

Nonlinear system identification of a MIMO quadruple tanks system using NARX model

Abstract. This paper has two main objectives. First, it gives an overview on the identification of MIMO nonlinear systems using NARX models. It covers the classical approach of the FROLS method, as well as the SEMP method. The second is to present some new useful results in model structure selection for NARX polynomial models applied to MIMO systems. It shows how to make a representation of MIMO systems from NARX polynomial models and the application of classical methods to identify these models. The study case used is a real didactic quadruple tank system manufactured by Quanser.

Streszczenie. Artykuł ma dwa cele. Po pierwsze przedstawia przegląd metod identyfikacji nieliniowych systemów MIMO przy użyciu modelu NARX. Przedstawiono klasyczną metodę FROLS a także metodę SEMP. Po drugie przedstawiono użyteczne wyniki selekcji struktury wielomianowego modelu NARX zastosowanego do systemów MIMO. Nieliniowy system identyfikacji systemu MIMO przy wykorzystaniu modelu NARX

Keywords: Nonlinear system identification, MIMO systems, Model Structure Selection, Coupled Tanks, quadruple tanks.

Słowa kluczowe: nieliniowy system identyfikacji, system MIMO, model NARX

Introduction

System identification is a knowledge area with the main objective of developing methods and techniques to find accurate and reliable mathematical representations for dynamic systems from observed data and available knowledge [1, 2]. A way to build these models is using information from collected system data to build ratios between inputs and outputs from the process, thus describing their dynamics. This method is called black-box identification since no information about internal details is used.

Some real applications, especially industrial applications, have multiple-inputs and multiple-outputs (MIMO) systems. Such systems present some difficulties in obtaining mathematical models by physical laws (white-box modelling) or even by black-box methods to represent their dynamic behaviors. Therefore, linear models are used to represent the dynamics of such MIMO systems but they present lower quality than non-linear models.

In black-box system identification, the representation of the polynomial nonlinear autoregressive moving average with exogenous input (NARMAX), proposed by [3], has a great performance in its ability to represent a nonlinear input-output relationship. In cases where the deterministic input-output relationship is the focus, a nonlinear autoregressive with exogenous input (NARX) model can be employed, using a simplification of the disturbance model [4]. Other researchers demonstrate these capabilities improved with applied techniques of structure selection and parameter estimation, enabling to find models with better representation and most robust performance, as explored in [5, 6].

Some papers propose the use of Artificial Neural Networks models [7, 8] or more elaborate methods [9, 10] for the task of identifying MIMO systems. However, NARX model-based identification methods, although generating less accurate models than the previously mentioned methods, have the advantage of generating simpler and easier models to apply to controllers such as PID or predictive, which are widely used in the industry.

Didactic systems are often used to investigate the behavior of parameter estimation methods, for instance the Quanser water level control quadruple tanks module. The physical meaning of this didactic system can be interpreted, in real applications, like oil-water separators or chemical processes such as bio-reactors and water treatment. In

most applications, this type of system is represented by linearized models around an operating point.

This study shows step-by-step application and compares the use of two black-box system identification techniques, the Forward Regression Orthogonal Least Squares (FROLS) [11, 12] and the Simulation Error Minimization with Pruning (SEMP) [5], for parameter estimation and model structure selection (MSS) of NARX models in a MIMO real study case, the Quanser quadruple tanks module. The obtained NARX model properly characterizes the real process dynamic and can be further used in simulations in future works.

The paper is organized as follows. The second section provides a basic framework for nonlinear system identification and briefly reviews two methods for the model structure selection problem. The third Section describes the coupled tanks used as a study case. The proposed method is illustrated in the fourth section and the results are shown in the fifth section. Finally, some concluding remarks are drawn in the sixth section.

Nonlinear system identification

System identification consists of four sub-steps [1, 13]:

- Experiment execution to acquire data and data pre-processing;
- Model structure selection;
- Parameter estimation;
- Model validation.

As discussed in [14], choosing the signals used to excite the system has great importance in the identification process. In some cases, a step signal is sufficient to identify some process dynamics. However, in cases where it is desired to estimate parameters of higher order models or non-linear models, it is necessary to use signals capable of exciting the various dynamics of the system. Another important fact about data acquisition is that, in some cases, the obtained data set is contaminated by noise with high values. In these cases a pre-filtering is necessary, like the use of a low-pass filter for example.

The model selection of structure (MSS) [15, 16, 17] problem for linear cases is often restricted to choosing the order of the model to be used, whereas for nonlinear models, this problem is more complex. This step is usually combined with parameter estimation.

The Least Squares (LS) Method is efficient to estimate systems that have linear regressors, but in nonlinear systems it is possible that the regressors matrix is badly

scaled, resulting in poor estimates, since they are susceptible to numerical error. In these cases the solution is to apply an orthogonalization method in the regressors matrix like the Classical Gram-Schmidt (CGS) or the Modified Gram-Schmidt (MGS) [11, 18, 19]. This approach is called Orthogonal Least-Squares, since the parameter estimation now will be done in an orthonormal basis space.

