Melt Index Prediction by Weighted Least Squares Support Vector Machines

Jian Shi, Xinggao Liu

National Laboratory of Industrial Control Technology, Institute of Systems Engineering, Zhejiang University, Hangzhou 310027, People' Republic of China

Received 20 May 2005; accepted 22 September 2005
DOI 10.1002/app.23311
Published online in Wiley InterScience (www.interscience.wiley.com).

ABSTRACT: Melt index is considered an important quality variable determining product specifications. Reliable prediction of melt index (MI) is crucial in quality control of practical propylene polymerization processes. In this paper a least squares support vector machines (LS-SVM) soft-sensor model of propylene polymerization process is developed to infer the MI of polypropylene from other process variables. Considering the use of a SSE cost function without regularization might lead to less robust estimates; the weighted least squares support vector machines (weighted LS-SVM) approach of propylene polymerization process is further proposed to obtain a robust estimation of melt index. The performance of standard SVM model is taken as a basis of comparison. A detailed comparison research among the standard SVM, LS-SVM, and weighted LS-SVM models is carried out. The research results confirm the effectiveness of the presented methods. © 2006 Wiley Periodicals, Inc. J Appl Polym Sci 101: 285–289, 2006

Key words: polypropylene; computer modeling; weighted least squares support vector machines; melt

INTRODUCTION

The melt index is considered crucial in determining the product's grade and quality control of polypropylene produced in practical industrial processes. It is defined as the mass rate of extrusion flow through a specified capillary under prescribed conditions of temperature and pressure, which is costly and time-consuming when carried out in the laboratory.[1–2] Such situations can lead to a significant production of off-grades, especially during the on-line operations involved to change product specifications. An alternative is to develop on-line estimators of product quality based on available process information that would allow the supervision of the overall process and to avoid mismatch of product quality during product grade transitions.

The mechanistic modeling approaches for the prediction of the melt index are often challenged by the engineering activity and the relatively high complexity of the kinetic behavior and operation of the polymer plants,[3–5] which makes it difficult to obtain detailed predictions in real time. Instead, some production plants use machine learning methods to provide information for product and process design, monitoring, and control[6−8] on the basis of real-time database systems where a considerable amount of data about the studied process is available. Several works have been carried out to predict melt indices with various types of modeling methods. Rallo et al.[9] provided a fuzzy ARTMAP neural system and two hybrid networks to infer the melt index of six different LDPE grades produced in a tubular reactor. Han et al. [10] compared the performance of support vector machines, partial least squares, and artificial neural networks for MI estimation of San and PP processes, and concluded that the standard SVM yields the best prediction among the three toward the studied problems. Unfortunately, further research on SVM regarding this topic has not been carried out.

In this paper, a LS-SVM model of propylene polymerization process is first developed to infer the MI of polypropylene from other readily measurable process variables. Considering the use of a SSE cost function without regularization, as it is in the case with LS-SVM, might lead to estimates which are less robust, e.g., with respect to outliers on the data or when the underlying assumption of a Gaussian distribution for the error variables is not realistic[11]; the weighted LS-SVM approach of propylene polymerization process is further presented to obtain a robust estimation of melt index of polypropylene. Up to now, litter, however, has appeared in the literature on these matters. The

Correspondence to: X. Liu (liuxg@iipc.zju.edu.cn).
Contract grant sponsor: National Natural Science Foundation of China; contract grant number: 20106008.
Contract grant sponsor: National HI-TECH Industrialization Program of China; contract grant number: 2004−2080.
Contract grant sponsor: Zhejiang University; contract grant number: 111000−581645.

Journal of Applied Polymer Science, Vol. 101, 285–289 (2006)

aspects concerning the methods are described, highlighting the basic architectures and algorithms. The standard SVM model of propylene polymerization process proposed by Han[10] is also developed as a basis of comparison research. The performance of all models have been illustrated and evaluated with an actual propylene polymerization. The results obtained are then discussed and concluding remarks about the methods are finally presented.

THE SUPPORT VECTOR METHOD OF FUNCTION ESTIMATION

Standard support vector machines

Support vector machine introduced by Vapnik[12,13] is a valuable tool for solving pattern recognition and classification problems. SVMs can be applied to regression problems by the introduction of an alternative loss function.

Consider regression in the following set of functions

$$f(x) = w^T \varphi(x) + b \tag{1}$$

with given training data $\{x_i, y_i\}_{i=1}^M$ where M denotes the number of training data, x_i is the input data, and y_i is the output data. The nonlinear mapping φ maps the input data into a so-called high dimensional feature space, where a linear regression problem is obtained and solved. In the support vector method one aims at minimizing the regularized risk

$$R(w,b) = \gamma \frac{1}{M}\sum_{i=1}^{M} L_\varepsilon(y_i, f(x_i)) + \frac{1}{2} w^T w \tag{2}$$

where

$$L_\varepsilon(y_i, f(x_i)) = \begin{cases} 0, & |y - f(x)| \le \varepsilon \\ |y - f(x)| - \varepsilon & \text{otherwise} \end{cases} \tag{3}$$

In eq. (2), L_ε is the so-called ε-insensitive loss function, which indicates that it does not penalize errors below ε. The second term, $\frac{1}{2} w^T w$ is used as a flatness measurement of function[1] and γ is a regularized constant determining the tradeoff between the training error and the model flatness. The estimation problem is formulated then as the optimization problem

$$\min_{w,b,\xi^*,\xi} R(w,\xi^*,\xi) = \frac{1}{2} w^T w + \gamma \left\{ \sum_{i=1}^{M} \xi_i^* + \sum_{i=1}^{M} \xi_i \right\} \tag{4}$$

subject to the constraints

$$\begin{cases} y_i - w^T\varphi(x_i) - b \le \varepsilon + \xi_i^* & i = 1,\ldots,M \\ -y_i + w^T\varphi(x_i) + b \le \varepsilon + \xi_i & i = 1,\ldots,M \\ \xi_i^*, \xi_i \ge 0 & i = 1,\ldots,M \end{cases} \tag{5}$$

where ξ, ξ^* are slack variables. One obtains $w = \sum_{i=1}^{M} (\alpha_i^* - \alpha_i)\varphi(x_i)$ where α^*, α are obtained by solving a quadratic program and are the Lagrange multipliers related to the first and second set of constraints. The data points corresponding to nonzero values for α^*, α are called support vectors. Typically, many of these values are equal to zero. Finally, one obtains the following model in the dual space

$$f(x) = \sum_{i=1}^{M} (\alpha_i^* - \alpha_i) K(x, x_i) + b \tag{6}$$

where the kernel function K corresponds to

$$K(x_i, x) = \varphi(x_i)^T \varphi(x) \tag{7}$$

according to Mercer's condition.[12] and RBF kernels will be employed in this study.

An extension studied in the context of the ε-insensitive loss function is

$$R(w,\xi^*,\xi) = \frac{1}{2} w^T w + \gamma \left\{ \sum_{i=1}^{M} (\xi_i^*)^p + \sum_{i=1}^{M} (\xi_i)^p \right\} \tag{8}$$

where p = 1 corresponds to eq. (4).

Least squares support vector machines

For the sequel, a least squares version[14] of the support vector method is employed for function estimation problems. It corresponds to p = 2 and the following form of ridge regression

$$\min_{w,b,\xi} R(w,\xi) = \frac{1}{2} w^T w + \frac{1}{2}\gamma \sum_{i=1}^{M} \xi_i^2 \tag{9}$$

subject to the equality constraints

$$y_i = w^T \varphi(x_i) + b + \xi_i \quad i = 1,\ldots,M \tag{10}$$

One defines the Lagrangian

$$L(w,b,\xi,\alpha) = R(w,\xi) - \sum_{i=1}^{M} \alpha_i (w^T\varphi(x_i) + b + \xi_i - y_i) \tag{11}$$

with α_i Lagrange multipliers. the conditions for optimality

$$\begin{cases} \dfrac{\partial L}{\partial w} = 0 \rightarrow w = \sum_{i=1}^{M} \alpha_i \varphi(x_i) & \\ \dfrac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^{M} \alpha_i = 0 & \\ \dfrac{\partial L}{\partial \xi_i} = 0 \rightarrow \alpha_i = \gamma \xi_i & i = 1,\ldots,M \\ \dfrac{\partial L}{\partial \alpha_i} = 0 \rightarrow w^T\varphi(x_i) + b + \xi_i - y_i = 0 & i = 1,\ldots,M \end{cases} \tag{12}$$

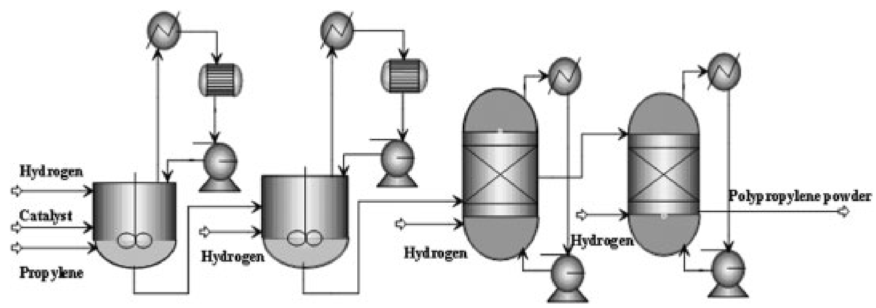

Figure 1 Schematic of propylene polymerization process.

can be written as the solution to the following set of linear equations after elimination of w and ξ_i

$$\begin{bmatrix} 0 & 1_v^T \\ 1_v & K + \gamma^{-1} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{13}$$

with $y = [y_1, \ldots, y_M]^T$, $1_v = [1, \ldots 1]^T$, $\alpha = [\alpha_1, \ldots, \alpha_M]^T$ and I is an identity matrix.

The support values α_i are proportional to the errors at the data points eq. (12), while in the case of eq. (4) most values are equal to zero.

The resulting LS-SVM model for function estimation becomes

$$f(x) = \sum_{i=1}^{M} \alpha_i K(x, x_i) + b \tag{14}$$

where α, b are the solution to eq. (13)

Weighted least squares support vector machines

To obtain a robust estimate[11] based upon the previous LS-SVM solution, in a subsequent step, one can weight the error variables $\xi_i = \alpha_i / \gamma$ by weighting factors v_i. This leads to the optimization problem:

$$\min_{w^*, b^*, \xi^*} R(w^*, \xi^*) = \frac{1}{2} w^{*T} w^* + \frac{1}{2} \gamma \sum_{i=1}^{M} v_i \xi_i^{*2} \tag{15}$$

subject to the equality constraints

$$y_i = w^{*T} \varphi(x_i) + b^* + \xi_i^* \quad i = 1, \ldots, M \tag{16}$$

The Lagrangian becomes

$$L(w^*, b^*, \xi^*, \alpha^*) = R(w^*, \xi^*) - \sum_{i=1}^{M} \alpha_i^* \{ w^{*T} \varphi(x_i) + b^* + \xi_i^* - y_i \} \tag{17}$$

The unknown variables for this weighted LS-SVM problem are denoted by the * symbol. From the conditions for optimality and elimination of w*, ξ^* one obtains the KKT system

$$\begin{bmatrix} 0 & 1_v^T \\ 1_v & K + V_\gamma \end{bmatrix} \begin{bmatrix} b^* \\ \alpha^* \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{18}$$

where the diagonal matrix V is given by

$$V_\gamma = \mathrm{diag}\left\{ \frac{1}{\gamma v_1}, \ldots, \frac{1}{\gamma v_M} \right\} \tag{19}$$

The choice of the weights v_i is determined based upon the error variables $\xi_i = \alpha_i / \gamma$ from the LS-SVM case.[13] Robust estimates are obtained then by taking

$$v_i = \begin{cases} 1 & \text{if } |\xi_i / \hat{S}| \le c_1 \\ \dfrac{c_2 - |\xi_i / \hat{S}|}{c_2 - c_1} & \text{if } c_1 \le |\xi_i / \hat{S}| \le c_2 \\ 10^{-4} & \text{otherwise} \end{cases} \tag{20}$$

where ŝ; is a robust estimate of the standard deviation of the LS-SVM error variables ξ_i:

$$\hat{S} = \frac{\mathrm{IQR}}{2 \times 0.6745} \tag{21}$$

The interquartile range, IQR, is the difference between the 75th percentile and 25th percentile. The constants c_1, c_2 are typically chosen as $c_1 = 2.5$ and $c_2 = 3$.[11] Eventually, the procedure (15)–(20) can be repeated iteratively.

