NEURAL NETWORK BASED APPROACH APPLIED TO FOR MODELING AND OPTIMIZATION AN INDUSTRIAL ISOPRENE UNIT PRODUCTION

Rita M. B. Alves[*] and Claudio A. O. Nascimento

CESQ- Center of Engineering of Chemical Systems
Department of Chemical Engineering - Polytechnic School - University of São Paulo
Av. Prof. Luciano Gualberto 380, tr3, CEP: 05508-900, São Paulo, SP, Brazil.
E-mail: rita@lscp.pqi.ep.usp.br; oller@usp.br

ABSTRACT

This work describes the application of a three-layer feed-forward neural network (NN) based approach for modeling, simulating and optimization of a real industrial plant. The industrial process studied is the isoprene production unit from BRASKEM. The purpose of this plant is to produce high purity Isoprene for obtaining synthetic and thermoplastic rubber from a C_5 cut arising from a pyrolysis gasoline unit. The chemical process consists basically of a dimerization reactor and a separation column train. The basic idea of the proposed methodology is to replace phenomenological models by an equivalent neural network, and use this NN model to carry out a grid search, mapping all the region of interest in order to perform the optimization procedure. Since NNs are able to extract information from plant data in an efficient manner, for this work, the neural network model was built directly from historical plant data, which were collected every 15 minutes during a period of one year. These data were carefully analyzed in order to identify and eliminate gross and systematic errors and establish steady state operational conditions. The modeling using NN was carried out by parts in order to get information on intermediate streams. Then, the global model was built, by interconnecting each individual neural network model as a sequential simulator, and used to simulate and optimize the process. The optimization procedure carries on a detailed grid search of the region of interest, by a full mapping of the objective function on the space of decision variables. Thus, it is easy to choose the optimum point, identify multiple optima, check constraints violation, and so on. A qualitative optimization procedure was used to take in account product quality, safe operations conditions, and energy consumption. Moreover, this work presents an example of how deal with cases in which problems of dimensionality arises. The optimization of the entire plant involves 21 variables to be optimized, and then the global model was divided into the parts in order to decrease the problem of dimensionality. Each part of the model was optimized separately, but sequentially, using the optimal conditions from the previous optimization procedure. Comparisons between the model's prediction and the experimental data were performed and reasonable results were achieved from an industrial point of view. Using neural network approach provides more comprehensive information for an engineer's analysis than the conventional procedure. NN is certainly a technique of interest due to its capability of learning the system without knowledge of the physical and chemical laws that govern it. Moreover, the NN readily deals with constraints, avoids several typical numerical problems of conventional optimization tools, and is not computationally time-consuming. However, success in obtaining a reliable and robust NN depends strongly on the choice of the variables involved, as well as the quality of available data set and the domain used for training purposes.

INTRODUCTION

Methods for optimization of a given process involve the measure of goodness of a design, or objective function, equality and inequality constraints, as well as constraints related to safety consideration and those that arise from the process model equations. The objective function is

---

[*] To whom all correspondence should be addressed

established from a technical and/or economic viewpoint. The equality constraints usually represent specifications while the inequality constraints usually represent the lower and upper bounds of the operating variables.

Formal methods of optimization can be utilized to optimize a superstructure of process units with streams that can be turned on and off using binary variables. The mixed-integer formulation of the optimization problem, in principle, permits the optimizer to select simultaneously the best flowsheet and optimize it with respect to its continuous variables, such as the pressure levels, reflux ratios, residence times, and split fractions. In practice, however, most design problems are not solved using superstructures and mixed-integer optimization algorithms. Rather, heuristics together with simulation and algorithmic methods are utilized to build and analyze synthesis trees. Although substructures, such as networks of heat exchangers, can be optimized conveniently using mixed-integer methods, it is impractical to attempt the optimization of entire process flowsheets in this manner (Seider et al, 1999).

Before performing optimization of a given process, it is necessary to analyze it, i.e., to model and simulate it. This step allows interpreting flowsheets, to locate malfunction, to predict the performance of processes, and to study potential changes in the operating conditions or the possibility of a retrofit to improve its profitability. The heart of analysis is the mathematical model, a collection of equations that relate the process variables, such as stream temperature, pressure, flow rate, and composition, to surface area, valve settings, geometrical configuration, and so on. There are several levels of analysis. In order of increasing complexity, they involve: material balances, material and energy balances, equipment sizing, and profitability analysis. Additional equations are added at each level. New variables are introduced, and the equation-solved algorithms become more complicated. So, an important prerequisite for process modeling and optimization is the availability of a valid and representative mathematical model that will be able to describe the system accurately. When available, a first-principle based model could be a natural choice and commercial simulators are important tools largely used on real plants and new design in order to analyze and optimize industrials processes. However, in most chemical engineering problems, the first-principle models are non-linear, complex, multivariate systems and computationally time-consuming. Moreover, when using the process simulators, it is important to recognize that, with some expectations, most stream are comprised of chemical species that distribute within one or more solution phases that are assumed to be in phase equilibrium. Then it is important, when using process simulators, to understand how they apply the theory of phase equilibrium in modeling streams as well as in some vapor-liquid equipment and it is known that in process design and optimization, up to 80% of the total computational time can be spent on evaluation of thermodynamic properties (Seider et al, 1999). Consequently to use commercial simulators is a very hard and difficulty task, mainly if fundamental knowledge about the process under consideration is missing or there is a lack of thermodynamics data or due to the complexity of the industrial system. For optimization purposes, that involve quite a high number of simulations, it is desirable that the model should be simulated in short times and, at the same time, should be able to describe the system accurately. In this sense some works have been proposed to use artificial neural network as a substitute for first-principle models (e.g. Nascimento and Giudici, 1998; Calderon et al., 1998; Guardani et al., 2001; Sternowsky et al., 2002).

Neural networks (NNs) have been claimed to be a universal non-linear approximator and their application in the field of chemical engineering has grown rapidly. NNs replace the phenomenological models due their capability of learning the system without knowledge of the physical and chemical laws that govern it. Moreover, the NN readily deal with constraints, avoid several typical numerical problems of conventional optimization tools, and are not computationally time-consuming (Nascimento and Giudici, 1998). The advantage of the NN model over other empirical approaches will depend strongly on the degree of non-linearity of the process. It is possible to obtain a representative NN model based on historical input/output data and/or designed plant experiments. However, the success in obtaining a reliable and robust NN depends strongly on choosing the process variables involved, as well as the quality of the available data set and the domain used for training purposes. If any important process behavior is not covered by the observations, the NN model or any other empirical model will be a poor representation and the optimization probably will fail.

The aim of this work is to optimize the Isoprene Unit by using the NN model built from historical plant data as a substitute of the phenomenological models. The optimization procedure carries out a detailed grid search of the region of interest, by a full mapping of the objective functions on the space of decision variables. The detailed grid search can be achieved in reasonable time if the dimensionality of the problem is not so high. The problem presented here is an example of how it is possible to deal with cases in which problems of dimensionality arise.

NEURAL NETWORK MODELS

In the context of chemical engineering applications, neural networks are of particular interest as predictive models (Nascimento et al., 1994, 1999 and 2000) and for pattern recognition (Nascimento et al., 1997). Moreover, neural networks have become a widely used tool in areas such as process analysis, simulation, and control, primarily due to their inherent advantages such as adaptability, non-linearity, and fault tolerance especially in combination with the relatively simple way of use. Neural networks possess the ability to 'learn' the behavior of the process without actually specifying the physical and chemical laws that govern the system. In addition, neural network models can simultaneously employ continuous and discrete input variables (Migliavacca et al., 1999). The success in obtaining a reliable and robust network strongly depends on the choice of appropriate input or process variables, as well as the available set of data and the domain used for training purposes.

