Taylor & Francis
Taylor & Francis Group

The application of a two-step AI model to an automated pneumatic drilling process

Q1

Javier Sedano[a], Emilio Santiago Corchado[b]*, Leticia Curiel Herrera[b],
José Ramón Villar Flecha[c] and Pedro Miguel Bravo[b]

[a]*Department of Electromechanical Engineering, University of Burgos, Burgos, Spain;* [b]*Department of
Civil Engineering, University of Burgos, Burgos, Spain;* [c]*Department of Computer Science,
University of Oviedo, Gijón, Spain*

Real-world processes may be improved through a combination of artificial intelligence and identification techniques. This work presents a multidisciplinary study that identifies and applies unsupervised connectionist models in conjunction with modelling systems. This particular industrial problem is defined by a data set relayed through sensors situated on a robotic drill used in the construction of industrial storage centres. The first step entails determination of the most relevant structures in the data set with the application of the connectionist architectures. The second step combines the results of the first one to identify a model for the optimal working conditions of the drilling robot that is based on low-order models such as black box that approximate 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.

Keywords: unsupervised learning; exploratory projection pursuit; black-box models

AMS 2000 Mathematics Subject Classification Codes: 68T99; 68U99

1. Introduction

The storage of goods under optimal conditions is an important industrial need that often calls for the modernization or the replacement of existing facilities. Large storage centres often need to be built, such as auto-carrier storage facilities. The drilling of reinforced concrete slabs, needed to install the racks in these storage centres, is generally a manual task performed by the workforce. However, this process may be automated, thereby reducing overall construction costs and shortening assembly times, through the use of a drilling robot that is able to perform the aforementioned operations and to use the drilling tool more economically and more efficiently. The robot must be able to calibrate the position of the drill bits and their diameters, to determine which operations are required (such as whether to take in, to filter or to recirculate the coolant while drilling), and to test the tool conditions when automatically changing the drill bit and their diameters. Performing these tasks

*Corresponding author. Email: escorchado@ubu.es

51 regularly improves assembly quality, and decreases drilling time and wear, all of which adds to
52 the duration of the drill bit and reduces corrective actions, saving both time and money. These
53 very real types of industrial problems may be analysed by a combination of conventional and
54 artificial intelligence (AI) models.
55   Unsupervised learning can be used initially, in the first step, prior to the establishment of the
56 model, to analyse the internal structure of the data sets. Consequentially, we need to know whether
57 the data sets are relevant, i.e., to know whether discrimination between any two different models
58 in the set is in fact possible. Exploratory projection pursuit (EPP) [5] is a statistical method aimed
59 at solving the difficult problem of identifying structures in high-dimensional data. In EPP, most
60 projections of high-dimensional data onto arbitrary lines through most multi-dimensional data
61 give almost Gaussian distributions [5], which is what usually defines a relevant structure.
62   Unsupervised learning, and specifically EPP, can be used in conjunction with classical identi-
63 fication techniques to define a model that can capture the dynamics of the real process. The
64 former is used to extract the relevant structures and relationships between features, in order to
65 guarantee that the data set obtained by the sensors during the experiments is sufficiently infor-
66 mative. Classical identification techniques model the operating conditions, in order to select the
67 optimum working parameters, which reduce the operating costs of the robot when performing the
68 aforementioned tasks.
69   This research presents a two-step procedure designed to identify the optimal conditions of a
70 pneumatic drill. This paper is organized as follows. Section 2 introduces the neural techniques
71 used to analyse the data in order to extract the relevant internal structures, which represents the
72 first step in the modelling process. Section 3 presents the classical identification techniques used
73 in the system modelling (second step). Section 4 provides details on the problem, the data set
74 description and an analysis and comparison of the best models and results, following which the
75 final section presents the conclusions and future research work.