To measure the quality of the identified model, the last step in system identification is model validation, a fundamental part of most studies, which consists of tests to confirm if the model can generalize the dynamics of the identified process, and furthermore, if it adequately represents the recorded data set [1, 20].

NARMAX models

Introduced by [11] NAR(MA)X models have been exploited in literature as good representations for non-linear systems as present in [21, 22, 23], although this kind of model increases its intricacy when either the number of regressors or the non-linearity degree increase. A large number of irrelevant regressors can cause over-fitting, resulting in a poor generalization of the model [24].

As cited before, black-box identification is an approach that explains the present output value based on past observations of the system's inputs and outputs. Since they are considered nonlinearities in the system, AR(MA)X models turn themselves into NAR(MA)X models, which can be described by Equation 1.

$$(1) \quad \begin{aligned} y(k) = & F^l[y(k-1), \dots, y(k-n_y), \\ & u(k-1-d), \dots, u(k-n_u-d), \\ & e(k-1), \dots, e(k-n_e)] \end{aligned}$$

where F is a nonlinear function with l non-linearity degree, n_y the maximum output order accepted, n_u the maximum input order accepted and n_e the maximum order for input noise regressors; or in polynomial form (Eq. 2).

$$(2) \quad y(k) = \sum_i^{n_\theta} c_i \prod_{j=1}^{n_y} y(k-j) \prod_{r=1}^{n_u} u(k-r) \prod_{q=0}^{n_e} e(k-q)$$

where n_θ is the total number of regressors parameters.

NARX models have a good representativeness for non-linear systems, although the total number of regressors increases according to the system's order. In MIMO systems (Equation 3) this characteristic appears strongly, since the number of inputs and outputs from the system makes the total number of regressors grow exponentially.

$$(3) \quad y_\alpha(k) = \sum_i^{n_\theta} c_i \prod_{j=1}^{n_y} y_\alpha(k-j) \prod_{r=1}^{n_u} u_\beta(k-r) \prod_{q=0}^{n_e} e(k-q)$$

with $\alpha=1, \dots, q_o$, where q_o is the total number of outputs and $\beta=1, \dots, q_i$, where q_i is the total number of inputs and $n_\theta=q_o n_y + q_i n_u$.

In the polynomial NARMAX and NARX models, there is a harder problem than parameter estimation. The number of candidates regressors increases rapidly when increasing the model's order, as well as when increasing the maximum delays of input and output signals [25]. In [26] it has been shown that an overparameterized model exhibits dynamical regimes which are not found in the original system. Thus, there is not only a numerical or practical reason to do structure selection, but there is also a dynamic justification [27].

The maximum number of candidates regressors (n_θ) grows when either the nonlinearity degree (l) or the maximum input and output delays (n_y and n_u) (and n_e , for NARMAX models) increase, according to $n_\theta=M+1$, where M can be obtained by the Equation (4) [18].

$$(4) \quad \begin{cases} M = \sum_{i=1}^l n_i \\ n_i = \frac{n_{i-1} (n_y \cdot q_o + n_u \cdot q_i + n_e \cdot q_e + i - 1)}{i} \end{cases}$$

Model structure selection

The MSS consists in choosing a subset of appropriate regressors from a set of candidates regressors that represent the dynamic of the dataset. Several studies have already worked on this problem [28, 26, 27, 29, 5, 30, 31, 32, 33, 34, 17].

Typical solutions for structure selection problems are the forward regression orthogonal least squares algorithm (FROLS) which uses an important index, the Error Reduction Ratio (ERR) [11], and the simulation error minimization with pruning (SEMP) method, associated with the simulation error reduction ratio (SRR) [5]. Since both methods are based in Least Squares it's possible to use the formulation present in equations (5,2) even as a generalization to MIMO case representation (Equation 3).

The FROLS algorithm consists of the basic structure of the OLS (Orthogonal Least Squares) algorithm with the addition of a full search for the models that were not selected at each new step, with the intent of finding the best candidate regressor [20].