The neural network employed in this study is a three-layer feed-forward network. This type of network, in which information propagates in only one direction, is particularly useful for steady-state modeling. Once the network topology is specified, a set of input-output data is used to train the network; i.e., to determine appropriate values for the weights associated with each interconnection. For a given topology, the magnitude of the weights defines the network characteristics and the structural properties of the model. Thus, an NN has the capability to represent complex systems whose structural properties are unknown (Willis et al., 1991).

The most extensive algorithm for the learning phase is the back-propagation algorithm, which is a generalization of the steepest descent method (Haykin, 1994; Rummelhart and J. McClelland, 1986). Since the NN model fitting is essentially a mathematical adaptive regression in which phenomenological considerations are not used, much care must be taken to validate its representation of the physical process and to prevent overfitting. A simple measure of the quality of the fitting for a given NN is based on comparisons between calculated values and experimental data from the test set (not used in fitting the NN).

The computer programs for data preparation, NN fitting, modelling and optimization used in this work were developed at the CESQ (Center of Engineering of Chemical Systems) of the Chemical Engineering Department of the University of São Paulo (Nascimento, 1991, Nascimento and Alves, 2003).

NEURAL NETWORK BASED APPROACH FOR OPTIMIZATION

The main idea is to replace the model equations with an equivalent neural network, and use this NN to carry on a detailed grid search of the region of interest, by a full mapping of the objective function on the space of decision variables. Using NN instead of a phenomenological model itself takes advantage of the comparative fast response by a neural network simulation. Moreover, the replacement of the first-principle model by an equivalent NN at the optimization step takes the advantage of high speed processing, since simulation with a NN involves only a few non-iterative algebraic calculations. In this way, even a detailed grid search can be achieved in reasonable time, as long as there are not too many variables being optimized (in which case problems of dimensionality arise). This approach is more reliable, readily deals with constraints, avoids several typical numerical problems of conventional optimization tools and is not computationally time-consuming. As an additional benefit, full mapping of the objective function allows one to identify multiple optima easily, an important feature not presented by conventional optimization methods. Moreover, the constraints are easily treated afterwards since points with

violated constraints can be recognized and classified (according to weak or hard constraints) (Nascimento et al., 2000).

Once the map is obtained, it is easy to choose the optimum point, to identify whether multiple optima are present, to check if constraints were violated, and so on. This approach definitely provides more comprehensive information for an engineer's analysis than the conventional non-linear programming procedure. This algorithm can be straightforwardly extended to treat a multi-objective optimization problem as well.

To solve the optimization problem the global structure of optimization is shown in Figure 1

PROCESS DESCRIPTION

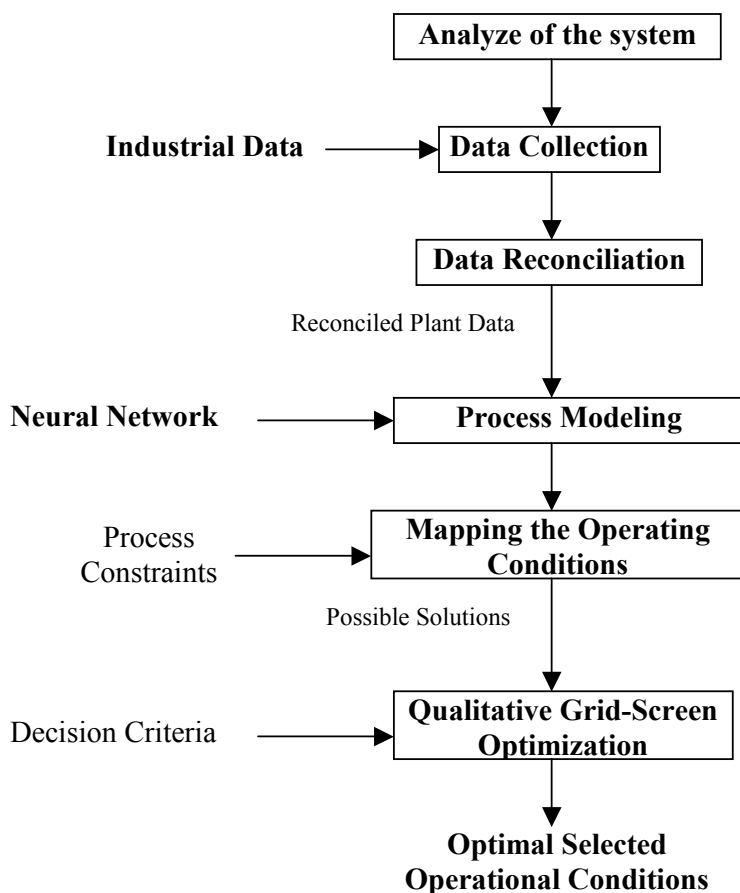The system studied is the Isoprene Production Unit from BRASKEM, the largest Brazilian petrochemical plant.

Fig. 1. Structure of Neural Network based approach

The objective of the Isoprene Extraction Unit is to yield high purity isoprene from a C_5 cut of pyrolysis gasoline. The isoprene produced is to be used to obtain synthetic and thermoplastic rubber. An extractive distillation process is used to perform the isoprene production. The process of isoprene production can be divided into five sections: feed preparation, extractive distillation, solvent recovery, sulfur removal and fractionating. The system consists basically of a dimerization reactor and a separation column train. At the feed preparation section, the unit feed stream flows through a tubular reactor to thermally dimerize cyclopentadiene (CPD) into dicyclopentadiene (DCPD). The reactor effluent is fed to a distillate column for removal of heavies. This first column – First Heavies

Removal Column - separates one stream rich in isoprene, as the distillate, from a heavy cut, rich in DPCD and piperylenes, as the bottom product. This bottom stream feeds another column – Depentanizer Column - to separate the heavy cut into two streams: one rich in DCPD and the other rich in piperylenes, both of them used as raw material in resin manufacture. The distillate product of the first column is the feed stream of the isoprene extraction section, which uses a mixed solvent to separate high purity isoprene from the other compounds of the C_5 stream. The first extractive distillation column is physically divided into two sections. The solvent is fed into the first section, and the distillate from the column for removal of heavies is fed into the second section. Paraffins are obtained as a top product and diolefins, whose main component is the isoprene, are obtained together with the solvent as a bottom product. The bottom product of the first column of extractive distillation is fed to solvent stripper column. The top stream of this column is fed into the last stage of the second extractive distillation column. The ratio solvent/feed for this column is less then those used in the first extractive distillation column in order to get, as the product top, a stream rich in isoprene, that is further washed with water in a liquid-liquid extraction column – product washing column - in order to remove the entrained solvent completely.

The bottom stream from the second extractive distillation column feed also the solvent stripper column, where the bottom recovers all the solvent used in the both extractive distillation columns. After heat recovery, this recycle stream is fed into the extractive distillation columns of the process.

The side stream of the solvent stripper column is sent to the CPD removal column, in which part of the butine-2, CPD, piperylenes and 1,4-pentadiene retained with the solvent is removed by the top. The bottom of this column is solvent quite pure, which is returned back to the solvent stripper column.

All the raffinate streams of the extraction section are washed with water in a liquid-liquid extractive column - C_5 washing column - in order to recover the entrained solvent. The bottom streams of the two washing columns, that contain the extracted solvent, are fed into a distillation column – first solvent recovery column – in order to separate water, which has practically no solvent, as the bottom stream from the stream rich in solvent as the top stream. This stream contains more water than the rate specified for the circulating solvent and returns back to the solvent circuit. The water control in this stream allows the system to keep the water concentration at the desired value.

The hydrocarbons from the top of C_5 washing column are sent to the C_5 raffinate vessel and then they go on to the raffinate spheres, from where they are sent to the other process unit and co-cracking in the pyrolysis furnace.

Isoprene Crude from the product-washing column is fed to the sulphur removal system. After that, the final isoprene purification is performed in two fractionating columns: the first one is for removal of lights and the second one is for removal heavies, and at the top of the last one, isoprene product with purity higher than 99,9% is obtained.

Figure 2 shows the Isoprene Unit schematically.


