NADEEM M. KHALFE
SANDIP KUMAR LAHIRI
SUNIL KUMAR SAWKE

Jubail Industrial City, Saudi Arabia

SOFT SENSOR FOR BETTER CONTROL OF CARBON DIOXIDE REMOVAL PROCESS IN ETHYLENE GLYCOL PLANT

Low carbon dioxide in cycle gas loop of ethylene glycol (EG) plant improves catalyst selectivity and overall economics of the plant. Carbon dioxide produced as a by-product in ethylene oxide reactor is removed by the Benfield process. In this process, the carbonate and bicarbonate ratio in lean carbonate solution is considered as an important quality control (QC) variable as the efficiency of carbon dioxide removal largely depends on it. In the event of a process malfunction or operating under suboptimal condition, the CO_2 content in the cycle gas loop will continue to rise until corrective action is taken after obtaining lab results for carbonate and bicarbonate ratio. This time consuming sampling process can be overcome by implementing a technological solution in form of an accurate and robust mathematical model capable of real time QC variable prediction. For well understood processes, the structure of the correlation for QC variables as well as the choice of the inputs may be well known in advance. However, the Benfield process is too complex and the appropriate form of the correlation and choice of input variables are not obvious. Here, knowledge of the processes, operating experience and statistical methods were applied in developing the soft sensor. This paper describes a systematic approach to the development of inferential measurements of carbonate and bicarbonate ratio using Support Vector Regression (SVR) analysis. Given historical process data, a simple SVR-based soft sensor model is found capable of identifying and capturing the cause and effect relationship between operating variables (model inputs) and QC variables (model outputs). Special care was taken to choose input variables, so that the final correlation and regression coefficient make senses from process engineering point of view. The developed soft sensor was implemented in commercial ethylene glycol plant in an Exaquantum interface and was found to satisfactorily predict the carbonate and bicarbonate ratio in real time.

Key words: support vector regression; soft sensor; modeling.

Soft sensors have been reported to supplement online instrument measurements for process monitoring and control. Both model-based and data-driven soft sensors have been developed. If a first principle model (FPM) describes the process sufficiently accurately, a model-based soft sensor can be derived [1]. However, a soft sensor based on detailed FPM is computationally intensive for real-time applications. Difficulties associated with the construction and solution of phenomenological models for soft sensors necessitate exploration of alternative modeling formalisms. Modern measurement techniques enable a large amount of operating data to be collected, stored and analyzed, thereby rendering data-driven soft sensor development a viable alternative.

In conventional empirical soft sensor modeling, appropriate linear or nonlinear models are constructed exclusively from the process input-output data without invoking the process phenomenology. A fundamental deficiency of the conventional empirical modeling approach is that the structure (functional form) of the data-fitting model must be specified a priori. Satisfying this requirement, especially for nonlinearly behaving processes is a cumbersome task since it involves selecting heuristically an appropriate nonlinear model structure from numerous alternatives.

In the last decade, artificial neural networks (ANNs) and more recently SVR have emerged as two attractive tools for nonlinear modeling especially in situations where the development of phenomenological or conventional regression models becomes impractical or cumbersome. In recent years, SVR [2-4], which is a statistical learning theory based machine, learning formalism is gaining popularity over ANN due to its many attractive features and promising empirical performance. Although the foundation of the SVR paradigm was laid down in mid 1990s, its chemical engineering applications such as fault detection have emerged only recently.

Building on these studies, the focus of this work is to develop a soft sensor for estimation of carbonate to bicarbonate ratio in carbon dioxide removal unit of ethylene glycol plant. It was reported in literature [3,4] that SVR based soft sensors can handle noise in process parameters (this type of noise in common in indication of DCS) and gives better performance than normal nonlinear regression based soft sensors. Based on the potential of SVR to regress complex function, an attempt has been made in the present study to explore the computational capability of SVR in the field of soft sensor development in petrochemical industry. This paper presents a systematic approach using robust SVR techniques to build a soft sensor from available process measurements.

