Computers and Chemical Engineering 183 (2024) 108591

Contents lists available at ScienceDirect

Computers and Chemical Engineering

journal homepage: www.elsevier.com/locate/compchemeng

Exploring nontraditional LSTM architectures for modeling demethanizer column operations

Marta Mandis [a], Roberto Baratti [a], Jorge Chebeir [b], Stefania Tronci [a,*], José A. Romagnoli [b]

[a] Dip. di Ingegneria Meccanica, Chimica e dei Materiali, Università degli Studi di Cagliari, Cagliari, 09123, Italy
[b] Department of Chemical Engineering, Louisiana State University, Baton Rouge, LA, 70809, USA

ARTICLE INFO

Keywords:
Natural gas liquids recovery
LSTM Neural Networks
Digital twin
Distillation column
Dynamic process simulation

ABSTRACT

Digital twins have recently attracted attention as a new technology that can facilitate the digital transformation of process industries. It may provide live, or near real-time, information and insights into the process and may be used for monitoring, control and optimization purposes. In this study, a digital twin has been developed for modelling the demethanizer column of a NGL separation plant. Based on a non-conventional Long Short-Term Memory (LSTM) neural network arrangement, the surrogate model has been trained and validated using data obtained by the process simulator Aspen HYSYS®. Model prediction can be obtained using only readily available variables as input data, ensuring easy and cost-effective implementation. Measurement noises have been considered in order to mimic real-world measurements in a real plant. In both steady-state and transient conditions, the developed demethanizer digital twin accurately reconstructs the separation operation, including compositions, temperatures, and pressures in the reboiler and all column stages.

1. Introduction

The NGL represent a mixture of different hydrocarbons that includes ethane, propane, butane, isobutane, and natural gasoline. These hydrocarbons constitute the heavier portion of the raw natural gas extracted from subsurface rock formations via drilling. After proper separation, the NGL can be marketed independently, with a higher market price compared to methane gas, and utilized as feedstock for various industrial applications such as plastics, synthetic rubber, gasoline blending, natural gasoline (He and You, 2014). The cryogenic expansion is one of the most common processes to separate NGL from the raw natural gas with a high rate of ethane recovery. Numerous cryogenic expansion processes have been proposed in the gas industry, including the gas subcooled process (GSP), cold residue recycle (CRR) process, recycle split-vapor (RSV) process, and recycle split-vapor with enrichment (RSVE) process among others. An extensive review of the different separation processes has been performed in literature (Manning and Thompson, 1991; Arnold and Stewart, 1999; Pitman et al., 1998; Kidnay and Parrish, 2006; Kidnay et al., 2011; Zhang et al., 2020).

The optimal design and operation under steady state conditions of different NGL separation processes have been studied by different authors, including Mehrpooya et al., (2006); Chebbi et al., (2008); Chebbi et al., (2010); Getu et al., (2013); Park et al., (2015) and Kherbeck and Chebbi (2015) among others. Recently, researchers have also focused on studying the intricacies involved in the dynamic aspects of the NGL recovery processes. The exploration of this field is generally focused on the systeḿs dynamics and the development of control strategies to achieve high product quality as well as reducing processing time and operating costs. Luyben (2013) studied the dynamic control of a demethanizer column with side reboilers, comparing inferential and direct composition control. Chebeir et al., (2019) examined NGL recovery processeś dynamics under typical disturbances and proposed control structures involving composition controllers in direct or cascade arrangements. In both studies, the delays caused by the utilization of composition analyzers as well as their implementation costs were overlooked. Tronci et al., (2020) addressed the high costs associated with purchasing and maintaining composition analyzers and the long-time delays in composition control loops by proposing a control strategy employing only indirect composition controllers for quality product control. Mandis et al., (2021) explored the use of a composition analyzer to achieve a hybrid cascade control using in-line delayed concentration measurements, aiming to eliminate the steady state offset of composition at the expense of response speed due to analyzer-related delays.

* Corresponding author.
E-mail address: stefania.tronci@unica.it (S. Tronci).

https://doi.org/10.1016/j.compchemeng.2024.108591
Received 19 October 2023; Received in revised form 18 December 2023; Accepted 10 January 2024
Available online 12 January 2024

To achieve optimal control of process production targets, alternative approaches have been investigated that do not rely on the use of composition analyzers. One widely adopted approach involves replacing measured concentrations with precise real-time estimations. Common techniques employed in the industrial domain for data-driven modelling include principal component regression, partial least squares, and artificial neural networks (ANNs) (Ge, 2017). ANNs, in particular, excel at capturing correlations in nonlinear physical systems and are recognized as universal approximators, capable of learning and representing complex input-output relationships (Lu et al., 2019). To address time-dependent problems, recurrent neural networks (RNNs) (Elman, 1990) were introduced. RNNs, with their recurrent loop and cell memory, can capture and retain temporal dependencies, even over long durations (Wu et al., 2019; Wu et al., 2022). Among RNNs, the Long Short-Term Memory network (LSTM) (Hochreiter and Schmidhuber, 1997) is specifically designed to overcome the vanishing gradient issue. Furthermore, LSTM can mitigate the influence of Gaussian noise on the estimation process due to the significant role played by the internal states of the network in predictions. However, its performance tends to diminish when dealing with datasets containing more realistic non-Gaussian noise (Wu et al., 2021).

LSTMs have been successfully applied to various chemical engineering problems, including fault diagnosis (Pang et al., 2022; Wang et al., 2022; Bi et al., 2022), process optimization (Zhu et al., 2020), and soft sensors (Yuan et al., 2019; Ke et al., 2017; Kwon et al., 2021). Kwon et al., (2021) presented a novel contribution to the literature as the first to utilize an LSTM for industrial distillation processes. The study introduced the development of a machine learning-based prediction model for forecasting temperatures in various sections of a distillation column with the goal of producing at least 99% normal butane from mixed butane. The focus of the modelling was a binary distillation column, where knowledge of temperatures in specific column locations could be effectively utilized for indirect composition control.

Another possible application for LSTMs is digital twin development for plant monitoring and control (Qu et al., 2020; Zhu and Ji, 2022; Li and Qin, 2023). The use of LSTM neural networks for creating digital surrogate models is of particular interest in the process industry. A digital twin represents a virtual model of an entire plant or individual units, closely integrated with real equipment as measurement devices. This tool facilitates real-time information exchange and enables predictions of process dynamics over time. Consequently, a digital surrogate serves as a valuable resource for performance monitoring, predictive asset maintenance, production optimization, and advanced process control. In comparison to conventional state estimators such as the extended Kalman filters (EKFs) (Kalman, 1960) and the extended Luenberger observer (Zeitz, 1987), data-driven approaches stand out due to their independence from a specific model and reliance on data. This proves to be a significant advantage, particularly because constructing a mathematical model for a multicomponent distillation column can be a daunting task, demanding a combination of theoretical knowledge, experimental data, and numerical techniques. Moreover, once digital twins are trained, they demonstrate lower computational demands and eliminate the need for knowledge about feed compositions. Additionally, they possess the capability to adapt offline by incorporating new historical plant data, if deemed necessary.

The objective of this study is to assess the performance of alternative neural network architectures exploring the possibility of using LSTM neural networks for the definition of an integrated surrogate model capable of accurately describing the dynamic operations of a demethanizer column. To achieve this goal, a digital surrogate column model was formulated and implemented by dividing the problem into three primary tasks. The first one involved the estimation of the column's top and bottom compositions, while the second was dedicated to the estimation of the column's internal and external flows. In the third task, three different neural model architectures for column internal operations are defined and compared. To develop the final data-driven model for the demethanizer column, known as the Demethanizer Digital Twin (DDT), these models were integrated into a comprehensive architecture. This DDT architecture relies solely on readily available variable measurements as input data, ensuring easy and cost-effective implementation.

The remainder of this manuscript is organized as follows. The next section provides a detailed description of the CRR separation process considered for the implementation of the proposed approach. Thereafter, the methodology implemented for the development of the DDT architecture is presented. Next, the results obtained with the proposed DDT architecture are depicted and analyzed. Finally, conclusions are drawn in the last section of this work.

2. Process description

The dynamic behavior of the CRR separation process (Fig. 1) was modelled using Aspen HYSYS® process simulator, by considering the Peng-Robinson equations of state (EOS) fluid package. The simulation was based on operating conditions from previous literature (Chebbi et al., 2010; Chebeir et al., 2019), considering the following inlet specification: natural gas nominal flowrate of 4980 kmol/h, at a pressure of 5818 kPa and temperature of 35 °C, and a gas inlet composition with low liquid content (Table 1). In this separation process, the demethanizer column (T-100) is the central unit, consisting of a 30-stage multicomponent distillation column with a reboiler (E-103). This unit is designed to separate the methane, the lightest component, from the heavier portion of the gas. For this reason, the CRR plant has to operate under cryogenic conditions with a nominal temperature in the top tray of approximately -116 °C.

In the CRR process (Fig. 1), the raw natural gas is initially cooled using the heat exchanger (E-100) and chiller (E-101). The resulting liquid portion is separated by the flash tank (TK-100) located before the demethanizer column (T-100). Part of the liquid is then mixed with vapor leaving the flash tank (TK-100), which is cooled by the heat exchanger (E-102), expanded through the Joule-Thompson valve (JTV-101), heated in heat exchanger (E-104) and fed into the 2nd stage of the column. The other portion of the liquid is depressurized by the Joule-Thompson valve (JTV-100) and fed into the 26th stage. The remaining vapor goes through a turboexpander (TE-100) and is fed into the 8th stage of the column. Part of the demethanizer top product is compressed with the cryogenic compressor (K-102), providing ethane-free reflux. This stream contacts the remaining ethane and heavier compounds in the vapor leaving the top of the column and entrains them into the descending liquid (Getu et al., 2013). This reflux is cooled by heat exchange with the inlet stream and expanded through the Joule-Thompson valve (JTV-102) to the column's top pressure. The overhead product, which is mainly methane, is then recompressed for commercial use after serving as a refrigeration fluid for E-100 and E-102 heat exchangers. Finally, the NGL bottom product of the demethanizer column undergoes further separation in the fractionation train.

Among the different control loops in the process, the temperature controllers have a key role in ensuring the product specifications in terms of ethane recovery and methane impurity in the bottom product of the demethanizer. The indirect control of ethane recovery is achieved by regulating the temperature (TIC-100) in the separator upstream of the demethanizer and manipulating the cooling fluid flowrate of the chiller (E-101). The indirect control of methane composition in the column bottom product is attained by controlling the column temperature at stage 28 (TIC-101) by manipulating the reboiler duty (Mandis et al., 2021).

3. Methodology for development of a DDT architecture

In this work, neural network architectures were developed using deep learning techniques for the achievement of a demethanizer surrogate data-driven model. In this sense, RNNs were employed for the

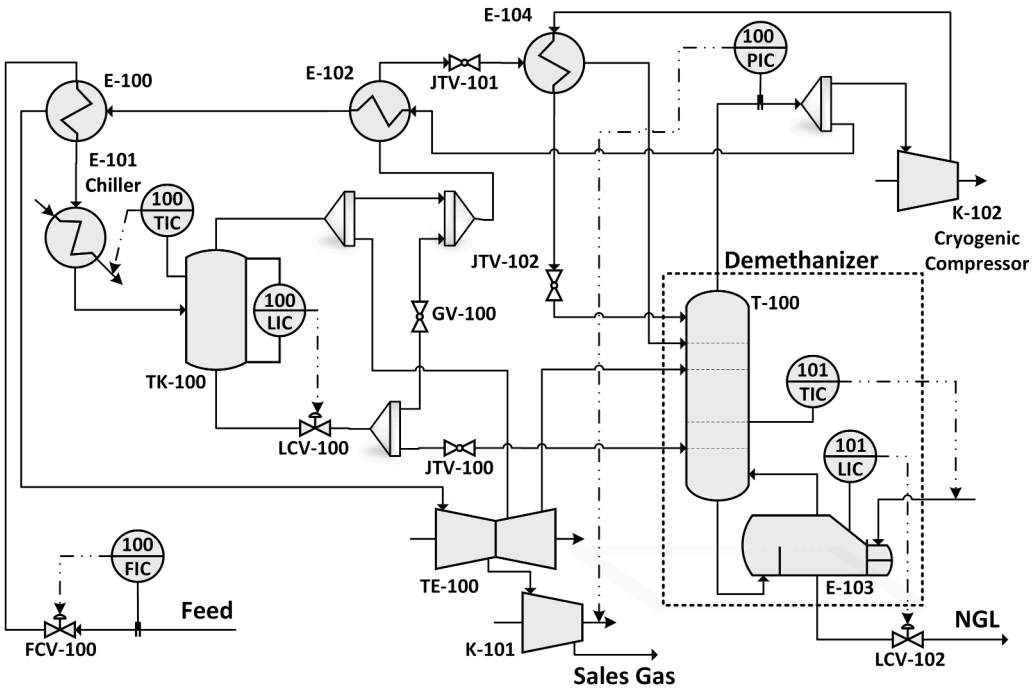

Fig. 1. CRR separation process scheme with control loops.

Table 1
Feed composition.

Components	Mole fractions
Nitrogen	0.01
Methane	0.93
Ethane	0.03
Propane	0.015
Butanes	0.009
Pentanes	0.003
Hexanes	0.003
% C_{2+}	6

development of the neural models due to their capability to account for the current values of relevant variables and their historical values and past outputs. Amid these types of neural networks, LSTM networks have the major advantage of avoiding a predominant drawback called the vanishing gradient problem. Therefore, LSTM neural networks have been implemented in the development of the DDT architecture.

3.1. Demethanizer digital twin architecture (DDT)

To formulate and implement a digital surrogate model for the demethanizer column, abundant data from the real process is required. Due to the lack of information from real separation plants, the data acquisition problem was addressed through the utilization of the process simulator Aspen HYSYS®. In this simulation environment, the CRR process operation was dynamically simulated (Mandis et al., 2022), which allowed the collection of all the data related to the demethanizer column, including compositions, temperatures, and pressures in the reboiler and all the column stages.

To simplify the complex task of modelling a multicomponent column for the realization of a DDT, the challenge was approached through decomposition into three distinct tasks, aimed at achieving a proper estimation considering as inputs only available data. The first task involved the development of a model for the estimation of the top and bottom product compositions. The second task was the formulation of a neural model for the column internal and external flows estimation. To achieve this, two distinct LSTM neural network models were developed. Finally, the third task consisted of the development of a neural network model capable of approximating the column's internal behavior. For this reason, three different neural network model architectures were created and compared in their ability to estimate the internal separation operations within the column. The objective was to determine the optimal structure in terms of estimation performance. These models were then integrated into a comprehensive architecture to develop the final data-driven model for the demethanizer column, referred to as the DDT. This architecture relies solely on readily available variable measurements as input data, ensuring easy and cost-effective implementation. Detailed descriptions of the neural model architectures used to realize each component are provided in the following subsections.

3.1.1. Neural model for product composition

To estimate the compositions of the demethanizer's top and bottom products, LSTM neural networks were used to develop a neural model called Neural Model for Product Composition (NMPC) estimation. As shown in Fig. 2a, the architecture of NMPC incorporates two LSTM cells. Each cell is responsible for calculating the compositions of one of the product streams. The input data shared by both cells comprises the pressure and temperature of all column feed streams, as well as the top and bottom product streams. Additionally, it encompasses the reboiler duty, power of the cryogenic compressor, and the pressure and temperature of the flash tank (TK-100). For each network, input data were selected based on a knowledge-based approach, considering variables with the strongest correlation with the desired outputs. Moreover, only easily measurable variables in real plant scenarios were provided as inputs for the integrated model.

3.1.2. Neural model for column flow

To estimate the top and bottom product streams as well as the reflux and the boilup streams, a neural network model, referred to as Neural Model for Column Flow (NMCF) estimation, was developed. The architecture of the NMCF estimation, schematized in Fig. 2b, involves the use of three interacting LSTM cells. The first cell is responsible for calculating the material balance around the cryogenic compressor and estimating the reflux stream. The second cell is responsible for calculating the total column material balance and gives as outputs the flows of