PROCESS MODELING

As said previously, modeling through neural networks consists, in general, in determining weights associated with each data input. These weights are values that minimize the quadratic error between the value calculated by neural networks and the experimental data provided as output. Thus, it is very important to keep in mind the procedures to be performed in order to adjust models by neural networks. The steps involved are described below:

1. **Analysis of the problem to be modeled by neural networks**

This step involves the consideration of each variable, which characterizes the process, including the binaries ones, since the training procedure is similar to the learning process of the human mind, lack of important information leads the network to determine incomplete non-linear

relationships in the same way and its predictive capability will be lost. The use of redundant information should also be avoided.

At this step, operational, environmental, safety restriction, product specifications, maximum and minimum limits and sensitivity of all the variables involved should be considered.
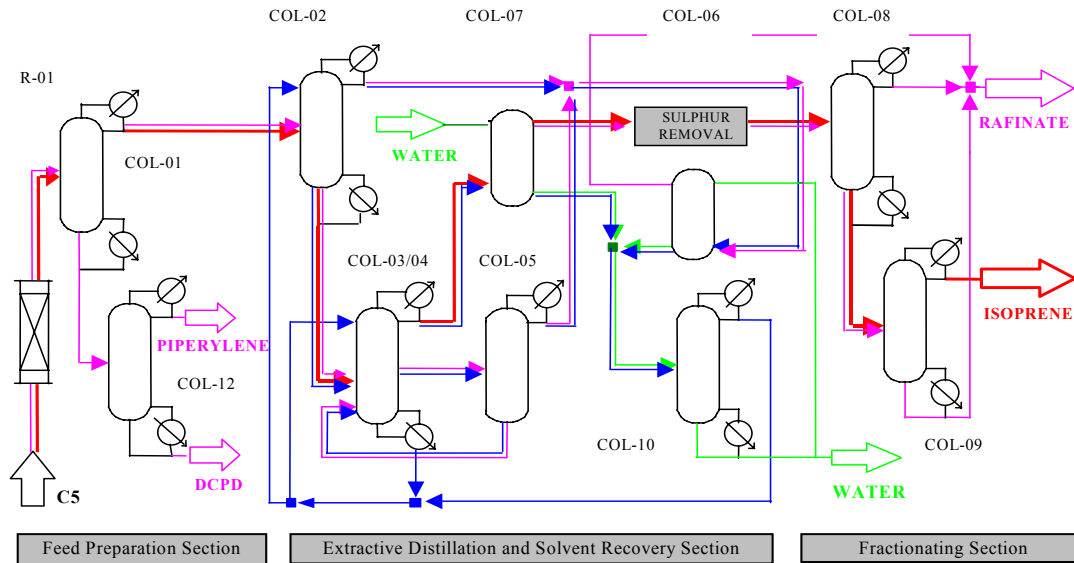
Fig. 2. - Isoprene Unit

2. Data Collection

Quality and quantity of the process data are relevant. Debatable confidence of the values could prejudice the network training and consequently the prediction step will show low quality.

The analysis of the process was undertaken by using a one-year database. The primary database consisting of about 34500 observation sets of 244 variables. The data were collected every 15 minutes.

3. Statistical Analysis and Data Reconciliation

The neural network modeling was done directly from industrial data. However presence of noise and gross errors, commonly found in industrial processes would disturb the modeling and lower its quality. Therefore, for this work, the first step was to prepare the industrial data to a model fitted to detection and elimination of data containing gross errors ("outliers") and systematic errors.

This step is extremely important since the quality of the data has direct influence on the quality of the fitting performed by the neural networks. Knowledge of the process, statistical procedures and others based in the first principles equations are used.

According to the average time considered for the data treatment, data fluctuation could be incorporated in the results. Many times, this could lead to unreliable information. In cases of errors with the measurement instruments over a long period of time, the average reflects this error.

The higher frequency of data collected allowed identifying periods of steady state operation. The criterion adopted was a constant feed flow fluctuation of 0.2-0.3 t/h during a period of 48 to 72 hours.

More details of these procedures can be seen in Alves and Nascimento (2001, 2002) and Alves et al. (2003).

4. Modeling of the System

With the system input and output previously selected at the treatment of the historical data step at the section 3, neural networks can model the process in distinct ways, among them: networks exist that model the whole process at once; networks that model the process by parts and than proceed with a connection among the several networks. In this case, the option was to divide the process in several subsystems, each one representing a given unit operation (reactor, distillation column, and so on). This methodology has as objective:

- To simplify the modeling by decreasing the neural network order.
- To make the plant data consistency better in each subsystem. Measurement instruments errors in one variable in one subsystem do not prejudice, most of time, the information to other subsystems.
- An important part of the process recycle stream is explicit, which facilitate the technical and economical analysis of the process.

In this way, for each subsystem, every process input and output measured variable could be analyzed concerning its importance and consistency. For this, Principal Components Analysis (PCA) was employed as an auxiliary tool in order to direct the judgement. However, the final decision followed the basic directives concerning the critical variables as suggested by the people of BRASKEM.

The Isoprene Process Unit was divided into 10 subsystems in order to carry out the modeling as described previously in the item Modeling of the System:

- ✓ Subsystem 1 – Reactor
- ✓ Subsystem 2 – Heavies Removal First Column
- ✓ Subsystem 3 – Depentanizer Column
- ✓ Subsystem 4 – First Column of Extractive Distillation
- ✓ Subsystem 5 – Second Column of Extractive Distillation and Stripper Column Solvent
- ✓ Subsystem 6 – C_5 Washing Column and Product Washing Column
- ✓ Subsystem 7 – CPD Removal Column
- ✓ Subsystem 8 – Solvent Recovery First Column
- ✓ Subsystem 9 – Lights Removal Column
- ✓ Subsystem 10 – Heavies Removal Second Column

After the modeling of the individual neural network, which represents each subsystem, the main focus is to interconnect sequentially these networks in order to generate data with the objective of building a response surface for all involved variables, as well analyzing the behavior of the most important process variables. Thus a "black box" was built, as shown as an example in Figure 3, containing all the networks interconnected by recycle and sequential streams.
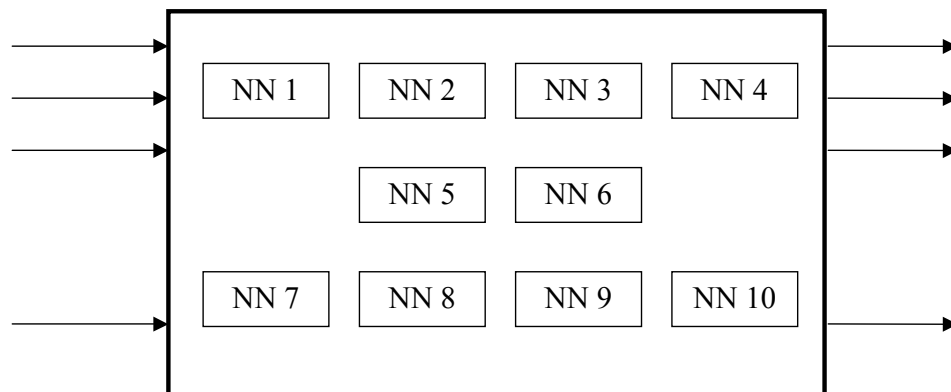

Fig. 3. "Black box" Scheme

The convergence procedure of the integrated system was performed interconnecting all the streams subsequentially at the corresponding unit operation or subsystem, in order to reproduce the industrial plant. The algorithm used for mathematical convergence was direct substitution with dumping factor.

Table 1 shows the variables selected and their applicability range for all subsystems.

4.1. Neural Network Training

After the plant data analysis, the process modeling was performed as explained previously. This step consisted of establishing an optimal neural network structure to represent each process unit individually, i.e., the number of neurons in the hidden layer, the number of iterations for the learning process and the weights associated.