SVR-BASED SOFT SENSOR MODELING

The problem of empirical data modeling is germane to many engineering applications. In empirical data modeling a process of induction is used to build up a model of the system, from which it is hoped to deduce responses of the system that have yet to be observed. Ultimately the quantity and quality of the observations govern the performance of this empirical model. By its observational nature data obtained is finite and sampled; typically this sampling is non-uniform and due to the high dimensional nature of the problem the data will form only a sparse distribution in the input space. Traditional neural network approaches have suffered difficulties with generalization, producing models that can over fit the data. This is a consequence of the optimization algorithms used for parameter selection and the statistical measures used to select the "best" model. The foundations of support vector machines (SVM) have been developed and are gaining popularity due to many attractive features, and promising empirical performance. The formulation embodies the structural risk minimization (SRM) principle, which has been shown to be superior to tra-

ditional empirical risk minimization (ERM) principle, employed by conventional neural networks. SRM minimizes an upper bound on the expected risk, as opposed to ERM that minimizes the error on the training data. It is this difference which equips SVM with a greater ability to generalize, which is the goal in statistical learning. SVMs were developed to solve the classification problem, but recently they have been extended to the domain of regression problems. The salient features of SVR are: *i*) like ANNs, SVR is an exclusively data based nonlinear modeling paradigm, *ii*) SVR based models are based on the principle of structural risk minimization, which equips them with greater potential to generalize, *iii*) automated abstraction: SVR can determine the essentials of input-output relationships automatically. We do not need a domain expert, that is, an expert in a particular problem solving domain (*e.g.* ethylene glycol plant specialist) to develop the "knowledge base" that expert systems require. Through training with direct (and sometimes imprecise) numerical data, the SVR can automatically determine cause-and-effect relations and develop its own knowledge base, *iv*) potential for online use: SVR may take a very long time to train, but once trained, they can calculate results from a given input very quickly. Since a trained SVR may take less than a second to calculate results, it has the potential to be used online in a control system, *v*) parameters of an SVR model are obtained by solving a quadratic optimization problem, *vi*) the objective function in SVR being of quadratic form, it possesses a single minimum thus avoiding the heuristic procedure involved in locating the global or the deepest local minimum on the error surface, *vii*) in SVR, the inputs are first nonlinearly mapped into a high dimensional feature space wherein they are correlated linearly with the output, *viii*) noise tolerance: Industrial process data and laboratory analysis data contains noise. This noise comes from various transmitters measuring thousands of process variables. The main advantage of SVR to build soft sensor model is its tolerance to noise.

In the following, the basic concepts of SVR are introduced. A more detailed description of SVR can be found in the literature [4,5].

Consider a training data set $g = \{(x_1, y_1), (x_2, y_2), ..., (x_P, y_P)\}$, such that $x \in R^N$ is a vector of input variables and $y \in R^N$ is the corresponding scalar output (target) value. Here, the modeling objective is to find a regression function, $y = f(x)$, such that it accurately predicts the outputs $\{y\}$ corresponding to a new set of input-output examples, $\{(x,y)\}$, which are drawn from the same underlying joint probability distribution,

$P(x,y)$, as the training set. To fulfill the stated goal, SVR considers following linear estimation function:

$$f(x) = \langle w, \varphi(x) \rangle + b \tag{1}$$

where w denotes the weight vector; b refers to a constant known as "bias"; $f(x)$ denotes a function termed feature, and $\langle w, \varphi(x) \rangle$ represents the dot product in the feature space, ℓ, such $\varphi : x \to \ell, w \in \ell$. In SVR, the input data vector, x, is mapped into a high dimensional feature space, ℓ, via a nonlinear mapping function, ϕ, and a linear regression is performed in this space for predicting y. Thus, the problem of nonlinear regression in lower dimensional input space R^N is transformed into a linear regression in the high dimensional feature space, ℓ. Accordingly, the original optimization problem involving nonlinear regression is transformed into finding the flattest function in the feature space ℓ and not in the input space, x. The unknown parameters w and b in Eq. (1) are estimated using the training set, g. To avoid over fitting and thereby improving the generalization capability, following regularized functional involving summation of the empirical risk and a complexity term, $\|w\|^2$, is minimized:

$$R_{reg}[f] = R_{emp}[f] + \lambda \|w\|^2 =$$
$$= \sum_{i=1}^{P} C(f(x_i) - y_i) + \lambda \|w\|^2 \tag{2}$$