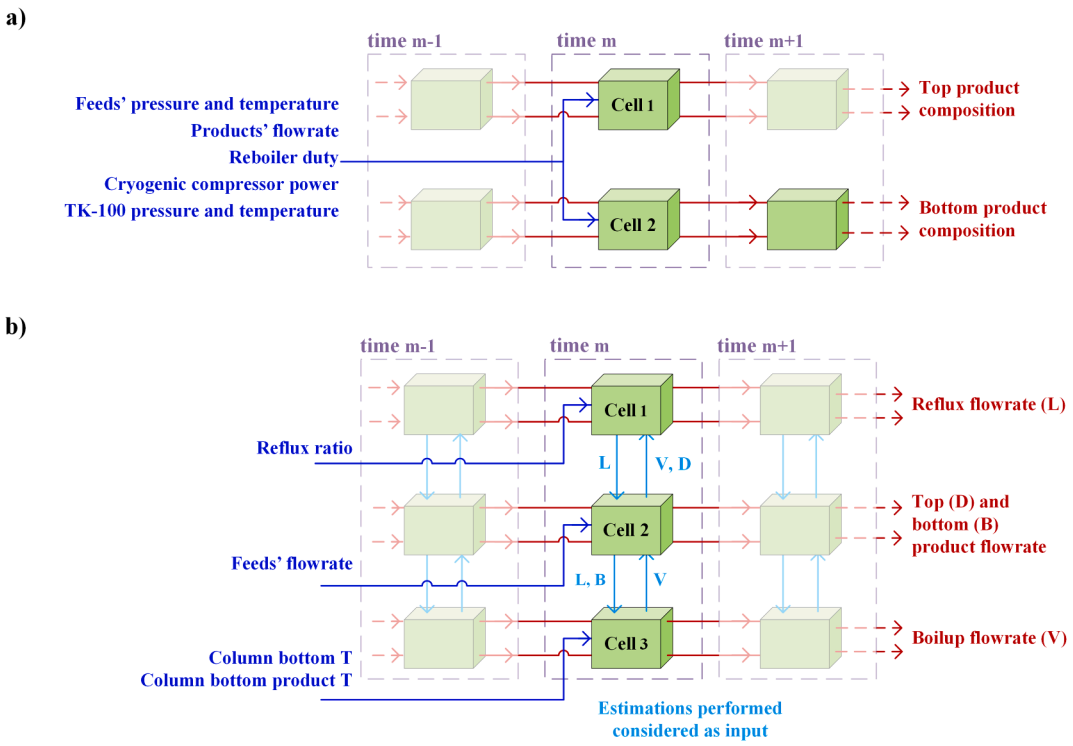

Fig. 2. Schematic representation of the NMPC (a) and the NMCF (b) recurrent structures.

the top and bottom column products. The third cell is responsible for estimating the material balance around the reboiler and is responsible for the estimation of boilup flow. The input data considered in the model was chosen with a knowledge-based approach and differed from cell to cell. For simulating the flow behavior in the column, the cells' outputs are also used as inputs. The first cell receives as input the reflux ratio in addition to the estimation of boilup and distillate. The second cell receives as input the measurement of the input flowrates and reflux and boilup streams estimations. The third cell receives as input column bottom temperature and bottom product temperature and bottom product and reflux streams estimations.

3.1.3. *Column neural model proposed architectures*

Different network architectures were considered to determine which could best emulate the dynamics of the demethanizer separation process. Three models were developed, each with a different way of exchanging information between LSTM units. The models share a common base layout that involves the utilization of a dedicated LSTM cell to mimic the operations of each stage of the distillation column and the reboiler. This allows the capture of the inherent dynamics of each column stage to be captured in the model. Furthermore, bidirectional connections have been considered to simulate the actions of internal flows within the distillation column, accounting for the interactions between adjacent stages. The bi-directional connection is realized by

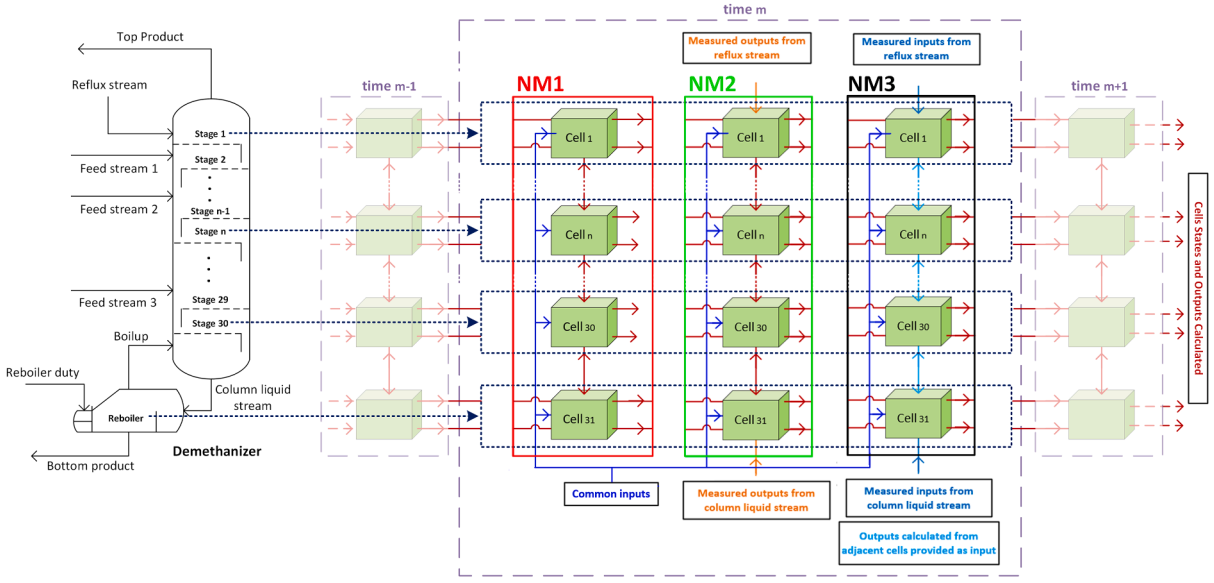

Fig. 3. Schematic representation of the column modeling highlighting the recursive structure of the three models NM1, NM2, and NM3 used alternately.

using two networks: a forward and a backward network. The forward network calculates outputs based on current inputs, and the hidden states and cell states are calculated to mimic the column's action from top to bottom. Similarly, the backward network calculates outputs based on current inputs, with hidden states and cell states calculated to mimic the column's action from the bottom to the top. The outputs of the two networks are then combined by exploiting two sigmoid activation functions (one for the hidden states and one for the cell states), resulting in the final output vectors and cell state vectors. This layout is the core of the neural network models developed to simulate the demethanizer operation at a specific instant of time over the length of the time sequence considered. The common basic layout is depicted in Fig. 3, along with the schematic of the three alternative recurrent architectures implemented for the estimation at time *m*.

The first developed model, referred to as neural model 1 (NM1), is designed with the aim of obtaining a single model capable of detecting and simulating the separation operations occurring in the distillation column. Considering developing individual neural models for each column stage using a dedicated LSTM network and then creating a single model representing the operations of the entire column, the interconnections between the utilized LSTM networks have been modified at the expense of the temporal connection between cells.

To better visualize the structure of the resulting model, a schematic of the corresponding inner LSTM cell is depicted in Fig. 4a, while a schematic representation of the NM1 layout is reported in Fig. 3. As can be seen in Fig. 4a, where the considered cell is referred to by the subscript *s*, the previous states for the cell are given by the hidden states and internal cell states calculated by the adjacent LSTM cells (indicated with subscripts $s-1$ and $s+1$) in the current time. Moreover, the usual LSTM time connection is only considered in the terminal LSTM cells of the network. For inner cells, the dependence on the states calculated by the same cell in previous instants is dropped and the only temporal dependencies are given by the information stored in the hidden cell state. Regarding the time-dependent input tensor, this retains the same information for each cell of the network.

The second neural model, referred to as neural model 2 (NM2), has been developed to account for the neglected time dependence in the inner cells of NM1. With this purpose the previous model layout has been modified while maintaining the preexisting connection that mimics the progression of the internal flows in the column. The modification applied are observable in Fig. 3 where both NM1 and NM2 layouts are shown, and in Fig. 4c where a specific inner LSTM cell is depicted. As can be seen in Fig. 4c, in the considered cell (also in this case referred to by the subscript *s*), the previous states for the cell are provided by combining the actual previous hidden states and internal cell states ($h_{t-1,s}$ and $c_{t-1,s}$) with the hidden states and internal cell states calculated by the adjacent LSTM cells (indicated with subscripts $s-1$ and $s+1$) in the current time. In addition, for the connections of the network terminal cells mimicking the column top tray and the reboiler, the calculated hidden states from the missing adjacent cell ($h_{t,0}$ and $h_{t,32}$) are replaced by the actual measurement of the corresponding variables of the reflux and the liquid stream entering the reboiler respectively for the first and the last cell of the network.

The third presented model for the demethanizer column, referred to as neural model 3 (NM3), is a modification of the NM2. In this case, the connection between a cell and the corresponding adjacent cells is realized by providing the calculated hidden states of cell_{s-1} and cell_{s+1} as inputs for the cell_s in addition to the time dependent common input vector. The temporal connection between the cells corresponds to the conventional LSTM cell connection, considering as previous outputs the hidden and internal cell states calculated in the previous instant of time by the same cell. The described cell structure is schematically shown in Fig. 4b, while the described model layout, responsible for estimating the demethanizer operation at a particular time, is shown in Fig. 3 along with the other models. In this model, the actual measurement of the corresponding variables of the reflux and the liquid stream entering the reboiler are provided as inputs, respectively, for the first and the last cell of the network instead of the calculated hidden states from the missing adjacent cell. To ensure comprehensiveness, we have included the mathematical equations for all cell architectures depicted in Fig. 4 within the supplementary material.

