COBEM 2023
27TH INTERNATIONAL CONGRESS OF MECHANICAL ENGENEERING

ABCM
Brazilian Society of Mechanical Sciences and Engineering
27th ABCM International Congress of Mechanical Engineering
December 04-08, 2023, Florianópolis-SC, Brazil

COB-2023-1242 DATA-DRIVEN MULTIPHASE FLOW PARAMETERS PREDICTION CAPABILITIES AND LIMITATIONS ON A REAL OIL WELL PRODUCTION DATA

Anderson Carlos Faller
Saon Crispim Vieira
Bernardo Pereira Foresti
Petrobras, Brazil
faller,saonvieira,foresti@petrobras.com.br

Adriano Todorovic Fabro
Departament of Mechanical Engineering, University of Brasilia, Brasília-DF, Brazil
fabro@unb.br

Marcelo Souza de Castro
School of Mechanical Engeneering - FEM, University of Campinas, Campinas-SP, Brazil
Center for Energy and Petroleum Studies - CEPETRO, University of Campinas-SP, Brazil.
mscastro@unicamp.br

Abstract: *Virtual Multiphase Flowmeters are model-based tools to estimate multiphase flow rates in pipelines that may replace physical flowmeters or test separators whenever those are not available, non existent or its use is not possible. They can be based on physics principles (simulation) or on data-driven models. We explore the application of the latter type across a few machine learning architectures by proposing a performance evaluation depending on the prediction goals. We state that data-driven models degrade over time due to changes in the operating conditions (short-term degradation) or due to slow changes in the fluid characteristics or reservoir inflow conditions (long-term degradation). Then, we apply the proposed methodology to quantify these degradations on a dataset extracted from a real oil well for over 3 years of its productive life. We show that, after training a model with a few hours of data, the prediction error increases on average 1-2 percentage points in the first 8 hours within the extrapolation range. When training a model with 2 years of historical data, the prediction error increases consistently with a rate of 30 percentage points in 10 months. We also show that larger models with time shifted input features yield better predictors.*

Keywords: *oil production, petroleum, virtual flowmeter, virtual sensing, machine learning*

1. INTRODUCTION

In petroleum production systems, oil, gas and water flow simultaneously (multiphase flow) from the reservoir to the processing facility through the production tubing, the subsea (if it is a satellite well) and surface pipelines. If necessary, an artificial lift method such as gas-lift may be used to supplement the energy from the reservoir (Thomas, 2004). Multiphase flow rates are periodically assessed via test separator alignment, which takes place with large time intervals between each other, or through multiphase flowmeters, which can provide real time online measurements with greater uncertainties. However, the latter is less available, so the engineering team must rely on sporadic test data and simulations to manage the field, optimize production, and diagnose anomalies related to, for example, flow assurance and integrity.

According to Varyan (2016), virtual multiphase flowmeters (VMFMs) are low-cost alternatives which can save significant capital expenditure in green and brown field projects. They are based on real time field instrumentation data (such as pressures and temperatures), and can be implemented using flow simulations, or can be purely based on black-box data-driven models. Simulation based VMFMs, while providing good extrapolation capabilities with respect to changes in operating conditions, require periodic human intervention and, fine tuning and accuracy of the model's physics. Data-driven VFMs, on the other hand, while having limited extrapolation performance due to the lack of physics principles, can provide a light weight and self-adjustable solution that can satisfy engineering requirements for production monitoring, optimization and anomaly detection. Also, both approaches can be used to estimate other parameters, such as pressures and temperatures (to detect deviations from expected values), void fraction, phase velocities, etc. VMFMs can be classified as an element of a digital twin instance, according to the definition by Grieves and Vickers (2017), as it is linked to a physical product (a production well) thoughout its life cicle, while providing real time estimations for physical quantities (liquid and gas flow rates).