2. EPP extraction of relevant internal structures

80 Principal component analysis (PCA) originated in work by Pearson [18], and independently by
81 Hotelling [12]. It describes multivariate data set variations in terms of uncorrelated variables, each
82 of which is a linear combination of the original variables. Its main goal is to derive new variables,
83 in decreasing order of importance, which are linear combinations of the original variables and
84 are uncorrelated with each other. PCA is a well-known technique that can be implemented by a
85 number of connectionist models [7,17]. It aims to find the orthogonal basis which maximizes the
86 variance of the data for a given dimensionality of basis. It has been the most frequently reported
87 linear operation involving unsupervised learning for data compression and feature selection.
88   The standard statistical EPP method [4,6] provides a linear projection of a data set. The data
89 projections make use of a set of basis vectors which best reveal the relevant structures of the data.
90 The relevancy is measured as interestingness, which is usually defined in terms of how far the
91 distribution is from the Gaussian distribution [19].
92   One neural implementation of EPP is maximum likelihood Hebbian learning (MLHL) [4,8].
93 The MLHL-based method has been widely used in the field of pattern recognition [2–4,8] as an
94 extension of PCA [7,17]. It identifies interestingness [4,6,8] by maximizing the probability of the
95 residuals using specific probability density functions that are non-Gaussian, which analyse the
96 fourth-order statistic or the kurtosis.
97   An extended version of this model is the cooperative MLHL (CMLHL) [2] model. CMLHL,
98 which is based on MLHL [4,8], adds lateral connections [2,3] which have been derived from the
99 rectified Gaussian distribution [19]. It is able to identify the independent factors of a data set, but
100 does so in a way that captures some type of global ordering in the data set.

Consider an N-dimensional input vector (x), an M-dimensional output vector (y) and a weight matrix W, where the element W_{ij} represents the relationship between input x_j and output y_i. Then, as shown in [2,3], CMLHL can be performed as a four-step procedure: first, the outputs are calculated, see Equation (1), then the prediction is obtained Equation (2). Finally, the generated error, calculated with Equation (3), is used to train the neural network by means of Equation (4).

$$\text{Feed-forward step} \qquad y_i = \sum_{j=1}^{N} W_{ij} x_j, \forall i \tag{1}$$

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

$$\text{Feedback step} \qquad e_j = x_j - \sum_{i=1}^{M} W_{ij} y_i, \forall j \tag{3}$$

$$\text{Weight adjustment step} \qquad \Delta W_{ij} = \eta \cdot y_i \cdot \text{sign}(e_j)|e_j|^{p-1} \tag{4}$$

Where t is the temperature, $[\ ]^+$ is necessary to ensure that the **y**-values remain within the positive quadrant, η is the learning rate, τ is the "strength" of the lateral connections, b is the bias parameter, p is a parameter related to the energy function [2,4,8] and A is 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. System modelling using conventional identification algorithms

3.1 *The identification criterion*

The identification criterion consists in evaluating which group of candidate models is best adapted to the data set prepared for the experiment and which best describes it; i.e., given a certain model $M(\theta_*)$ its prediction error may be defined by Equation (5). The aim is to obtain a model that complies with the following premise [13]: a good model is one that makes good predictions, and which produces small errors when it is applied to the observed data, i.e., on any one data set 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 $t = 1, 2, 3, \ldots N$, is minimized.

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

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

$$V_N(\theta, Z^N) = \frac{1}{N} \sum_{t=1}^{N} \frac{1}{2}(y(t) - \hat{y}(t|\theta))^2 \tag{6}$$

$$\hat{\theta} = \hat{\theta}_N(Z^N) = \arg\min_{\theta \in D_M} V_N(\theta, Z^N) \tag{7}$$

Black-box structures [13] have the advantage of only requiring very few explicit assumptions with regard to the pattern that is 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 a combination of

deterministic and stochastic models, as in Equation (8), in which, $u(t)$ 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.

$$y(t) = G(q^{-1})u(t) + H(q^{-1})e(t) \tag{8}$$

The structure of a black-box model depends on how the noise affects the model [13]: $H(q^{-1})$. Thus, if this term is 1, then the FIR (finite impulse response) and OE (output error) models are applicable; whereas a great range of models are applicable if it differs from zero: ARX (autoregressive with external input), ARMAX (autoregressive moving average with external input), BJ (Box–Jenkins) and ARMA (autoregressive moving average) being the most common.