RESULTS AND DISCUSSION

Figure 1 shows the schematic diagram of a propylene polymerization process, which consists of four reactors in series. The polymerization reaction takes place in a liquid phase in the first two reactors and is completed in a vapor phase in the third and fourth reactors to produce the powdered polymer product. The modeling data used for training and validating the soft-sensor have been acquired from the historical logs

TABLE I
Performance for the Testing Dataset

Modeling method	MAE	MRE (%)	RMSE	STD	TIC
Weighted LS-SVM	0.0754	3.27	0.0198	0.1055	0.0223
LS-SVM	0.0842	3.66	0.0214	0.1116	0.0240
SVM	0.1105	4.80	0.0274	0.1394	0.0307

recorded in a real propylene polymerization plant. A total of 9 process variables (t, p, l, a: process temperature, pressure, level of liquid, percentage of hydrogen in vapor phase; f_1, f_2, f_3: flow rate of 3 streams of propylene; f_4, f_5: flow rate of catalyst and aid-catalyst respectively) are chosen as the input variables according to the reaction mechanism. The average residence time of the products is taken into account in the process of datasets construction. Data are filtered to discard abnormal situations and to improve the quality of the predictive system. The input and output variables are normalized with respect to their maximum operation values. Data from the records of the process variables and MI are separated into training, test and generalization sets that are constructed from the time series of recorded plant data. And the test set is obtained from the same batch as the training set, while the generalization set is derived from another batch.

The detailed comparison of test performance between LS-SVM and standard SVM is listed in Table I and the comparison between LS-SVM and weighted LS-SVM is also presented. The difference between the output of the models and the desired output is referred to as the error and can be measured in different ways. Here, mean absolute error (MAE), mean relative error (MRE), root mean squared error (RMSE), standard deviation of absolute errors (STD), and Theil's Inequality Coefficient (TIC), are adopted as derivation measurements between measured and predicted values. They are defined as the following, respectively:

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{22}$$

$$\mathrm{MRE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{23}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{24}$$

$$\mathrm{STD} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{N}(e_i - \bar{e})^2}\ i = 1,\ldots,N \tag{25}$$

$$\mathrm{TIC} = \frac{\sqrt{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}}{\sqrt{\sum_{i=1}^{N}y_i^2} + \sqrt{\sum_{i=1}^{N}\hat{y}_i^2}} \tag{26}$$

where, $e_i = y_i - \hat{y}_i$, $\bar{e} = \frac{1}{N}\sum_{i=1}^{N} e_i$, and y_i, $\hat{y}_i$ denote the measured value and predicted result, respectively.

The data listed in Table I indicates that the weighted LS-SVM model functions best on the overall, with mean absolute error of 0.0754, compared with those of 0.0842 and 0.1105 obtained from the corresponding LS-SVM and SVM models, respectively. The RMSE listed also in Table I have confirmed the prediction accuracy of the proposed methods. Weighted LS-SVM yields the smallest STD among the three, which indicates the predictive stability of the method. TIC of weighted LS-SVM is quite acceptable when compared with those of standard SVM and LS-SVM, which indicates a good level of agreement between the proposed model and the studied process.[15] A visual impression of the agreement between the measured MI and the models output can be obtained from Figure 2, where the weighted LS-SVM model yields consistently good predictions.