Initially, the industrial data was divided into two sets: a learning set and a test set. The test set contained variables within the range of the learning set variables. Thus, no extrapolation takes place in the model in comparison with the test set data. Then, the NN was run with several different specifications, i.e., varying the number of neurons in the hidden layer and the number of data presentations (iterations), in order to select the best one. In this way several networks with different errors are generated. To avoid an overfitting of the model, choosing a minimum neuron number in the hidden layer (NH), as well as an adequate number of training cycles that accomplishes the lowest error in the test set is recommended (Pollard, 1992).

The phase of learning or training is, basically, a problem of parameters estimation and, in general, the same difficulty associated to conventional optimization, such as, convergence, local minima and time consuming can be found.

Results for the sub-system 1 will be presented as an example of the methodology employed. Figure 4 shows the neural network 1, which represents the reactor.
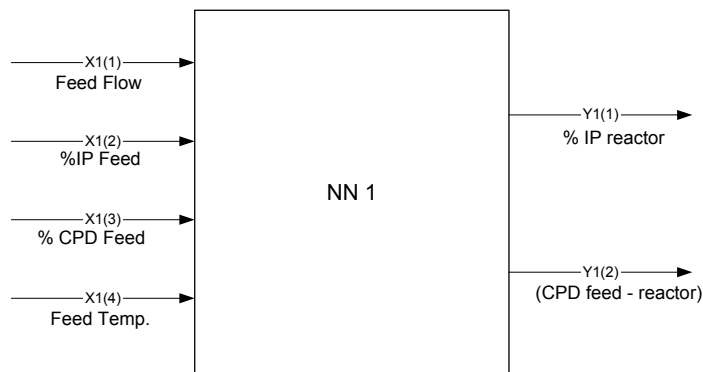
Fig. 4. Neural Network for the reactor

Table 1. Process variables and their application range

Variable	Description	Min	Max	Variable	Description	Min	Max
NN 1				**NN 6**			
x1(1)	Feed Flow	9.5	14.52	x6(1)	Top Temperature	54.81	76.96
x1(2)	%IP in feed	11.57	20.86	x6(2)	Distillate Temperature	27.45	32.03
x1(3)	%CPD in feed	2.31	19.59	x6(3)	%IP - Distillate	1.56	18.71
x1(4)	Feed Temperature	55.79	80.18	x6(4)	%Butyne-2 - Distillate	10.16	19.87
				x6(5)	%CPD - Distillate	20.84	58.91
y1(1)	%IP in the reactor	9.8	20.24	x6(6)	Bottom Flow	0.79	1.37
y1(2)	CPD feed - CPD reactor	0.2	17.45	x6(7)	Bottom Temperature	89.4	95.75
NN 2				y6(1)	Feed Flow	1.24	2
				y6(2)	Reflux Flow	1.17	1.84
x2(1)	Feed Flow	11	14.02				
x2(2)	Feed Temperature	62.61	66.51				
x2(3)	%IP - feed	14.5	20.16	**NN 7**			
x2(4)	%CPD - feed	0.83	7.04				
x2(5)	Reflux Flow	31.22	53.23	x7(1)	Feed Flow I	4.45	8.71
x2(6)	Reboiler Steam	7.94	11.59	x7(2)	Feed Flow II	2	3.15
				x7(3)	%Butyne-2 - feed	0.34	2.18
y2(1)	Distillate Flow	5.57	8.39	x7(4)	%CPD - feed	0.5	6.61
y2(2)	%IP - Distillate	24.82	37.74	x7(5)	Water Flow I	1.9	3.08
y2(3)	%CPD - Distillate	0.85	3.49	x7(6)	Water Flow II	1.3	3.05
y2(4)	%2M2B - Distillate	0.07	1.24				
y2(5)	Bottom Flow	4.56	6.4	y7(1)	Distillate Flow	1.68	2.83
				y7(2)	Top Temperature	26.8	31.6
				y7(3)	Bottom Flow	4.2	7.3
NN 3							
				NN 8			
x3(1)	Feed Flow	4.61	6.11				
x3(2)	Reflux Flow	1.6	3.41				
x3(3)	Feed Temperature	55.46	66.38	x8(1)	Feed Flow	3.31	7.34
x3(4)	Reboiler Steam	0.83	1.33	x8(2)	Feed Temperature	76.78	87.16
				x8(3)	Reflux Flow	0.6	1.95
y3(1)	Distillate Flow	2.21	3.68	x8(4)	Reboiler Steam	1.26	1.77
y3(2)	Top Temperature	48.7	51.52				
y3(3)	Bottom Flow	1.33	2.99	y8(1)	Distillate Flow	1.17	1.79
y3(4)	Bottom Temperature	110.75	126.98	y8(2)	Top Temperature	101.27	106.43
				y8(3)	%H2O - Distillate	21.33	30.13
				y8(4)	Temperature# 10	105.96	110.21

Table 1. (cont.). Process variables and their application range

Variable	Description	Min	Max	Variable	Description	Min	Max
NN 4							
x4(1)	Feed Flow	5.79	8.54	**NN 9**			
x4(2)	%IP - feed	21.3	38.54				
x4(3)	%CPD - feed	0.81	3.66	x9(1)	Feed Flow	1.47	3.06
x4(4)	%2M2B - feed	0.03	1.38	x9(2)	Feed Temperature	27.58	35.52
x4(5)	Solvent Flow	53.37	71.49	x9(3)	Reflux Flow	9.36	13.88
x4(6)	Reflux Flow	10.62	17.67	x9(4)	Reboiler Solvent Flow	38.26	63.98
x4(7)	Reboiler Steam	2.48	5.68	x9(5)	Reboiler Solvent Input Temperature	102.05.0	104.3
y4(1)	Distillate Flow	1.61	3.7	x9(6)	Reboiler Solvent Output Temperature	66.08	79.63
y4(2)	Distillate Flow	3.15	5.13				
y4(3)	%i-C5 - Distillate	19.26	39.73				
y4(4)	%n-C5 - Distillate	22.48	39.61	y9(1)	DistillateTemperature	31.73	42.73
y4(5)	Bottom Flow	51.43	71.1	y9(2)	Bottom Temperature	58.72	64.06
y4(6)	Bottom Temperature	84.95	91.23	y9(3)	Bottom Flow	1.38	2.73
				NN 10			
NN 5							
				x10(1)	Feed Temperature	60.11	63.77
x5(1)	Feed Flow I	57.27	68.63	x10(2)	Feed Flow	2.01	2.75
x5(2)	Feed Temperature I	85.1	91.23	x10(3)	Reflux Flow	11.39	15
x5(3)	Feed Flow II	0.77	1.33	x10(4)	Reboiler Condensate Flow	6.95	22.7
x5(4)	Feed Temperature II	90.78	95.14	x10(5)	Reboiler Solvent Flow	7.85	23.24
x5(5)	Solvent Flow	13	16.9	x10(6)	Reboiler Solvent Output Temperature	55.21	66.94
x5(6)	Reflux Flow	12.4	15.91				
x5(7)	Reboiler Steam	4.82	6.2				
				y10(1)	Distillate Flow	1.86	2.6
y5(1)	Distillate Flow	2.2	3	y10(2)	%CPD - Distillate	0.07	0.52
y5(2)	%Butyne-2 – Distillate	0.88	1.68	y10(3)	Bottom Temperature	49.2	53.59
y5(3)	%CPD – Distillate	1.2	5.94				
y5(4)	Size Draw	1.22	1.89				
y5(5)	Temperature # 55	98.22	101.07				
y5(6)	Bottom Flow	68.77	80.15				
y5(7)	Bottom Temperature	101.9	104.06				
y5(8)	%H2O – Solvent (Column 04 - Bottom)	15.09	26.17				