The structure may be represented in the form of a general model Equations (9) and (3.1), where $B(q-1)$ is a polynomial of degree n_b, which can incorporate pure delay nk in the inputs, and $A(q-1)$, $C(q-1)$, $D(q-1)$ and $F(q-1)$ are autoregressive polynomials of degree n_a, n_c, n_d and n_f, respectively. It is also possible to use a predictor expression, for the on-step prediction ahead of the output $\hat{y}(t|\theta)$ (10). The values of n_a, n_b, n_c, n_d, n_f and n_k are parameterized.

$$A(q^{-1})y(t) = q^{-n_k}\frac{B(q^{-1})}{F(q^{-1})}u(t) + \frac{C(q^{-1})}{D(q^{-1})}e(t) \tag{9}$$

$$\hat{y}(t|\theta) = \frac{D(q^{-1})B(q^{-1})}{C(q^{-1})F(q^{-1})}u(t) + \left[1 - \frac{D(q^{-1})A(q^{-1})}{C(q^{-1})}\right]y(t) \tag{10}$$

$$A(q^{-1}) = \sum_{i=1}^{n_a} 1 + a_i(q^{-1}), \quad B(q^{-1}) = \sum_{i=1}^{n_b} b_i(q^{-i}), \quad C(q^{-1}) = \sum_{i=1}^{n_c} 1 + c_i(q^{-1}),$$

$$D(q^{-1}) = \sum_{i=1}^{n_d} 1 + d_i(q^{-1}), \quad F(q^{-1}) = \sum_{i=1}^{n_f} 1 + f_i(q^{-1}) \tag{11}$$

3.2 *Modelling procedure for optimal performance of the drilling robot*

The best model must be selected in order to estimate the optimal settings of the drilling robot. The identification procedure includes establishing the identification techniques [9,10,13,15,16,20], the selection of the model structure, the estimation of the suitable polynomial degree [11,21], the identification criterion, and the optimization techniques in order to generate the final model.

Furthermore, the identification procedure includes a validation stage, which ensures that the selected model meets the necessary conditions for estimation and prediction. Three tests were performed to validate the model: 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 proposed by Akaike [1]; and finally, the graphical comparison between desired outputs and the outcome of the models through the simulation of one (or k) previous step(s).

4. **Real-world application: the drilling robot industrial problem**

This section presents the real-world application solved in this study. In this case, a company that constructs storage centres needs reinforced concrete slabs to be drilled in order to erect the facilities. A pneumatic drilling robot fitted diamond drill bits is used to drill holes in the concrete slabs. The problem lies in how to select the most appropriate operating conditions in each case,

201 so as to maximize the lifetime of the diamond drill bits and to reduce the operating costs. Up until
202 the present, company technicians have manually fixed the operational settings.
203   Several steps must be carried out, in order to model the optimal conditions for the robot.
204 First, a representative real-world data set must be generated by using sensors. Second, in the pre-
205 processing stage, the relevant selection of internal structures must be made. Finally, the model
206 must be generated on the basis of the most important variables and the relationships that have
207 been identified. Each of these steps is detailed as follows.

4.1 *Generation of the data set*

The first task involved the generation of the data set that reflected the most appropriate operating
conditions. Several data gathering experiments were performed using the drilling robot to drill
into reinforced concrete slabs. The number of samples of the experimental data set is limited
by economic constraints due to the high cost of the diamond drill bits. The guidelines on cutting
reinforced concrete considered in [14] were followed for the design and capture of the experimental
data, which were gathered from both the drilling tool and the drilling material, in order to identify
the optimal conditions related to aspects such as:

219 • A drill bit that only engages a concrete slab.
220 • A drill bit that engages a concrete slab and a small portion of a steel rebar.
221 • A drill bit that engages a steel rebar and a concrete slab.

