Novel soft sensor modeling and process optimization technique for commercial petrochemical plant

sandip k lahiri
Novel soft-sensor modeling and process optimization technique for commercial petrochemical plant

S. K. Lahiri* and Nadeem M. Khalfe

Saudi Arabian Basic Industries Corporation, Kingdom of Saudi Arabia

Received 14 January 2009; Revised 14 August 2009; Accepted 19 August 2009

ABSTRACT: Soft sensors have been widely used in industrial process control to improve the quality of product and assure safety in production in real-time basis. The core of a soft sensor is to construct a soft sensing model. This paper proposes a new soft sensing modeling method based on a recent advanced computational technique called support vector regression (SVR). The major advantage of the strategies is that soft sensor modeling can be conducted exclusively from the historic process data wherein the detailed knowledge of process phenomenology (reaction mechanism, kinetics, etc.) is not required.

Ultraviolet (UV) transmittance is one of the most important quality variables of monoethylene glycol (MEG) product that has impact on the polyester product quality. UV transmittance measures the presence of undesirable compounds in MEG that absorb light in the UV region of the spectrum and indirectly measures the impurity of MEG product. Off-line laboratory method for MEG UV measurement is common practice among the manufacturer, where a sample is withdrawn several times a day from the product stream and analyzed by time-consuming laboratory analysis. In the event of a process malfunction or operating under suboptimal condition, the plant continues to produce off-specification (off-spec) product until laboratory results become available. It results in enormous financial losses for a large-scale commercial plant. In this paper, a soft sensor was developed to predict the UV transmittance on real-time basis and an online hybrid SVR-differential evolution (DE) technique was used to optimize the process parameters so that UV is maximized. This paper describes a systematic approach for the development of inferential measurements of UV transmittance using SVR analysis. After predicting the UV accurately, model inputs are optimized using DEs to maximize the UV. The optimized solutions when verified in actual commercial plant resulted in a significant improvement in the MEG quality. © 2009 Curtin University of Technology and John Wiley & Sons, Ltd.

KEYWORDS: SVR; DE; soft sensor; modelling; optimization
products. Even when hardware sensors can be used, operators and engineers have found the problems listed in Table 1. These problems with hardware sensors were identified as the results of a questionnaire to 26 companies in Japan (Process System Engineering 143 Committee, 2004). Soft sensors are judged to be useful for addressing these problems.

For successful monitoring and control of chemical plants, there are important quality variables that are difficult to measure online, because of limitations such as cost, reliability and long dead time. These measurement limitations may cause important problems such as product loss, energy loss, toxic byproduct generation and safety problem. A soft sensor, an inferential model, can estimate the qualities of interest online using other available online measurements such as temperatures and pressures. An inferential sensor provides valuable real-time information that is necessary for effective QC. The major purpose of using soft sensors is to (1) stabilize product quality through its online estimation, (2) reduce energy and material consumption through effective operation close to specifications/constraints and (3) validate online analyzers by comparison with the soft sensors. The soft sensor can be derived from the first-principle model when the model offers the sufficient accuracy within the reasonable computation time. However, owing to complexity in industrial processes, there are cases when the first-principle model is not available, or sometimes it takes too much time to compute. As a result, empirical data driven models are the most popular ones to develop soft sensors. Empirical models are usually obtained based on various modeling techniques such as multivariate statistics, artificial neural network and support vector regressions (SVR). In recent years, SVR has been widely used as a useful tool to the nonlinear soft sensing modeling. SVR is a computer modeling approach that learns from examples through iterations without requiring a prior knowledge of the relationships of process parameters and QC variable. It is also capable of dealing with uncertainties, noisy data and nonlinear relationships. SVR modeling have been known as ‘effortless computation’ and readily used extensively because of their model-free approximation capabilities of complex decision-making processes. Once an SVR based process model is developed, it can be used for predicting the important quality variable. Also it can be interfaced with online DCS and continuous monitoring can be achieved to yield the better process control. This SVR-based process model can also be used for process optimization to obtain the optimal values of the process input variables that maximize the quality of product. In such situations, an efficient optimization formalism known as differential evolution (DE) can be used. The DEs were originally developed as the genetic engineering models mimicking population evolution in natural systems. Specifically, DE like genetic algorithm (GA) enforce the ‘survival-of-the-fittest’ and ‘genetic propagation of characteristics’ principles of biological evolution for searching the solution space of an optimization problem.