where R_{reg} and R_{emp} denote the regression and empirical risks, respectively; $\|w\|^2$ is the Euclidean norm; C is a cost function measuring the empirical risk, and $\lambda > 0$, is a regularization constant. For a given function, f, the regression risk (test set error), $R_{reg}(f)$, is the possible error committed by the function f in predicting the output corresponding to a new (test) example input vector drawn randomly from the same sample probability distribution, $P(x,y)$, as the training set. The empirical risk $R_{emp}(f)$, represents the error (termed "training set error") committed in predicting the outputs of the training set inputs. Minimization task described in Eq. (2) involves: i) minimization of the empirical risk function $R_{emp}(f)$ and ii) obtaining as small a w as possible, using the training set g. The commonly used loss function is the "ε-insensitive loss function" given as:

$$C(f(x) - y) = |f(x) - y| - \varepsilon \qquad \text{for } |f(x) - y| \geq \varepsilon$$
$$= 0 \qquad\qquad \text{otherwise} \tag{3}$$

where ε is a precision parameter representing the radius of the tube located around the regression function (Figure 1); the region enclosed by the tube is known as "ε-intensive zone".
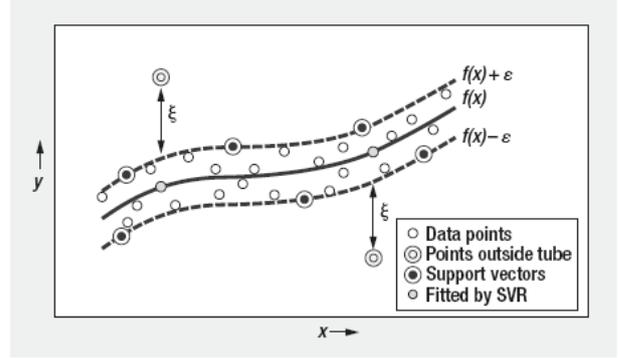

Figure 1. Support vector scematic regression using ε-sensitive loss function.

The SVR algorithm attempts to position the tube around the data as shown in Figure. 1.

The optimization criterion in Eq. (3) penalizes those data points whose y values lie more than ε distance away from the fitted function, $f(x)$. In Figure 1, the sizes of the stated excess positive and negative deviations are depicted by ζ and ζ^*, which are termed "slack" variables. Outside of the $[-\varepsilon, \varepsilon]$ region, the slack variables assume nonzero values. The SVR fits $f(x)$ to the data in a manner such that: i) the training error is minimized by minimizing ζ and ζ^* and ii) w^2 is minimized to increase the flatness of $f(x)$ or to penalize over complexity of the fitting function [3] showed that the following function possessing finite number of parameters can minimize the regularized function in Eq. (2):

$$f(x, \alpha, \alpha^*) = \sum_{i=1}^{P} (\alpha_i - \alpha_{i*}) K(x, x_i) + b \tag{4}$$

where α_i and α_{i*} (≥ 0) are the coefficients (Lagrange multipliers) satisfying $\alpha_i \alpha_{i*} = 0$, $i = 1, 2, ..., P$, and $K(x, x_i)$ denotes the so called "kernel" function describing the dot product in the feature space. The kernel function is defined in terms of the dot product of the mapping function as given by:

$$K(x_i, x_j) = \langle \varphi(x_i), \varphi(x_j) \rangle \tag{5}$$

The advantage of this formulation (Eqs. (4) and (5)) is that for many choices of the set $\{\varphi_i(x)\}$, including infinite dimensional sets, the form of K is analytically known and very simple. Accordingly, the dot product in the feature space (i, j) can be computed without actually mapping the vectors x_i and x_j into that space (i.e., computation of $\varphi(x(i))$ and $\varphi(x(j))$). There exist several choices for the kernel function K; the two commonly used kernel functions, namely, radial basis function (RBF) and n^{th} degree polynomial is defined below in Eqs. (6) and (7), respectively.

$$K(x_i, x_j) = \exp\left(\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \qquad (6)$$

$$K(x_i, x_j) = \left[1 + [(x_i, x_j)]^n\right] \qquad (7)$$

In Eq. (4), the coefficients α_i and α_{i^*} are obtained by solving following quadratic programming problem.