We implemented and analyzed state-of-the-art machine-learning model architectures, such as recurrent models like

LSTM from Andrianov (2018) and Bi-LSTM from Ali *et al.* (2021), as well as non-recurrent models using Random Forests Breiman (2001) (an implementation similar to Bikmukhametov and Jäschke (2019)'s Gradient Boosting Machine), that predict oil production multiphase flow rates and other parameters based on historical data. Since these models don't consider any physics principles and rely purely only on field data, it is expected that their extrapolation capabilities are limited and become degraded due to changes in operating conditions (short term degradation) and slow variations in fluid properties, such as oil composition, water cut and gas-oil-ratio, and reservoir inflow conditions (long term degradation) over time. The goal of this study is to provide a methodology to quantify each model's degradation through time, measuring the evolution of its prediction error, then visualizing the results graphically.

In this paper, we tested this methodology on real well production data spanning several years, then showed that larger models with time-shifted features and larger training windows are better predictors than smaller and time-independent models, as they show smaller prediction errors when extrapolating in time. The extrapolation requirements and the available training data are taken into account when one has to select the proper model to the application. This paper aims at evaluating the data-driven VMFMs capability of accurately predicting the multiphase flow rates when the production well is no longer aligned to a test separator. The possibility of occurence of both intermittent and dispersed flow regimes is also taken into account. We present methodologies for the evaluation considering both prediction goals in short and long-term, which have specific characteristics summarized in Tab. 1.

Table 1. Comparison between both short and long-term prediction goals

Characteristic	Short-Term Prediction	Long-Term Prediction
Extrapolation Time	Hours (or days)	Months
Training Set Time Span	Contiguous period of production through test separator containing at least one cycle	Historical data contaning several production tests
Model Evaluation Method	Prediction error in the hours following the training data	Error trend when predicting entire future production tests
Error Sources	Small changes in operating conditions and random errors	Changes in fluid properties and reservoir inflow conditions

2. METHODS

In order to recreate a real application for a VMFM on a production well, we generated a dataset spanning over 3 years of its productive life, encompassing changes on its fluid composition (such as water cut) and on its stability conditions. Within this time span, we selected periods on which the well was aligned through a test separator and the liquid and gas flow rates are available. This dataset is described in Sec. 2.1.

The model architectures are described in Sec. 2.2. They share the same input/output structure, Fig. 1, with all the n input features (sensors) sampled in k equally spaced time instants, and m outputs at the present time t. The models were trained with the objective of predicting the instantaneous liquid and gas flow rates ($m = 2$), given a window of a multivariate time series containing the field sensors data. In this study, we used all the available sensors described in Sec. 2.1 as input features to the model ($n = 9$). However, one could reduce the input size by eliminating highly correlated or non-important features, or calculate engineered features as a function of the raw data (pressure drop along the pipeline, for example), which was not covered in the present work. The specific objectives evaluated include short-term and long-term extrapolation, to evaluate the model degradation over time, using the metrics described in Sec. 2.3.

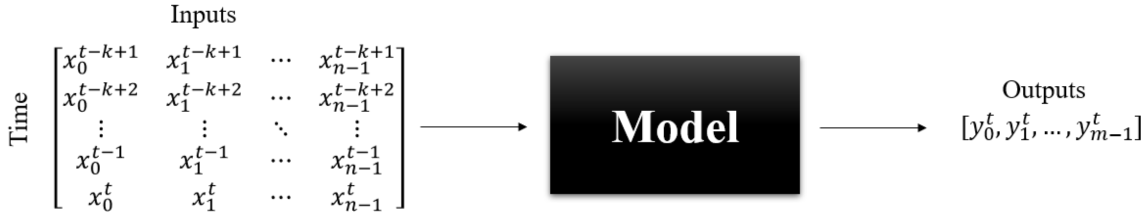

Figure 1. Model's inputs and outputs representation.

2.1 DATASET

The dataset prepared to evaluate the VMFMs extrapolation capabilities consists of a set of 70 production tests data from a single subsea satellite well with continuous gas-lift injection as its artificial lift method. The instrumentation available comprises bottomhole, wellhead, production choke (upstream and downstream) and gas-lift pressures and temperatures, annular pressure, gas-lift flow rate and production choke relative opening. The sample rate is 2 Hz and missing values were filled using linear interpolation.