Considering the prediction error form used in Least Squares (Equation 5), the solution using the CGS method is to decompose the regressors matrix by $\Psi=QA$ where Q is the new matrix regressor, A is the base for this space and $\Psi = [\psi_1, \psi_2, \dots, \psi_{n_\theta}]$.

$$(5) \quad \mathbf{y} = \Psi \hat{\theta} + \xi$$

$$(6) \quad A = \begin{bmatrix} 1 & \alpha_{1,2} & \cdots & \cdots & \cdots & \alpha_{1,n_\theta} \\ & 1 & \alpha_{2,3} & \cdots & \cdots & \alpha_{2,n_\theta} \\ & & \ddots & \ddots & \cdots & \cdots \\ & & & \ddots & \ddots & \cdots \\ & & & & 1 & \alpha_{n_\theta-1,n_\theta} \\ & & & & & 1 \end{bmatrix}$$

Where the components of the matrix and the composition of the matrix are given by the Equations 7-10 [1].

$$(7) \quad \mathbf{q}_1 = \psi_1$$

$$(8) \quad \alpha_{j,i} = \frac{\langle \mathbf{q}_j, \psi_i \rangle}{\langle \mathbf{q}_j, \mathbf{q}_j \rangle}, 1 \leq j \leq i$$

$$(9) \quad \mathbf{q}_j = \psi_j - \sum_{i=1}^{j-1} \alpha_{j,i} \mathbf{q}_i$$

$$(10) \quad Q = [q_1, q_2, \dots, q_{n_\theta}]$$

This method gives parameters in an orthonormal basis space $\hat{g} = A\hat{\theta}$, so the parameters in the original basis space are given by $\hat{\theta} = A^{-1}\hat{g}$, where \hat{g} is given by Equation 11.

$$(11) \quad \hat{g}_i = \frac{\mathbf{q}_i^T \mathbf{y}}{\mathbf{q}_i^T \mathbf{q}_i}, i = 1, \dots, n_\theta$$

In order to select the best model structure which fits the data set, a criterion for regressor selection must, at each iteration, select the regressor with the best value.

The ERR criteria provides an effective way of determining a group of significant regressors [20]. These significant terms are selected based on the value of ERR_m , which is defined by Equation 12.

$$(12) \quad [ERR]_m = \frac{(\hat{g}_m)^2 (\mathbf{q}_m^T \mathbf{q}_m)}{\sigma}$$

where $\sigma = \mathbf{y}^T \mathbf{y}$.

The FROLS method increments the ERR (Error Reduction Ratio) to evaluate the regressor significance for the model then builds a reduced regressors matrix and finally estimates the parameters for the model, for this an orthogonal matrix is built from the candidates matrix using GS (Equations 7-10). The process is illustrated in Algorithm 4.

Algorithm 1 FROLS

```

1:  $Q = \Psi$ 
2:  $P \text{ hi} = \Psi$ 
3:  $A = I$ 
4: for  $m=1:n_\theta$  do
5:   Calculation of the transformation base  $\hat{g}$ 
6:   Calculation of the
ERR 7:  $k = \text{position of}$ 
max ERR 8:  $Q^{(1)} = q_k$ 
9:  $P \text{ hi} = \Psi$  without  $k$ -column
10:  $s = 2$ 
11:  $aux = 1$ 
12: while  $s \leq n_\theta$  do
13:   Calculation of  $Q$  matrix from GS using  $Phi$ 
14:   Calculation of transformation base  $\hat{g}$ 
15:   Calculation of ERR

```

In SEMP method, initially, it is considered a matrix that has all the candidate regressors Ψ_M with dimension $N \times n_\theta$, where N is the number of observations and n_θ is the number of regressors. The matrix $\mathbf{P} = [p_1 \dots p_j]$, with $j \leq n_\theta$, represents the set of candidate regressors that are in the current model of iteration j , and is initialized as $\mathbf{P} = \{\}$. The matrix $\mathbf{Q} = [q_1 \dots q_{n_\theta - j}]$ represents the set of candidate regressors that are outside of the current model

At each iteration, a term j of matrix \mathbf{Q} is sequentially added to matrix \mathbf{P} . For each term added to the model, the algorithm finds the parameters $\theta \in \mathbb{R}^p$ using an estimator. At the end of the $j \leq n_\theta$ iterations, the matrix \mathbf{P} is a subset of Ψ_M with the best set of regressors chosen by the SRR criterion.

In [5] the ERR criterion of the j regressor added to the current model is defined as:

$$(13) \quad [ERR]_j = \frac{MSPE(M_{i-1}) - MSPE(M_i)}{1/N \sum_{i=1}^y y^2(k)}$$

where $MSPE(M_i)$ is the mean square prediction error of the model at the j^{th} iteration including the j^{th} regressor. This criterion uses the one step ahead prediction error to calculate the influence of the j^{th} regressor in the current model.

The SRR (Equation) criterion is defined in close analogy to the ERR. But, instead of using the MSPE, it uses the mean square simulation error (MSSE). The main idea is to simulate the current model and collect the error, instead of using the one step ahead prediction error.

$$(14) \quad [SRR]_j = \frac{MSSE(M_{i-1}) - MSSE(M_i)}{1/N \sum_{i=1}^y y^2(k)}$$

The pruning procedure occurs after adding a regressor to the model. In this step, the regressors of the best chosen model are drawn sequentially one by one and the model without the current regressor is evaluated. If the model has a better SRR value than the complete model, then it becomes the best model.

