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

Enhanced Hybrid Model for Gas-Lifted Oil Production

Ruan R. Faria * **Fernando Arrais R. D. Lima** * **Argimiro R. Secchi** ** **Maurício B. de Souza Jr.** *

* *Escola de Química, EPQB, Universidade Federal do Rio de Janeiro, P.O. Box 68542, Rio de Janeiro, RJ 21941-909, Brazil (e-mail: rrfaria@eq.ufrj.br, farrais@eq.ufrj.br, mbsj@eq.ufrj.br).*
** *Programa de Engenharia Química, PEQ/COPPE, Universidade Federal do Rio de Janeiro, P.O. Box 68502, Rio de Janeiro, RJ 21941-972, Brazil (e-mail: arge@peq.coppe.ufrj.br)*

Abstract: Gas-lift is a strategy to enhance oil production from oil wells by reducing the hydrostatic pressure of the fluid. Efficient modeling of this process is a key point for the oil and gas industry to maximize oil output while minimizing gas consumption. This study introduces a novel hybrid model for the gas-lift process, using the Universal Differential Equation (UDE) approach. Neural networks replace algebraic equations, trained based on physical laws. The UDE method eliminates dependence on unmeasured variables, enhancing accuracy. Tested with simulated data, the hybrid model outperforms traditional models, demonstrating effective prediction of state variables and efficient handling of algebraic variables. This approach holds promise for gas-lift process modeling, control, and optimization.

Keywords: Universal Differential Equation, Neural Networks, Oil and Gas, Machine Learning.

1. INTRODUCTION

During oil production, the energy from the reservoir can not be enough to transfer the fluid to the surface. One strategy commonly used to overcome this issue is the gas lift, which is the injection of compressed natural gas into the production tubing of an oil well (Namdar, 2019). This action allows an increase in the pressure gradient between the reservoir and well fluid that pushes the fluid to the surface (Misener et al., 2009).

Many studies proposed optimization strategies for the gas-lift process to maximize oil production and minimize gas consumption. Some examples of this application are the studies developed by Peixoto et al. (2015), Krishnamoorthy et al. (2016), Carvalho et al. (2016), Krishnamoorthy et al. (2018), and Carpio et al. (2021). There are also applications of the nonlinear model predictive control (NMPC) approach for the gas-lift process (Soares et al., 2022; Miyoshi et al., 2018).

An efficient model is essential to developing the optimization and NMPC proposals. However, the gas-lift models usually rely on unmeasured parameters, needing a tool to predict these values (Krishnamoorthy et al., 2016). State estimator is one of these tools, which were used by Krishnamoorthy et al. (2016) and Krishnamoorthy et al. (2018) to estimate the gas-oil ratio (GOR). Moreover, Delou et al. (2023) used an Extended Kalman Filter (EKF) to estimate the reservoir valve flow coefficients and the top valve flow coefficient in a gas-lift process.

Artificial neural networks are another tool used in the gas-lift process to predict uncertain parameters. Teixeira et al. (2014) developed a soft sensor for measuring the downhole pressure in a gas-lift oil well. Shokir et al. (2017) trained neural networks to predict the bottom pressure and fluid flowrate from synthetic data in gas-lift oil wells, and used this data-driven model to obtain the optimum gas injection and oil production rate.

Khan et al. (2020) developed neural network models that were able to predict oil flowrates in a gas-lift well. Soares et al. (2022) developed a neural network to predict the mass of gas in the annulus and in the tubing, and the mass of oil in the tubing. They compared the neural network's performance to an EKF, considering a gas-lift process controlled by an NMPC, demonstrating an efficient performance of the machine learning strategy. Furthermore, Dias et al. (2019) efficiently used Echo State Networks for modeling gas-lift oil wells, being able to predict the process behavior even for large prediction horizons.

Hybrid models consist of models that combine machine-learning approaches with phenomenological laws. Franklin et al. (2022) used this strategy to develop a virtual sensor for application in an oil well system. They combined the phenomenological model with recurrent neural networks, and could efficiently predict the flow rates and pressures of the system. In the Universal Differential Equation (UDE) hybrid modeling strategy, one or more terms in a differential equation are replaced by universal approximators, which can be neural networks. During this neural network training, the predictions of the differential equations from the phenomenological model are considered in the loss function (Rackauckas et al., 2020). Even with the successful application of this approach for modeling some chemical processes (Nogueira et al., 2022; Bangi et al., 2022; Lima et al., 2023), this approach has no application in the oil and gas field.

