Inferential control strategies using neural soft sensor in a high purity distillation column

Estratégias de controle inferencial usando sensor macio neural em uma coluna de destilação de alta pureza

Estrategias de control inferencial utilizando un sensor neural blando en una columna de destilación de alta pureza

DOI: 10.54033/cadpedv21n5-160

Originals received: 04/19/2024
Acceptance for publication: 05/10/2024

Arioston Araújo de Morais Júnior
Doctor in Chemical Engineering
Institution: Universidade Federal da Paraíba (UFPB) - campus I
Address: João Pessoa, Paraíba, Brazil
E-mail: aamj@ct.ufpb.br

José Carlos Finiz Filho
Bachelor’s Degree in Chemical Engineering
Institution: Universidade Federal da Paraíba (UFPB) - campus I
Address: João Pessoa, Paraíba, Brazil
E-mail: carlosdeniz14@gmail.com

Leopoldo Oswaldo Alcazar Rojas
Doctor in Chemical Engineering
Institution: Universidade Federal do Rio Grande do Norte, Universidade Federal da Paraíba (UFPB) - campus I
Address: João Pessoa, Paraíba, Brazil
E-mail: leopoldo@ct.ufpb.br

Romildo Pereira Brito
Doctor in Chemical Engineering
Institution: Universidade Estadual de Campinas, Universidade Federal da Paraíba (UFPB) - campus I
Address: João Pessoa, Paraíba, Brazil
E-mail: brito.romildo@outlook.com

ABSTRACT

High-purity distillation columns are processes in which you want to minimally increase the purity of a key component by separation. These processes are
sensitive to disturbances, where small changes in the feed flowrate stream cause drastic changes in product compositions. Furthermore, when one wants to apply traditional control and optimization techniques to these processes, some difficulties have to be faced: the process is generally non-linear and has a long response time, there are many immeasurable disturbances, it is difficult to keep the process in steady-state, and the bottom and top compositions are highly coupled. Therefore, this paper presents a methodology for the development of soft sensors (SS), here in applied to an industrial 1,2-Dichloroethane separation plant. The process was simulated and validated with real data from the industrial plant. In the step second an algorithm based on multivariate statistics was developed for the selection of inputs SS variables. The configuration for training multilayer artificial neural network (ANN's) was optimize, accounting for the ANN's prediction capacity of responses and the rejection of disturbances inserted in the process. Thereafter, a temperature control was implemented to maintain the impurity compositions within specifications but failed due to the small temperature variations in the high purity column. To overcome these difficulty three inferential control strategies (impurity composition estimated by the SS) were implemented: a classic feedback control, a cascade control, and a ratio-cascade control. All control strategies acted to minimize the process disturbances effects. However, the inferential ratio-cascade control shown to be the most robust, because of the lowest integral error criteria value and percentage overshoot.

**Keywords**: High Purity Column. Soft Sensor. Artificial Neural Network. Inferential Control.

**RESUMO**
Colunas de destilação de alta pureza são processos os quais se deseja aumentar minimamente a pureza de um componente chave por separação. Esses processos são sensíveis às perturbações, onde pequenas alterações na corrente de alimentação provocam mudanças drásticas nas composições dos produtos. Além disso, quando se deseja aplicar técnicas de controle e otimização tradicionais nesses processos, algumas dificuldades têm de ser enfrentadas: o processo é geralmente não linear e com tempo de resposta longo, há muitos distúrbios imensuráveis, é difícil manter o processo em estado estacionário, e as composições de base e do topo são altamente acopladas. Sendo assim, este artigo apresenta uma metodologia para o desenvolvimento de sensores virtuais (SV), aplicados a uma planta industrial de separação de 1,2-Dicloroetano, que foi simulada e validada com dados reais desta planta. Na segunda etapa foi desenvolvido um algoritmo baseado em estatística multivariada para a seleção das variáveis de entrada do SV. A configuração para treinamento de redes neurais artificiais multicamadas (RNA) foi otimizada, levando em consideração a sua capacidade de predição de respostas e a rejeição de perturbações inseridos no processo. Posteriormente, um controle de temperatura foi implementado para manter as composições de impurezas dentro das especificações, mas falhou devido às pequenas variações de temperatura na coluna de alta pureza. Para superar esta dificuldade foram implementadas três estratégias de controle inferencial (composição de impurezas estimada pelo SS): um controle clássico
realimentação, um controle em cascata e um controle em cascata-razão. Todas as estratégias de controle atuaram no sentido de minimizar os efeitos dos distúrbios do processo. No entanto, o controle inferencial razão-cascata mostrou-se o mais robusto, devido ao menor valor dos critérios de erro integral e à sobre-elevação percentual.


**RESUMEN**

Las columnas de destilación de alta pureza son procesos en los que se desea aumentar mínimamente la pureza de un componente clave mediante separación. Estos procesos son sensibles a las perturbaciones, donde pequeños cambios en el caudal de alimentación provocan cambios drásticos en la composición del producto. Además, cuando se quieren aplicar técnicas tradicionales de control y optimización a estos procesos, hay que enfrentar algunas dificultades: el proceso generalmente no es lineal y tiene un tiempo de respuesta largo, hay muchas perturbaciones incommensurables, es difícil mantener el proceso en estado estacionario, y las composiciones inferior y superior están altamente acopladas. Por lo tanto, este artículo presenta una metodología para el desarrollo de soft sensores (SS), aquí aplicados a una planta industrial de separación de 1,2-dicloroetano. El proceso fue simulado y validado con datos reales de la planta industrial. En el segundo paso se desarrolló un algoritmo basado en estadística multivariada para la selección de variables de entrada SS. Se optimizó la configuración para el entrenamiento de redes neuronales artificiales (RNA) multicapa, teniendo en cuenta la capacidad de predicción de respuestas de las RNA y el rechazo de perturbaciones insertadas en el proceso. Posteriormente, se implementó un control de temperatura para mantener las composiciones de impurezas dentro de las especificaciones, pero falló debido a las pequeñas variaciones de temperatura en la columna de alta pureza. Para superar estas dificultades se implementaron tres estrategias de control inferencial (composición de impurezas estimada por el SS): un control de retroalimentación clásico, un control en cascada y un control en cascada de relación. Todas las estrategias de control actuaron para minimizar los efectos de las perturbaciones del proceso. Sin embargo, el control inferencial en cascada de relación demostró ser el más robustas, debido al valor de criterio de error integral más bajo y al porcentaje de sobre-paso.

**Palabras clave:** Columna de Alta Pureza. Soft Sensors. Redes Neurales Artificiales. Control Inferencial.
1 INTRODUCTION

Most distillation columns are designed to achieve a specific separation between the two main components, the heavy component impurity at the top and the light component at the bottom. Thus, the operation and control of a tower would measure the compositions of the two products and manipulate two inputs to keep them within a desired value. However, rare processes employ this structure, as the compositions cannot be measured directly and when measured, it is necessary to use online analyzers, which are expensive, have dead time and are unfeasible for some applications (Zambrogna et al., 2015; Luyben, 2013).

When monitoring composition in an industrial process, samples are customary collected and sent to external labs. Depending on the lab results, the process engineering team will tune the process variables in order to return it to the setup state, but often with a remarkable delay. This is the case for distillation (for example, binary, complex, azeotropic, and high purity) operations, where the outputs, top and bottom compositions, are monitored. Online analyzers can be an alternative but have high capital and operating costs, and sometimes with an unacceptable delay (Campos et al., 2013; Fortuna et al., 2005 and Kano et al., 2009). The developments in process automation and control, acquiring and processing huge amounts of data, led to an accepted way of estimating process variables that are difficult to measure. The algorithm, known as soft sensor, is based on a mathematical model where the output variables ($\hat{y}$) are functions of secondary variables ($u_1, u_2, u_3, ..., u_n$), which are measurable (Morais Jr, 2015).

