Available online at www.sciencedirect.com

ScienceDirect

IFAC-PapersOnLine 49-7 (2016) 371–376

Practical Use of Savitzky-Golay Filtering-Based Ensemble Online SVR

Hiromasa Kaneko*, Takuya Matsumoto, Shigeki Ootakara***, Kimito Funatsu***

** Department of Chemical System Engineering, The University of Tokyo, Tokyo, Japan (e-mail: hkaneko@chemsys.t.u-tokyo.ac.jp, funatsu@chemsys.t.u-tokyo.ac.jp).*
*** Process Technology Center, Mitsui Chemicals Inc., Chiba, Japan (e-mail: Takuya.Matsumoto@mitsuichemicals.com)*
**** Osaka Works, Mitsui Chemicals, Inc., Osaka, Japan (e-mail: Shigeki.Ootakara@mitsuichemicals.com)*

Abstract: As a result of collaboration between Mitsui Chemicals, Inc. and the University of Tokyo, a soft sensor tool was developed and implemented in several plants in Mitsui Chemicals, Inc. A soft sensor is an inferential model constructed between process variables that are easy to measure (X) and process variables that are difficult to measure (y). y-values can be estimated in real time by inputting X-values into a soft sensor. To maintain predictive ability of a soft sensor to be high, we employ ensemble online support vector regression (EOSVR) model as an adaptive soft sensor model, which can adapt to both nonlinear changes and time-varying changes. Additionally, to reduce noise in estimated y-values, Savitzky-Golay (SG) filtering is used for estimated y-values. Our proposed method is called EOSVR-SG and implemented as a soft sensor tool. In this paper, we show our soft sensor tool used in real chemical plants and its execution results in which the EOSVR-SG model could estimate y-values accurately and smoothly.

Keywords: Process control, Soft sensor, Support vector regression, Adaptive model, Ensemble, Noise reduction, Smoothing, Savitzky-Golay filtering.

1. INTRODUCTION

Soft sensors are widely used to predict process variables that are difficult to measure online (Kano and Nakagawa, 2008; Kadlec *et al.*, 2009). An inferential model is constructed between process variables that are easy to measure online, which are called X-variables, and process variables that are difficult to measure online, which are called y-variables. The values of y can then be predicted using that model. Through the use of soft sensors, values of y can be predicted with a high degree of accuracy. Both lab samples and measurements of online analyzers are examples of y-variables.

One of the crucial difficulties of soft sensors is that predictive accuracy drops because of changes in the state of chemical plants. This is called soft sensor model degradation (Kaneko and Funatsu, 2013a). To reduce degradation, the model is reconstructed with the newest data. For instance, a moving window (MW) model (Qin, 1998; Kadlec and Gabrys, 2010) is constructed with the data that were measured most recently, and a just-in-time (JIT) model (Schaal *et al.*, 2002; Cheng and Chiu, 2004) is constructed with data that are more similar to a query than other data. Models such as MW, JIT, and time difference (Kaneko and Funatsu, 2011) models that have adaptive mechanisms are called adaptive models (Kadlec *et al.*, 2011). Each adaptive model has strengths and weaknesses for each type of degradation of soft sensor models (Kaneko and Funatsu, 2013a).

As an MW approach, ensemble online support vector regression (EOSVR) has been proposed recently to adapt to nonlinear relationships between X-variables and an y-variable and to time-varying changes in process states (Kaneko and Funatsu, 2014). Through several case studies including the use of real industrial data set, it was confirmed that EOSVR had higher predictive ability than the other adaptive soft sensors did.

However, operation data include not only important variations caused by process changes but also noise such as measurement noise and sensor noise, which leads to decrease predictive ability of soft sensors in both model construction and prediction. In addition, noisy estimated y-values are inappropriate for process control. For instance, it is difficult to judge the end of transition when estimated y-values include noise. We have to handle noise appropriately for soft sensors to estimate important variations in y-values accurately.

Chemometric methods can handle noise in preprocessing data and in regression analysis. Principal component analysis (PCA) (Wold, 1987) and partial least squares (PLS) regression (Wold *et al.*, 2001) are statistical methods projecting data from an original space of m process variables to a new space of n uncorrelated variables (components) while reducing the dimension ($m > n$). By using only the first some principal components that include main variations in a data set, we can reduce effect of noise to further analysis. In support vector regression (SVR) (Bishop, 2006), robust models for noise can be constructed by using the ε-insensitive loss function and handling y-errors whose absolute values are lower than ε are set as zero. PCA, PLS and SVR can be

10.1016/j.ifacol.2016.07.364

combined with kernel functions and modified to nonlinear methods (Muller et al., 2001).

