Valeriy ROGOZA, Joanna BOBKOWSKA

Westpomeranian Technological University in Szczecin

Using sensitivity functions to simulation of complex processes

Abstract. *The paper deals with the use of sensitivity functions that allow us to develop mathematical models of complex processes using short-time experimental samples. The process is said to be complex if variables which describe the states of the process in time are interrelated. The models of complex processes proposed in the paper are presented in the form of regression equations which can be used for the analysis of mutual influences of process variables as well as for the short-time prediction of future process states. The discussed approach is based on the assumption that the process to be studied exhibits the regularity property. As is shown in the paper, under this condition, it is sufficient to have five or six experimental samples to start synthesis of models which can be further modified during simulation.*

Streszczenie. *Artykuł dotyczy wykorzystania funkcji czułości đo konstrukcji matematycznych modeli złożonych procesów w celu dokonywania krótkoterminowej predykcji. Proces nazywamy złożonym, kiedy zmienne opisujące stany tego procesu w czasie są wzajemnie zależne. Modele złożonych procesów, które proponujemy są przedstawione w postaci układów rownan, które mogą być używane do analizowania wzajemnych wpływów zmiennych tych procesów oraz krótkoterminowej prognozy przyszłych stanów tych procesów. Proponowane podejście oparte jeśt na założeniu, że obserwowany proces wykazuje pewne regularności. W takiej sytuacji wystarczy pięć, sześć próbek daných eksperymentalnych, aby rozpocząc syntezę mođeli, które następnie będą modyfikowane podczas symulacji. (**Wykorzystanie funkcji czułości do symulacji procesów**)*

Keywords: sensitivity functions, mathematical simulation, analysis of complex processes, short-time prediction.
Słowa kluczowe: funkcje czułości, symulacja matematyczna, analiza złożonych procesów.

Introduction

Traditionally, computer investigations of any process, be it chemical, technological, economical or ecological process, is beginning with the development of mathematical models which offer a view of how the process is formed, which process state variables are dominant, and what variable values might be in the future. By convention, the types of mathematical models used in the simulation of processes can be divided into the two wide classes: stochastic models and deterministic models. Building stochastic models is based on the analysis of a great body of experimental data using probabilistic methods and methods of mathematical statistics. Such an approach has gained wide-spread acceptance in the solution of problems everywhere when we deal with a huge data sets gathered during a relatively long time, what allow us to calculate statistical parameters (the parameter expectations, the statistical dispersions, etc.) and use them for purposes of deep analysis of sophisticated processes in nature and society [1].

The deterministic approach is based on building models in the form of equations which approximate certain functional relationships between the state variables. Such an approach is the basis of the theory of identification of systems [2]. In solving the classical identification problem, it is assumed that the structure of the analysed system and the class of models which describe the system states, are predetermined. Such a formulation of the problem is inherent in the theory of system control, where the approach finds its major use. In this case, a typical situation is that the researcher possesses a little information on the previous states of the system (or process) to be studied, but it is required to generate a high-rate response on the varying external conditions by making near to optimal decision at least one or two time steps ahead.

In this paper, we would like to discuss the deterministic method of analysis of complex process which as distinct from the classical methods of system identification does not require any a priori information on the structural and functional relationships between the state variables.

As was mentioned above, it is suggested that the method to be considered works if some conditions are met. We assume that in practice the researcher may possess a very limited number of experimental data and it is required to study a process although one knows nothing on its regularity. Statistical methods give no way of overcoming the lack data is not capable to offer extra-data on the object considered. Among other things it means that if experimental data samples, which are at the researcher's disposition, contain data that are significantly irregular (because of the noise, the low accuracy of measurements or experimental data indeterminacy), there is no way to convert irregular data to regular, and the method may be failed.

On the other hand, if the data samples exhibit behavior near to regular, have the low level of measurement errors, and contain almost no gaps in data, a regularity within them must be immediately exhibited at the beginning of investigations. In this case, there is a principal possibility to construct mathematical models of the studied process using the deterministic approach.

There are some special features in simulation of processes depending on whether the process is trivial or complex. The process is called simple if the variables which describe the state of the process in time (state variables) are not interrelated, and the process is complex if the mentioned interrelation between the state variables takes place. The most frequently situation is complicated by the fact that the interrelation between the state variables is sophisticated in character and can be approximated by nonlinear functions. It is clear that in this situation, interpolation methods give no way of simulating such complex processes, and special methods are required.

General strategy of model synthesis

The problem is that a small quantity of experimental points does not allow us, in principle, to specify the behavior of complex process during a long-time interval. It is difficult to detect periodical changes of state variables within a long-time interval.

