Conventional Methods and AI models for Solving an Industrial Problem

Andrés Bustillo[1], Javier Sedano[2], José Ramón Villar[3], Leticia Curiel[1], Emilio Corchado[1]
1. *Department of Civil Engineering, University of Burgos, Burgos, Spain*
[2,] *Department of Electromechanical Engineering. University of Burgos, Burgos, Spain*
[3,] *Department of Computer Science, University of Oviedo, Spain*

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

Abstract

This study presents a research that identifies and applies unsupervised connectionist models in conjunction with modelling systems, in order to determine optimal conditions to perform laser milling of metallic components. This industrial problem is defined by a data set relayed through sensors situated on a laser milling centre that is a machine-tool used to manufacture high value micro-molds and micro-dies. The results of the study and the application of the connectionist architectures allow the identification, in a second phase, of a model for the milling machine process based on low-order models such as Black Box, which are capable of approximating the optimal form of the model. Finally, it is shown that the most appropriate model to control these industrial tasks is the Box-Jenkins algorithm, which calculates the function of a linear system from its input and output samples.

1. Introduction

Laser milling consists on the controlled evaporation of material due to its interaction with a high-energy pulsed laser beam. A conventional milling machine knows in every moment the amount of material removed (the whole volume of its mill) but that is not so easy for a laser milling machine. Therefore, for the quick industrial spread of this technology is necessary to develop a model that can predict the exact amount of material that each laser pulse is able to remove. This model will allow the control of laser milling with the accuracy that microtools required and, also, the optimization of its manufacture. In this article we have developed such a model using a combination of conventional and AI models.

Unsupervised learning can be used as a preliminary phase before the model is established. It is used to analyse the internal structure of the data sets in order to know whether the data sets are informative enough.

Then, this research presents a two-phase process in order to identify the optimal conditions of industrial laser milling.

2. Study of the Initial Data Set

In this study, Exploratory Projection Pursuit (EPP) [1] connectionist methods as Cooperative Maximum Likelihood Hebbian Learning (CMLHL) [2] and also statistics ones as Principal Component Analysis (PCA) [3], [4], [5], [6] are applied in order to know whether the data is "informative enough". In the worse case, experiments have to be performed again.

EPP [7], [8] provides a linear projection of a data set, but it projects the data onto a set of basic vectors which best reveal the interesting structure in data; interestingness is usually defined in terms of how far the distribution is from the Gaussian distribution [9].

One neural implementation is Maximum Likelihood Hebbian Learning (MLHL) [8], [10]. It identifies interestingness by maximising the probability of the residuals under specific probability density functions that are non-Gaussian. An extended version is the

978-0-7695-3325-4/08 $25.00 © 2008 IEEE
DOI 10.1109/EMS.2008.106

Cooperative Maximum Likelihood Hebbian Learning (CMLHL) [2] model. CMLHL is based on MLHL [8], [10] adding lateral connections [11], [2] which have been derived from the Rectified Gaussian Distribution [9].

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 [12], [13] as:

1. Feed-forward step:

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

2. Lateral activation passing:

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

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 [2], [8], [10] and A a symmetric matrix used to modify the response to the data [2]. The effect of this matrix is based on the relation between the distances separating the output neurons.

3. Modelling the Process

3.1. The Identification Criterion

The identification criterion consists in evaluating which of the group of candidate models is best adapted and best described the group of data collected in the experiment; i.e., given a certain model $M(\theta_*)$ its prediction error may be defined by equation (5); what wish to obtain is a model that complies with the following premise [12]: a good model is one that makes good predictions, and which produces tiny errors when the observed data is applied, i.e., on any one data group Z^t it will calculate the prediction error $\varepsilon(t,\theta)$, equation (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 made as small as possible.