Figure 5 shows the result of the variation of the residual error from the test set as function of the number of iterations and neurons in the hidden layer. The error is represented by the square mean deviation between the experimental and calculated values. As the optimal number of neurons in the hidden layer (NH) corresponds to the lowest error from the test set, NH should be 14. However, as a heuristic rule, is better to choose a neural network with the lowest numbers of neurons as possible in order to avoid overfitting. Moreover, the higher is the number of neurons in the hidden layer the higher is the number of weights to be adjusted. For all of these reasons, for this case, a neural network with 7 (seven) neurons was used without any lost of quality in the performance in the representation of systems by the neural network chosen as it can be seen from the results of the test validation and consistency. It can also be verified the influence of the number of iterations on the residual error. At about over 40000 iterations the decreasing of the error is so insignificant to justify using more iteration.
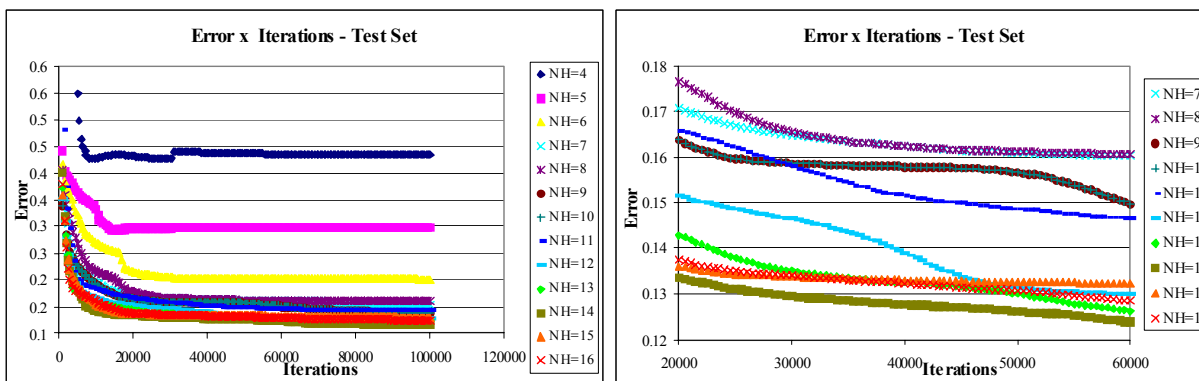

**Fig. 5.** Minimum error for the test set

4. 2. **Validation of the training neural network**

Once the best neural network that represents the output variables was selected, tests of consistency were carried out in order to validate the network. The first evaluation is the analysis of dispersion by comparison between the calculated values and the measured data. Figures 6a and 6b show some examples of the comparison between measured and calculated output variable values for the test data set. Good agreement between modeled and measured values can be observed and the uniform distribution indicates absence of any tendency on fitting.
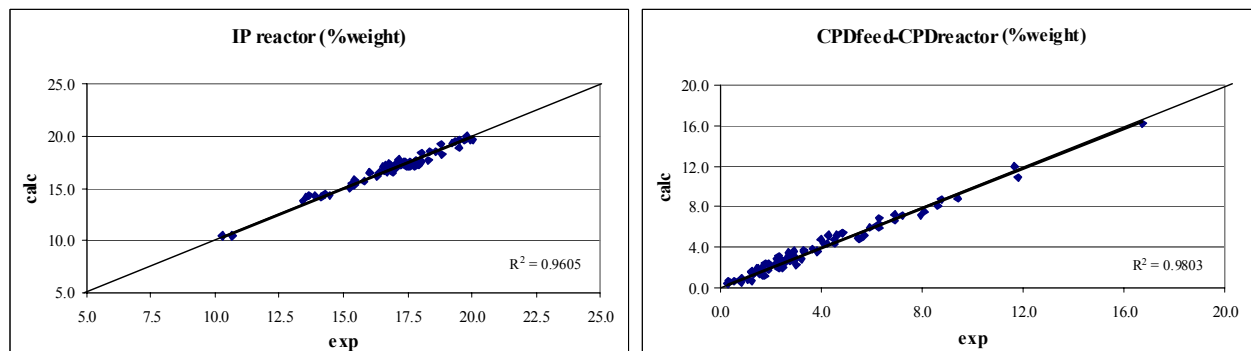
Figs. 6a,b. Comparison between experimental and calculated values

The second test of consistency of the results is concerning the distance function defined as the difference between the experimental and calculated values, $\left(y_{\exp} - y_{calc,RN}\right)$. These information provide the distribution of the difference of the absolute error at each point and evaluate the quality of adjust. More uniform distribution signifies better adjust. Figures 7a and 7b show the graphical of the distance function for the output variables from network 1. It can be observed a good symmetry in the distribution of the errors, however the model shows a good representation of the system.
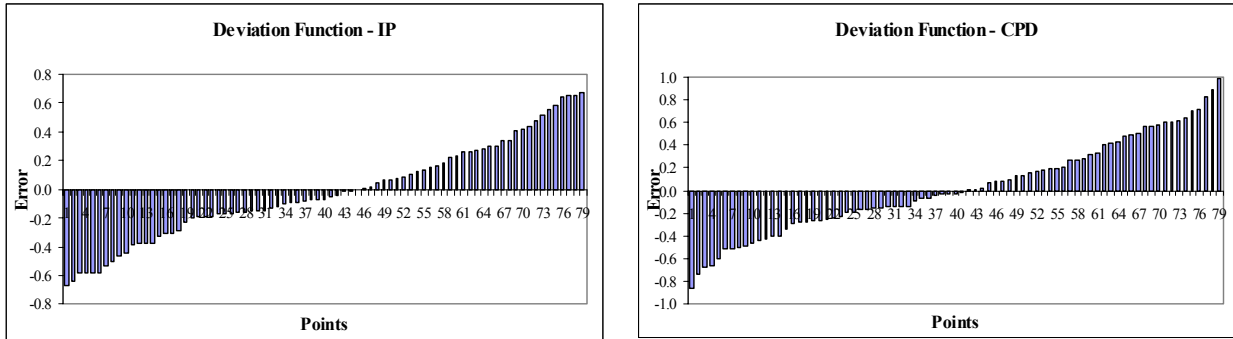

Figs. 7a,b. Distance Function

Then, the normal distribution of errors between the calculated and measured values is verified. This information evaluates the quality of the fit and the normal distribution demonstrates the equilibrium of frequency of the values located around the medium zero. The normal distribution shows the good quality of the adjustment. The histograms of the errors for the output variables of the network 1 are shown in the Figures 8a and 8b.

It can be observed from the histograms a very good fit, since the distribution shows a normal shape and media approximately zero.

In general, the errors evaluated show a small variability, which is an evidence of a good representation of the systems, by neural network model built. This fact is proving that the choice of a lower number of neurons in the hidden layer, in this case, does not prejudice the expected results. An important factor to be considered is the quality of the data. In this case, a better control of the variables in the industrial plant could leave also a better result.
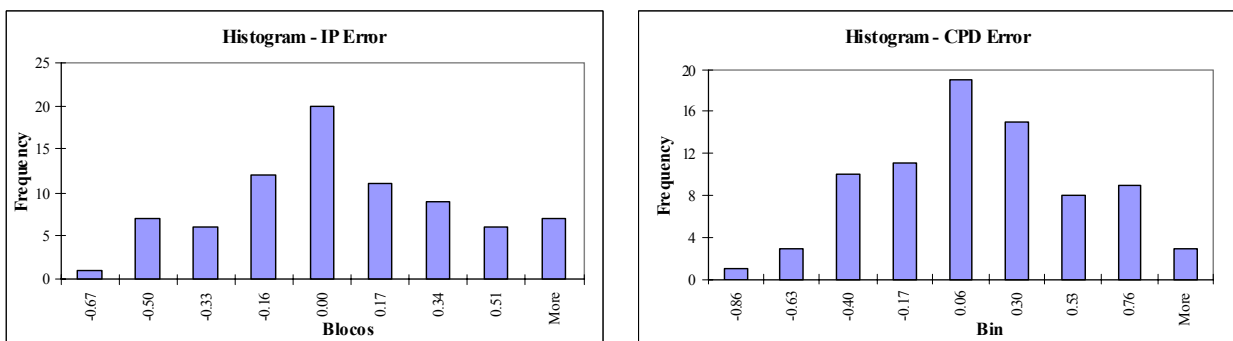
Figs. 8a,b. Histograms

Similar procedure was executed for all others sub-systems and, however each one has its particularities and valuables information, they will not be shown in this. For the present work, the selected neural network configurations are shown are in Table 2 for each process unit involved in this study.