Soft sensor building techniques are already known and the usage selection depends on the process nature and the application model. The models may be based on first-principles models (FPM), data-based models, and hybrid techniques. Although FPM algorithms generally present a range of validation, a complete dynamic process model may become a complex task. Thus, in these cases hybrid models may be used such as the extended Kalman Filter or state observer (Kadlec et al., 2009; Morais Jr et al., 2019). In complex systems or when the deterministic model is difficult to obtain, Artificial Neural Networks (ANNs) have proven to be a robust tool for implementing soft sensors. Different ANN
models are reported in the literature; one extensively used due to its versatile architecture is the Multi-Layer Perceptron (MLP). An ensemble approach for Soft Sensor development based on MLP was published in Kadlec & Gabrys (2008). This Soft Sensor was applied to an industrial drier process. In this work, the optimal MLP topology was established by training several models with different complexities and assessing their relative performance. Error Back Propagation (EBP) was an important advance in ANNs, but the algorithm slows down convergence. Thus, much work was done to accelerate the EBP algorithm. One of the developed algorithms was the Levenberg-Marquart (LMA) algorithm which captures the high velocity of the Newton algorithm and the stability of the Gradient Descent Methods (Wangdong et al., 2014; Rumelhart et al., 2002; Battiti, 1992), and is used for ANN training and convergence.

The soft sensor technology appears as a robust and cheap alternative to the application of inferential control of compositions. Specifically in high-purity distillation columns, there is great difficulty in controlling the top or bottom compositions. High-purity distillation columns are processes in which it is desired to minimally increase the purity of a key component by separation (Luyben, 2013). These processes are sensitive to disturbances, where small changes in the feed flowrate stream cause drastic changes in product compositions. Furthermore, when one wants to apply traditional control and optimization techniques to these processes, some difficulties have to be faced: the process is generally non-linear and has a long response time, there are many immeasurable disturbances, it is difficult to keep the process in steady-state, and the bottom and top compositions are highly coupled (Kano et al. 2009; Wenxiang et al., 2010).

Therefore, the aim of this work is to analyze the dynamic behavior and control impurities in a high purity column of 1-2-dichloroethane (EDC). The column is part of the vinyl chloride monomer (VCM) process, where impurities are to be removed before entering the thermal cracker reactor, where the VCM is obtained. Controlling the impurities output of the column is critical and their monitoring is very difficult (only by sampling and external analysis). Thus, monitoring and control through soft sensors is proposed to overcome this difficulty. Comparing with other works available in the literature (Kano et al. 2009,
Udugama et al. 2019, Hsiao et al. 2021, Kalbani & Zhang, 2023), here the soft sensors are developed with artificial neural networks (data-driven modeling) who used available data from a robust nonlinear model that was developed in the Aspen Plus and Aspen DynamicsTM simulators, validated with real process data. Finally, with estimation through soft sensors, inferential control strategies were implemented: a classic feedback control, a cascade control, and a ratio-cascade control.

2 PROCESS DESCRIPTION AND PROBLEM DEFINITION

The global PVC Resins market was valued at USD 52,241.26 million in 2022 and is expected to reach USD 56,665.83 million by the end of 2029 (G II-Global information, 2023). The PVC is part of the chlorine (Cl₂) industry, obtained by electrolysis of a brine, that’s reacts with ethylene (C₂H₄) to produce 1,2-dichloroethane (DCE), and DCE is also formed by oxychlorination of C₂H₄, i.e. reactions of the (Equations 1 and 2), respectively.

\[ \text{C}_2\text{H}_4 + \text{Cl}_2 \rightarrow \text{C}_2\text{H}_4\text{Cl} \quad \text{Eq. 1} \]

\[ 2\text{C}_2\text{H}_4 + 4\text{HCl} + \text{O}_2 \rightarrow 2\text{C}_2\text{H}_4\text{Cl} + 2\text{H}_2\text{O} \quad \text{Eq. 2} \]

The formed EDC is thermally cracked into vinyl chloride monomer (VCM) but must be previously purified to remove undesirable products like acetylene and other by-products of the C₂H₄ and hydrogen chloride reaction, in addition to ethane with Cl₂. The EDC loading into the VCM reactor must be without humidity to prevent corrosion and present high purity (above 99.5% molar), avoiding contaminants that can partially inhibit the reaction and promote coke formation on the tube walls of the furnace, shortening production time of the reactor (Dimian and Sorin, 2008). These problems were studied by Oliveira et al. (2017) and Nyeng (2015). They reported that the extended kinetic model is very sensitive to carbon tetrachloride (CCl₄) content and CHCl₃ is a coke and corrosion promoter. Thus, the EDC must be purified before entering the VCM reactor. This is done in
three distillation columns to remove humidity and chlorinated organic compounds (COCs) of low boiling points, as shown in the process flow diagram (Figure 2). The first column (C1) is a high purity azeotropic distillation column, water and other light products are recovered at the top distillate stream (D1), and dried EDC with heavy products at the bottom stream (B1). C1 is fed rates by two streams: F1, which comes from the direct chlorination reaction and feed the column in tray 32 (numbered from top to bottom); while F2 is the output of the oxychlorination reactor and feed the column tray 54. Both feed flow rates present high mass fraction of EDC (98.5-99.3 %) with low humidity and about 17 other COCs. Water leaves the column at the top because of the inversion of the relative volatility; this effect is mainly seen above the F1 feed. The decanter drum is a three-phase separator: gaseous (non-condensable) phase and two liquid immiscible phases. The aqueous phase stream is discharged and the organic phase splits into a reflux stream (R1), a distillate stream (D1), and the non-condensable gases are vented and flared.

The objective of the second column (C2) is to produce high purity EDC in the distillate (D2), free of heavy components. This column has three feed streams, one is the bottom stream of C1 (B1), the second comes from the distillate of the third column (D3), and lastly a recycle stream F3 which recovers the unreacted EDC. The third column (C3) operates under a vacuum and its purpose is to eliminate the dried chlorine-hydrocarbons which leave from the bottom of C3 (B3).
The third column is fed with the bottom stream of C2 (B2) and an EDC recovery stream (F4), coming from a purification unit of VCM. The CCl4 and CHCl3 contents in the first column (C1) are monitored in B1 stream so as to keep their compositions below 400 and 3000 ppm, respectively. This composition control procedure is complex and extremely critical for CHCl3. The compositions are determined in an external lab where samples (one to three samples per day) are delivered. So, the proposal is to build up a soft sensor and put an inferential control in B1 in order to minimize the delays and offsets in the transient processes.

3 BUILDING THE SOFT SENSOR

The soft sensors are virtual devices which estimate output variables through an algorithm based on treating input variables (Figure 1). We can easily encounter thousands of variables in an industrial process, so deciding how many and selecting which variables are to be used is an extremely important task. Sometimes the practical expertise in the chemical process or in theoretical model considerations can help to specify the set of secondary variables to be used. In practice, rejecting specific variables is a usual approach in order to end up with just a few input variables and a small and concise model. On the other hand, increasing many of the secondary variables will unnecessarily make the model more rigid. In fact, the inappropriate selection of estimator inputs may lead to numerical problems, such as singularity and over-parameterization, or may markedly reduce the estimation accuracy (Kano et al., 2000; Zamprogna et al., 2005). In this section we will describe the stage of selection of variables and what was the structure of ANNs defined for the construction of soft sensors.
3.1 VARIABLE SELECTION

For the selection of secondary variables, the use of multivariate statistical techniques is commonplace, with aid in prediction by Multiple Linear Regression (MLR), All-Possible Regressions (APR), Forward Selection (FS), Backward Elimination (BE), Stepwise Regression (SR) and Principal Component Analysis (PCA). Ming-Da Ma et al. (2009) used stepwise regression, while Zamprogna et al. (2005) used PCA in selecting input variables of soft sensors. A more detailed work on variable selection was proposed by Wang et al. (2015), where seven variable selection methods were compared: stepwise regression (SR), partial least squares with regression coefficients (PLS), genetic algorithm with PLS, least absolute shrinkage and selection operator (Lasso), and competitive adaptive Reweighted sampling with PLS (CARS-PLS).