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

The estimated parametrical vector $\hat{\theta}$ that minimizes the error, equation (8), is obtained from the minimization of the error function (6). This is obtained by applying the least-squares criterion for the linear regression, i.e., by applying the quadratic norm $\ell(\varepsilon) = \frac{1}{2}\varepsilon^2$, equation (7).

$$V_N(\theta, Z^N) = \frac{1}{N}\sum_{t=1}^{N} \ell(\varepsilon_F(t,\theta)) \cdot \tag{6}$$

$$V_N(\theta, Z^N) = \frac{1}{N}\sum_{t=1}^{N} \frac{1}{2}(y(t) - \hat{y}(t \mid \theta))^2 \cdot \tag{7}$$

$$\hat{\theta} = \hat{\theta}_N(Z^N) = \arg\min_{\theta \in D_M} V_N(\theta, Z^N) \cdot \tag{8}$$

3.2. Black-box Models

The methodology of black-box structures 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 between a deterministic and a stochastic part, equation (9); 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.

$$y(t) = G(q^{-1})u(t) + H(q^{-1})e(t) \, . \tag{9}$$

The structure of a black-box model depends on the way in which the noise is modelled $H(q^{-1})$; thus, if this value is 1, then the OE (Output Error) model is applicable; whereas if it is different from zero a great range of models are applicable; one of the most common is the BJ (Box Jenkins). This structure may be represented in the form of a general model, where $B(q^{-1})$ is a polynomial of grade n_b, which can incorporate pure delay n_k in the inputs, and $A(q^{-1})$, $C(q^{-1})$, $D(q^{-1})$ y $F(q^{-1})$ are autroregressive polynomials ordered as n_a, n_c, n_d, n_f, respectively (10). In the same way, it is possible to use a predictor expression, for the on-step prediction ahead of the output $\hat{y}(t \mid \theta)$ (11). In Table 1, the generalized polynomial expressions are presented, as well as those that represent the polynomials used in the case of each particular model.

$$A(q^{-1})y(t) = q^{-n_k}\frac{B(q^{-1})}{F(q^{-1})}u(t) + \frac{C(q^{-1})}{D(q^{-1})}e(t) \tag{10}$$

$$\hat{y}(t \mid \theta) = \frac{D(q^{-1})B(q^{-1})}{C(q^{-1})F(q^{-1})}u(t) + \left[1 - \frac{D(q^{-1})A(q^{-1})}{C(q^{-1})}\right]y(t) \tag{11}$$

3.3. Procedure for Modelling the Laser Milling

The identification procedure followed to obtain a parametrized model *M*, selected as the best of those that model the laser milling characteristics based on the variable measurements, is carried out in accordance with two fundamental patterns: a first pre-analytical and then an analytical stage that assists with the determination of the parameters in the identification process and the model estimation. The pre-analysis test is run to establish the identification techniques [12], [13], [14], [15], [16], [17], the selection of the model structure and its order estimation [18], [19], the identification criterion and search methods that minimize it and the specific parametrical selection for each type of model structure.

A second validation stage 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) estimated as explained by Akaike [20] and the graphical comparison between desired outputs and the outcome of the models through simulation one (or k) steps before.

4. An Industrial Problem: Laser Milling of Metallic Microtools

The purpose of this work presented is the study of the best conditions for laser milling of Alum 7010, often used in the mold industry, using a commercial Nd:YAG laser with a pulse length of 10μs [21]. Three parameters of the laser process can be controlled: laser power (u1), laser milling speed (u2) and laser pulse frequency (u3). The laser is integrated in a laser milling centre (DMG Lasertec 40).

To simplify the industrial problem a test piece has been designed. This test piece consists on an inverted truncated pyramid profile that should be laser milled on a flat metallic piece of Alum 7010. As the laser parameters are not know for Alum 7010, both parameters will show an error called angle error (y1) and depth error (y2). We have applied different modelling systems to achieve the optimal conditions, those that provide the minimum errors of laser milling.

The experiment design has been performed using a Taguchi L25 with 3 input parameters and 5 levels to include all the range of the laser milling process that can be change from the control of the machine, and therefore that can be modify by the end user of this machine. Table 2 summarizes input and output variables of the experiment. The experiment was performed on the test piece already described. After the laser milling, actual inverted pyramid depth and wall angle were measured by means of proper optical means.

