$ and $\mathbf{w} = \left[\mathbf{v}^T, \mathbf{u}^T\right]^T$. The neurons output N-by-n matrix is $\mathbf{O} = [\varphi_1(x_j), \ldots, \varphi_n(x_j)]_{j=1}^N$. The outputs of the network can be described by eq. (7), where the linear dependence of the network on the output weights and the dependence of $\mathbf{O}$ on $\mathbf{v}$ have been made explicit.

$$\mathbf{y} = \mathbf{O}(\mathbf{v})\,\mathbf{u} \tag{7}$$

Eq. (6) now becomes:

$$\mathcal{E}(\mathbf{w}) = \frac{1}{2} \|\mathbf{t} - \mathbf{O}(\mathbf{v})\,\mathbf{u}\| \tag{8}$$

The formulation presented so far involves all the network parameters in the optimisation procedure. As already mentioned the output weights can be optimally determined by the *least squares* (LS) solution. Substituting the target values vector, $\mathbf{t}$, in eq. (7), denoting matrix $\mathbf{O}(\mathbf{v})$ by $\mathbf{A}$ and solving for $\mathbf{u}$ yields,

$$\hat{\mathbf{u}} = \mathbf{A}^+ \mathbf{t}$$

where $\mathbf{A}^+$ stands for the pseudo-inverse of matrix $\mathbf{A}$. Substituting this result in eq. (8) gives the new training criterion:

$$\mathcal{E}(\mathbf{v}) = \frac{1}{2} \left\|\mathbf{t} - \mathbf{A}\mathbf{A}^+ \mathbf{t}\right\| \tag{9}$$

This new training criterion does not depend on the linear parameters, $\mathbf{u}$, and explicitly incorporates the finding that, whatever values the nonlinear parameters $\mathbf{v}$ take, the $\mathbf{u}$ parameters employed are the optimal ones. For non-linear LS problems the LM algorithm is recognised as the best method, as it exploits the sum-of-the squares characteristic of the problem (Ruano *et al.*, 1992). Let Ω_k denote the training criterion in iteration k. The optimisation procedure is iterated until a set of termination criteria is met (Gill *et al.*, 1981). Assume θ_k is a measure of absolute accuracy, where τ_f is a measure of the desired number of correct figures in the objective function:

$$\theta_k = \tau_f (1 + \Omega_k)$$

The optimisation stops when all the following conditions are met:

$$\begin{aligned} \Omega_{k-1} - \Omega_k &< \theta_k \\ \|\mathbf{v}_{k-1} - \mathbf{v}_k\| &< \sqrt{\tau_f}\,(1 + \|\mathbf{v}_k\|) \\ \|\mathbf{g}_k\| &\leq \sqrt[3]{\tau_f}\,(1 + |\Omega_k|) \end{aligned} \tag{10}$$

$\mathbf{g}_k$ is the gradient vector involved in the LM optimisation method.

5.2 *On-line learning*

The on-line learning algorithm considered comes from the LM method presented in the previous section and the reasoning behind its implementation follows. The LM optimisation method is iterated a certain number of times, at each time step k, over a subset $\mathcal{Z}_M(k) =$

Table 1. Values involved in signal pre-processing

Signal	DC term	Interval
Inside air temperature ($^{\circ}C$)	13.1	$[0\ldots24]$
Outside air temperature ($^{\circ}C$)	10.8	$[0\ldots24]$
Outside solar radiation (W/m^2)	0	$[0\ldots1070]$
Inside relative humidity (%)	19	$[0\ldots100]$

$\{\mathcal{X}_j, \mathcal{Y}_j\}_{j=k-M+1}^{k}$ of the training data, until the termination criteria (10) is met. At $k+1$ the first input-output pair in $\mathcal{Z}$ is discarded and the one pertaining time step $k+1$ is added. Assuming that the dimension of $\mathcal{Z}$ is large enough two conclusions can be drawn: its statistical properties at $k+1$ are essentially the same as in k, and its distribution on input space is representative of the process data to some extent. As a consequence, the point $\mathbf{w}$ in parameter space that minimises Ω at time $k+1$ will be the same as in k with a slight correction. The choice of M is application and problem dependent and care should be taken with its choice in order to satisfy the assumptions made. A more detailed explanation and analysis of both the off-line and on-line algorithms can be found in Ferreira *et al.* (2001) and Ferreira and Ruano (2000).