Selecting experimental data on the process studied, we form simplified process models building regression equations. This is a training stage of model synthesis in which models are adjusted to the given data. Next, using created models, we try to predict the values of process state variables at the few subsequent time points. This is a testing stage of model synthesis. At this stage we can run into the following situations: if the predicted system state variables at the next time points, say, t_{i+1} and t_{i+2} are sufficiently close to the actual values given experimentally, we do not need any adjustments of models for a few next time points. This could mean that either the period of changes of the state variables is too long that it is not visible yet, or this period has already been covered by our models. In the both cases, we expand the scope of prediction for about four-five time steps using the existing models. Then we check the accuracy of the models using the testing data set once more, etc.

If we revealed drastic discrepancy between the predicted and experimental values for subsequent time points t_{i+1} and t_{i+2}, this might mean that actual state values of the process are significantly noisy and/or irregular, and/or has a large period. Any of the mentioned situations cannot be detected using short-time sampling. In this case, we can try to improve the accuracy of models extending the range of training time point by including data on points t_{i+1} and t_{i+2} to the training set and repeating the training stage once more. This stage would be named *the stage of model correction*.

Next we proceed by the same way through prediction and, if need be, correct the models in the following time points t_{i+3} and t_{i+4}, and so on. Finally, if the process to be studied is near to regular, we approach step by step to the models which practically satisfy the requirements of accuracy.

Thus, as can be seen, the strategy of constructing models is dynamic and includes the following three stages: training, testing/prediction, and correction. To perform this process we divide the set of samples at least into the two (not necessarily equal) parts: a subset of training samples that initialize the process of model synthesis and a subset of testing samples. And as we noted above, some samples of the second set may pass into the first one. The difficulty of practical implementation of this strategy lies in the complexity of the process of control over the model synthesis procedure. The researcher does not know a priori what number of time points he really needs to build the desired models and, in particular, what number of samples should be selected for each of the above two subsets. In this situation the researcher should supplement the insufficient data (if necessary) in the course of solving the problem of prediction comparing the results of prediction with the actual values of variables.

Another problem is that even if created models seem to be accurate within the observed time interval, there is no assurance that the accuracy would be saved during the entire interval of process study because the process might be regular only at a certain one. Thus, the specific strategy of building models which are designed dynamically using short-time samples are conditioned by the specific features of processes for which these models are formed. And the third problem is caused by the specific feature of fast-acting processes. This specific implies that state variables might have essentially different influences on the process progress at different time intervals of observation. In these circumstances it makes sense to build a series of relatively simple partial models that include combinations of a few state variables rather than a single surplus model containing all the state variables. Different partial models can be used at different partial time intervals corresponding to their accuracy within the certain partial time intervals.

Governing by the above reasoning we selected the Kolmogorov-Gabor regression equations[3] for our analysis as is described below.

Building partial regression models for the analysis of complex processes

Let $x_1, x_2, ..., x_n$ be the set of variables of a process (or an object or a system) describing the states of the process in different time instances. These variables we call the state variables of the process. The set of n variables allows us to describe the states of the process in the n-dimensional real metric space $(x_1, x_2, ..., x_n)\epsilon\Re^n$. At each instant $t_0, t_1, ..., t_i, ...$ state variables determine points $P(0), P(1), ..., P(i), ...$ in the $\Re^n$ space. The geometric interpretation of this situation is convenient because it allows us to specify the distance between the states in terms of the metric spaces. According to our problem we have the series of experimental samples at instances $t_0, t_1, ..., t_i, ...$, each can be presented by a certain point $P(0), P(1), ..., P(i)$ in n-dimensional state space $\Re^n$.

The solution of the problem can be divided into two steps. At the first step we create models of the process and use them to the solution of the analysis problem, and at the second step we take advantage of the design model to solve problems of prediction.

In this paper we would like to show a way to the solution of the first described problem by the construction and application the regression equations using the sensitivity functions [4]. Determining the impact of each state variable on the overall process is the goal of analysis. If we know the impact of each state variable on the process we can control the behaviour of the process changing some of the state variables.

If the state variables of the process are mutually independent we deal with the simple process as it is called. But if in the simple process dependences between input and output variables are nonlinear, standard interpolation methods can be used to approximate these dependences.

