Proceedings of the International Conference
on Computational and Mathematical Methods
in Science and Engineering, CMMSE2009
Gijón, Spain, June 30, July 1-3 2009

1. Soft Computing for detecting thermal insulation failures in buildings

Javier Sedano[1], Enrique de la Cal[2], Leticia Curiel[3], José Ramón Villar[2], Emilio Corchado[3]

[1,] *Department of Electromechanical Engineering. University of Burgos, Burgos, Spain*
[2,] *Department of Computer Science, University of Oviedo, Spain*
[3] *Department of Civil Engineering, University of Burgos, Burgos, Spain*

emails: jsedano@ubu.es, delacal@uniovi.es , lcuriel@ubu.es, villarjose@uniovi.es, escorchado@ubu.es

Abstract

Detecting thermal insulation in buildings represents a big challenge to reduce the impact of domestic heating system in the global warming, but also in the economy. Thermal insulation failure detection is usually considered in the contract and rarely it is analysed in the normal situations, when the building is in operation. In this case, the lighting, occupancy and temperature profiles and ventilation services make this process quite complex. Identifying this kind of failures in operation would reduce the energy consumption in heating systems. In this paper, soft computing techniques are hybridized to detect the insulation thermal failures whenever the building is in operation. The study considered the local regulations in building and heating systems and the specific climate zone analysed. The three steps procedure is detailed. Firstly, the thermal dynamics have been modelled for different building characteristics and climate zones. Secondly, the Cooperative Maximum-Likelihood Hebbian Learning is used to extract the relevant features. Finally, neural projections and identification techniques have been applied in order to detect bias in the room temperatures and, consequently, thermal insulation failures. Although a vast amount of research has to be carried out, it is expected to outperform the energy consumption in the Spanish buildings.

1. Introduction

Soft computing represents a collection or set of computational techniques and intelligent systems principles in machine learning, computer science and some engineering disciplines, which investigate, simulate, and analyze very complex issues and phenomena in order to solve real-world problems. Specifically, the thermal insulation failure detection in buildings is one of such kind of complex problems [1], [2]. In thermal insulation detection several different disciplines conjunct [3].

On the one hand, current local regulations need to be analysed and considered in order to accomplish with all the premises and to establish the legal bounds of the physical variables. In the case of Spain, the building regulations and also the heating systems regulations define up to five winter climate zones and five summer climate zones in the whole country territory. Considering the pair <winter zone, summer zone> the building materials, the insulation width and materials and so on are determined. Nevertheless, there is another factor given by market behaviour that should be included: the building design –geometrical and orientation aspects–, the aesthetics aspects and the inner or in-door distributions aspects have a high impact in the thermal dynamic of a building. This is called the building topology. A topology and a pair of <winter zone, summer zone> forms what is called configuration. Given a configuration the environmental variables history –such as sun radiation, outdoor temperature, wind speed, etc.– can be gathered. Also, the heating and ventilation systems can be designed and the occupancy profiles, lighting profiles, small power devices profiles, ventilation profiles and temperature set point profiles can be estimated.

On the other hand, predicting the thermal dynamic in a building is a complex task. Thermal dynamic has been used mainly to estimate the power requirements in buildings. As an example, in [4] the difficulties in obtaining the black box model for a generic building is documented. Further more, in [5] the errors with different kind of techniques and a possible solution are given. In both cases, the aim of the study is the design an electrical energy distribution device to control electrical heating systems in order to bound the electrical energy consumption while the comfort is kept in the building. Also, in [5] the influence in the thermal efficiency is analysed for a specific building component. In this case, the thermal dynamics of an aluminium roof is analysed and compared with standard roof materials. The mentioned works are all involved in the design of new buildings, but -up to our knowledge- nothing has been said about the efficiency in buildings that are in operation.

