I&EC research
Industrial & Engineering Chemistry Research
Article
pubs.acs.org/IECR

Optimization of Heat-Integrated Crude Oil Distillation Systems. Part I: The Distillation Model

Lluvia M. Ochoa-Estopier and Megan Jobson*

Centre for Process Integration, School of Chemical Engineering and Analytical Science, The University of Manchester, Manchester, M13 9PL, U.K.

* Supporting Information

ABSTRACT: This work presents a methodology for optimizing heat-integrated crude oil distillation systems. Part I of this three-part series presents a modeling strategy where artificial neural networks are used to represent the distillation process. Part II presents a new methodology to retrofit heat exchanger networks (HENs) and Part III presents the application of this distillation model to perform operational optimization of the crude oil distillation unit while proposing retrofit modifications to the associated HEN. Independent variables of the distillation model include flow rates of products, stripping steam, pump-around specifications, and furnace exit temperature. Dependent variables include those related to product quality, and temperatures, duties, and heat capacities of process streams involved in heat integration. The resulting neural network model is able to overcome convergence problems presented by rigorous or simplified models. Simulation time is significantly improved using neural networks, compared to rigorous models, with practically no detriment to model accuracy.

1. INTRODUCTION

Crude oil distillation is a very important step in the refining process. The refining process begins with distillation, where crude oil is separated into a series of fractions based on boiling ranges. These distillation products are further processed in downstream operations (e.g., hydrotreating, hydrocracking, catalytic reforming, fluidized catalytic cracking, etc.) before being blended into final products. Crude oil distillation systems typically consist of a preheat train, a flash unit or a prefractionator, an atmospheric distillation unit, and a vacuum distillation unit, as illustrated in Figure 1. The distillation units are configured as a main distillation column with side-strippers and intermediate coolers (i.e., pump-arounds or pump-backs). Crude oil first enters a set of heat exchangers (i.e., the preheat train) before entering the atmospheric distillation unit, where the crude oil is fractionated. The heaviest product stream from the atmospheric distillation unit, the atmospheric residue, is further separated in the vacuum distillation unit. A flash unit or a prefractionator may also be present upstream of the atmospheric distillation unit.

Crude oil distillation is an energy-intensive process. It is reported[3] that the energy consumed in the overall refining process is equivalent to between 7% and 15% of the crude oil processed, of which 35−45% is consumed by the atmospheric and vacuum distillation units. Heat integration is vital for an energy-efficient operation of the crude oil distillation process. Heat integration is achieved by exchanging heat between hot and cold process streams. The main cold process stream, the crude oil feed, is heated to an intermediate temperature by cooling distillation process streams, such as the pump-arounds and product streams. The crude oil then enters a furnace to reach the required processing temperature. Increasing energy recovery between process streams reduces fuel consumption in the furnace, thus operating costs. A flash unit or a prefractionator may be placed before the crude oil furnace, which can also help to reduce energy consumption.

Increasing concerns related to carbon emissions and process economics, such as the rise in fuel prices, have motivated the implementation of grass-roots design, retrofit, and operational optimization projects aimed to improve the energy and separation performance of distillation systems. Operational optimization is more frequently implemented than retrofit, while grass-roots design projects are carried out least often. Methodologies for grass-roots design, retrofit, and operational optimization purposes have been developed. These design methodologies can either focus on the distillation process, the heat exchanger network, or the heat-integrated distillation system (i.e., distillation process and HEN). Considering the distillation process and HEN together allows interactions within the system to be exploited, increasing the chances of finding better designs, compared to considering the distillation process and HEN separately.[4]

Early design procedures[5,6] used simple mass and energy balances, guidelines, and empirical correlations to design distillation units. Over time, new computational tools have facilitated the development of design methodologies that employ optimization algorithms at some point of the design procedure. When the scope of these methodologies is the design of heat-integrated crude oil distillation systems, models for the crude oil distillation process and heat recovery network need to be considered in the optimization framework. This paper reviews crude oil distillation models, while Part II[1] of this series discusses HEN design models.

Received: September 25, 2014
Revised: March 13, 2015
Accepted: March 23, 2015
Published: March 23, 2015

DOI: 10.1021/ie503802j
Ind. Eng. Chem. Res. 2015, 54, 4988−5000

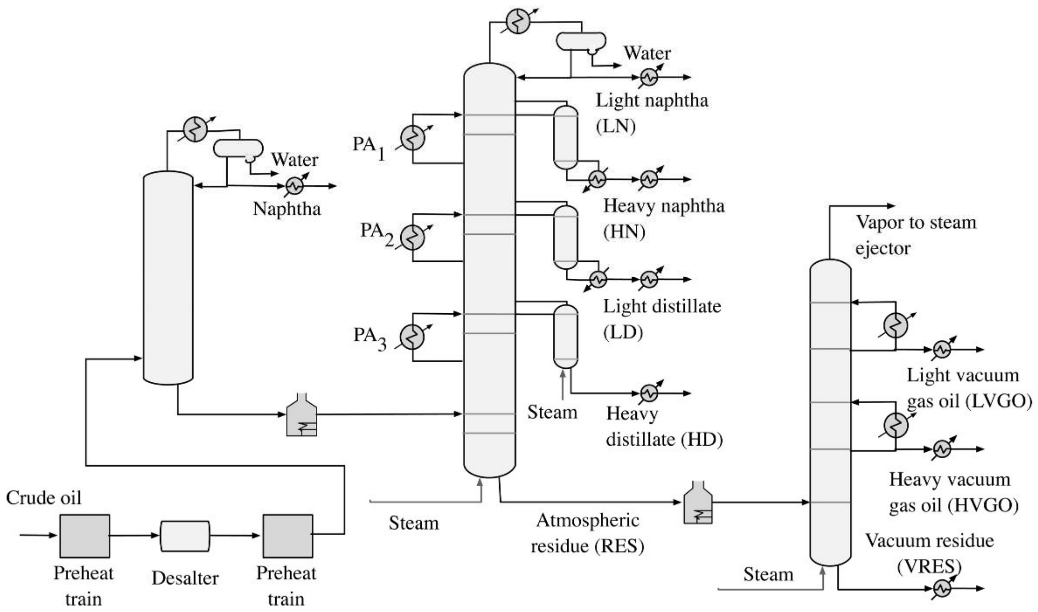


Figure 1. Crude oil distillation system.

Crude oil distillation models can be classified in three main categories: rigorous, simplified and statistical models.[7] These models have been incorporated in approaches that consider grass-roots design, retrofit or operational optimization. Rigorous distillation models were used in the work of Liebmann et al.[8] and Bagajewicz and Ji[9,10] to design heat-integrated crude oil distillation units. Starting from initial column designs, their approaches use guidelines and composite curves or heat diagrams to propose modifications to the distillation units that increase energy recovery. Liebmann et al.[8] focused on selecting the vaporization mechanism (stripping steam or reboiling), pump-around duties, and degree of thermal coupling. Bagajewicz and Ji[9,10] dealt with the placement and duties of pump-arounds for columns that process different types of crude oils. These design procedures[8−10] are iterative; trial and error is required to achieve the final designs.

Rigorous models have also been used to perform operational optimization of crude oil distillation units.[11−14] However, no HEN details were considered in these methodologies. HEN details refer to the topology of the HEN (i.e., connections between heat exchanger units and stream splitters in the HEN), available heat transfer area of each unit, overall heat transfer coefficients, etc. Seo et al.[11] optimized the feed stage location and pump-around duties to reduce operating and capital costs. Basak et al.[12] developed a methodology to perform operational optimization to increase product revenue while reducing operating costs. Inamdar et al.[13] solved multiobjective design problems, such as maximizing total profit while minimizing energy costs. Optimization variables have included flow rates of products, pump-arounds and column reflux, and coil outlet temperature[12,13] (furnace exit temperature). Al-Mayyahi et al.[14] performed multiobjective optimization to maximize net profits and minimize CO_2 emissions. Crude oil blending fractions, stripping steam flow rates, and pump-around duties were considered as optimization variables.

Chen[4] used simplified models based on the Fenske−Underwood−Gilliland (FUG) method to perform grass-roots design, retrofit and operational optimization of crude oil distillation systems. In the work of Chen,[4] structural variables of the distillation column (e.g., stage distribution, location of feed stage, and pump-arounds) and HEN details are taken into account. Simplified models based on the FUG method have the advantage of being more robust and converging faster than rigorous models. However, simplified models are also highly sensitive to initial guesses.

Statistical models are extremely robust and are simpler than rigorous and simplified distillation models. Many approaches are available to develop statistical models; these include linear regression, polynomial regression, support vector regression, artificial neural networks (ANNs), etc. Liau et al.[15] proposed a methodology to perform operational optimization of a distillation unit using ANN models. In their work, the ANN distillation model is regressed against plant measurements. Operational optimization is applied to increase the yields of diesel or kerosene products. No HEN details are considered in the approach of Liau et al.[15]

Yao and Chu[16] used a nonlinear regression method known as support vector regression to model crude oil distillation units. The distillation model was implemented into an optimization framework to find the operating conditions (e.g., pump-around specifications and flow rates of products and stripping steam) that maximize profits. Lopez C. et al.[17] regressed second-order polynomial functions to model a system of several crude oil distillation units. These models were used to perform operational optimization to maximize profits. Optimization variables comprised crude oil blending fractions, pump-around specifications, coil outlet temperature, and flow rates of products and stripping steam. Mass and energy balances were used in the work of Lopez C. et al.[17] to simulate the associated HEN and constrain the inlet and outlet temperatures of the heat exchangers.

The artificial neural networks concept is the most widely used framework in recent years to develop nonlinear models.[18] Many applications of ANN models in chemical engineering can been found; reviews of such applications are presented by Pirdashti et al.[19] and Himmelblau.[20] ANN models have been developed in the crude oil refining industry to simulate processing units or predict properties. Processing units represented by ANN models include crude oil distillation,[15] fluid catalytic cracking,[21,22] hydrodesulfurization,[22,23] catalytic

DOI: 10.1021/ie503802j
Ind. Eng. Chem. Res. 2015, 54, 4988−5000

reforming,[24] delayed coking [25] and hydrocracking. [26] ANN models have also been used to predict properties of crude oil[27] and distillation side-draws;[28] and to predict fouling in crude oil heat exchangers.[29]

Models to be used for design and optimization of heat-integrated crude oil distillation systems need to be robust and accurate. Robustness refers here to the ability to find solutions to the model equations when their inputs are varied. Accuracy refers to the agreement between model predictions and the actual process represented by the model. Additionally, these models should provide sufficient information to represent the synergy between the distillation process and heat recovery network. In this context, ANN distillation models have the advantage of being extremely robust, compared to rigorous and simplified models.[7] The accuracy of ANN models depends mainly on the data used for regression and the ability of the designer to correctly choose the ANN model parameters (e.g., type of regression function, number of neurons, etc.). If this is the case, ANN model predictions can be as or more consistent with the actual process as rigorous or simplified distillation models. Furthermore, ANN distillation models can be tailor-made to only include the variables that represent the separation and energy performance of the heat-integrated crude oil distillation system. Thus, it is possible to avoid the use of complex formulations found in rigorous and simplified models.

This paper presents a methodology to develop ANN models for crude oil distillation units. The resulting distillation model considers operational variables such as coil outlet temperature, pump-around temperatures, and flow rates, and flow rates of distillation products and stripping steam. The ANN distillation model also calculates the information required to simulate the associated HEN. An example is presented to illustrate the accuracy and computational performance of the resulting ANN model and its application to optimize a distillation unit considering a simple heat recovery model (i.e., the grand composite curve).

Part II[1] of this series describes an approach to simulate and retrofit HENs considering network details (e.g., structure, heat transfer areas, etc.), while Part III[2] presents the implementation of the resulting distillation and HEN models to optimize heat-integrated crude oil distillation systems. The overall method-ology described in this series focuses on operational optimization of the distillation process, taking into account the necessary retrofit modifications to the associated HEN.

2. ARTIFICIAL NEURAL NETWORKS

Artificial neural networks are statistical modeling tools inspired by the structure of biological neural networks. ANNs relate the independent variables or inputs with the dependent variables or outputs through a set of neurons organized into layers. A neuron is the basic computing unit in the network. Even though many ANN models are similar to established statistical models, the terminology used to describe ANNs is very different to that used in statistical modeling. Sarle[30] provides a comparison between ANN and statistical modeling terminology.

Many ANN structures or architectures can be found in the literature.[31] Feed-forward networks are the most commonly used network arrangements and are the ones used in this work. Figure 2 illustrates a feed-forward artificial neural network consisting of two layers, namely a hidden layer and an output layer.

Each layer comprises one or more neurons, with their corresponding transfer f unctions. Feed-forward networks can

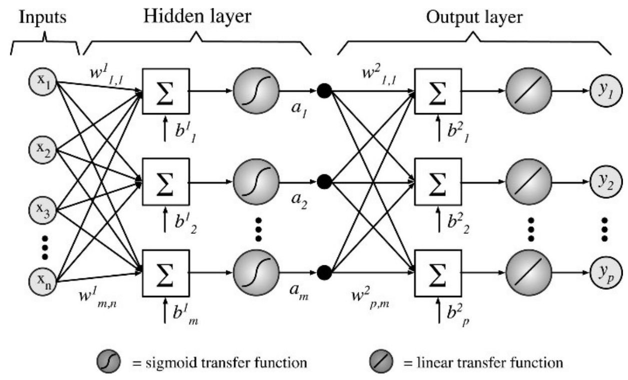


Figure 2. Structure of a feed-forward neural network.[33]

contain more than one hidden layer, but the most typically used networks consist of only one hidden layer. It is found[30,32] that networks with one or more hidden layers are universal approximators, that is, they are able to represent any function to any desired degree of accuracy. This work employs networks with two hidden layers and one output layer. Equations 1 and 2 represent the feed-forward network depicted in Figure 2:

$$\mathbf{a} = f^1\,(\mathbf{W}^1\mathbf{x} + \mathbf{b}^1) \tag{1}$$

$$\mathbf{y} = f^2\,(\mathbf{W}^2\,\mathbf{a} + \mathbf{b}^2) \tag{2}$$

where x is the vector of independent variables, and a and y are the output vectors of the hidden and output layers, respectively. W is the matrix of weights, and b is the vector of biases. Superscripts 1 and 2 indicate the hidden and output layers, respectively. Transfer function f^1 is a hyperbolic tangent function while f^2 is an identity function (i.e., $g(x) = x$). This selection of transfer functions is suitable for data fitting, while using hyperbolic tangent functions in f^1 and f^2 is suitable for pattern recognition.[33]

The weights and biases required to calculate the outputs are determined by a training algorithm, which is an algorithm that regresses the model equations against a set of samples or targets. The most common training algorithm is called backpropagation and is the one used in this work. This type of training algorithm minimizes the mean squared errors between model predictions and their targets using a gradient descent optimization method:[34]

$$F = \frac{1}{N}\sum_{i=1}^{N}(t_i - y_i)^2 \tag{3}$$

where t is the vector of targets and N is the number of observations used to train the model.

The neural network modeling process involves the following steps: (a) data collection, (b) data processing, (c) model selection, (d) training and validation.[33] Data collection is vital to create accurate models, since samples are the only source of information used to represent the process being modeled. Neural networks are not able to accurately extrapolate beyond the upper and lower limits of the data set used for training. Therefore, it is important that the samples cover the range of input values for which the network will be used. Data processing, particularly normalization, is applied in this work to facilitate the training process.

Model selection refers to specifying the structure of the neural network (i.e., number of layers, neurons, and choice of transfer functions). It is reported[35] that networks with more than two hidden layers generally do not perform better, in terms of accuracy, than networks with fewer hidden layers. Furthermore, it becomes more difficult to train the network as the number of layers and neurons increases.[33]

There are no universally valid guidelines to choose the number of neurons in a neural network. Some authors suggest specifying the number of neurons depending on the sizes of the input and output vectors,[36,37] or as a fraction of the number of samples used for training.[38] Methodologies to automatically determine the ANN structure have also been developed. These methodologies include pruning and growing the networks,[37,39,40] and modifying the ANN structure using genetic algorithms.[41,42] Training is the process of calculating the weights and biases of the ANN equations using an optimization algorithm. The training algorithm employed in this work divides the data into three sets, namely training, validation and test sets. The training set is used to calculate the weights and biases, while the validation set prevents overfitting of the network. The test set is used after training to validate the resulting model.

3. DISTILLATION MODELING APPROACH

A modeling framework using neural networks is proposed in this work to facilitate optimization of heat-integrated crude oil distillation systems. This modeling framework can also be applied to simple distillation columns.

3.1. Data Collection and Processing.

Data collection is the first step of the neural network design process. The inputs are selected in this work as variables that can be manipulated during operation and that impact on heat recovery and product revenue. Thus, the inputs of the ANN model consist of flow rates of distillation products and stripping steam, coil outlet temperature, and pump-around specifications. Other variables can also be included, such as feed flow rate or crude oil bending ratio. The outputs of the ANN model include the variables needed to evaluate the objective function and constraints. In this work, the outputs comprise variables related to product quality (e.g., true boiling point temperatures and flow rates), to column flooding conditions, and to the simulation of the heat exchanger network, in particular temperatures, enthalpies, and heat capacities of process streams passing through the heat exchanger network.

Samples to build statistical models can be obtained from plant measurements[15] or from other models.[16,17] Plant measurements often present measurement errors and process variability over time, which need to be accounted for by processing the data. Furthermore, models applied for optimization purposes need to include scenarios that may differ from current and past operation. Models regressed with plant measurements can fail to represent these new operating scenarios.

One way to overcome these limitations is using samples from rigorous simulations to build the ANN distillation model. Even though this approach involves solving the highly nonlinear system of equations repeatedly to obtain the samples, the resulting ANN model is simple, accurate, and extremely robust. These model characteristics are desirable for implementation in optimization. In this work, rigorous models embedded in the commercial simulation package Aspen HYSYS (V7.3)[43] were used to obtain the samples.

The sampling method selected in this work is the Latin Hypercube Sampling (LHS) method. Other types of sampling methods are also available, such as Monte Carlo sampling and Hammersley sequence sampling.[44] LHS is a stratified sampling method that can be applied to multiple variables.[45] Consider a set of variables $x_1, x_2, x_k, \ldots, x_K$ for which N samples need to be created. The range of each variable is divided into N equal intervals. Then, one value from each interval is randomly selected. For each sample ($i = 1,2,\ldots,N$), the values of x_k are determined by eq 4:[46]

$$x_{k,i} = P^{-1}\frac{i - 1 + \xi}{N} \tag{4}$$

where ξ is a random variable such that $\xi \in [0,1]$ and is different for each value i of variable k. P is the cumulative probability distribution of x_k. Equation 4 produces sample values $x_{k,i}$ scaled to the interval [0,1]. The N values obtained for variable x_1 are then paired at random with each N values of variables $x_2, x_k, \ldots, x_K$.

The LHS method is applied to the inputs of the ANN model (i.e., flow rates of products and stripping steam, coil outlet temperature, and pump-around specifications) to obtain the sampling plan. The samples were obtained through rigorous simulations in Aspen HYSYS (V7.3), while the distillation model was developed using the Artif icial Neural Network Toolbox in MATLAB.[47] An automation code was developed in MATLAB to assist data collection from Aspen HYSYS (V7.3). This automation code works as an interface between both softwares, allowing MATLAB to automatically run Aspen HYSYS (V7.3) and collect the simulation results. Note that, in practice, the rigorous model needs to be validated first with the actual process before being used to generate the samples.

3.2. ANN Structure.

The large number of outputs requires several neural networks to create the distillation model. To facilitate training of the ANNs, the outputs are grouped according to their dimensions and order of magnitude. For example, output group 1 for ANN_1 consists of stream temperatures, output group 2 for ANN_2 consists of heat capacities, etc. These neural networks are structured as feed-forward networks with two hidden layers. The transfer function for the hidden layers is a hyperbolic tangent function, while an identity function is used for the output layer.

The number of neurons in the hidden layers of each network is determined by manually growing the network and selecting the number of neurons that achieves the desired accuracy. The accuracy indicator used in this work to select the number of neurons is the coefficient of determination R^2.

Another ANN was built to predict whether a specified set of inputs constitute a feasible operating scenario. In this work, inputs are deemed to be feasible if they result in converged rigorous simulations in Aspen HYSYS (V7.3). This feasibility ANN is a feed-forward network with one hidden layer and one output layer. The output of this pattern recognition network is an integer that has a value of 0 for unfeasible inputs, and a value of 1 for feasible inputs. Hyperbolic tangent functions are used for the hidden and output layers. The number of neurons in the hidden layer was also selected by manually growing the network.

3.3. ANN Training and Validation.

Backpropagation training is applied in this work to calculate the values of the weights and biases that minimize the error between model predictions and targets. This training method minimizes eq 3 using an optimization algorithm. The ANNs related to the distillation model (i.e., networks that calculate variables related to product quality, column diameters, and stream information for HEN simulation) use the Levenberg−Marquardt algo-rithm[48,49] to minimize eq 3. For the feasibility ANN, the scaled

DOI: 10.1021/ie503802j
Ind. Eng. Chem. Res. 2015, 54, 4988−5000

conjugate gradient algorithm[50] is employed to minimize eq 3. During training, the samples obtained from rigorous simulations are randomly divided into three sets. The ratios for the training, validation, and testing sets are 0.70, 0.15, and 0.15, respectively.

The goodness of fit of the ANNs of the distillation model is assessed with the coefficient of determination R^2 and residual analysis. The residuals of each output are plotted against their targets to identify nonrandom patterns, which may indicate a poor regression. The mean and standard deviation of the errors between ANN predictions and their targets for each output are also calculated to provide quantitative measures of confidence intervals. In this work, the absolute error is calculated for the temperature-related variables, while the relative error is calculated for variables other than temperatures.

The goodness of fit of the feasibility ANN is assessed with a conf usion matrix. This confusion matrix reports the number of false positive, false negative, true positive, and true negative predictions, which allows a more detailed analysis of accuracy of the feasibility ANN.

3.4. Modeling Temperature-Dependent Heat Capaci-ties. Heat integration is achieved in the crude oil distillation system by exchanging heat between hot and cold process streams. To obtain a reasonable representation of the energy efficiency of the distillation process, details of the heat exchanger network need to be taken into account. These details include the topology of the network, heat loads, heat transfer areas, and heat transfer coefficients of each exchange unit, etc. Design and simulation of a detailed HEN also requires that at least supply and target temperatures, heat capacities, and enthalpies from each process stream are known. The terms supply and target temperatures refer to stream temperatures before and after passing through the HEN, respectively.

HEN design approaches typically implement simplifying assumptions, such as considering constant thermal properties for all process streams. In particular, heat capacity is often assumed constant. This assumption is valid when streams do not undergo phase changes, considerable temperature changes, or when heat capacity is not significantly affected by temperature. However, process streams from crude oil distillation experience significant temperature variations that affect heat capacity, among other properties (e.g., density, viscosity, etc.). For this reason, it is necessary to develop a model for calculating heat capacity as a function of temperature.

Three situations occur for process streams exchanging heat in the HEN: (a) streams that only exchange sensible heat, (b) streams that mostly exchange latent heat, and (c) streams that exchange sensible and latent heat. For the crude oil distillation system considered in this work, the process streams that transfer only sensible heat (a) are the products and pump-around streams. It is assumed that phase equilibrium is achieved on each stage in the column. Thus, the products leave the distillation unit as saturated liquids. These products are cooled down to be stored or processed downstream. Likewise, pump-around streams are withdrawn from the main column as saturated liquids and returned as subcooled liquids to an upper stage.

Streams that absorb heat through experiencing phase change (b) are process streams undergoing vaporization in reboilers. Similarly, if the outlet of the condenser is specified as a vapor or saturated liquid, then this process stream belongs to situation (b). In contrast, if the outlet of the condenser is specified as a subcooled liquid, then the process stream in the cooler rejects both latent and sensible heat (c). Crude oil feed belongs to situation (c), increasing its temperature from storage temperature to approximately 330−385 °C, depending on crude composition. The crude oil is gradually vaporized throughout the temperature range before entering the flash zone of the main distillation tower.

Chen[4] used multisegmented stream data to represent temperature-dependent thermal properties, namely stream enthalpy. For each stream, a new segment was generated each 40 °C and when a phase change took place. Then, a fourth-order polynomial was regressed for each stream using the segmented stream data. The fourth-order polynomials used by Chen[4] correlated temperature as a function of enthalpy. One of the shortcomings of this approach is that the fourth-order polynomials are not suitable for representing streams of situation (c). Moreover, it is computationally demanding to regress the parameters of the fourth-order polynomials each time a simulation is run. Another shortcoming of the approach of Chen[4] is that by using these fourth-order polynomials, the HEN energy balance becomes a system of nonlinear equations, which needs to be solved sequentially.

In this work, different heat capacity flow rate models are considered for each situation. An understanding of the relationship between temperature and enthalpy was gained using rigorous simulations to select each type of model (e.g., linear, polynomial, etc.). Thus, simple linear heat capacity models are developed for streams that exhibit a linear dependence with respect to temperature. Similarly, nonlinear equations are used for streams that exhibit nonlinear relationships between heat capacity flow rates and temperatures. This approach is simpler and more computationally efficient than the approach of Chen.[4] Moreover, the heat capacity correlations developed in this work can be implemented in the HEN energy balance without altering its linearity. Thus, the HEN energy balance can still be solved as a system of linear equations. Part II[1] of the series describes how the heat capacity models are implemented in the HEN model.

It was found that a linear equation is adequate to represent temperature-dependent heat capacity of products and pump-around streams (a). For these streams, the ANN distillation model provides predictions of supply and target temperatures, enthalpy change, and the ratio between heat capacities at supply and target temperatures. Thus, the heat capacity flow rate for case (a) is calculated using linear interpolation:

$$CP = \frac{CP_s(\psi - 1)}{T_t - T_s} T + \frac{CP_s(T_t - \psi T_s)}{T_t - T_s} \tag{5}$$

where T_s and T_t are the supply and target temperatures, respectively. CP_s is the heat capacity flow rate at the supply temperature, ψ is the ratio between heat capacities at supply and target temperatures (i.e., $\psi = Cp_t/Cp_s$), and CP is the heat capacity flow rate at a given temperature T. Equation 5 represents a straight line. CP_s is related to the mean heat capacity $\overline{CP}$ and total enthalpy change H:

$$\overline{CP} = \frac{H}{T_t - T_s} = \frac{CP_s + CP_t}{2} = \frac{CP_s(1 + \psi)}{2} \tag{6}$$

where CP_t is the heat capacity flow rate at the target temperature, H is the stream enthalpy change predicted by the ANN distillation model, and $\overline{CP}$ is the average heat capacity flow rate. The value of CP_s can also be expressed as

DOI: 10.1021/ie503802j
Ind. Eng. Chem. Res. 2015, 54, 4988−5000

Table 1. Input Variables of Atmospheric Distillation Unit

item	base case	lower bound	upper bound	optimized case	
LN flow rate (bbl/h)	465.9	326.1	605.7	498.9	(+7%)
HN flow rate (bbl/h)	483.6	338.5	628.7	462.6	(−4%)
LD flow rate (bbl/h)	921.9	645.3	1198.5	922.9	(~0%)
HD flow rate (bbl/h)	285.7	200.0	371.4	336.9	(+18%)
RES flow rate (bbl/h)	2009.6			1945.4	(−3%)
PA_1 duty (MW)	11.20	7.84	14.56	10.06	(−10%)
PA_2 duty (MW)	17.89	12.52	23.26	19.22	(+7%)
PA_3 duty (MW)	12.84	8.99	16.69	16.59	(+29%)
PA_1 temp drop (°C)	20.0	14.0	26.0	24.0	(+20%)
PA_2 temp drop (°C)	50.0	35.0	65.0	54.5	(+9%)
PA_3 temp drop (°C)	30.0	21.0	39.0	30.0	(~0%)
HD steam flow rate (kmol/h)	1200.0	840.0	1560.0	1213.0	(+1%)
RES steam flow rate (kmol/h)	250.0	175.0	325.0	261.8	(+5%)
coil outlet temp (°C)	365.0	330.0	370.0	366.9	(+2 °C)

Equation 5 is a simple, yet suitable representation of the relationship between heat capacity flow rate and temperature, and applies to distillation products and pump-around streams.