The SEMP method is summarized in Algorithm 2. Note that all the procedure is repeated for each of the q_o equations that represent the dynamics of each system output.

Algorithm 2 SEMP

```

Initialization:  $\sigma, \mu, \rho$ 
1: for  $N_{\text{Output}} = 1:q_o$  do
2:    $\mathbf{P} = []$ 
3:    $\mathbf{Q} = \Psi_M$ 
4:   for  $i = 1 : n_\theta$  do
5:     for  $j = 1 : \text{size of } \mathbf{Q}$  do
6:        $\mathbf{P}_{\text{test}} = [\mathbf{P} \ \mathbf{q}_j]$ 
7:       Calculation of  $\theta_j$  using  $LS$ 
8:       Calculation of  $J_j$  using  $SRR$ 
9:        $l = \text{position of argmax } J$ 
10:      if  $J_l > J_{\text{old}} \ \& \ |J_{\text{old}} - J_l| > \rho$  then
11:         $\mathbf{P} = [\mathbf{P} \ \mathbf{q}_l]$ 
12:         $\mathbf{q}_l = []$ 
13:      else
14:        END
15:      for  $k = 1 : \text{size of } \mathbf{P}$  do
16:         $\mathbf{R} = \mathbf{P}$  without  $p_k$ 
17:        Calculation of  $\theta_k$  using  $LS$ 
18:        Calculation of  $J_{p_k}$  using  $SRR$ 
19:         $m = \text{position of argmax } J_p$ 
20:        if  $J_{p_m} > J_{\text{old}}$  then
21:           $\mathbf{P} = [\mathbf{P}$  without  $p_m]$ 
22:          GOTO 16
23:        else
24:          GOTO 4

```

Being iterative algorithms, both FROLS and SEMP need a stop criterion. The search is terminated at the M step when the ESR (see Equation 15) is less than a pre-specified threshold ρ .

$$(15) \quad ESR = 1 - \sum_{s=1}^M RR(s) \leq \rho$$

Quadruple tanks system

In this section the quadruple tanks system will be explored. This kind of problem is oftentimes used in industrial facility as water treatment, bio-reactors, distillation process and in processes like cereal storage as rice, soy and corn. [35, 36, 37].

Even with the quadruple tanks system being a simplified version of bigger systems, its complexity can be increased by adding more stages, and this characteristic allows for this kind of system to represent a large scale of industrial process, and your simplicity in mathematical modelling becomes attractive to study techniques in control and identification.

In this paper the Quanser quadruple tanks system (Figure 1) is being used to investigate the efficiency of the previous cited algorithms, FROLS and SEMP, to detect structure and estimate parameters for a nonlinear black-box model.



Fig. 1. Quanser's quadruple tanks system.

In the configuration shown in Figure 2 pump 1 pumps water to tanks 1 and 4, while pump 2 pumps water to tanks 2 and 3. The water in the upper tanks flows to the lower tanks, and then flows into the water reservoir so that it is pumped back into the upper tanks, becoming an autonomous water re-circulation system.

To build the mathematical model, Equation 16, it was assumed the pumps could be modelling by a gain, using mass balance to describe the water flowing through the tanks. This mathematical modelling can be expanded to any quantity as explored in [38].

It is important to note that, over time, some of the parameters in Table 1 begin to change, mainly due to the wear of the mechanical elements or due to the accumulation of residues in the holes and hoses. In this way, the modeling using the data provided by the manufacturer may present a divergent behavior in relation to the real system.

$$(16) \quad \begin{cases} \frac{dH_1}{dt} = \frac{K_1 v_1 - d_1 \sqrt{2gh_1}}{D_1} \\ \frac{dH_2}{dt} = \frac{K_2 v_2 + d_1 \sqrt{2gh_1} - a_2 \sqrt{2gh_2}}{D_2} \\ \frac{dH_3}{dt} = \frac{K_2 v_2 - d_3 \sqrt{2gh_3}}{D_3} \\ \frac{dH_4}{dt} = \frac{K_1 v_1 + d_3 \sqrt{2gh_3} - a_4 \sqrt{2gh_4}}{D_4} \end{cases}$$

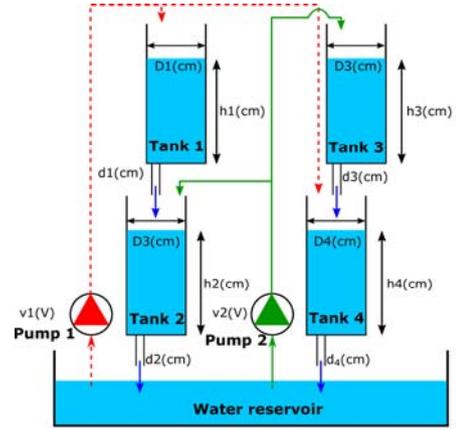


Fig. 2. Schematic diagram of the quadruple tanks process.