This work aims to use the UDE approach for modeling gas-lift oil wells. We used the model from Krishnamoorthy et al. (2018) for gas-lift oil wells and replaced the equations to calculate the reservoir oil and gas flowrates by neural networks. The proposed model requires less parameter tuning for accurate predictions, which can be further improved by incorporating experimental production data obtained from the plant.

2. METHODOLOGY

2.1 Process Description

Artificial gas-lift injection increases oil well productivity by reducing bottom hole hydrostatic pressure through gas-lift injection into the well annulus. This process, illustrated for two wells in Fig. 1, facilitates the upward flow of oil to topside facilities.

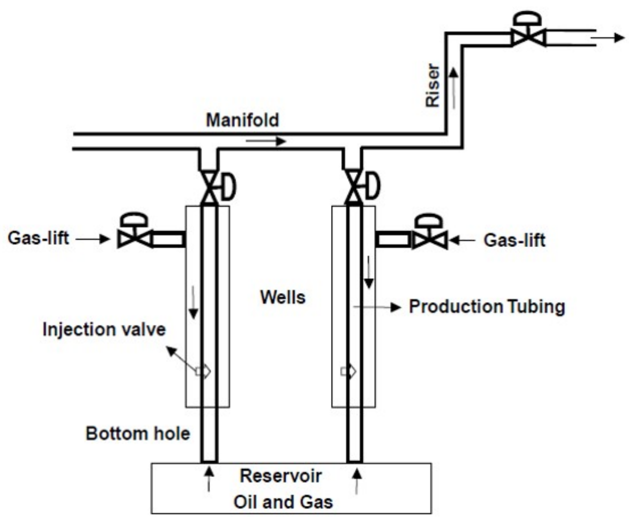

Fig. 1. A gas-lifted well network with two wells.

Our study utilizes the dynamic model from Krishnamoorthy et al. (2018) to simulate well-production tests and generate data for hybrid modeling. For the sake of brevity, we present only the differential equation model and algebraic equations concerning gas and oil flow rates from the reservoir to the production tubing, as their dynamics remain unknown for hybrid model training. We refer to Krishnamoorthy et al. (2018) for additional algebraic equations and parameters.

The specific differential equation model for N_w wells is detailed in (1).

$$\dot{m}_{ga_i} = w_{gl_i} - w_{iv_i} \tag{1a}$$

$$\dot{m}_{gt_i} = w_{iv_i} - w_{pg_i} + w_{rg_i} \tag{1b}$$

$$\dot{m}_{ot_i} = w_{ro_i} - w_{po_i} \tag{1c}$$

$$\dot{m}_{gr} = \sum_i w_{pg_i} - w_{tg} \tag{1d}$$

$$\dot{m}_{or} = \sum_i w_{po_i} - w_{to}, \quad i \in N_w \tag{1e}$$

where $\dot{m}_{ga_i}$ represents the gas-lift holdup rate in the annulus, $\dot{m}_{gt_i}$ and $\dot{m}_{ot_i}$ denote the gas and oil mass holdup rate in the production tubing, and $\dot{m}_{gr}$ and $\dot{m}_{or}$ denotes the gas and oil mass holdup rate in the riser/manifold. The gas-lift injection flow rate is denoted by w_{gl}, and w_{iv} represents the gas-lift flow rate through the injection valve into the production tubing. w_{pg} and w_{po} stand for the gas and oil flow rates into the production tubing. w_{rg} and w_{ro} represent the gas and oil flow rates from the reservoir to the production tubing. Finally, w_{to} and w_{tg} represent the oil and gas flow rate to the separator.

The algebraic equations for N_w wells are given in (2).

$$w_{ro_i} = PI_i(p_r - p_{bh_i}) \tag{2a}$$

$$w_{rg_i} = GOR_i\, w_{ro_i}, \quad i \in N_w \tag{2b}$$

where uncertain parameters include the reservoir productivity index (PI_i) and gas-oil ratio (GOR_i). Reservoir pressure (p_r) is a fixed parameter for each well, while bottom hole pressure (p_{bh_i}) is an algebraic state.

