A Neural Approach of Multimodel Representation of Complex Processes

Nesrine Elfelly, Jean-Yves Dieulot, Pierre Borne

Abstract: The multimodel approach was recently developed to deal with the issues of complex processes modeling and control. Despite its success in different fields, it still faced with some design problems, and in particular the determination of the models and of the adequate method of validities computation.

In this paper, we propose a neural approach to derive different models describing the process in different operating conditions. The implementation of this approach requires two main steps. The first step consists in exciting the system with a rich (e.g. pseudo random) signal and collecting measurements. These measurements are classified by using an adequate Kohonen self-organizing neural network. The second step is a parametric identification of the base-models by using the classification results for order and parameters estimation. The suggested approach is implemented and tested with two processes and compared to the classical modeling approach. The obtained results turn out to be satisfactory and show a good precision. These also allow to draw some interpretations about the adequate validities’ calculation method based on classification results.

Keywords: complex processes, modeling, multimodel approach, Kohonen map

1 Introduction

Nowadays, technological developments increase the complexity of systems. This complexity can be caused by non linearity, non stability, wide operating domain, variations of system parameters or external perturbations. As a result, it is often difficult or even impossible to propose a simple model which could reckon with the whole process complexity by using physical laws. Consequently, it is very useful to focus on advanced and practical approaches in order to handle this complexity.

The multimodel approach has recently been developed and applied in several science and engineering domains. It was proposed as an efficient and powerful method to cope with modeling and control difficulties when complex non linear and/or uncertain processes are concerned. The multimodel approach supposes the definition of a set of models. Then, it becomes possible to replace the unique model by a set of simpler models thus making a so-called models’ base. Each model of this base describes the behavior of the considered process at a specific operating point. The multimodel approach objective is to decrease the process complexity by its study under certain specific conditions.

Several researchers have been interested in multimodel analysis and control approaches [7, 8, 19] and many applications have been proposed in different contexts.

In spite of its success in many fields (academic, biomedical, . . . ), the multimodel approach remains confronted with several difficulties such as the calculation of models’ validities, the adequate technique of fusion or switching between models as well as the determination of the models’ base.

Indeed, in 1985, Takagi and Sugeno [21] suggested a fuzzy process representation by the contribution of local models. This approach has been applied in many fields of activities but often faced problems related to the lack of information about the system structure or the incertitude of its parameters. Besides, it can sometimes lead to a large number of models which generates a high computational burden when designing the control algorithm. This has led several researchers to develop other approaches in order to cope with these difficulties.

For uncertain complex systems with bounded parameters, some approaches were developed for the determination of the models’ base [9, 11]. An extension for uncertain discrete systems has been proposed by Mezghani [13]. The case of multivariable systems was addressed by Raissi [16]. The approaches,
previously named, require the knowledge of parameters’ variations limits, which is generally not possible, in particular in case of uncertain systems for which parameters variations domains are unknown or ill-known.

Other related studies [12, 19] suggest the system be described by a set of local models often defined by using a reference model and some linearization methods.

In another context, both fuzzy logic and neural networks were carried out for the multimodel control [1, 18]. An idea was to use neural approaches for complex systems modeling. These methods have the advantage of requiring very little information about the considered process and are useful for uncertain systems. Within this context, very few studies [2, 3, 15, 17, 20, 22] were proposed but they don’t address strongly non-linear systems.

In this paper, a neural approach for the determination of the models’ base for uncertain complex systems is proposed, in particular those which exhibit strong non-linearities. The proposed approach requires a priori little knowledge about the considered system; only input/output information can be sufficient.

In the following section, the different steps of the proposed modeling approach are detailed. Two simulation examples and some interpretations are then presented. The evaluation of the suggested modeling strategy is the topic of the last section. We finish the present paper by a conclusion.

2 Complex systems’ modeling: Neural approach

In this section, a models’ base-determination approach for uncertain complex systems is described. This approach requires neither a global model definition nor the knowledge of parameters variations domains; only input/output information are needed.

The suggested approach allows the generation of the base-models’ structure and parameters. The application of this approach requires first to classify the numerical data by exploiting a Kohonen map and to determine the number of models. Secondly, a structural and parametric identification of different base-models is carried out by using classification results. Then, the validity index of each model is computed. Finally, the multimodel output is obtained by the fusion of the models’ outputs weighted by their validity indexes.

2.1 Classification of the numerical data by using a Kohonen map

As a first step, the output or input/output measurements collected from the considered process have to be classified in order to identify operating clusters from which the models’ base will be deduced. This classification is carried out by using a self-organizing Kohonen map.