Table 1. The parameters of the sensor

Variable	Definition
$k_1 = 3$	Pump 1 equivalent gain
$k_2 = 3$	Pump 2 equivalent gain
$D_1 = 15.5\text{cm}$	Diameter Tank 1
$D_2 = 15.5\text{cm}$	Diameter Tank 2
$D_3 = 15.5\text{cm}$	Diameter Tank 3
$D_4 = 15.5\text{cm}$	Diameter Tank 4
$d_1 = 0.123\text{cm}$	Output diameter Tank 1
$d_2 = 0.123\text{cm}$	Output diameter Tank 2
$d_3 = 0.123\text{cm}$	Output diameter Tank 3
$d_4 = 0.123\text{cm}$	Output diameter Tank 4
$g = 981\text{ cm/s}^2$	Gravity acceleration

Experimental Results

It is a common practice to excite the system of interest and to use the resulting data to build the model, and the choice of an input signal used for the system's activation is critical in the task of model building and parameter identification [14]. Thus, to identify the quadruple tanks system, a PRS (pseudo random signal) was used to excite each of the pumps that compose the system.

Figure 3 shows the generated input signal used in estimation step. An important feature of the Quanser system is that it uses an amplifier to condition the signals applied to the pumps in the tanks system. This amplifier multiplies each of the signals in Figure 3 by 3 and then applies them to the pumps.

For the identification of processes with discrete-time signals, the sample rate must be chosen prior to the measurement. If the chosen sample time is too small, then a badly conditioned system of equations results, as the difference equations for different values of sample become nearly linearly dependent. Hence, when decreasing the sample time, one suddenly witnesses a big increase in the parameter variances. However, choosing the sample time is rather uncritical as the range between too small and too large sample times is relatively broad [13]. In this experiment the used sample time is $T_s = 0.1$ seconds.

With the data set that will be used in the identification process, we proceed to the stage of structure selection and parameter estimation. At this point, a candidate model is chosen, which has all the candidate regressors. For the experiment, a model with the following configuration was used: output delay $n_y=2$, input delay $n_u = 2$ and nonlinear degree $l = 2$. With this configuration, the NARX model of each output equation has a total of 91 candidate regressors (see Equation 4).

This number of candidates regressors of the NARX model evidences the necessity of the structure selection step along with the parameter estimation.

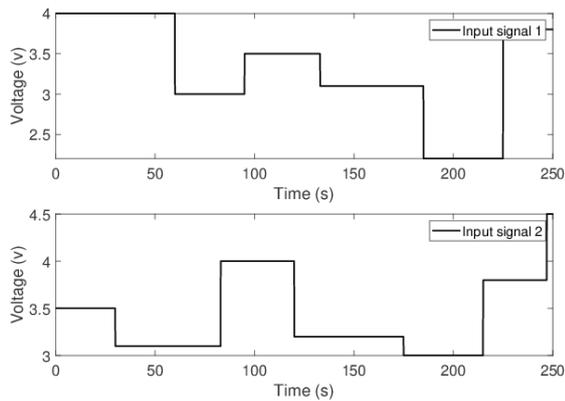


Fig. 3. PRS input signal applied to each pump of the system in estimation step of identification processing.

Equation 17 shows the model obtained by the FROLS method. Note that for each equation representing the dynamics of each output, only two regressors of the 91 candidates are used to represent the system dynamics.

$$\begin{aligned}
 (17) \quad & y_1(k) = 0.994 \cdot y_1(k-1) + 0.008 \cdot u_1(k-1)^2 \\
 & y_2(k) = 0.733 \cdot y_2(k-1) + 0.267 \cdot y_2(k-2) \\
 & y_3(k) = 0.994 \cdot y_3(k-1) + 0.006 \cdot u_2(k-2)^2 \\
 & y_4(k) = 0.720 \cdot y_4(k-1) + 0.280 \cdot y_4(k-2)
 \end{aligned}$$

Equation 18 shows the model obtained by the SEMP method. Similar to FROLS, the SEMP method can select a set with the best regressors capable of representing the dynamics of the tanks system. In both cases, the threshold of the stop criteria is $\rho = 10^{-4}$.

$$\begin{aligned}
 (18) \quad & y_1(k) = 0.995 \cdot y_1(k-1) \\
 & \quad + 0.008 \cdot u_1(k-1) \cdot u_1(k-2) \\
 & y_2(k) = 0.698 \cdot y_2(k-1) + 0.298 \cdot y_2(k-2) \\
 & \quad + 0.008 \cdot u_1(k-2) \cdot u_2(k-2) \\
 & y_3(k) = 0.747 \cdot y_3(k-1) + 0.008 \cdot u_2(k-2)^2 \\
 & \quad + 0.246 \cdot y_3(k-2) \\
 & y_4(k) = 0.681 \cdot y_4(k-1) + 0.314 \cdot y_4(k-2) \\
 & \quad + 0.008 \cdot u_1(k-2) \cdot u_2(k-1)
 \end{aligned}$$

The generated signals in Figure 3, after being amplified, are applied to the pumps 1 and 2, as shown on the scheme of Figure 2, as well as in the models of Equations 17 and 18, and generate the output signals shown in Figure 4. It is important to note that in this case the system is in an open loop.

