Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 50-1 (2017) 1559–1564

Black-box identification of a pilot-scale dryer model: A Support Vector Regression and an Imperialist Competitive Algorithm approach

Robert Sałat*, Michał Awtoniuk**, Krzysztof Korpysz***

Warsaw University of Life Sciences, Faculty of Production Engineering, Nowoursynowska 164, 02-787 Warsaw, Poland
* Tel.: +48 22 59 34 608; e-mail: robert_salat@sggw.pl
** e-mail: michal_awtoniuk@sggw.pl
*** e-mail: krzysztof_korpysz@sggw.pl

Abstract: This report describes system identification by means of the hybrid black-box method. The identification was carried out on the regulation object (a dryer model) representing the pilot-scale processes occurring during air conditioning and drying. The new approach proposed in the paper is the use of the imperialist competitive algorithm (ICA) as a tool for selecting the best parameters for support vector regression (SVR) and for selecting an optimal set of regressors. The advantage of this method is that the selection of an optimal set of regressors and the optimal parameters of SVR for this set is performed automatically, which reduces the time needed for identification. The results of the SVR with the ARX, ARMAX, OE, Box-Jenkins (BJ) and low-order transfer function (Tf) models were compared. The research was conducted for two fan speeds, equal to 40% and 60%. The Fit and MSE indicators for the SVR achieved a higher value with respect to those of the ARX, ARMAX, OE, BJ and Tf models. This method is sufficiently universal and can be applied to any plant as an efficient alternative method. This report is supported by National Science Centre grant UMO-2015/17/B/NZ7/02937

Keywords: black box, system identification, support vector machine, imperialist competitive algorithm.

1. INTRODUCTION

Good knowledge of the behavior of the system in the control process is essential for providing high quality products and for achieving low production costs. Object (plant) identification (considered as components of the system) consists of constructing a mathematical description of the object dynamics on the basis of measured data and forecasting the behavior of the object for new input data. Mathematical models resulting from the identification are used for the purpose of controlling the simulation tests and forecasting. However, it is difficult to find a mathematical model that is capable of accurately mapping the dynamic behavior of the system. Most of the entity industrial systems are non-linear, resulting in very complex problems in identification and control. In industry, linear models are usually built around the set points. These models have been well tested and described in the literature. Among scientists, there is now a tendency to use and test the non-linear models of systems (Juditsky et al., 1995). If we have a priori knowledge and know the laws of physics that occur in the system, then we can use identification based on a white box. However, if the knowledge is incomplete, i.e., a model is created on the basis of the laws of physics but a number of parameters must be estimated from the measurement data, then we can use a grey-box model. However, if we do not have sufficient knowledge about the system or if the system is very complex, it is convenient to use black-box type identification i.e., identification based only on the measured input and output signals (Sjoberg et al., 1995). Black-box type identification may involve both linear systems and non-linear systems. There are many algorithms that use the black-box method, e.g., gradient-based methods, which, however, do not guarantee the achievement of the global minimum. In recent years, several authors (Gretton et al., 2001; Ranković and Nikolić, 2008) suggested using the method of artificial intelligence in black-box identification. This suggestion resulted in a large number of works in which different techniques were presented, such as artificial neural networks (ANN) (Rivals et al., 1999), wavelets (Zhang and Benveniste, 1992), fuzzy-neural networks (FNN) (Bernd et al., 1999), and an imperialist competitive algorithm (ICA) (Winiczenko et al., 2013). Models created on the basis of identification can be used both as simulators (e.g., stability analysis of the plant) and as models for use in control systems. Among the neural networks used in black-box identification, the most frequently reported networks are the multi-layer perceptron (MLP), fuzzy-neural networks (FNN), radial basis function (RBF), Runge-Kutta (RK) method, and digital recurrent network (DRN). Few authors used support vector machines (SVM) (Gretton et al., 2001) as an identification tool. Typically, the authors also ignore the rules of the selection of the regressors and parameters of both the ANN and SVM. In this paper, a black-box hybrid method was proposed for the identification of a pilot-scale dryer model (plant). Preliminary research for this method was presented in (Salat et al, 2013) and was limited to the simple algorithm. The new approach proposed in the paper is the use of an evolutionary algorithm (e.g., ICA) as a tool for selecting the best parameters of support vector regression (SVR) and for selecting an optimal

10.1016/j.ifacol.2017.08.309

set of regressors. The advantage of this method is that the selection an optimal set of regressors as well as the optimal parameters of SVR for this set is performed automatically, which reduces the time needed for identification. Because linear models are very often used in the identification of an industrial process, the results of using SVR and ICA with ARX, ARMAX, OE, BJ (Ljung, 1999) and Tf models, which are available in the Matlab System Identification Toolbox R2010b, were compared. The research was conducted for two fan speeds, equal to 40% and 60%. This method is sufficiently universal and can be applied to any plant.

2. BLACK-BOX IDENTIFICATION

Because SVR was taken as a non-linear network, black-box identification will be described with a focus on non-linear models (Juditsky *et al.*, 1995). Assuming that $u(t)$ is a discrete input of a non-linear system, and $y(t)$ is an output, the data set can be defined as:

$$\mathbf{u}=\left[u(1),u(2),\ldots,u(t)\right],\ \mathbf{y}=\left[y(1),y(2),\ldots,y(t)\right] \quad (1)$$

where $u(1)$, …, $u(t)$, $y(1)$, …, $y(t)$ are sequential samples in time.