The final data set included several input variables: wear on the drill bit, the force applied by the
drilling robot, and so on. The data set was gathered by drilling reinforced concrete slabs. Eighty
samples were taken, which is a relatively small number due to aforementioned costs.

4.2 *Extraction of the relevant internal structures*

229 As detailed in Section 2, PCA and CMLHL, which were both applied to this real-life problem,
230 are instrumental in identifying the internal structure of the data. They were applied to the data set
231 in order to select the features that best described the relationships between the robotic conditions
232 and in order to establish whether the data set was sufficiently informative. The results of applying
233 both methods (PCA to the left, and CMLHL to the right) are shown in Figure 1.
234   As may be seen, both methods have identified four different clusters and a clear internal
235 structure based on an initial classification relating to the speed factor. It may be affirmed that
236 CMLHL (Figure 1 – right-hand figure) provides a sparser representation than PCA (Figure 1 –
237 left-hand figure), due to the combined use of the MLHL-based method and the application of
238 lateral connections.
239   Having analysed the overall global results, an interesting internal structure was identified from
240 the existence of four clearly different groups (Figure 1 – left-hand figure). When the data set was
241 considered sufficiently informative, the second step in the process began. This step accurately
242 and efficiently optimizes the model of the robotic drilling system, through the application of
243 several conventional modelling systems. In the first phase, it was evident from the original data
244 set that only four variables were relevant when estimating wear on the tool: the force applied by
245 the drilling robot, the rotation speeds of the drill bits, the flow of coolant to the bit head and the
246 drilling time. These four variables were used as input in the second step of the process.
247   Both methods were able to perform an initial classification based on data relating mainly to the
248 speed factor. Evidently, CMLHL (Figure 1 – right-hand figure) provides sparser representation
249 than PCA (Figure 1 – left-hand figure) as the four more expansive clusters revealed by CMLHL
250 provide more visual information. After studying the internal structure of the four CMLHL clusters,
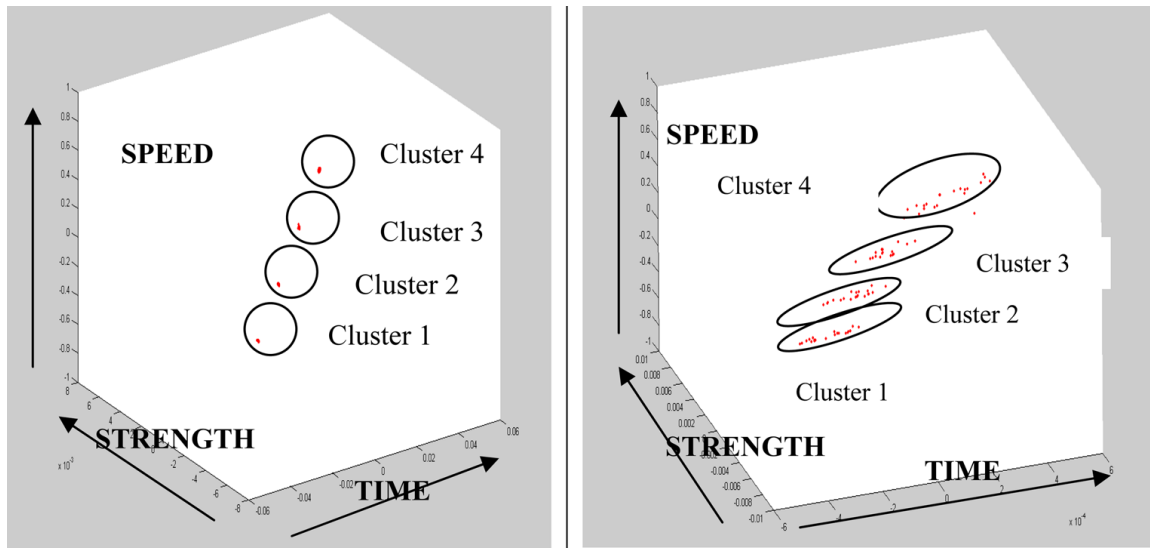
Figure 1. The CMLHL projection on the right shows a wider data spread than the PCA projection on the left. The three axes of this 3D representation are the "the strength applied to the drilling robot", "the drilling time used" and "the speed of the bits of the drilling robot".