Self-organizing Kohonen map methodology

The Self-Organizing Map (SOM) represents a specific kind of neural networks. In its original form, the SOM was invented by the founder of the Neural Networks Research Center, Professor Teuvo Kohonen in 1981-82. The special property of the SOM is that it effectively creates spatially organized internal representations of various features of input signals and their abstractions [10]. The schematic representation of this network is given in figure 1.

Neurons in the target layer are ordered and correspond to cells of a bi-dimensional map. Every neuron of the input layer is connected to every neuron of the output layer. The classification strategy consists in applying the Kohonen rule. This rule is characterized by an unsupervised competitive learning where a competition takes place before the modification of the network-weights. Only the neurons which win the competition have the right to change their weights.

The Kohonen rule works as follows:
A Neural Approach of Multimodel Representation of Complex Processes

Figure 1: Kohonen map

- weights are initialized to random values;
- an input vector is presented to the network;
- the distance between the input vector and weights connecting inputs to each output neuron is computed;
- the neuron corresponding to the smallest distance, i.e. the nearest to the input vector, wins the competition, weights connecting inputs to this neuron are modified accordingly [5].

This procedure is repeated several times until weights stabilize. At the end of the learning stage, the representative vectors of different clusters and their centers are obtained.

**Determination of the operating-system clusters**

The determination of the operating-system clusters requires firstly that the considered system be excited. Secondly, the number of clusters has to be determined. The third step consists in classifying data by using a Kohonen neural network.

The first step consists in applying an input signal and then collecting (output or input/output) measurements that will be used for classification. The excitation signal must be rich enough and persistently exciting with well-chosen parameters in order to allow a full excitation of the operating dynamics, and to take in consideration the non-linear aspect of the considered process.

For the second step, the method proposed by Talmoudi has been adopted for the determination of the adequate number of clusters which corresponds to the number of models [22]. Within this scope, a Kohonen network, with an important number $n$ of neurons in the output layer, has been considered. At the end of the learning procedure, if the repartition of classes is not good, the clusters $i$ having a number of elements $N_{Ci}$ verifying the relation (1), will have to be removed [22].

$$N_{Ci} \leq \frac{1}{2} \frac{N_H}{n},$$

(1)

where $N_H$ represents the number of the considered measures.

Else, the number of neurons in the output layer is increased and training is restarted. The same procedure is repeated over and over until the satisfactory number of clusters is obtained. Afterwards, the data classification is tackled by using a Kohonen network for which the number of neurons in the input layer equals the number of system-variables to be considered and the number of neurons in the output layer equals the number of clusters determined with the help of the method previously described. The classification results will then be exploited for the identification of the different base-models.

**2.2 Parametric identification of the base-models**

In this section, the orders of the models are estimated in a first step. The chosen method is the so-called instrumental determinants’ ratio-test. This method is mainly based on the conditions concerning
a matrix called "information matrix" which contains the input/output measurements [4]. This matrix is described as follows:

\[
Q_m = \frac{1}{N_H} \sum_{k=1}^{N_H} \begin{bmatrix}
    u(k) & u(k+1) & \cdots & y(k+1) & \cdots & u(m+k) & u(k+m)
  
end{bmatrix},
\]

(2)

where \(N_H\) is the number of observations. The instrumental determinants’ ratio (RDI) is given by:

\[
RDI(m) = \frac{\det(Q_m)}{\det(Q_{m+1})}.
\]

(3)

For every value of \(m\), the determination procedure of the order consists in building the matrices \(Q_m\) and \(Q_{m+1}\) and in evaluating the ratio \(RDI(m)\), the retained order \(m\) is the value for which the ratio \(RDI(m)\) quickly increases for the first time.

As a second step, the parametric identification issue consists in calculating the values of the parameters of the corresponding model-equation, given several experimental measures which describe the dynamic behavior of the system. As previously mentioned, the data classification gives a certain repartition of clusters. For each cluster, input/output measurements are collected. These measurements allow the identification of the corresponding model. For this, the recursive least-squares method (RLS) [4] was applied to achieve the parameters estimation.

2.3 Computation of Validities

The validity coefficient is a number belonging to the interval \([0, 1]\). It represents the relevance degree of each base-model calculated at each instant. In literature, several methods have been proposed to deal with the validity issue. In our study, the residues’ approach was adopted for the calculation of validities. This method is based on the distance measurement between the process and the considered model. For example, the residue can be given by the following expression:

\[
r_i = |y - y_i|, \quad i = 1, \ldots, N
\]

(4)

where: \(N\): number of base-models;

\(y\): process output;

\(y_i\): output of the model \(M_i\).