Table 2. Neural network configurations

	Description	NEXP learn	NEXP test	NVAR1	NVAR2	NH	NO	NCMAX	ETA	ALFA
NN 1	**Reactor**	171	79	4	2	7	2	30000	1.00	1.00
NN 2	**Column 01**	56	19	6	5	13	5	25000	1.00	1.00
NN 3	**Column 12**	97	37	4	4	16	4	45000	1.00	1.00
NN 4	**Column 02**	109	42	7	6	19	6	30000	1.00	1.00
NN 5	**Columns 03-04**	79	27	7	8	14	8	67000	0.85	1.00
NN 6	**Column 05**	225	106	7	2	12	2	60000	1.00	1.00
NN 7	**Columns 06-07**	282	136	6	3	13	3	59500	1.00	1.00
NN 8	**Column 10**	202	95	4	4	16	4	55000	1.00	1.00
NN 9	**Column 08**	163	70	6	3	15	3	50000	1.00	1.00
NN 10	**Column 09**	136	58	6	3	16	3	26000	1.00	1.00

4.3. Global Model

The global structure of the modeling using the neural network approach is shown in Figure 9 and Table 3 shows the connections considered.

The global model is complex due to the great number of variables and recycles involved. The global model generated and the structure of convergence employed (direct substitution with dumping factor) show good efficiency in representing the process.

Once the neural network was trained and the global model was well adjusted, the model could be used to simulate the process with great speed. A great number of simulation procedures can be performed in seconds, which allows its use in procedures of process optimization, process control, on-line process control, analysis of process restriction, sensitivity analysis of variables, virtual analyzers, among other applications.
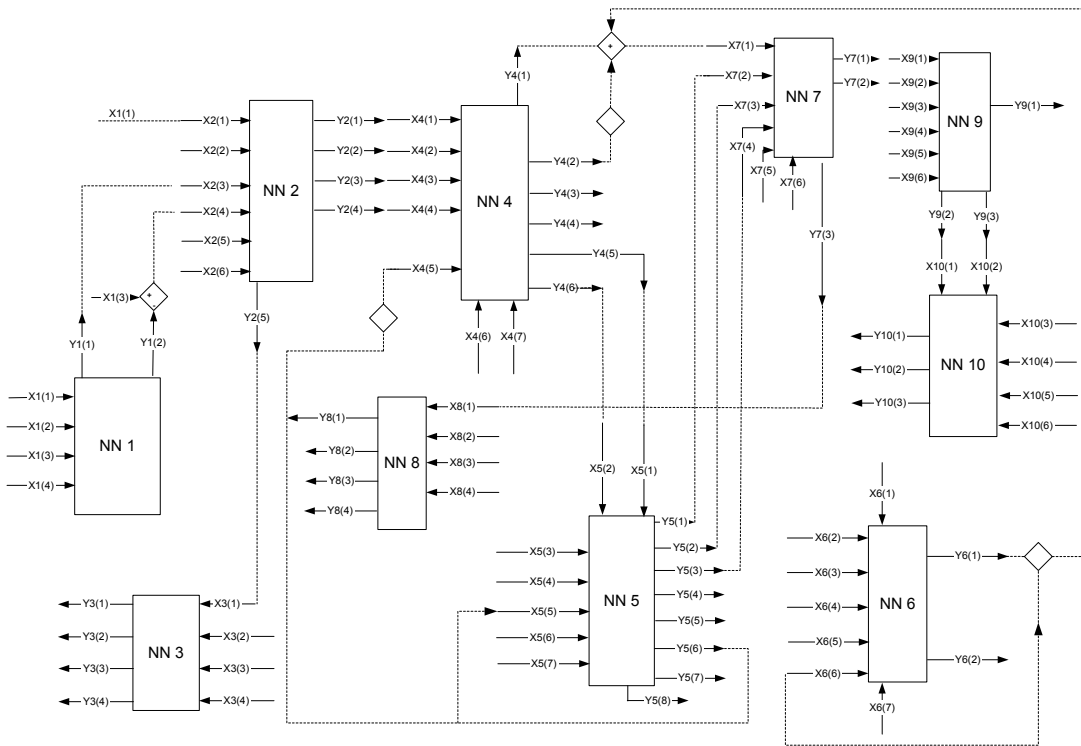
Fig. 9. Scheme of the NN modeling

Table 3. Connections for modeling scheme

Connections
$x2(1) = x1(1)$
$x2(3) = y1(1)$
$x2(4) = x1(3) - y1(2)$
$x3(1) = y2(5)$
$x4(1) = y2(1)$
$x4(2) = y2(2)$
$x4(3) = y2(3)$
$x4(4) = y2(4)$
$x4(5) = y8(1) + y5(6) - x5(5) + 0.20$
$x5(1) = y4(5)$
$x5(2) = y4(6)$
$x5(3) = x6(6)$
$x5(4) = x6(7)$
$y5(4) = y6(1)$
$x7(1) = y4 (1) + y4 (2) + 0.2 + y6 (1) - x6(6)$
$x7(2) = y5(1)$
$x7(3) = y5(2)$
$x7(4) = y5(3)$
$x8(1) = y7(3)$
$x9(1) = y7(1)$
$x9(2) = y7(2)$
$x10(1) = y9(2)$
$x10(2) = y9(3)$

PROCESS OPTIMIZATION

The constraints of quality and safety required by the process and the independent variables are shown in Table 4. All the units of the variables are arbitrary.

The choice of the constraints shown in the Table 4 has the following objectives:

- To specify the composition of CPD and 2M2B at the Feed Preparation Section in order to guarantee the specification of the final product.
- To concentrate IP at the Feed Preparation Section in order to feed the Extractive Distillation Section as required.
- To specify the temperature of the bottom of the Depentanizer Column in order to avoid degradation of DCPD.
- To specify the composition of 2-Butine and CPD at the distillate of the Second Extractive Distillation Column in order to guarantee the specification of the final product.
- To specify the composition of Water at the top of the First Column of Solvent Recovery. The content of Water in the solvent circuit is a critical point of this unit.

Depending of the main objective to be achieved, others constraints could be considered. The cases that will be shown here are some examples among several possibilities of optimization by the model built for this system.

Table 4. – Independent Variables and Constraints for the Process Optimization.

Notation	Variable Description	
	Independent Variables	
x1(1)	Feed Flow of the Unit	
x1(4)	Feed Temperature of the Reactor	
x2(5)	Reflux Flow of the COL-01	
x2(6)	Vapor Flow for the Reboiler of the COL-01	
x3(2)	Reflux Flow of the COL-12	
x3(4)	Vapor Flow for the Reboiler of the COL-12	
x4(6)	Reflux Flow of the COL-02	
x4(5)	Solvent Flow for the COL-02	
x4(7)	Vapor Flow for the Reboiler of the COL-02	
x5(5)	Solvent Flow of the COL-03	
x5(6)	Reflux Flow of the COL-03	
x5(7)	Vapor Flow for the Reboiler of the COL-03	
x7(5)	Water Flow for the COL-07	
x7(6)	Water Flow for the COL-06	
x8(3)	Reflux Flow of the COL-10	
x8(4)	Vapor Flow for the Reboiler of the COL-10	
x9(3)	Reflux Flow of the COL-08	
x9(4)	Solvent Flow for the Reboiler of the COL-08	
x9(5)	Solvent Inlet Temperature - Reboiler of the COL-08	
x9(6)	Solvent Outlet Temperature - Reboiler of the COL-08	
x10(3)	Reflux Flow of the COL-09	
	Constraints	
x1(3)-y1(2)	Composition of CPD at R-01 Outlet	< 4
x1(2)-y1(1)	IP Loss at R-01	< 4
y2(2)	Composition of IP at the distillate of the COL-01	> 23
y2(3)	Composition of CPD at the distillate of the COL-01	< 4
y2(4)	Composition o of 2M2B at the distillate ot the COL-01	< 2
y3(4)	Bottom Temperature - COL-12	< 120
y5(2)	Composition of Butine-2 at the distillate of the COL-03	< 2
y5(3)	Composition of CPD at the distillate of the COL-03	< 2
y8(3)	Composition of H2O at the distillate of the COL-10	< 23

1. Objective Function and Optimization Criteria

The optimization procedure employed in this work does not require necessarily a formal objective function. It should perform either qualitative optimization from a set of process constraints and economical analysis or quantitative optimization by using an objective function that describes the economical goal.