2.2 UDAE model

This study addresses the modeling and estimation of algebraic states in the artificial gas-lift injection process. We employ a hybrid modeling approach, illustrated in Fig. 2, where neural networks (NNs) approximate reservoir gas and oil flow rates, replacing two algebraic equations, i.e., from (2), in the system of differential-algebraic equations (DAEs). This substitution aims to enhance the modeling of algebraic states dependent on uncertain parameters like GOR and PI.

We extend the concept of UDE to DAEs (we refer to universal differential-algebraic equation - UDAE), optimizing neural networks based on the availability of training data. Production data simulation involves the model from Section 2.1, introducing uncertainty in GOR and PI values and measurement noise to mimic real-world conditions. Plant measurements, represented by y_1^m and y_2^m (where the subscripts 1 and 2 refer to the measured differential states, i.e., $\dot{m}_{gt}$ and $\dot{m}_{ot}$), guide the optimization process aiming to minimize mean squared error (MSE), as shown in (3), while aligning with the DAE model.

$$MSE(\theta) = \frac{1}{N}\sum_{k=1}^{N} ||y_1^p(w_{rg}(\theta,k), w_{ro}(\theta,k)) - y_1^m(k)||^2 + ||y_2^p(w_{rg}(\theta,k), w_{ro}(\theta,k)) - y_2^m(k)||^2 \tag{3}$$