While the variables of the process are mutually dependent, the process is called complex. The control of such a process considerably complicates. As a result of the interdependence of states variables, changing the value of the variable $x_1(t_i)$ may cause changes in the values of some other states variable $x_2, x_3, ...$, etc., and the latter may affect the value of the variable $x_1(t_i)$ increasing or reducing its value. This effect is similar to the effect of the impact of the complex feedback in the electronic amplifiers.

Unfortunately, our problem is more complex because, as a rule, we deal with a lot number of variables which can change simultaneously.

The current problem is creating a method we can use in control of complex processes and prognosis that allows to create the model with a small number of testing samples. The method we will use is based on the sensitivity functions.

We will use semi-relative sensitivity functions. We will be searching mathematical models in the class of polynomial models for these functions in relation to the classical GMDH method in the form of Kolmogorov-Gabor polynomials (1):

$$Y = a_0 + \sum_{i=1}^{n} a_i x_i + \sum_{j=1}^{n} \sum_{i \leq j} a_{ij} x_i x_j + \\ + \sum_{i=1}^{n} \sum_{j \leq i} \sum_{k \leq j} a_{ijk} x_i x_j x_k + ..., \tag{1}$$

where the coefficients $a_0, a_i, a_{ij}, a_{ijk}x_i x_j x_k, ...$ are unknown, under calculations, and as Y we understand any output variable $y_1(t), y_2(t), ..., y_n(t)$. Among all the possible versions of equations (1)in which the interdependence of variables $x_1(t), x_2(t), ..., x_n(t)$ may be taken into account, the simplest is second-order one of the form (2):

$$Y = a_0 + a_1 x_i + a_2 x_j + a_3 x_i^2 + a_4 x_j^2 + a_5 x_i x_j, \tag{2}$$

where, for simplicity, we use the vertical notation of coefficients $a_0, a_1, ..., a_5, i \neq j$.

The results of the experiment with regular dataset

Our method can work on a particular set of data and it may not work, unfortunately, on any of the datasets for the reasons given above. Let the (*n*-dimensional) process that we are examining be possible to control. It means we can change the values of each variable attempting to make changes to these parameters in the future. The question arises: if we change the values of some variables in *i*-point of a time $(P(i))$, what are the expected values of these parameters in the next time $P(i+1)$, or even $P(i+2), P(i+3)$, etc.? We cannot change the mutual dependencies between variables of the process but we can control this process by changing some or all of the variables in a right direction. Therefore, the problem is: what will be the actual values of the variables $x_k, k = 1, ..., n$, in following points $P(i+1), P(i+2), ...$, where each of these variables will be determined by us by the increment $ddx_k = \frac{\triangle x_k}{x_k}$, $k = 1, ..., n$, at the time point $P(i)$, and what impact has each of the variables of the process on the changes of these values. Using the terminology adopted from the general theory of control systems we can describe the process mathematically by the sets of equations:

(3)
$$\begin{aligned}(x_1(t_{i+1}), x_2(t_{i+1}), ..., x_n(t_{i+1})) = \\ = F(x_1(t_i + ddx_1(t_i), x_2(t_i) + ddx_2(t_i), \\ ..., x_n(t_i) + ddx_n(t_i))\end{aligned}$$

where F is an operator of the transition from one state of the process to another in n-dimensional metric space $(x_1, x_2, ..., x_n)$. This operator specifies the state of the process in the next time point $P(i+1)$ when we modify the variable values in $P(i)$ by the relative increases $ddx_k(t_i)$. For simplicity we analyse the process in three-dimensional space in this paper. Let x, y, z be the parameters of a certain real process. First we create the model for partial function $X(x, y)$ in the form of equation (2):

(4)
$$X(x, y) = a_0 + a_1 x + a_2 y + a_3 x^2 + a_4 y^2 + a_5 xy,$$

where $X(x, y)$ is the value of variable x in the next time point presented as a partial two-variable function (x and y) with values of previous time point. $X(x, y)$ is partial function because it does not contain all variables but only various two variable combination pairs (like in the classical GMDH method). To this end we create the system of six equations to determine the values of the coefficients $a_0, a_1, ..., a_5$.