This work represents a step forward the development of techniques to improve the thermal efficiency in buildings through the detection of thermal insulation failures while the building is in operation. The proposal is based in determining when the behaviour of the available variables is biased with respect the steady state. Although this can be seen as a simple task, the noise introduced due the occupancy or the lighting profiles of use in the building can distort to make such detection much more complex than it seems. The proposal includes a three steps procedure for testing and validating the model. Firstly, the thermal dynamics of a building in a specific configuration is calculated by means of the HTB2 software [6]. The outcome of the HTB2 should be post processed in order to obtain a suitable data set. Then the data set is analysed using the Cooperative Maximum-Likelihood Hebbian Learning (CMLHL) to extract the structure of the data and the relationships between the variables. Afterwards, the modelling stage will produce a model to estimate the temperature in the room for the specific configuration. Finally, if the temperature error, which is measured as the difference between the room temperature and the model output temperature, is higher than a certain threshold it can be concluded that there are thermal insulation failures in that room.

This paper is organised as follows. Section 2 introduces the used unsupervised connectionist techniques for analysing the data in order to extract the relevant internal structures. Section 3 deals with the classical identification techniques used in the system modelling. In Sect. 4, the problem details and the multi-step procedure are detailed. Finally, conclusions and future work are included.

2. System Analyses Using Unsupervised Connectionist Techniques

2.1. Data structure analysing using connectionist techniques

The CMLHL [7] is used in this research in order to analyse the internal structure of the data set, which describe the heaing process to establish whether it is "sufficiently informative". In the worse case, the experiments have to be performed again in order to collect a proper and informative data set.

CMLHL is a Exploratory Projection Pursuit (EPP) method [8] [9]. In general, EPP provides a linear projection of a data set, but it projects the data onto a set of basic vectors which help reveal the most interesting data structures; interestingness is usually defined in terms of how far removed the distribution is from the Gaussian distribution [10].

One connectionist implementation is Maximum-Likelihood Hebbian Learning (MLHL) [11], [12]. It identifies interestingness by maximising the probability of the residuals under specific probability density functions that are non-Gaussian. An extended version is the CMLHL [7] model, which is based on MLHL [11],[12] but adds lateral connections [13], [14] that have been derived from the Rectified Gaussian Distribution [10].

Considering an N-dimensional input vector (x), and an M-dimensional output vector (y), with W_{ij} being the weight (linking input j to output i), then CMLHL can be expressed [15], [16] as:

1. Feed-forward step:

$$y_i = \sum_{j=1}^{N} W_{ij} x_j, \forall i \quad . \tag{1}$$

2. Lateral activation passing:

$$y_i(t+1) = \left[y_i(t) + \tau (b - Ay) \right]^{+} \quad . \tag{2}$$

3. Feedback step:

$$e_j = x_j - \sum_{i=1}^{M} W_{ij} y_i, \forall j \quad . \tag{3}$$

4. Weight change:

$$\Delta W_{ij} = \eta . y_i . sign(e_j) | e_j |^{p-1} \quad . \tag{4}$$

Where: η is the learning rate, τ is the "strength" of the lateral connections, b the bias parameter, p a parameter related to the energy function [10], [11], [12] and A a symmetric matrix used to modify the response to the data [10]. The effect of this matrix is based on the relation between the distances separating the output neurons.

2.2. Feature selection and extraction

Feature Selection and extraction [15], [16] includes feature construction, space dimensionality reduction, sparse representations and feature selection among others. All these techniques are commonly used as pre-processing tools to machine learning tasks including pattern recognition. Although researchers have tackled such problems for many years, there has been recently a renewed interest in feature extraction. A large number of new applications with very large input spaces critically need space dimensionality reduction for efficiency and efficacy of the predictors. Some of these applications include new and classical topics as bioinformatics (DNA microarrays, etc.), remote sensing multi- and hyperspectral imagery, pattern recognition (e.g. handwriting recognition, text processing, web mining), speech processing, artificial vision, industrial applications and so on.

In this study, our approach to feature selection is based mainly on the dimensionality reduction issue, using initially a projectionist method called Cooperative Maximum Likelihood Hebbian Learning (CMLHL)[7], characterized for its capability to enforce a more sparse representation in each weight vector than other classical methods as PCA [17], [18] or Maximum Likelihood Hebbian Learing (MLHL) [11], [12] and its capability of preserving some global ordering [14], due to the effect of the lateral connections.