In this step, the most important is the main objective to be achieved, for example, higher production for a given product specification. Thus, the objective function chosen may attempt operational features only, such as to specify feasible modification in certain process variables in order to satisfy a given requirement for others variables.

The main objective of this work is to establish optimal operational conditions in order to obtain higher production of Isoprene within of the specifications defined by the costumers.

2. Mathematical Definition of the Process

This step involves a choice of an appropriate optimization technique, which can be able to define mathematically the process. Since this problem deals with non-linear optimization with constraints and the interest is to show the capability of the model in providing some optimal operational conditions in a feasible range, the option was to use the direct search method.

The optimization of the entire plant involves 21 variables to be optimized. The range of applicability of each of one is divided into equal intervals, which generates a great number of cases when mapping the grid of possible solutions. This makes the problem so big to be solved easily. Then the global model was divided into parts in order to decrease the problem of dimensionality. Each part of the model was optimized separately, but sequentially using the optimal conditions from the previous optimization procedure. The division corresponds to the following sections: 1- Feed Preparation; 2- Extractive Distillation and Solvent Recovery; 3- Fractionating. Figure 10 shows schematically this procedure.
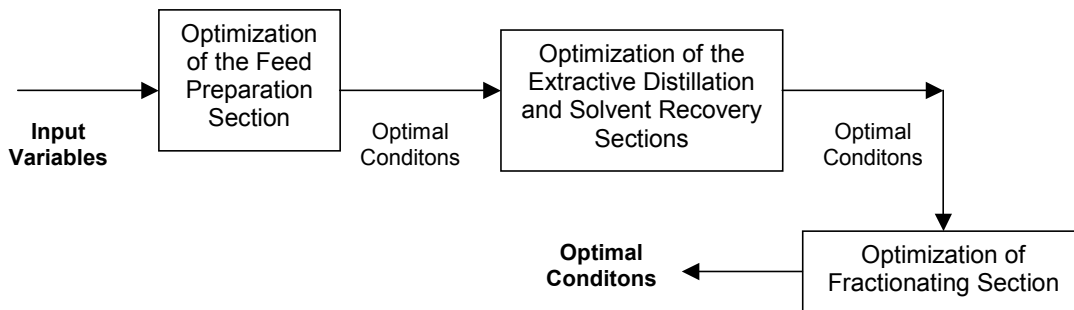

Fig. 10. Optimization Procedure

From this way, initially was performed the optimization of the Feed Preparation Section, i.e., NN 1, NN 2 and NN 3, followed by the optimization of the Extractive Distillation and Solvent Recovery Section, i.e., NN 4, NN 5, NN 6, NN 7 and NN 8 and finally it was performed the optimization of the Fractionating Section, i.e., NN 9 and NN 10. The optimal values generated at each step of the optimization procedure were used as input data for the next step. The optimization procedure at each step was carried out as described previously, i.e., the model built is used to generate a detailed grid search of the region of interest, by a full mapping of the objective functions on the space of decision variables.

The optimization procedure at each step was carried out as described previously, i.e., the model built is used to generate a detailed grid search of the region of interest, by a full mapping of the objective functions on the space of decision variables. For this, each input variable (NVAR) is divided into an equal number of intervals and by according to number of intervals (NSTEP), the neural network model generates a number of cases equal to $NVAR^{NSTEP}$. From these, only the cases that satisfy the specified constraints are selected. From these, the most convenient are selected in order to satisfy the criteria of a less expensive operation, i.e., the lowest energetic cost (qualitative optimization). In general, numerous cases that satisfy the established restriction are observed, but the final solution of each case studied presents only a few operational conditions that satisfy the criteria of a less expensive operation. Of course, many times, to satisfy all these conditions is not possible, and to make use of some hierarchical rules is necessary.

RESULTS AND DISCUSSION

As an example, the Table 5 shows the input variables used as the base case for the optimization study.

Table 5. Input variables for the Optimization Procedure – Base Case

x1(1)	13,99	**x5(1)**	64,67	**x8(1)**	4,39
x1(2)	20,72	**x5(2)**	88,40	**x8(2)**	78,26
x1(3)	14,90	**x5(3)**	1,00	**x8(3)**	0,79
x1(4)	59,80	**x5(4)**	90,89	**x8(4)**	1,63
x2(1)	13,50	**x5(5)**	14,02	**x9(1)**	2,67
x2(2)	62,61	**x5(6)**	15,01	**x9(2)**	29,05
x2(3)	17,33	**x5(7)**	5,54	**x9(3)**	13,05
x2(4)	4,79	**x6(1)**	63,34	**x9(4)**	53,92
x2(5)	44,24	**x6(2)**	28,48	**x9(5)**	102,28
x2(6)	9,40	**x6(3)**	6,71	**x9(6)**	70,68
x3(1)	4,89	**x6(4)**	17,65	**x9(1)**	2,67
x3(2)	2,80	**x6(5)**	47,83	**x9(2)**	29,05
x3(3)	65,64	**x6(6)**	1,18	**x9(3)**	13,05
x3(4)	1,08	**x6(7)**	92,33	**x9(4)**	53,92
x4(1)	8,24	**x7(1)**	7,35	**x9(5)**	102,28
x4(2)	24,37	**x7(2)**	2,55	**x9(6)**	70,68
x4(3)	2,87	**x7(3)**	1,46	**x10(1)**	61,58
x4(4)	1,06	**x7(4)**	1,38	**x10(2)**	2,34
x4(5)	69,91	**x7(5)**	2,45	**x10(3)**	13,22
x4(6)	12,75	**x7(6)**	2,50	**x10(4)**	13,05
x4(7)	3,98			**x10(5)**	18,16
				x10(6)	64,19

Optimization of the Feed Preparation Section

The first step is to give the feed specification to the reactor, i.e., total flow rate and feed concentration of isoprene (IP) and CPD. Then the mapping of solution is initiated by changing the manipulated variables and at this step only the solutions, which satisfy some operational conditions are accepted. These conditions are CPD outlet concentration and IP loss from the reactor; IP, CPD and 2M2B concentrations in the overhead product from the distillation column and the temperature of the bottom of the depentanizer column as shown at the Table 5. All of these variables must be in a given range as specified by the personnel of BRASKEM

The next step is to choose qualitatively the operational conditions, which could minimize the energy costs, i.e., conditions that lead to a less expensive operation. These conditions are the lowest reflux flow, the lowest reboiler steam flow and the lowest feed temperature. The low temperature is also important because it avoids dimerization reactions, which increases IP loss and this is a not desirable situation.

Another point to be observed is the great variability of the feed conditions, which leads a different optimal operational condition.

Table 6 shows the number of situations that satisfy all these conditions for the section 1.

Each one of the best operational conditions selected in order to satisfy all requirements are then considered as input data for the next section optimization procedure, i.e., extractive distillation and Recovery Solvent Sections represented by the set of followings neural network: NN 4, NN 5, NN 6, NN 7 and NN 8.