For the reboilers and condenser (where the outlet of the latter specified as a saturated liquid), a pseudo-heat-capacity flow rate CP* is introduced to represent latent heat in terms of sensible heat:

$$CP_S = \frac{2}{(1+\psi)}\frac{H}{(T_t - T_S)} \quad (7)$$

$$CP^* = \frac{H}{T_t - T_S} \quad (8)$$

Equation 8 has no physical meaning and is only used to represent latent heat in a manner that is consistent with the HEN model formulation. H represents the latent heat of vaporization and condensation for the reboilers and condenser, respectively. If the stream is a pure component and only a phase change occurs without any change in pressure, then $T_t - T_s = 0$ °C . In this case, the temperature needs to be corrected to a small value different from zero to avoid undefined CP* values, for example $T_t - T_s = 0.1$ °C.

Finally, a polynomial equation was developed for the crude oil feed. This equation predicts the relationship between enthalpy and temperature for the interval $[T_s,T_t]$. The equation was regressed using data of enthalpy H measured at different temperatures T. The reference enthalpy in eq 9 is the enthalpy of the crude oil measured at feed conditions. Aspen HYSYS (V7.3) was used to generate such data. It was found that a third degree polynomial equation provides an accurate representation:

$$H = p_4 T^3 + p_3 T^2 + p_2 T + p_1 \quad (9)$$

where p_i (for i = 1,...,4) are the model parameters found by least-squares regression. Considering constant pressure and mass flow rate, the heat capacity flow rate can be calculated as CP = dH/dT. Differentiation is applied to eq 9 to obtain a pseudo-heat-capacity flow rate that represents both sensible and latent heat:

Once the ANNs are regressed and validated, they can be used to simulate the distillation column and to calculate the stream heat capacity flow rates. The ANN distillation model, the feasibility ANN, and eq 5 to 10 provide all the information from the distillation process that is needed for optimization. For details on how eq 5 to 10 are implemented in the HEN model, the reader is referred to Part II[1] of this series.

4. CASE STUDY

The neural network modeling approach developed to enable operational optimization of the distillation column and presented in the previous section is illustrated by a case study. The distillation system comprises an atmospheric crude oil distillation unit and its associated heat exchanger network. However, this case study focuses only on the distillation unit. The case study demonstrates how the distillation model can be applied for operational optimization of the distillation column in a simplified study, where the heat exchanger network is not explicitly modeled but heat recovery opportunities are estimated using the grand composite curve (GCC). The model is applied for optimization, while the GCC is used in this paper for simplicity. Part III[2] uses a detailed model of an existing HEN in the optimization and retrofit of the crude oil distillation system.

The atmospheric distillation unit processes 100 000 bbl/day (0.184 m^3/s) of crude oil into five products, namely light naphtha, heavy naphtha, light distillate, heavy distillate, and residue. The crude oil mixture is Venezuela Tia Juana Light crude;[6] the true boiling point curve and properties are presented in Tables S1 and S2 (see Supporting Information). The initial operating conditions are taken from Chen[4] based on a case study presented by Watkins.[6] This information is used in Aspen HYSYS (V7.3) to create a set of pseudocomponents that represent the crude oil mixture. The characterization technique embedded in Aspen HYSYS (V7.3) calculates the physical and thermodynamic properties (i.e., molecular weight, density, viscosity, vapor pressure, enthalpy, etc.) of each pseudocomponent. The normal boiling point, molecular weight, density, and liquid volume composition of these pseudocomponents are presented in Table S3 (see Supporting Information).

$$CP^* = \frac{p_4}{3}T^2 + \frac{p_3}{2}T + p_2 \qquad (10)$$

In summary, the ANN modeling approach starts by setting up a rigorous simulation, from which the samples are obtained. The samples are used to regress the parameters of the ANNs.

The atmospheric distillation unit is structured as a main column with three side strippers and three pump-arounds, as in Figure 1. Steam is used as a stripping agent for the main column and HD stripper, while reboiling is employed for the HN and LD strippers. The stage distribution of the distillation

DOI: 10.1021/ie503802j
Ind. Eng. Chem. Res. 2015, 54, 4988−5000

unit and initial operating conditions are based on an optimized design presented by Chen.[4] Table 1 and Table S4 (see Supporting Information) show the initial operating conditions and the structure of the distillation unit, respectively. The reflux ratio in the main distillation column at the initial operating conditions is 4.57. Operating pressure is 2.5 bar. Product flow rates in Table 1 are reported as ideal liquid flow rates at standard conditions (15 °C and 1 atm) on a water-free basis. Table S8 (see Supporting Information) presents the flooding conditions for the crude oil distillation unit at the initial operating conditions. These flooding conditions are calculated using the tray sizing utility in Aspen HYSYS (V7.3), with the default parameters for sieve trays.

4.1. Distillation Model. A rigorous simulation was set up in Aspen HYSYS (V7.3) using the information from the crude oil assay and the structure of the distillation unit. A total of 3000 sample points were obtained using the LHS technique. The lower and upper bounds for each independent variable are presented in Table 1. These sample points were simulated in Aspen HYSYS (V7.3) using the automation program developed in MATLAB. From the 3000 simulated scenarios, 2845 converged.

The results of the converged simulations were used to train the ANNs related to the distillation process information, while all 3000 sample points were used to train the feasibility ANN. The architecture of each neural network and the description of their inputs and outputs are presented in Tables S5 to S7 (see Supporting Information). Data collection was the most time-consuming step of the ANN model development, compared to the data processing, model selection, regression, and validation steps. Data collection involved setting up the rigorous simulation in Aspen HYSYS (V7.3), developing the automation program, and carrying out multiple rigorous simulations to obtain the samples. On average, each simulation took 0.5 s using a computer with an Intel Core processor of 3.40 GHz and 8.00 GB of installed RAM.

The coefficients for eq 9 and 10 were calculated using least-squares regression and results from calculating crude oil feed enthalpy over a suitable range of temperatures:

$$H = -1.1402 \times 10^{-6} T^3 + 1.1194 \times 10^{-3} T^2 + 0.1778T - 4.1333 \quad (11)$$

The reference enthalpy is the enthalpy of crude oil at feed conditions, that is, 25 °C and 2.5 bar. Figure 3 compares the regressed enthalpies and the enthalpies calculated with Aspen HYSYS (V7.3), showing good agreement.

4.2. Model Validation. The first validation of the ANNs was performed by the ANN Toolbox in MATLAB after training. Randomly selected points (15% of the total number of samples initially provided, i.e., results of converged simulations) comprised the validation set used in this case.

To gain more confidence of the ANN model, a second validation was carried out using a new set of data from rigorous simulations. This new data set was used to calculate the coefficient of determination R^2, the error and standard deviation σ of the error. The residual was calculated for temperature-related variables, while the relative error was calculated for variables other than temperature. LHS was used to obtain 1000 new data points, of which 955 simulations converged in Aspen HYSYS (V7.3). The lower and upper bounds in Table 1 were used to generate the LHS samples. Table 2 shows the values of R^2, the error and standard deviation

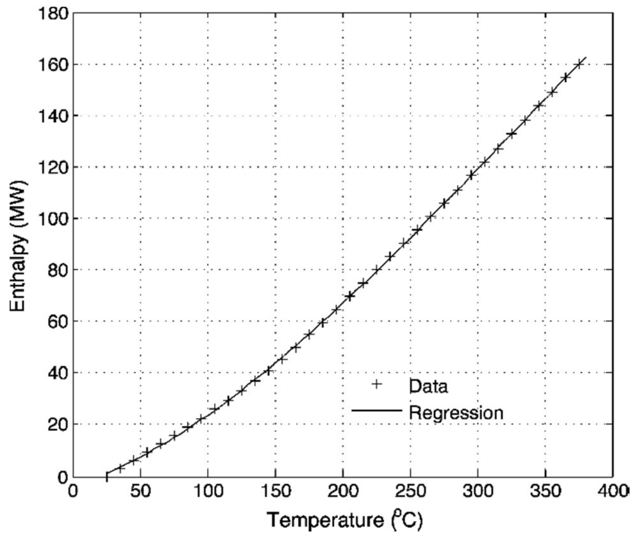


Figure 3. Enthalpy−temperature relationship of crude oil.

for each output of the neural networks. Results from validating the feasibility ANN are shown in Figure 4.

For temperature-related variables, the coefficient of determination was at least 0.97, the residual no greater than 0.22 °C and the standard deviation up to 3.6 °C. The maximum error and standard deviation of the error was found for the T95% temperature of HD stream. For this variable, the mean of the residual equals 0.22 °C and the standard deviation equals 3.63 °C. The T5% temperature of HN stream presented the lowest value of the coefficient of determination ($R^2 = 0.97$). The standard deviation represents data variability, while the coefficient of determination is an indicator of goodness of fit. A good fit should show low values for the standard deviation and values for the coefficient of determination close to 1. The confidence interval for the mean error of the T95% temperature predictions of the HD stream was calculated. This variable showed the greatest standard deviation for temperature-related variables, thus its confidence interval is also the greatest among these variables. The calculated confidence interval is ±0.3 °C from the mean error, using a 99% confidence level.