3. System Modelling Using Classical Identification Algorithms

3.1. The identification criterion

The identification criterion consists in evaluating which of the group of candidate models is the best adapted and that best described the data set gathered for the experiment; i.e., given a certain model $M(\theta_*)$, its prediction error may be defined as in Eq. (5). The aim is to obtain a model that complies with the following premise [9]: a good model is one that makes good predictions, and which produces small errors when the observed data is applied, i.e., on any one data set Z^t it will calculate the prediction error $\varepsilon(t,\theta)$, Eq. (5), in such a way that for any one t=N, a particular $\hat{\theta}_N$ (estimated parametrical vector) is selected so that the prediction error $\varepsilon(t,\hat{\theta}_N)$ in t=1,2,3…N, is minimized.

$$\varepsilon(t,\theta_*) = y(t) - \hat{y}(t \mid \theta_*) \tag{5}$$

The methodology of black-box structures [9], is used for linear systems or for systems that are linear in the working area; this methodology has the advantage of only requiring very few explicit assumptions on the pattern to be identified, but that in turn makes it difficult to quantify the model that is obtained. The discrete linear models may be represented through the union of a deterministic model and a stochastic model, Eq. (6). In Eq. (6), u(t) is the input, y(t) is the output, $G(q^{-1})$ is the transfer function from u(t) to y(t), $H(q^{-1})$ is the transfer function from e(t) to y(t) and q, q^{-1} are forward and backward shift operators. The term e(t) (white noise signal) includes the modelling errors and is associated with a series of random variables, of mean null value and variance λ.

$$y(t) = G(q^{-1})u(t) + H(q^{-1})e(t) \tag{6}$$

The structure of a black-box model depends on how the noise influences the model [12], that is, the term $H(q^{-1})$. Thus, if this term is 1, then the FIR (Finite Impulse Response) and OE (Output Error) models are applicable; whereas if it is different from zero a great range of models are applicable; the most common being: ARX (Autoregressive with external input), ARMAX (Autoregressive Moving Average with external input), BJ (Box Jenkins) and ARMA (Autoregressive Moving Average).

3.2. The ANN in the identification process

The use of ANN in the process of identification requires the selection of several parameters: the number of layers, the number of neurons per layer and the activation functions. The methods to set up the parameters have been documented in depth in the literature, and it was found the ANN with two layers using sigmoidal of hyperbolic functions in the hidden layer are universal approximators or predictors [20], [21].

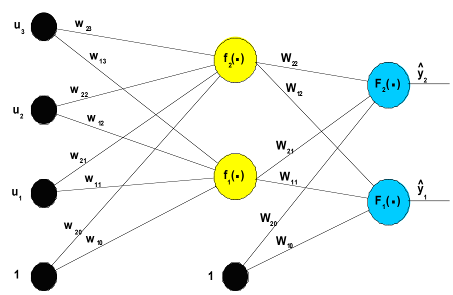

Fig. 1. A MLP network with two layers, with two nodes per layer, and three inputs. *W* is the weight matrix between the hidden and output layer, while *w* is the weight matrix between the inputs and the h has two bias nodes with value 1.

Also, the number of neurons per layer is a relevant design parameter that should be analyzed in order to avoid over fitting [22], [23]; which state that the number of neurons and the size of the weight matrix depends on the ANN training algorithm. Each algorithm will introduce some restrictions in the weight matrix. The most widely used training algorithms in system identification are the Lenvenberg-Marquardt method [24], recursive Gauss-Newton method [19], the batch and the recursive versions of the back-propagation algorithm [25].

3.3. The process of identification

When using ANN, the purpose of a identification process is determining the weight matrix based on the observations Z^t, so that the relationships between nodes in the network are given. The weight matrix is usually referred as w, W or θ.
Then the supervised learning algorithm is carried out to find the estimator θ, so that the identification criterion is obtained. In the case of this work, the minimization of the mean square error criterion is used as defined in Equations (7) and (8). The iterative minimization scheme is defined in Eq. (9), where f(t) is the search direction and $\mu(t)$ step size.