In this paper, SVR formalism is integrated with DEs to arrive at soft sensor modeling and process optimization strategies. The strategy (henceforth referred to as ‘SVR-DE’) use an SVR as the nonlinear process modeling paradigm for development of soft sensor and the DE for optimizing the input space of the SVR model such that an improved process performance is realized. This paper describes a systematic approach to the development of inferential measurements of ultraviolet (UV) transmittance [QC variable of monoethylene glycol (MEG) product in glycol plant] using SVR regression analysis. After predicting the UV accurately, model inputs describing process operating variables are optimized using DEs with a view to maximize the UV. The SVR-DE is a new strategy for chemical process modeling and optimization. The major advantage of the strategies is that modeling and optimization can be conducted exclusively from the historic process data wherein the detailed knowledge of process phenomenology (reaction mechanism, kinetics, etc.) is not required. The optimized solutions when verified in actual commercial plant resulted in a significant improvement in the MEG quality.

Table 1. Problems with hardware sensors.

<table>
<thead>
<tr>
<th>Percentage</th>
<th>Recognized problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>27</td>
<td>Time consuming maintenance</td>
</tr>
<tr>
<td>21</td>
<td>Need for calibration</td>
</tr>
<tr>
<td>15</td>
<td>Aged deterioration</td>
</tr>
<tr>
<td>13</td>
<td>Insufficient accuracy</td>
</tr>
<tr>
<td>10</td>
<td>Long dead time, slow dynamics</td>
</tr>
<tr>
<td>8</td>
<td>Large noise</td>
</tr>
<tr>
<td>2</td>
<td>Low reproducibility</td>
</tr>
<tr>
<td>4</td>
<td>Others</td>
</tr>
</tbody>
</table>

The results of a questionnaire to 26 companies in Japan.

© 2009 Curtin University of Technology and John Wiley & Sons, Ltd.
and percentage of the UV light transmitted through the MEG sample is measured. UV transmittance measures the presence of impurities in MEG that absorb light in the UV region of the spectrum. These undesirable compounds are in trace quantities in the parts per billion ranges and primarily unknown in chemical structure and they have hardly any effect on process parameters. Samples showing higher transmittance are considered to be of a greater quality grade. The low UV MEG product gives an undesirable color during the making of white polyester fiber. In glycol plant, the MEG is drawn off from MEG distillation column as product, its UV transmittance is affected by many things such as impurity formation in upstream ethylene oxide reactor, impurity formation and accumulation in MEG column bottoms due to thermal degradation of glycol, nonremoval and accumulation of aldehyde in the system, etc. That is why it is very difficult for any phenomenological model for UV prediction to succeed in industrial scenario. Normally, online UV analyzers are not available to monitor product MEG UV analysis in ethylene glycol plant. Offline methods for MEG QC is common practice among the manufacturer, where a sample is withdrawn from the process and product stream for laboratory analysis several times a day and analyzed by time-consuming laboratory analysis. In the event of a process malfunction or operating under suboptimal condition, the plant will continue to produce off-spec product until laboratory results become available. For a big world-class capacity plant, this represents a huge amount of off-spec production and results in enormous financial losses.

This necessitates the online UV sensors or analyzers, which can give UV continuously on real-time basis. Accurate, reliable and robust UV soft sensors can be a viable alternative in this scenario. Making of UV soft sensor is not an easy task, as rigorous mathematical model for MEG product UV is still not available in literature which can predict UV transmittance to minimize the dependency on laboratory analysis. The comprehensive process model is expected to take into account the various subjects, such as chemistry, chemical reaction, UV deteriorating compound generation and accumulation, which consequently become very complex. Industry needs this mathematical model to predict MEG UV on real-time basis so that the process parameters can be adjusted before the product goes off-spec.

In this study, the SVR-DE strategy has been used to model and optimize the MEG product UV for a commercial plant. The best sets of operating conditions obtained thereby when subjected to actual plant validation indeed resulted in significant enhancements in UVs.

The process optimization objective under consideration is expressed as:

Given the process data comprising values of the multiple process inputs and the corresponding values of the process outputs (MEG UV in this case), find the optimal values of the process inputs such that the prespecified measures of process performance is maximized.

