Korean J. Chem. Eng., **27**(6), 1662-1668 (2010)
DOI: 10.1007/s11814-010-0280-x

Prediction of the melt flow index using partial least squares and support vector regression in high-density polyethylene (HDPE) process

Tae Chang Park, Tae Young Kim, and Yeong Koo Yeo†

Department of Chemical Engineering, Hanyang University, Seoul 133-791, Korea
(*Received 4 January 2010 • accepted 2 March 2010*)

Abstract−In polyolefin processes the melt flow index (MFI) is the most important control variable indicating product quality. Because of the difficulty in the on-line measurement of MFI, a large number of MFI estimation and correlation methods have been proposed. In this work, mechanical predicting methods such as partial least squares (PLS) method and support vector regression (SVR) method are employed in contrast to conventional dynamic prediction schemes. Results of predictions are compared with other prediction results obtained from various dynamic prediction schemes to evaluate predicting performance. Hourly MFIs are predicted and compared with operation data for the high density polyethylene process involving frequent grade changes. We can see that PLS and SVR exhibit excellent predicting performance even for severe operating situations accompanying frequent grade changes.

Key words: Melt Flow Index, Partial Least Square, Support Vector Regression, Prediction

INTRODUCTION

Demand for petrochemical products with high quality and low price is greatly increasing, especially in modern high density polyethylene (HDPE) markets. Consumers' requirements have become increasingly sophisticated and strict for new HDPE products. In HDPE production operations the melt flow index (MFI) is the most important control variable indicating product quality. To meet the various and stringent requirements for HDPE products such as MFI, it is imperative to maintain uniformity of HDPE properties during grade change operations. However, long settling time for each grade and large overshoots as well as large amount of off-specification products prohibit achievement of efficient grade transitions in actual plant operations. It is very difficult to measure MFI on-line during plant operations. Thus, it is a usual practice to estimate MFI values in terms of measurable key variables such as temperature, pressure, feed rates of each reactants, etc.

So far a large number of MFI estimation and correlation methods have been proposed. The estimation method presented by McAuley and MacGregor [1,2] is based on the logarithm of the linear combination of flow rates of each component. To take into account unmeasured impurities or disturbances, they employed a constant term that is updated iteratively. In the estimation model proposed by Ohshima [3,4] an additional term representing temperature is included. Oh [5] developed an off-line Excel-based program for estimating polymer properties. They proposed an empirical prediction model for estimation of instantaneous MFI. Recently, construction of dynamic model predictive control system was proposed based on the selection of the most significant operating variable [6].

In this study mechanical predicting methods are employed to predict MFI in contrast to the application of conventional dynamic predicting schemes. Mechanical predicting methods have found wide application areas. Artificial neural network (ANN) was used to forecast the melting point, which is one of the fundamental physicochemical properties of molecules [7]. ANN was also used to predict polymer properties in batch polymerization reactors [8] as well as to estimate the glass transition temperature of polymer materials [9]. Partial least squares (PLS) was used to predict solubility parameters of several materials [10]. For the estimation of MFI, results obtained by using ANN were reported [11], and the application of weighted least squares support vector machines (weighted LS-SVM) was analyzed [12].

In this work, partial least squares (PLS) and support vector regression (SVR) methods are employed to predict MFIs in the high density polyethylene process. Results of predictions are compared with operation data as well as other conventional dynamic predicting schemes to validate effectiveness of the proposed predicting methods.

1. Theoretical Background of PLS and SVR

1-1. Partial Least Squares (PLS)

The partial least squares (PLS) method is a multiple linear regression model which can be used to model relations between inputs and outputs with correlations and/or restricted number of data points [13]. For this capability PLS has been widely used as a powerful modeling technique for constructing so-called black-box models from experimental or operational data. In the PLS method the high dimensional spaces of the input and output data obtained from a plant are projected onto the low dimensional feature spaces followed by identification of the best relationship between the feature vectors. More specifically, the high dimensional data matrices X and Y are projected onto several key factors (T, U) and linear regression is performed for the relation between these factors. The outer relations for the output block Y and input block X are given by:

$$X = TP^T + E = \sum t_h p_h^T + E \quad (1)$$

$$Y = UQ^T + F = \sum u_h q_h^T + F \quad (2)$$

The coefficient b_h combining each block can be found from the following relation:

†To whom correspondence should be addressed.
E-mail: ykyeo@hanyang.ac.kr

$$\hat{u}_h = b_h t_h \tag{3}$$

Usually, the least squares method is used to compute the regression coefficient $b_h(=u_h^T t_h / t_h^T t_h)$. In computation of key factors the nonlinear iterative partial least squares scheme is widely used as a computing algorithm. The nonlinear partial least squares method can be classified into two types: in type I, the inner relations among blocks are nonlinear and in type II, both the inner and outer relations are nonlinear. The general form of the type I is given by:

$$\hat{u}_h = f(t_h) + R \tag{4}$$

Various models according to the function f have been proposed. For example, f can take a quadratic form or artificial neural network can be used.