Maximize:

$$R(\alpha^*, \alpha) = -0.5\sum_{i,j=1}^{P}(\alpha_{i^*} - \alpha_i)(\alpha_{j^*} - \alpha_j)K(x_i, x_j) -$$
$$-\varepsilon\sum_{i=1}^{P}(\alpha_{i^*} + \alpha_i) + \sum_{i=1}^{P}y_i(\alpha_{i^*} - \alpha_i)$$

subject to constraints:

$$0 \le \alpha_i, \alpha_{i^*} \le C, \forall i \text{ and } \sum_{i=1}^{P}(\alpha_{i^*} - \alpha_i) = 0 \qquad (8)$$

Having estimated α, α^* and b, using a suitable quadratic programming algorithm, the SVR-based regression function takes the form:

$$f(x, w) = f(x, \alpha, \alpha^*) = \sum_{i=1}^{P}(\alpha_{i^*} - \alpha_i)K(x_i, x) + b \qquad (9)$$

where vector w is described in terms of the Lagrange multipliers α and α^*. Owing to the specific character of the above-described quadratic programming problem, only some of the coefficients, $(\alpha_i - \alpha_{i^*})$ are non-zero and the corresponding input vectors, x_i, are called support vectors (SVs). The SVs can be thought of as the most informative data points that compress the information content of the training set. The coefficients α and α^* have an intuitive interpretation as forces pushing and pulling the regression estimate f (x_i) towards the measurements, y_i. In Eq. (9), the bias parameter, b, can be computed as follows:

$$b = \begin{cases} y_i - f(x_i)_{b=0} - \varepsilon \\ y_i - f(x_i)_{b=0} + \varepsilon \end{cases} \text{ for } \alpha_i, \alpha_{i^*} \in (0, C) \qquad (10)$$

where x_i and y_i denote the ith support vector and the corresponding target output, respectively. In the SVR formulation, C and ε are two user-specified free parameters; while C represents the trade-off between the model-complexity and the approximation error, ε signifies the width of the ε-insensitive zone used to fit the training data. The stated free parameters together with the specific form of the kernel function control the accuracy and generalization performance of the regression estimate. The procedure of judicious selection of C and ε is explained by [6,7].

Process description

In an ethylene glycol plant, ethylene oxide is produced by the gas phase catalytic reaction of ethylene and oxygen (Eq. (11)) in a shell and tube type reactor at 20 bar pressure and high temperature. The side reaction (Eq. (12)) also occurrs at that temperature and pressure and undesired CO_2 is produced.

Primary reaction:

$$C_2H_4 + \tfrac{1}{2}O_2 \rightarrow C_2H_4O \qquad (11)$$

Secondary reaction:

$$C_2H_4 + 3O_2 \rightarrow 2CO_2 + 2H_2O \qquad (12)$$
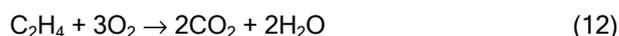
This CO_2 has to be removed from the cycle gas system otherwise it will decrease the catalyst selectivity and increase the system pressure.

The CO_2 from the cycle gas system is removed by the Benfield process where CO_2 is absorbed by contacting cycle gas with hot potassium carbonate solution in a CO_2 absorber column called a contactor (Eq. (13)). The outlet carbonate solution from the contactor (contains unreacted potassium carbonate (K_2CO_3), potassium bicarbonate ($KHCO_3$) and water (H_2O)) is then flashed in two low pressure flash vessels to remove dissolved ethylene and methane. The liquid outlet from the flash vessel is fed to regenerator column, where the potassium bicarbonate is again converted back to carbonate using heat (Eq. (14)).