This manuscript proposes to select variables based on statistical analysis because this selection process cannot be only made based on theoretical bases. All-Possible Regressions (APR) technique is proposed to be used when the number of candidates regressors is not too large. This approach has the advantage that the “best” regression equation with respect to certain criteria can be attained when compared to other stepwise-type methods (e.g., FS, BE and SR) which offer no such assurance. Furthermore, the APR procedure is not
distorted by dependencies among the regressors as stepwise-type methods are (Ming-Da Ma et al., 2009; Montgomery & Runger, 2014). The algorithm was implemented in Matlab® and the main steps are described as follows:

- **Step 1.** Enter the $Z_{mxn}$ matrix data and $Y_{mx1}$ vector; where $n$ number of secondary variables (regressors) and $m$ the number of samples.

- **Step 2.** Calculate the square matrix $W_{nxn}$ (Equations 3).

$$W = Z^T Z \quad \text{Eq. 3}$$

- **Step 3.** Calculate the inverse of matrix $W$, $W^{-1}$.

- **Step 4.** Calculate the vector $Q_{nx1}$, $Q = Z^T Y$ (Equations 4).

$$Q = [\sum_{i=1}^{m} y_i \quad \sum_{i=1}^{m} z_{i1} y_i \quad \sum_{i=1}^{m} z_{i2} y_i \quad \cdots \quad \sum_{i=1}^{m} z_{iN} y_i]^T. \quad \text{Eq. 4}$$

- **Step 5.** Find the vector of coefficients ($\beta_{nx1}$) (Equations 5).

$$\beta = W^{-1}.Q = [\beta_0, \beta_1, \beta_2, \ldots, \beta_N]^T. \quad \text{Eq. 5}$$

- **Step 6.** Calculate the combinations ($C_{RN,CN}$) for a $RN$ collection of regressors taken $CN$ elements; where the collection varies as $NR=1, 2, \ldots, n$ and the number of elements taken $CN= 1, 2, \ldots, p \ (p \leq RN)$ (Equation 6).

$$C_{RN,CN} = \frac{RN!}{CN!(RN-CN)!} \quad \text{Eq. 6}$$

- **Step 7.** All the possible regression models ($APR$), $APR = 2^n - 1$, are obtain by summing every $C_{RN,CN}$ (Equation 7).

$$APR = \sum_{RN=1}^{n} \sum_{CN=1}^{p=RN} C_{RN,CN} \quad \text{Eq. 7}$$
Step 8. Calculate the \(i\)-th estimated output (\(\hat{y}_i\)) for each model, with the following multiple linear regression (Equation 8), where \(\beta_n\) is calculated by (Equation 5).

\[
\hat{y}_i = [\beta_0 \ \beta_1 \ \beta_2 \ \cdots \ \beta_n] \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}
\]

Step 9. Implement the mean square error (MSE) (Equation 9) and the Adjusted R-Squared (\(R^2_{adj}\)) (Equation 10) criteria for selecting the models to be used by the neural soft sensors.

\[
\text{MSE} = \frac{SS_E}{(n - p)} \quad \text{Eq. 9}
\]

\[
R^2_{adj} = 1 - \frac{SS_E(n - 1)}{SS_T(n - p)} \quad \text{Eq. 10}
\]

Where \(SS_E\) is the sum of squares error, \(SS_R\) the sum of squares regression and \(SS_T\) is the total corrected sum of squares of \(y\), (Equation 11, 12, 13 and 14, respectively).

\[
SS_E = \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \quad \text{Eq. 11}
\]

\[
SS_R = \sum_{i=1}^{N} (\hat{y}_i - \bar{y})^2 \quad \text{Eq. 12}
\]

\[
SS_T = \sum_{i=1}^{N} (y_i - \bar{y})^2 \quad \text{Eq. 13}
\]

\[
SS_T = SS_R + SS_E \quad \text{Eq. 14}
\]
3.2 LOCATION OF SENSITIVE-TRAYS

The control structure in distillation columns is usually single-end or dual-composition based. Single-end control scheme is frequently used due to its simplicity and efficiency, where a single composition or temperature is controlled. The temperature stages of a column are related to composition and are simple to measure with low cost and confidence, and almost instantaneous response time (Mejdell & Skogestad, 1993; Rani et al., 2013). However, high purity columns present low temperature changes along the stages, making it difficult to identify the most composition sensitive stage for inference purposes. In a relevant work about soft sensors Kano et al. (2009) used data from fifteen variables, including ethane composition in the feed rate to estimate the ethane composition at the exit of an industrial ethylene fractionator. Real time concentration data, as well the composition feed rate, are scarcely available.

The proposal in this work is to use variables that are easy to measure as temperatures and flow rates. Column C1 presents 72 stages, including the condenser and reboiler, with only three temperature sensors: the reflux drum ($T_0$), the first tray ($T_1$) and the bottom tray ($T_{71}$). Further important information referent to the contaminants composition in the bottom may be retrieved from the temperatures of the rest of the column trays. Among the available selection methods, the Singular Value Decomposition (SVD) by Moore (1992) is the most frequently used. The SVD method is based on a linear algebra theorem that states that a rectangular matrix, $K$, can be expressed by the product of an orthogonal matrix, $U$, a diagonal matrix, $S$, and the transposition of an orthogonal matrix, $V$. The theorem is stated in (Equation 15).

$$K_{mn} = U_{mm}S_{mn}V_{nn}^T$$  \hspace{1cm} \text{Eq. 15}
The columns of $U$ are the orthonormal eigenvectors of $KK^T$, the columns of $V$ are orthonormal eigenvectors of $K^TK$, and $S$ is a diagonal matrix containing the square roots of eigenvalues of $U$ or $V$ in decreasing order. The gain matrix $K$ stores the changes in two manipulated variables (2 columns) due to changes in $m$ tray temperatures ($m$ rows). The $K$ matrix in this work was decomposed using Matlab® software. The diagonal entries in $S$ are the singular values ($\sigma$) of $K$, the columns in $U$ are called left singular vectors, and the columns in $V$ are called right singular vectors. One important aspect of the physical significance of the singular value analysis can be seen in the Condition Number (CN). This number is the ratio of the largest to the smallest, non-zero, singular value and is used in numerical computation to gauge the "condition" of a set of (Equation 16).

$$CN = \frac{\sigma_{Max}}{\sigma_{Min}}$$  
Eq. 16

3.3 MULTILAYER PERCEPTRON (MLP) NETWORKS

The MLP is a neural network made up of one or more hidden internal layer(s) placed between the input and output layers; therefore, the MLP network has at least two layers and their neurons are distributed among the internal and external layers (Haykin, 2009; Silva et al., 2017). Backpropagation is the workhorse algorithm of learning in neural networks by doing successive applications of the forward and backward propagation. As shown in (Figure 3A) where the MLP has two hidden layers with $n$ signals; $h_1$ and $h_2$ are the neurons of the first and second layer, respectively, and $m$ is the signal of the output (third) neural layer. In (Figure 3 A and B) each network input ($x_1$, $x_2$, $x_3$, ..., $x_n$), the secondary variables of the soft sensor, will forward propagate through the layers toward the output layer. In this way, the neuron outputs of the first neural layer will be the neuronal inputs of the second hidden layer, and so on until the last neural exit layer. The resulting inference ANN is compared with real (desired) data and outgoing departure errors. A backward propagation scheme is then applied based on these errors, tuning the synaptic weights of every neuron to
minimize the derivations (Silva et al., 2017; Goldberg, 2017; Haykin, 2009). The neurons weights and thresholds are interactively tuned successively between the forward and backward phases, thus minimizing the sum of the deviation errors.