Table 1. Black-box models structures.

Polynomials in (10)	Polynomials used in equation 10	Name of model structure
$A(q^{-1})=1+a_1(q^{-1})+a_2(q^{-2})+\cdots+a_{n_a}(q^{-n_a})$ $B(q^{-1})=b_1(q^{-1})+b_2(q^{-2})+\cdots+b_{n_b}(q^{-n_b})$ $C(q^{-1})=1+c_1(q^{-1})+c_2(q^{-2})+\cdots+c_{n_c}(q^{-n_c})$ $D(q^{-1})=1+d_1(q^{-1})+d_2(q^{-2})+\cdots+d_{n_d}(q^{-n_d})$ $F(q^{-1})=1+f_1(q^{-1})+f_2(q^{-2})+\cdots+f_{n_f}(q^{-n_f})$	BF BFCD	OE BJ

Table 2. Variables, units and values used during the experiments.

All values are common to this laser milling process. Output y(t), Input u(t).

Variable (Units)	Range
Angle error of the test piece, $y_1(t)$	
Depth error on the test piece, $y_2(t)$	
Laser power in percent of the maximum power performed by the laser (%), $u_1(t)$.	20-100
Speed (mm/s), $u_2(t)$.	200-800
Laser pulse frequency (kHz), $u_3(t)$.	20-100

4.1. Application of the two Phases of the Modelling System

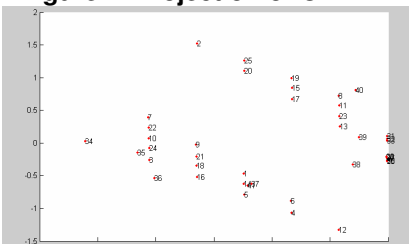
The experiments have been organized into two phases.

Phase 1. Initial identification of the internal structure of the data set. Application of several unsupervised neural models.

Phase 2. Final identification of the model that best defines the dynamic of the laser milling process.

4.1.a. Phase 1. Figure 1 shows the results obtained by CMLHL. We can see that it has identified a clear structure, meaning that the data analysed is informative enough.

Figure 1. Projection of CMLHL



We have also applied PCA. Both methods have identified a clear internal structure based on an initial classification but CMLHL (Figure 1) provides a sparser representation than PCA.

4.1.b. Phase 2. Modelling Laser Milling by means of Classical Models.

Figures 2 and 3 show the results of the output $y_1(t)$, angle error, and $y_2(t)$, depth error, respectively for different models, in relation to the polynomial order and the delay in the inputs. To obtain the maximum accuracy, we have considered various delays for all inputs and various polynomials order, which is [na, nb_1, nb_2, nb_3, nc, nd, nf, nk_1, nk_2, nk_3], in accordance with the structure of the models that have been used; see Table 1. X-axis of Figures 2 and 3 shows the number of samples used in the validation of the model, whereas the Y-axis represents the range of output variable.

Table 3 shows a comparison of the qualities of estimation and prediction of the models obtained, as a function of the model, the estimation method, and the indicators, which are defined as follows:

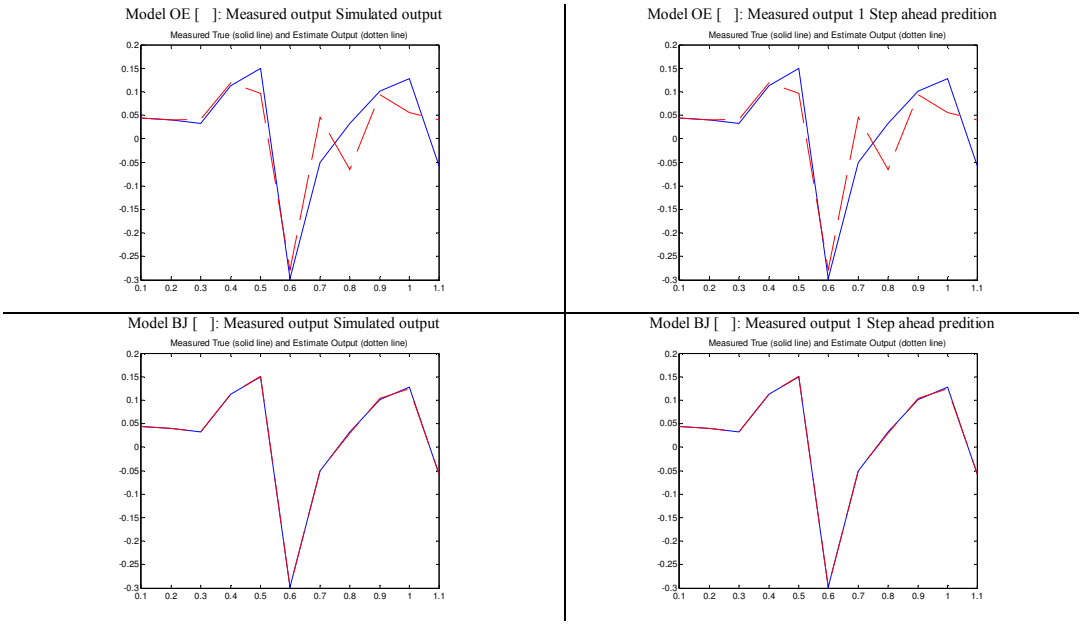
The percentage representation of the estimated model (expressed in "%") in relation to the true system. This is the numeric value of the normalized mean error that is computed with one-step prediction (FIT1), with ten-step prediction (FIT10), or by means of simulation (FIT). Also, the graphical representation of true system output and the one-step prediction $\hat{y}_1(t \mid m)$, the ten-step prediction $\hat{y}_{10}(t \mid m)$,or the model simulation $\hat{y}_{\infty}(t \mid m)$.

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 average generalization error value (FPE). This is the numeric value of the criterion FPE that is computed with the estimation data set.

Figure 2. Representation of measured output, simulated output and one-step-ahead prediction for two black–box models.



The results showed in Figure 2 correspond to the output y1(t), angle error. The model generated by the OE model is shown in the upper row. On the left, measured ouput vs. simulated output, on the right, measured output vs. one-step-ahead prediction. The BJ model is shown in row 2. The validation data set was not used for the estimation of the model. The order of the structure of the model is [2 3 1 1 3 2 2 2 1 1] according to the model type. The solid line represents true measurements and the dotted line represents estimated output.

Figure 3. Representation of measured output, simulated output and one-step-ahead prediction for two black–box models.