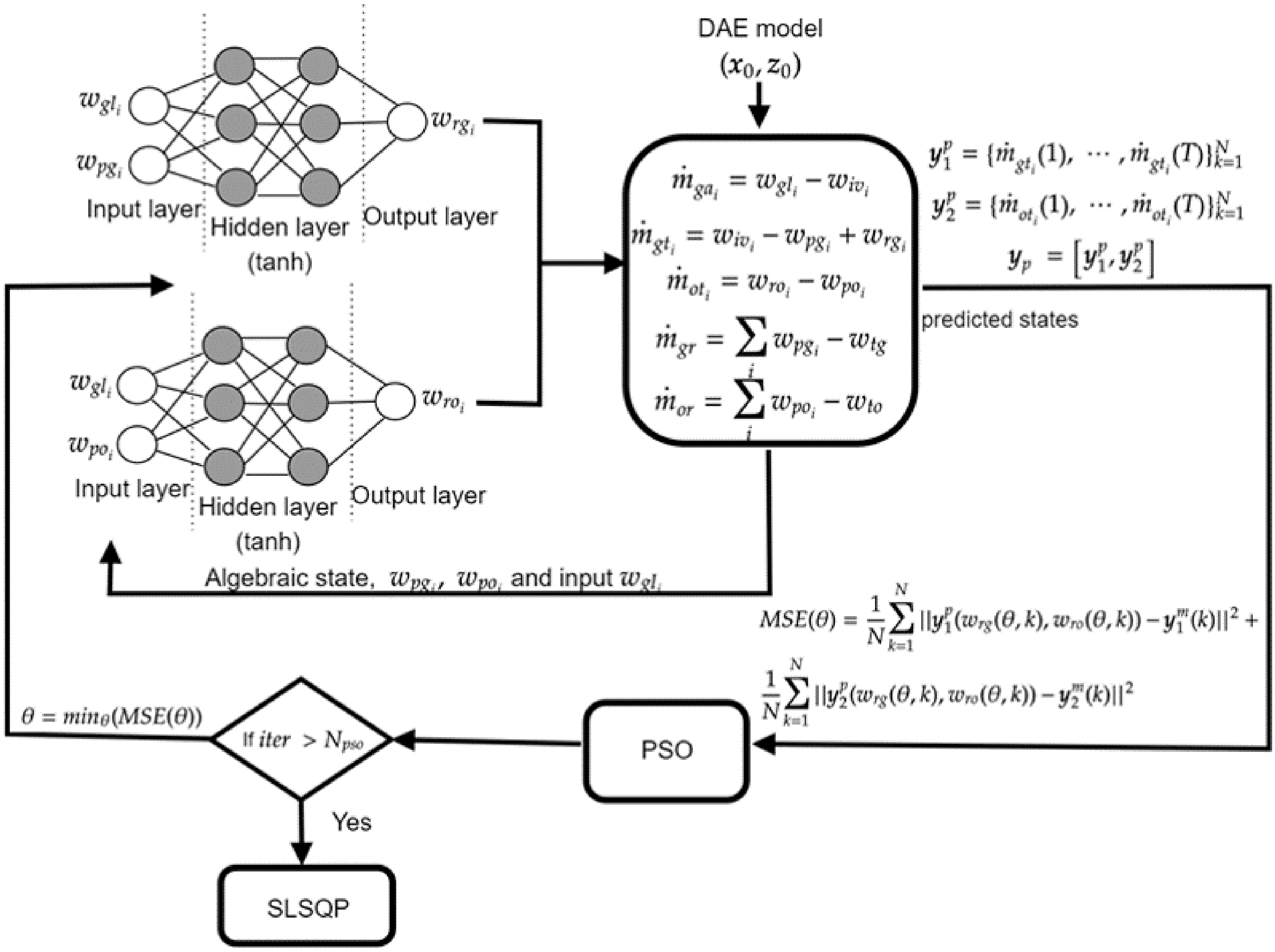

Fig. 2. UDE model structure.

Equation (3) refers to the objective function, with y_1^p and y_2^p denoting trajectories obtained from integrating the system of DAEs for each steady-state condition (i.e., $k = 1, \cdots, N$) with a dimension of $N_w \times T \times N$. At each time step (i.e., $i = 1, \cdots, T$), NNs-approximated algebraic states are input to the DAE model, with some resulting algebraic states feeding back into the NNs.

We employ two algorithms sequentially: particle swarm optimization (PSO) for a solution close to the global optimum and sequential least squares programming algorithm (SLSQP) for refinement. This dual-method approach aims to accelerate convergence, mitigating the risk of local minimum. It is important to note that the optimization process, illustrated in Fig. 2, proceeds with PSO exploring the solution space for N_{pso} iterations. Then, PSO is replaced by SLSQP, which is terminated when no further improvements in the optimization solution are obtained, assuming the final solution is close to a global optimum.

3. SIMULATION RESULTS

The system of DAEs outlined in Section 2.1 was integrated to obtain experimental data. Seven steady states with different GOR values and measurement noise were simulated using the open-source Casadi software (Andersson et al., 2019) and the IDAs integrator with an absolute tolerance of 10^{-8}, assuming an integration time of $T = 7200\,s$ and interval of $\Delta t = 60\,s$, with $t = [0\,s, 60\,s, \ldots, 7200\,s]$, is sufficient for reaching steady-state. The inputs (w_{gl_i} (kg/s), w_{pg_i} (kg/s), w_{po_i} (kg/s)) and outputs (m_{gt_i} (ton) and m_{ot_i} (ton)) of these batches are provided in Table 1, along with their statistical properties.

Table 1. Statistical analysis of the input and output data of the seven batches.

Properties	w_{gl_i}	w_{pg_i}	w_{po_i}	m_{gt_i}	m_{ot_i}
Minimum	0.2	3.67	31.55	0.77	2.97
Maximum	10.0	13.44	37.45	1.21	6.86
Mean	5.20	8.89	34.32	1.04	4.25
Standard deviation	2.47	2.44	1.01	0.11	0.87

The UDAE model consists of two neural networks with 1 hidden layer, 10 hidden neurons, and the *tanh* activation function. The activation function, number of neurons, and hidden layers were chosen after a sensitivity analysis to find the proper hyperparameters. The neural networks replace the equations related to w_{rg} and w_{ro}, as depicted in Fig. 2. The Casadi software with the IDAs integrator integrates the system of DAEs with the mentioned absolute tolerance.

To train the UDAE model, we use the open-source PySwarms software to execute the PSO algorithm (Miranda, 2018). One hundred particles are generated for each 100 iterations, using the standard global-best PSO optimizer with an absolute tolerance of 10^{-8} and built-in hyperparameter specifications. The PSO optimization solution is then refined using the successive quadratic programming (SLSQP) algorithm from the SciPy library (Virtanen et al., 2020) with an absolute tolerance of 10^{-8}.

3.1 The UDAE Model Training

The training process of the UDAE model is outlined in Fig 3. The optimization process involved 24,304 iterations to reach this solution, with 10,000 iterations performed by the PSO algorithm and the remaining by the SLSQP algorithm.

