12th IFAC International Symposium on
Advanced Control of Chemical Processes
July 14-17, 2024. Toronto, Canada

# Predicting nonlinear dynamics of a gas-lift oil production system through hybrid decomposition-recurrent neural networks models

**Gustavo de O. e Souza, Jorge O. Trierweiler, and Marcelo Farenzena**

*Group of Intensification, Modeling, Simulation, Control, and Optimization of Processes (GIMSCOP), Chemical Engineering Department, Federal University of Rio Grande do Sul, Porto Alegre 90040060, Brazil (e-mail: {souza, jorge, farenz}@enq.ufrgs.br)*

**Abstract**: The present paper introduces an innovative approach integrating recurrent neural networks, static models, and signal decomposition into base and residual behavior components for system nonlinear dynamic modeling and identification. The proposed methodology divides a nonlinear single-input single-output gas-lift oil production system into base response and residual components, assessing the first with a first-order dynamic model with variable gain and the latter with an Encoder-Decoder (E-D) GRU model. The study evaluates the methodology under various conditions, including noiseless and noisy data and scenarios with unmeasured disturbance. The percentage of stationary residues and the normalized root-mean-squared error (NRMSE) are applied to assess the model's performance. Overall, the proposed methodology demonstrates significant effectiveness, with NRMSE lower than 5% and percentages of stationary residues ranging from 90 % to 100 % across the study scope. The results stand out when compared to the direct application of E-D GRU model without decomposition, where the percentage of stationary residues was equal to 83%, and the training mean squared error was 10 times higher than that of noiseless scenarios.

*Keywords*: nonlinear systems, recurrent neural networks, machine learning, hybrid models, chemical engineering applications, oil and gas production system

## 1. INTRODUCTION

Nonlinear systems are typically characterized as systems that do not satisfy the superposition principle (Billings, 2013). This category of systems englobes cases among numerous fields of knowledge, especially process modeling and control (Xie et al., 2011, Badillo-Hernandez et al. 2019). Regarding oil and gas production, these systems are found to be crucial processes in different steps of their chain, such as crude oil extraction and oil cracking (Ray and Villa (2000), Diehl et al. (2018), Yang et al. (2004)). Considering that the oil and gas market size alone represents approximately 7% of global GPD, with a CAGR of nearly 5% from 2022 to 2023 (The Business Research Company, (2023), World Bank (2022)), efficiently controlling these processes is crucial.

Generally, there are two main categories of approaches by which nonlinear systems can be assessed: phenomenological and empirical models. The first are based on mass and energy balances, usually described by PDE and ODE models, while the latter consists of system identification from data. When conceptually applied jointly, a hybrid approach is obtained (Henson, 1998). Fundamental models, although presenting many advantages compared to empirical models, require extensive process knowledge and may present dynamics too complex to be applied, resulting in the need for reduction techniques. System identification, on the other hand, applies to any system, regardless of its underlying causes and internal processes (Henson (1998), Billings (2013), but depends highly on available and trusted data – these risks, however, are currently being increasingly mitigated, due to the widespread adoption of Industry 4.0 technologies, increasing the application range of empirical approaches (Jagatheesaperumal et al., 2022).

Within empirical methodologies, machine learning algorithms have gained significant prominence, experiencing substantial application growth, mainly propelled by advancements in computational power (notably in the case of deep learning models) (Sharifani and Amini, 2023). It is noteworthy to emphasize that despite being already integrated into the routine of the oil industry for decades (Alkinani et al., 2019), these algorithms, as noted by Sharifani and Amini (2023), are still in the early phases of development, harboring untapped potential. In system identification, the theoretical utilization of these algorithms has been extensively discussed as an alternative to established NARX models, with a specific focus on recurrent neural networks (refer to Section 2) (Pham and Liu, 1995). The practical implementation of these discussions is becoming increasingly viable due to the aforementioned technological advances.