(5)
$$\begin{aligned}P(i) = a_0 + a_1 x(t_{i-1}) + a_2 y(t_{i-1}) + \\ + a_3 x^2(t_{i-1}) + a_4 y^2(t_{i-1}) + a_5 x(t_{i-1}) y(t_{i-1})\end{aligned}$$

After solving the system of equations (5) with our dataset we obtain: $a_0 = -0.0003; a_1 = 1.1172; a_2 = -0.0155; a_3 = 0.0172; a_4 = 0.0027; a_5 = -0.0126$. Therefore, the searched model takes the form

(6)
$$\begin{aligned}X(x, y) = -0.0003 + 1.1172x - 0.0155y + \\ + 0.0172x^2 + 0.0027y^2 - 0.0126xy.\end{aligned}$$

Similarly for model $Y(y, x)$:

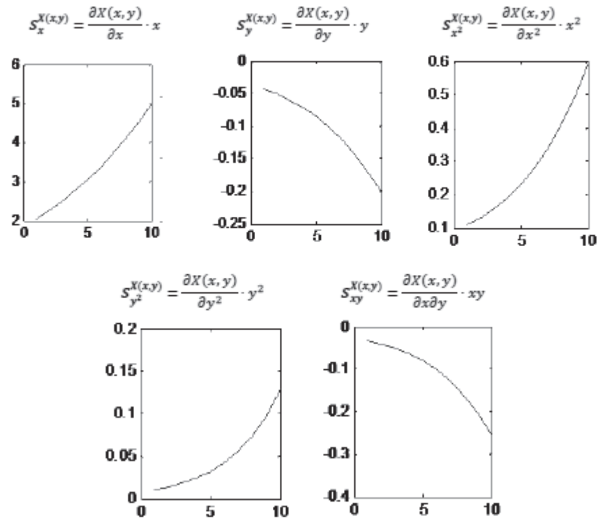

Fig. 1. The plots of $X(x, y)$ semi-relative sensitivity functions for ten first time points.

(7)
$$Y(xy, x) = b_0 + b_1 y + b_2 x + b_3 y^2 + b_4 x^2 + b_5 yx,$$

$b_0 = 0.0017; b_1 = 1.1826; b_2 = -0.0159; b_3 = -0.0023; b_4 = -0.0119; b_5 = 0.0088$

and the rest of them: $X(x, z), Z(z, x), Y(y, z), Z(z, y)$.

We calculate the values of the semi-relative sensitivity functions (first and second order differentials of equation (4)). Our dataset seems like regular. We can observe some regularities in the values of the semi-relative sensitive functions of that process on the below plots (Fig 1.) So we could try to extrapolate the values of these functions using, for example, Lagrange's interpolation polynomial. Next we find the values of increases $ddx(t_i)$ and $ddx(t_i)$ for $i = 1...6$ solving the linear set of equations (8) presenting the simplified formula for calculating the values of the variables x and y in a point $P(i+1)$ e.i functions $X_{t_{i+1}}(x, y)$ and $Y_{t_{i+1}}(y, x)$.

(8)
$$\begin{aligned}X_{t_{i+1}}(x(t_i), y(t_i)) = x(t_i) + \\ + S_{x(t_i)}^{X(x,y)} \cdot ddx(t_i) + S_{y(t_i)}^{X(x,y)} \cdot ddy(t_i) \\ Y_{t_{i+1}}(y(t_i), x(t_i)) = y(t_i) + \\ S_{y(t_i)}^{Y(y,x)} \cdot ddy(t_i) + S_{x(t_i)}^{Y(y,x)} \cdot ddx(t_i)\end{aligned}$$

We check the accuracy of the model calculating the values of all variables in the first six time points of the dataset. If the results are satisfying we calculate several next values of the partial increases $ddx(t_i)$ and $ddy(t_i)$ and sensitivity functions using Lagrange's interpolate polynomial. Next we can predict/calculate the future values of our dataset using the set of the equations (8). The results of our researches are presented in the Table 1.

If the first model or predictions are not satisfying we calculate the partial increments for the more precise model containing the sensitivity functions of the first and second order solving nonlinear set of equations (9). We proceed it for all partial functions.

Time point	Real x values	Real y values	Calc. x	Calc. y
P(0)	1.771561	1.387836	-	-
P(1)	1.948717	1.596012	1.9487	1.596
P(2)	2.143589	1.835414	2.1436	1.8354
P(3)	2.357948	2.110726	2.3579	2.1107
P(4)	2.593742	2.427335	2.5937	2.4273
P(5)	2.853117	2.791435	2.8531	2,7914
P(6)	3.138428	3.21015	**3.1384**	**3.2091**
P(7)	3.452271	3.691673	**3.4523**	**3.6903**
P(8)	3.797498	4,245423	**3.7975**	**4.2437**