However, the above chemometric methods do not consider characteristics of operating data or time-series data (Lütlepohl, 2005). For example, temporally close data have strong relationships and correlations. Kaneko and Funatsu (2015a) proposed smoothing-combined soft sensors for noise reduction and improvement of predictive ability. Before both model construction and prediction in soft sensor analysis, a smoothing method such as a simple moving average, a linearly weighted moving average, an exponentially weighted moving average and Savitzky-Golay (SG) filtering (Savitzky and Golay, 1964) is used for time-series data of each X-variable, which is measured frequently, and noise can be reduced while considering relationships between temporally close data. Case studies using simulated and industrial datasets confirm that the use of the proposed methods enables soft sensors to predict y-values smoothly and accurately and SG filtering had higher performance or lower prediction errors than the other smoothing methods.

In this study, we developed a method combining EOSVR and SG filtering, which is called EOSVR-SG and implemented a MATLAB® tool performing the method in real plants in Mitsui Chemicals, Inc. After a y-value of a query is estimated by the EOSVR model, SG filtering is employed to an estimated y-variable. The smoothed y-value is the final estimation result.

As case studies, we show the prediction results of our soft sensor tool in real plants of Osaka Works, Mitsui Chemicals, Inc. The EOSVR-SG model produces smooth estimated y-values and achieves highly predictive ability.

2. METHOD

2.1 Ensemble Online Support Vector Regression (EOSVR)

The SVR method (Bishop, 2006) applies a support vector machine (SVM) to regression analyses, and can be used to construct nonlinear models by applying a kernel trick along with the SVM. The primal form of the SVR can be given as the following optimization problem.

Minimize

$$\frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{N}\left|y_i - f\left(\mathbf{x}_i\right)\right|_{\varepsilon}, \tag{1}$$

where y_i and $\mathbf{x}_i$ are training data, f is the SVR model, $\mathbf{w}$ is a vector of weights of X to y, ε is a threshold, N is the number of training data, and C is a penalty factor that controls the trade-off between model complexity and training errors. The second term of Eq. (1) is the ε-insensitive loss function, which is given as follows:

$$\left|y_i - f\left(\mathbf{x}_i\right)\right|_{\varepsilon} = \max\left(0, \left|y_i - f\left(\mathbf{x}_i\right)\right| - \varepsilon\right). \tag{2}$$

By minimizing Eq. (1), we can construct a regression model with a good balance between its generalization capabilities and its ability to adapt to the training data. A y-value estimated by inputting data $\mathbf{x}$ is represented as follows:

$$f\left(\mathbf{x}\right) = \sum_{i=1}^{N}\left(\alpha_i - \alpha_i^*\right)K\left(\mathbf{x}_i, \mathbf{x}\right) + b, \tag{3}$$

where b is a constant and K is a kernel function. The kernel function in our application is a radial basis function given as follows:

$$K\left(\mathbf{x}_i, \mathbf{x}\right) = \exp\left(-\gamma\left\|\mathbf{x}_i - \mathbf{x}\right\|^2\right), \tag{4}$$

where γ is a tuning parameter that controls the width of the kernel function.

Basic concept of EOSVR is shown in Fig. 1. Different SVR models mean different sets of the SVR hyperparameters C, ε, and γ values. First in EOSVR, to obtain multiple combinations of C, ε, and γ offline for various states in a plant, the window size ws is set, and then, the SVR hyperparameters are optimized by moving the window. The data in each window are then as follows:

$$\left(\mathbf{X}_1, \mathbf{y}_1\right), \left(\mathbf{X}_2, \mathbf{y}_2\right), \ldots, \left(\mathbf{X}_i, \mathbf{y}_i\right), \ldots, \left(\mathbf{X}_n, \mathbf{y}_n\right), \tag{5}$$

where $\mathbf{X}_i$ and $\mathbf{y}_i$ are the i^{th} data set of the X-variables and that of a y-variable, respectively. When the window is moved by h data points, the i^{th} dataset is from the $h(i-1)+1^{\text{th}}$ data point to the hi^{th} data point. For each dataset, the SVR hyperparameters are optimized based on cross-validation, which can be performed very quickly (Kaneko and Funatsu,

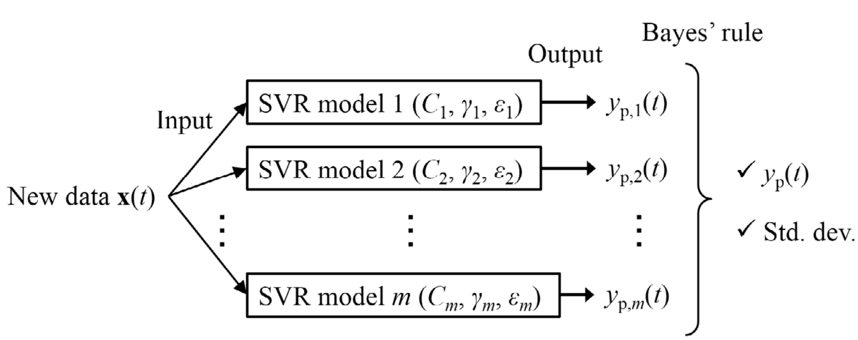

Fig. 1. Basic concept of EOSVR (Kaneko and Funatsu, 2014).