Figure 2 shows a superposition of every production test's (identified by each color) liquid and gas flow rates. It is noticeable the wide range of flow rates and the diversity of flow regimes, based on the different amplitudes and shapes for each test.

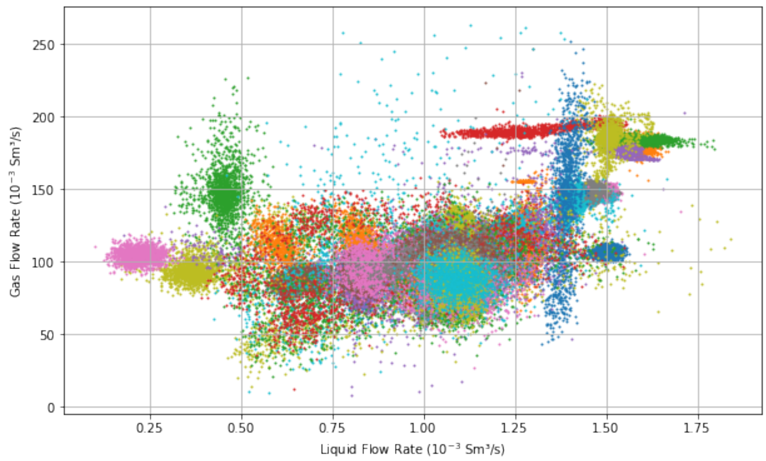

Figure 2. Instant liquid (water plus oil) and gas flow rates for different production tests.

2.2 MODELS

We performed the performance evaluation over the following model architecture types:

- Long Short-Term Memory (LSTM) Recurrent Neural Network (RNN) with an attached Multi Layer Perceptron (MLP), implemented similarly to Andrianov (2018), and Bidirectional LSTM (Bi-LSTM), as in Ali *et al.* (2021);
- Random Forest (RF) Regressor, using time-shifted input features (sliding windows), so the inputs and outputs match the ones from the RNN model and their performance can be directly compared;
- Random Forest (RF) Regressor without time-shifted inputs, so the outputs must be predicted based on the instantaneous sensor readings. This architecture was tested in order to assess the performance of memoryless models.

We tested several combinations of architecture variations, such as: both LSTM an Bi-LSTM RNNs, different numbers of recurrence cells, wide and narrow, shallow and deep MLPs, with and without dropout for regularization, and different connectivity strategies concerning the LSTM hidden states; different depth limitations on the Random Forests, pruning techniques, and other hyperparameters tuning. When applicable, the data was scaled for better learning performance.

2.3 METRICS

Usually, regression models are obtained by minimizing an estimation of the Mean Squared Error (MSE). However, the Mean Absolute Percentage Error (MAPE) is often used complementarly as an evaluation metric because of its intuitive interpretation as a relative error according to de Myttenaere *et al.* (2016). Moreover, it is used when the predicted quantity is sufficiently greater than zero, which is ensured in this study due to the dataset characteristics. The MAPE, for estimated values $\hat{y}_i$ and true values $y_i > 0$, is defined in Eq. 1.

$$\text{MAPE}(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^{n} \frac{|y_i - \hat{y}_i|}{y_i} \tag{1}$$

The coefficient of determination R^2 represents the model's ability to explain the variance of the data. While the goal is to obtain a model to explain most of the signal variance, this metric should be assessed carefully when evaluating stationary and noisy data. For instance, when a well is flowing within a dispersed flow regime, or with low amplitude hydrodynamic slugging, the signal-to-noise ratio can be very small, so the noise accounts for most of the variance. Good models should be able to filter out noisy data and explain high frequency variance only if it is correlated to the inputs. Due to the amount of nearly stationary and noisy data contained in our dataset, the R^2 was not evaluated.

3. RESULTS

We will explore the results considering both short and long-term prediction goals for our models. These are different modeling approaches and different training window sizes, therefore their degradations should be assessed differently, also considering different time windows. In short, as described in the next sections, short term prediction degradation (how much the MAPE increases) occurs over the course of hours, while long term prediction degradation occurs over the course of months after the last time instant where training data was available.

3.1 SHORT-TERM PREDICTION

The optimal time window and the model's important features (most relevant inputs) may vary depending on the well characteristics, the flow regime and the operating condition. In the following example, we trained a LSTM recurrent network attached to a MLP, using all sensors (pressures, temperatures and lift gas flow rate) and a 60 minutes time window as input features. After training with about 2 hours of data, the model was able to predict the next 9 hours with a MAPE of 11.7% for the liquid flow rates and 10.9% for the gas flow rates. Fig. 3 shows the model's predictions over both the training and the testing sets for the liquid flow rate, and Fig. 4 shows the model's predictions for the gas flow rate.

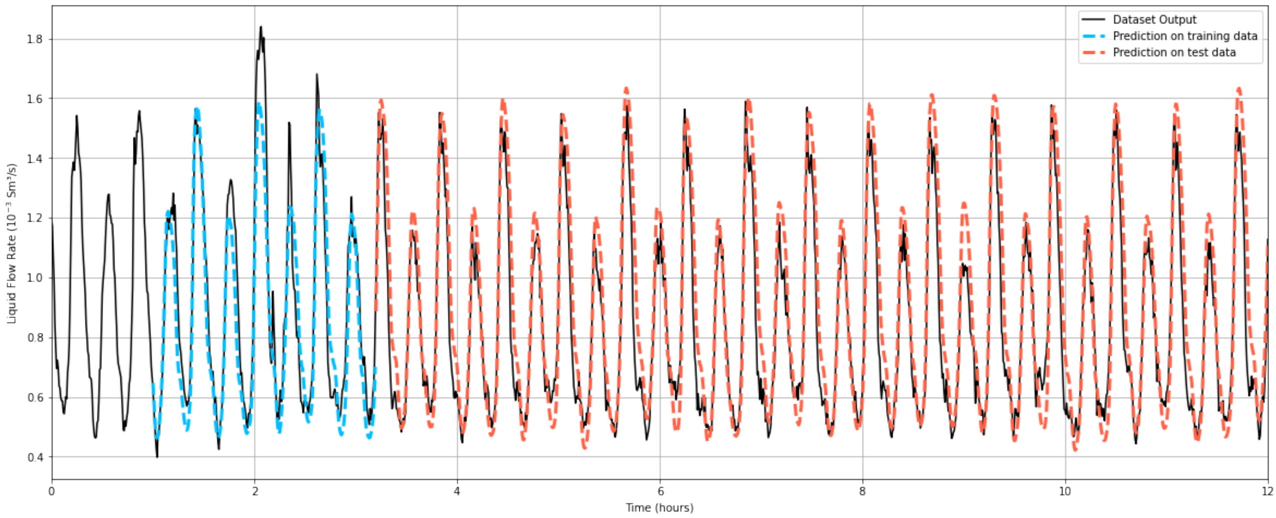

Figure 3. Liquid flow rate extrapolation with intermitent flow.