Table 6. - Grid Search Results for Feed Preparation Section

	Number of conditions
Feasible Optimal Solution	4100
Process Conditions Satisfied	255
Qualitative Constraints Satisfied	5
NSTEP	4

NSTEP- Number of intervals

Optimization of the Extractive Distillation and Solvent Recovery Sections

For optimizing this section, the mapping of solution is initiated by changing the manipulated variables and only the solutions that satisfy the requested process conditions are accepted. The next step was to choose qualitatively the operational conditions that lead to less operational costs or less energy consumption. These conditions are: the lowest reflux flow, the lowest reboiler steam and the lowest temperatures. The flow rate conditions that maximize the isoprene product are also considered.

Each different optimal condition from the previous step provided a new set of optimal conditions for the current optimization step.

Table 7 shows the number of situations that satisfy all these conditions for the section 2.

Table 7. - Grid Search Results for Extractive Distillation and Solvent Recovery Section

	Number of conditions				
	Case 1	Case 2	Case 3	Case 4	Case 5
Feasible Optimal Solution	60000				
Process Conditions Satisfied	17658	12447	4552	7380	4837
Qualitative Constraints Satisfied	5	5	6	5	7
NSTEP	3				

NSTEP- Number of intervals

Each one of the best operational conditions selected in order to satisfy all requirements for this section are then considered as input data for the next section optimization procedure, i.e., the Fractionating Section represented by the set of followings neural network: NN 9, NN 10.

Optimization of the Fractionating Section

For optimizing this section, the mapping of solution is initiated by changing the manipulated variables and only the solutions that satisfy the requested process conditions are accepted. The next step was to choose qualitatively the operational conditions that lead to less operational costs or less energy consumption. These conditions are: the lowest reflux flow, the lowest reboiler steam and the lowest temperatures. The flow rate conditions that maximize the isoprene product are also considered.

In this optimization, some optimal cases from the previous section was rejected due to relation solvent/feed is higher then the process ideal value. This situation was observed for the cases with low feed flow rate. From the remaining optimal cases for the sections 1 and 2, it was chosen only some conditions in order to carry out the optimization of this section.

Table 8 shows the number of situations that satisfy all these conditions for the section 3.

Table 8. - Grid Search Results for the Fractionating Section

	Number of conditions
Feasible Optimal Solution	1024
Process Conditions Satisfied	256
Qualitative Constraints Satisfied	3
NSTEP	4

NSTEP- Number of intervals;
Cases from previous sections: 3-2; 4-2; 4-4; 5-1

CONCLUSION

For the studied case, the adopted methodology was to construct a global model as a sequential simulator by interconnecting the individual neural networks corresponding to the unit operations involved in the process. This approach shows very good results and makes possible a good understanding of the process behavior.

The simulation results of the global model show a good coherence with the industrial plant behavior, which evidences the possibility to employ the developed methodology to represent accurately the process of isoprene production studied.

The optimization procedure based on a grid search of the problem using the neural network as an industrial plant model has an advantage of mapping all the solutions. This procedure allows one to define the domain where the optimal solution is located. Quantitative and qualitative solutions can be promptly analyzed. To have an explicit objective function in order to get optimal operation points is not necessary since it is possible to work only with inequalities.

The results obtained from this work provide suggestions for the feasible operational conditions to apply to real unit, and by this way it is possible to demonstrate the practical features of this study.

Acknowledgements

Appendix A. Notation

ALFA	Mommentum Parameter
ETA	Dumping Factor
NCMAX	Number of iterations
NEXP learn	Number of points for the learning set
NEXP test	Number of Points for the test set
NH	Number of neurons in the hidden layer
NO	Number of neurons in the output layer
NVAR1	Number of Input Variables
NVAR2	Number of Output Variables
2M2B	2-Methyl-Butene-2

C5	Hydrocarbon with number of carbons equal to five
CPD	Cyclopentadiene
DCPD	Dicyclopenadiene
H2O	Water
i-C5	Isopentane
IP	Isoprene
n-C5	n-Pentane
X	Input Variables
Y	Output Variables

References

Alves, R.M.B. and Nascimento, C.A.O. (2001). Gross Errors Detection of Industrial Data by Neural Network and Cluster Techniques. *Proceedings of the ENPROMER 2001, Santa Fé-Argentina.*

Alves, R.M.B. and Nascimento, C.A.O. (2002). Gross Errors Detection of Industrial Data by Neural Network and Cluster Techniques. *Braz. J. Chem. Eng.*, vol.19, no.4, p.483-489.

Alves, R.M.B, Nascimento, C.A.O. and Carneiro, W.L. (2003). Analysis and Detection of Outliers and Systematic Errors from an Industrial Data Plant, *Proceedings of the FOCAPO 2003, Coral Springs, USA*.

Guardani, R., Onimaru, R.S. and Creso, F.C.A. (2001), Neural Network Model for the On-Line Monitoring of a Crystallization Process. *Brazilian Journal of Chemical Engineering*, v. 18(3), 267-275.

Haykin, S., (1994). Neural Networks – A Comprehensive Foundation. *Macmillan College Publisihing Co.*, New York.

Migliavacca, S.C.P., Nascimento, C.A.O. & Rodrigues, C., (1999). Use of Neural Network for the Simulation of a Gas Centrifuge. *J. Nuclear Science. & Technology*. 36(4), 364-370.

Pollard, J.F.; Broussard, M.R.; Garrison, D.B.; San, K.Y., Process Identification Using Neural Networks, Computers Chem. Engng., 1992, 16, pp. 253-270.

Nascimento, C.A.O., (1991). Programa de Redes Neurais, Neuro14, v.14, *Departamento de Engenharia Química da Escola Politécnica da Universidade de São Paulo*-Brazil

Nascimento, C.A.O., Oliveros, E. and Braun, A.M., (1994). Neural Network Modeling for Photochemical Process. *Chem. Engineering and Processing*, 33, 319-324.

Nascimento, C.A.O., Guardani, R. and Giulietti, M. (1997). Use of Neural Networks in the Analysis of Particle Size Distribution by Laser Diffraction. *Powder Technol.*, v.80 (1), 89-94.

Nascimento, C. A. O. and Giudici, R., (1998). Neural Network Based Approach for Optimization Applied to an Industrial Nylon-6,6 Polymerization Process. *Comput. Chem. Eng.*, 22, S595-600.

Nascimento, C.A.O., Giudici, G. and Scherbakoff, N. (1999). Modeling of Industrial Nylon-6,6 Polycondensantion Process. Part 2: Neural Networks and Hybrid Models. *J. Applied Polymer Science*, 72, 905-912.

Nascimento, C.A.O., Giudici, R., and Guardani, R. (2000). Neural Network Based Approach for Optimization of Industrial Chemical Processes. *Comput. Chem. Eng.*, 24, 2303-2314.

Nascimento, C.A.O ; Alves, R.M.B., (2003). Programa para Simulação do Processo de Produção de Isopreno via Redes Neurais, v.1. *Departamento de Engenharia Química da Escola Politécnica da Universidade de São Paulo.*

Nascimento, C.A.O; Alves, R.M.B., (2003). Programa para Otimização do Processo de Produção de Isopreno via Redes Neurais, v.1. *Departamento de Engenharia Química da Escola Politécnica da Universidade de São Paulo.*

Rumelhart, D.E.; McClelland, J.L., (1986). Parallel Distributed Processing: Explorations in the Microestructure of Cognition. *MIT Press*, Cambridge, MA.

Seider, W.D., Seader, J.D. & Lewin, D.R., (1999) Process Design Principles: Synthesis, Analysis, and Evaluation, *John Wiley & Sons, Inc*.

Sternowsky, S, Schulte G., Guardani R. and Nascimento C.A.O. (2002). Experimental Study and Neural Network Modeling of the Ligament Disintegration in Rotary Atomization. *Atomization and Sprays*, v. 12, 1-15.

Willis, M.J., Di Massimo, C., Montague, G.A., Tham, M.T. and Morris, A.J., (1991). Artificial Neural Networks in Process Engineerings. *IEE Proceedings-D*, 138 (3), 256-266.