$$V_N(\theta, z^t) = \frac{1}{2N}\sum_{t=1}^{N}\left[y(t) - \hat{y}(t \mid \theta)\right]^T\left[y(t) - \hat{y}(t \mid \theta)\right] \tag{7}$$

$$\hat{\theta} = \arg\min_{\theta} V_N(\theta, Z^t) \quad (8)$$

$$\theta(t+1) = \theta(t) + \mu(t) f(t) \quad (9)$$

Several different well-known model structures are used when merging system identification with ANN. If the ARX model is used as the regresion vector θ, the model structure is called NNARX, as can be seen in Eq. (10). NNARX stands for neural network ARX. Likewise, NNFIR, Eq. (11), NNARMAX, Eq. (12) and NNOE structures, Eq. (13), are also used in depth. In the same way, it posible to use a estimator for the on-step prediction ahead of the output $\hat{y}(t|\theta)$, i.e., the NNARX, Eq. (14), the NNFIR, Eq. (14), the NNARMAX, Eq. (15) and the NNOE, Eq. (15). The value of the degree of each polynomio -n_a, n_b, n_c, n_d, n_f and n_k- are given as parameters.

$$\varphi(t) = \left[y(t-1)\ldots y(t-n_a)u(t-n_k)\ldots u(t-n_b-n_k+1)\right]^T \quad (10)$$

$$\varphi(t) = \left[u(t-n_k)\ldots u(t-n_b-n_k+1)\right]^T \quad (11)$$

$$\varphi(t) = \left[y(t-1)\ldots y(t-n_a)u(t-n_k)\ldots u(t-n_b-n_k+1)e(t-1)\ldots e(t-n_c)\right]^T \quad (12)$$

$$\varphi(t) = \left[\hat{y}(t-1|\theta)\ldots \hat{y}(t-n_a|\theta)u(t-n_k)\ldots u(t-n_b-n_k+1)\right]^T \quad (13)$$

$$\hat{y}(t|\theta) = \hat{y}(t|t-1,\theta) = g(\varphi(t),\theta) \quad (14)$$

$$\hat{y}(t|\theta) = g(\varphi(t),\theta) \quad (15)$$

3.4. The system identification methodology

The best model for estimating the conditions of the thermal dynamic must be chosen. The identification procedure includes establishing the identification techniques [19], [26], [27], [28], [29]], the selection of the model structure, the estimation of the suitable polynomials degree [30], [31], the identification criterion, and the optimization techniques in order to generate the final model.

Also, the identification procedure includes a validation stage, which ensures that the selected model meets the necessary conditions for estimation and prediction. In order to validate the model, three tests were performed: residual analysis $\varepsilon(t,\hat{\theta}(t))$, by means of a correlation test between inputs, residuals and their combinations; final prediction error (FPE) estimate as explained by Akaike [32] and finally the graphical comparison between desired outputs and the outcome of the models through simulation one (or k) steps before.

4. The multi-step method for detecting thermal insulation failures in buildings.

4.1. Thermal dynamics data gathering by means of simulation

The methodology that follows is proposed to detect thermal insulation failures in buildings. Firstly, a model of the dynamics of a building should be determined. In order to determine the dynamics, the building conditions in normal operational stage should be obtained. This data can be gathered by two methods: by distributing sensors among the building or by simulation. Once the thermal dynamics data is gathered, then a model of normal operation can be obtained. Finally, thermal insulation failures can be detected using this model when a significant bias in the room temperature is identified.

This section deals with the whole method to detect the thermal insulation failures, In the next subsection, the procedure to gather thermal dynamics data from simulation is detailed. Then, it is shown how the CMLHL is used to extract features from the thermal dynamics data with the relevant information of the process. Finally, the system identification techniques are applied to obtain the model which must detect the thermal insulation failures.

Thermal dynamics data gathering by meas of simulation There are two methods to obtain data for the dynamics of a building: by a network of sensors placed in each space of a scaled or real size building measuring the desired variables, and other method is based on the simulation of the dynamics by specific software [6]. We think that this second choice is more suitable than physical modelling because it's a more flexible and scalable solution that shaves a lot of resources (money and time).