1-2. Support Vector Regression (SVR)

The support vector machine (SVM) has been known as very powerful modeling algorithm used to solve classification problems. Because the formulation of SVMs is based on structural risk minimization rather than empirical risk minimization, which is employed by other conventional black-box modeling algorithms such as ANN and PLS, the SVMs typically show better performance than the conventional algorithms. In the SVM, a hyperplane dividing the multidimensional learning data into two groups is identified first. The hyperplane is used as a determining function to predict the group to which the unknown data should belong [14]. For given SVM learning data, the hyperplane is found so that two classes are away from each other as long as possible according to the label y_i (see Fig. 1). In this case the hyperplane $f(x)=w^T x+b$ becomes the determining function. If the value of f(x) is greater than 0 for x with unknown label, it is classified as +1 while it is classified as −1 class when f(x) is less than 0. In this case, the value of $2/\|w\|$ is called a margin which represents the minimum distance between two classes separated by the hyperplane. Basically, an SVM uses the hyperplane that maximizes the margin as the determining function.

The problem of estimation of w and b in the determining function $f(x)=w^T x+b$ of a SVM can be formulated as a 2^{nd}-order programming problem (SP_0) given by:

$$\begin{aligned} &\min w^T w \\ (SP_0):\ &\text{s.t. } y(w^T x^i + b) \geq 1, \text{ for } i=1, \ldots, l \end{aligned} \tag{5}$$

The learning data with different labels may not be separated clearly by the hyperplane. To take into account these overlapping cases, the relaxed model (SP) may be used instead of the original model (SP_0).

$$\begin{aligned} &\min w^T w + c\sum_i^l \varepsilon_i \\ (SP):\ &\text{s.t. } y_i(w^T x^i + b) \geq 1 - \varepsilon_i,\ \forall i \\ &\varepsilon_i \geq 0,\ \forall i \end{aligned} \tag{6}$$

$$\begin{aligned} &\min e^T \alpha + \frac{1}{2}\alpha^T Q \alpha \\ (SD):\ &\text{s.t. } y^T \alpha = 0 \\ &0 \leq \alpha_i \leq C,\ \forall i \end{aligned} \tag{7}$$

The second term $\Sigma\varepsilon_i$ represents the amount of the error which represents the degree of mismatch between the learning data and the determining function obtained from experienced errors. C is the relative weight between the margin and the experienced error. In short, the objective in (SP) identifies the determining function f(x) which maximizes the margin and minimizes the experienced error.

$$\begin{aligned} u &= \sum_i c_i^* y_i x^1 \\ f(x) &= w^T x + b = \sum_i \alpha_i^* y_i (x^i)^T x + b \end{aligned} \tag{8}$$

Transformation to a specific pace is performed by a kernel function. The most widely used kernel functions include RBF (radial basis function) and polynomial kernel. The RBF kernel is given by:

$$k(x^i, x^j) = \exp\left(-\frac{\|x^i - x^j\|^2}{\sigma^2}\right) \tag{9}$$

And the polynomial kernel is given by:

$$k(x^i, x^j) = ((x^i)^T x^j)^p \tag{10}$$

SVM is used to classify the learning data into '+1' class and '−1' class (see Fig. 1). The support vector regression (SVR) can be considered as generalization of SVM so that arbitrary real values might be estimated. In the SVR ε-insensitive loss function (Eq. (11)) is used instead of $\Sigma\varepsilon_i$.

$$L_\varepsilon(x, y, f) = \max(0, |y - f(x)| - \varepsilon) = \max(0, |y - (w^T x + b)| - \varepsilon) \tag{11}$$

The value of the ε-insensitive loss function depends on the size of the estimation error defined by the difference between the actual value y and the estimated value $f(x)=w^T x+b$. If the estimation error is less than ε, we have $L_\varepsilon(x, y, f)=0$. Otherwise, the loss function

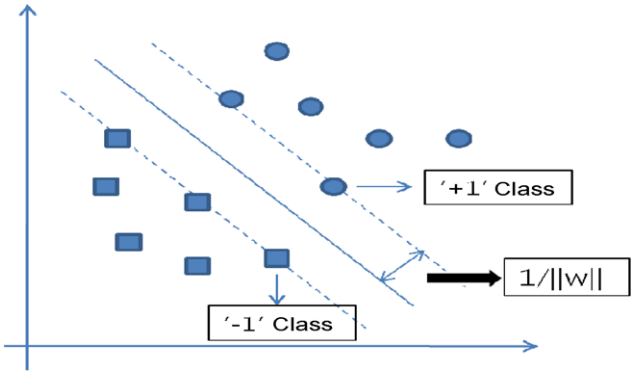

Fig. 1. A decision function of SVM (support vector machine).

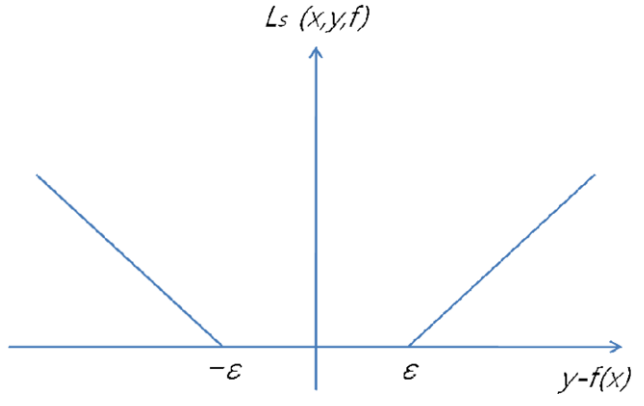

Fig. 2. ε-Insensitive loss function.

takes the value of the difference between the absolute error and ε as given by $L_\varepsilon(x, y, f)=|y-f(x)|-\varepsilon$ (see Fig. 2). In SVR, the margin is maximized while maintaining y and $f(x)=w^Tx+b$ values within ε, which can be described as an optimization model given by:

$$\begin{aligned} &\min w^Tw+c\sum_i^l(\varepsilon_i^+\varepsilon_i^-)\\ \text{(RP): s.t. } &y_i(w^Tx^i+b)-y_i\geq\varepsilon+\varepsilon_i^+,\ \forall i\\ &y_i-(w^Tx^i+b)\geq\varepsilon+\varepsilon_i^-,\ \forall i\\ &\varepsilon_i^+\varepsilon_i^-\geq 0,\ \forall i \end{aligned} \tag{12}$$

$$\begin{aligned} &\max\sum_i y_i(\alpha_i^- -\alpha_i^+)-\varepsilon\sum_i(\alpha_i^- +\alpha_i^+)-\frac{1}{2}\sum_{i,j}(\alpha_i^- -\alpha_i^+)(\alpha_i^- -\alpha_i^+)Q_{i,j}\\ \text{(RD): s.t. } &\sum_i(\alpha_i^- -\alpha_i^+)=0\\ &0\leq\alpha_i^+,\ \alpha_i^-\leq C,\ \forall i \end{aligned} \tag{13}$$

2. HDPE Process

The HDPE plant at LG petrochemicals (LGPC) consists of two processes where the polymerization reaction occurs. Fig. 3 shows a typical slurry polymerization process for the production of HDPE considered in the present study. Two polymerization reactors are involved in each process. The polymerization reaction is highly exothermic with reaction heat of about 1,000 kcal per 1 kg ethylene. Thus, it is required to provide internal cooling coils and appropriate external cooling systems which remove about 80% of polymerization heat. Ethylene, comonomer (1-butene or higher alpha-olefin), hydrogen, catalyst, activator and hexane as well as continuously recycled mother liquid are fed into reactors as reactants. About 90-95% of reactor volume is occupied by reaction slurry. As the reaction pressure builds up, the polyethylene slurry is transported to subsequent process equipment and the level in the reactor is maintained within a permissible range. Cake containing diluents is obtained from the separation of the reaction slurry in the centrifugal separator followed by removal of diluents by hot nitrogen gas in the dryer. Then appropriate additives are added according to final uses. Pelletization is performed in water. Dried pellets are transferred to an homogenizer and cooled.

3. MFI Prediction Model

McAuley and MacGregor [1,2] proposed an inferential system for the on-line prediction of MFI of low density polyethylene (LDPE) produced in the UNIPOL fluidized-bed reactor. They employed two models: (1) an instantaneous polymer property model that describes the relationship between process variables and the current polymer property; and (2) a cumulative property model that describes the relationship between the instantaneous and cumulative polymer properties in the reactor. The instantaneous MFI model can be represented as

$$\ln(MFI_i)=3.5\ln\left\{k_0+k_1\frac{[H_2]}{[C_2]}+k_2\frac{[C_3]}{[C_2]}+k_3\frac{[C_4]}{[C_2]}+k_4\frac{[R]}{[C_2]}\right\}+k_5\left(\frac{1}{T}-\frac{1}{T_0}\right) \tag{14}$$

where the subscript i denotes the i-th reactor, MFI is the instantaneous MFI, k_j (j=0, 1, …, 5) denotes parameters, T_0 is the reference temperature, T is the temperature of the reactor, and $[C_2]$, $[C_2]$ and $[C_2]$ are the concentrations of ethylene, 1-butene and high alpha-olefin, respectively. And $[H_2]$ and [R] are the concentrations of hydrogen and co-catalyst.

The corresponding cumulative MFI model is given by

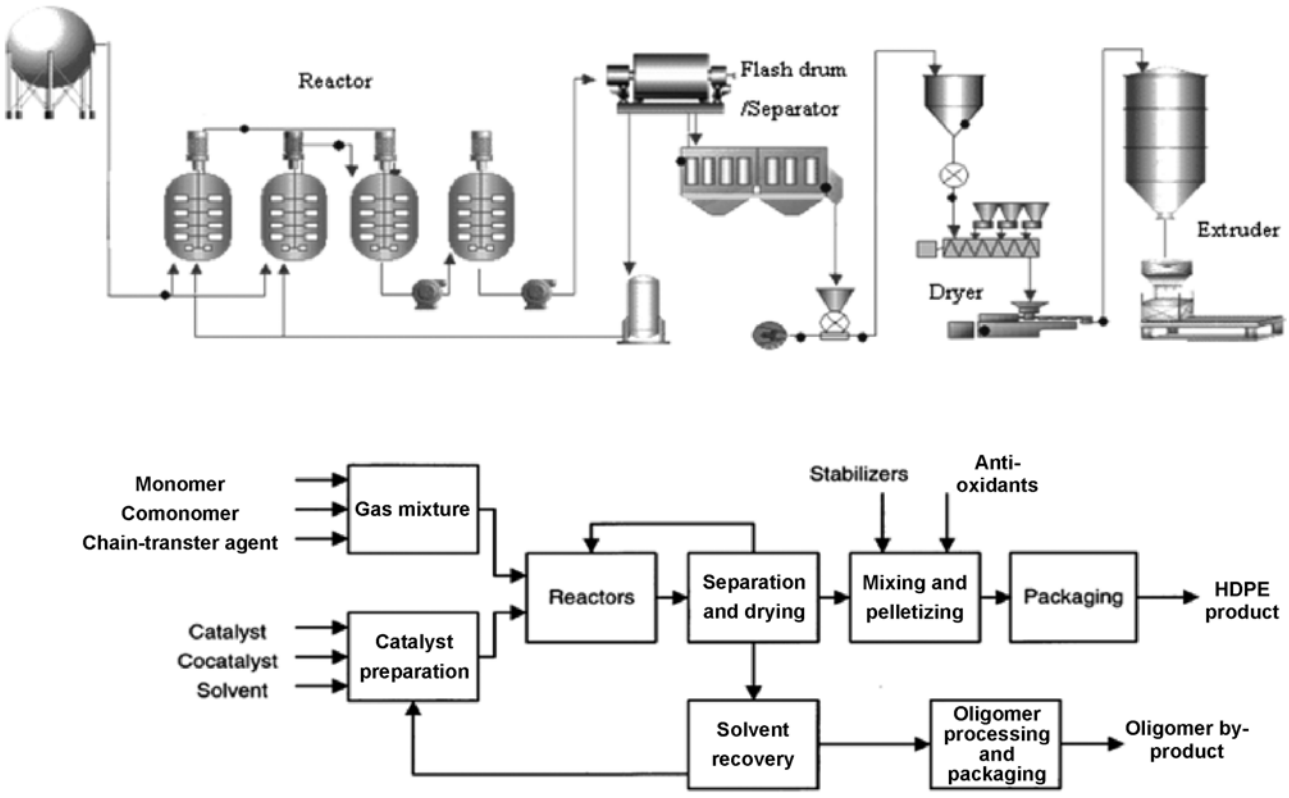

Fig. 3. A schematic of slurry polymerization process for the production of HDPE.

Table 1. Scaled parameters used in the MFI estimation models'

Parameter	Scaled value [-]	Equation
k_1	0.0726	(14)
k_2	0.3298	(14)
α_1	2.2	(16), (21)
α_2	0.6	(16), (21)
α_3	1.1	(16), (21)

$$\frac{d(MFI_c(t)^{-0.286})}{dt} = \frac{1}{\tau(t)}MFI_i(t)^{-0.286} - \frac{1}{\tau(t)}MFI_c(t)^{-0.286} \quad (15)$$

where MFI_c and MFI_i denote MFIs of cumulative polymer and instantaneous polymer, respectively.

Ohshima et al. [3] suggested an instantaneous MFI estimation model based on the assumption that reaction time is much shorter than the residence time in the reactor and the typical polymer structure is similar with each other [4]. Their model is a linear combination of concentration ratios of each reactants and logarithm of catalyst and temperature terms and is represented as

$$\log(MFI_i) = \beta + \alpha_1\frac{[H_2]}{[C_2]} + \alpha_2\frac{[C_3]}{[C_2]} + \alpha_3\frac{[C_4]}{[C_2]} + \alpha_4\log[R] + \alpha_5\log(T) \quad (16)$$

where the subscript i denotes the i-th reactor, MFI is the instantaneous MFI, β and α_j (j=0, 1, …, 5) denotes parameters.

Again the corresponding cumulative MFI model can be given by

$$\frac{d\log(MFI_c(t))}{dt} = \frac{1}{\tau_1(t)}\log(MFI_i(t)) - \frac{1}{\tau_1(t)}\log(MFI_c(t)) \quad (17)$$

where τ_1 represents the residence time of polymer in the reactor. Table 1 shows key scaled parameters used in Eq. (14), (17) and (21).

Cumulative MFI models for each process unit except the extruder are given by

$$\frac{d\log(MFI_q(t))}{dt} = \frac{1}{\tau_q(t)}\log(MFI_{q-1}(t)) - \frac{1}{\tau_q(t)}\log(MFI_q(t)) (q=1, 2 \text{ and } 3) \quad (18)$$

By assuming time delay of T_L, the extruder can be represented as

$$\log(MFI_4(t)) = \log(MFI_3(t - T_L)) \quad (19)$$

All of the four units (reactor, flash drum/separator, dryer and extruder) described above are shown in Fig. 4, where each unit is denoted by 1, 2, 3 and 4. In Fig. 4, MFI_q is the MFI of polymer produced at unit q, MFI_0 is the instantaneous MFI produced in the reactor and MFI_p is the MFI of polymer powder.

The MFI prediction model proposed by Sato [15] consists of two inputs (flow rates of hydrogen and butane) and two outputs (cumulative MFI and density):

$$MFI_i = \exp\left\{b_T\left(\frac{1}{T} - \frac{1}{T_0}\right)\right\}\cdot\left\{b_0 + b_1\frac{[C_3]}{[C_2]} + b_3\frac{[H_2]}{[C_2]}\right\}^{3.5} \quad (20)$$

where $b_{0, 1, \ldots, 7}$ are model parameters.

Oh [5] presented an empirical instantaneous MFI estimation model that predicts instantaneous properties of polymer. The ratio of the amount of feeds instead of concentrations of each component is used in this model, which is given by

$$\log(MFI_{inst,i}) = \alpha + \alpha_0\log(T)_i + \alpha_1\log\left(\frac{[H_2]}{[C_2]}\right)_i + \alpha_2 FC_3C_{2,i} + \alpha_3 FC_4C_{2,i} \quad (21)$$

where FC_3C_2 represents the ratio of the feed rate of propylene to that of ethylene and FC_4C_2 denotes the ratio of the feed rate of 1-butene to that of ethylene. Again, α, $a_{0, 1, \ldots, 3}$ are model parameters.

RESULTS AND DISCUSSION

In the present study one of the parallel reactors of the HDPE plant (LG petrochemical Co.) was selected as the target process. The range of the predicting interval was from 12 A.M. on December 7, 2008, to 3 P.M. on December 9, 2008. The ratio of hydrogen rate to ethylene rate (Q), the rate of pure hexane fed into the reactor (HDH), the amount of recycled hexane (HMH), and the reaction temperature (RxT) were found to be affecting the MFI most significantly and were chosen as input variables. Fig. 5 shows general trend of input variables as well as that of MFIs.

The configuration of input data for training in PLS is the same as that in SVR. First, the input data were divided into three cases (CASE 1, CASE 2 and CASE 3) before prediction and the predicted MFIs were compared with measured MFIs during operation. Table 2 shows the configuration of input data for each case. MFI(t−1) and MFI(t−2) represents past MFIs at 1 and 2 hours before, respectively.

The prediction performances of the three cases of PLS, SVR and

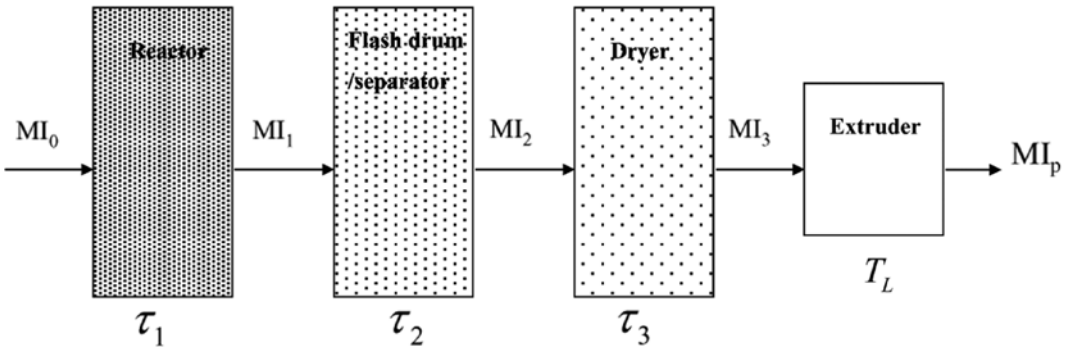

Fig. 4. Four sequential process units with time delays.