it was evident that the classification was mainly based on the speed factor, and also to a lesser extent on the strength applied and the drilling time.

4.3 *System identification to model the optimal conditions for the drilling robot*

The different training methods for the model were implemented in Matlab© making use of tool-boxes – function libraries for Matlab – the System Identification Toolbox and the Control System Toolbox. The experiment followed the identification procedure detailed in Section 3.2: the model structures were analyzed in order to obtain the models that best suited the data set. The Akaike information criterion (AIC) was used to obtain the best model and the delay for each model structure. A total of 36 techniques were carried out to obtain the models, which included:

- The frequency response analysis based on the spectrum analysis, and the fast Fourier transform (FFT) to determine the data dynamics.
- The FIR method correlation analysis, which was used to determine the steady-state conditions.
- The black-box models synthesis: up to 31 different combinations of model structures and optimization techniques were considered – such as least squares, the QR factorization method for ARX models and the RARMAX recursive normalized gradient algorithm [13,20].
- Three different residual analyses based on cross-correlation were performed: 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 $\hat{R}_{\varepsilon^2 u^2}^N(\tau)$.

Several well known and widely used indexes in system identification [13,16,20] were applied to validate the models:

- 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): the numeric value of the mean square error that is computed with the estimation data set.

- The generalization error value (NSSE): the numeric value of the mean square error that is computed with the validation data set. **Q2**
- The FPE: the average generalization error value computed with the estimation data set.