The process of simulating the dynamics of a building requires the definition of the following elements:

- The topology of the building: the distribution of spaces and orientation.
- The climate zone: the Spanish regulation sets five winter/summer zones, from E1 (more severe climate zone) to A3 (more soft climate zone).
- Meteorological data for the climate zone and the simulated time period: such as sun radiation, outdoor temperature, wind speed, etc. These data can be taken from meteorological stations placed in a climate zone representative city.
- The building materials that fulfill with local regulations for the chosen climate zone.
- Heating profile: the installed power in each space according to the local regulations and the set point temperature profiles defined by the user.
- Lighting profile: the lighting system of a building produces a sensible heat that influences in its dynamics, so it's necessary to define the operation profile of the lighting system for each space.
- Small power devices profile: as for lighting, small power sources (typically electric appliances) produce heat while they are in use, so they affect the thermal dynamics in the buildings.
- Occupancy profile: occupant heat gains for each space are to be used. Also the metabolic rates (sensible heat gain per adult in W) suitable for the activities held in the space must be considered. So, the timetable of occupancy for each space must be defined.
- Ventilation profile: ventilation and infiltration are important factors in the energy transport in buildings. There are several ventilation specifications for a building; from specifying air change rates for spaces, through specifying space to space air flows, to specifying individual window openings.

Once, these data are defined and/or collected, we can carry out the simulation with the chosen simulation tool and obtain the output data needed for the next step: the air temperature in each space and required power output for each heater to accomplish the heating profiles.

4.2. Extraction of the relevant internal structures

As we can see in the figure 2, CMLHL is instrumental in identifying the internal structure of the data. It was applied to the data set in order to select the features that best described the relationships between the heating conditions and in order to establish whether the dataset was sufficiently informative. The results of applying the method is shown in Figure 2.

.

Having analysed the overall global results, we can see in Figure 2 has classified in two different clusters, then we can say the data set has an interesting internal structure. When the data set was considered sufficiently informative, the third step in the process began. This step accurately and efficiently optimizes the model of the heating system for detecting thermal insulation failures in buildings, through the application of several conventional modelling systems.

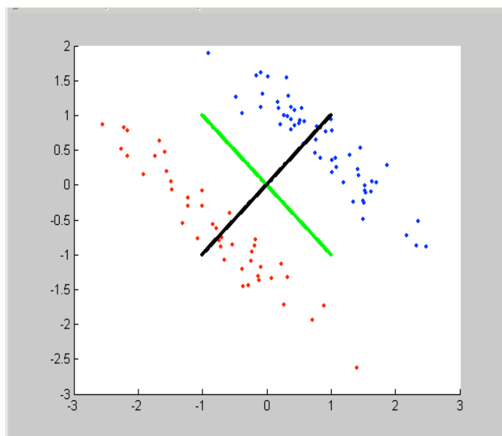

Fig. 2. The CMLHL projection shows the original data set and the directions found by our probabilistic rule with p=2. The first direction is green. The projections of the data are shown in the right diagram in which we see that the first projection totally smears the clusters which the second projection completely separates the clusters.

4.3. System identification to model normal building operation

Once the relevant variables and their transformation have been extracted from the termal dynamics data, then a model to fit the normal building operation should be obtained in order to identify the bias in the room temperatura, which, at the end, is used for failure detection. The different model learning methods used were implemented in Matlab© making use of several toolboxes: the System Identification Toolbox, the Control System Toolbox, the Neural Network Toolbox and the Neural network Based System Identification Toolbox [33]. The experiment followed the identification procedure detailed in Section 3.4: the model structures were analyzed in order to obtain the models that best suite the data set. The Akaike Information Criterion (AIC) was used to obtain the best degree of the model and its delay for each model structure. A total of 70 techniques were carried out to obtain the models, including:

- The frequency response analysis based on the spectrum analysis and the Fourier Fast Transform (FFT) were used to determine the data dynamics,
- The finite impulse response method (FIR) correlation analysis was used to determine the steady state conditions,
- The black-box models synthesis: there were considered up to 31 different combinations of model structure and optimization technique -such as the least squares method, the QR factorization of ARX models, or the recursive normalized gradient algorithm of RARMAX models [19], [27],
- The nonlinear model structures synthesis: there were considered up to 34 different combinations of model structures and optimization technique -such as the Lenvenberg-Marquardt method, the batch version of the back-propagation algorithm or recursive Gauss-Newton method for NNARX, NNFIR, NNARMAX and NNOE models [19], [24], [25],
- Three different residual analysis based on cross correlation were carried out: the residual analysis between the residual $\hat{R}_{\varepsilon}^{N}(\tau)$, between the residual and the input $\hat{R}_{\varepsilon u}^{N}(\tau)$ and the non linear residual correlation $\hat{R}_{\varepsilon^2 u^2}^{N}(\tau)$.