If this residue value is equal to zero, the corresponding model \(M_i\) perfectly represents the process at that time. On the contrary, a non null value translates the fact that the model \(M_i\) represents the system partially. The normalized residues are given by:

\[
r'_i = \frac{r_i}{\sum_{j=1}^{N} r_j}.
\]

(5)

Within the context of the residues’ approach, several methods have been proposed for the calculation of validities [6, 13, 14]. Only two methods will be considered: the simple and the reinforced validities. The validities are given by:

\[
v_i = 1 - r'_i.
\]

(6)

The simple and reinforced validities are defined by using the following formulas.
Simple validities: the normalized simple validities are defined so that their sum must be equal to 1 at each time:

\[ v_{i}^{\text{simp}} = \frac{v_i}{N-1} \] (7)

Reinforced validities: for this type of validities, the reinforcement expression is introduced as:

\[ v_{i}^{\text{renf}} = v_i \prod_{j=1, j \neq i}^{N} (1 - v_j) \] (8)

The normalized reinforced validities could be written as follows:

\[ v_{i}^{\text{renf}} = \frac{v_{i}^{\text{renf}}}{\sum_{j=1}^{N} v_{j}^{\text{renf}}} \] (9)

2.4 Computation of the multimodel output

The multimodel output is calculated by a fusion of the models’ outputs weighted by their respective validity indexes, as illustrated by the following expression:

\[ y_{\text{mm}}(k) = \sum_{i=1}^{N} y_{i}(k) v_{i}(k) \] (10)

\(v_{i}(k)\) could be a simple or a reinforced validity for which \(\sum_{i=1}^{N} v_{i}(k) = 1\).

3 Simulation examples

In order to underline the interest and the performance of the proposed approach, some simulation examples are carried out.

3.1 Example 1: second order discrete system

The considered system is a complex discrete system whose evolution is described by the following equation:

\[ y(k) = -a_1(k) y(k-1) - a_2(k) y(k-2) + b_1(k) u(k-1) + b_2(k) u(k-2) \] (11)

The variation laws of different parameters of the process are given by:

\[ a_1(k) = 0.04 \sin (0.035k) - 0.8 \] , \[ a_2(k) = 0.005 \sin (0.03k) + 0.1 \] , (12)

\[ b_1(k) = 0.02 \sin (0.03k) + 0.5 \] , \[ b_2(k) = 0.01 \sin (0.035k) + 0.2 \] . (13)

First, the system is excited by a uniform random signal \(u(k)\). Then, the measurements \(y(k)\) and \(y(k-1)\) are collected at different instants.

The adequate number of clusters determined by using the method described in paragraph 3.1.2, is equal to three. The numerical data are fed into a Kohonen network which presents an input layer with two neurons and three neurons in the output layer. The classification results are given in figure 2.

From each of the three data sets relative to the various clusters, the orders and the parameters of the transfer functions relative to the three base-models are estimated. Figure 3 shows the evolutions of the \(RDI_i(m)\) for the three models.
It appears clearly that the estimated orders of the three models are equal to two. This result is predictable since the considered system is a second order one. The application of the recursive least-squares method allows us to write the following transfer functions:

\[
F_1(z^{-1}) = \frac{0.33549 z^{-1} - 0.047793 z^{-2}}{1 - 0.83697 z^{-1} - 0.039754 z^{-2}},
\]

\[
F_2(z^{-1}) = \frac{0.41086 z^{-1} - 0.021659 z^{-2}}{1 - 0.8633 z^{-1} - 0.055068 z^{-2}},
\]

\[
F_3(z^{-1}) = \frac{0.36985 z^{-1} - 0.0079934 z^{-2}}{1 - 0.9234 z^{-1} - 0.060727 z^{-2}}.
\]

In order to evaluate the obtained global model, a validation step is worked out where other inputs which are different from those used for classification are fed into the system. Then, the real and the multimodel outputs are compared. Let us consider the following input sequence:

\[
u(k) = 1 + \sin(0.08k).
\]

The validation results are given in figures 4 and 5.

\(y_{mm_r}, y_{mm_s}\): multimodel outputs obtained by using respectively the methods of reinforced and simple validities.

\[
y_{mm_r}(k) = \sum_{i=1}^{N} y_i(k) \nu_i^{renf}(k), y_{mm_s}(k) = \sum_{i=1}^{N} y_i(k) \nu_i^{simp}(k).
\]

\(e_r, e_s\): relative errors between the real and the multimodel outputs.

\[
e_r(k) = \left| \frac{y(k) - y_{mm_r}(k)}{y(k)} \right|, e_s(k) = \left| \frac{y(k) - y_{mm_s}(k)}{y(k)} \right|.
\]
It can be seen that the multimodel output, obtained by the fusion of the base-models’ outputs weighted by the reinforced validities, follows the real output with a negligible error ($e_r(t)$). This error is more important when applying the simple validities method ($e_s(t)$). This allows to conclude that, for this kind of system, the reinforced validities method is more appropriate than the simple validities one.