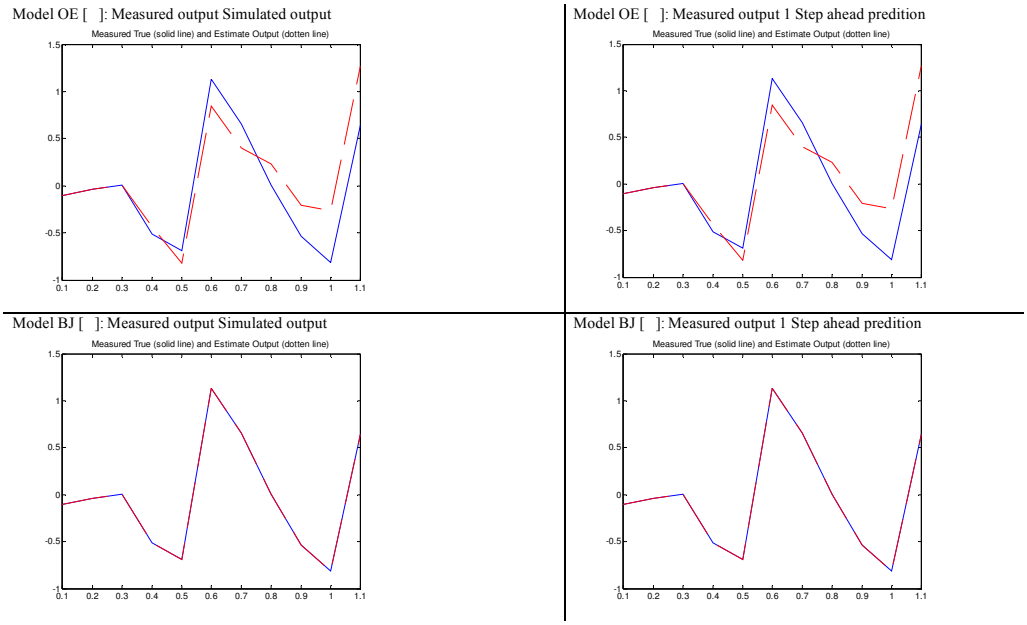


Figure 3 presents the results corresponding to the output y2(t), depth error. The model generated by the OE model is shown in the upper row. On the left, measured output vs. simulated output, on the right, measured output vs. one-step-ahead prediction. The BJ model is presented in row 2. The validation data set was not used for the estimation of the model. The order of the structure of the model is [2 1 3 1 3 2 2 1 3 1] according to the model type. The solid line represents true measurements and the dotted line represents estimated output.

It may be seen from Fig. 2 and 3 that the BJ model is capable of simulating and predicting the behaviour of the errors of the laser milling process as it meets the indicators and is capable of modelling better than 98% of the true measurements. This is also evident from Table 3. Table 4 shows the function and the parameters that define the laser milling process, on the basis of the BJ model. The tests were performed using Matlab and the System Identification Toolbox. Similar results were found for the depth error.

Table 3. Indicator values for OE and BJ models

	Indicators and order [na, nb_1, nb_2, nb_3, nc, nd, nf, nk_1, nk_2, nk_3]					
	Angle error			***Depth error***		
Model	[1 2 1 1 3 2 1 2 2 2]	[2 2 2 1 3 2 2 2 2 1]	[2 3 1 1 3 2 2 2 1 1]	[1 2 1 1 3 2 1 2 2 2]	[2 2 2 1 3 2 2 2 2 1]	[2 1 3 1 3 2 2 1 3 1]
Black-box model, OE model is estimated using the prediction error method	FIT:12.87% FIT1:12.87% FIT10:12.8% V: 0.020 FPE:0.047 NSSE:0.010	FIT:16.6% FIT1 :16.6% FIT10:16.6% V: 0.0124 FPE:0.049 NSSE:0.099	FIT:51.76% FIT1 :51.76% FIT10:51.7% V: 0.021 FPE:0.088 NSSE:0.033	FIT:5.21% FIT1:5.21% FIT10:5.21% V: 0.362 FPE:0.844 NSSE:0.31	FIT:61.1% FIT1:61.1% FIT10:61.1% V: 0.29 FPE:1.18 NSSE:0.052	FIT:48.7% FIT1:48.7% FIT10:48.7% V: 0.297 FPE:1.233 NSSE:0.0926
Black-box model, BJ model is estimated using the prediction error method.	FIT:38.3% FIT1:41.88% FIT10:39.3% V: 0.023 FPE:0.106 NSSE:0.0048	FIT:41.65% FIT1:55.21% FIT10:27.2% V: 0.0066 FPE:0.073 NSSE:0.0029	FIT:**98.9**% FIT1:**98.62**% FIT10:98.9% V: 0.0111 FPE:0.155 NSSE:**2.7e-6**	FIT:51.55% FIT1:54.82% FIT10:44.5% V: 0.196 FPE:0.875 NSSE:0.071	FIT:68.12% FIT1:63.02% FIT10:58.3% V: 0.138 FPE:1.527 NSSE:0.047	FIT:**99.8**% FIT1: **99.8**% FIT10:99.8% V: 0.023 FPE:0.45 NSSE:**2e-27**

Table 4. Function and parameters that represent the behavior for the angle error.