4. Results and discussion

The presented neural models were developed using the open-source library PyTorch and implemented and trained using Python

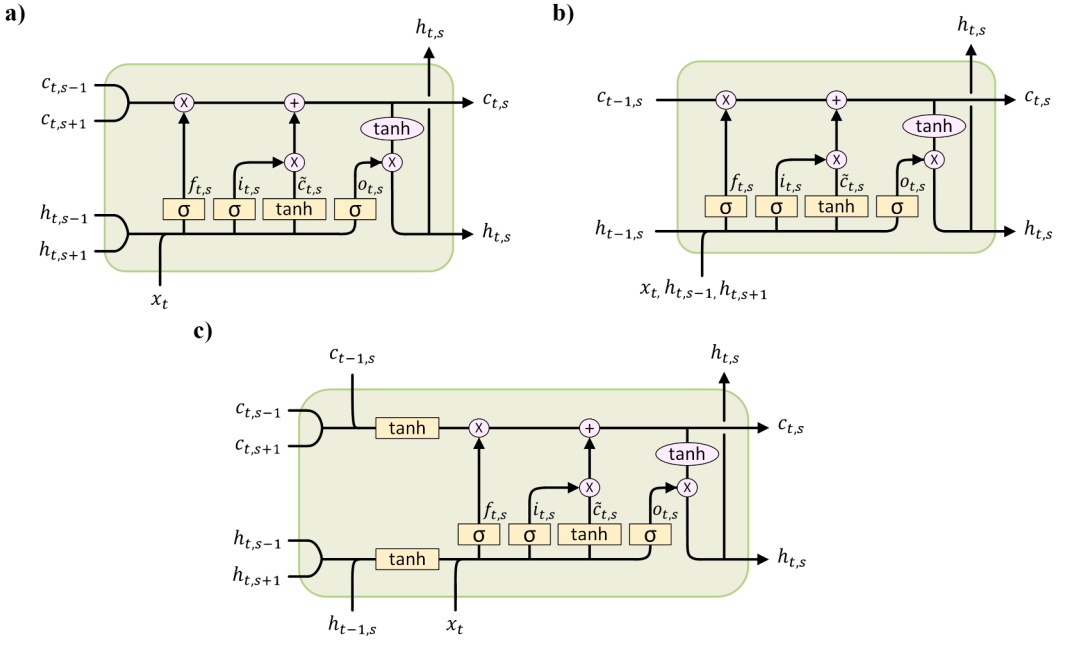

Fig. 4. Schematic representations of an inner LSTM cell for NM1 (a), an inner LSTM cell for NM3 (b), and an inner LSTM cell for NM2 (c). The schemes reported show a combination of two cells combined with the bi-directional LSTM method. The vectors with subscript $s+1$ are forwarded to the backward network cell, while the vectors with subscript s-1 are forwarded to the forward network cell.

programming (Paszke et al., 2019). The data used for training, testing, and validating the models along with the functions used for training the different models are described in the following subsections.

4.1. Datasets

The datasets employed for training and testing the developed neural models were simulated by considering two months of plant operation of the CRR separation process under temperature control (Fig. 1) simulated in the process simulator Aspen HYSYS®. To generate these sets of data, ramp changes of varying amplitude and duration in the plant feed flowrate were applied in the process simulator. The imposed variations were used to emulate the daily demand for natural gas, which exhibits an increasing trend during daylight hours and a decreasing trend during night hours (with variation peaks of 10% of the plant feed nominal value). Of all recorded datasets, 80% were used for network training, while the remaining 20% were used for the testing campaign. To validate the performance of presented neural network models, two days (48 h) of plant operation were simulated by applying ramp variations of 5% in the plant feed flowrate nominal value. All the datasets used were registered by considering a sampling time of 20 s, with a total of 211,600 time observations for all recorded variables in the training set, 52,900 time observations in the test set, and 8,640 time observations in the validation set. No measurement delays were considered as the data were treated as historical plant data. To adopt a more realistic approach, we introduced white noise into the dataset to simulate measurement noise, utilizing the maximum values outlined in Table 2. Furthermore, we applied Max-Min normalization to both input and output data, aiming to improve model accuracy. This normalization helped in mitigating the potential disproportionate impact of certain variables over others, stemming from variations in their magnitudes.

4.2. Model training

The loss function considered in the training optimization problem is given by the Mean Square Error (MSE) defined as reported in Eq. (1).

$$MSE = \frac{1}{N_b}\sum_{j=1}^{O}\sum_{n=1}^{N}\sum_{m=1}^{M}\left(\hat{y}_{j,n,m} - y_{j,n,m}\right)^2 \quad (1)$$

where N_b corresponds to the number of observations in the current iteration; O is the number of considered outputs in a given cell, while the general output is denoted by j; N represent the length of the selected time sequence; M is the number of cells employed by the different models; $\hat{y}$ and y denote a general hidden state calculated by the considered model and the corresponding measured value, respectively.

The loss function is calculated at each iteration in which the back-propagation through the time algorithm is used to calculate the associated gradient with respect to weights and biases. The obtained gradient is then used for the parameter updating performed by the Adaptive Moment Estimation optimizer (ADAM) (Kingma and Ba, 2017). The network hyperparameters were obtained by multiobjective optimization using the NSGA-II algorithm available in the pymoo package (Blank and Deb, 2020). The hyperparameters that were optimized include batch size, sequence length, hidden size and learning rate. The batch size determines the number of time-sequential data series processed by the network in each iteration. The sequence length corresponds to the length of the time sequence considered and determines the number of cells performing calculations in each iteration. The hidden size determines the number of cell states considered by the network cells. Finally, the learning rate defines the step size used by the optimization algorithm in the learning process. The resulting hyperparameters used are shown in Table 3.

4.3. Demethanizer digital twin estimation performance

In this section, the results obtained from the validation campaign for the NMPC and NMCF models are presented. Additionally, it is included a comparison of the estimation performance of the NM1, NM2, and NM3 models. Following this, the results for the DDT are compared to those obtained with the process simulator.

4.4. NMPC estimation performances

The results of the control performance achieved in the validation campaign from the NMPC neural model are depicted in Fig. 5 for the estimation of the distillate and bottom product compositions. For the sake of brevity, only the composition profiles of the two components depicting the higher variations are here included. A complete set of results, comprehending the omitted variables is provided in the supplementary material. The results are obtained for inlet feed flowrate variations (increasing and decreasing ramp variation with an amplitude of 5% in the CRR process unit).

The left panels of Fig. 5 depict the comparison of the measured (light blue line) and the actual (dashed red line) profiles respectively of ethane and propane top composition, with the estimated (black line) profiles obtained with the NMPC. In spite of significant measurement noise compared with the actual concentration changes, the model is effective at filtering it out. The obtained estimations accurately approximate the actual transient trend in ethane and propane concentrations, as well as the new steady state conditions reached in the column top product. The comparison among the measured, actual, and estimated profiles of methane and ethane concentrations at the bottom is depicted in the right panels of Fig. 5. Again, the estimation obtained by the developed model shows good capabilities in filtering the considered measurement noise and the estimate is perfectly capable of reconstructing the transient evolution of ethane in the bottom product of the demethanizer. In this case, the model also manages to filter out the measurement noise, providing a more accurate estimate of the actual profile of the plant at steady-state as well as at transient times following the plant's response caused by the increased flow rate. Nevertheless, the performance of the model in estimating the concentrations in the top product slightly deteriorates as heavier components are considered and thus as the concentrations of the compounds considered in the estimation approach zero (see Supplementary Material). Overall, the NMPC model was able to successfully and accurately reconstruct the transient profiles obtained for all the compositions in the bottom product of the demethanizer column except nitrogen whose presence is, however, negligible.

Table 3
Neural Models' optimized hyperparameters

	NMPC	NMCF	NM1	NM2	NM3
Batch size	16	5	11	26	17
Sequence length	11	10	12	7	29
Hidden size	18	6	14	15	16
Learning rate	10^{-2}	10^{-2}	$10^{-1.51}$	$10^{-1.54}$	$10^{-1.53}$

Table 2
Values of maximum measurement noise considered for input and output data in train, test, and validation datasets

	Duty	Temperatures	Pressures	Flows	Concentrations
maximum measurement noise	1.3%*	0.1 [°C]	1%*	1%*	2%*

* Referred to the maximum value of the considered variable for the given column stage.

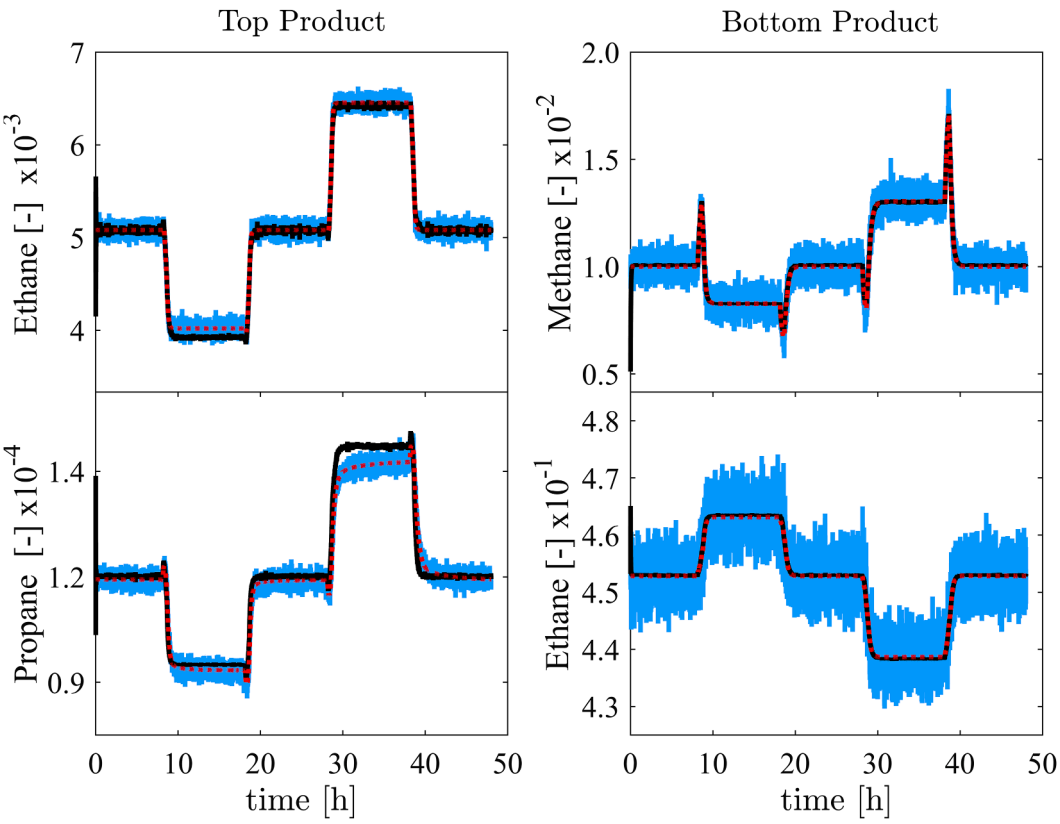

Fig. 5. Comparison of the estimated transient profiles (black lines) obtained for top product and bottom product compositions with the actual (red lines) and measure profiles (light blue lines) during the validation campaign.

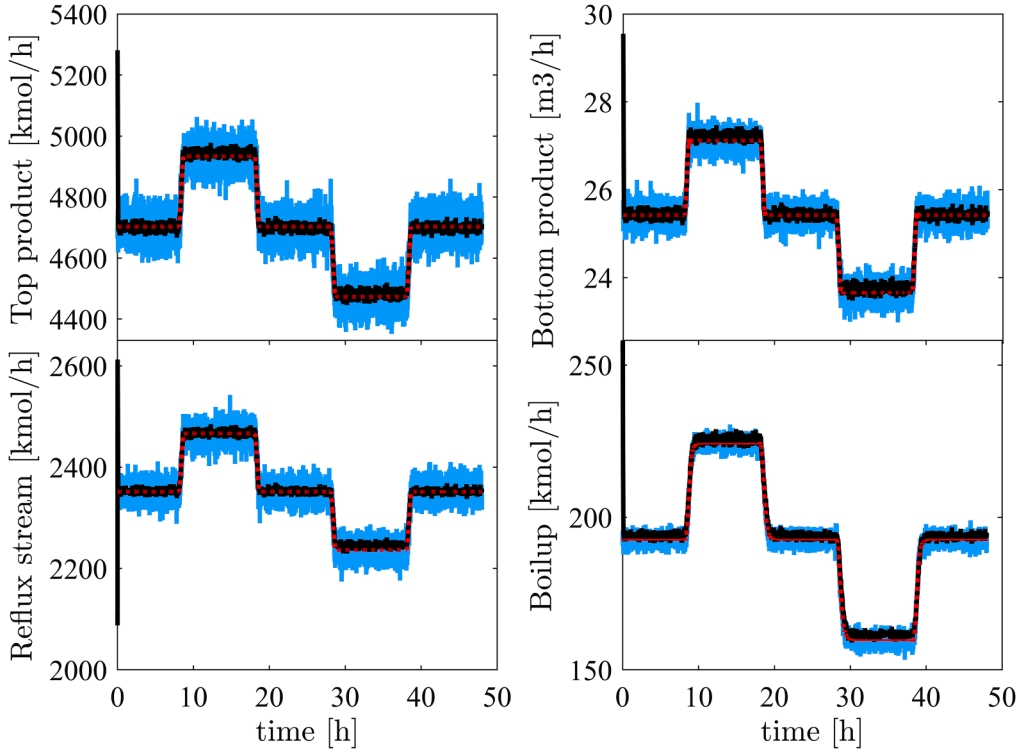

Fig. 6. Comparison of the estimated transient profiles (black lines) obtained for distillate, bottom, reflux and boilup streams with the actual (red lines) and measure profiles (light blue lines) during the validation campaign.

4.5. NMCF estimation performances

The validation campaign results obtained with the NMCF model for the reconstruction of the dynamic profiles of the top and bottom product flowrates as well as the reflux and boilup flowrates are depicted in Fig. 6.

The outcomes obtained with this neural network model depict similar results to those obtained for the model for the estimation of product compositions. Again, it can be observed that, for all the considered estimation outputs, the model demonstrates good filtering capabilities in the estimated profiles with errors lower than the measurement noise considered. The obtained estimates prove to be very accurate and able to follow the actual output profiles of the demethanizer column.

4.6. Selection of column neural model architecture

The performances of the three models developed to select the best architectures to estimate the column's internal behavior are evaluated considering the loss function (Eq. (1)) as well as comparing the temperature and concentration profiles obtained at steady state and during transients for the validation campaign. For the sake of brevity, only the results obtained in the presence of the decreasing flow disturbances are shown. The values obtained for the loss functions during training, test, and validation campaign for the considered models, are reported and confronted in Table 4.

As it is possible to visualize considering the loss function values obtained at the end of the training campaign, the NM2 model registers a smaller error, followed by the NM3 model, while the highest value is obtained for NM1. The same results are obtained with the loss function values obtained during the test and validation campaign. This suggests that the changes applied to the NM1 model for maintaining the temporal connection between cells led to an improvement in the column model. Considering that the NM2 show the lowest loss function values in all the campaigns this may be the best model among those proposed to approximate the operations of the demethanizer column.

The ability of NM1, NM2 and NM3 to reconstruct the temperature and key-component profiles are evaluated under the worst-case variation represented by a decrease in the feed plant nominal value. Fig. 7 illustrates the predicted temperature (top-left panel) and ethane composition (bottom-left panel) steady state column profiles, along with the related predicted transient profiles in the tray where the deviation from the true steady state value is maximum. Again, complete set of results including other components composition and dynamics behavior of temperature in selected trays are provided in the supplementary material. Considering the upper left panel of Fig. 7, it is possible to visualize that all three models are generally able to reconstruct the actual column temperature profile. However, in the last stages of the column, the estimation deviates from the trend, with a major deviation registered for NM1.

To analyze further the temperature estimation obtained in this section of the column, the time evolution of the temperature in the 29th column tray is reported in the right upper panel of Fig. 7. It is evident that none of the presented models can reconstruct the transient temperature dynamics in the tray with an error lower than the measurement noise. Specifically, the estimation obtained with NM1 not only exhibits a noisier estimate but also deviates from the actual profile, reaching a maximum deviation of approximately 5 °C. The deviations in the estimations obtained with NM3 are slightly worst compared to NM2, with NM3 showing a maximum deviation of around 2 °C and NM2 around 1.8 °C. The data obtained for stage 28th and 30th are reported in the supplementary materials. Here, it is important to highlight that NM2 always outperforms NM3 except in 28th, where NM2 exhibits a maximum error of 0.8 °C while NM3 shows a maximum error of about 0.7 °C. NM2 has always a deviation lower than 1 °C except in the 29th tray where, as observed, the maximum error reaches the value of 1.8 °C. NM3 exceeds the value of 1 °C in the 29th tray with a maximum deviation of 2 °C and in the 30th tray with a maximum deviation of 1.5 °C. It is worth noting that most of the column temperature variation occurs in the latter three stages of the column, with about 80 °C of variation due to the presence of methane in the gas phase, causing the temperature estimation problem to become more challenging.

Bottom panels of Fig. 7 show the results of the estimation performed by the three compared models on ethane composition steady state column profiles, and transient profiles obtained in the 29th tray where the registered deviations from the actual profiles are maximum. As it is possible to visualize, again, models NM2 and NM3 are perfectly able to accurately reconstruct the column concentration steady state profiles, however, model NM1 shows a poorer prediction. Upon observing the transient profiles acquired for the 29th tray (located at the right bottom of Fig. 7), it becomes apparent that NM1 is capable of qualitatively tracking the transient behavior. However, it deviates from the composition trend, exhibiting an average value that demonstrates greater variation as it approaches the new steady-state value, beyond what can be attributed to measurement noise.

Analyzing the transient profiles corresponding to ethane concentration profiles in the 28th and 30th tray of the column (see Supplementary Materials), similar results are obtained for the 30th tray. NM1 model prediction shows an error greater than the considered measurement error, whereas the other two models produce estimations that remain within the measurement range and exhibit a less noisy response. For what concern the results obtained for the 28th tray, NM1 also shows poorer estimation performances when compared with the other two models. Indeed, the prediction obtained with NM2 and NM3 has lower noisy estimations compared to NM1. However, NM1′s ability to mitigate measurement noise allows its prediction to lie within the measurement profiles.

In the present results, the NM1 turns out to have the worst prediction performance when considering most of the target variable estimations. This is probably due to the reduced temporal dependency in the structure of this model, as well as the impossibility of benefiting from the presence of the actual measures of the output variables used by the end cells of NM2 and NM3. Given the marginal differences in the performance of NM2 and NM3, the choice of the optimal model to represent the internal behavior of the column relied on comparing the temperature profiles predicted by these two models. This choice was driven by the acknowledgment that noisier estimations are associated with this variable. Notably, NM2 exhibits superior temperature estimation performance in the final stages of the column and the reboiler, which is the most critical section. It provides less noisy temperature estimations compared to NM3. This observation aligns with the higher MSE values observed for NM3.

The NM2 seems to be a more suitable neural network model since it can adequately represent and predict the dynamic behavior of the whole column. Therefore, this model architecture represents a viable candidate for estimating the demethanizer column's internal behavior in the digital twin development.

Table 4
Loss function evaluated for the three neural models test and validation campaign.

	NM1	NM2	NM3
Training loss	1.518E-03	1.324E-03	1.368E-03
Test loss	1.636E-03	1.324E-03	1.366E-03
Validation loss	2.273E-03	2.125E-03	2.133E-03

4.7. DDT estimation performances

As NM2 showed the best estimation performances of the target variables, this model architecture was thus selected for the development of the demethanizer digital twin (DDT). To mitigate potential delays associated with online composition measurements (which could arise

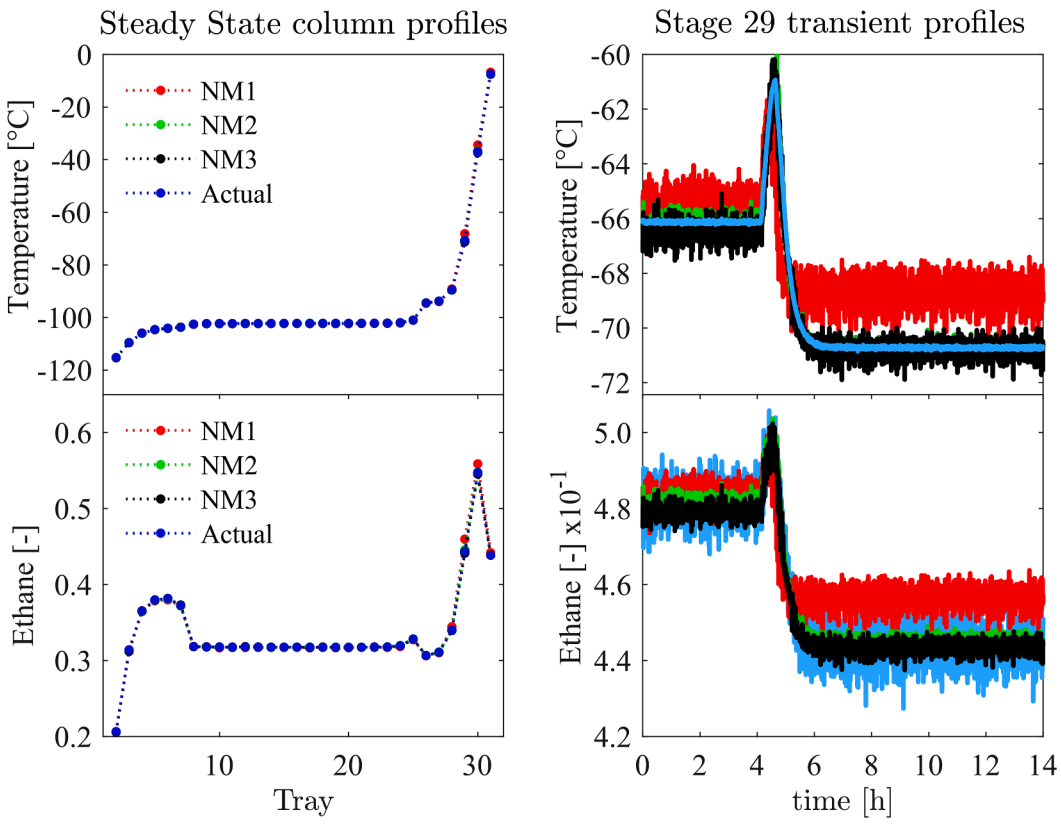

Fig. 7. Temperature and ethane profiles at the new steady state (left panels); transient profiles of 29th stage temperature, 29th stage ethane (right panels) obtained in the validation campaign for measure (light blue lines), NM1 (red lines), NM2 (green lines) and NM3 (black lines) under decreasing variation of 5% in the plant feed nominal value.

from the use of chromatography devices) in the digital twin and consequently degrade estimation performance, the NM2 architecture has been integrated with NMPC. Furthermore, the NMCF was employed to estimate reflux and boilup flows, and these estimates were incorporated into the model inputs. This addition enhances the model's capacity to provide essential information regarding the column's internal flows. It also leverages the filtering capabilities of the NMCF, ensuring a more reliable and accurate estimation. In this way, it was possible to obtain a complete model of the column, whereby it was possible to estimate the demethanizer column operations employing only readily available measurements.

To compare the results obtained by the DDT, the profiles obtained in the most critical area of the column, namely the column's bottom and the reboiler, are analyzed. For the sake of brevity, only the profiles obtained in the validation campaign in response to the 5% decrease in the plant inlet flowrate are reported. In Fig. 8, the plots on the upper panel show the methane concentration estimations, clearly demonstrating the effective reconstruction of variable's dynamics with reduced noise achieved by the DDT model. Similar results were obtained in the temperature estimation for tray 29 and the reboiler, as shown in the plots reported in the bottom panels, where an excellent agreement between measure data and the values obtained with the digital twin is observed.

4.7.1. Validation transient profiles

To have a complete overview of the results obtained by the DDT model, developed for the prediction of demethanizer column operations, the comparison of the transient profiles obtained for column pressure, temperature and key components composition and the actual transient profiles has been evaluated. For sake of brevity, estimated temperature and ethane concentration along with the deviations from to the actual value are reported in Fig. 9, whereas the results regarding the estimation of the other variables are represented in the supplementary material.

The bottom panels of Fig. 9 display the temperature profile and the corresponding estimation error obtained during the transient time considered. The temperature profile obtained with the DDT consistently predicts the overall trend of the actual transient column temperature profile, with estimation errors ranging between 0.1 °C and below. The first tray shows a slightly higher error, but it still remains below 0.2 °C. Progressing towards the ends of the column, starting from the 28th tray, the estimation error increases but stays below 1 °C. Although the error is larger in the later stages of the column, it is essential to consider that in this region, there is a temperature variation of 80 °C. Therefore, an error of less than one degree represents a small percentage compared to the total variation.

Regarding the ethane concentration, bottom panels of Fig. 9 showcase the profile and estimation error, respectively. This reveals that the DDT model accurately predicts the transient profiles of the actual ethane composition for all the column trays. The estimation error increases in correspondence to the column region where most of the variations for the considered component occur. Table 5 shows the MSE and the maximum deviation, obtained by considering the predictions of the transient composition profiles in all the column trays and the actual profiles, for non-key components and methane in the validation campaign. The proposed neural model was able to accurately reconstruct all the composition transient column profiles, with extremely low values of MSE and maximum estimation error. The latter always results under the assumed measurement error, confirming the model's ability to filter the measurement noise.

While the DDT model effectively predicts the time evolution of the separation in the demethanizer column, its application for real-time process monitoring and control may be susceptible to delays in communication or errors in information exchange, potentially impacting its robustness. This vulnerability arises from its reliance on real-time

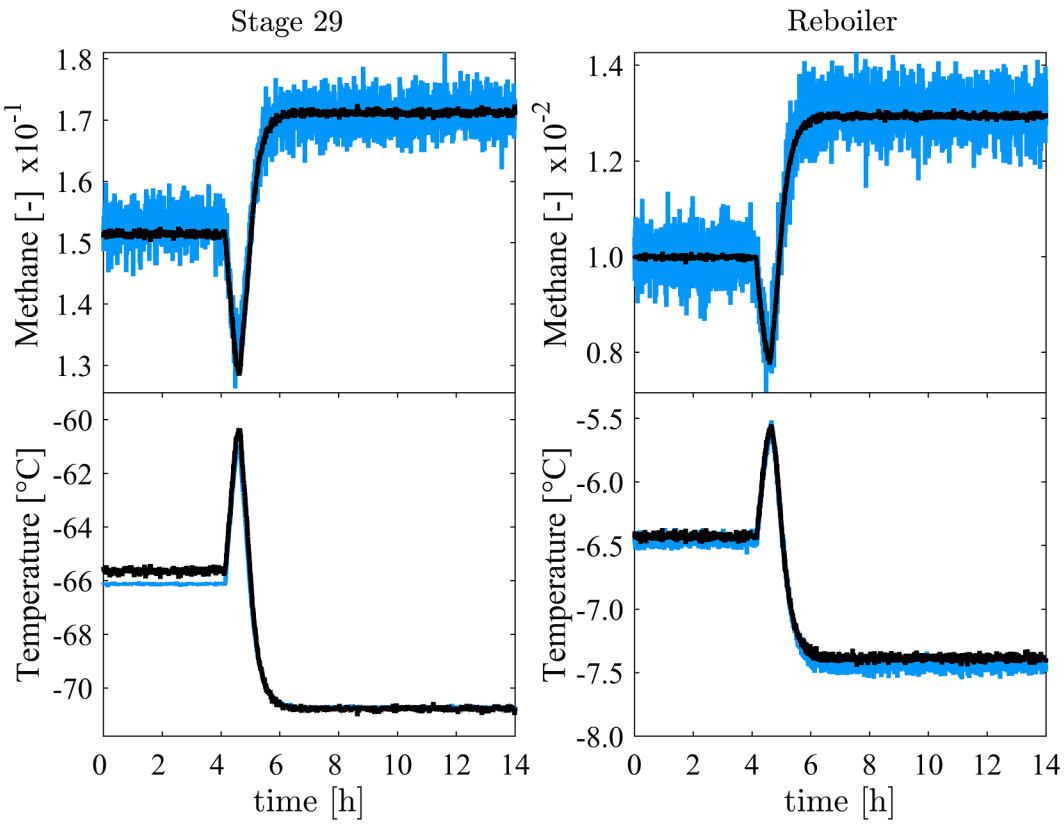

Fig. 8. Comparison of transient profiles of methane composition and tray temperature: 29th stage (left panels respectively), reboiler (right panels) obtained in the validation campaign for measure (light blue lines) and DDT (black lines) under a decreasing variation of 5% in the plant feed nominal value.

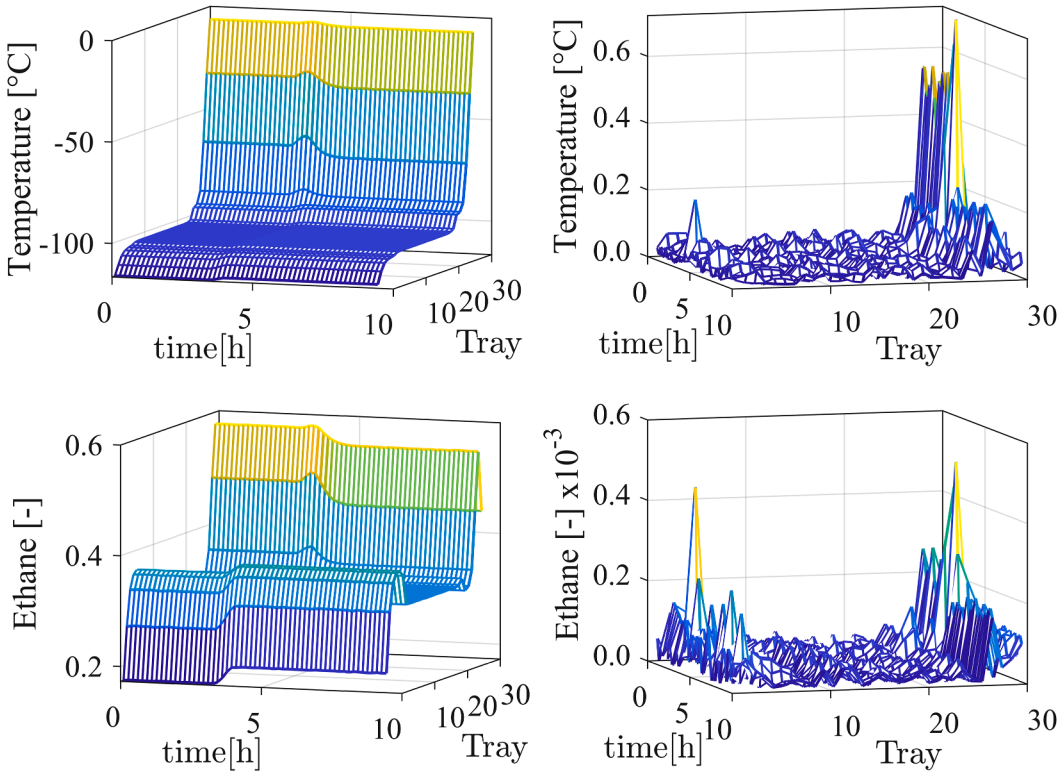

Fig. 9. Transient column profiles obtained under a decreasing variation of 5% in the plant feed nominal value for: estimated temperature and corresponding estimation error (upper panels); estimated ethane concentration and corresponding estimation error (bottom panels).

Table 5
Mean square error and max error obtained for methane and non-key components in the validation campaign of the DTT under a decreasing variation of 5% in the plant feed nominal value.

	Methane	Propane	Butanes	Pentanes	Hexane
MSE	3.69E-7	7.24E-8	1.22E-8	1.49E-9	2.57E-9
Max Error	0.0039	0.0011	7.62E-4	2.60E-4	2.64E-4

data and communication for accurate estimations. To mitigate the potential impact of such issues on the correct functioning of the digital twin, continuous comparisons should be made between the model's estimations and the currently measured outputs in the plant, including product pressures and temperatures, as well as available temperature measurements in the column, and there should also be a mechanism to automatically disconnect the neural model from real-time operations if the estimation of critical variables exceeds an acceptable deviation.

4.7.2. Runtime comparison

The development of the demethanizer column digital twin (DDT) is driven by the need to have a real-time process monitoring and control tool. To assess the achievement of this goal, a comparison of calculation times between the Aspen HYSYS® simulation and the DDT was performed. The DDT calculation time was recorded within the Python environment during the validation campaign. On the other hand, the results reported for Aspen HYSYS® represent average calculation times, as the actual time recording is not possible in the Aspen HYSYS® environment, making the registration of actual computation times impractical. For the Aspen HYSYS® simulation, the average computation time was calculated by dividing the total time taken for the simulation by the number of samplings considered for each variable. The average computation time obtained for the Aspen HYSYS® simulation was 0.3 s, while for the DDT was 0.0183 s. This indicates that the DDT offers a substantial reduction in calculation time, approximately 93.9% faster than Aspen HYSYS®. Specifically, the DDT was able to estimate around 48 h of plant operation in about 2.7 min, whereas the Aspen HYSYS® simulation required approximately 43 min. The machine employed to perform the calculations has an Intel Xeon Processor operating at 4.00 GHz with 32.0 GB of installed RAM.

5. Conclusions

This paper delves into the potential of employing a data-driven methodology alongside with neural networks to model the dynamic operation of a demethanizer column. The main objective of this work is the development of a DDT architecture, which is capable of enabling a more precise monitoring and control of the column's operating performance. Moreover, the data acquisition is critical for the generation of the different neural models that constitute the DDT architecture. For this, a CRR separation process is modelled in the process simulator Aspen HYSYS©.

The proposed DDT architecture encompasses three distinct neural models: one for approximating the column's material balance (NMCF), another for estimating the column product composition (NMPC), and a model designed to predict demethanizer separation across all trays and reboiler. To identify the most suitable network architecture capable of accurately estimating the separation processes across all the trays of the column, various neural network models with diverse interconnections between network cells were compared. The NM2 model, which accounted for temporal dependencies within its inner cells and integrated real-time flow and concentration measurements, demonstrated superior estimation performance and was selected for the implementation of the DDT. In order to craft a practical digital twin applicable to real-world plant scenarios and reliant on easily measurable variables, the DDT integrated NMCF's estimations of reflux and boil-up flows as inputs with the NM2. This strategic approach ensures the digital twin's effective implementation in industrial settings while utilizing readily available measurements.