6. EXPERIMENTAL SETUP

The data set used in the experiments is composed of 4257 points acquired with a sample rate of 5 minutes. All DC terms were subtracted from the signals, which were then scaled to an amplitude one, $[-0.5, 0.5]$, interval. Table 1 shows the values of the DC terms subtracted from the signals and the amplitude interval from which they were scaled. The prediction horizon considered is 24 time steps. For the off-line trained NNs the complete test data set will also be considered as a prediction horizon. The first 1000 points compose the training set and the rest is used for testing purposes. Over the prediction horizon the model predicted values are used for the inputs related with past values of the inside air temperature. The size of the networks was chosen as the best performing one from a previous comparison study (Ferreira *et al.*, 2001). 6 and 8 neurons are used for the on-line adapted and off-line trained NNs, respectively. A value of $\tau_f = 0.001$ is employed for off-line trained NNs and $\tau_f = 0.01$ for on-line adapted ones, in the termination criteria (10). In the on-line case M takes the size of one day of data. The initial values for the centers of the RBF NNs are obtained by one iteration of the *optimal adaptive k-means* clustering algorithm (Chinrungrueng and Séquin, 1995). The spreads of the Gaussian activation functions are determined as in Haykin (1998). The initial linear weight vector for the LM minimising the new criterion is determined using the LS optimal values with a small perturbation. By the time the data was acquired the greenhouse environment was not controlled by any means. It should be noticed that when this happens the models of environmental quantities in the greenhouse will need to be revised in order to account for the control actions. For the present study it will be assumed that the relevant quantities for the model are those already mentioned as used in previous work. Table 2 presents the models that are analysed and the terms used as their inputs. Model M7 is included for comparison purposes. It has been used in previous studies and is referred in section 2. The addition or remotion of a term was based on the prediction error analysis and by visual inspection of the validation tests (1). SR_o, H_i, AT_o and AT_i stand for the outside solar radiation, inside relative humidity, outside air temperature and inside air temperature, respectively. k is the time instant. In the following, models will be referred to by the names in the first column of table 2.

Table 2. Tested models and input terms employed

Model	SR_o	H_i	AT_o	AT_i
M1	k	k	k	k-1
M2	k	k	k	k-1, k-2
M3	k	k, k-3	k	k-1, k-2
M4	k, k-21	k, k-3	k	k-1, k-2
M5	k, k-21	k	k	k-1, k-2
M6	k, k-21	k, k-3	k, k-2	k-1, k-2
M7	k-1,k-2	k, k-1	k, k-1	k-1, k-2

Table 3. Summary of error values obtained with the off-line trained NNs.

Model	E_t	$min(\mathrm{E}_p)$	$mean(\mathrm{E}_p)$	$max(\mathrm{E}_p)$
M1	0.006366	0.0015	0.0642	0.5220
	0.2740	0.0000	0.1766	0.6929
M2	0.005806	0.0017	0.0433	0.2558
	0.0818	0.0000	0.0542	0.4036
M3	0.005896	0.0019	0.0371	0.1851
	0.0709	0.0000	0.0515	0.2715
M4	0.005212	0.0021	0.0357	0.1461
	0.0650	0.0000	0.0513	0.2338
M5	0.005461	0.0014	0.0312	0.1424
	0.0731	0.0000	0.0480	0.2888
M6	0.005030	0.0013	0.0280	0.1632
	0.0591	0.0000	0.0428	0.2301
M7	0.005359	0.0015	0.0405	0.1866
	0.0795	0.0000	0.0574	0.2617