Hybrid methodologies have been designed by combining neural networks and fundamental models to mitigate potential naïve data-driven results and boost model performance in general. Among them, physics-informed neural networks stand out, as highlighted by Raissi et al. (2019). Essentially, these models leverage neural networks' significant function approximation capabilities (Haykin, 1999) and prior knowledge of the system incorporated through differential equations. Despite the increasing popularity of this approach, it is essential to note that it has limitations when applied to specific systems and may encounter challenges in accurately capturing their dynamics (Dwivedi et al., 2021).

A viable approach to handling a complex system involves dividing it into smaller, and consequently simpler, subsystems and addressing them individually. The stratification or decomposition of a problem is a fundamental aspect of Cartesian problem-solving analysis as proposed by the philosopher Descartes (Campos, 2010). While this rationale is frequently employed in time series analyses to independently examine trend, seasonal, and noise components (Brockwell and Davis, 2016), it also proves beneficial in diverse applications, including optimizing large-scale plants (Zhang and Zhu, 2000).

This paper proposes a hybrid methodology combining system decomposition and neural network system identification for predicting the dynamics of a nonlinear gas-lift oil production system. This system consists of applying natural gas injection to reduce the column weight of oil in deepwater oil wells, enhancing productivity by reducing production column pressure through the expansion of injected gas. As offshore oil sources represent around 30% of the industry production nowadays, and gas-lift systems are heavily nonlinear, with limit-cycle behavior in certain operation conditions, this system is an excellent option for model validation (Diehl, 2018). The decomposition aims to split the system's original data into two components: a base response component, that is, an approximation obtained by a given model, and its residual, containing the characteristics of the system not englobed by the first model. The first component is based on a first-order response for the present study, while the latter corresponds to the system's damping characteristics. This paper will focus mainly on the residual analysis, as, for the case study, it represents the system's damping characteristics.

The rest of this paper is structured as follows. Section 2 provides necessary definitions regarding the applied neural network architectures. Section 3 explains the proposed methodology, and Section 4 details the case study. Section 5 presents the results, and to sum up, Section 6 capitulates the key conclusions and lists the extended possibilities of the proposed approach.

## 2. BACKGROUND

Recurrent Neural Networks (RNN) consist of a class of neural networks that essentially consider the sequence or order of input data, rendering them particularly advantageous for analyzing sequential data, such as time series (Aggarwal, 2018).

As the input window size increases to achieve a satisfactory predictive model, the number of units required in the neural network must also increase. However, as input window size grows, so does the likelihood of issues such as exploding or vanishing gradients. These phenomena consist of error backpropagation malfunction, and both scenarios can compromise the model's accuracy. To prevent these problems, the LSTM (Long Short-Term Memory) network was developed, introducing the concept of network state (or memory), as well as adding input, output, and forget gate layers as new units (Aggarwal, 2018).

While LSTM networks occasionally mitigate vanishing or exploding gradient issues, their application may demand significant computational memory due to their intricate architecture (Salehinejad et al., 2018). Consequently, utilizing Gated Recurrent Unit (GRU) networks, a simplified version of LSTM, emerges as a viable solution to overcome this potential barrier. The primary distinctions between these architectures lie in the absence of an explicit cell state and utilizing only two internal gates within the GRU network: the reset gate ($r_t$) and the update gate ($z_t$). Equations (1) to (4) and Figure 1 delineate the components of the GRU network (Aggarwal, 2018). W and b are the weights and the bias, $x_t$ is the input data, $h_t$ and $h_{t-1}$ are the input and output hidden states, and $\odot$ corresponds to the Hadamard product. The activation functions described are sigmoid ($\sigma$) and hyperbolic tangent (tanh) functions.

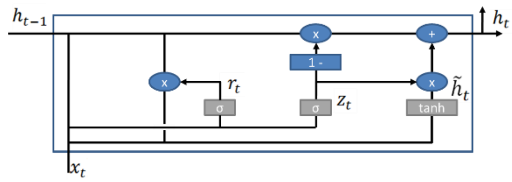


Figure 1. GRU cell unit.

$$z_t = \sigma\left(W_f[h_{t-1}, x_t] + b_z\right) \tag{1}$$

$$r_t = \sigma(W_i[h_{t-1}, x_t] + b_r) \tag{2}$$

$$\tilde{h}_t = tanh(W_h[\, r_t \odot h_{t-1}, x_t] + b_h) \tag{3}$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \tag{4}$$

In addition to the architectures mentioned above, many variations have been proposed to improve RNN performance in diverse problems. One such configuration involves including a backward copy of the input sequence through an additional layer. This arrangement is called Bidirectional RNN, or BRNN (Graves and Schmidhuber, 2005).

In problems where both network input and output are sequential data, commonly referred to as sequence-to-sequence, simple recurrent networks are typically not employed. Instead, the Encoder-Decoder neural network structure emerges as an alternative (Lyu et al., 2020). As its name suggests, this network consists of two primary components: the encoder, tasked with extracting information from the input sequence, and the decoder, which transforms the encoder's hidden state ($h_t$) into output data. A state vector interconnects these components.

## 3. METHODOLOGY

This study introduces a hybrid methodology that combines Encoder-Decoder neural networks, regression models, and time series decomposition to forecast the behavior of nonlinear systems. The proposed model aims to predict the response of a given system to a disturbance (step or set-point changes response) by utilizing only the current values of the variables ($y_t$ and $u_t$) and the future disturbance information ($u_{t+h}$ or set-point value) as input. To achieve this goal, the variable of interest undergoes decomposition into two distinct components: the first one derived from a first-order transfer function with a fixed time constant (base component), and the corresponding residual, representing the system's deviation from a first-order response (referred to as the residual component).

Two distinct models are employed in this process: an Encoder-Decoder neural network model that integrates GRU and Bidirectional GRU structures, with the residual component designated as the desired output (referred to as the dynamic model – due to the residuals representing the damping characteristics of the system); and a first-order transfer function model with a fixed time constant of $\tau$ =10 min and variable gain tasked with forecasting the future stationary state value of the controlled variable (referred to as the base response model – B). The model's gain is obtained by the difference from the output variable value at time t = 0 $y_0$ and its stationary value $y_{ss}$ correspondent to $u_{t+h}$, obtained by a polynomial regression model. The base response component transfer function is given by:

$$B(s) = \frac{y_{ss}(u_{t+h}) - y_0}{10s+1} \quad (5)$$

The proposed decomposition involves individually designing models to effectively capture variations in both the system's damping and nonlinear gain. Figure 2 provides a detailed illustration of the decomposition approach. For the depicted example, the first response's base response component transfer function (at time = 10h) was obtained with a gain constant of $K_1 = y_{ss} - y_1$, with yss ≈ $y_2$ ideally. Analogously, the same proceeding is repeated to predict the second response, with $K_2 = y_{ss} - y_2$, with yss ≈ $y_3$ and so on.

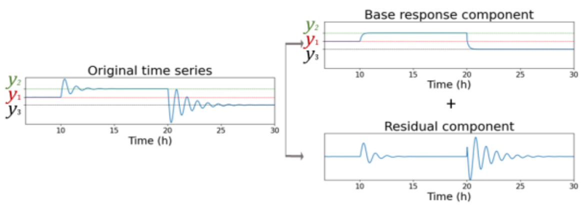


Figure 2. Example of the proposed decomposition approach.

The dynamic model consists of an Encoder-Decoder network with a two-layered encoder (a GRU layer followed by bidirectional GRU layer) and a single GRU layer decoder, trained by MSE minimization with Adam (Kingma and Ba, 2015) and Glorot weight initialization (Glorot and Bengio, 2010).

The performance of the dynamic model is evaluated using two key metrics: the normalized root mean squared error (NRMSE), which measures the disparity between the validation data and its corresponding prediction, and the percentage of stationary residues determined by the Augmented Dickey-Fuller (ADF) test. If the test's null hypothesis is rejected (i.e., p-value < 0.05), indicating statistical significance, the analyzed residual is considered stationary (resembling white noise). This outcome suggests that the model has effectively captured the deterministic aspect of the system (Brockwell and Davis, 2016).

## 4. CASE STUDY

As mentioned, this paper study focuses on a gas-lift oil production system. Gas-lift is an oil production technique involving natural gas injection to alleviate the column weight in a deepwater oil well production system. Natural gas is injected into the production column through annular pipes, sourced from surface facilities via a gas lift supply line to the wellhead at Christmas Tree. The injection causes the gas to expand, reducing the specific mass of the extracted mixture. These effects decrease production column pressure, enhancing well productivity (Lea et al., 2008; Diehl et al., 2018).

However, as previously mentioned, this system can exhibit limit cycles under specific operating conditions, leading to persistent oscillations in system pressures and flows. This situation poses potential risks to equipment integrity and may diminish well productivity. Therefore, a precise understanding and prediction of the system are crucial for risk mitigation. Stable operation and limit cycles are determined by a Hopf bifurcation, with limit cycles potentially occurring at low gas lift flow rates and/or high topside choke valve openings (Diehl et al., 2018).

This paper considers a Single-Input-Single-Output (SISO) approach to this system, ranging from stable to near Hopf bifurcation states. The study focuses on a constant gas lift flow rate of 165,000 $Sm^3$/day, with the choke valve opening as the input variable and PDG pressure as the output variable. The proposed methodology undergoes evaluation in distinct scenarios. For the first one, PDG pressure response to open-loop choke valve variation is predicted, as for the second one, both open-loop and closed-loop variations are assessed. Closed-loop data was obtained with a Proportional-Integral (PI) controller. In addition, a new Boolean input variable is incorporated into the dynamic model (0 for open-loop, 1 for closed-loop), and the $u_{t+h}$ input variable for the closed-loop state is derived from the set-point values using the static model. The evaluation metrics for the model include the percentage of stationary residues generated.

Furthermore, the open-loop analysis encompasses two additional scenarios. The first scenario involves evaluating the proposed methodology in the presence of noise and unmeasured disturbance on the gas lift flow rate within the system. The second scenario assesses the model's performance using validation data characterized by shorter and more frequent variations in the input variable, hinting at potential applications for predictive control. In this case, the output variable considered is the output flow rate, and the evaluation metric is the Normalized Root Mean Squared Error (NRMSE). Table 1 provides a summary of the paper's scope.

**Table 1. System analysis scope**

| ID | Scenario | Output variable | Metrics |
|---|---|---|---|
| 1 | Open Loop | PDG pressure | Percentage of stationary residues (%) |
| 2 | Open and closed loop | | |
| 3 | Open loop (noise and disturbance) | | |
| 4 | Open loop (multistep model) | Output flow rate | NRMSE |

The data used for this paper was simulated using FOWM model developed by Diehl *et al.* (2017).

## 5. RESULTS

This section is divided into 6 parts. Firstly, we present the results of directly applying the dynamic model to the studied system without the decomposition process. Section 5.2 provides a general view of the results obtained from applying the methodology for scenarios 1 to 4 (see Table 1). Sections 5.3, 5.4, 5.5, and 5.6 cover the results for scenarios 1, 2, 3, and 4, respectively.

### *5.1 Non-decomposed system results*

In order to fully understand the benefits of applying system decomposition through base response modeling, we applied the exact same E-D GRU network to the original output data. The results, described in Figure 3, show that the model fails to capture the system's dynamics for most responses and the nonlinear gain variations. The training mean squared error obtained was 6,8 x $10^{-4}$, and only 83% of the prediction residues presented stationarity according to ADF test. The following sections show that this problem is not found in undisturbed scenarios when output data is segregated into two distinct components.

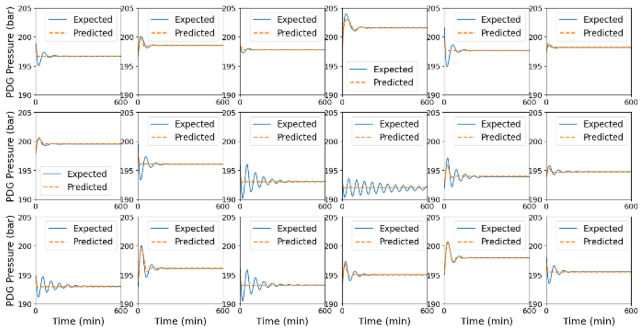


Figure 3. Prediction results for non-decomposed system.

### *5.2 General view of results*

Table 2 summarizes the general hyperparameter settings for dynamic model training, the training mean squared error, and the percentage of stationary residue identified in the validation response prediction for each scenario.

**Table 2. General results for dynamic model training and validation for gas-lift oil production system.**

| Scenario / Parameter | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Units on Encoder Layer 1 (GRU) | 20 | 24 | 20 | 20 |
| Units on the Encoder Layer 2 (BGRU) | 40 | 48 | 40 | 40 |
| Units on the Decoder layer (GRU) | 20 | 24 | 20 | 20 |
| Number of trainable parameters | 10,221 | 14,641 | 10,221 | 10,221 |
| Training mean squared error x $10^{-5}$ | 4.8 | 6.1 | 4.8 | 170 |
| NRMSE (%) | - | - | - | 4.7% |
| Percentage of stationary residues (%) | 92% | 90% | 100% | - |

All open-loop models were configured with an identical number of neurons across all layers (20, 40, and 20 for the encoder first layer, encoder second layer, and decoder layer, respectively) and were trained for an equal number of epochs (600), except for the open and closed-loop models. Due to the anticipated higher complexity, the number of hyperparameters for each layer in these models was increased by 5%.

Every noiseless data scenario exhibited a notably low training error, with the scenario model for the output flow rate demonstrating the lowest value (approximately 20% lower than the PDG pressure models). Upon analyzing validation data predictions, the open-loop results for PDG pressure showed a percentage of 92% for the open-loop-only scenario and 90% for the multi-loop scenario model. Despite the varying levels of complexity in the PDG pressure cases, the results were deemed satisfactory for both scenarios.

Additionally, the results of residue stationarity for noisy and disturbed data were excellent – all residues obtained were stationary. In terms of training error, the error was approximately 100 times higher than that of other models. However, a high error value is not necessarily undesirable in this context. It suggests that the model effectively rejected noise and unmeasured disturbance as deterministic components of the system. The rest of this section is divided into four parts, each describing a scenario.

### *5.3 Open loop PDG pressure prediction*

The training set for open-loop pressure models (as well as for the output flow rate) comprised a 42-day time series, equivalent to 30 continuous 33-hour-long step changes, as illustrated in Figure 4. The validation dataset (depicted in Figure 5) encompasses an approximately 7-day-long sequence of input variable alterations with randomized magnitudes and durations.

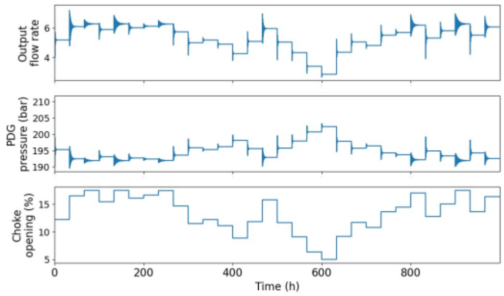


Figure 4. Training data set for open loop scenarios.

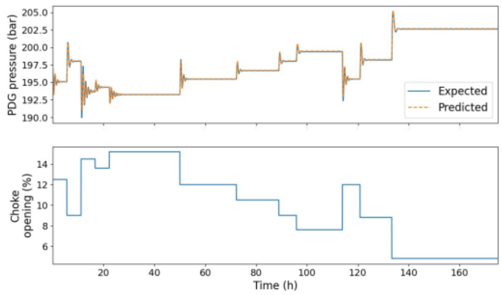


Figure 5. Validation data and prediction for open loop PDG pressure.

It can be concluded that the model excels in predicting the PDG response to choke valve step variations. Figure 5 indicates that the model's performance remains satisfactory even for the highest step in the input variable (around time = 10 hours).

### *5.4 Open and closed loop PDG pressure prediction*

For the open and closed-loop models, the training dataset was developed using half of the data employed for the training of the open-loop model (see Figure 4) and closed-loop data, simulated by an analog similar approach (Figure 6), resulting in approximately 42 days of data for each case. The validation dataset (depicted in Figure 7) spans 21 days and was obtained analogously to that described in Section 5.3.

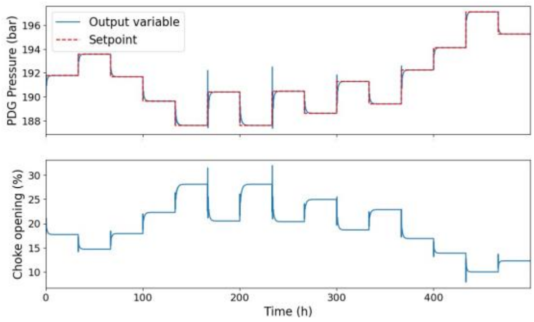


Figure 6. Closed loop training data.

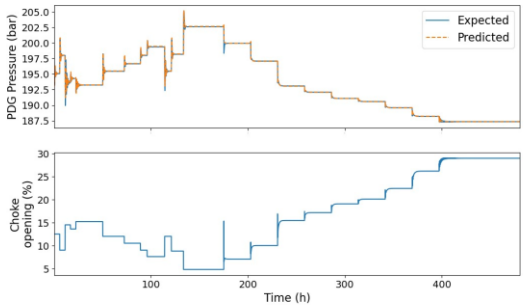


Figure 7. Multi-scenario validation data and prediction.

It is evident that the model, for most responses, effectively captured variations in the degree of damping with consistent precision. The exception is observed during the transition of the loop state at approximately time 170 hours, which is anticipated given that the training set did not encompass this particular switch.

### 5.5 *Noise and unmeasured disturbance effect for PDG pressure prediction*

The same method for defining training and validation datasets was used in sections 5.1 and 5.2, with a training set (Figure 8) that was 42 hours long and a validation set (Figure 9) that consisted of 16 days of random responses.

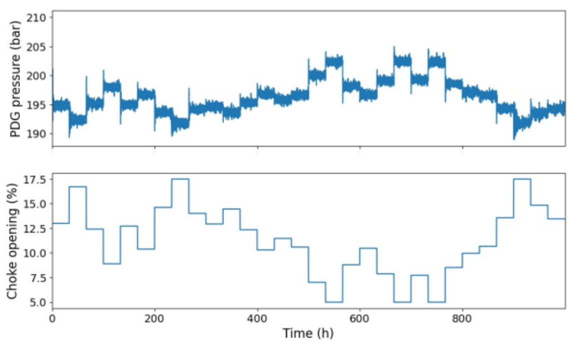


Figure 8. Training data for open loop scenario with noise and unmeasured disturbance.

As shown in Figure 9, the dynamic model unequivocally dismissed noise as a deterministic component of the analyzed system. Furthermore, predictions often exhibited deviations in instability values. It is crucial to emphasize that the stability value is directly linked to the implemented static model (in this case, polynomial regression). The model proved itself incapable of adjusting for the impact of unmeasured disturbances on pressure; these results, however, are expected, since static model tuning and robustness, although crucial to overall model performance, do not constitute the primary focus of this paper – the decomposition strategy and dynamic model structure.

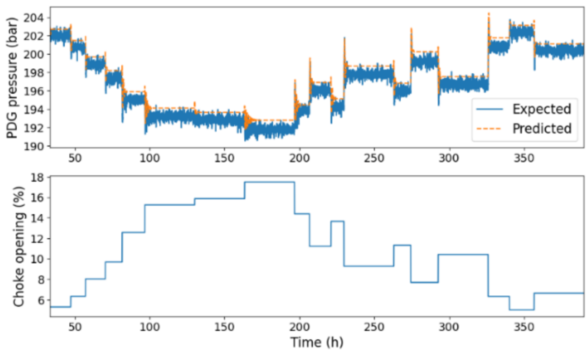


Figure 9. Validation and prediction for noisy data and unmeasured disturbance scenarios.

### 5.6 *Open loop output flow rate prediction*

The dataset used for this scenario is depicted in Figure 3. Similarly, the validation dataset (Figure 9) comprises an approximately 8-day-long sequence with shorter and more frequent step alterations in the choke valve opening.

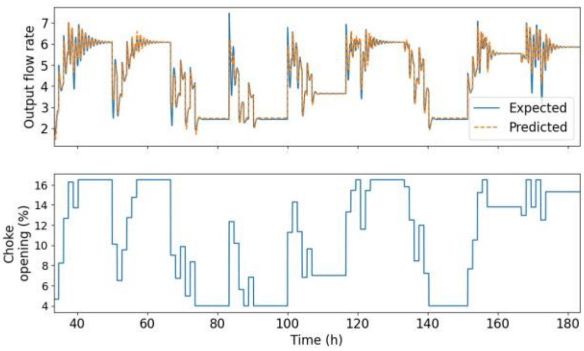


Figure 10. Multistep validation data and predictions for output flow rate (open loop).

An analysis of the prediction results depicted in Figure 10 shows that the proposed methodology predicts the output effect of sequential steps in the input variable. This conclusion suggests that the developed method is well-suited for practical application in predictive control strategies. Furthermore, using the output flow rate as an output variable instead of the PDG pressure has not compromised model performance.

## 6. CONCLUSIONS

In this paper, we have examined the results obtained from applying a proposed hybrid methodology for predicting the behavior of a nonlinear gas-lift oil production system across various scenarios. The method is deemed satisfactory, demonstrating promising results for all evaluated scenarios.

Our study makes a significant contribution to the field by introducing an innovative approach to address the challenges of nonlinear systems modeling and control. The methodology presented in this study provides fresh perspectives and opportunities that extend beyond individual processes or sectors, holding the potential for application in various scenarios. Thus, this contribution extends to the broader domains of process modeling and control, providing valuable insights and tools for practitioners and researchers.

It is worth noting that the proposed methodology may have its limitations. Despite demonstrating robust performance for the system analyzed in this study, the number of hyperparameters set for each case was not uniform, and the precise tuning of these hyperparameters may influence future applications. Moreover, the overall performance relies on the static model's high quality, making the methodology's application feasible only when fundamental knowledge or a comprehensive static data history is available. As this is a novel approach, there is a possibility that the methodology is susceptible to other factors or impacts not explicitly addressed in this study.

The applicability of the methodology in even more challenging and complex scenarios can be explored. Additionally, there is potential for real-world application in nonlinear model predictive control strategies for the cases studied in this paper.

## REFERENCES

Aggarwal, C. C. (2018). *Neural Networks and Deep Learning: A Textbook (1st ed.).* Heidelberg: Springer-Verlag, Berlin.

Alkinani, H. H. et al. (2019). Applications of Artificial Neural Networks in the Petroleum Industry: A Review. In SPE Middle East Oil and Gas Show and Conference.

Badillo-Hernandez, U., Alvarez, J., & Alvarez-Icaza, L. (2019). Efficient modeling of the nonlinear dynamics of tubular heterogeneous reactors. *Computers and Chemical Engineering*, 123, 389-406.

Billings, S. A. (2013). *Nonlinear System Identification* (1st ed.). Wiley, Chichester.

Brockwell, P. J., & Davis, R. A. (2016). *Introduction to Time Series and Forecasting* (3rd ed.). Springer Texts in Statistics, Switzerland.

Campos, V. F. (2010). *True Power* (1st ed.). Falconi, Nova Lima.

Diehl, F. C., Anzai, T. K., Almeida, C. S., Von Meien, O. F., Neto, S. S., Rosa, V. R., Campos, M. C. M. M., Reolon, F., Gerevini, G., Ranzan, C., Farenzena, M., & Trierweiler, J. O. (2017). Fast Offshore Wells Model (FOWM): A practical dynamic model for multiphase oil production systems in deepwater and ultra-deepwater scenarios. *Computers and Chemical Engineering*, 99(6), 304-313.

Diehl, F. C., Almeida, C. S., Anzai, T. K., Gerevini, G., Neto, S. S., Von Meien, O. F., Campos, M. C. M. M., Farenzena, M., & Trierweiler, J. O. (2018). Oil production increase in unstable gas lift systems through nonlinear model predictive control. *Journal of Process Control*, 69, 58-69.

Dwivedi, V., Parashar, N., & Srinivasan, B. (2021). Distributed learning machines for solving forward and inverse problems in partial differential equations. *Neurocomputing*, 420(8), 299-316.

Glorot, X., & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the International Conference on Artificial Intelligence and Statistics (AISTATS'10)*. Society for Artificial Intelligence and Statistics, 2010.

Graves, A., & Schmidhuber, J. (2005). Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Networks*, 18(5-6), 602-610.

Haykin, S. (1999). *Neural Networks: A Comprehensive Foundation* (2nd ed.). Prentice-Hall, Englewood Cliffs.

Henson, M. A. (1998). Nonlinear model predictive control: current status and future directions. *Computers and Chemical Engineering*, 23, 187-202.

Jagatheesaperumal, S. K., Rahouti, M., Ahmad, K., Al-Fuqaha, A., & Guizani, M. (2022). The Duo of Artificial Intelligence and Big Data for Industry 4.0: Applications, Techniques, Challenges, and Future Research Directions. *IEEE Internet of Things Journal*, 9(15).

Kingma, D., & Ba, J. (2014). Adam: A Method for Stochastic Optimization. In *Proceedings of the 3rd International Conference for Learning Representations*, San Diego, 2015.

Lea, J. F., Nickens, H. V., Wells, M. R., & Wells, M. R. (2008). *Gas Well Deliquification (Gulf Drilling Guide)* (2nd ed.). Gulf Professional Publishing, Amsterdam. ISBN: 978-0-7506-8280-0.

Lyu, P., Chen, N., Mao, S., & Li, M. (2020). LSTM based encoder-decoder for short-term predictions of gas concentration using multi-sensor fusion. *Process Safety and Environmental Protection*, 137, 93-105.

Pham, D. T., & Liu, X. (1995). *Neural Networks for Identification, Prediction and Control* (1st ed.). Springer-Verlag, London.

Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-Informed Neural Networks: A Deep Learning Framework for Solving Forward and Inverse Problems Involving Nonlinear Partial Differential Equations. *Journal of Computational Physics*, 378, 686-707.

Ray, W. H., & Villa, C. M. (2000). Nonlinear dynamics found in polymerization processes - a review. *Chemical Engineering Science*, 55(2), 275-290.

Salehinejad, H., Sankar, S, Barfett, J., Colak, E., Valaee, S. *Recent advances in recurrent neural networks*. arXiv e-prints. DOI: 10.48550/arXiv.1801.01078.

Sharifani, K., & Amini, M. (2023). Machine Learning and Deep Learning: A Review of Methods and Applications. *World Information Technology and Engineering Journal*, 10(7), 3897-3904.

Souza, G. O. Trierweiler, J. O. & Farenzena. M. (2024). Combination of recurrent neural networks with static models: a practical solution for nonlinear dynamic systems prediction. Submitted to Computers and Chemical Engineering.

The Business Research Company. (2023). Oil And Gas Global Market Report 2023. Available at: https://www.thebusinessresearchcompany.com/report/oil-and-gas-global-market-report.

The World Bank, (2023). GPD (current US$). In The World Bank. Available at: https://data.worldbank.org/indicator/NY.GDP.MKTP.CD.

Xie, W., Bonis, I., & Theodoropoulos, C. (2011). Off-line model reduction for on-line linear MPC of nonlinear large-scale distributed systems. *Computers and Chemical Engineering*, 35(5), 750-757.

Yang, Q., Li, S., & Tian, X. (2004). Nonlinear Control of a Fluid Catalytic Cracking Unit. *IFAC Proceedings Volumes*, 37(1), 119-124.

Zhang, N. & Zhu X.X. (2000). A novel modelling and decomposition strategy for overall refinery optimisation. *Computers and Chemical Engineering*, 24, 1543-1