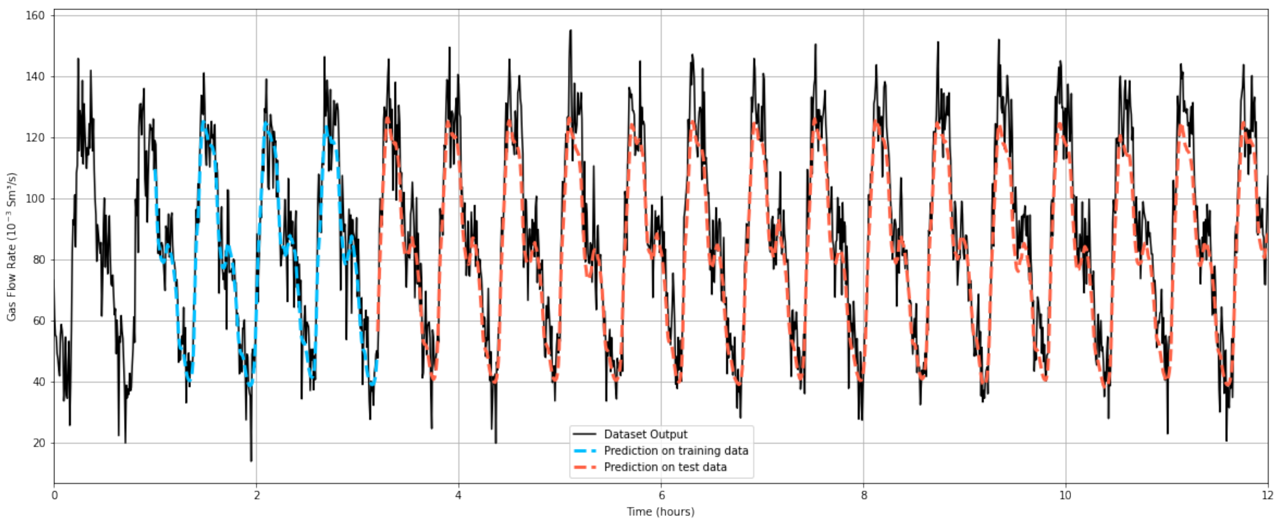

Figure 4. Gas flow rate extrapolation with intermitent flow.

While the absolute percentage errors may seem to be large in this specific example, it is important to point out that the outputs ranges are wide ($Q_{liquid}^{max}/Q_{liquid}^{min} = 3.6$ and $Q_{gas}^{max}/Q_{gas}^{min} = 7.8$) due to the intermittent flow, then this error measure is acceptable.

To evaluate how the different architectures behave and how the models degrade over time, we performed the same experiment described above with the entire dataset and calculated the MAPE for each hour in the prediction ranges. The selected models were the best performers for each architecture type: a Random Forest with no time window (no time shifted features), a Random Forest with a 30 minute time window, and a LSTM neural network with a 30 minute time window. Figure 5 shows how the prediction error increases, on average, as a function of time after the model training took place. Random Forest based models are more likely to overfit (smaller errors within the training data, at $t = 0$, immediately increasing in the first hour within the test data), but are still good at extrapolating. The LSTM model, on the other hand, yields consistent scores (between training and test data) due to its regularization, but were not able to extrapolate as well as the RF models. It is also important to point out that, after the extrapolation period begins (first hour), the MAPE increases about 1-2 percentage points in the next 8 hours, regardless of the model architecture.

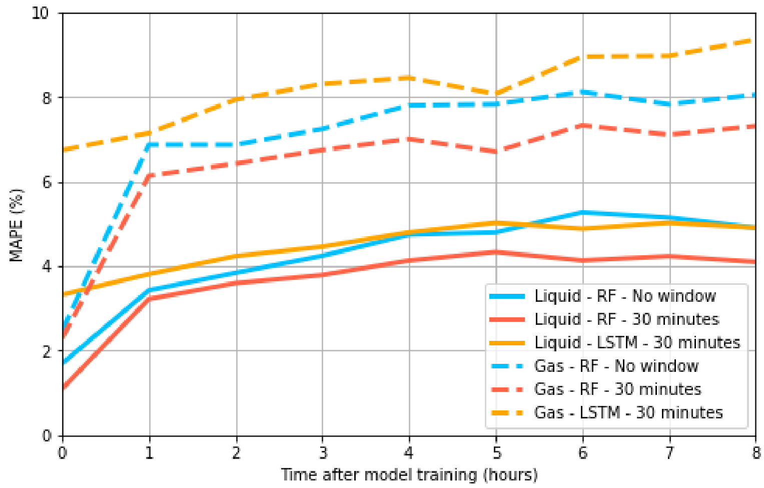

Figure 5. Hourly degradation of each model, averaged across the dataset.