From Figure 4 it is possible to notice that the models are able to represent well the system dynamics from the estimation data.

After the identification of the models obtained by the methods described above, a validation step is required. This step is important because it shows the generalization capacity the obtained models have.

To validate the models, a data set different from the estimation step is used for a free simulation. Figure 5 shows the set of excitation signals applied in the real system and in the models. Note that, once again, a PRS signal was used to excite the system.

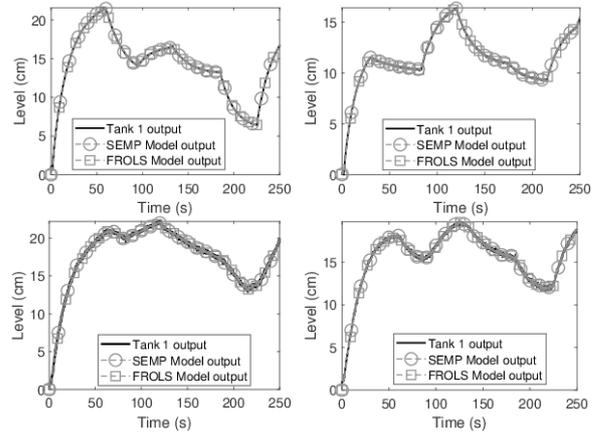


Fig. 4. Response of each tank and its estimated models to the PRS signal of Figure 3.

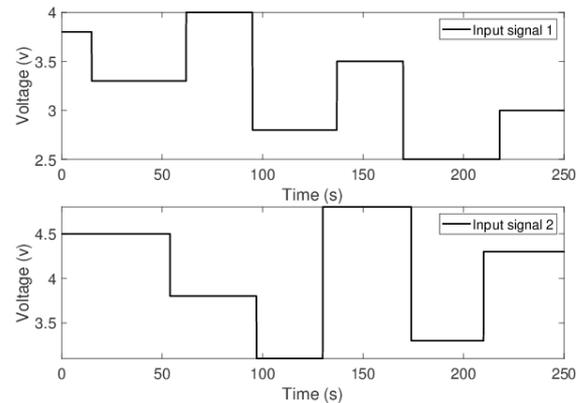


Fig. 5. PRS input signal applied to each pump of the system in validation step of identification processing.

Figure 6 shows the system's output after the application of the validation signal of Figure 5.

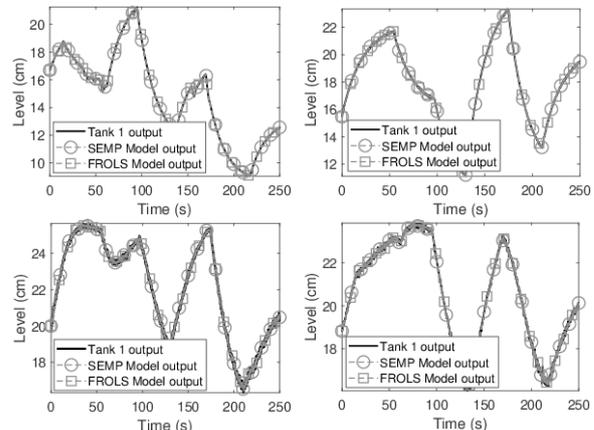


Fig. 6. Response of each tank to PRS signal in Figure 5.

To judge the quality of each of the obtained models, the root mean square error (RMSE) is used in each of the output vectors. The RMSE indicates the standard deviation of the residues or prediction errors and how these residues are spread. The RMSE is applied according to the following expression:

$$(19) \quad RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}$$

where the number of observations (or discrete points) is $N=2500$ and the error is calculated using the actual output y and the estimated output \hat{y} . The Table 2 shows the obtained results.

Table 2. RMSE Values for each output

Method/Output	Tank 1	Tank 2	Tank 3	Tank 4
FROLS	0.0343	0.0946	0.0425	0.0885
SEMP	0.0342	0.0923	0.0408	0.0867

According to Table 2, it is possible to observe that both methods have low error values. As expected, the SEMP method has lower error values than FROLS, however, it has a significantly higher computational cost.

Conclusion

The analysis of the results obtained in the MIMO quadruple tank system suggests that such methods of identification of nonlinear systems are quite efficient, especially regarding the selection of NARX models. The identified models have low error values.

Although NARX models are quite simple, such models show a very high growth in the number of candidate regressors when increasing the degree of the system and the degree of non-linearity, which consequently causes an increase in the computational cost of identification methods.