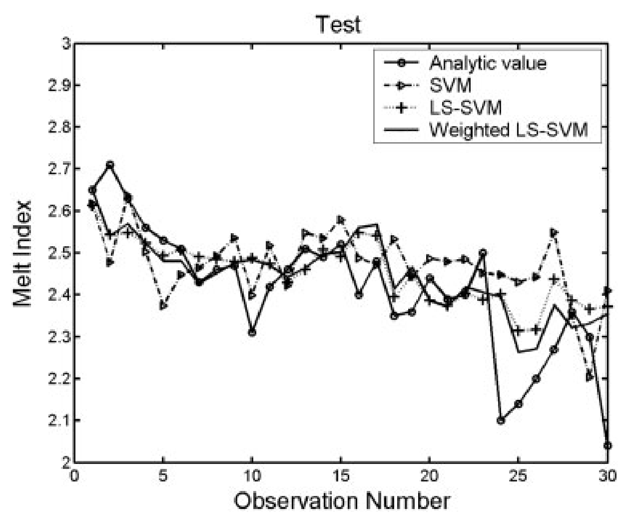

Figure 2 Estimation via three approaches for test dataset.

TABLE II
Performance for the Generalization Dataset

Modeling method	MAE	MRE (%)	RMSE	STD	TIC
Weighted LS-SVM	0.0635	2.49	0.0312	0.0695	0.0138
LS-SVM	0.0662	2.60	0.0313	0.0751	0.0138
SVM	0.0818	3.21	0.0410	0.0962	0.0181

To illustrate the universality of the proposed model, a detailed comparison of the generalization data set is presented in Table II. It is noted that the performance is consistent with the above test results, with a slightly increase in predictive precision. The mean absolute error of weighted LS-SVM is 0.0635, compared with 0.0818 of SVM, showing an error decrease of approximately 20%. Similar behaviors are observed in terms of MXAE, MXRE, RMSE, STD, and TIC.

CONCLUSIONS

This paper has presented methods for using LS-SVM and weighted LS-SVM to infer MI of polypropylene from other process variables. Comparing with standard SVM, LS-SVM involves equality instead of inequality constraints, which greatly simplifies the problem in such a way that the solution is characterized by a linear system instead of a quadratic programming problem. More robust estimates for regression can be obtained by a weighted version of LS-SVM, which is done by first applying an un-weighted LS-SVM and, in the second stage, associate weighting values to the error variables based upon the resulting error variables from the first stage.

The weighted LS-SVM model predict MI with mean relative error of approximately 3.27% when appropriately trained, compared with those of 3.66% and 4.80% obtained from the corresponding LS-SVM and SVM models, respectively. The results indicate that the proposed method provides prediction reliability and accuracy and supposed to have promising potential for practical use.

References

1. Bafna, S. S.; Beall, A.-M. J Appl Polym Sci 1997, 65, 277.
2. Yi, H.-S.; Kim, J. H.; Han, C.; Lee, J.; Na, S.-S. Ind Eng Chem Res 2003, 42, 91.
3. McAuley, K. B.; MacGregor, J. F. AICHE J 1991, 6, 825.
4. Sarkar, P.; Gupta, S. K. Polym Eng Sci 1993, 6, 368.
5. McKenna, T. F.; Soares, J. B. P. Chem Eng Sci 2001, 56, 3931.
6. Hunt, K. J.; Sbarbaro, D.; Zbikowski, R.; Gawthrop, P. J. Automatica 1992, 28, 1083.
7. Barto, A. G.; Sutton, R. S.; Anderson, C. W. IEEE Trans Syst Man Cybern 1983, 13, 834.
8. Xiong, Z. H.; Zhang, J. Neurocomputing 2004, 61, 317.
9. Rallo, R.; Ferre-Giné, J.; Arenas, A.; Giralt, F. Comput Chem Eng 2002, 26, 1735.
10. Han, I.-S.; Han, C.; Chung, C.-B. J Appl Polym Sci 2005, 95, 967.
11. Suykens, J. A. K.; Vandewalle, J. Neurocomputing 2002, 48, 85.
12. Vapnik, V. The Nature of Statistical Learning Theory; Springer: New York, 1995.
13. Vapnik, V. Statistical Learning Theory; John Wiley: New York, 1998.
14. Suykens, J. A. K.; Vandewalle, J. Neural Process Lett 1999, 9, 293.
15. Murray-Smith, D. J. Math Comput Model Dyn Syst 1998, 4, 5.