Some auxiliary variables and parameters must be defined to better understand the backpropagation algorithm. The variables for MLP topology and the configuration for the artificial neuron, in which $W_{ji}^{(L)}$ represents the weight matrix for the $L^{th}$ layer with ($j$) neurons (Figure 3A and 3B, respectively). Each matrix element denotes the synaptic weight that connects the $j^{th}$ neuron of the $L^{th}$ layer with the $i^{th}$ neuron of the $(L-1)^{th}$ layer. Two MLPs were tested with one and two hidden layers. The vector $I_j^{(L)}$ elements represent the weighted inputs of the $j^{th}$ neuron in the $L^{th}$ layer (Equation 17).

$$I_j^{(L)} = \sum_{i=0}^{p} W_{ji}^{(L)} \cdot y_i^{(L-1)} \quad \text{Eq. 17}$$

Where: $\mathbf{v}$ (can be $n$), $h_1$, $h_2$ and $y_j^{(L)}$ are vectors whose elements represent the output of the $i^{th}$ neuron in the $L^{th}$ layer determined (Equation 18)
The activation functions used in this manuscript were sigmoid logarithm functions and hyperbolic tangent (Equations 19 and 20, respectively), described in the Matlab® software as `logsig` and `tansig`, respectively. The constant $\alpha$ is associated to the slope of the functions in relation to the inflection point.

$$f(u) = \frac{1}{1 + e^{-\alpha u}} \quad \text{Eq. 19}$$

$$f(u) = \frac{1 - e^{-\alpha u}}{1 + e^{-\alpha u}} \quad \text{Eq. 20}$$

The backpropagation algorithm is an efficient algorithm, but slows down the convergence, which is why alternative methods were developed to speed up the convergence. The Levenberg-Marquardt algorithm (LMA) represents a bridge between the Newton’s and the Gradient Descent methods. The first provides a very high convergence speed due to the quadratic properties but may arrive at a local instead of global minimum. Therefore, the LMA proves to be a robust algorithm capable of ANN training (Demuth & Beale, 2002; Kermani et al., 2005).

For the LMA, the quadratic error and the mean quadratic error can (Equation 21).

$$V = \frac{1}{2p} = \sum_{k=1}^{p} E(k).E(k) \quad \text{Eq. 21}$$

where $E(k)=[DV_j(k)-y(k)^{(3)}]$ and $E=[E(1) E(2),...,E(p)]$ is the error vector of the $p$ samples used for training; and where $y(k)^{(3)}$ is the value produced by the $j$-th output neuron of the network for the $p$-th training sample, while $DV_j(k)$ is the corresponding desired value ($DV$). The recurrence expression for the LMA, tuning the weight matrix ($W^{(L)}$) (Equation 22).
\[ \Delta w = w_{K+1} - w_K = [\nabla^2 V(W) + \mu I]^{-1} \nabla V(W) \]  

Eq. 22

where \( \nabla^2 V(W) \) and \( \nabla V(W) \) are the Hessian and Jacobian matrixes, respectively; \( I \) is the identity matrix, with same dimension as the Hessian; \( \mu \) is a tuning (or loading) parameter whose value guarantees that the matrix \([\nabla^2 V(W) + \mu I]\) is always positive definite and well-posed over all the computational process. The tuning parameter \( \mu \) has an important role in the LMA. The LMA is reduced to the Newton’s method for a low \( \mu \), and conversely to the Gradient Descent Algorithm for high \( \mu \) values (Kermani et al., 2005; Haykin, 2009), where \( \mu \) value was fixed at \( 10^{-3} \).

Some difficulties can occur in the ANN training process with the LMA or other algorithms when trying to determine the endpoint (a global minimum). This is since that error falls rapidly at the beginning and slows progressively to very slow change rates, going toward a local minimum on the error surface. Thus, some criteria must be followed in order to generalize the procedure for stopping the training (Taylor et al., 2006; Silva et al., 2017). Therefore, the early stopping method with cross-validation was employed in this work. The method stops the training every \( C \) cycle and an estimation of the error is performed with the used data. The process is finally ended when an increase in the error is detected. The study by Coulibaly et al. (2000) used of the early stopping training with LMA in a real-time application concluded that its use reduced the training time of ANNs by four times.

The total data in this paper was distributed into three groups: the training group (70%), used for parameter determination (ANN weights and bias); test group (15%), used to verify the stop training criteria efficiency and future performance; and the third group for validation purposes. The stop training was triggered when the criteria of six successive increments of the validation error \( (E_{\text{val}}) \) were attained (Equation 23).
\[ E_{\text{val}} = \sum_{i=1}^{p} \sum_{t=1}^{o} (y_i - \hat{y}_t)^2 \]  

Eq. 23

Where

\( y_i \) and \( \hat{y}_t \) are the real and estimated outputs, respectively;
\( p \) number of samples;
\( O \) is the number nodes in the ANN exit layer, meaning the number of output variables.

The real output variables \( (y_i) \) were supplied by performing Aspen Dynamics™ simulations and validated by real plant data. Some noise was inserted at this point, which can be treated as uncertainties in observations or errors in the measurement. Two estimation errors, the mean squared error (MSE) and the roots mean squared error (RMSE), were calculated (Equations 24 and 25, respectively).

\[ MSE = \sum_{i=1}^{N} \frac{1}{N} (y_i - \hat{y}_i)^2 \]  

Eq. 24

\[ RMSE = \sqrt{\sum_{i=1}^{N} \frac{1}{N} (y_i - \hat{y}_i)^2} \]  

Eq. 25

Where

\( N \) is the total number of samples.

4 PROCESS SIMULATION

For processing purposes, the bottom product composition of column C1 should be monitored so the liquid mass fraction of EDC \( (x_{\text{EDC}}) \), water \( (x_{ \text{H2O}}) \), CHCl\(_3\) \( (x_{\text{CHCl3}}) \), and CCl\(_4\) \( (x_{\text{CCl4}}) \) must be known. The process was simulated in the steady and transient states using Aspen Plus and Aspen Plus Dynamics™,
respectively. The Aspen Plus™ steady-state simulations were parameter-optimized so that the simulated results were in accordance with the lab data composition of \(x_{\text{EDC}}, x_{\text{CHCl}_3}\) and \(x_{\text{CCl}_4}\), for the top and bottom column products. The flowsheet diagram under study in this manuscript (Figure 4).

![Flowsheet of the EDC drying and purification column.](source)

All the equipment and accessories were inserted in the simulation: distillation column, reboiler, condenser, reflux tank, stream splitter, control valve and pumps. The distillation column was simulated using model RadFrac/strip2, which allows to inset two feed streams. All stages had a 70%, Anderson et al. (1976), Murphee efficiency except for the reboiler and condenser, which was considered 100%. The \(F_1\) and \(F_2\) feeds may contain up to 18 compounds, but only the significant (above 10 ppm) ones were considered: EDC, carbon tetrachloride (\(\text{CCl}_4\)), chloroform (\(\text{CHCl}_3\)), water (\(\text{H}_2\text{O}\)), 1,1,2-trichloroethane (\(\text{CH}_3\text{CCl}_3\)), tetrachloroethane (\(\text{C}_2\text{H}_2\text{Cl}_4\)), ethyl chloride (\(\text{C}_2\text{H}_5\text{Cl}\)), MVC (\(\text{C}_2\text{H}_3\text{Cl}\)), chloroethanol (\(\text{C}_2\text{H}_5\text{ClO}\)) and chloral (\(\text{C}_2\text{HCl}_3\text{O}\)). The operational variables of the feed were fixed, and the simulation results are shown (Table 1).
Table 1. Values of mass flow and stream compositions in the EDC drying process.