The practical experiment presented in this paper suggests that application of MIMO nonlinear systems identification techniques should always be considered. If the system presents nonlinearity behaviour then linear models will have lower quality representing the dynamic of the system, especially in MIMO cases.

Acknowledgments

This work was conducted during a scholarship supported by the Cooperation Program CAPES at the University of Rio Grande do Norte

Authors: Prof. Ícaro B. Q. Araújo, Yan L. S. Lúcio, Computer Institute, Federal University of Alagoas Av. Lourival de Melo Mota, Bloco 12, Tabuleiro do Martins 57072-970, Maceió, AL, Brazil, email: icaro@ic.ufal.br, Gabriel B. Cavalcante, Department of Automation and Systems Federal University of Santa Catarina Campus Universitário Reitor João David Ferreira Lima - R. Delfino Conti, s/n - Trindade, Florianópolis - SC, 88040-900 M. Sc. Ícaro B. Q. Araújo, Fábio M. U. Araújo, Department of Computer Engineering and Automation, Federal University of Rio Grande do Norte - Technology Center, UFRN Campus Universitário Lagoa Nova, 59078-970, Natal, RN, BRAZIL.

REFERENCES

- [1] L. Ljung, System identification: theory for the user. Prentice-hall, 1987.
- [2] R. Pintelon and J. Schoukens, System Identification: A Frequency Domain Approach. Wiley, 2004.
- [3] I. Leontaritis and S. A. Billings, "Input-output parametric models for non-linear systems part i: deterministic non-linear systems," International journal of control, vol. 41, no. 2, pp. 303–328, 1985.
- [4] W.-X. Zhao and H.-F. Chen, "Identification of wiener, hammerstein, and narx systems as markov chains with improved estimates for their nonlinearities," Systems & Control Letters, vol. 61, pp. 1175–1186, 12 2012.
- [5] L. Piroddi and W. Spinelli, "An identification algorithm for polynomial narx models based on simulation error minimization," International Journal of Control, vol. 76, no. 17, pp. 1767–1781, 2003.
- [6] I. Lind and L. Ljung, "Regressor and structure selection in narx models using a structured anova approach," Automatica, vol. 44, no. 2, pp. 383–395, 2008.
- [7] T. Orlowska-Kowalska and C. T. Kowalski, "Neural network application for flux and speed estimation in the sensorless induction motor drive," in ISIE '97 Proceeding of the IEEE International Symposium on Industrial Electronics, pp. 1253–1258 vol.3, July 1997.
- [8] A. Ghadirian and M. Zekri, "Mimo nonlinear dynamic systems identification using fully recurrent wavelet neural network," in The 2nd International Conference on Control, Instrumentation and Automation, pp. 1113–1118, Dec 2011.
- [9] J. S. R. Jang, "Anfis: adaptive-network-based fuzzy inference system," IEEE Transactions on Systems, Man, and Cybernetics, vol. 23, pp. 665–685, May 1993.
- [10] N. Mathur, I. Glesk, and A. Buis, "Comparison of adaptive neuro-fuzzy inference system (anfis) and gaussian processes for machine learning (gpml) algorithms for the prediction of skin temperature in lower limb prostheses," Medical Engineering & Physics, vol. 38, no. 10, pp. 1083 – 1089, 2016.
- [11] S. A. Billings, M. J. Korenberg, and S. Chen, "Identification of non-linear output-affine systems using an orthogonal least-squares algorithm," International Journal of Systems Science, vol. 19, pp. 1559–1568, Apr 1988.
- [12] S. A. Billings, S. Chen, and M. J. Korenberg, "Identification of mimo non-linear systems using a forward-regression orthogonal estimator," International Journal of Control, vol. 49, no. 6, pp. 2157–2189, 1989.
- [13] R. Isermann and M. Münchhof, Identification of Dynamic Systems: An Introduction with Applications. Advanced Textbooks in Control and Signal Processing Series, Springer Berlin Heidelberg, 2010.
- [14] W. Jakowluk, "Design of an optimal excitation signal for identification of inertial systems in time domain," Przegląd Elektrotechniczny, vol. R. 85, nr 6, pp. 125–129, 2009.
- [15] A. Falsone, L. Piroddi, and M. Prandini, "A randomized algorithm for nonlinear model structure selection," Automatica, vol. 60, pp. 227 – 238, 2015.
- [16] F. Bianchi, A. Falsone, M. Prandini, and L. Piroddi, "A randomised approach for narx model identification based on a multivariate bernoulli distribution," International Journal of Systems Science, vol. 48, no. 6, pp. 1203–1216, 2017.
- [17] M. Avellina, A. Brankovic, and L. Piroddi, "Distributed randomized model structure selection for narx models," International Journal of Adaptive Control and Signal Processing, vol. 31, no. 12, pp. 1853–1870, 2017.
- [18] M. KORENBERG, S. Billings, Y. Liu, and P. McIlroy, "Orthogonal parameter estimation algorithm for non-linear stochastic systems," International Journal of Control, vol. 48, no. 1, pp. 193–210, 1988.
- [19] S. Chen and S. A. Billings, "Representations of non-linear systems: the narmax model," International Journal of Control, vol. 49, no. 3, pp. 1013–1032, 1989.
- [20] S. A. Billings, Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains. John Wiley & Sons, 2013.
- [21] N. Chiras, C. Evans, and D. Rees, "Nonlinear gas turbine modeling using narmax structures," IEEE Transactions on Instrumentation and Measurement, vol. 50, no. 4, pp. 893–898, 2001.
- [22] E. H. Fung, Y. Wong, H. Ho, and M. P. Mignolet, "Modelling and prediction of machining errors using armax and narmax structures," Applied Mathematical Modelling, vol. 27, no. 8, pp. 611–627, 2003.
- [23] S. K. Pradhan and B. Subudhi, "Narmax modeling of a two-link flexible robot," in 2011 Annual IEEE India Conference, pp. 1–5, IEEE, 2011.
- [24] R. Salat, M. Awtoniuk, and K. Korpysz, "Black-box system identification by means of support vector regression and imperialist competitive algorithm," Przegląd Elektrotechniczny, vol. 89, pp. 223–226, 09 2013.
- [25] Y. Cheng, L. Wang, and J. Hu, "A two-step method for nonlinear polynomial model identification based on evolutionary optimization," in 2009 World Congress on Nature Biologically Inspired Computing (NaBIC), pp. 613–618, Dec 2009.
- [26] L. A. Aguirre and S. Billings, "Dynamic effects of overparametrization in nonlinear models," Physica D: Nonlinear Phenomena, vol. 80, no. 1, pp. 26 – 40, 1995.
- [27] L. A. Aguirre and C. R. F. Jacome, "Cluster analysis of narmax models for signal-dependent systems," IEE Proceedings - Control Theory and Applications, vol. 145, pp. 409–414, Jul 1998.
- [28] W. Luo and S. Billings, "Adaptive model selection and estimation for nonlinear systems using a sliding data window," Signal Processing, vol. 46, no. 2, pp. 179 – 202, 1995.