7. RESULTS AND DISCUSSION

Table 3 presents the predictive error values obtained by the models in table 2. The value in the E_t column on the first line for each model is the *root mean square error* (RMSE) of the one-step-ahead prediction obtained during training. For all points in the test data set one prediction horizon (24 points) was calculated and the RMSE over it was computed, resulting in a vector, E_p. The last three values in the first line of each model are the minimum, mean and maximum values of this vector. The second line presents absolute prediction errors considering the whole test data set as prediction horizon. The first value is the RMSE of the error obtained and the following values are the minimum, mean and maximum of the absolute error values. Figures 2 to 6

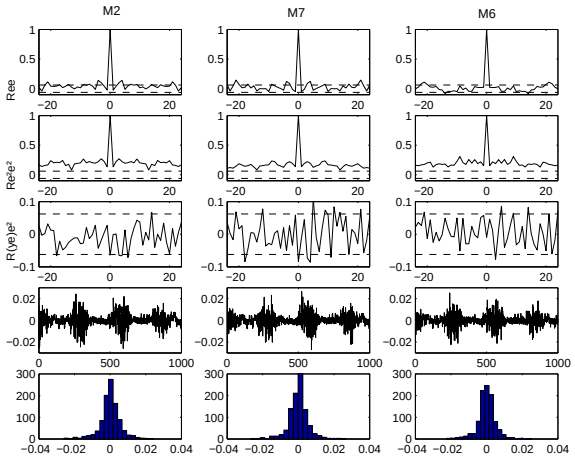

Fig. 2. Correlation tests involving error and output. Prediction error. Error distribution.

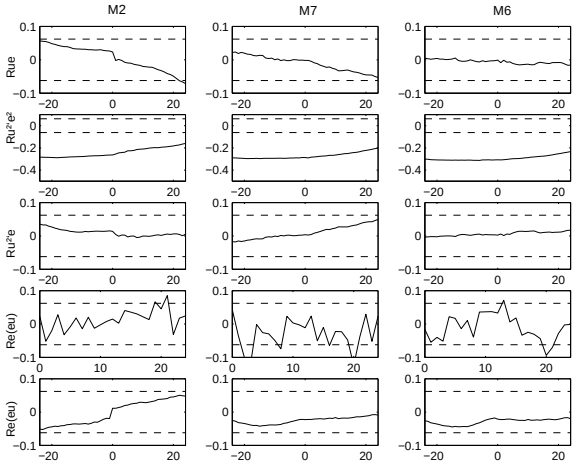

Fig. 3. Correlation tests involving error, output and outside solar radiation.

depict the correlation based tests. All these figures are organized by columns, each related to one model as indicated. The first figure presents the tests involving the error and output, the error sequence and its distribution. Subsequent figures present the tests relating the error, output and inputs. These are the outside solar radiation, inside relative humidity, outside air temperature and inside air temperature, respectively.

Models M1 and M2 were used as starting models, compared in terms of error performance and validation tests. Looking at table 3 there is no doubt that M2 behaves better than M1. This was also the case with the validation tests. Although the tests for M2 are sometimes outside the confidence bands, for M1 there were more test violations. Inspecting the validation tests for M2 it can be observed that $R_{ee}(\tau)$ is violated at $\tau = \{3,4,21\}$ and that the tests involving the inputs, particularly $R_{e(eu)}$, are violated close to those values of τ. For the smaller values of τ the biggest violation occurs for H_i. This term was included and comparing M3 with M2 using table 3, it is possible to affirm that its inclusion was beneficial. The violations in the test at small values of τ also disappeared in $R_{e(eu)}$ for H_i and were reduced in $R_{ee}(\tau)$. The same type of