<table>
<thead>
<tr>
<th>Stream</th>
<th>Mass flow (kg/hr)</th>
<th>Mass fraction (kg/kg) of the components</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(x_{CCl_4})</td>
</tr>
<tr>
<td>(F_1)</td>
<td>19050.88</td>
<td>0.002169</td>
</tr>
<tr>
<td>(F_2)</td>
<td>22432.82</td>
<td>0.000029</td>
</tr>
<tr>
<td>(D_1)</td>
<td>40.00</td>
<td>0.012000</td>
</tr>
<tr>
<td>(B_1)</td>
<td>41421.36</td>
<td>0.001000</td>
</tr>
<tr>
<td>(VENT)</td>
<td>3.60</td>
<td>0.005000</td>
</tr>
<tr>
<td>(W_1)</td>
<td>38.71</td>
<td>0.000053</td>
</tr>
</tbody>
</table>

Source: Prepared by the authors themselves.

When the “Heater” model is used, it can be verified that the top vapors are partially condensed and two liquid phases appear: a saturated organic water and an aqueous organic, separated in the decanter tank (Decanter). Two uncoupled blocks, a column block and a decanter block were used to better represent the system. The splitter divides the organic stream into the reflux (R1) and light product (D1) streams. The \(\gamma-\phi\) thermodynamic model was used for the gas-liquid equilibrium, and the Non-Random Two Liquid (NRTL) and Hayden-O’Connell models were used to determine the activity and fugacity coefficients, respectively. The NRTL model was selected based on Poling et al. (2001), which indicates this model for systems with partial miscibility, as in this study. The temperature and composition (the four most represent components) profiles along the column. Temperatures \(T_0\), \(T_1\) and \(T_{71}\), measured by sensors, are also included (Figure 5), and are in good agreement with those calculated in the simulation. The stages 1 and 71 represents the top and bottom of the column, the \(CCl_4\) can be identified as the intermediate component due to the characteristic concentration inversion in the rectification region (Figure 6). An important feature is the EDC concentration of the feed stages (31, 53) of approximately 98.8 and 99.7%, respectively, which distinguish this process as a high purity distillation column. This means that column C1 needs 71 stages to increase about 0.3% and 2% in the feed purity of stages 31 53, respectively, in order to have the bottom composition of EDC above 99.5%.

The C1 features a heterogeneous azeotropic distillation column due to the formation of two liquid immiscible phases in some stages (one rich in COCs and saturated with water, and the other with water saturated with COCs). However, there is no solvent addition to modify the relative volatility, leaving this function to
the reflux rate (R1). Some components of the mixture to be separated in C1 do not have ideal behavior: the EDC-CCl₄ systems present a homogeneous azeotropic behavior, but CCl₄-H₂O instead presents a heterogeneous azeotropic system. They are shown temperature-composition diagram (Txy) and xy diagram for the EDC-CCl₄ system (Figure 7A and B) and CCl₄-H₂O system (Figure 7C and D).

Figure 5. Profile of temperatures in column C₁.

Source: Prepared by the authors themselves.

Figure 6. Composition profile in column C₁.

Source: Prepared by the authors themselves.
4.1 VARIABLE SELECTION IN SOFT-SENSORS

The results obtained by neural networks depend upon the quality and quantity of the available data containing relevant information which enables the model to reproduce the process. Some relevant works about soft sensor, such as those by Luo et al. (1995) and Kano et al. (2009), estimate the composition of interest considering the feed compositions. However, these are inconvenient variables when trying to monitor the process transient due to the difficulty in the composition measurement. For this reason, some tests for variable selection were carried out in this study to establish the need for using these feed compositions, or others which more easily measure variables such as temperatures or pressures in the stages.
4.1.1 Sensitivity mesh analysis and stage temperature selection

The temperature profile along the stages and the distillate and bottom composition are closely related. However, in high purity processes where temperature changes in the stages are very small, the substitution of the feed composition by those temperatures is very risky. Therefore, a group of candidate variables to be evaluated in the variable selection process is proposed herein. It is fair to notice that only three temperature sensors are available along the column C1: the bottom stage (\(T_7\)), the top stage (\(T_1\)) and the reflux drum (\(T_0\)). However, other temperature stages could be important for purpose and control inference. Then it is possible to detect the most sensitive variables with the SVD method and the steady state simulations. The following manipulated variables (\(MV_i\)) were selected for the SVD analysis: \(F_1\), \(F_2\), \(F_3\), \(QR_1\) and \(R_1\), generating 10 combination pairs. Each variable pair is analyzed with respect to the process variables (stage temperatures), \(PV_j\). For the sensitive analysis each pair of manipulated variables were perturbed to +10% of the steady stage value and the changes in the \(PV_j\) variables were monitored. These procedures generate 10 gained matrices \(K_{2,71}\). The combined pairs and the respective conditional number are presented (Table 2).

<table>
<thead>
<tr>
<th>Combined Pairs</th>
<th>Conditional Number (CN)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(F_1 - F_2)</td>
<td>10.31390</td>
</tr>
<tr>
<td>(F_1 - F_3)</td>
<td>4.2030</td>
</tr>
<tr>
<td>(F_1 - R_1)</td>
<td>222.8583</td>
</tr>
<tr>
<td>(F_1 - QR_1)</td>
<td>26.1591</td>
</tr>
<tr>
<td>(F_2 - F_3)</td>
<td>9.9220</td>
</tr>
<tr>
<td>(F_2 - QR_1)</td>
<td>441.3733</td>
</tr>
<tr>
<td>(F_2 - R_1)</td>
<td>115.3110</td>
</tr>
<tr>
<td>(F_3 - QR_1)</td>
<td>35.3539</td>
</tr>
<tr>
<td>(F_3 - R_1)</td>
<td>342.5086</td>
</tr>
<tr>
<td>(R - QR_1)</td>
<td>12565</td>
</tr>
</tbody>
</table>

Source: Prepared by the authors themselves.
The four least CNs highlighted in (Table 2) are related to the three stream flows $F_1$, $F_2$ and $F_3$. The results were evaluated for the lowest CN, pair $F_1$-$F_3$. The left singular gain matrix $(U)$ gives a more adequate coordinate system to visualize the process sensitivity. In this sense, the graph, depicted in (Figure 8A e B) shows the stationary gain and sensitive $(U)$ matrixes for the $(F_1$-$F_3)$ pair, which has the highest grid interaction. This last statement confirms that the SVD analysis indicates the temperatures of stages 17 and 27 ($T_{17}$, $T_{27}$) has the highest variability along the column. Therefore, only the two most sensitive variables were chosen for the SVD validation in order to make the model of candidate variables robust: $T_{17}$ and $T_{27}$, and the base column temperature (stage 71, $T_{71}$). It should be noted that from the process control point of view, the $F_1$-$F_3$ pair could not be chosen because $F_1$, the feed flow, is a disturbance, and cannot be a manipulated variable, only being used here for sensitivity analysis reasons.

Figure 8. Representation of the SVD method. A - Sensitivity analysis. B - SVD analysis, , sensitivity matrices $U_1$ ($F_1$) e $U_2$ ($F_3$).
4.1.2 Models selection

As shown in (Table 3), eleven secondary variables were considered. The total combination models calculated was 2047 for outputs, \(x_{CCl_4}(B1)\) and \(x_{CHCl_3}(B1)\). The outputs were estimated from the algorithm described in section 3.1, and the performance criteria were presented (Equations 9 and 10).