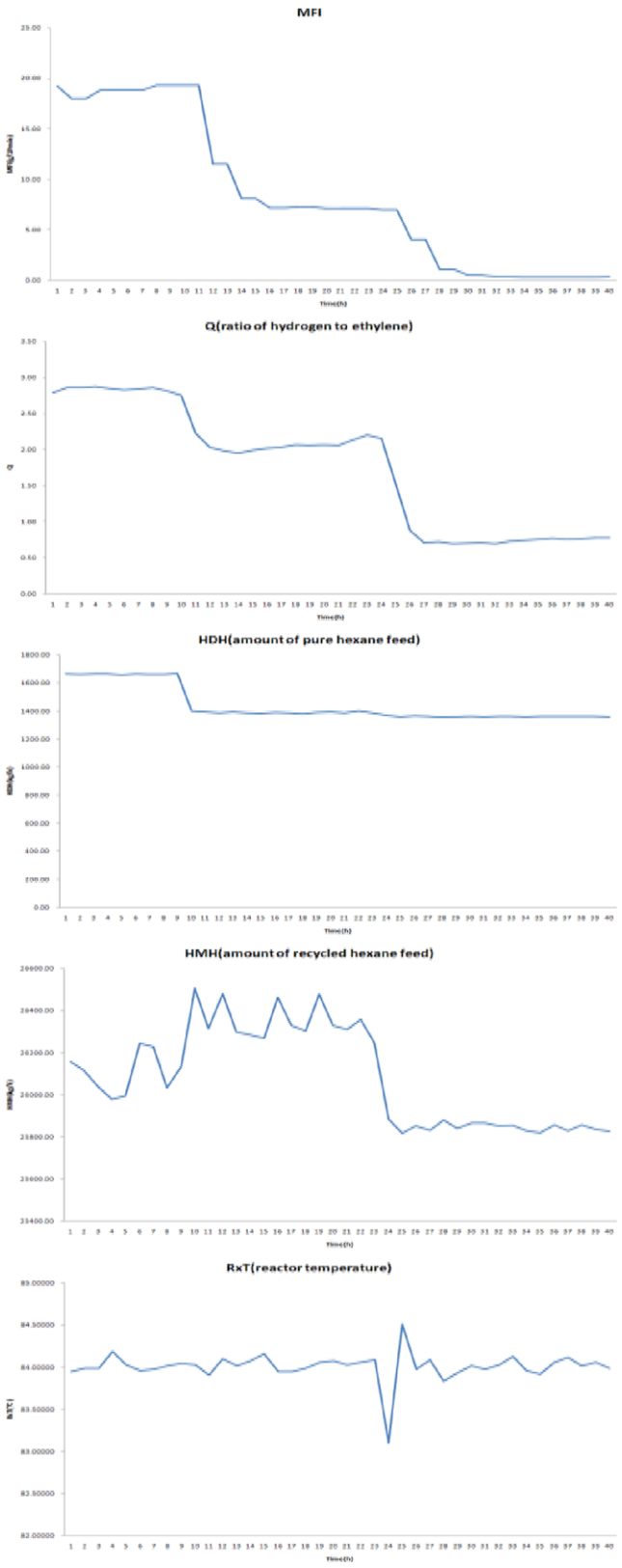

Fig. 5. Trends of 4 input variables and MFI data.

four different MFI prediction models can also be assessed from Table 3, which lists the root-mean-squared errors(RMSEs) between the measured and predicted values of the melt indices computed by:

Table 2. The Configuration of input and output data for training: (a) CASE 1, (b) CASE 2 and (c) CASE 3

(a)		(b)		(c)	
Input data	Output data	Input data	Output data	Input data	Output data
ETH	MFI	ETH	MFI	ETH	MFI
HDH		HDH		HDH	
HMH		HMH		HMH	
Q		Q		Q	
RxT		RxT		RxT	
		MFI(t−1)		MFI(t−1)	
				MFI(t−2)	

Table 3. RSMEs for the Models for Melt Indices

Modeling method		RSMEs for the testing data sets
CASE.1	PLS	2.72
	SVR	3.08
CASE.2	PLS	2.92
	SVR	2.88
CASE.3	PLS	2.87
	SVR	2.85
Model 1 (McAuley)		3.04
Model 2 (Ohshima)		2.83
Model 3 (Sato)		2.98
Model 4 (Oh)		2.99

$$\mathrm{RMSE} = \sqrt{\sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)^2}{2}} \tag{22}$$

Fig. 6 shows results of prediction by PLS and SVR as well as measured MFIs for each case. As can be seen, inclusion of past data in the prediction gives better predicting performance. But results of prediction for CASE 2 (includes past data at 1 hour before) do not show significant difference from those for CASE 2 (includes past data at 1 and 2 hours before). This means that inclusion of long range past data in the prediction does not necessarily improve predicting performance. From Fig. 6, we can see that prediction by SVR exhibits better tracking performance than that by PLS.

For the interval of rapid changing MFI, results of prediction by PLS and SVR were compared with those obtained from existing prediction models. Fig. 7 shows the performance of prediction using the model by McAuley (Model 1) as well as that using proposed PLS and SVR methods. We can see that PLS and SVR exhibit better predicting performance than Model 1. For the prediction from 14^{th} hour to 25^{th} hour, Model 1 shows significant predicting errors compared to PLS and SVR. Fig. 8 shows the performance of prediction using the model by Ohshima (Model 2) as well as that using proposed PLS and SVR methods. We can see that PLS and SVR show better tracking performance than Model 2. Model 2 exhibits significant predicting error for rapid changes in MFIs. Fig. 9 shows the performance of prediction using the model by Sato (Model 3) as well as that using proposed PLS and SVR methods. We can see that PLS and SVR exhibit better predicting performance than Model