For better visualization of the optimization process, the evolution of the objective function is shown for PSO during 5,000-10,000 iterations and for SLSQP during 20,000-24,304 iterations. Notably, the PSO algorithm played a crucial role in the initial exploration of the solution space for the optimization problem—a common feature in meta-heuristic-based algorithms—thus avoiding the possibility of getting stuck in local minima. The SLSQP optimization algorithm subsequently refines the initial solution obtained through PSO. The latter benefits from a good initial solution, improving convergence and achieving a final loss function value of 3.25×10^{-4}. Since no further improvements were obtained in the optimization solution, this value is considered close to a global optimum.

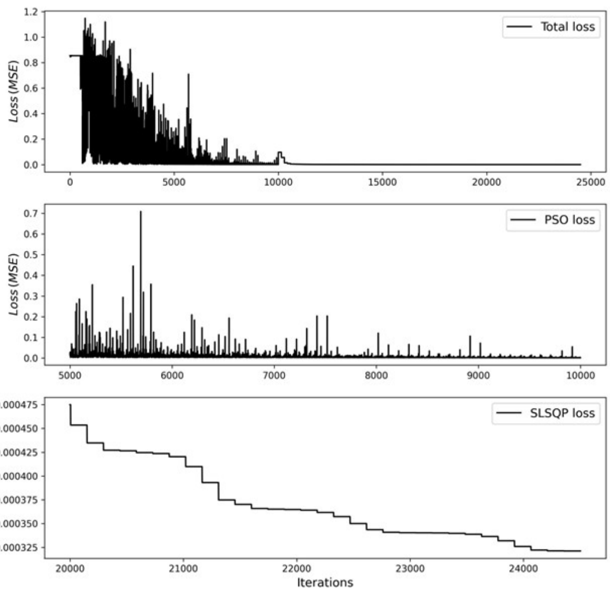

Fig. 3. UDAE model training loss.

3.2 Modeling and Prediction Study

The modeling and prediction study aims to evaluate the UDAE model's performance compared to plant data under simulated conditions that may occur in an actual process.

In the first simulated condition, the plant model is assumed to be identical to the one used to obtain the training data. Fig. 4 illustrates the estimated algebraic states of each neural network (i.e., w_{rg} and w_{ro}), showing its ability to capture the complexity and dynamic nature of the process. The good fit is evident in the predicted differential states shown in Fig 5. This shows the viability of the proposed methodology for estimating unknown dynamics, such as the reservoir gas and oil flow rates for the production tubing, while respecting the system of DAEs. Therefore, the methodology shows potential for estimation and modeling purposes, being an alternative to purely data-driven methods. Additionally, it can be extended to leverage experimental production data from the plant, available through the multiphase flowmeter, to continuously adjust the neural networks and enhance the modeling of the system of DAEs.

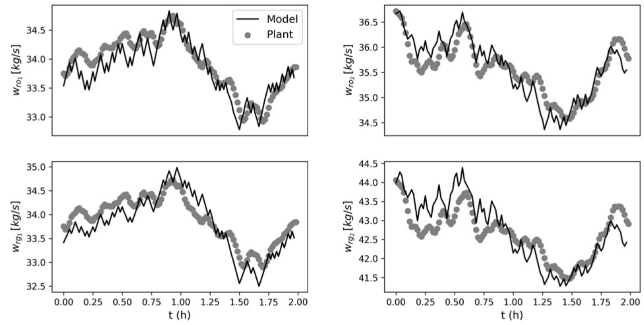

Fig. 4. The estimated reservoir gas and oil flow rates using the UDAE model.

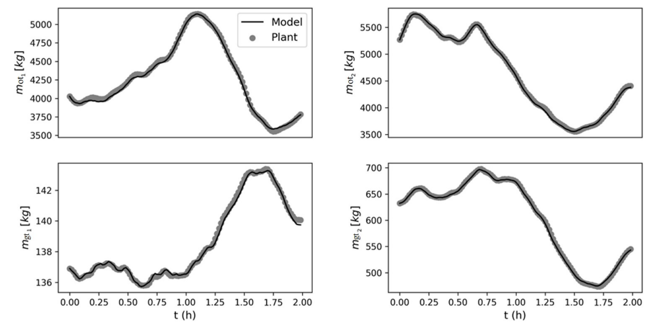

Fig. 5. The estimated gas and oil production mass flowrates using the UDAE model.

The following simulated condition assumes a Gaussian probability distribution for sampling GOR, a parameter with high uncertainty that could influence the modeling of artificial gas-lift injection, as shown in (4). In this study, $E_0\left(GOR_i\right)$ is defined as $GOR = (0.10, 0.12)$, with an uncertainty $\sigma = (0.005, 0.005)$.

$$GOR_i = E_0\left(GOR_i\right) + \sigma_i, \quad \forall i \in N_w \tag{4}$$

While some works use the extended Kalman filter for dynamic estimation of such a parameter, our methodology treats it as unknown. It models it directly using the neural network approximating the reservoir gas and oil flow rates. Again, the algorithm successfully estimates the dynamic nature of unknown algebraic states, as shown in Fig. 6 and confirmed by the predicted differential states (Fig. 7). This result highlights the generalization ability of neural networks to new process conditions, enhanced by integration with a system of DAEs, ensuring robust solutions even in extrapolated situations. This represents an advantage over purely data-driven methods, which require much training data to achieve a model with good generalization capability.

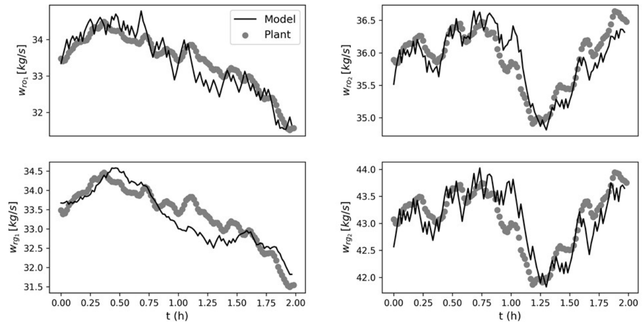

Fig. 6. The estimated reservoir gas and oil flow rates using the UDAE model, considering uncertainty in the value of GOR for a new process condition.

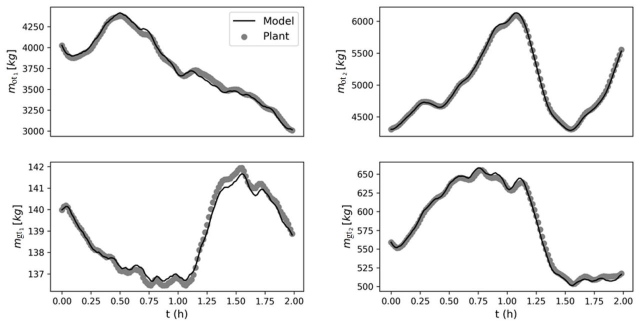

Fig. 7. The estimated gas and oil production mass flowrates using the UDAE model, considering uncertainty in the value of GOR for a new process condition.

Finally, we simulated a condition in which, along with the uncertainty in the GOR value, measurement noise was introduced into the measured plant states (i.e., y_i^m), as depicted in (5), where the measurement noise depends on a random variable ($w(t) \sim N(0,1)$) and the standard deviation of measurement errors ($\boldsymbol{\sigma_v} = 0.001\boldsymbol{y_i^m}$).

$$\tilde{\boldsymbol{y}}_{\boldsymbol{i}}^{\boldsymbol{m}} = \boldsymbol{y_i^m} + 0.001\boldsymbol{y_i^m} \tag{5}$$

Fig. 8 illustrates that the UDAE model could estimate the algebraic states adequately. However, there was a slight decrease in estimation performance compared to the previous cases, which was expected given the complexity of the addressed process dynamics. Nevertheless, the UDAE model could still reasonably estimate the predicted differential states (as shown in Fig. 9), highlighting its generalization capability.

4. CONCLUSION

A hybrid model for the gas-lift process was developed using UDAE. The approach was trained with offline data and validated with online data under conditions of model-plant mismatch. The sequential use of PSO and SLSQP to obtain the neural network weights and biases was adequate to avoid local minimum and accelerate the convergence. The

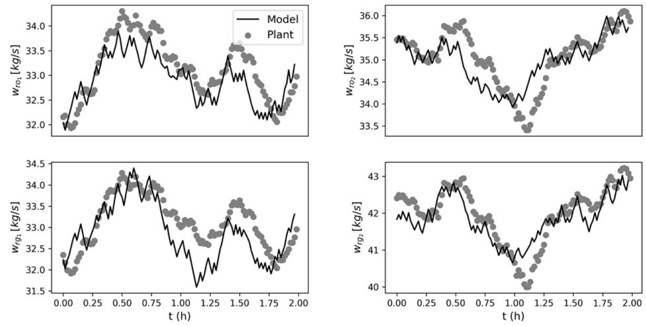

Fig. 8. The estimated reservoir gas and oil flow rates using the UDAE model, considering uncertainty in the value of GOR and measurement noise for a new process condition.