To validate the obtained models, several different indexes have been used. The indexes are well-known and widely used measures in system identification [19], [26], [27]:

- The percentage representation of the estimated model. This index is calculated as the normalized mean error for the one-step prediction (FIT1), for the ten-step prediction (FIT10) and with the ∞-step prediction (FIT). The FIT is known as simulation in classical system identification.
- The graphical representation of the FIT1 –$\hat{y}_1(t\,|\,m)$–, the FIT10 –$\hat{y}_{10}(t\,|\,m)$– or the FIT –$\hat{y}_{\infty}(t\,|\,m)$–.
- The loss function or error function (V). This is the numeric value of the mean square error that is computed with the estimation data set.
- The generalization error value (NSSE). This is the numeric value of the mean square error that is computed with the validation data set.
- The FPE calculated as the average generalization error value computed with the estimation data set.

5. Conclusions and future work

Thermal insulation is essential in the energy efficiency of the heating systems in buildings. The higher the buildings closure the lower the energy losses due insulation failures. This thermal insulation failure detection represents a new challenge in building energy efficiency. In this paper, a new methodology for detecting thermal insulation failures is proposed. The procedure includes the possibility of using the sensors network in the building or the using specialized software in case there is no such network available. Finally, different techniques are applied to obtain the suitable model which will be the responsible of detecting the failures as a bias in the predicted room temperature. Future work include testing the proposal in different climate zones and with different buildings, so a generic methodology could be obtained.

Acknowledgments. We would like to give thanks to Phd. Magnus Nørgaard for the marvellous and freeware Matlab Neural Network Based System Identification Toolbox.

References

[1] Han, J., Lua, L. and Yang, H.: Investigation on the thermal performance of different lightweight roofing structures and its effect on space cooling load, Applied Thermal Engineering, Volume 29 (11-12), pp. 2491-2499, (2009)

[2] Yu, J., Yang, C., Tian, L. and Liao, D.: Evaluation on energy and thermal performance for residential envelopes in hot summer and cold winter zone of China, Applied Energy Volume 86 (10), pp. 1970-1985, (2009)

[3] Villar, J. R., de la Cal, E., Sedano, J.: A Fuzzy Logic Based Efficient Energy Saving Approach for Domestic Heating Systems, Integrated Computer Aided Engineering 16 (2), pp. 151-163 (2009)

[4] Villar, J. R., de la Cal, E. and Sedano, J.; Minimizing energy consumption in heating systems under uncertainty. Lecture Notes in Artificial Intelligence 5271, pp. 583–590, (2008)

[5] Enrique de la Cal, E., Villar, J. R. and Sedano, J.; A thermodynamical model study for an energy saving algorithm, to be published in Lecture Notes in Artificial Intelligence (2009)

[6] Lewis, P. T. and Alexander, D. K.; Htb2: A flexible model for dynamic building simulation. Building and Environment, 1, pp. 7-16, (1990)

[7] E. Corchado and C. Fyfe, Connectionist Techniques for the Identification and Suppression of Interfering Underlying Factors. Int. Journal of Pattern Recognition and Artificial Intelligence, 17(8) (2003), pp. 1447-1466.

[8] J.H. Friedman and J.W. Tukey, Projection Pursuit Algorithm for Exploratory Data-Analysis, IEEE Transactions on Computers, 23(9) (1974), pp. 881-890.

[9] E. Corchado, D. MacDonald and C. Fyfe, Maximum and Minimum Likelihood Hebbian Learning for Exploratory Projection Pursuit. Data Mining and Knowledge Discovery, 8(3) (2004), pp. 203-225.

[10] H.S. Seung, N.D. Socci and D. Lee, The Rectified Gaussian Distribution, Advances in Neural Information Processing Systems, 10 (1998), pp. 350-356.

[11] E. Corchado, D. MacDonald and C. Fyfe, Maximum and Minimum Likelihood Hebbian Learning for Exploratory Projection Pursuit. Data Mining and Knowledge Discovery, 8(3) (2004), pp. 203-225.