The SVR-DE strategy fulfills the above-stated objective in two steps. In the first step, an SVR-based process model is developed. This model has the inputs describing process operating parameters and variables such as reflux ratio, reflux flow, MEG column top pressure, MEG column condenser pressure, MEG column control temperature, MEG column feed flow, upstream drying column control temperature, drying column bottom temperature, crude glycol reprocessing flow and its outputs represent process output variable MEG UV. In the second step of the SVR-DE procedure, the input space of the SVR model is optimized using a DE algorithm such that the optimized process inputs result in the enhanced values of the output variables.

This optimization problem can be formulated as:

Maximize

\[ UV = f(\text{reflux ratio}, \text{reflux flow}, \text{MEG column top pressure}, \text{MEG column condenser pressure}, \text{MEG column control temperature}, \text{MEG column feed flow}, \text{Drying column control temperature}, \text{drying column bottom temperature}, \text{off-spec glycol reprocessing flow}) \] (1)

SVR based modeling

Industrial data contain noise. Normally different transmitter, signal transmissions, etc. add these noises with process parameters. Normal regression techniques try to reduce the prediction error on noisy training data. This empirical risk minimization (ERM) principle is generally used in the classical methods such as the least-square methods, the maximum likelihood methods and traditional artificial neural networks (ANN). 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 by Vapnik (1995) and are gaining popularity because of many attractive features and promising
empirical performance. The formulation embodies the structural risk minimization (SRM) principle, which has been shown to be superior to traditional ERM principle, used 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 that 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.\(^{[3]}\) The salient features of SVR are: (1) like ANNs, SVR is an exclusively data-based nonlinear modeling paradigm, (2) SVR based models are based on the principle of SRM, which equips them with greater potential to generalize, (3) parameters of an SVR model are obtained by solving a quadratic optimization problem, (4) 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 and (5) in SVR, the inputs are first nonlinearly mapped into a high dimensional feature space wherein they are correlated linearly with the output.

SVMs have been successfully applied to a number of applications such as handwriting recognition, particle identification (e.g. muons), digital images identification (e.g. face identification), text categorization, bioinformatics (e.g. gene expression), function approximation and regression, database marketing and so on.

The SVM methodology and theory is well documented and details can be found in number of research papers and technical reports.\(^{[4-7]}\) Although the foundation of the SVR paradigm was laid down in the mid-1990s, its chemical engineering applications such as fault detection\(^{[5]}\) have emerged only recently.

In SVM, the ERM is replaced by the SRM principle, which seeks to minimize an upper bound of the expected risk. The SVR algorithm attempts to position a tube around the data as shown in Fig. 1. \(\varepsilon\) is a precision parameter representing the radius of the tube located around the regression function (Fig. 1); the region enclosed by the tube is known as ‘\(\varepsilon\)-intensive zone.’ The diameter of the tube ideally should be the amount of noise in the data. The optimization criterion in SVR penalizes those data points whose \(y\) values lie more than \(\varepsilon\) distance away from the fitted function, \(f(x)\). There are two basic aims in SVM. The first is to find a function \(f(x)\) that has at most \(\varepsilon\) deviation from each of the targets of the training inputs.

At the same time, we would like this function to be as flat as possible. This second aim is not as immediately intuitive as the first, but nevertheless important in the formulation of the optimization problem used to construct the SVR approximation.
above, the optimization object becomes:

$$\text{Minimize} \quad \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{n} (\xi_i + \xi_i^*)$$

With the constraints, Subject to

$$\begin{align*}
y_i - \langle w, x_i \rangle - b & \leq \varepsilon + \xi_i \\
\langle w, x_i \rangle + b - y_i & \leq \varepsilon + \xi_i^*
\end{align*}$$

$$\xi_i, \xi_i^* \geq 0$$

Where the constant $C > 0$ stands for the penalty degree of the sample with error exceeding $\varepsilon$. Two positive slack variables $\xi_i, \xi_i^*$ represent the distance from actual values to the corresponding boundary values of $\varepsilon$-tube. The SVR fits $f(x)$ to the data in a manner such that: (1) the training error is minimized by minimizing $\xi_i, \xi_i^*$ and (2) $w^2$ is minimized to increase the flatness of $f(x)$ or to penalize over complexity of the fitting function. A dual problem can then be derived by using the optimization method to maximize the function,

Maximize

$$\begin{align*}
-\frac{1}{2} \sum_{i,j=1}^{n} (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)(x_i, x_j) \\
- \varepsilon \sum_{i=1}^{n} (\alpha_i + \alpha_i^*) + \sum_{i=1}^{n} y_i (\alpha_i - \alpha_i^*)
\end{align*}$$

Subject to

$$\sum_{i=1}^{n} (\alpha_i - \alpha_i^*) = 0 \text{ and } 0 \leq \alpha_i, \alpha_i^* \leq C$$

Where, $\alpha_i, \alpha_i^*$ are Lagrange multipliers. 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 $\alpha$ and $\alpha^*$ have an intuitive interpretation as forces pushing and pulling the regression estimate $f(x_i)$ towards the measurements, $y_i$.