The idea of black-box identification is to find the connection between past observations, i.e., [$u(t-1)$,.., $u(t-n_u)$, $y(t-1)$,…, $y(t-n_y)$], and the future $y(t)$. The primary goal is estimating a non-linear function as follows:

$$y(t)=f\left(\left[u(t-1),\ldots,u(t-n_u),y(t-1),\ldots,y(t-n_y)\right],\boldsymbol{\theta}\right)+e(t) \quad (2)$$

Introducing $\mathbf{x}(t)$=[$u(t-1)$, …, $u(t-n_u)$, $y(t-1)$, …, $y(t-n_y)$], (2) can be expressed as follows:

$$y(t)=f\left(\mathbf{x}(t),\boldsymbol{\theta}\right)+e(t) \quad (3)$$

where $f(\cdot)$ is the unknown basis function, $\mathbf{x}(t)$ is the regression vector, n_u is the number of past observation samples of the input, n_y is the number of past observation samples of the output, $\boldsymbol{\theta}$ is a vector of associated parameters, $e(t)$is the noise and its value should be as small as possible.

Parameterizing the unknown basis function $f(\cdot)$ can usually be approximated by a finite dimensional vector $\boldsymbol{\theta}$. The vector of associated parameters ($\boldsymbol{\theta}$) can be assessed by means of a fit between the model and the data set:

$$\sum_{i=1}^{N}\left\|y(t)-f\left(\left[u(t-1),\ldots,u(t-n_u),y(t-1),\ldots,y(t-n_y)\right],\boldsymbol{\theta}\right)\right\|^2 \quad (4)$$

The elements of vector $\mathbf{x}(t)$ are called regressors and, depending on the choice of the vector of regressors, we can build a variety of models: NFIR, NARX, NARMAX, NOE, NBJ (Sjoberg *et al.*, 1995). The problem of identification for each model comes down to finding the number of regressors, i.e., the number of past observation samples and the non-linear function $f(\cdot)$, to obtain the model with the best fit to the object. Finding a suitable number of regressors is very important. An excessive number of regressors may lead to the creation of a very complex model and variable duplication. A large number of irrelevant regressors can cause overfitting, which results in a poor generalizability of the model.

3. SUPPORT VECTOR REGRESSION (SVR) IN PROBLEM IDENTIFICATION

Support vector machines (SVMs) that address modeling and prediction are called support vector regressions (SVRs) (Smola and Scholkopf, 1998). Because support vector machines (SVMs) and support vector regression (SVR) are often presented in the literature (Salat and Salat, 2013; Salat and Osowski, 2011; Wen *et al.*, 2006), only a basic presentation of SVR in problem identification was presented. To present the algorithm of SVR, it must first be introduced on the projection function F from the regressors' space $\Re$ to a hypothetical feature space $\aleph$. Assume that the training set is as follows:

$$(\mathbf{x}_1,y_1),\ldots,(\mathbf{x}_n,y_n)\in\Re^N\times\Re \quad (5)$$

where x_n is the input vector composed of the regressors for the NARX model, and y_n is output.

By defining the ε – insensitive loss function as

$$\left|y-f(\mathbf{x})\right|_\varepsilon=max\left\{0,\left|y-f(\mathbf{x})\right|-\varepsilon\right\} \quad (6)$$

the procedure can be treated as quadratic programming (QP), and the estimation function can be expressed in the "standard" SVR as

$$f(\mathbf{x})=(\mathbf{w}\cdot F(\mathbf{x}))+b,\quad \mathbf{w},\mathbf{x}\in\Re^n,b\in\Re \quad (7)$$

where $\mathbf{w}$ is the weight vector and b is the bias.

Then, $f(\mathbf{x})$ can be determined from the minimization problem as follows:

$$min\ \frac{1}{n}\sum_{i=1}^{n}(\left|y_i-f(\mathbf{x}_i)\right|-\varepsilon)=\frac{1}{n}\sum_{i=1}^{n}(\left|y_i-\mathbf{w}\cdot F(\mathbf{x}_i)+b\right|-\varepsilon) \quad (8)$$

where n is the number of training data pairs.

Introducing slack variables ξ_i, ξ_i* into (8), an optimization problem can be formulated as follows:

$$\min_{\mathbf{w},\xi,\xi_i^*}\ \frac{1}{2}\left\|\mathbf{w}\right\|^2+C\sum_{i=1}^{n}\xi_i+C\sum_{i=1}^{n}\xi_i^* \quad (9)$$

which is subject to:

$$y_i-\mathbf{w}\cdot F(\mathbf{x}_i)\le\varepsilon+\xi_i,\ \mathbf{w}\cdot F(\mathbf{x}_i)-y_i\le\varepsilon+\xi_i^*,\ \xi_i,\xi_i^*\ge 0 \quad (10)$$

where C, ε are the user-specified constants.

The optimization problem can be easily solved by the introduction of the Lagrange function and the formulation of the so-called dual problem regarding the Lagrange multiplier α and the introduction of the kernel term

$$k(\mathbf{x}_i,\mathbf{x}_j)=F^T(\mathbf{x}_i)F(\mathbf{x}_j) \quad (11)$$

that is defined in accordance with Mercer's theorem (Vapnik and Chervonenkis, 1974). The formulation of the dual problem is equivalent to finding an expression as follows:

$$\min_{\alpha,\alpha^*}\ \frac{1}{2}(\alpha-\alpha^*)^T Q(\alpha-\alpha^*)+\varepsilon\sum_{i=1}^{p}(\alpha_i+\alpha_i^*)+\sum_{i=1}^{p}y_i(\alpha_i-\alpha_i^*) \quad (12)$$

which is subject to: