PROCEEDINGS OF THE 33rd European Symposium on Computer Aided Process Engineering (ESCAPE33), June 18-21, 2023, Athens, Greece

A Demethanizer column Digital twin with non-conventional LSTM neural networks arrangement

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*

Abstract

This work aims to develop a digital twin for a demethanizer column and provide a useful tool for monitoring and quality control of the NGL recovery process. For this purpose, a digital data-driven model was proposed to mimic real dynamics of a cold residue reflux (CRR) unit through the incorporation of physical knowledge. A non-conventional LSTM network arrangement was developed considering training test and validation data sets generated by the process simulator Aspen HYSYS®®. This simulation model was built by considering realistic measurement noises to mimic the actual measures in a real plant. The obtained surrogate model was evaluated considering its ability to recreate the operation of the actual distillation column, estimating the temperature and composition transient profiles of the bottom column product and every stage of the column. Overall, the model developed with the proposed LSTM network arrangement proves capable of successfully reconstructing the actual profiles of all the variables considered.

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

1. Introduction

Natural gas is a hydrocarbon mixture constituted mostly of methane and a variable fraction of heavier hydrocarbons known as natural gas liquids (NGL). The latter constitute a source of additional profit as can be sold separately with a higher selling price and used as industrial feedstock materials. Additionally, with NGL separation it is possible to reduce CO_2 emissions related to natural gas combustion. The NGL recovery process under steady-state conditions has been extensively studied in the literature (Kidnay et al., 2011; Chebbi et al., 2010; Getu et al., 2013; Kherbeck & Chebbi, 2015). The associated dynamics of the NGL separation process have been the subject of recent studies and generally focused on the dynamics and control of product quality indices (Luyben, 2013; Chebeir et al., 2019; Mandis et al., 2022). For this, it is essential to have information on the plant composition targets, which can be obtained with the use of expensive composition analyzers. These analyzers also introduce long-time delays in the control loops due to measurements delays. To overcome these drawbacks, the replacement of measured concentrations with accurate real-time estimations represents one of the most widespread approaches. In this type of task, temporal dependencies play a key role to obtain adequate estimations and with the goal of addressing time-dependent problems Recurrent neural networks (RNNs) (Elman, 1990) were developed. One of the most useful RNNs when addressing long time dependency is the Long-Short Term Memory network (LSTM) (Hochreiter & Schmidhuber, 1997), designed to overcome the problem of the vanishing gradient. LSTM networks have been applied to different types of problems in chemical engineering including digital twin realization for plant monitoring and control (Qu et al., 2020; Zhu & Ji, 2022). A digital twin (Grieves, 2015) consists of a virtual model of whole plants or individual units which is interconnected with the actual equipment. The availability of reliable digital surrogate models is of great interest in the process industry.This tool exchanges information in real-time, and it can predict the dynamic evolution of a process over time. For this reason, a digital surrogate can provide a useful tool for performance monitoring, predictive asset maintenance, production optimization and advanced process control. This work aims to develop a data-driven digital model for a demethanizer column of a CRR unit, which is realized by using LSTM neural network with a non-conventional network architecture. The main goal is to obtain a surrogate model capable of approximating the real dynamics of the column with a much shorter computation time than the process simulator itself. In addition, it is intended to improve NGL recovery control strategies with real-time estimations of control variables.

2. Process Description

The dynamic operation of the CRR unit has been modelled in the process simulator Aspen HYSYS® considering realistic operating conditions (Chebeir et al., 2019). The raw gas is fed to the plant with a nominal flowrate of 4980 kmol/h, pressure and temperature of 5818 kPa and 35 °C, respectively, and a gas composition characterized by a low content of liquids (Chebbi et al., 2010). The main unit of the separation process, responsible for the removal of the light component of the raw gas, is the demethanizer column, given by a 30 stages distillation column with a reboiler. The raw gas entering the CRR unit is first cooled and then the liquid fraction is separated utilizing a flash tank. Three resulting feed streams enter the demethanizer column in the 2nd the 8th and the 26th trays. The column reflux is provided by compressing part of the demethanizer top product with a cryogenic compressor, ensuing a reflux stream of nearly pure methane. Finally, the remaining part of the column product is recompressed for commercialization while the NGL bottom product of the demethanizer column is sent to further separation in the fractionation train.

3. Methods

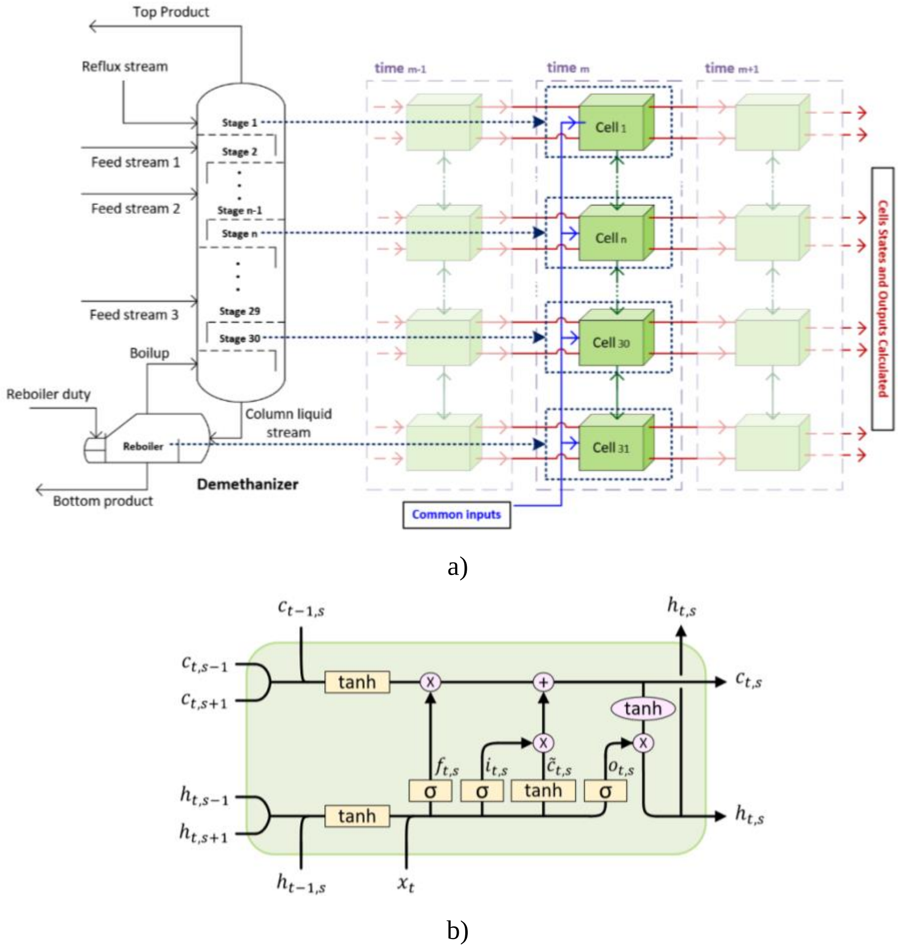
In this work, deep learning methods for the development of RNNs were used to build the surrogate data-driven model. The goals of the neural model were the achievement of accurate estimations for the prediction of the time evolution of temperature and compositions of all column trays and reboiler of the demethanizer column. The inputs considered were given by only easy-to-measure variables: pressure and temperature of all column feeds, reboiler duty, power of the cryogenic compressor and separator pressure and temperature. The selection of RNNs is based on the need to obtain an adaptive neural model capable to capture the existing temporal correlations between the involved variables. In particular, thanks to their ability to overcome the vanishing gradient problem, LSTM networks were used for digital twin realization.

3.1. Neural model architecture

In order to design a suitable model to represent the dynamics of the demethanizer column, the LSTM neural networks are arranged considering a particular architecture.

The proposed network layout involves the utilization of a dedicated LSTM cell to predict the operations of each stage of the distillation column and the reboiler. Thus, each cell was demanded at modelling the dynamic separation of the corresponding tray, allowing the dynamics inherent in the separation process occurring in each column stage to be captured in the model. Furthermore, bidirectional connections have been applied by modifying the flow of information of a standard LSTM, interconnecting adjacent cells one to another to account for the interactions between adjacent stages and emulate the action of internal flows within the distillation column. The described model layout, whose structure is schematically represented in Figure 1a, is responsible to approximate the demethanizer operation at a certain instant of time. A schematic representation of a modified LSTM cell is depicted in Figure 1b. As can be seen, in the considered cell (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.

For what concern the terminal cells of the network, mimicking the column top tray and the reboiler, the hidden states from the missing adjacent cell are replaced by the actual measurement of the corresponding variables of the reflux in a case and the liquid stream entering the reboiler in the other.



a)

b)

Figure 1: Schematic representation of: Overall architecture a) and the modified LSTM cell b).