Where the constant $C > 0$ determines the tradeoff between flatness (small $w$) and the degree to which deviation larger than $\varepsilon$ are tolerated.

The SVM for function fitting obtained by using the above-mentioned maximization function is then given by,

$$f(x) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) \langle x_i, x \rangle + b$$

As for the nonlinear cases, the solution can be found by mapping the original problems to the linear ones in a characteristic space of high dimension, in which dot product manipulation can be substituted by a kernel function, i.e. $K(x_i, x_j) = \varphi(x_i)\varphi(x_j)$. In this study, different kernel function is used in the SVR. Substituting $K(x_i, x_j) = \varphi(x_i)\varphi(x_j)$ in above equation allows us to reformulate the SVM algorithm in a nonlinear paradigm. Finally, we have,

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

Training and testing

Training an SV machine consists of an iterative process in which the SVR is given the desired inputs along with the correct outputs for those inputs. It then seeks to alter its margin ($w$) and bias ($b$) to try and produce the correct output (within a reasonable error margin). If it succeeds, it has learned the training set and is ready to perform upon previously unseen data. If it fails to produce the correct output it re-reads the input and again tries to produce the correct output. The margins and bias are slightly adjusted during each iteration through the training set (known as a training cycle) until the appropriate margins and bias has been established. Depending upon the complexity of the task to be learned, many thousands of training cycles may be needed for the SVR to correctly identify the training set. Once the output is correct the margins ($w$) and bias ($b$) can be used with the same SVM on unseen data to examine how well it performs. SVM learning is considered successful only if the system can perform well on test data on which the system has not been trained. This capability of an SVM is called generalizability.

DE based optimization of SVR models

Having developed an SVR-based process model, a DE algorithm is used to optimize the $N$-dimensional input space of the SVR model. Conventionally, various deterministic gradient-based methods are used for performing optimization of the phenomenological models. Most of these methods require that the objective function should simultaneously satisfy the smoothness, continuity and differentiability criteria. Although the nonlinear relationships approximated by an SVR model can be expressed in the form of generic closed-form expressions, the objective function(s) cannot be guaranteed to satisfy the smoothness criteria. Thus, the gradient-based methods cannot be efficiently used for optimizing the input space of an SVR model and, therefore, it becomes necessary to explore alternative optimization formalisms, which are lenient towards the form of the objective function.

The principal features possessed by the DEs are: (1) they require only scalar values and not the second and/or first-order derivatives of the objective function, (2) capability to handle nonlinear and noisy objective functions, (3) they perform global search and thus are.
more likely to arrive at or near the global optimum and (4) DEs do not impose preconditions, such as smoothness, differentiability and continuity, on the form of the objective function.

DE, an improved version of GA, is an exceptionally simple evolution strategy that is significantly faster and robust at numerical optimization and is more likely to find a function's true global optimum. Unlike simple GA that uses a binary coding for representing problem parameters, DE uses real coding of floating point numbers. The mutation operator here is addition instead of bit-wise flipping used in GA. And DE uses nonuniform crossover and tournament selection operators to create new solution strings. Among the DEs’ advantages are its simple structure, ease of use, speed and robustness. It can be used for optimizing functions with real variables and many local optima. In the recent years, DEs that are members of the stochastic optimization formalisms have been used with a great success in solving problems involving very large search spaces,\cite{8,9}

This paper demonstrates the successful application of DE to the practical optimization problem. As already stated, DE in principle is similar to GA. Therefore, as in GA we use a population of points in our search for the optimum. The population size is denoted by \( NP \). The dimension of each vector is denoted by \( D \). The main operation is the NP number of competitions that are to be carried out to decide the next generation.