### 3.2 Example 2: second order continuous system with input/output-dependent parameters

As a second simulation example, we consider the system whose evolution is described by the following equations:

$$a_0 y + a_1 y + \dot{y} = u + b \dot{u},$$  \hspace{1cm} (20)

$$a_0(y) = 0.3 + \text{sigm}(y - 2), \quad a_1(y) = \text{sat}(y^2), \quad b(u) = \text{sat}(u),$$  \hspace{1cm} (21)

$\text{sigm}$ is the sigmoid function; $\text{sat}$ is the saturation function.

The considered system is complex and strongly non-linear with parameters being functions of both the input and the output, which makes the modeling task difficult.

The chosen excitation signal (figure 6) is a sine curve distorted by a random uniform signal since this input is richer than a simple random signal and allows considering the complex and non-linear aspects of the system. It is worthy to note that signal parameters (frequency, amplitude) need to be adjusted in order to obtain good results.

Moreover, in order to describe the system dynamics, the number of variables used for classification is increased. In fact, not only the output data $y(k)$ and $y(k-1)$ are considered but also the input data...
After generating the output, a sampling of the input and output signals followed, with an adequate sampling period, in order to collect the different measurements: $y(k)$, $y(k-1)$, and $u(k-1)$.

Once the number of clusters determined, the numerical data are presented to a Kohonen network owning three neurons in both input and output layers. At the end of the learning procedure, three data sets (figure 7) are obtained, each of which being used for the identification of the corresponding model.

The application of the instrumental determinants’ ratio-test method allows us to determine the three models’ orders which are respectively: two, three and two. The corresponding transfer functions are given by the following expressions:

$$F_1(z^{-1}) = \frac{0.058303 z^{-1} + 0.052302 z^{-2}}{1 - 0.43246 z^{-1} - 0.45748 z^{-2}}.$$  \hspace{1cm} (22)

$$F_2(z^{-1}) = \frac{-0.023815 z^{-1} + 0.0013593 z^{-2} + 0.031481 z^{-3}}{1 - 0.68802 z^{-1} - 0.24025 z^{-2} - 0.07212 z^{-3}}.$$  \hspace{1cm} (23)

$$F_3(z^{-1}) = \frac{0.12995 z^{-1} - 0.014359 z^{-2}}{1 - 0.64023 z^{-1} - 0.37958 z^{-2}}.$$  \hspace{1cm} (24)

The application of the following input sequence is the purpose of the validation step:

$$u(k) = 1 + \sin(k).$$  \hspace{1cm} (25)

The validation results are illustrated by the figures 8 and 9.

Contrary to the first example, the results obtained by the application of the simple validities’ method are much better than those given by the reinforced validities’ method. In fact, the figures 8 and 9 show that the multimodel output $y_{mm}(t)$, deducted by fusion of the base-models’ outputs weighted by the simple validities, follows almost perfectly the real output with an error $e_s(t)$ nearly null compared to the
error \( e_r(t) \) which is relatively important.  

In order to give prominence to the capacity of the identified models to reproduce the operating system in different domains, let us consider another input sequence given by:

\[
 u(k) = 1.2 + 1.5 \sin(2k) .
\]  

(26)

The multimodel output is generated by application of the simple validities’ method. The result is given in figure 10.

![Figure 10: Real and multimodel outputs (second validation)](image)

3.3 Interpretations

Referring to the obtained simulation results, it can be noticed that the application of the suggested approach allows a good modeling of the considered systems.

We can also make an important interpretation about the use of the validities’ calculation methods. As remarked before, in some cases simple validities give better results and in other cases it is preferable to use reinforced validities. So, more simulations and observations were worked out to conclude that the choice of the validities’ calculation method depends on the classification results i.e. the clusters structure and repartition. Thus, it can be noted that when there are several variations in the same cluster and when an overlapping between clusters occurs, which is the case in the second simulation example (figure 7), it is worth to use the simple validities’ method since it takes account of different models’ outputs referring to the expression (7). In this case, no model could represent ideally the process at any time. But when the clusters present very few variations and are well separated (figure 2), the reinforced validities’ method is better-adapted. The application of this method, thanks to the reinforcement expression (8), promotes the contribution of the most dominant model which represents at best the process behavior.