This regenerated carbonate solution is then pumped and recycled back to contactor column for reuse.

$$K_2CO_3 + CO_2 + H_2O \rightarrow 2KHCO_3 \qquad (13)$$

$$2KHCO_3 \rightarrow K_2CO_3 + CO_2 + H_2O \qquad (14)$$
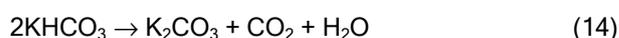
This CO_2 absorption by carbonate solution and CO_2 stripping in regenerator column is a continuous process and has a long term effect on reactor performance and overall glycol plant economics. The CO_2 removal capacity of the carbonate solution depends mainly on carbonate flow, total carbonate strength in solution, carbonate to bicarbonate ratio and inlet cycle gas CO_2 concentration. The CO_2 removal capacity of carbon dioxide removal section has to be gradually increased throughout Catalyst life (usually 3 years) as the catalyst selectivity drops from Start of run (SOR) to End of run (EOR) and consequently the CO_2 generation in EO reactor gradually increases. That is why carbonate flow gradually increases from SOR to EOR of catalyst life. The carbonate to bicarbonate ratio represents the regenerator column performance and has a profound effect on CO_2 absorption in the contactor. If carbonate to bicarbonate ratio in lean carbonate solution is low it will reduce the CO_2 absorption capacity in contactor and contactor outlet gas CO_2 will increase. One of the main reasons for low carbonate to bicarbonate ratio is that less heat is applied in regenerator column and thus bicarbonate to carbonate

conversion (Eq. (14)) is low. Unlike other parameters which affect the CO_2 removal unit (CRU) performance, the carbonate to bicarbonate ratio is not measured on a continuous basis and only offline laboratory sampling is available. There is a need for continuous online analysis of this ratio as it will help to monitor and adjust the regenerator performance.

Due to the nonlinearity of the dynamics and involvement of complex electrolyte chemistry, it is very difficult to develop a first principle based model for the regenerator and the contactor. Also, the few phenomenological based models which were available in literature were too complex and take long execution time and thus make them unsuitable for use in online soft sensors. Based on this study SVR was tried for making online soft sensor.

MAKING OF ONLINE SOFT SENSOR FOR CARBONATE TO BICARBONATE RATIO

Input selection

After an extensive literature survey all physical parameters that influence the carbonate-bicarbonate ratio are put in a so-called "wish-list". Data were collected for all the possible inputs. Principal component analysis (PCA) method and plant operating experience were used to shortlisted the input variables which qualify for final model building. Finally, those inputs were exposed to SVR algorithm and the input set which gives minimum AARE in test set is finally selected.

Based on the above approach and plant operating experience, following input variables were short listed which have an effect on carbonate to bicarbonate ratio (Figure 2):

1. Density of lean carbonate solution (D1).
2. Steam flow in reboiler (F1).
3. Direct steam flow in regenerator (F2).
4. Regenerator bottom temperature (T1).
5. Lean Carbonate flow (F3).

Data collection

Hourly average of actual plant operation data was collected from Exaquantum (Yokogawa Process Database Management System) for whole year. All the data were collected at different age of catalyst and at different carbonate flow to cover a wide range operating data. The carbonate/bicarbonate ratio data was collected from lab sample (thrice per week frequency) for one year.

Data regression

Initially, all input and output data was put in a spreadsheet and data alignment was done. Based on process experience and looking into the dynamics of the process, plant operating data are averaged for 3 h
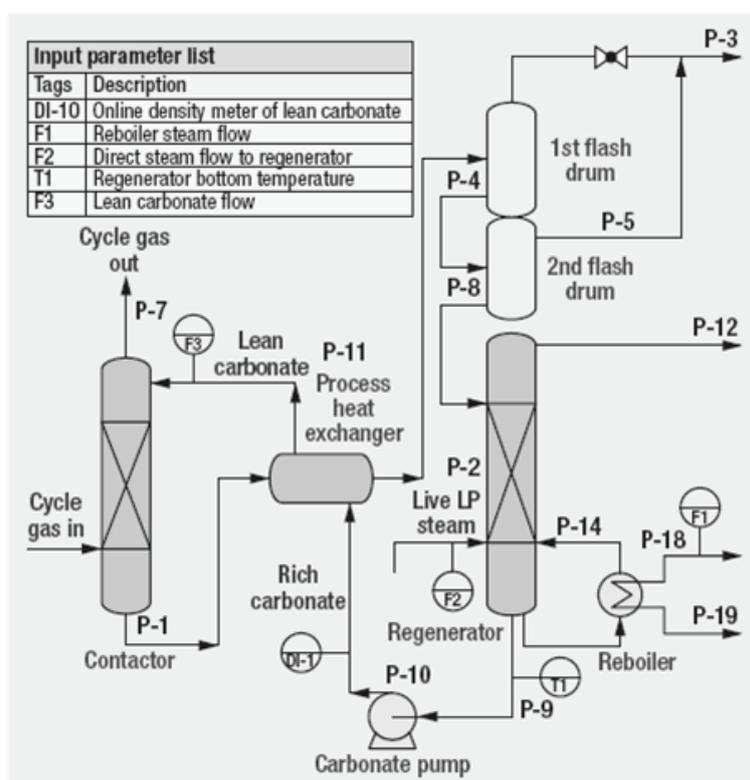