The graphical representations of the $\hat{y}_1(t|m)$ and $\hat{y}_\infty(t|m)$ for different models are shown in Figure 2. The x-axis shows the number of samples used in the validation of the model and the y-axis represents the normalized output variable range. That is, the normalized wear of the tool. The validation data set includes 20 samples, which are distributed in 0.1 s intervals.
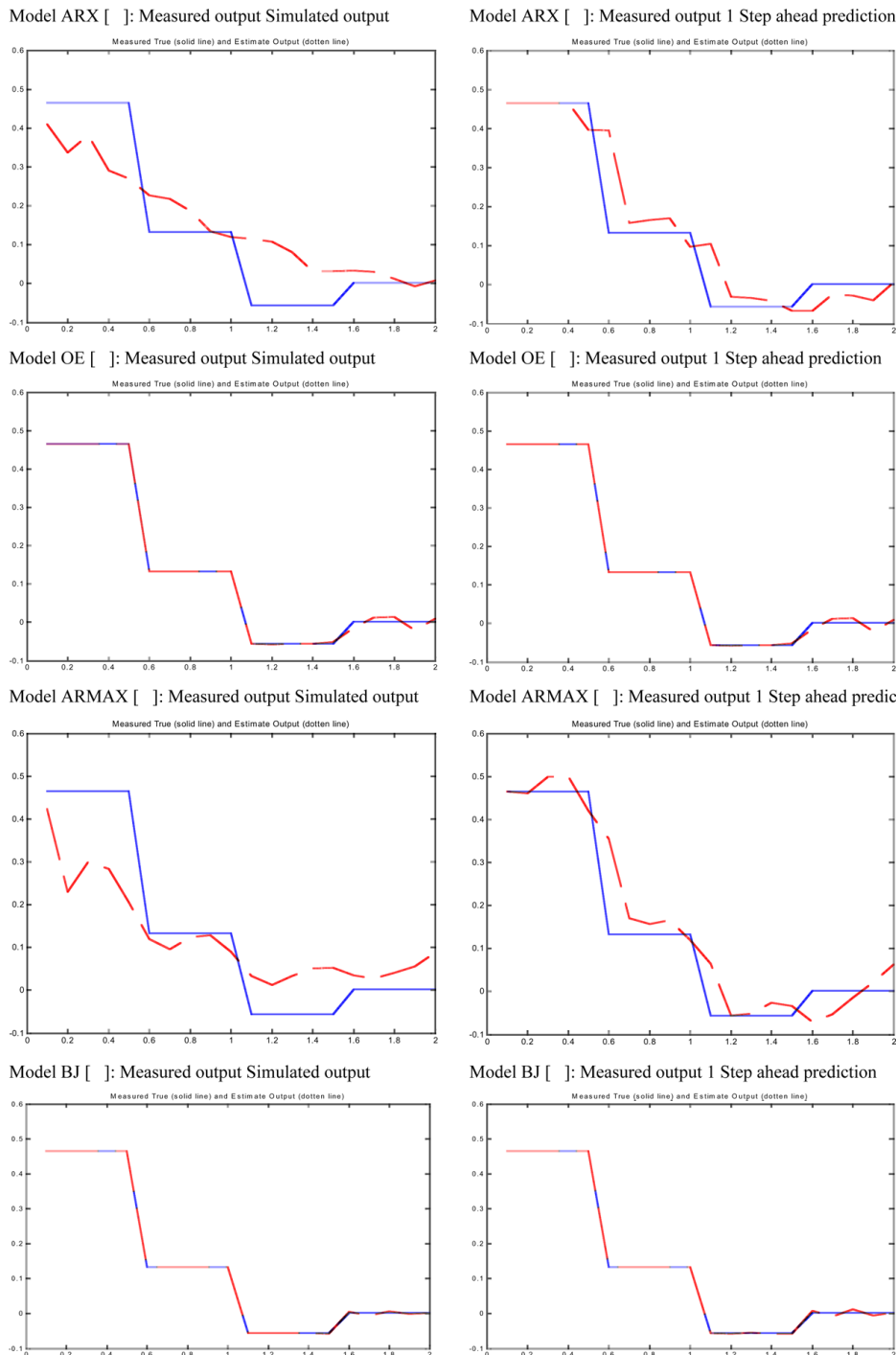
Figure 2. Representation of system output, simulated output and one-step-ahead prediction for four black-box models. The order of the structure of the model is $n_a = 4, n_b = 2, n_c = 3, n_d = 2, n_f = 4, n_k = 3$ [4, 2, 3, 2, 4, 3] according to the model type. The solid line represents the output variable and the dashed line represents estimated output.

Table 1. Indicator values for several proposed models.

Model	Indexes
Black-box OE model with $n_b = 2$, $n_f = 4$, $n_k = 2$, which is estimated using the prediction error method. The model selection is based on the best AIC (the structure that minimizes the AIC).	FIT: 91.56%; FIT1: 91.56% FIT10: 91.5%; V: 0.04 FPE: 0.02; NSSE: 0.0003
Black-box OE model with $n_b = 2$, $n_f = 4$, $n_k = 3$, which is estimated using the prediction error method. The model selection is based on the best AIC (the structure that minimizes the AIC).	FIT: 96.33%; FIT1: 96.33% FIT10: 96.33%; V: 0.05 FPE: 0.02; NSSE: 0.00005
Black-box BJ model with $n_b = 2$, $n_c = 3$, $n_d = 2$, $n_f = 4$, $n_k = 2$, which is estimated using the prediction error method. The model selection is based on the best AIC (the structure that minimizes the AIC).	FIT: 94.56%; FIT1: 94.69% FIT10: 94.1%; V: 0.005 FPE: 0.009; NSSE: 0.0001
Black-box BJ model with $n_b = 2$, $n_c = 3$, $n_d = 2$, $n_f = 4$, $n_k = 3$, which is estimated using the prediction error method. The model selection is based on the best AIC (the structure that minimizes the AIC).	FIT: 97.7%; FIT1: 98.98% FIT10: 98.98%; V: 0.02 FPE: 0.02; NSSE: 0.00001