It is shown that a Random Forest based model benefits from time shifted features, which means that the system dynamics observed in the sensors are relevant when predicting the instant flow rate for a given phase. This is expected, due to the transport phenomena involved in the multiphase flow through the wellbore and the long pipeline and riser, which causes delays in the inputs/outputs causal relations.

3.2 LONG-TERM PREDICTION

In this experiment, we chose an arbitrary time in the well's productive life where we trained a model with the entire historical data, then evaluated how the model degraded (how much the MAPE increased) on average for each production test after this given time.

Considering a Random Forest model trained with the last 35 available production tests (which happened over about 2 years), we performed the beforementioned evaluation for both liquid and gas flow rates and plotted the MAPE evolution, as can be seen in Fig. 6.

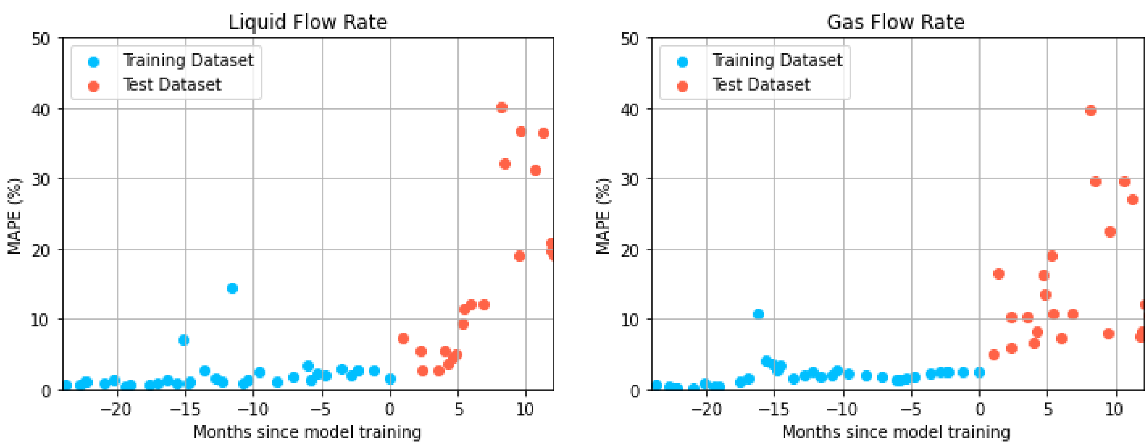

Figure 6. Long-term degradation after training.

There is a visible increasing trend on the model's extrapolation MAPE, despite the large variance. For both model's outputs (liquid and gas flow rates), as the model degrades over time, it can be said that the MAPE increases roughly, on average, 30 percentage points in 10 months.

3.3 RESULTS COMPARISON

Both short and long-term predictions measured degradations should be assessed separately, as they represent different goals. In the short-term goal, while the model is trained with a small amount of data, the measured MAPE on average increases minimally in the following hours (Fig. 5) because the operating conditions are not expected to change within the production test alignment period. In this case, the error increases due to small variations in the slug signature, as measured in the separator outputs. This can be observed in Fig. 3, where the model visibly fails to match a few slug peaks and tails.

When assessing the long-term prediction degradation, we should take into account that the operating conditions and the fluid characteristics may have changed in a way the model becomes unaware and starts to output wrong predictions, simply because it was not trained in those conditions. In the case presented in Sec. 3.2the extrapolation period comprises a water cut increase and also a change in the production choke opening, taking the input variables outside the training range. This explains the increasing MAPE.

Regarding the quality of the results, our work shows two case studies, one for each prediction goal, with a visualization for the proposed metric and how it evolves in time, with no intended comparison between those.