4 Evaluation of the suggested modeling strategy

In order to highlight the interest and the performance of the proposed modeling strategy, the classical modeling approach involving the identification of a global model was carried out. Then, the results given by the suggested approach were compared to those given by classical modeling strategy.

Let us consider the second simulation example described by the equations (20) and (21) and the excitation signal given by figure 6. By using the instrumental determinants’ ratio-test method for the estimation of the order and the recursive least-squares method for the parametric identification, the transfer function \( F(z^{-1}) \) of the global model is given by the following expression:

\[
 F(z^{-1}) = \frac{0.03601 z^{-1} + 0.0035236 z^{-2}}{1 - 0.55726 z^{-1} - 0.42595 z^{-2}} .
\]  

(27)

By using the same numerical data, the classification and the identification steps give the three models described by equations (22), (23) and (24). The input sequence given by the equation (25) was applied
again. Figure 11 represents the evolutions of the real, the multimodel and the global model outputs. The multimodel output is obtained by fusion of the models’ outputs weighted by their simple validities degrees calculated at each instant.

![Figure 11: Evolutions of the real, multimodel and global model outputs](image)

The adopted multimodel approach offers a very satisfactory precision compared to the case of the classical modeling approach based on the unique global model.

## 5 Conclusion

In this paper, a neural approach of multimodel representation is proposed. This approach is applicable when dealing with complex, strongly non-linear and uncertain processes. It allows the determination of the models’ base by using a Kohonen network and two methods of structural and parametric identification. The different steps were detailed. The multimodel output is obtained by using the technique of fusion and the adequate validities’ computation method. The suggested approach has been implemented and tested for different complex systems. Simulation results, two of which were described in this paper, prove the efficiency and the precision of the proposed modeling strategy and show that the method works well with various processes even when highly complex. Some interpretations have been made about the choice of the adequate validities’ calculation method to be applied. Furthermore, in order to demonstrate the performance and the relevance of the suggested approach, a comparison with the classical modeling approach has been made.

## Bibliography


Nesrine Elfelly¹, Jean-Yves Dieulot², Pierre Borne³

¹Université des Sciences et Technologies de Lille (USTL), ²Ecole Polytechnique de Lille, ³Ecole Centrale de Lille (EC Lille)

Laboratoire d’Automatique, Génie Informatique et Signal
Ecole Centrale de Lille, Cité scientifique BP 48
59651 Villeneuve d’Ascq Cedex, France
E-mail: nesrine.elfelly@ed.univ-lille1.fr, jean-yves.dieulot@polytech-lille.fr, pierre.borne@ec-lille.fr

Received: December 4, 2007

Nesrine Elfelly was born in Monastir, Tunisia in 1982. She received the Engineer Diploma degree in Computer Engineering from the "Ecole Nationale d’Ingénieur de Tunis" (Tunisia) in 2006. She obtained the Master of Automatic Control, Computer Engineering and Image Processing from the "Ecole Centrale de Lille" (France) in 2007 and the Master of Automatic Control and Signal Processing from the "Ecole Nationale d’Ingénieur de Tunis" in the same year. She is currently a PhD student in Automatic Control and Applied Computing within the framework of LAGIS-USTL and LARA-ENIT cooperation. Her current research interests are the applications of neural approaches for complex systems analysis and control.

Jean-Yves Dieulot graduated from French Ecole d’Ingénieurs Institut Industriel du Nord in 1990, and obtained a PhD from the University of Lille in 1993. He is currently Associate Professor at Ecole Polytechnique Universitaire de Lille, France, and with LAGIS (Laboratory of Automatic Control). His main interests are in fuzzy, nonlinear and robot control.

Pierre Borne was born in Corbeil, France in 1944, he received the Master degree of Physics in 1967, the Masters of Electronics, of Mechanics and of Applied Mathematics in 1968. The same year he obtained the Diploma of "Ingénieur IDN" (French "Grande Ecole"). He obtained the PhD in Automatic Control of the University of Lille in 1970 and the DSc in Physics of the same University in 1976. He became Doctor Honoris Causa of the Moscow Institute of Electronics and Mathematics (Russia) in 1999, of the University of Waterloo (Canada) in 2006 and the Polytechnic University of Bucharest (Romania) in 2007. He is author or coauthor of about 200 Journal articles and book chapters, of 35 plenary lectures and of about 250 communications in international conferences. He has been the supervisor of 69 PhD thesis and is author of 20 books. He is Fellow of IEEE and has been President of the IEEE/SMC society in 2000 and 2001. He is presently Professor "de classe exceptionnelle" at the Ecole Centrale de Lille and director of the French "plan pluriformation" of research in Automatic Control.