Input parameter list	
Tags	Description
DI-10	Online density meter of lean carbonate
F1	Reboiler steam flow
F2	Direct steam flow to regenerator
T1	Regenerator bottom temperature
F3	Lean carbonate flow

Figure 2. Block diagram of a CO_2 removal unit.

basis and align with the less frequent lab analysis. Nonlinear data regression was tried to get a correlation which will represent the ratio as a function of all input variables listed above. All types of equations were tried (polynomial with varying degree, exponential radial basis function, spline, etc.). Input parameters were arranged in those equations which make engineering sense. The statistical analysis of prediction is based on the following performance criteria:

1. The average absolute relative error (AARE) should be minimum:

$$AARE = \frac{1}{N}\sum_{1}^{N}|(\frac{y_{predicted} - y_{actual}}{y_{actual}})| \qquad (15)$$

2. The correlation coefficient (CC) should be maximum (as near to unity):

$$CC = \frac{\sum_{1}^{N}(y_{actual} - \overline{y}_{actual})(y_{predicted} - \overline{y}_{predicted})}{\sqrt{\sum_{1}^{N}(y_{actual} - \overline{y}_{actual})^2 \sum (y_{actual} - \overline{y}_{actual})^2}} \qquad (16)$$

After several trials and errors it was found that the performance of simple nonlinear regression (polynomial regression) was very poor and none of the equations gave acceptable AARE and CC. Also, the prediction was unable to catch the increase and reduction of ratio (Figure 3) which is considered as a basic requirement of a soft sensor. After extensive literature survey, SVR was found promising to correlate such type of difficult parameters. So, a computer program was made for SVR based on the above description.

RESULTS AND DISCUSSIONS

After collecting around 160 sets of data, we segregate 80% data as a training set and the rest 20% (chosen randomly) as the test set. SVR was performed on the training set data but performance was judged by the test set data.

To obtain an optimal SVR model, it is necessary to examine the effects of the kernel function and other algorithm [6] specific parameters; the three kernel functions that were tested are polynomial, RBF and spline. Among these, RBF resulted in the least AARE values for test sets (Table 1). The number of support vectors used by the SVR algorithm for fitting the carbonate to bicarbonate ratio model is 144 which corresponds to 90% of data. The optimal values of the three SVR-specific parameters namely, width of RBF kernel (s), regularization parameter (λ), risk function parameter (ε) that minimized the training and test error are listed in Table 1. These values were arrived to after running the SVR numerous times for all the different combinations of the above three parameters. Also listed are the values of correlation coefficients for the test set predictions along with the corresponding AARE values for different models. A comparison of the SVR model predicted and the corresponding target values of carbonate and bicarbonate ratio is depicted in Figure 4. From Figure 4 it is evident that there is a very good agreement in SVR prediction and actual lab analysis of carbonate and bicarbonate ratio.