The optimization objective underlying the DE-based optimization of an SVR model is defined as: find the \( N \)-dimensional optimal decision variable vector, \( \mathbf{x}^* = [x_1^*, x_2^*, \ldots, x_N^*]^T \), representing optimal process conditions such that it maximizes process outputs, \( y_k \); \( k = 1, 2, \ldots, K \). The corresponding single objective function \( J \) to be maximized by the DE is defined in Eqn. (1). In the DE procedure, the search for an optimal solution (decision) vector, \( \mathbf{x}^* \), begins from a randomly initialized population of probable (candidate) solutions. To start with, we have a population of NP vectors within the range of the objective function. We select one of these NP vectors as our \textit{target vector}. We then randomly select two vectors from the population and find the difference between them (vector subtraction). This difference is multiplied by a factor \( F \) (specified at the start) and added to third randomly selected vector. The result is called the \textit{noisy random vector}. Subsequently, crossover is performed between the target vector and noisy random vector to produce the \textit{trial vector}. Then, a competition between the trial vector and target vector is performed and the winner is replaced into the population. The same procedure is carried out NP times to decide the next generation of vectors. This sequence is continued till some convergence criterion is met. This summarizes the basic procedure carried out in DEs. The details of this procedure are in Appendix 1.

Figure 2. Flowchart for differential evolution based optimization of support vector regression model. This figure is available in colour online at www.apjChemEng.com.

The stepwise procedure for the DE-based optimization of an SVR model is provided in flowchart in Fig. 2.

CASE STUDY OF MEG PRODUCT UV TRANSMITTANCE

Figure 3 describes a brief process description of glycol section of MEG plant where glycol (90%) and water solution (10%) were fed to the drying column to remove the water from drying column top. The bottom of drying column was fed to MEG column to distill...
MEG from heavier glycols (namely diethylene glycol and triethylene glycol). MEG product (99.9% weight purity) is withdrawn from the MEG column below the top packing bed. An overhead vapor purge of up to 10% of the product is taken out to purge the light compounds.

Figure 3 shows the location of input parameters from drying column and MEG column which were used to build the model of UV.

Development of the SVR based correlation

The development of the SVR-based correlation had been started with the collection of a large databank. The next step was to perform a neural regression and to validate it statistically.

Collection of data

The quality and quantity of data is very crucial in SVR modeling as learning is primarily based on these data. Hourly average of actual plant operating data at steady state was collected for approximately one year. Data was checked and cleaned for obvious inaccuracy and retains those data when plant operation was in steady state and smooth. Finally, 6273 records are qualified for neural regression. This wide range of database includes plant operation data at various capacities starting from 75% capacity to 110% of design capacity.

Identification of input and output parameters

On the basis of the operating experience in glycol plant, all physical parameters that influence UV are put in a so-called ‘wish-list.’ Out of the number of inputs in ‘wish list’ several sets of inputs were made and tested through rigorous trial-and-error on the SVR. Finally, the nine input variables (Table 2) have been finalized to predict UV.

SVR

For modeling purposes, the column operating conditions data (Table 2) can be viewed as an example input matrix (X) of size (6273 × 9), and the corresponding UV data as the example output matrix (Y) of size (6273 × 1). For SVR training, each row of X represents a nine-dimensional input vector $x = [x_1, x_2, \ldots, x_9]^T$.

<table>
<thead>
<tr>
<th>Table 2. Input and output variable for support vector regression model.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input variables</strong></td>
</tr>
<tr>
<td>Reflux ratio (product flow/reflux flow)</td>
</tr>
<tr>
<td>Reflux flow (MT/h)</td>
</tr>
<tr>
<td>MEG column top pressure (mmHg)</td>
</tr>
<tr>
<td>MEG column condenser pressure (Barg)</td>
</tr>
<tr>
<td>MEG column control temperature (°C)</td>
</tr>
<tr>
<td>MEG column feed flow (MT/h)</td>
</tr>
<tr>
<td>Drying column control temperature (°C)</td>
</tr>
<tr>
<td>Drying column bottom temperature (°C)</td>
</tr>
<tr>
<td>Crude glycol reprocessing flow (MT/h)</td>
</tr>
</tbody>
</table>

© 2009 Curtin University of Technology and John Wiley & Sons, Ltd.
and the corresponding row of matrix $Y$ denotes the one-dimensional desired (target) output vector $y = [y_1]^T$. As the magnitude of inputs and outputs greatly differ from each other, they are normalized in $-1$ to $+1$ scale. About 80% of total dataset was chosen randomly for training and rest 20% was selected for validation and testing.