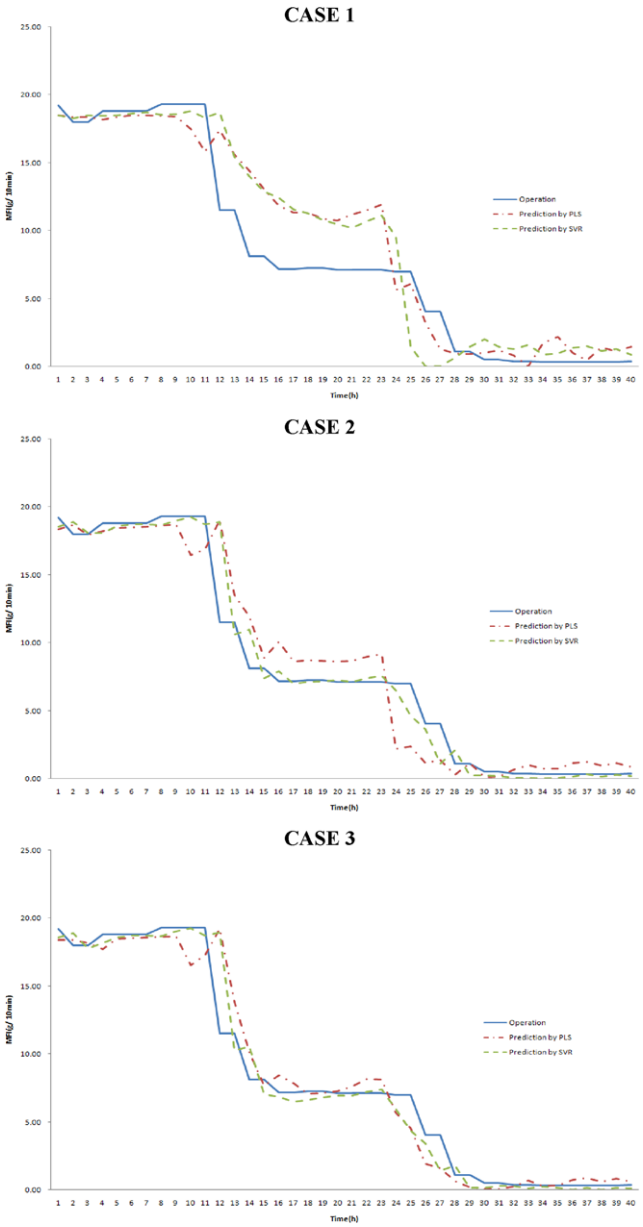

Fig. 6. Comparison of predicted MFI for each case.

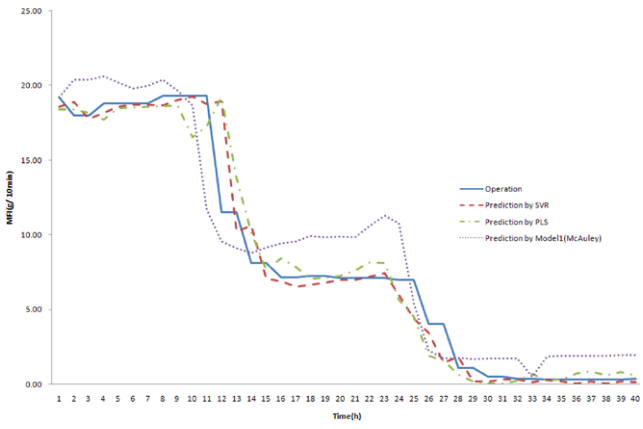

Fig. 7. Comparison of predicted MFI by PLS, SVR and Model 1 (McAuley).

3. For the prediction from 14th hour to 25th hour, Model 3 shows significant predicting errors compared to PLS and SVR as in the case of Model 1. Fig. 10 shows the performance of prediction using the model by Oh (Model 4) as well as that using proposed PLS and SVR methods. Again, we can see that PLS and SVR exhibit better predicting performance than Model 4. For the prediction from 14th

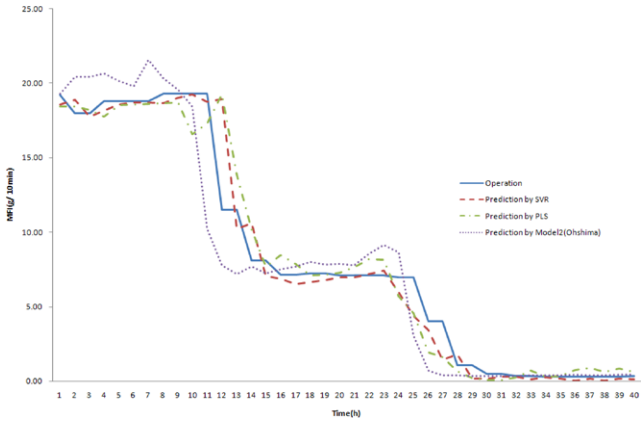

Fig. 8. Comparison of predicted MFI by PLS, SVR and Model 2 (Ohshima).