Table 1. The comparison of achieved results for partial model $X(x, y)$

Time point	Real x values	Real z values	Calc. x	Calc. z
P(0)	1.771561	4.464961	-	-
P(1)	1.948717	3.907302	1.9487	3.9073
P(2)	2.143589	3.660721	2.1436	3.6607
P(3)	2.357948	3.590333	2.3579	3.5903
P(4)	2.593742	3.638608	2.5937	3.6386
P(5)	2.853117	3.778778	2.8531	3.7788
P(6)	3.138428	3.998615	**3.1384**	**3.9986**
P(7)	3.452271	4.293792	**3.4523**	**4.2992**
P(8)	3.797498	4.664874	**3.7972**	**4.7021**

Table 2. The comparison of achieved results for partial model $X(x, z)$

$$(9)\quad \begin{aligned} X_{t_{i+1}}(x(t_i), y(t_i)) = x(t_i) + S^{X(x,y)}_{x(t_i)} \cdot ddx(t_i) + \\ + S^{X(x,y)}_{y(t_i)} \cdot ddy(t_i) + 0.5 \cdot S^{X(x,y)}_{x^2(t_i)} \cdot (ddx(t_i))^2 + \\ +0.5 \cdot S^{X(x,y)}_{y^2(t_i)} \cdot (ddy(t_i))^2 + S^{X(x,y)}_{x(t_i)y(t_i)} \cdot (ddx(t_i)) \cdot (ddy(t_i)) \\ Y_{t_{i+1}}(y(t_i), x(t_i)) = y(t_i) + S^{Y(y,x)}_{y(t_i)} \cdot ddy(t_i) + \\ + S^{Y(y,x)}_{x(t_i)} \cdot ddx(t_i) + 0.5 \cdot S^{Y(y,x)}_{y^2(t_i)} \cdot (ddy(t_i))^2 + \\ +0.5 \cdot S^{Y(y,x)}_{x^2(t_i)} \cdot (ddx(t_i))^2 + S^{Y(y,x)}_{y(t_i)x(t_i)} \cdot (ddy(t_i)) \cdot (ddx(t_i)) \end{aligned}$$

Next we check the accuracy of the models again and choose that model which fits the value of each variable most. Then we use these models for prediction.

The results we obtained for the partial model $X(x, y)$ are presented in Table 1. In Table 2 and 3 we present the results od model $X(x, z)$ and $Y(y, z)$ respectively. First six values were used for training the model, bolded values are predicted ones.

We achieved similar accuracy for the rest of models. We were using that method beginning with various time points and the results were also satisfying.

Conclusion

We showed that it is sufficient to have five or six experimental samples to start synthesis of models capable to deal with prediction problem with satisfying accuracy. The reason that our method works (for the sufficiently regular processes) is that the sensitivity functions and partial increments of the variables can be extrapolated as any other functions. Moreover, if the processes are almost regular, sensitivity functions are sufficiently smooth and thus their future values can be found with fairly high accuracy. Substituting extrapolated values of the sensitivity functions and the partial increments for general formulas you can roughly specify the values of all variables in the future moments of the time.

Time point	Real y values	Real z values	Calc. y	Calc. z
P(0)	1,387836	4.464961	-	-
P(1)	1.596012	3.907302	1.5960	3.9073
P(2)	1.835414	3.660721	1.8354	3.6607
P(3)	2.110726	3.590333	2.1107	3.5903
P(4)	2.427335	3.638608	2.4273	3.6386
P(5)	2.791435	3.778778	2.7914	3.7788
P(6)	3.21015	3.998615	**3.2102**	**3.9986**
P(7)	3.691673	4.293792	**3.6917**	**4.3737**
P(8)	4.245423	4.664874	**4.2454**	**5.4103**

Table 3. The comparison of achieved results for partial model $Y(y, z)$

REFERENCES

[1] Barsegian A. A., Kuprijanow M. S. et al: Analiza danych i procesów, Sankt-Petersburg. BHW-Petersburg, 2009, (in Russian).

[2] Bellman R.: Dynamic programming, system identyfication and suboptimization, Journal of the SIAM on Control, 1966, v. 4. No. 1.

[3] H.R. Madala, A.G. Ivakhnenko. Inductive Learning Algorithms for Complex Systems Modeling. CRC Press, Boca Raton, 1994.

[4] H. T. Banks, S. Dediu and S.E. Ernstberger Sensitivity functions and their uses in inverse problems. CRSC-TR07-12, July, 2007; J. Inverse and Ill-posed Problems, 15 (2007), 683-708.

***Authors**: Prof. Valeriy Rogoza, Ph.D. student Joanna Bobkowska Department of Information Technologies, Westpomeranian Technological University in Szczecin, ul. Żołnierska 52, 71-210 Szczecin, Poland, email: vrogoza@wi.zut.edu.pl*