Figure 5 represents the prediction success of SVR algorithm against actual lab analysis. One or two outlier data in the parity plot may be due to the
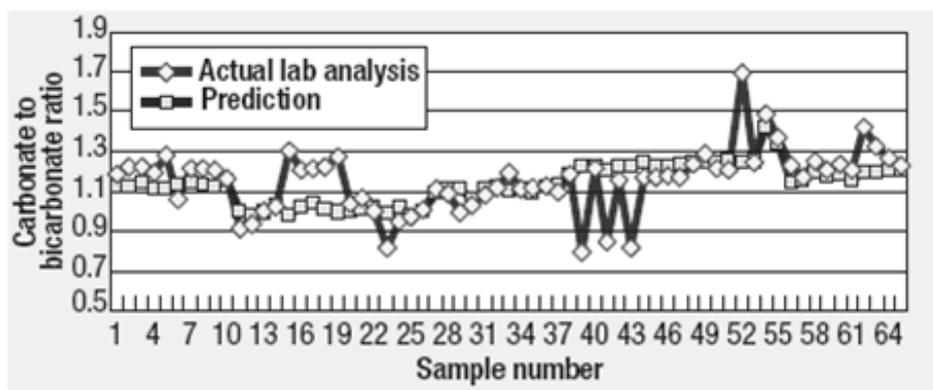

Figure 3. Actual vs prediction for simple nonlinear regression model.

Table 1. SVR algorithm performance for different kernel types (CC = correlation coefficient, C = cost function, ε = risk function parameter)

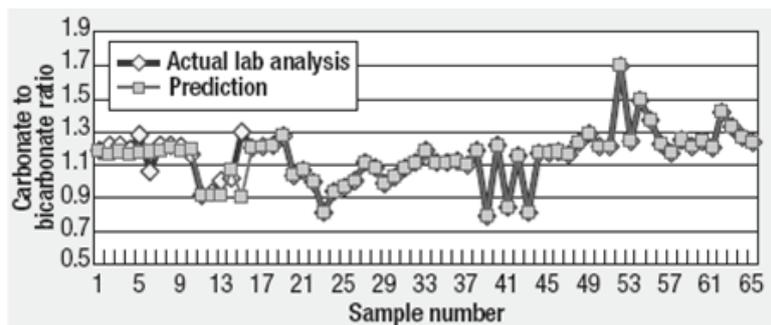| No.	Kernel type	AARE	CC	C	ε	Kernel parameter
1	Polynomial	0.067	0.637	10000	0	2
2	RBF	0.012	0.937	10000	0.01	3
3	Spline	0.068	0.611	10000	0.01	3

22

Figure 4. Actual vs prediction of carbonate to bicarbonate ratio using RBF kernel in SVR.
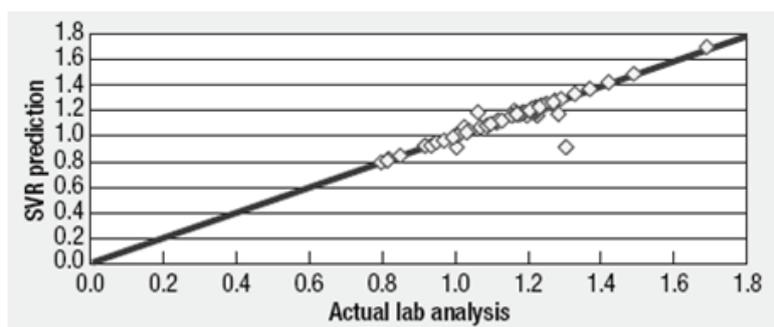

Figure 5. Parity plot for SVR prediction and actual lab analysis of carbonate to bicarbonate ratio.

misrepresentative lab analysis which also shows the tolerance of SVR algorithm against outliers.

Implementation and benefit

After getting the satisfactory performance of SVR prediction, the formula was applied to a same type but another ethylene glycol plant data and performance was found quite satisfactory. This proves the generalization capability and robustness of developed soft sensor. The soft sensor formula was put in the Exaquantum interface to operate it on a real time basis and performance was monitored in commercial ethylene glycol plant for one month against lab analysis. It was found to be in good agreement with lab results and was capturing the downward and upward trend of this ratio properly. After getting the online real time indication for carbonate and bicarbonate ratio, it is now becoming easy for a panel operator to vary steam in the regenerator to maintain this ratio within the acceptable limits. By controlling this ratio, the overall performance of the carbon dioxide removal unit is improved, which in turn has a long term benefit in catalyst selectivity and overall economics of EG plant.

CONCLUSION

This paper presents a systematic approach for building a soft sensor which is useful to optimize the carbon dioxide removal unit in EG plant. This paper shows how the recent support vector based regression can be advantageous and effective in complex processes where the normal nonlinear regression based approach fails. The small prediction error (within 1.2%) and high correlation coefficient (0.937) make the SVR algorithm promising for future soft sensor development.