[12] C. Fyfe and E. Corchado, Maximum Likelihood Hebbian Rules, Proc. of the 10th European Symposium on Artificial Neural Networks (ESANN 2002), (2002), pp. 143-148.

[13] E. Corchado and C. Fyfe, Connectionist Techniques for the Identification and Suppression of Interfering Underlying Factors. Int. Journal of Pattern Recognition and Artificial Intelligence, 17(8) (2003), pp. 1447-1466.

[14] E. Corchado, Y. Han and C. Fyfe, Structuring Global Responses of Local Filters Using Lateral Connections, Journal of Experimental & Theoretical Artificial Intelligence, 15(4) (2003), pp. 473-487.

[15] L. Ljung, System Identification, Theory for the User, Prentice-Hall, 1999.

[16] M. Nögaard, O.Ravn, N. K. Poulsen, and L. K. Hansen, Neural Networks for Modelling and Control of Dynamic Systems, Springer-Verlag, London, U.K., 2000.

[17] K. Pearson, On Lines and Planes of Closest Fit to Systems of Points in Space, Philosophical Magazine, 2(6) (1901), pp. 559-572.

[18] H. Hotelling, Analysis of a Complex of Statistical Variables Into Principal Components, Journal of Education Psychology, 24 (1933), pp. 417-444.

[19] Ljung, L. (1999). System Identification, Theory for the User. Upper Saddle River, NJ 07458,USA : second edition Prentice-Hall, ISBN 0-13-656695-2.

[20] Cybenko, G.: Aproximation by superpositions of sigmoidal function, Math. Control, Sygnals and System, 2(4), pp. 473-487, pp. 303-314, (1989)

[21] Hornik, K., Stinchcombe, M. and White H.: Multilayer Feedforward Networks are Universal Aproximators, Neural Networks, 2(5), pp. 359-366, (1989).

[22] Hansen, L.K. and Pedersen, M. W. : Controlled Growth of Cascade Correlation Nets, Proc. ICANN 94, Sorrento, Italia, Eds. M. Marinaro & P.G. Morasso, pp. 797-800, (1994).

[23] Hassibi, B. and Stork D.G.: Second Order Derivatives for Network Pruning: Optimal Brain Surgeon, Advances in neural Information Processing System 5, Eds. S.J. Hanson et al., Proceedings of the 1992 Conference, San Mateo, CA. Morgan Kaufmann, pp. 164-171 (1993).

[24] Fletcher, R. :Practical Methods of Optimization, Wiley & Sons, Chichester, 2nd edition, (1987).

[25] Hertz, J., Krogh, A. and Palmer, R.G.: Introduction to the Theory of Neural Computation, Addison-Wesley, (1991).

[26] Nögaard, M., Ravn, O., Poulsen, N. K. and Hansen, L. K.: Neural Networks for Modelling and Control of Dynamic Systems, Springer-Verlag, London, U.K., (2000).

[27] Söderström, T. and Stoica , P. (1989). System identification, Englewood Cliffs, NJ: Prentice Hall

[28] Nelles, O. (2001). Nonlinear System Identification, From Classical Approaches to Neural Networks and Fuzzy Models, Berlin: Springer. ISBN: 978-3-540-67369-9

[29] Haber, R. and Keviczky, L. (1999). Nonlinear System Identification, Input-Output Modeling Approach, Part. 2: Nonlinear System structure Identification, Dorchrecht/Boston/London: Kluwer Academic Publishers.

[30] Stoica, P. and Söderström, T.: A useful parametrization for optimal experimental design, In IEEE Trans. Automatic. Control, AC-27 (1982).

[31] He, X. and Asada, H.: A new method for identifying orders of input-output models for nonlinear dynamic systems, In Proc. Of the American Control Conf., S. F., pp. 2520–2523, California, (1993).

[32] Akaike, H.: Fitting autoregressive models for prediction, Ann. Inst. Stat. Math., 20, pp. 425–439, (1969).

[33] Nørgaard, M.: Neural network Based System Identification Toolbox, Report Technical Tecnico 00-E-891, Department de Automation Technical University of Denmark (2000).