For variables other than temperatures, the coefficient of determination was at least 0.99, the mean error no greater than 0.23%, and the standard deviation up to 2.71%. The maximum error corresponds to flooding condition predictions for stages 25 to 32 of the main column. For this variable, the mean of the relative error and standard deviation of the error equal −0.23% and 0.80%, respectively. The maximum standard deviation corresponds to enthalpy change predictions for the LN stream. For this variable, the mean of the relative error and standard deviation of the error equal 0.15% and 2.71%, respectively. The coefficient of determination for variables other than temper-atures is equal or greater than 0.99. The confidence interval for the mean error of the enthalpy change predictions of the LN stream was calculated. The calculated confidence interval is ±0.2% from the mean error, using a 99% confidence level. Results presented in Table 2 show that predictions from the neural networks are in very good agreement with results from rigorous models.

The feasibility ANN was validated using a confusion matrix, as shown in Figure 4. The output of this ANN is an integer equal to 1 in case inputs lead to a converged simulation in

DOI: 10.1021/ie503802j
Ind. Eng. Chem. Res. 2015, 54, 4988−5000

Table 2. Coefficient of Determination, Error, and Standard Deviation of the Error for Outputs of the Neural Networks

variable	R^2	$T - T_{calc}$	$\sigma(T - T_{calc})$	variable	R^2	error %	σ(error%)
T5% (°C)				Column flooding %			
LN	0.99	0.00	0.17	column stages 1−5	1.00	0.01	0.83
HN	0.99	−0.02	1.23	column stages 6−14	1.00	0.08	1.02
LD	0.99	−0.13	1.88	column stages 15−24	1.00	0.00	0.77
HD	0.99	−0.09	2.15	column stages 25−32	1.00	−0.23	0.80
RES	1.00	−0.04	1.38	column stages 33−41	1.00	−0.08	0.81
T95% (°C)				HD stripper	0.99	−0.01	1.64
LN	0.99	0.02	1.57	LD stripper	1.00	0.00	0.87
HN	0.97	−0.15	3.12	HN stripper	1.00	−0.05	1.04
LD	0.99	0.05	2.51	Enthalpy change			
HD	0.99	0.22	3.63	HN reboiler	1.00	0.05	0.86
RES	1.00	0.03	0.57	LD reboiler	1.00	0.05	0.66
Supply temp (°C)				condenser	1.00	0.01	0.80
HN reboiler	0.99	−0.03	1.14	LN	0.99	0.15	2.71
LD reboiler	1.00	−0.01	1.08	HN	1.00	−0.03	1.40
condenser	0.98	0.10	1.10	LD	1.00	−0.03	0.63
PA_1	0.99	−0.04	0.93	HD	1.00	−0.01	1.04
PA_2	1.00	0.01	0.84	RES	1.00	0.00	0.18
PA_3	1.00	0.04	0.59	Heat capacity ratio ψ			
LN	0.99	0.00	0.29	PA_1	1.00	0.00	0.04
HN	0.99	−0.03	1.09	PA_2	1.00	0.01	0.04
LD	1.00	−0.02	1.02	PA_3	1.00	0.00	0.04
HD	1.00	0.00	0.90	LN	0.99	0.00	0.07
RES	1.00	−0.01	0.50	HN	0.99	0.00	0.21
Target temp (°C)				LD	0.99	0.00	0.17
HN reboiler	0.99	−0.03	1.09	HD	0.99	−0.02	0.16
LD reboiler	1.00	−0.02	1.02	RES	0.99	0.01	0.11
condenser	0.99	0.00	0.29				

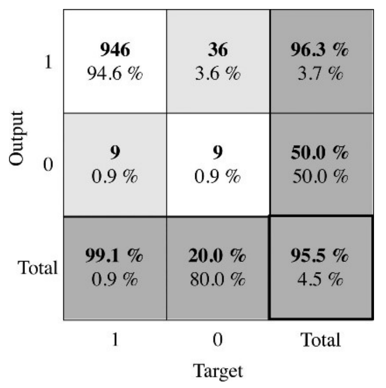


Figure 4. Confusion matrix for the feasibility ANN: no shading, correct predictions; light shading, incorrect predictions; double shading, summary of results.

Aspen HYSYS (V7.3). Otherwise, a value of zero is allocated. Values in boxes with no shading indicate the number and percentage of correct predictions. The values in boxes with light shading indicate incorrect predictions. The boxes in double shading summarize results for these predictions. It can be seen that 95.5% of predictions were correct, with 3.6% false positive predictions and 0.9% false negative predictions. From these two types of false predictions, false positive ones are the least desirable. A false positive prediction will allow unfeasible inputs to be considered by the algorithm used to optimize the operating conditions of the distillation column, which can lead to unrealistic results. On the contrary, a false negative prediction will make the optimization algorithm reject a set of feasible inputs, which decreases the number of available design options.

Computation times to simulate the distillation process using rigorous models in Aspen HYSYS (V7.3) and the ANN model built in MATLAB were compared. The computer used in this work has an Intel Core processor of 3.40 GHz and 8.00 GB of RAM. The maximum number of iterations for the solver embedded in Aspen HYSYS (V7.3) was changed to 20. The rest of the solver settings were kept to their default values. One hundred samples were generated using the LHS technique. These sample points were simulated with both distillation models. Total computation time for the ANN model was 1 s. Total computation time for rigorous models in Aspen HYSYS (V7.3) was 182 s; the exchange of information with MATLAB took another 4 s. From these results it is evident that computation times are significantly greater when using rigorous models. Furthermore, the convergence rate of rigorous models was 95%, compared to a 100% convergence rate of the ANN model. In the context of process optimization, where many thousands of simulations may be carried out, the advantages of the ANN model would be even greater. The ANN model is considerably more computationally efficient than rigorous models. Moreover, the validation tests proved that results from the ANN model are in very good agreement with results from rigorous models.

4.3. Optimization Considering Minimum Energy Requirements. The ANN distillation model, the feasibility ANN and eq 5 to 11 are implemented in an optimization framework to improve profit. The grand composite curve is used in this case study to calculate minimum energy requirements. Specifically, the requirements for fired heating

DOI: 10.1021/ie503802j
Ind. Eng. Chem. Res. 2015, 54, 4988−5000

and cooling water. The consideration of a detailed HEN model to calculate energy requirements is presented in Parts II[1] and III[2] of the series.

The objective function for optimization can be expressed mathematically as

$$\min f = -\sum_{i=1}^{N_{\text{prod}}} C_{\text{prod},i} F_{\text{prod},i} + C_{\text{crude}} F_{\text{crude}} + \sum_{j=1}^{N_{\text{stm}}} C_{\text{stm},j} F_{\text{stm},j} + \sum_{k=1}^{N_{\text{util}}} C_{\text{util},k} U_{\min,k} \quad (12)$$

where C and F refer to prices and flow rates, respectively. Subscripts prod, crude, stm, and util refer to the product streams, crude oil, stripping steam, and utilities, respectively. N_{prod} is the total number of distillation products; N_{stm} is the number of steam streams (i.e., stripping steam in the main column and HD stripper); and N_{util} is the total number of utilities (i.e., cooling water and fired heating). $U_{\min}$ is the minimum requirement of utility k, calculated from the grand composite curve. The product unit prices used in this case study are based on the crude oil price of 2010[51] and calculated using the procedure presented by Maples.[52] Unit prices of stripping steam and utilities are taken from Chen.[4] Capital costs to retrofit the HEN are not included in eq 12, but they are considered in Parts II[1] and III[2] of the series. Table 3 presents the unit prices of crude oil, products, steam, and utilities used in eq 12.

Table 3. Prices of Crude Oil, Distillation Products, Utilities and Stripping Steam

item	price	units
crude oil	79.6	\$/bbl
LN	103.5	\$/bbl
HN	92.7	\$/bbl
LD	99.0	\$/bbl
HD	96.6	\$/bbl
RES	61.3	\$/bbl
fired heating (1500−800 °C)	150.0	\$/kWy
cooling water (10−40 °C)	5.25	\$/kWy
stripping steam (260 °C, 4.5 bar)	0.14	\$/kmol

4.3.1. Process Constraints. Process constraints are employed to make sure that solutions are practicable and sensible. The objective function for optimization described in eq 12 is subjected to constraints on the optimization variables, product quality, and column flooding. Column flooding is calculated using the tray sizing utility in Aspen HYSYS (V7.3), with the default parameters for sieve trays.

An additional constraint is used to exclude operating conditions that are unfeasible. This feasibility constraint employs the output of the feasibility ANN, which is equal to 1 for inputs that lead to converged simulations in Aspen

$$T_{\text{PA}k}^{\text{lb}} \le T_{\text{PA}k} \le T_{\text{PA}k}^{\text{ub}} \quad k = 1, 2, 3 \quad (16)$$

$$\text{COT}^{\text{lb}} \le \text{COT} \le \text{COT}^{\text{ub}} \quad (17)$$

$$T5_i^{\text{lb}} \le T5_i \le T5_i^{\text{ub}} \quad i = 1, 2, ..., N_{\text{prod}} \quad (18)$$

$$T95_i^{\text{lb}} \le T95_i \le T95_i^{\text{ub}} \quad i = 1, 2, ..., N_{\text{prod}} \quad (19)$$

$$\text{Flooding}_m \le \text{Flooding}_m^{\text{ub}} \quad m = 1, 2, ..., N_{\text{sections}} \quad (20)$$

$$\alpha = 1 \quad (21)$$

where Q_{PA} and T_{PA} denote the duty and temperature drop for pump-around k, COT is the coil outlet temperature; T5 and T95 indicate the T5% and T95% true boiling point (TBP) temperatures for product i; "Flooding" refers to the flooding percentage for sections of the main column and strippers; and α is the output of the feasibility ANN. Superscripts lb and ub indicate the specified lower and upper bounds of the variables, respectively.