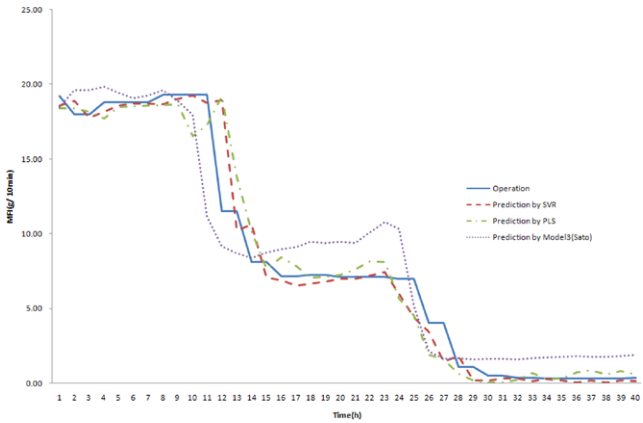

Fig. 9. Comparison of predicted MFI by PLS, SVR and Model 3 (Sato).

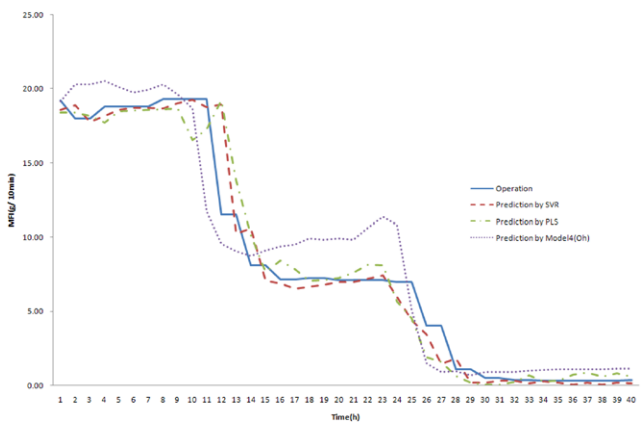

Fig. 10. Comparison of predicted MFI by PLS, SVR and Model 4 (Oh).

hour to 25th hour, Model 4 shows significant predicting errors compared to PLS and SVR. In short, the proposed predicting methods based on PLS and SVR show better tracking performance than other predicting methods reported so far.

CONCLUSION

In high density polyethylene (HDPE) processes the melt flow index (MFI) is the most important control variable indicating product quality. Because of the difficulty in the on-line measurement of MFI, many MFI estimation and correlation methods have been proposed. In this work, mechanical predicting methods such as partial least squares (PLS) method and support vector regression (SVR) method are employed in contrast to conventional dynamic prediction schemes. Results of predictions are compared with other prediction results obtained from various dynamic prediction schemes to evaluate predicting performance. We evaluated four different predicting models reported so far to demonstrate the effectiveness of the proposed predicting methods based on PLS and SVR. Hourly MFIs are predicted and compared with operation data for the high density polyethylene process involving frequent grade changes. We can see that PLS and SVR exhibit excellent predicting performance even for severe operating situations accompanying frequent grade changes. On-line application of the proposed predicting methods in the plant operation is yet to be investigated.

ACKNOWLEDGEMENTS

This work is the outcome of a Manpower Development Program for Energy supported by the Ministry of Knowledge and Economy (MKE).

REFERENCES

1. K. B. McAuley and J. F. MacGregor, *AIChE J.*, **37**, 825 (1991).
2. K. B. McAuley and J. F. MacGregor, *AIChE J.*, **39**, 855 (1993).
3. M. Ogawa, M. Ohshima, K. Morinaga and F. Watanabe, *J. Process Control*, **9**, 51 (1999).
4. M. Ohshima and M. Tanigaki, *J. Process Control*, **10**, 135 (2000).
5. S. J. Oh, J. S. Lee and S. W. Park, *Ind. Eng. Chem. Res.*, **44**, 8 (2005).
6. E. H. Lee, T. Y. Kim and Y. K. Yeo, *Korean J. Chem. Eng.*, **25**(4), 613 (2008).
7. M. Karthikeyan, *J. Chem. Inf. Model.*, **45**, 581 (2005).
8. J. Zhang, *Neural Networks*, **12**, 927 (1999).
9. A. Afantitis, G Melagraki, K. Makridima, A. Alexandridis, H. Sarimveis and O. Iglessi-Markopoulou, *J. Molecular Structure*, **716**, 193 (2005).
10. V. Tantishaiyakul, N. Worakul and W. Wongpoowarak, *International J. Pharm.*, **325**, 8 (2006).
11. J. Shi, X. Liu and Y. Sun, *Neurocomputing*, **70**, 280 (2006).
12. J. Shi and X. Liu, *J. Appl. Polymer Sci.*, **101**, 285 (2006).
13. K. G Min, C. H. Han and K. S. Chang, *Korean J. Chem. Eng.*, **5**(2), 2437 (1999).
14. C. K. Park, *Korean Operations Research and Management Society*, **23**(2), 75 (2006).
15. C. Sato, T. Ohtani and H. Nishitani, *Comput. Chem. Eng.*, **24**, 945 (2000).