The resulting digital twin exhibited remarkable accuracy in predicting desired output trends and demonstrated excellent noise filtering capabilities. Additionally, a comparison between calculation times of the Aspen HYSYS® simulation and the DDT was conducted, revealing a substantial reduction in calculation time offered by the DDT. This establishes the DDT as a valuable tool with the capacity to accurately predict the dynamic behavior of the demethanizer column. Moreover, the digital twin can make a significant contribution to the optimization, monitoring, and control of the NGL recovery process.

A future direction of this research will be focused on the possibility of embedding the developed digital twin with economic analysis for the online optimization of the process. This will enable a more comprehensive decision-making operation that enhances the sustainability of the NGL separation process by ensuring optimum resource management. Additionally, future works may assess the estimation performances in presence of even more realistic situations when datasets are corrupted with non-Gaussian noise, as in the case of real industrial measurements.

CRediT authorship contribution statement

Marta Mandis: Formal analysis, Methodology, Software, Visualization, Writing – original draft. **Roberto Baratti:** Conceptualization, Formal analysis, Writing – review & editing. **Jorge Chebeir:** Writing – review & editing. **Stefania Tronci:** Conceptualization, Formal analysis, Supervision, Writing – original draft, Writing – review & editing. **José A. Romagnoli:** Visualization, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.compchemeng.2024.108591.

References

Arnold, K., Stewart, M., 1999. Surface Production Operations, Design of Gas-Handling Systems and Facilities. Elsevier. Vol. 2.
Bi, X., Qin, R., Wu, D., Zheng, S., Zhao, J., 2022. One step forward for smart chemical process fault detection and diagnosis. Comput. Chem. Eng. 164, 107884.
Blank, J., Deb, K., 2020. Pymoo: multi-objective optimization in python. IEEE Access 8, 89497–89509.
Chebbi, R., Al Mazroui, K.A., Abdel Jabbar, N.M., 2008. Study compares C2-recovery for conventional turboexpander. GSP. Oil Gas J. 106 (46), 50–54.
Chebbi, R., Al-Amoodi, N.S., Abdel Jabbar, N.M., Husseini, G.A., Al Mazroui, K.A., 2010. Optimum ethane recovery in conventional turboexpander process. Chem. Eng. Res. Des. 88, 779–787.
Chebeir, J., Salas, S.D., Romagnoli, J.A., 2019. Operability assessment on alternative natural gas liquids recovery schemes. J. Nat. Gas Sci. Eng. 71, 102974.
Elman, J.L., 1990. Finding structure in time. Cogn. Sci. 14 (2), 179–211.
Ge, Z., 2017. Review on data-driven modeling and monitoring for plant-wide industrial processes. Chemom. Intell. Lab. Syst. 171, 16–25. Additional Information.
Getu, M., Mahadzir, S., Long, N.V.D., Lee, M., 2013. Techno-economic analysis of potential natural gas liquid (NGL) recovery processes under variations of feed compositions. Chem. Eng. Res. Des. 91 (7), 1272–1283.
He, C., You, F., 2014. Shale gas processing integrated with ethylene production: novel process designs, exergy analysis, and techno-economic analysis. Ind. Eng. Chem. 53 (28), 11442–11459.
Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9 (8), 1735–1780.

Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. J. Basic Eng. 82 (1), 35–45. https://doi.org/10.1115/1.3662552.
Ke, W., Huang, D., Yang, F., Jiang, Y., 2017. Soft sensor development and applications based on LSTM in deep neural networks. In: 2017 IEEE Symposium Series on Computational Intelligence (SSCI). IEEE, pp. 1–6.
Kherbeck, L., Chebbi, R., 2015. Optimizing ethane recovery in turboexpander processes. J. Ind. Eng. Chem. 21, 292–297.
Kidnay, A.J., Parrish, W.R., 2006. Fundamentals of Natural Gas Processing, 1st edition. Taylor and Francis.
Kidnay, A.J., Parrish, W.R., McCartney, D.G., 2011. Fundamentals of Natural Gas Processing, 2nd Edition. CRC Press, Boca Raton, FL.
Kingma, D.P., Ba, J. (2017). Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980.
Kwon, H., Oh, K., Choi, Y., Chung, Y., Kim, J., 2021. Development and application of machine learning-based prediction model for distillation column. Int. J. Intell. Syst. 36 https://doi.org/10.1002/int.22368.
Li, J., Qin, S.J., 2023. Applying and dissecting LSTM neural networks and regularized learning for dynamic inferential modeling. Comput. Chem. Eng. 175, 108264.
Lu, L., Jin, P., Karniadakis, G.E. (2019). Deeponet: learning nonlinear operators for identifying differential equations based on the universal approximation theorem of operators. arXiv preprint arXiv:1910.03193.
Luyben, W.L., 2013. NGL demethanizer control. Ind. Eng. Chem. Res. 52, 11626–11638.
Mandis, M., Chebeir, J., Tronci, S., Baratti, R., Romagnoli, J.A., 2021. Control of a natural gas liquid recovery plant in a GSP unit under feed and composition disturbances. IFAC PapersOnLine 54, 182–187. -3.
Mandis, M., Baratti, R., Chebeir, J., Tronci, S., Romagnoli, J.A., 2022. Performance assessment of control strategies with application to NGL separation units. J. Nat. Gas Sci. Vol. 106, 104763.
Manning, F., Thompson, R., 1991. Oilfield Processing: Natural Gas. PennWell Publishing Co., Tulsa, OK.
Mehrpooya, M., Gharagheize, F., Vatani, A., 2006. An optimization of capital and operating alternatives in a NGL recovery unit. Chem. Eng. Technol. 29 (12), 1469–1480, 2006.
Pang, C., Duan, D., Zhou, Z., Han, S., Yao, L., Zheng, C., Yang, J., Gao, X., 2022. An integrated LSTM-AM and SPRT method for fault early detection of forced-oxidation system in wet flue gas desulfurization. Process Saf. Environ. Prot. 160, 242–254.
Park, J.H., Khan, M.S., Andika, R., Getu, M., Bahadori, A., Lee, M., 2015. Techno-economic evaluation of a novel NGL recovery scheme with nine patented schemes for offshore applications. J. Nat. Gas Sci. Eng. 27, 2–17.
Paszke, A. et al. (2019). PyTorch: an imperative style, high-performance deep learning library. arXiv:1912.01703.
Pitman, R.N., Hudson, H.M., Wilkinson, J.D., Cuellar, K.T., 1998. Next generation processes for NGL/LPG recovery. In: Proceedings of the Seventy-seventh Annual GPA Convention Dallas.
Qu, X., Song, Y., Liu, D., Cui, X., Peng, Y., 2020. Lithium-ion battery performance degradation evaluation in dynamic operating conditions based on a digital twin model. Microelectron. Reliab. 114, 113857.
Tronci, S., Chebeir, J., Mandis, M., Baratti, R., Romagnoli, J.A., 2020. Control strategies for natural gas liquids recovery plants. Comput. Aided Chem. Eng. 48, 1291–1296.
Wang, P., Zhang, J., Wan, J., Wu, S., 2022. A fault diagnosis method for small pressurized water reactors based on long short-term memory networks. Energy 239, 122298.
Wu, Z., Alnajdi, A., Gu, Q., Christofides, P.D., 2022. Statistical machine-learning-based predictive control of uncertain nonlinear processes. AIChE J 68, e17642.
Wu, Z., Tran, A., Rincon, D., Christofides, P.D., 2019. Machine learning-based predictive control of nonlinear processes. Part I: theory. AIChE J 65, e16729, 2019.
Wu, Z., Rincon, D., Luo, J., Christofides, P.D., 2021. Machine learning modeling and predictive control of nonlinear processes using noisy data. AIChE J. V. 67, e17164.
Yuan, X., Li, L., Wang, Y., 2019. Nonlinear dynamic soft sensor modeling with supervised long short-term memory network. IEEE Trans. Ind. Inf. 16 (5), 3168–3176.
Zhang, S., Jiang, H., Jing, J., Qin, M., Chen, D., Chen, C., 2020. Comprehensive comparison of enhanced recycle split vapour processes for ethane recovery. Energy Rep. 6, 1819–1837.
Zeitz, M., 1987. The extended Luenberger observer for nonlinear systems. Syst. Control Lett. 9 (2), 149–156. VolumeIssuepages.
Zhu, W., Chebeir, J., Romagnoli, J.A., 2020. Operation optimization of a cryogenic NGL recovery unit using deep learning based surrogate modelling. Comput. Chem. Eng. 137, 106815.
Zhu, X., Ji, Y., 2022. A digital twin–driven method for online quality control in process industry. Int. J. Adv. Manuf. Technol. 119 (5), 3045–3064.