3.2. Datasets

All the data used in the work were generated using the simulated demethanizer column described in section 2 and collected in different datasets for the training, testing and validation campaigns. The datasets employed for training and testing of the neural models, considering a split ratio of 80:20, were two months of simulated plant operation. Step changes of varying amplitude in the plant feed flow rate were applied with variation peaks of 10% of the plant feed nominal value. To validate the performance of the presented neural model, 2 days of plant operation were simulated by applying step variations of 5% in the plant feed flow rate nominal value.

All the datasets used in this work were registered by considering a sampling time of 20 seconds and without measurement delays, as the data were treated as historical plant data. Moreover, a measurement noise was applied to provide more realistic measured data. The values of the maximum allowed measurement noise applied are shown in Table 1.

Table 1: Values of maximum measurement noise considered for input and output data in train, test and validation datasets.

	Duty	Temperatures	Pressures	Concentrations
maximum measurement noise	1.3%*	0.1 [◦C]	1.5%*	2%*

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

3.3. Neural model training

The LSTM neural networks described in section 3.1, have been implemented and trained in the Python programming environment by using the functions available in the open-source library PyTorch. The loss function considered in the optimization problem is given by the Mean Square Error (MSE), while the parameter updating was performed by the Adaptive Moment Estimation optimizer (ADAM) (Kingma & Ba, 2015). The network hyperparameters were obtained by multiobjective optimization using the NSGA-II algorithm available in the pymoo package (Blank & Deb, 2020).

4. Results

The performances of the developed neural model are evaluated by considering the column transient profiles predicted by the proposed neural model in the validation campaign for temperature, and key components composition. For sake of brevity, only the results obtained under the worst-case variation represented by the application of a 5% decrease in the feed plant nominal value for ethane composition are reported and shown in Figure 2. The ability of the model to estimate the remaining target variable is assessed by considering the Mean square error (MSE) between the actual and the predicted profiles in the same campaign. Figures 2a and 2b show respectively the actual and estimated ethane transient column profiles. The comparison of these two graphs gives a qualitative indication of the estimation performance of the model for the transient column profile of ethane composition. As it is possible to visualize the proposed model is able to predict the trend of ethane time evolution profile providing an estimation with no detectable deviations from the actual profile. To provide a quantitative measure of the accuracy of the obtained estimation, the estimation error of the column transient profiles for ethane composition is depicted in Figure 2c. Here, is it shown that the proposed neural model can accurately predict the ethane composition transient profiles for all the column trays. The estimation error increase where most of the ethane variations occur, indeed, it increases in the column terminal regions, where most of the separation takes place. The deviations of the predicted transient column profile are always within the maximum error reported in Table 1.

To visualize in detail the estimation obtained, Figure 2d shows the comparison between the estimated ethane composition profile and the measured profile in the tray where the higher estimation error is registered, the 29th column tray. As can be seen from the figure, the profile obtained with the neural model can correctly predict the evolution of the composition, with considerably lower variance than the considered measurement.

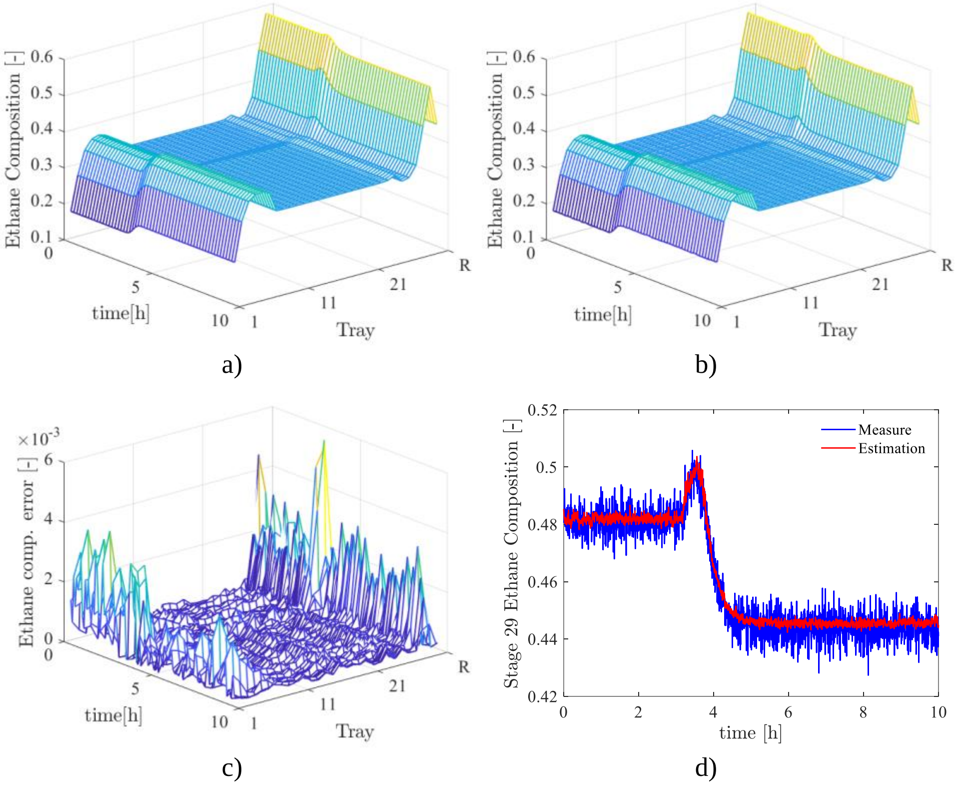


Figure 2: Ethane composition results obtained in the validation campaign under the decreasing variation of plant feed nominal value: actual column profile (a); predicted column profile (b); estimation error profile (c) and transient profiles of 29th tray (d).

Table 2: Mean square error obtained in the validation campaign under decreasing variation of 5% in the plant feed nominal value.

	Temp.	**Methane**	**Ethane**	**Propane**	**Butanes**	**Pentanes**	**Hexane**
MSE	0.0072	7.33e-7	4.31e-7	7.49e-8	1.22e-8	1.42e-9	2.27e-9

Analogue results are obtained for the temperature and the composition of the remaining components. Table 2 shows the MSE values obtained by considering the predictions of the transient composition profiles in all column trays and the actual profiles, for all the target variables in the validation campaign. As it can be noticed from the obtained MSE value, the use of the proposed model provides an accurate prediction of dynamic temperature profiles. Although it is not apparent from the MSE, it is reasonable to highlight that larger errors have been observed in the last three stages and in the reboiler, where most of the column temperature variation takes place. Concerning the prediction of the remaining components, also in these cases, the model is capable of accurately estimating the evolution over time With estimation errors resulting always lower than the assumed measurement errors, as can be inferred from the obtained MSE values, model's ability to filter the measurement noise is furthermore assessed.

5. Conclusions

In this work, a demethanizer column of a simulated CRR unit has been modelled by employing modified bidirectional LSTM neural networks exploring the possibility of its use as digital twins for the column. The results show that the developed neural model can approximate the column trend profiles of the target variables and has a good ability to filtrate measurement noise for the column composition profiles. It also proves to be able to predict the actual trend of the temperature transient profiles quite accurately for most of the column trays and to be a useful tool for temperature monitoring in the final stages of the column and the reboiler. As a result, the presented model may be a suitable candidate as a digital twin of the demethanizer column. On the other hand, the need for concentration measurements yields less attraction to be used as digital twins in a real plant.

References

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. Cognitive science, 14(2), 179-211.

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.

Grieves, M. (2015). Digital Twin: Manufacturing Excellence through Virtual Factory Replication.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.

Blank, J., Deb, K. (2020). pymoo: Multi-Objective Optimization in Python, in IEEE Access, vol. 8, pp. 89497-89509.

Kherbeck, L., Chebbi, R. (2015). Optimizing ethane recovery in turboexpander processes. J. Ind. Eng. Chem. 21, 292–297.

Kidnay, A.J., Parrish, W.R., McCartney, D.G. (2011). Fundamentals of Natural Gas Processing, 2nd Editio. ed, Fundamentals of Natural Gas Processing. CRC Press, Boca Raton, FL.

Kingma, D. P., & Ba, J. (2015). Adam: A Method for Stochastic Optimization. ICLR. 2015. arXiv preprint arXiv:1412.6980, 9.

Luyben, W.L. (2013). NGL demethanizer control. Ind. Eng. Chem. Res. 52, 11626–11638.

Mandis, M., Baratti, R., Jorge, A., Chebeir, J.A., Tronci, S., Romagnoli, J. A. (2022). Performance assessment of control strategies with application to NGL separation units. J Natural Gas Sci Eng vol. 106, 104763.

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. Microelectronics Reliability, 114, 113857.

Tao, F., Zhang, L., Nee, A. Y. C., & Pickl, S. W. (2016). Editorial for the special issue on big data and cloud technology for manufacturing. The International Journal of Advanced Manufacturing Technology, 84(1-4), 1-3.

Zhang, S., Liu, C., Jiang, H., Wei, S., Dai, L., & Hu, Y. (2015). Feedforward sequential memory networks: A new structure to learn long-term dependency. arXiv preprint arXiv:1512.08301.

Zhu, X., & Ji, Y. (2022). A digital twin–driven method for online quality control in process industry. The International Journal of Advanced Manufacturing Technology, 119(5), 3045-3064.