Table 3. Secondary variables used.

<table>
<thead>
<tr>
<th>Variables Used for Selection</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>(u_1)</td>
<td>(F_1) Feed EDC oxychlorination (kg/h)</td>
</tr>
<tr>
<td>(u_2)</td>
<td>(F_2) Feed EDC direct chlorination (kg/h)</td>
</tr>
<tr>
<td>(u_3)</td>
<td>(D_1) or (F_3) Distillate flow rate (kg/h)</td>
</tr>
<tr>
<td>(u_4)</td>
<td>(Q_{R1}) Reboiler Duty (GJ/h)</td>
</tr>
<tr>
<td>(u_5)</td>
<td>(x_{CCl_4}(F_1)) Composition of (CCl_4) in (F_1) (kg/kg)</td>
</tr>
<tr>
<td>(u_6)</td>
<td>(x_{CCl_4}(F_2)) Composition of (CCl_4) in (F_2) (kg/kg)</td>
</tr>
<tr>
<td>(u_7)</td>
<td>(x_{CHCl_3}(F_1)) Composition of (CHCl_3) in (F_1) (kg/kg)</td>
</tr>
<tr>
<td>(u_8)</td>
<td>(x_{CHCl_3}(F_2)) Composition of (CHCl_3) in (F_2) (kg/kg)</td>
</tr>
<tr>
<td>(u_9)</td>
<td>(T_{17}) tray #17 temperature (°C)</td>
</tr>
<tr>
<td>(u_{10})</td>
<td>(T_{27}) tray #27 temperature (°C)</td>
</tr>
<tr>
<td>(u_{11})</td>
<td>(T_{71}) bottom temperature (°C)</td>
</tr>
</tbody>
</table>

Source: Prepared by the authors themselves.

Data was generated from 100 hours of Aspen Dynamics™ simulations for the selection variables process with the following considerations: pseudo-random-binary sequence (PRBS) were applied for variables \(u_6\), amplitude of 10%; \(u_7\), \(u_8\), \(u_9\) e \(u_{10}\), amplitude of 20%; variables \(u_1\), \(u_2\) and \(u_3\) underwent steps with amplitudes of +5 and -5% in a non-simultaneous manner; and white Gaussian noise was inserted in the controller measure signals of \(u_1\), \(u_2\) and \(u_3\). The results of the variable selection process are presented for output \(x_{CCl_4}(B1)\) and \(x_{CHCl_3}(B1)\) (Table 4). The criterion for selecting the better models is based in the lower MSE and greater \(R^2_{\text{Adj}}\), since a lower MSE increases the numerator of (Equation 10), consequently increasing \(R^2_{\text{Adj}}\). Models \(M_1\)-\(M_5\) and \(M_{11}\)-\(M_{15}\) respectively, are the five best inferential models containing temperatures (\(u_9\), \(u_{10}\), \(u_{11}\)) for the outputs \(x_{CCl_4}(B1)\) and \(x_{CHCl_3}(B1)\). While models \(M_6\)-\(M_{10}\) and \(M_{11}\)-\(M_{15}\) respectively, are the five best inferential models containing feed composition (\(u_5\), \(u_6\), \(u_7\) and \(u_8\)) for the same \(x_{CCl_4}(B1)\) and \(x_{CHCl_3}(B1)\) outputs. From these it can be concluded that both \(M_1\)-\(M_5\) (output \(x_{CCl_4}(B1)\)) and \(M_6\)-\(M_{10}\) (output
xCHCl₃(B1)) models without and with composition variables had very similar results. This indicates that the composition feeds may be disregard in order to improve the estimation robustness.

Table 4. Best inference models.

<table>
<thead>
<tr>
<th>Model</th>
<th>Best Combinations</th>
<th>MSE</th>
<th>R²Adjust</th>
<th>Outputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>M₁</td>
<td>[u₁ u₂ u₃ u₄ u₅ u₆ u₇]</td>
<td>140.6138</td>
<td>0.6852</td>
<td>XCCl₄(B1)</td>
</tr>
<tr>
<td>M₂</td>
<td>[u₁ u₂ u₃ u₄ u₅ u₆ u₇]</td>
<td>140.6352</td>
<td>0.6739</td>
<td></td>
</tr>
<tr>
<td>M₃</td>
<td>[u₁ u₂ u₃ u₄ u₅ u₆ u₇]</td>
<td>140.7473</td>
<td>0.6596</td>
<td></td>
</tr>
<tr>
<td>M₄</td>
<td>[u₁ u₂ u₃ u₄ u₅ u₆ u₇]</td>
<td>140.7811</td>
<td>0.6464</td>
<td></td>
</tr>
<tr>
<td>M₅</td>
<td>[u₁ u₂ u₃ u₄ u₅ u₆ u₇]</td>
<td>142.4610</td>
<td>0.6182</td>
<td></td>
</tr>
<tr>
<td>M₆</td>
<td>[u₁ u₂ u₃ u₄ u₅ u₆ u₇]</td>
<td>271.3035</td>
<td>0.4740</td>
<td></td>
</tr>
<tr>
<td>M₇</td>
<td>[u₁ u₂ u₃ u₄ u₅ u₆ u₇]</td>
<td>271.3112</td>
<td>0.4736</td>
<td></td>
</tr>
<tr>
<td>M₈</td>
<td>[u₁ u₂ u₃ u₄ u₅ u₆ u₇]</td>
<td>271.5238</td>
<td>0.4698</td>
<td></td>
</tr>
<tr>
<td>M₉</td>
<td>[u₁ u₂ u₃ u₄ u₅ u₆ u₇]</td>
<td>271.5637</td>
<td>0.4662</td>
<td></td>
</tr>
<tr>
<td>M₁₀</td>
<td>[u₁ u₂ u₃ u₄ u₅ u₆ u₇]</td>
<td>272.2005</td>
<td>0.4480</td>
<td></td>
</tr>
<tr>
<td>M₁₁</td>
<td>[u₁ u₂ u₃ u₄ u₅ u₆ u₇]</td>
<td>117.4310</td>
<td>0.6851</td>
<td>XCHCl₃(B1)</td>
</tr>
<tr>
<td>M₁₂</td>
<td>[u₁ u₂ u₃ u₄ u₅ u₆ u₇]</td>
<td>117.4503</td>
<td>0.6740</td>
<td></td>
</tr>
<tr>
<td>M₁₃</td>
<td>[u₁ u₂ u₃ u₄ u₅ u₆ u₇]</td>
<td>119.6843</td>
<td>0.6064</td>
<td></td>
</tr>
<tr>
<td>M₁₄</td>
<td>[u₁ u₂ u₃ u₄ u₅ u₆ u₇]</td>
<td>119.6984</td>
<td>0.5982</td>
<td></td>
</tr>
<tr>
<td>M₁₅</td>
<td>[u₁ u₂ u₃ u₄ u₅ u₆ u₇]</td>
<td>123.4800</td>
<td>0.4787</td>
<td></td>
</tr>
<tr>
<td>M₁₆</td>
<td>[u₁ u₂ u₃ u₄ u₅ u₆ u₇]</td>
<td>261.7077</td>
<td>0.2851</td>
<td></td>
</tr>
<tr>
<td>M₁₇</td>
<td>[u₁ u₂ u₃ u₄ u₅ u₆ u₇]</td>
<td>261.7943</td>
<td>0.2739</td>
<td></td>
</tr>
<tr>
<td>M₁₈</td>
<td>[u₁ u₂ u₃ u₄ u₅ u₆ u₇]</td>
<td>261.9902</td>
<td>0.2675</td>
<td></td>
</tr>
<tr>
<td>M₁₉</td>
<td>[u₁ u₂ u₃ u₄ u₅ u₆ u₇]</td>
<td>262.0176</td>
<td>0.2596</td>
<td></td>
</tr>
<tr>
<td>M₂₀</td>
<td>[u₁ u₂ u₃ u₄ u₅ u₆ u₇]</td>
<td>262.7865</td>
<td>0.2372</td>
<td></td>
</tr>
</tbody>
</table>

Source: Prepared by the authors themselves.