There are five different parameters to be evaluated to design a successful regression model. These five parameters are (1) kernel type, (2) type of loss function, (3) kernel parameter i.e. degree of polynomial, etc., (4) $C$ (represents the trade-off between the model-complexity and the approximation error) and (5) $\varepsilon$ (signifies the width of the $\varepsilon$-insensitive zone used to fit the training data.)

As the prior knowledge is not there regarding the suitability of particular value of any of the above five parameter, the strategy adopted here is holistic and summarized in Fig. 4. The SVR performance was evaluated exhaustively for all combinations of above parameters. All the kernel type available in literature (namely linear, polynomial, Gaussian radial basis function, exponential radial basis function (ERBF), Splines and B Splines) is tested with all combinations of loss function (namely $\varepsilon$- insensitive loss function, quadratic loss function).

The degree of kernel was varied from 1 to 6, capacity control varied from 10,000 to 0.1 (typically six values 10,000, 1000, 100, 10, 1 and 0.1) and $\varepsilon$ varies from 0 to 25 (typically six values 0, 0.1, 1, 10 and 25). Each run was exposed with same training and testing data and average absolute relative error (AARE) and sigma was calculated for each run.

The statistical analysis of ANN prediction is based on the following performance criteria:

1. The AARE on test data should be minimum

$$AARE = \frac{1}{N} \left[ \sum_{i=1}^{N} \left( \frac{|y_{pred(i)} - y_{exp(i)}|}{y_{exp(i)}} \right) \right]$$

2. The standard deviation of error ($\sigma$) on test data should be minimum

$$\sigma = \sqrt{\frac{1}{N-1} \left[ \sum_{i=1}^{N} \left( \frac{|y_{pred(i)} - y_{exp(i)}|}{y_{exp(i)}} \right)^2 - AARE \right]^2}$$

3. The cross-correlation coefficient ($R$) between input and output should be around unity.

$$R = \frac{\sum_{i=1}^{N} (y_{exp(i)} - \bar{y}_{exp})(y_{pred(i)} - \bar{y}_{pred})}{\sqrt{\sum_{i=1}^{N} (y_{exp(i)} - \bar{y}_{exp})^2 \sum_{i=1}^{N} (y_{pred(i)} - \bar{y}_{pred})^2}}$$

Where $y_{exp(i)}$ is the actual MEG product UV in plant for ith samples and $y_{pred(i)}$ is the predicted UV for the ith samples.

ANN learning is considered successful only if the system can perform well on test data on which the system has not been trained.

RESULTS AND DISCUSSIONS

SVR model developments for UV soft sensor

Although the training set was used for the iterative updation of the SVM parameters ($b$ and $w$), the test set was used for simultaneously monitoring the generalization ability of SVR model. For developing an optimal SVR model, its five structural parameter described...
above was varied systematically. For choosing an overall optimal model, the criterion used was least AARE for the test set. The optimal model that satisfied this criterion has ERBF, $\varepsilon$ insensitive loss function, width of ERBF function = 2, $C = 1000$, and $\varepsilon = 0.1$. The AARE for training and test set is calculated as 0.04 and 0.042% and corresponding $R$ calculated as 0.84 and 0.83, respectively. The low and comparable training and test error AARE values indicate good prediction and generalization ability of the trained SVR model. Good prediction and generalization performance of the model is also evident from the high and comparable $R$ values corresponding to both the outputs of training and test sets.

The final equation is as follows: MEG product UV = $K(i, j) \times Beta + Bias$

Where $K(i, j) = \text{kernel function} = e^{-\frac{((tstx(i, 9) - trnx(j, 9)) \times (tstx(i, 9) - trnx(j, 9)))^2}{2p_1^2}}$

for $i = 1 \text{ to } n$ and $j = 1 \text{ to } m$

Where, $n = \text{number of test record (supplied by user)}$ and $m = \text{number of training record (6273 in this case)}$.

$tstx(i, 9)$ is a $(n \times 9)$ test matrix for testing input (which user wants to test for calculating MEG product UV) and should be arranged in a sequence similar to Table 2, $trnx(j, 9)$ is a $(6273 \times 9)$ training matrix for training input data which is arranged in a sequence similar to Table 2, $beta(6273, 1)$ is a $(6273 \times 1)$ matrix calculated during training and bias is zero for this case and $p_1$ is width of rbf function (optimum value of $p_1$ is found 2 here).