Model BJ [2 3 1 1 3 2 2 2 1 1]	
$y_1(t) = q^{-n_k}\frac{B_1(q^{-1})}{F_1(q^{-1})}u_1(t) + q^{-n_k}\frac{B_2(q^{-1})}{F_2(q^{-1})}u_2(t) + q^{-n_k}\frac{B_3(q^{-1})}{F_3(q^{-1})}u_3(t) + \frac{C(q^{-1})}{D(q^{-1})}e(t)$	
Parameters and polynomials.	
$B_1(q) = -0.001573\ q^{-2} - 0.001705\ q^{-3} - 0.001045\ q^{-4}$	$D(q) = 1 - 1.801\ q^{-1} + 0.9602\ q^{-2}$

$B_2(q) = -4.85e\text{-}5\ q^{-1}$	$F_1(q) = 1 + 0.4735\ q^{-1} + 0.1797\ q^{-2}$
$B_3(q) = 0.001323\ q^{-1}$	$F_2(q) = 1 - 0.4525\ q^{-1} + 0.8133\ q^{-2}$
$C(q) = 1 - 1.548\ q^{-1} + 0.5537\ q^{-2} + 0.2632\ q^{-3}$	$F_3(q) = 1 - 0.5514\ q^{-1} + 0.09725\ q^{-2}$
	$e(t)$ is white noise signal with variance 0.0834

5. Conclusions and Futures lines of Work

This research has presented an investigation to identify the most appropriate modelling system to solve a real-life industrial problem such as laser milling of metallic components. Several methods were investigated to achieve the best practical solution to this interesting problem. The paper shows why, from among the classical models, the BJ model is the one that is best adapted to this case in terms of identifying the best conditions and predicting future circumstances.

The novelty of the paper lies in the use of a two-phase model for modelling of the laser milling process: a first phase, which applies EPP connectionist processes to establish whether the data set is "informative enough". As a consequence, the first phase eliminates one of the problems that these identification systems have, which is that of not knowing beforehand if the experiment that generates the data group can be considered acceptable and will present sufficient information in order to identify the overall nature of the problem. Future work will be focus on the study and application of another kind of materials, such us copper or steel.

Acknowledgments. This work has been possible thanks to the disinterested help of ASCAMM Centro Tecnológico (http://www.ascamm.es) by providing the laser milling data and has performed all laser tests. The authors would like to thank specially Mr. Pol Palouzie and Mr. Javier Diaz. This research has been partially supported by the project BU006A08.- JCyL.

References

[1] P. Diaconis and D. Freedman, Asymptotics of Graphical Projections, The Annals of Statistics, 12(3), pp. 793-815, 1984.

[2] 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.

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

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

[5] C. Fyfe, PCA Properties of Interneurons: from Neurobiology to Real World Computing, Proc. of the Int. Conf. on Artificial Neural Networks, ICANN 1993, Verlag, S., 93 (1993), pp. 183-188.

[6] E. Oja, A Simplified Neuron Model as a Principal Component Analyzer, Journal of Mathematical Biology, 15(3) (1982), pp. 267-273.

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

[8] 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.

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

[10] 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.

[11] 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.

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

[13] 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.

[14] T. Söderström and P. Stoica , System identification, Prentice Hall, 1989.

[15] O. Nelles, Nonlinear System Identification, From Classical Approaches to Neural Networks and Fuzzy Models, Springer, 2001.

[16] R. Haber and L. Keviczky, Nonlinear System Identification, Input-Output Modeling Approach, Part2: Nonlinear System structure Identification, Kluwer Academic Publishers, 1999.

[17] R. Haber and L. Keviczky, Nonlinear System Identification, Input-Output Modeling Approach, Part1: Nonlinear System Parameter Estimation, Kluwer Academic Publishers, 1999.

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

[19] X. He and H. Asada, 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.

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

[21] J. Diaz, P. Palouzie and G. Arias, Caracterización de materiales para micromecanizado mediante el uso de un láser Nd:YAG, Actas del IV Taller Nacional Procesado Materiales con Laser, Valencia, Spain, 2007.