5 SOFT SENSORS FOR THE AZEOTROPIC COLUMN

The M₁-M₅ and M₁₁-M₁₅ models were selected to estimate the mass fraction of xCCl₄(B1) and xCHCl₃(B1) by the ANNs. The input data was normalized according to the activation function to improve the inference: hyperbolic tangent (interval [-1 1]) and sigmoid logarithm (normalized in the interval [0 1]). As mentioned above (section 3.4.2), the data was divided into three groups: the training data group (70% of data), the test group, and the validation group (15% of data each). The MLP type was used regarding the ANN’s topology. ANN with 1 to 3 hidden layers were tested containing between 5 to 25 neurons per layer. The training procedure consists in off-line (using Matlab®) and on-line (Simulink®) routines, described in the following algorithm:
- **Step 1.** Input the total data of the selected models, select the variables for each output, \( M_1 - M_5 \) for \( xCCl_4 \) and \( M_{11} - M_{15} \) for \( xCHCl_3 \).

- **Step 2.** Select the trainee and test standards.

- **Step 3.** Define standards for the normalizing activation functions.

- **Step 4.** Define the ANN’s topology candidates to be used.

- **Step 5.** Train the candidate topologies.

- **Step 6.** Use the Early Stopping Method of Training.

- **Step 7.** Select the ANN’s with the lowest MSE and RMSE or return to **Step 1**.

The MSE and RMSE errors criteria were used to measure the ANN’s performance. A mean MSE was calculated for three simulation runs. These mean MSE performances are depicted for different hidden layers and layer neurons (Figure 9). The figures show the ten best ANN models for \( xCCl_4 \) (Figure 9A) and \( xCHCl_3 \) (Figure 9B) outputs. In comparing the models for \( xCCl_4 \) output (Figure 9A), model \( M_1 \) has the best performance, independently of the number of hidden layers. The lowest MSE (7.0387E-4) and RMSE (2.65E-2) errors for this model were for the three hidden layer model [10 15 10]. Similarly, the \( M_{11} \) model presents the same characteristics, the lowest MSE and RMSE values (1.15E-3, 3.39E-2) with three hidden layers [10 15 10] for the \( xCHCl_3 \) output (Figure 9B). It is noteworthy to state that these better performances were also attained with the test and validation data. The values of the “Initial mu", "mu decrease factor", "mu increase factor" and "Maximum mu", which produced the best results for ANN's simulations were 1E-3; 0.1, 10 and 1E10, respectively.
Figure 9. ANN that presented the best results. A - Output \( x_{CCl_4} (B_1) \). B - Output \( x_{CHCl_3} (B_1) \).

Therefore, \( M_1 \) (for output \( x_{CCl_4} (B_1) \)) and \( M_{11} \) (for output \( x_{CHCl_3} (B_1) \)) models with their respective network topologies were used for inferring the outputs due to their good performance, as presented above. Moreover, Column 1 (C1) is provided with physical sensors for variables \( u_1, u_2, u_3, u_4 \) and \( u_{11} \) remaining to monitor variables \( u_9 \) and \( u_{10} \), which is an easy to install and cheap task. The training results for the two outputs (\( x_{CCl_4} \) and \( x_{CHCl_3} \)) are shown (Figure A and B, respectively) in which the inference variable is compared with that obtained by simulation.
As can be shown (Figure 10) both soft sensors were submitted to the disturbances and noise; even so, inference values were overlapped to those provided by the *Aspen Dynamics™* simulation (take as real value), and the data used for the ANN trainee. The mean MSE for each sensor were given for a ten sequential training period as 5.9515E-04 and 3.4095E-04 respectively. Similar values were obtained for the validation error ($E_{val}$). In observing the dynamics simulations (Figure 11 A and B) MLP ANN’s with LMA models $M_1$ and $M_{11}$ with [10 15 10] topologies were able to create learning patterns and provide an efficient composition estimation. The regression analysis compares the estimated and real values, its results represent 5001 data values used in the $xCCl_4(B_1)$ and $xCHCl_3(B_1)$ studies (Figure 12 A and, respectively). The regression analysis is presented only as a qualitative analysis of the results. Therefore, in showed at (Figure 12) it can be concluded that the results are consistent with a dispersion coefficient very near to unity.
Figure 12. Dispersions of data estimated by soft sensors. A - Regression of the data to $x_{CCl4}$ ($B_1$), B - Regression of the data to $x_{CHCl_3}$($B_1$)

Source: Prepared by the authors themselves.

5.1 INFERENTIAL CONTROL STRATEGIES

For the unit classical control structure (Figure 13 A) includes: a controller for each feed flowrate and controller for the distillation flow rate (FC-101, FC-102, and FC-103, respectively, both with reverse control action), three controllers for the reflux drum (two control levels for the aqueous and organic phase, and the third for controlling the tank pressure) LC-101, LC-102 and PC-101, both with direct control action, and a sump level controller (LC-103).

An inferential control was proposed for column C1 for minimizing disturbances. As there are two compositions to be controlled, the most critical variable, $x_{CHCl_3}$, was chosen. Thus, an inference feedback controller (AC-101) was implemented (Figure 13B). It was also verified that temperature fluctuations in stage 27 ($T_{27}$) can be minimized before they can affect the composition, so a cascading inferential strategy was proposed, where $x_{CHCl_3}$($B_1$) is the primary variable, AC-101, and $T_{27}$ the secondary, T-101, as shown in (Figure 13C). Nevertheless, an inferential cascade-ratio control strategy was inserted (Figure 13D) to minimize the feed rate variations $F_1$, the major process disturbance. The output signal from controller AC-101 changes the $T_{27}$ set point (Stage 27 temperature) of controller TC-101, which in turn changes the ratio of the reboiler heat load to the feed flow rate. The steady-state value of the second input to the multiplier, which is the TC-101 output signal, is calculated as $Q_{R1}(GJ/h)/F_1(kg/h) = 5E-4$. 
The feed flowrates $F_1$ and $F_2$ only change their set-points in abnormal process situations such as the unit start up or shut down, or in a schedule production change, characterizing them as great process disturbances. The $xC_4Cl_4$(in $F_1$ or $F_2$) and $xCHCl_3$(in $F_1$ or $F_2$) feed compositions are also disturbances to be considered. Thus, dynamic open-loop disturbances in the composition and flow rate variables of $F_1$ and $F_2$ feeds were performed in order to study how to lower the effect of these variables in the process. The composition was +20% step disturbed at 5h and -20% at 20h; similarly, a +5% step disturbance was at 5h and -5% at 20h for the flowrates. The transient behaviors for the outputs variables $xCHCl_3$ and $xCCl_4$ by soft sensors are presented (Figure 14 A and B).
Figure 13. Control structures of the column C1. A - Installed control structure. B - Inferential feedback control. C - Inferential cascade control. D - Inferential ratio-cascade control.

Source: Prepared by the authors themselves.

Figure 14. Transient behavior for open loop compositions to disturbances. A - Transient response of xCHCl₃ in open loop. B - Transient response of xCCl₄ in open loop.

Source: Prepared by the authors themselves.