The $trnx()$ and $beta()$ matrix are not reproduced here for sake of brevity and can be supplied upon request.

To validate the reliability of model, actual plant data were taken from DCS at different plant load at different point of time and actual laboratory measured UV was compared with the model predicted UV.

Figure 5 depicts a comparison of the outputs as predicted by the SVR model and their target values. Considering the fact that all the input-output data are from real plant with their inherent noise, the very low prediction error can be considered as an excellent SVR model. Once developed, this SVR model can be used to quantitatively predict the effects of all input parameters on the MEG product UV transmittance.

**DE-based optimization of the SVR model**

After development of successful SVR model of glycol column, next step is to find out the best set of operating conditions, which lead to maximum UV. DE-based hybrid model was run and optimum parameters were evaluated (within their permissible operating limit). Figure 6 depicts the actual versus the optimum UV. From Fig. 6 it is clear that by making a small change in the nine input parameters, the 1–2% rise in UV can be made. Refer Table 3 for optimum value of input variables calculated by DE algorithm. Drying column control temperature was found to have a significant effect on MEG product UV. This temperature will help to strip out UV deteriorating compounds from drying column itself before they enter to MEG column. Three cases were run with three different limits of this temperature. Optimum value is shown in Table 3. The program was made online where it gives the operator what should be the nine input parameters at different time to maximize the UV in real-time basis.

After verifying all the calculations, the optimum input parameters were maintained in actual plant and benefit was found exactly same as calculated. This ensures the validation and accuracy of this calculation.

**CONCLUSION**

This paper introduces SVR into soft sensing modeling and proposes a new soft sensing modeling method based on SVR. In the strategy, a soft sensor model is developed using SVR method after which the input space of that model is optimized using DEs such that the process performance is maximized. The major
Table 3. Optimum value of input variables calculated by differential evolution algorithm.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Case 1</th>
<th>Case 2</th>
<th>Case 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Minimum value</td>
<td>Optimum value</td>
<td>Maximum value</td>
</tr>
<tr>
<td>Input Reflux ratio (product flow/reflux flow)</td>
<td>0.70</td>
<td>0.78</td>
<td>0.78</td>
</tr>
<tr>
<td>Reflux flow (MT/h)</td>
<td>110.0</td>
<td>114.52</td>
<td>115.0</td>
</tr>
<tr>
<td>MEG column top pressure (mmHg)</td>
<td>92.0</td>
<td>94.75</td>
<td>97.0</td>
</tr>
<tr>
<td>MEG column condenser pressure (Bar)</td>
<td>1.60</td>
<td>1.66</td>
<td>1.67</td>
</tr>
<tr>
<td>MEG column control temperature (°C)</td>
<td>168.80</td>
<td>169.0</td>
<td>169.0</td>
</tr>
<tr>
<td>MEG column feed flow (MT/h)</td>
<td>98.0</td>
<td>98.0</td>
<td>98.0</td>
</tr>
<tr>
<td>Drying column control temperature (°C)</td>
<td>85.0</td>
<td>90.92</td>
<td>100.0</td>
</tr>
<tr>
<td>Drying column bottom temperature (°C)</td>
<td>165.0</td>
<td>166.0</td>
<td>166.0</td>
</tr>
<tr>
<td>Crude glycol reprocessing flow (output)</td>
<td>0.0</td>
<td>0.0</td>
<td>8.0</td>
</tr>
<tr>
<td>Output Optimized ultraviolet</td>
<td>97.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actual ultraviolet</td>
<td>96.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
advantage of the SVR-DE strategy is that modeling and optimization can be conducted exclusively from the historic process data wherein the detailed knowledge of process phenomenology (reaction mechanism, kinetics, etc.) is not required. Effective results indicate that SVR modeling method provides a new tool for soft sensing modeling and has promising application in industrial process applications. Using SVR-DE strategy, a number of sets of optimized operating conditions leading to maximized product UV was obtained. The optimized solutions when verified in actual plant resulted in a significant improvement in the product UV.

APPENDIX

Steps performed in DE

Assume that the objective function is of D dimensions and that it has to be optimized. The weighting constants F and the crossover constant (CR) is specified.