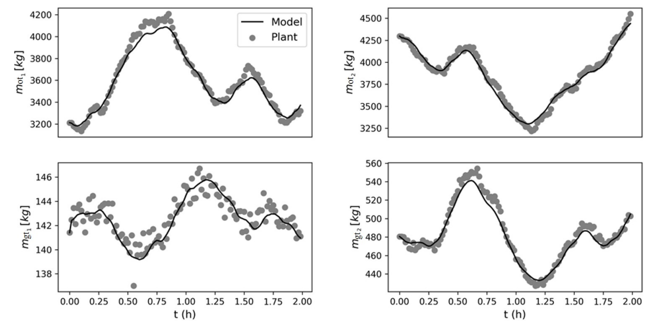

Fig. 9. The estimated gas and oil production mass flowrates using the UDAE model, considering uncertainty in the value of GOR and measurement noise for a new process condition.

hybrid strategy could efficiently model the dynamic behavior of the reservoir flow rates without needing information about the GOR and the PI. The hybrid model presented similar prediction to the model in Krishnamoorthy et al. (2018) considering the differential states. Furthermore, the UDAE model efficiently accounted for noise in the measurements, showing its generalization capability. The proposed UDAE model showed a potential for being used in optimization problems in future developments.

ACKNOWLEDGEMENTS

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. Professor Maurício B. de Souza Jr. is grateful for financial support from CNPq (Grant No. 311153/2021-6) and Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) (Grant No. E-26/201.148/2022).

REFERENCES

Andersson, J.A.E., Gillis, J., Horn, G., Rawlings, J.B., and Diehl, M. (2019). CasADi – A software framework for nonlinear optimization and optimal control. *Mathematical Programming Computation*, 11(1), 1–36. doi: 10.1007/s12532-018-0139-4.

Bangi, M.S.F., Kao, K., and Kwon, J.S.I. (2022). Physics-informed neural networks for hybrid modeling of lab-scale batch fermentation for β-carotene production using saccharomyces cerevisiae. *Chemical En-*

gineering Research and Design, 179, 415–423. doi: https://doi.org/10.1016/j.cherd.2022.01.041.

Carpio, R.R., dAvila, T.C., Taira, D.P., Ribeiro, L.D., Viera, B.F., Teixeira, A.F., Campos, M.M., and Secchi, A.R. (2021). Short-term oil production global optimization with operational constraints: A comparative study of nonlinear and piecewise linear formulations. *Journal of Petroleum Science and Engineering*, 198, 108141. doi: https://doi.org/10.1016/j.petrol.2020.108141.

Carvalho, M.A.d., Secchi, A.R., and Bagajewicz, M.J. (2016). Model reformulation and global optimization of oil production using gas lift. *Industrial Amp; Engineering Chemistry Research*, 55, 10114–10120. doi: 10.1021/acs.iecr.6b00223.

Delou, P.A., Matias, J., Jäschke, J., de Souza, M.B., and Secchi, A.R. (2023). Steady-state real-time optimization using transient measurements and approximated hammerstein dynamic model: A proof of concept in an experimental rig. *Journal of Process Control*, 132, 103111. doi:https://doi.org/10.1016/j.jprocont.2023.103111.

Dias, A.C.S.R., Soares, F.R., Jäschke, J., Souza Jr., M.B., and Pinto, J.C. (2019). Extracting valuable information from big data for machine learning control: An application for a gas lift process. *Processes*, 7(5). doi: https://doi.org/10.3390/pr7050252.

Franklin, T.S., Souza, L.S., Fontes, R.M., and Martins, M.A. (2022). A physics-informed neural networks (pinn) oriented approach to flow metering in oil wells: an esp lifted oil well system as a case study. *Digital Chemical Engineering*, 5, 100056. doi: https://doi.org/10.1016/j.dche.2022.100056.