From the behavior of the system (Figures 14A and B), the responses to flowrate disturbances are inverse to those obtained for composition disturbances.
This means there is an increasing tendency of $x_{CCl_4(B1)}$ and $x_{CHCl_3(B1)}$ after the increasing step of $F_1$ and $F_2$ (at 5hr) and inversely in the decreasing step, but then the tendency is reversed again after the disturbance. When the feed composition is disturbed, again in (Figure 14 A and B), the output compositions were more sensitive to composition changes in the $F_1$ feed stream. This is relevant as this stream is less pure in EDC and contains more contaminants. Comparatively, the percent overshoot was 30% higher in both outputs, when the composition disturbances were inserted in $F_1$.

5.1.1 Analysis of compositions with direct temperature control

To minimize the effects of disturbances, some control strategies were implemented in the system and compared. The first one involves the control of the temperature of stage 27 ($T_{27}$) and then the temperature of the tower bottom ($T_{71}$) was controlled (Figures 15A and B, respectively) testing its effectiveness when disturbances were inserted in the flowrate $F_1$. According to the graphs of (Figures 15C and D), it is found that the control of $T_{27}$ was more efficient in maintaining the compositions of the column base, $x_{CCl_4(B1)}$ and $x_{CHCl_3(B1)}$, compared to $T_{71}$. These results are in accordance with those presented in section 4.1.1., where it was observed by the SVD analysis that stage 27 is part of the region with the highest column temperature gradient.

5.1.2 Analysis of compositions with inferential control strategies

To improve the rejection of the disturbances, three inferential control structures were proposed, the neural soft-sensors online training were conducted in the Simulink® software and linked online with Aspen Dynamics™ through the AMSimulation interface block. The flowsheet of inferential ratio-cascade control structure was building in the Simulink® (Figure 16 A), the “AM Simulation block” (Aspen Dynamics™ simulation) (Figure 16 B), which simulates Column 1 and the implemented neural soft sensor block (Figure 16 C).
Figure 15. Transient behavior for open loop compositions to disturbances. A - Direct temperature control of stage 27 ($T_{27}$). B - Direct temperature control of stage 71 ($T_{71}$). C - Transient behavior of $x_{CHCl_3}$ with $T_{27}$ control. D - Transient behavior of $x_{CCl_4}$ with $T_{71}$ control.

Source: Prepared by the authors themselves.
The graphs (Figure 17) showed a comparative performance of the proposed control structures when disturbances were applied in the F₁ and feed composition. The flowrate and temperature controllers were tuned using the relay test for close loop, while the composition controllers used the internal control method (ICM). When F₁ flowrate and xCHCl₃ composition are disturbed (Figure 17A and B, respectively), the inferential ratio-cascade controller performance is significantly better than the cascade alone and the feedback controllers; this is because it can get ahead of changes in F₁ so to reduce abrupt fluctuations of the xCHCl₃(B₁) around its set-point. It is also still noticeable that the cascade controller also presents good performance due to the fast-secondary temperature controller response that avoids xCHCl₃(B₁) departures from its set-point value. All controllers presented similar comparable performances for the disturbances over
the feed composition (Figure 17B); nevertheless, the cascade and ratio-cascade were superior.

The integral of the absolute value of the error (IAE), integral of absolute error multiplied by time (ITAE) and integral of the square error (ISE) are used as a quantitative performance criterion to indicate the best control structure to be used (Equations 26, 27 and 28, respectively).

\[
IAE = \int_0^\infty |e(t)| \, dt \tag{26}
\]

\[
ITAE = \int_0^\infty t |e(t)| \, dt \tag{27}
\]

\[
ISE = \int_0^\infty e^2(t) \, dt \tag{28}
\]

Where

\( t \) is the time and \( e(t) \) is difference between setpoint and controlled variable.

The results of these evaluations are shown in (Table 5), and once again the inferential rate-cascade control presents the lower IAE, ITAE and ISE values.
for both flowrate and composition disturbances, meaning it is the better control structure.

Table 5. Controller performances.

<table>
<thead>
<tr>
<th>Control Strategy</th>
<th>Disturbances</th>
<th>IAE</th>
<th>ITAE</th>
<th>ISE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inferential feedback</td>
<td>F₁</td>
<td>4.22E-5</td>
<td>8.34E-6</td>
<td>8.40E-4</td>
</tr>
<tr>
<td>Inferential Cascade</td>
<td>CHCl₃(B₁)</td>
<td>6.48E-6</td>
<td>6.42E-8</td>
<td>1.22E-5</td>
</tr>
<tr>
<td>Inferential Cascade-ratio</td>
<td>2.56E-8</td>
<td>5.12E-11</td>
<td>4.73E-6</td>
<td></td>
</tr>
<tr>
<td>Inferential feedback</td>
<td>CHCl₃(B₁)</td>
<td>9.10E-8</td>
<td>3.69E-10</td>
<td>8.96E-6</td>
</tr>
<tr>
<td>Inferential Cascade</td>
<td>1.10E-9</td>
<td>6.66E-11</td>
<td>3.56E-7</td>
<td></td>
</tr>
<tr>
<td>Inferential Cascade-ratio</td>
<td>1.55E-12</td>
<td>5.48E-14</td>
<td>1.56E-8</td>
<td></td>
</tr>
</tbody>
</table>

Source: Prepared by the authors themselves.

6 CONCLUSIONS

This manuscript aimed to develop a methodology for constructing soft sensors with the objective of estimating and controlling the compositions of impurities that predate the cracking of 1,2-dichloroethane (EDC). The impurities decrease the operating time of the thermal cracking reactors, and they should be minimized in the EDC purification step, specifically around the high purity azeotropic distillation column of the EDC. The soft sensor was a viable alternative for estimating carbon tetrachloride (CCl₄) and trichloromethane (CHCl₃) compositions in real time, even with the typical simulation of disturbances and noise which are characteristic of physical sensor measurements, for example, the white Gaussian noise and pseudo random binary sequence (PRBS). In this paper, the molar fractions of the estimated compounds are diluted in the mixture and have values in the magnitude order of ppm; a fact which could compromise the estimation. Therefore, an algorithm for selecting secondary variables was developed based on the all-possible regressions (APR), which proved effective for the number of candidate variables and using statistical criteria for performance evaluation.

A significant contribution of this paper to the literature is the fact that the inference for very low molar fraction contaminants for high-purity columns can be made by dispensing with feed compositions as secondary variables and replacing
them with temperature measurements along the tower. In addition, all secondary variables except temperatures T17 and T27 have their measurements available in the process. The ten (10) best models in the selection were trained and thus confirmed by ANNs as those with the best estimates, all of whom employed the Levenberg-Marquart algorithm, which accelerated training and improved the ANN’s online implementation.

Another relevant contribution in this manuscript concerns inferential control in high-purity towers, three control strategies were implemented with communicability of the Simulink® and Aspen Dynamics™ software: inferential feedback, inferential cascade and inferential cascade-ratio. Lastly, the controllers presented satisfactory performances in rejecting disturbances over major process disturbances, flow rate and feed composition. Meanwhile, the inferential ratio-cascade control showed to be superior with the lowest overshoot and integral criteria (IAE, ITAE and ISE) values.

For future work, it is recommended to use soft sensors in supervisory industrial in addition to simulation environments. Furthermore, model-based predictive controllers (MPC) can be used to replace the PID controllers used in inferential control strategies, this can avoid controller retuning problems and increase the robustness of the control system.
REFERENCES


LUYBEN, W. L. Distillation design and control using Aspen simulation. 2nd ed., John Wiley & Sons, Inc. All rights reserved. 2013.


NYENG, I. Modeling of a 1,2-Dichloroethane Cracker. Thesis of the doctoral, Chemical Engineering and Biotechnology, IKP. University of Science and Technology Norwegian University of Science and Technology. 2015.