Step 1. Generate NP random vectors as the initial population: generate \((NP \times D)\) random numbers and linearize the range between 0 and 1 to cover the entire range of the function. From these \((NP \times D)\) numbers, generate NP random vectors, each of dimension \(D\), by mapping the random numbers over the range of the function.

Step 2. Choose a target vector from the population of size NP: first generate a random number between 0 and 1. From the value of the random number decide which population member is to be selected as the target vector \(X_i\) (a linear mapping rule can be used).

Step 3. Choose two vectors at random from the population and find the weighted difference: generate two random numbers. Decide which two population members are to be selected \(X_a, X_b\). Find the vector difference between the two vectors \((X_a−X_b)\). Multiply this difference by \(F\) to obtain the weighted difference.

Weighted difference = \(F \times (X_a−X_b)\)

Step 4. Find the noisy random vector: generate a random number. Choose a third random vector from the population \(X_c\). Add this vector to the weighted difference to obtain the noisy random vector \(X'c\).

Step 5. Perform crossover between \(X_i\) and \(X'c\) to find \(X_t\) the trial vector: generate \(D\) random numbers. For each of the \(D\) dimensions, if the random number is greater than CR, copy the value from \(X_i\) into the trial vector; if the random number is less than CR, copy the value from \(X'c\) into the trial vector.

Step 6. Calculate the cost of the trial vector and the target vector: for a minimization problem, calculate the function value directly and this is the cost. For a maximization problem, transform the objective function \(f(x)\) using the rule \(F(x) = 1/[1 + f(x)]\) and calculate the value of the cost. Alternatively, directly calculate the value of \(f(x)\) and this yields the profit. In case cost is calculated, the vector that yields the lesser cost replaces the population member in the initial population. In case profit is calculated, the vector with the greater profit replaces the population member in the initial population.

Step 1–6 are continued till some stopping criterion is met. This may be of two kinds. One may be some convergence criterion that states that the error in the minimum or maximum between two previous generations should be less than some specified value. The other may be an upper bound on the number of generations. The stopping criterion may be a combination of the two. Either way, once the stopping criterion is met, the computations are terminated.

Choosing DE key parameters NP, F and CR is seldom difficult and some general guidelines are available. Normally, NP ought to be about 5–10 times the number of parameters in a vector. As for \(F\), it lies in the range of 0.4–1.0. Initially \(F\) = 0.5 can be tried then \(F\) and/or NP is increased if the population converges prematurely. A good first choice for CR is 0.1, but in general CR should be as large a possible (Price and Storn, 1997).

REFERENCES

QUERIES TO BE ANSWERED BY AUTHOR

IMPORTANT NOTE: Please mark your corrections and answers to these queries directly onto the proof at the relevant place. Do NOT mark your corrections on this query sheet.

Queries from the Copyeditor:
AQ1 We have provided ‘Soft sensor modelling and optimization’ as the ‘Running head’ for this article. It will appear on the top of the third page and all the recto pages of the article. Please confirm if this is appropriate.
AQ2 Please confirm whether the changes made in the affiliation and correspondence address are ok.
AQ3 Please confirm whether the expansion of ‘QC’ as ‘quality control’ is ok. Else please provide the same.
AQ4 Please confirm whether the expansion of ‘PSE’ as ‘Process System Engineering’ is ok. Else please provide for the same.
AQ5 Please provide the expansion of ‘DCS.’
AQ6 Please confirm whether the changes made in the sentence ‘For a big world-class capacity…’ are ok.
AQ7 Please confirm whether the expansion of ‘ANN’ as ‘artificial neural networks’ is ok. Else provide the same.
AQ8 Figure(s) F1 to F4 is low resolution. Kindly provide better replacement figure(s) or confirm if we may use the existing one(s).
AQ9 There is no appropriate citation of equation and even it is not numbered in the text. Please clarify.
AQ10 Please note that the Equations are not numbered throughout. So please specify to which Equation does this 2.33 refer to.
AQ11 Please provide the expansion of ‘NP.’
AQ12 Please confirm whether the ‘wt’ here refers to ‘weight.’ Else please provide the significance of the same.
AQ13 Please provide the expansion of ‘MT’ is that refers to ‘meter.’
AQ14 Please check there is reference to 6 values but there are only 5 values provided in the sentence ‘0 to 25 (typically six values…’
AQ15 Please provide the report number and publisher for reference 2.
AQ16 Please provide the volume number for references 7 and 9.