For flooding calculations within Aspen HYSYS (V7.3), the main distillation column is divided into five sections as shown in Table 2, while only one section is considered per stripper. The maximum flooding condition considered for all sections and strippers is 85%.

The lower and upper bounds of the optimization variables (eq 13 to 17) are parameters of the optimization algorithm, and are chosen by the designer. Constraints in eq 18 to 21 are included in the objective function as penalty functions. Thus, the objective function in eq 12 is adapted to the form:

$$\min F(\mathbf{x}) = f(\mathbf{x}) + \gamma_1 |h(\mathbf{x})| + \sum_{j=1}^{n} \gamma_{1+j} \max(0, g_j(\mathbf{x})) \quad (22)$$

$$h(\mathbf{x}) = 0$$

$$g_j(\mathbf{x}) \le 0$$

where x are the optimization variables (i.e., F_{prod}, F_{stm}, Q_{PA}, T_{PA}, and COT); F(x) is the unconstrained objective function; f(x) is the constrained objective function (eq 12); h(x) is the equality constraint of eq 21; g_j(x) are the inequality constraints (eq 18 to 20); and γ are penalty factors that ensure that constraints are scaled and given the corresponding importance during optimization. The values for the lower and upper bounds of the optimization variables are shown in Table 1. The values of the T5% and T95% TBP temperatures are allowed to vary ±10 °C from the base case conditions to maintain product quality. Variables related to product quality are presented in Table 4.

4.3.2. Calculation of Minimum Energy Requirements. The distillation system in this case study comprises three cold process streams (i.e., crude oil, LD, and HN reboiler streams) and nine hot process streams (i.e., pump-arounds, condenser, and distillation product streams). Cooling water and fired

HYSYS (V7.3). The constraints considered for the distillation unit can be expressed as follows:

$$F_{\mathrm{prod},i}^{\mathrm{lb}} \leq F_{\mathrm{prod},i} \leq F_{\mathrm{prod},i}^{\mathrm{ub}} \quad i = 1, 2, ..., N_{\mathrm{prod}} - 1 \tag{13}$$

$$F_{\mathrm{stm},j}^{\mathrm{lb}} \leq F_{\mathrm{stm},j} \leq F_{\mathrm{stm},j}^{\mathrm{ub}} \quad j = 1, 2 \tag{14}$$

$$Q_{\mathrm{PA}k}^{\mathrm{lb}} \leq Q_{\mathrm{PA}k} \leq Q_{\mathrm{PA}k}^{\mathrm{ub}} \quad k = 1, 2, 3 \tag{15}$$

heating are used as cold and hot utilities, respectively. The values of supply and target temperatures (T_s, T_t), enthalpy change (ΔH), and CP ratio (ψ) for the base case conditions are presented in Table S9 (see Supporting Information).

Stream information in Supporting Information, Table S9 is used in eq 5 and 7 to model heat capacity flow rate as a function of temperature for the pump-around and product streams. Stream information in Table S9 is also employed in eq 8 to calculate the pseudo-heat-capacity flow rate (CP*) for the

DOI: 10.1021/ie503802j
Ind. Eng. Chem. Res. 2015, 54, 4988−5000

Table 4. Product Quality Results

item	base case	lower bound	upper bound	optimized case
LN T5% (°C)	6	−4	16	6 (+1)
HN T5% (°C)	102	92	112	107 (+5)
LD T5% (°C)	174	164	184	176 (+2)
HD T5% (°C)	289	279	299	290 (+1)
RES T5% (°C)	358	348	368	366 (+8)
LN T95% (°C)	111	101	121	112 (+1)
HN T95% (°C)	187	177	197	189 (+1)
LD T95% (°C)	312	302	322	313 (+1)
HD T95% (°C)	363	353	373	366 (+3)
RES T95% (°C)	889	879	899	894 (+5)

condenser and reboiler streams. Finally, eq 11 is used to calculate the pseudo-heat-capacity flow rate for the crude oil feed.

Figure 5 shows the GCC for the heat-integrated system considering (a) constant thermal properties, and (b) temperature-dependent thermal properties (eq 5 to eq 11). The minimum approach temperature is 25 °C. Hot and cold utility requirements for case (a) are 41.89 MW and 47.33 MW, respectively. For case (b), hot and cold utility requirements are 51.53 MW and 57.97 MW. These results show that utility requirements are very different in the two cases. For these operating conditions, hot and cold utility requirements are underestimated by over 18% when thermal properties are assumed constant. The work of Chen[4] shows that it is just as important to consider the temperature dependence of thermal properties when taking into account the details of the HEN. Chen[4] compared the results of simulating a HEN with and without considering temperature-dependent heat capacities. The comparison showed that the calculated HEN temperatures assuming constant heat capacities were underestimated by up to 27 °C. Thus, it is very important to consider the temperature dependence of thermal properties to obtain meaningful estimations of energy requirements and temperatures.

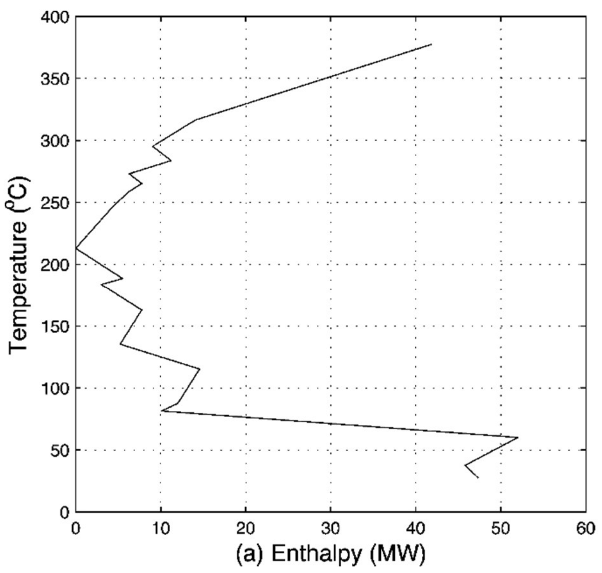
4.3.3. Optimization Framework. The optimization framework employed in this case study is illustrated in Figure 6. Starting with a distillation unit with feasible operating conditions, the optimization algorithm calculates the values of F_{prod}, F_{stm}, Q_{PA}, T_{PA} and COT that increase profits. Profit improvement can be achieved by increasing the flow rates of the most valuable products (i.e., LN, LD, and HD), decreasing steam flow rates and/or decreasing utility consumption, especially that of fired heating.

For each iteration of the optimization algorithm, the distillation process is simulated using neural networks. Then, stream information (i.e., T_s, T_t, H and ψ) is passed to the heat recovery model. In this case study, the heat recovery model is the grand composite curve. The GCC provides minimum utility requirements considering temperature-dependent thermal properties. Results from simulating the distillation column and the GCC are used to calculate the objective function in eq 12 and penalty functions for constraints in eq 18 to eq 21.

Simulated annealing is used as the optimization algorithm. Function simulannealbnd embedded in the Global Optimization Toolbox in MATLAB is employed in this case study. The parameters of the SA algorithm were kept to their default values except for the TolFun (i.e., function tolerance) and TimeLimit (i.e., time limit) parameters, which were selected by trial and error and by taking into account the order of magnitude of the objective function. The function tolerance was set to 10, while the time limit was set to 5 min.

4.3.4. Optimization Results. The optimized variables are presented in Table 1. The optimized calculated reflux ratio is 4.57, where the initial value was also 4.57. A summary of optimization results is presented in Table 5. Table S10 (see Supporting Information) shows stream information used to calculate minimum energy requirements. The GCC at the optimized conditions is illustrated in Figure 7. Table S8 (see Supporting Information) presents the flooding conditions for the crude oil distillation unit at optimized operating conditions.

The optimization algorithm was run 10 times to obtain a range of solutions, from which the best was chosen. A total of 1299 function evaluations were carried out for the best solution found. The new operating conditions were validated in Aspen HYSYS (V7.3), showing good agreement.



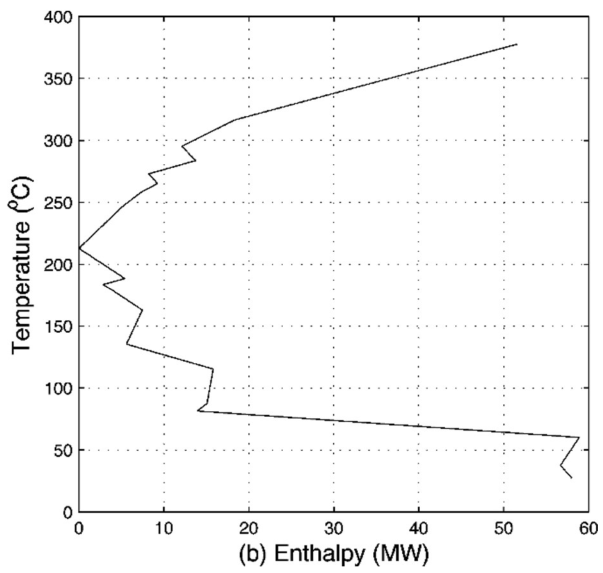


Figure 5. Grand composite curves for base case conditions considering: (a) constant thermal properties and (b) temperature-dependent thermal properties.