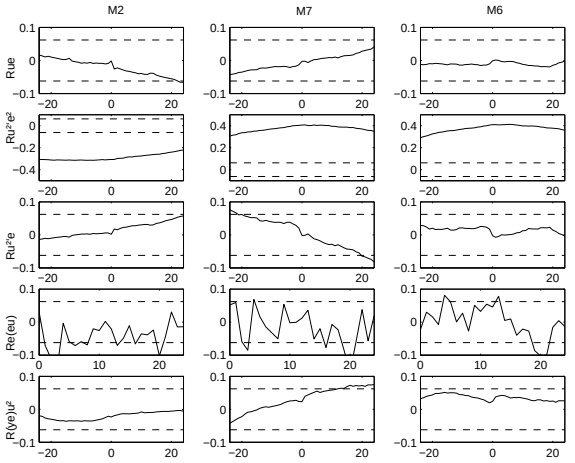

Fig. 4. Correlation tests involving error, output and inside relative humidity.

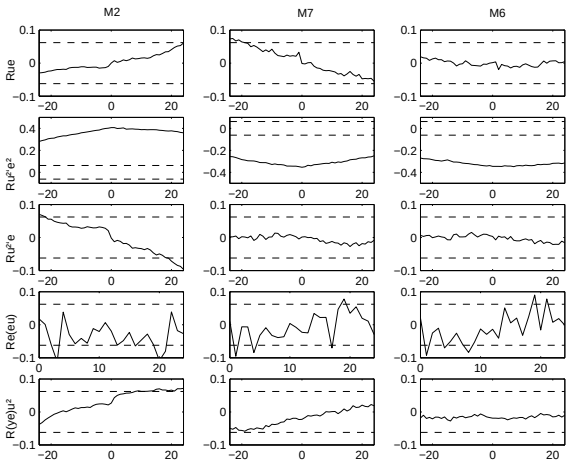

Fig. 5. Correlation tests involving error, output and outside air temperature.

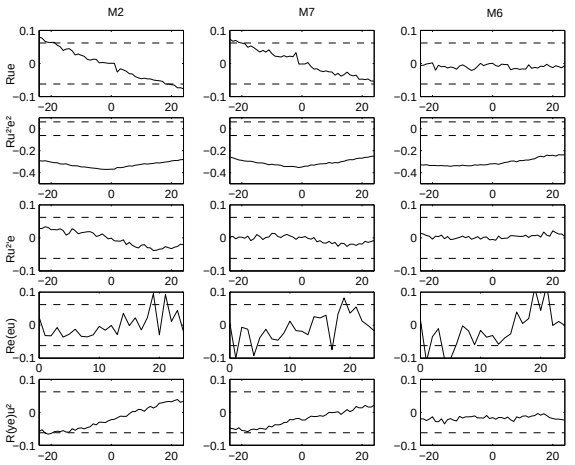

Fig. 6. Correlation tests involving error, output and inside air temperature.

behaviour was verified when the term at $\tau = 21$ for SR_o was introduced in M5. From the tests and comparisons of M2, with M3 and M5, the two terms were included in M4. The resulting tests were more correct and the error performance was, overall, better. The test $R_{e(eu)}$ for AT_o and R_{ee} still presented violations around $\tau = 2$

Table 4. Summary of RMSE values obtained with the on-line NNs.

Model	$min(E_p)$	$mean(E_p)$	$max(E_p)$
M6	0.0016	0.0305	0.3063
	0.0069	0.0389	0.0677
M7	0.0018	0.0334	0.5596
	0.0069	0.0531	0.1020

resulting in the inclusion of that term to form M6. This model has the same complexity as M7 and in terms of prediction error is always better. Also the validation tests present better behaviour. Those tests which presented some slope are for M6 always near zero and in general the tests are better fitted within the confidence bands. The tests $R_{e^{2\prime}e^2}(\tau)$ and $R_{u^{2\prime}e^2}(\tau)$ are always outside the confidence bands. They could fit but for some reason a bias is always present. Table 4 presents the prediction errors obtained by models M6 and M7 in simulated on-line operation. For each model the first line presents data related with the RMSE values over the prediction horizon, computed at every time instant. The second line regards the RMSE values computed for each prediction instant over all simulation steps. Model M6 is significantly better than M7.