Table 2. Function and parameters that represent the behaviour of the drilling robot. The polynomial degrees of the BJ model are $n_b = 2$, $n_c = 3$, $n_d = 2$, $n_f = 4$, $n_k = 3$ [2, 3, 2, 4, 3].

Parameters and polynomials

$B_1(q) = -0.04883\,q^{-3} - 0.01538\,q^{-4}$ $D(q) = 1 - 0.4843\,q^{-1} - 0.3384\,q^{-2}$
$B_2(q) = -0.000169\,q^{-3} + 9.223\mathrm{e} - 005\,q^{-4}$ $F_1(q) = 1 - 0.4102\,q^{-1} - 0.1256\,q^{-2} - 0.03927\,q^{-3} + 0.04401\,q^{-4}$
$B_3(q) = -0.145\,q^{-3} + 0.1575\,q^{-4}$ $F_2(q) = 1 - 1.158\,q^{-1} + 0.3967\,q^{-2} + 0.01603\,q^{-3} - 0.1476\,q^{-4}$
$B_4(q) = -0.0001203\,q^{-3} + 0.0001846\,q^{-4}$ $F_3(q) = 1 - 1.241\,q^{-1} + 0.412\,q^{-2} - 0.006801\,q^{-3} - 0.1674\,q^{-4}$
$C(q) = 1 + 0.6335\,q^{-1} + 0.5132\,q^{-2}$ $F_4(q) = 1 - 1.556\,q^{-1} + 0.6325\,q^{-2} - 0.3091\,q^{-3} + 0.3692\,q^{-4}$
$\quad + 0.3435\,q^{-3}$ $e(t)$ is white noise signal with variance 0.11

From the above graphs, it may be concluded that both the OE and the BJ models are capable of simulating and predicting the behaviour of the drilling robot, as they both comply with the indicators and, respectively, are capable of modelling 96.33% and 97.70% of the true measurements. Table 1 shows a comparison of the estimations and predictions for the best models, as a function of the model, the estimation method and the indexes. It was also found that the longer the system delay n_k, the higher the percentage representation of the estimated model. Table 2 shows the final BJ model.

The model may be used to predict the wear of the drill bits and also to determine the optimal conditions needed to minimize the same parameter. Bearing in mind that this is a polynomial model, if all but one of the input variables are fixed, then the remaining variable can be calculated and fixed in order to minimize wear on the tool.

5. Conclusions and future work

This paper has presented an investigation to identify the most appropriate modelling system to solve a real-life industrial problem: the drilling robot. Several methods were investigated to achieve the best practical solution to this challenging problem. This paper shows that the BJ model is better adapted than the other models in this case, insofar as it is able to identify the best conditions and to predict future circumstances. The BJ model can be used to estimate the wear of the tool, but also to fix an input variable in order to minimize the wear of the tool when the rest of the input variables are given.

Furthermore, a two-step procedure has been designed to obtain and to validate the model: a first step, which applies unsupervised connectionist architectures to identify the relevant internal

structures of a data set, to establish whether it is informative enough, and then to identify the most significant variables to be used in the second step. The second step uses conventional system identification techniques to identify the most appropriate model. The aim of this procedure is to reduce the time spent searching for the most appropriate model, which is achieved in the first phase, when the relevant structures are identified. Were the data set not informative enough, a further set of experiments would have to be performed with the drilling robot.

Future work will focus on the study and application of other modelling systems. This modelling system will also be applied to the automation of similar industrial problems that are of interest.