DOI: 10.1021/ie503802j
Ind. Eng. Chem. Res. 2015, 54, 4988−5000

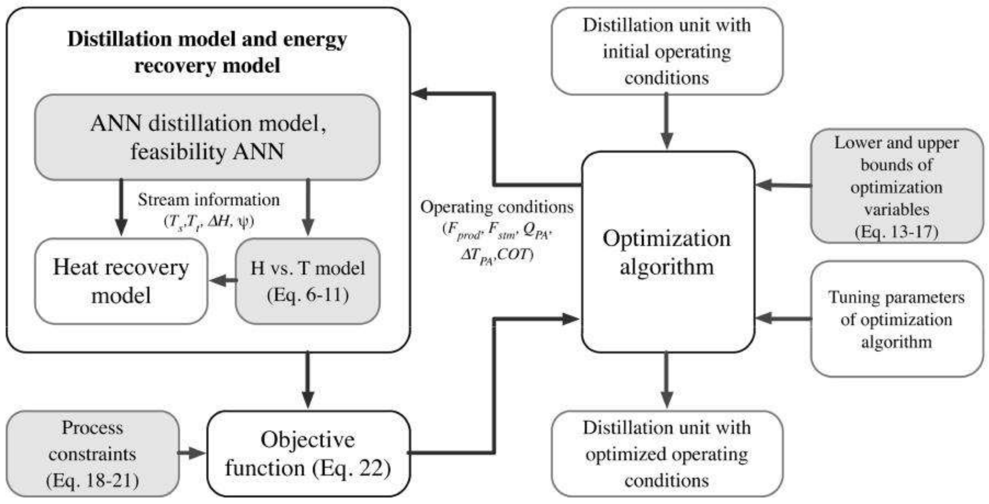


Figure 6. Optimization framework.

Table 5. Optimization Results for the Crude Oil Distillation System

item	base case	optimized case	
Summary of utility consumption and operating costs			
hot utility (MW)	51.5	48.2	(−6%)
cold utility (MW)	58.0	55.1	(−5%)
utility costs (M$/y)[a]	8.0	7.5	(−6%)
steam cost (M$/y)	1.7	1.8	(+2%)
crude oil cost (M$/y)	2852.3	2852.3	
operating cost (M$/y)	2862.1	2861.6	(~0%)
Product income and profit			
product income (M$/y)	2881.9	2904.0	(+1%)
profit (M$/y)	19.8	42.4	(+22.6 M$/y)

[a]M$ denotes millions of dollars.

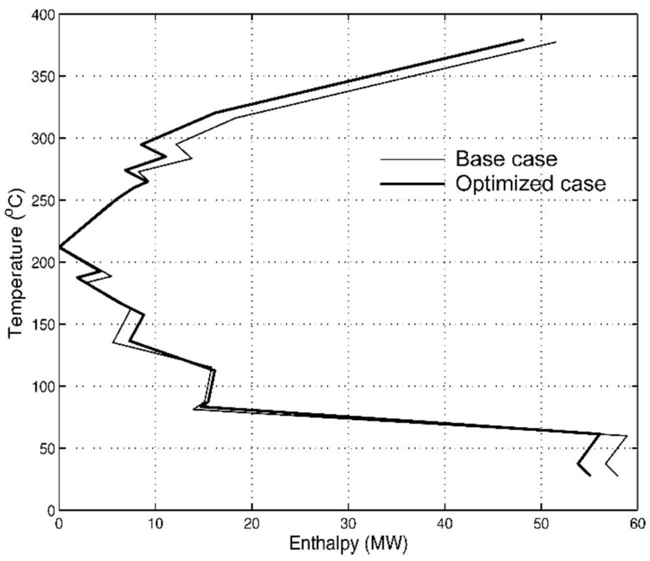


Figure 7. Grand composite curve for optimized conditions considering temperature-dependent thermal properties.

The product income for the optimized operating conditions is 2904.0 M$/y (millions of US$ per year), which represents an increase of 22.1 M$/y (~1%) from the base case. The flow rates of the most valuable distillation products (LN, LD, and HD) are increased at the expense of the least valuable ones (HN and RES). Minimum hot and cold utility consumption is decreased by 3.3 MW and 2.9 MW, respectively. For this case study, results indicate that product income has a dominant effect in process economics. Utility costs together amount to less than 1% of product revenue. Even though utility costs are marginal, the optimizer was able to reduce these operating costs by around 6% while achieving considerable improvements in product revenue.

Table 4 shows T5% and T95% TBP temperatures for the optimized operating conditions; these results confirm that product quality is within specified ranges. As the flow rates of the lightest components increase, more heavy material is moved from the bottom stages to the upper part of the column; as a consequence, temperatures in the column increase. The increased temperatures of the product streams allow more useful heat from the column to be transferred to the preheat train. On the other hand, if the flow rates of the top products are diminished, the temperatures of all distillation products decrease, thus available energy for heating exists at a lower temperature. This energy at a lower temperature is less useful for preheating the crude oil feed. Therefore, more fuel needs to be supplied to the furnace, and the hot utility demand increases.

Results from Tables 1, 4, and 5 should be treated with caution. Although the new operating conditions of the distillation system meet the constraints; that is, a mathemati-cally feasible solution is found, there is no guarantee that the new operating conditions are feasible in practice. A key reason for this caution is the lack of consideration of HEN details in the optimization problem. The new set of operating conditions presents considerable changes in flow rates (e.g., HD and PA3) and cooling and heating requirements that may not be achievable in the existing HEN. Other factors such as the capacity of the column internals (e.g., downcomers, distrib-utors, nozzles), pumps and furnace have not been taken into account in the optimization problem. One strategy to increase the chances of obtaining practicable results using this simple optimization strategy is to restrict the values of the optimization variables to relatively small ranges. This reduced range would lead to results with less significant perturbations to the system, but could also limit the improvements that could be achieved. Part III[2] of this series includes a detailed HEN model into the optimization problem.

DOI: 10.1021/ie503802j
Ind. Eng. Chem. Res. 2015, 54, 4988−5000

5. CONCLUSIONS

Design and optimization of heat-integrated crude oil distillation systems is a complex task. The challenge lies not only in developing optimization frameworks that consider both the distillation process and heat recovery network simultaneously, but also in developing models that are accurate and computationally efficient for these purposes. Detailed distil-lation models have been developed for applications that do not consider interactions between the distillation unit and heat exchanger network. However, these models are computation-ally demanding and often present convergence problems that make them unsuitable for optimization of heat-integrated systems.

This paper presents a new methodology to model crude oil distillation processes in the context of the heat recovery system. This modeling approach employs artificial neural networks to regress the variables that describe the separation and energy performance of the distillation process. Results from rigorous simulations are used to train the ANNs. The independent variables (inputs of the ANN model) are the operating conditions of the distillation process, namely flow rates of products and stripping steam, pump-around duties, and temperature drops and coil outlet temperature. The dependent variables (outputs of the ANN distillation model) are variables related to product quality (i.e., T5% and T95% TBP temperatures), flooding conditions, and stream information to calculate energy requirements. This stream information consists of supply and target temperatures, enthalpies, and heat capacity variation of process streams passing through the heat exchanger network.

Predictions from the ANN distillation model and rigorous models are compared and shown to be in good agreement. The resulting ANN distillation model is demonstrated to be accurate and significantly faster in convergence than rigorous models. ANN models do not require the specification of initial guesses or complex solution algorithms that rigorous models do, since ANN models have simpler formulations and are more robust. These characteristics make the ANN distillation model suitable for implementation in optimization strategies, as illustrated in a case study presented in this paper. A comparison between minimum energy requirements assuming constant and temperature-dependent thermal properties is made. As expected, minimum energy requirements were significantly underestimated when thermal properties are assumed constant. Part III[2] of this series implements the distillation models presented in this paper and the HEN models presented in Part II[1] to optimize the overall heat-integrated crude oil distillation system.

The proposed distillation modeling strategy can be extended to include plant measurements to train the neural networks or to include different distillation variables to the ones selected in this work. Future work includes developing ANN distillation models that consider structural variables of the distillation unit, such as feed or pump-around location; or additional variables, such as crude oil blending ratio.

■ ASSOCIATED CONTENT

* Supporting Information

Additional data and results for the optimization of the atmospheric distillation unit presented in the case study. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: megan.jobson@manchester.ac.uk.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank the Mexican National Council of Science and Technology (CONACYT, no. 204362/308621) and the Roberto Rocca Education Program for financial support granted for the development of this work.

■ REFERENCES

(1) Ochoa-Estopier, L. M.; Jobson, M.; Chen, L.; Rodriguez, C.; Smith, R. Optimization of heat-integrated crude oil distillation systems. Part II: Heat exchanger network retrofit model. Ind. Eng. Chem. Res. 2015, DOI: 10.1021/ie503804u.

(2) Ochoa-Estopier, L. M.; Jobson, M. Optimization of heat-integrated crude oil distillation systems. Part III: Optimization framework. Ind. Eng. Chem. Res. 2015, DOI: 10.1021/ie503805s.

(3) Szklo, A.; Schaeffer, R. Fuel specification, energy consumption and CO_2 emission in oil refineries. Energy 2007, 32, 1075−1092.