REFERENCES

[1] W.W. Yan, H.H. Shao, X.F. Wang, Comput. Chem. Eng. **28**(8) (2004) 1489-1498.

[2] F. Chauchard, J. Svensson, J. Axelsson, S. Andersson-Engels, S. Roussel, Chemom. Intell. Lab. Syst. **91**(1) (2008) 34-42.

[3] N.S. Kaveh, F. Mohammadi, S.N. Ashrafizadeh, Chem. Eng. J. **147**(2-3) (2009) 161-172.

[4] P.P. Gupta, S.S. Merchant, A.U. Bhat, Ind. Eng. Chem. Res. **48**(21) (2009) 9631-9654.

[5] Y. Liu, N.P. Hu, H.Q. Wang, P. Li, Ind. Eng. Chem. Res. **48**(12) (2009) 5731-5741.

[6] H.Q, Wang, P. Li, F.R Gao, Z.H. Song, S.X. Ding, AIChE J. **52**(10) (2006) 3515-3531.

[7] V. Vapnik, S. Golowich, A. Smola, Neural Inform. Proc. Syst. **9** (1996) 281-287.

NADEEM M. KHALFE

SANDIP KUMAR LAHIRI

SUNIL KUMAR SAWKE

Jubail Industrial City, Saudi Arabia

NAUČNI RAD

VIRTUELNI SENZORI ZA BOLJU KONTROLU PROCESA OTKLANJANJA UGLJENIK(IV)-OKSIDA IZ POSTROJENJA ZA PROIZVODNJU ETLEN-GLIKOLA

Mala koncentracija ugljenik(IV)-oksida u gasnoj petlji postrojenja za etilen-glikol pobolj-šava selektivnost katalizatora i ukupnu ekonomiku postrojenja. Ugljenik(IV)-oksid dobijen kao sporedni proizvod u reaktoru etilen-oksida uklanja se procesom po Benefildu. U ovom procesu odnos karbonata i bikarbonata u slabom rastvoru karbonata se smatra va-žnim pokazateljem kvaliteta kontrole, jer efikasnost uklanjanja ugljenik(IV)-oksida uglav-nom zavisi od njega. U slučaju greške u procesu ili rada pod suboptimalnim uslovima, sadržaj ugljenik(IV)-oksida u gasnoj petlji će nastaviti da raste sve dok se ne preduzme korektivna akcija posle dobijanja laboratorijskih rezultata za odnos karbonata i bikarbo-nata. Vreme koje se troši na uzorkovanje može se uštedeti implementiranjem tehnološ-kog rešenja u obliku tačnog i robustnog matematičkog modela koji može predvideti po-kazatelj kvaliteta kontrole u realnom vremenu. Za proces koji se dobro razume, struktura korelacije za pokazatelje kvaliteta kontrole, kao i izbor ulaznih promenljivih, može biti unapred poznat. Međutim, proces po Benefildu je isuviše složen, pa pogodan oblik kore-lacije i izbor ulaznih promenljivih nisu jasni. U radu su primenjeni znanja o procesu, rad-no iskustvo i statističke metode da bi se razvio virtuelni senzor. Rad opisuje sistematski pristup razvoju deduktivnih merenja odnosa karbonata i bikarbonata koristeći SVR (sup-port vector regression) analizu. Na osnovu istorijskih procesnih podataka, definisan je jednostavan model virtuelnog senzora zasnovanog na SVR koji može da identifikuje i hvata zavisnost uzroka i efekta između procesnih promenljivih (ulazi modela) i poka-zatelja kontrole kvaliteta (izlazi modela). Posebna pažnja je posvećena izboru ulaznih promenljivih, tako da krajnja korelacija i regresioni koeficijent imaju smisla sa inžinjerske tačke gledišta. Razvijeni virtuelni senzor je iskorišćen u komercijalnom postrojenju za eti-len-glikol u interfejsu sistema Exaquantum-a. Utvrđeno je da on zadovoljavajuće pred-viđa odnos karbonata i bikarbonata u realnom vremenu.

Ključne reči: regresija pomoćnog vektora; virtuelni senzor; modelovanje.