Khan, M.R., Tariq, Z., and Abdulraheem, A. (2020). Application of artificial intelligence to estimate oil flow rate in gas-lift wells. *Natural Resources Research*, 29(6), 4017–4029. doi:https://doi.org/10.1007/s11053-020-09675-7.

Krishnamoorthy, D., Foss, B., and Skogestad, S. (2016). Real-time optimization under uncertainty applied to a gas lifted well network. *Processes*, 4(4). doi: https://doi.org/10.3390/pr4040052.

Krishnamoorthy, D., Foss, B., and Skogestad, S. (2018). Steady-state real-time optimization using transient measurements. *Computers & Chemical Engineering*, 115, 34–45.

Lima, F.A.R., Rebello, C.M., Costa, E.A., Santana, V.V., de Moares, M.G., Barreto, A.G., Secchi, A.R., de Souza, M.B., and Nogueira, I.B. (2023). Improved modeling of crystallization processes by universal differential equations. *Chemical Engineering Research and Design*. doi: https://doi.org/10.1016/j.cherd.2023.11.032.

Miranda, L.J. (2018). Pyswarms: a research toolkit for particle swarm optimization in python. *Journal of Open Source Software*, 3(21), 433. doi:10.21105/joss.00433.

Misener, R., Gounaris, C.E., and Floudas, C.A. (2009). Global optimization of gas lifting operations: A comparative study of piecewise linear formulations. *Industrial & Engineering Chemistry Research*, 48(13), 6098–6104. doi:https://doi.org/10.1021/ie8012117.

Miyoshi, S.C., Nunes, M., Salles, A., Secchi, A.R., de Souza, M.B., and Brandão, A.L. (2018). Nonlinear model predictive control application for gas-lift based oil production. In *28th European Symposium on Computer Aided Process Engineering*, volume 43 of *Computer Aided Chemical Engineering*, 1177–1182. Elsevier. doi: https://doi.org/10.1016/B978-0-444-64235-6.50205-9.

Namdar, H. (2019). Developing an improved approach to solving a new gas lift optimization problem. *Journal of Petroleum Exploration and Production Technology*, 9, 2965–2978. doi:https://doi.org/10.1007/s13202-019-0697-7.

Nogueira, I.B.R., Santana, V.V., Ribeiro, A.M., and Rodrigues, A.E. (2022). Using scientific machine learning to develop universal differential equation for multicomponent adsorption separation systems. *The Canadian Journal of Chemical Engineering*, 100(9), 2279–2290. doi:https://doi.org/10.1002/cjce.24495.

Peixoto, A.J., Pereira-Dias, D., Xaud, A.F., and Secchi, A.R. (2015). Modelling and extremum seeking control of gas lifted oil wells. *IFAC-PapersOnLine*, 48(6), 21–26. doi:https://doi.org/10.1016/j.ifacol.2015.08.004. 2nd IFAC Workshop on Automatic Control in Offshore Oil and Gas Production OOGP 2015.

Rackauckas, C., Ma, Y., Martensen, J., Warner, C., Zubov, K., Supekar, R., Skinner, D., Ramadhan, A., and Edelman, A. (2020). Universal differential equations for scientific machine learning. 1–55. doi: https://doi.org/10.48550/arXiv.2001.04385.

Shokir, E.M.E.M., Hamed, M.M.B., Ibrahim, A.E.S.B., and Mahgoub, I. (2017). Gas lift optimization using artificial neural network and integrated production modeling. *Energy & Fuels*, 31(9), 9302–9307. doi: https://doi.org/10.1021/acs.energyfuels.7b01690.

Soares, F.D.R., Secchi, A.R., and Souza Jr., M.B. (2022). Development of a nonlinear model predictive control for stabilization of a gas-lift oil well. *Industrial & Engineering Chemistry Research*, 61(24), 8411–8421. doi:10.1021/acs.iecr.1c04728. URL `https://doi.org/10.1021/acs.iecr.1c04728`.

Teixeira, B.O., Castro, W.S., Teixeira, A.F., and Aguirre, L.A. (2014). Data-driven soft sensor of downhole pressure for a gas-lift oil well. *Control Engineering Practice*, 22, 34–43. doi: https://doi.org/10.1016/j.conengprac.2013.09.005.

Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., et al. (2020). Scipy 1.0: fundamental algorithms for scientific computing in python. *Nature methods*, 17(3), 261–272. doi: https://doi.org/10.1038/s41592-019-0686-2.