(4) Chen, L. Heat-Integrated Crude Oil Distillation Design. Ph.D. Thesis, The University of Manchester, Manchester, UK, 2008.

(5) Nelson, W. L. Petroleum Refinery Engineering, 4th ed.; McGraw-Hill: New York, USA, 1958.

(6) Watkins, R. N. Petroleum Refinery Distillation, 2nd ed.; Gulf Publishing Company: Houston, USA, 1979.

(7) Ochoa-Estopier, L. M.; Jobson, M.; Smith, R. The use of reduced models for design and optimisation of heat-integrated crude oil distillation systems. Energy 2014, 75, 5−13.

(8) Liebmann, K.; Dhole, V. R.; Jobson, M. Integrated design of a conventional crude oil distillation tower using pinch analysis. Chem. Eng. Res. Des. 1998, 76, 335−347.

(9) Bagajewicz, M.; Ji, S. Rigorous procedure for the design of conventional atmospheric crude fractionation units. Part I: Targeting. Ind. Eng. Chem. Res. 2001, 40, 617−626.

(10) Ji, S.; Bagajewicz, M. Design of crude distillation plants with vacuum units. I. Targeting. Ind. Eng. Chem. Res. 2002, 41, 6094−6099.

(11) Seo, J. W.; Oh, M.; Lee, T. H. Design optimization of crude oil distillation. Chem. Eng. Technol. 2000, 23, 157−164.

(12) Basak, K.; Abhilash, K. S.; Ganguly, S.; Saraf, D. N. On-line optimization of a crude distillation unit with constraints on product properties. Ind. Eng. Chem. Res. 2002, 41, 1557−1568.

(13) Inamdar, S. V.; Gupta, S. K.; Saraf, D. N. Multi-objective optimization of an industrial crude distillation unit using the elitist nondominated sorting genetic algorithm. Chem. Eng. Res. Des. 2004, 82, 611−623.

(14) Al-Mayyahi, M. A.; Hoadley, A. F. A.; Smith, N. E.; Rangaiah, G. P. Investigating the trade-off between operating revenue and CO_2 emissions from crude oil distillation using a blend of two crudes. Fuel 2011, 90, 3577−3585.

(15) Liau, L. C. K.; Yang, T. C. K.; Tsai, M. T. Expert system of a crude oil distillation unit for process optimization using neural networks. Expert Syst. Appl. 2004, 26, 247−255.

(16) Yao, H.; Chu, J. Operational optimization of a simulated atmospheric distillation column using support vector regression models and information analysis. Chem. Eng. Res. Des. 2012, 90, 2247−2261.

(17) Lopez C, D. C.; Hoyos, L. J.; Mahecha, C. A.; Arellano-García, H.; Wozny, G. Optimization model of crude oil distillation units for an optimal crude oil blending and operating conditions. Ind. Eng. Chem. Res. 2013, 52, 12993−13005.

(18) Abdullah, Z., Aziz, N., Ahmad, Z. Nonlinear modelling application in distillation column Chem. Prod. Process Model. 20072, Art. 12.

(19) Pirdashti, M.; Curteanu, S.; Kamangar, M. H.; Hassim, M. H.; Khatami, M. A. Artificial neural networks: applications in chemical engineering. Rev. Chem. Eng. 2013, 29, 205−239.
(20) Himmelblau, D. Applications of artificial neural networks in chemical engineering. Korean J. Chem. Eng. 2000, 17, 373−92.
(21) Zhao, W.; Chen, D.; Hu, S. Optimizing operating conditions based on ANN and modified GAs. Comput. Chem. Eng. 2000, 24, 61− 65.
(22) Bollas, G. M.; Papadokonstantakis, S.; Michalopoulos, J.; Arampatzis, G.; Lappas, A. A.; Vasalos, I. A.; Lygeros, A. A computer-aided tool for the simulation and optimization of the combined HDS-FCC processes. Chem. Eng. Res. Des. 2004, 82, 881−894.
(23) Bellos, G. D.; Kallinikos, L. E.; Gounaris, C. E.; Papayannakos, N. G. Modelling of the performance of industrial HDS reactors using a hybrid neural network approach. Chem. Eng. Process. 2005, 44, 505− 515.
(24) Zahedi, G.; Mohammadzadeh, S.; Moradi, G. Enhancing gasoline production in an industrial catalytic-reforming unit using artificial neural networks. Energy Fuels 2008, 22, 2671−2677.
(25) Zahedi, G., Lohi, A., Karami, Z. A neural network approach for identification and modeling of delayed coking plant Int. J. Chem. React. Eng. 20097Art. A16.
(26) Alhajree, I.; Zahedi, G.; Manan, Z. A.; Zadeh, S. M. Modeling and optimization of an industrial hydrocracker plant. J. Pet. Sci. Eng. 2011, 78, 627−36.
(27) Al-Marhoun, M. A., Osman, E. A. Using Artificial Neural Networks to Develop New PVT Correlations for Saudi Crude Oils. Proceedings of the 10th Abu Dhabi International Petroleum Exhibition and Conference, Abu Dhabi, Saudi Arabia, October 13−16, 2002; pp 1−15.
(28) Wu, K.; He, X.; Chen, B. On-line training method of ANN in DMS for side draw quality of refinery fractionator. Comput. Chem. Eng. 2000, 24, 1585−1589.
(29) Aminian, J.; Shahhosseini, S. Evaluation of ANN modeling for prediction of crude oil fouling behavior. Appl. Therm. Eng. 2008, 28, 668−674.
(30) Sarle, W. S. Neural Networks and Statistical Models. Proc. Ninet. Annu. SAS Users Gr. Int. Conf. 1994; pp 1538−1550.
(31) Krose, B., van der Smagt, P. An Introduction to Neural Networks, 8th ed.; University of Amsterdam: Amsterdam, The Netherlands, 1996.
(32) Nascimento, C. A. O.; Giudici, R.; Guardani, R. Neural network based approach for optimization of industrial chemical processes. Comput. Chem. Eng. 2000, 24, 2303−2314.
(33) Beale, M. H., Hagan, M. T., Demuth, H. B. Neural Network Toolbox. User's Guide; The MathWorks, Inc.: Natick, MA, 2011.
(34) Cavazzuti, M. Optimization Methods: From Theory to Design; Springer: Modena, Italy, 2013.
(35) Bengio, Y., LeCun, Y. In Large-Scale Kernel Mach.; Bottou, L., Chapelle, O., DeCoste, D., Weston, J., Eds.; M.I.T. Press: 2007; pp 321−360.
(36) Smith, S. W. Scientists and Engineer's Guide to Digital Signal Processing, 2nd ed.; California Technical Publishing: San Diego, USA, 1999; Chapter 26, pp 451−480.
(37) Heaton, J. Introduction to Neural Networks with Java, 2nd ed.; Heaton Research, Inc.: Chesterfield, USA, 2005.
(38) Wang, W.; Van Gelder, P. H. A. J. M.; Vrijling, J. K. Some issues about the generalization of neural networks for time series prediction. Artif. Neural Networks Form. Model. Appl. 2005, 559−564 ICANN 2005.
(39) Angus, J. E. Criteria for Choosing the Best Neural Network: Part I; Naval Medical Research and Development Command: Bethesda, MD, 1991.
(40) Narasimha, P. L.; Delashmit, W. H.; Manry, M. T.; Li, J.; Maldonado, F. An integrated growing−pruning method for feedforward network training. Neurocomputing 2008, 71, 2831−2847.
(41) Nolfi, S., Parisi, D. In Handbook of Brain Theory Neural Networks, 2nd ed.; Arbib, M., Ed.; The MIT Press: Madison, USA, 2002; Chapter 3, pp 418−421.
(42) Stanley, K. O.; Miikkulainen, R. Efficient Evolution of Neural Network Topologies. Proceedings of the 2002 Congress on Evolu-tionary Computation, Piscataway, NJ, 2002, pp 1757−1762
(43) Aspen HYSYS, version 7.3; Aspen Technology: Bedford, MA, 2011.
(44) Diwekar, U. M.; Kalagnanam, J. R. Efficient sampling technique for optimization under uncertainty. AIChE J. 1997, 43, 440−447.
(45) Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S. Global Sensitivity Analysis: The Primer; John Wiley & Sons, Ltd.: Chichester, UK, 2008.
(46) Cheng, J.; Druzdzel, M. Latin Hypercube Sampling in Bayesian Networks. Proc. 13th Int. Florida Artif. Intell. Res. Soc. Conf. 2000, 287− 292.
(47) MATLAB, version R2012a; The MathWorks Inc.: Natick, MA, 2012.
(48) Levenberg, K. A method for the solution of certain problems in least squares. Q. Appl. Math. 1944, 2, 164−168.
(49) Marquardt, D. W. An algorithm for least-squares estimation of nonlinear parameters. J. Soc. Ind. Appl. Math 1963, 11, 431−441.
(50) Møller, M. F. A. Scaled conjugate gradient algorithm for fast supervised learning. Neural Networks 1993, 6, 525−533.
(51) Data of petroleum and other liquids. www.eia.gov/dnav/pet/pet_pri_spt_s1_d.htm (accessed March 18, 2012).
(52) Maples, R. Petroleum Refinery Process Economics, 2nd ed.; Pennwell Corp.: Tulsa, USA, 2000.

DOI: 10.1021/ie503802j
Ind. Eng. Chem. Res. 2015, 54, 4988−5000