4. CONCLUSION

In this paper, we proposed the application of black-box data-driven Virtual Multiphase Flow Meters (VMFM) using different machine learning architectures by proposing a performance evaluation depending on the prediction goals. It is shown that when the production engineer is designing a black-box data-driven VMFM, the specific objective should be taken into account, because it affects not only the model's architecture and training method, but also how it's performance should be assessed.

We investigated two different objectives, then proposed the assessment of the model's degradation measuring the evolution of the MAPE over time. For the short-term prediction objective, the model is trained with a few hours from the last available data and tries to predict the flow rates. In this case, we showed that from the beginning of the prediction periods, on average, the MAPE increases 1-2 percentage points in 8 hours for both liquid and gas flow rates. For the long-term prediction objective, the model is trained with all the historical data available. Then, we showed that, for a particular case, the degradation increases following a clear trend of roughly 30 percentage points in 10 months for the MAPE.

The degradation, measured as the MAPE slope, will vary depending on the specific scenario. However, we have also shown that, given enough historical data from the well (or a similar one), the MAPE can be estimated, so the production engineers will be able to know beforehand how much prediction error can be expected a certain time after the model's training.

5. ACKNOWLEDGEMENTS

We would like to acknowledge PETROBRAS (grant number 2017/00778-2) and ANP for the financial support of this research, the Center for Energy and Petroleum Studies (CEPETRO) and School of Mechanical Engineering (FEM), both at the University of Campinas and the University of Brasilia (UnB).

6. REFERENCES

Ali, N., Viggiano, B., Tutkun, M. and Cal, R.B., 2021. "Data-driven machine learning for accurate prediction and statistical quantification of two phase flow regimes". *Journal of Petroleum Science and Engineering*, Vol. 202, p. 108488. ISSN 0920-4105. doi:https://doi.org/10.1016/j.petrol.2021.108488. URL `https://www.sciencedirect.com/science/article/pii/S0920410521001479`.

Andrianov, N., 2018. "A machine learning approach for virtual flow metering and forecasting". *IFAC-PapersOnLine*, Vol. 51, No. 8, pp. 191–196. ISSN 2405-8963. doi:https://doi.org/10.1016/j.ifacol.2018.06.376. URL `https://www.sciencedirect.com/science/article/pii/S2405896318307067`. 3rd IFAC Workshop on Automatic Control in Offshore Oil and Gas Production OOGP 2018.

Bikmukhametov, T. and Jäschke, J., 2019. "Oil production monitoring using gradient boosting machine learning algorithm". *IFAC-PapersOnLine*, Vol. 52, No. 1, pp. 514–519. ISSN 2405-8963. doi:https://doi.org/10.1016/j.ifacol.2019.06.114. URL `https://www.sciencedirect.com/science/article/pii/S2405896319302009`. 12th IFAC Symposium on Dynamics and Control of Process Systems, including Biosystems DYCOPS 2019.

Breiman, L., 2001. "Random forests". *Machine learning*, Vol. 45, pp. 5–32.

de Myttenaere, A., Golden, B., Grand, B.L. and Rossi, F., 2016. "Mean absolute percentage error for regression models". *Neurocomputing*, Vol. 192, pp. 38–48. doi:10.1016/j.neucom.2015.12.114. URL `https://doi.org/10.1016/j.neucom.2015.12.114`.

Grieves, M. and Vickers, J., 2017. "Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems". *Transdisciplinary perspectives on complex systems: New findings and approaches*, pp. 85–113.

Thomas, J.E., ed., 2004. *Fundamentos de Engenharia de Petróleo [Fundamentals of Petroleum Engineering]*. Editora Interciência, Rio de Janeiro, Brazil, 2nd edition.

Varyan, R., 2016. "Cost Saving of Implementing Virtual Flow Metering at Various Fields and Engineering Phases - A Case Study". Vol. Day 3 Thu, March 24, 2016 of *Offshore Technology Conference Asia*, p. D032S003R006. doi: 10.4043/26637-MS. URL `https://doi.org/10.4043/26637-MS`.

7. RESPONSIBILITY NOTICE

The authors are solely responsible for the printed material included in this paper.