- [29] L. A. Aguirre, U. S. Freitas, C. Letellier, and J. Maquet, "Structure-selection techniques applied to continuous-time nonlinear models," *Physica D: Nonlinear Phenomena*, vol. 158, no. 1, pp. 1 – 18, 2001.
- [30] M. Farina and L. Piroddi, "An iterative algorithm for simulation error based identification of polynomial input-output models using multi-step prediction," *International Journal of Control*, vol. 83, no. 7, pp. 1442–1456, 2010.
- [31] M. Hirsch and L. del Re, "Iterative identification of polynomial narx models for complex multi-input systems," *IFAC Proceedings Volumes*, vol. 43, no. 14, pp. 445 – 450, 2010. 8th IFAC Symposium on Nonlinear Control Systems.
- [32] M. Bonin, V. Seghezza, and L. Piroddi, "Lasso-enhanced simulation error minimization method for narx model selection," in *Proceedings of the 2010 American Control Conference*, pp. 4522–4527, June 2010.
- [33] S. A. M. Martins, E. G. Nepomuceno, and M. F. S. Barroso, "Improved structure detection for polynomial narx models using a multiobjective error reduction ratio," *Journal of Control, Automation and Electrical Systems*, vol. 24, pp. 764–772, Dec 2013.
- [34] A. M. Barbosa, R. H. C. Takahashi, and L. A. Aguirre, "Equivalence of non-linear model structures based on pareto uncertainty," *IET Control Theory Applications*, vol. 9, no. 16, pp. 2423–2429, 2015.
- [35] W. Fuchs, H. Binder, G. Mavrias, and R. Braun, "Anaerobic treatment of wastewater with high organic content using a stirred tank reactor coupled with a membrane filtration unit," *Water Research*, vol. 37, no. 4, pp. 902–908, 2003.
- [36] A. Numsomran, V. Tipsuwanporn, and K. Tirasesth, "Modeling of the modified quadruple-tank process," in *2008 SICE Annual Conference*, pp. 818–823, IEEE, 2008.
- [37] S. Xie, Y. Xie, W. Gui, and C. Yang, "Weighted-coupling cstr modeling and model predictive control with parameter adaptive correction for the goethite process," *Journal of Process Control*, vol. 68, pp. 254–267, 2018.
- [38] A. Maxim, C. Ionescu, and R. De Keyser, "Modelling and identification of a coupled sextuple water tank system," in *Automation, Quality and Testing, Robotics (AQTR), 2016 IEEE International Conference on*, pp. 1–6, IEEE, 2016.