8. CONCLUSIONS AND FUTURE WORK

Analysing the results presented in the previous section it can be concluded that the correlation based validation tests were useful to strengthen the quality of the greenhouse inside air temperature predictive model. The procedure of iteratively inspecting the tests results, correcting the model structure and comparing the new error values, conducted to a better model in terms of error performance and more correct in terms of correlation tests. The comparison of this model with one used in previous work accounts in benefit of this technique: a more correct model of the same complexity was achieved, with better error performance in both off-line and on-line operation. As the model will be revised once the greenhouse becomes controlled, future work will focus on the selection of input variables from a set of candidate ones and on a means to predict the external perturbations and inside relative humidity.

9. REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **AC-19**, 716–723.

Billings, S.A. and Q.M. Zhu (1993). Nonlinear model validation using correlation tests. Research report n^o 463. Department of Automatic Control and Systems Engineering. University of Sheffield, Sheffield S14DU, UK.

Billings, S.A. and W.S.F. Voon (1986). Correlation based model validity tests for non-linear models. *International Journal of Control* **44**(1), 235–244.

Broomhead, D.S. and D. Lowe (1988). Multivariable functional interpolation and adaptive networks. *Complex Systems* **2**, 321–355.

Chinrungrueng, C. and C.H. Séquin (1995). Optimal adaptive *K-M*eans algorithm with dynamic adjustment of learning rate. *IEEE Transactions on Neural Networks* **6**(1), 157–169.

Cunha, J. Boaventura, A.E. Ruano and E.A. Faria (1996). Dynamic temperature models of a soilless greenhouse. In: *Proceedings of the* 2^{nd} *Portuguese Conference on Automatic Control*. Vol. 1. Portuguese Association of Automatic Control. Porto, Portugal. pp. 77–81.

Ferreira, P.M. and A.E. Ruano (2000). Exploiting the separability of linear and nonlinear parameters in radial basis function networks. In: *IEEE Symposium 2000: Adaptive Systems for Signal Processing, Communications, and Control*. Lake Louise, Alberta, Canada. pp. 321–326.

Ferreira, P.M. and A.E. Ruano (2001). Predicting the greenhouse inside air temperature with rbf neural networks. In: *Preprints of the* 2^{nd} *IFAC-CIGR Workshop on Intelligent Control for Agricultural Applications*. Bali, Indonesia. pp. 67–72.

Ferreira, P.M., E.A. Faria and A.E. Ruano (2000*a*). Application of radial basis function neural networks to a greenhouse inside air temperature model. In: *IFAC Agricontrol 2000, International Conference on Modelling and Control in Agriculture, Horticulture and Post-harvested Processing*. Vol. 2. Wageningen, The Netherlands. pp. 172–177.

Ferreira, P.M., E.A. Faria and A.E. Ruano (2000*b*). Neural network models in greenhouse environmental control. In: *Proceedings of the* 6^{th} *International Conference on Engineering Applications of Neural Networks (EANN'2000)* (Dimitris Tsaptsinos, Ed.). Kingston University, England. pp. 100–107.

Ferreira, P.M., E.A. Faria and A.E. Ruano (2001). Neural network models in greenhouse air temperature prediction. *Neurocomputing*. To appear in a forthcoming special issue on *Engineering Applications of Neural Networks*.

Gill, Philip E., Walter Murray and Margaret H. Wright (1981). Practicalities. In: *Practical Optimization*. Chap. 8, p. 306. Academic Press, Inc.

Haykin, Simon (1998). Learning strategies. In: *Neural Networks: A Comprehensive Foundation*. second ed.. Chap. 5, p. 299. Prentice Hall.

Ljung, L. (1987). *System Identification: Theory for the user*. Prentice-Hall, Inc.. Englewood Cliffs, New Jersey.

Ruano, A.E., P.J. Fleming and D.I. Jones (1992). A connectionist approach to pid autotuning. In: *IEE Proceedings, Part D*. Vol. 139. Brighton, U.K.. pp. 279–285.