Acknowledgements

References

[1] H. Akaike, *Fitting autoregressive models for prediction*, Ann. Inst. Stat. Math. 20 (1969), pp. 425–439. **Q3**
[2] E. Corchado and C. Fyfe, *Connectionist techniques for the identification and suppression of interfering underlying factors*, Int. J. Pattern Recogn. Artif. Intell. 17(8) (2003), pp. 1447–1466.
[3] E. Corchado, Y. Han, and C. Fyfe, *Structuring global responses of local filters using lateral connections*, J. Experiment. Theoret. Artif. Intell. 15(4) (2003), pp. 473–487.
[4] E. Corchado, D. MacDonald, and C. Fyfe, *Maximum and minimum likelihood Hebbian learning for exploratory projection pursuit*, Data Mining Knowl. Discov. 8(3) (2004), pp. 203–225.
[5] P. Diaconis and D. Freedman, *Asymptotics of graphical projections*, Ann. Statist. 12(3) (1984), pp. 793–815.
[6] J.H. Friedman and J.W. Tukey, *Projection pursuit algorithm for exploratory data-analysis*, IEEE Trans. Comput. 23(9) (1974), pp. 881–890.
[7] C. Fyfe, *PCA Properties of Interneurons: From Neurobiology to Real World Computing*, Proceedings of the International Conference on Artificial Neural Networks (ICANN 1993), Vol. 93, Verlag, S., 1993, pp. 183–188. **Q4**
[8] C. Fyfe and E. Corchado, *Maximum Likelihood Hebbian Rules*, Proceedings of the 10th European Symposium on Artificial Neural Networks (ESANN 2002), 2002, pp. 143–148. **Q4**
[9] R. Haber and L. Keviczky, *Nonlinear System Identification, Input–Output Modeling Approach*, Part 2: *Nonlinear System Structure Identification*, Kluwer Academic Publishers, 1999a. **Q4**
[10] ———, *Nonlinear System Identification, Input–Output Modeling Approach*, Part 1: *Nonlinear System Parameter Estimation*, Kluwer Academic Publishers, 1999b. **Q4**
[11] X. He and H. Asada, *A New Method for Identifying Orders of Input–Output Models for Nonlinear Dynamic Systems*, In Proceedings of the American Control Conference, S. F., California, 1993, pp. 2520–2523.
[12] H. Hotelling, *Analysis of a complex of statistical variables into principal components*, J. Educ. Psychol. 24 (1933), pp. 417–444.
[13] L. Ljung, *System Identification, Theory for the User*, Prentice-Hall, 1999. **Q4**
[14] M. Lorenzo, P. Bravo, and M. Preciado, Parametrización del Corte en Hormigón Fuertemente Armado Mediante Coronas de Diamante, IX Congreso Nacional de Propiedades Mecánicas de Sólidos, Spain, 2004.
[15] O. Nelles, *Nonlinear System Identification, from Classical Approaches to Neural Networks and Fuzzy Models*, Springer, 2001. **Q4**
[16] M. Nögaard, O. Ravn, N.K. Poulsen, and L.K. Hansen, *Neural Networks for Modelling and Control of Dynamic Systems*, Springer-Verlag, London, UK, 2000.
[17] E. Oja, *A simplified neuron model as a principal component analyzer*, J. Math. Biol. 15(3) (1982), pp. 267–273.
[18] K. Pearson, *On lines and planes of closest fit to systems of points in space*, Philos. Mag. 2(6) (1901), pp. 559–572.
[19] H.S. Seung, N.D. Socci, and D. Lee, *The rectified Gaussian distribution*, Adv. Neural Inf. Processing Syst. 10 (1998), pp. 350–356.
[20] T. Söderström and P. Stoica, *System Identification*, Prentice Hall, 1989. **Q4**
[21] P. Stoica and T. Söderström, *A useful parametrization for optimal experimental design*, IEEE Trans. Automatic. Control AC-27 (1982).