Data-driven Estimation of Flowing Bottom-Hole Pressure in Petroleum Wells using Long Short-Term Memory

Mateus A. Fernandes
Dept. of Reservoir and Production Engineering
Petrobras
Santos, Brazil
matfernan@petrobras.com.br

Eduardo Gildin
Dept. of Petroleum Engineering
Texas A&M University
College Station, USA
egildin@tamu.edu

Marcio A. Sampaio
Dept. of Mining and Petroleum Engineering
University of São Paulo
São Paulo, Brazil
marciosampaio@usp.br

***Abstract*— Monitoring bottom-hole pressure (BHP) during flowing periods in petroleum wells is critically important for reservoir management, formation evaluation, lift optimization, and flow assurance. The advent of Permanent Downhole Gauges (PDGs) has changed paradigms regarding the availability and accessibility of this information. However, the costs associated with installing or replacing these gauges can sometimes be prohibitive. In this work, we propose a data-driven framework to fill the information gap when PDG data is unavailable. We used surface measurement data and wellhead gauge readings as inputs and train machine learning methods to accomplish this task. We implemented and evaluated several configurations of Long Short-Term Memory (LSTM) and compared their results with non-deep learning methods as Multi-Layer Perceptron (MLP) Neural Networks and Ridge Regression, taking as case study a real dataset from an offshore oilfield in the Brazilian Pre-salt. We show that our framework can provide reliable estimated values of flowing BHP, with error metrics MAPE and SMAPE consistently under 3%. We highlight the following main contributions: (1) the utilization of time-related data; (2) the development of a framework applicable across a wide range of reservoir and flow conditions; and (3) the potential for practical applications, including real-time monitoring.**

Keywords—soft sensors, long short-term memory, neural networks, regression analysis, petroleum industry

I. Introduction

Monitoring bottom-hole pressure (BHP) during the operation of oil and gas producing wells is fundamental for several disciplines of petroleum engineering. Reservoir management and formation evaluation rely on this information to monitor pressure decline, calculate the productivity index, estimate the skin factor, and anticipate phenomena that can cause production losses [1]. Meanwhile, operational areas focus on flow assurance and lift optimization [2].

Ideally, continuous, real-time pressure monitoring is achieved through sensors installed at the bottom of the wells—Permanent Downhole Gauges (PDGs)—positioned as close as possible to the perforated intervals, where the inflow of fluids from the reservoir to the production tubing occurs. Although PDGs are more commonly used in high-productivity offshore wells, the installation costs for these sensors and the associated monitoring systems can be prohibitive in other cases, particularly in wells with older completions or those expected to have low productivity and longevity. In instances where a PDG signal is not available, bottom-hole pressure measurements can only be performed sporadically through workover operations, which require interruptions to regular production while positioning temporary downhole test tools. Even in wells originally equipped with PDGs, information may not always be readily available due to potential failures of the sensor itself or its communication with topside monitoring systems.

The use of surface and wellhead data, combined with knowledge of the flow dynamics of produced fluids, can help bridge the gap in bottom-hole pressure information in wells without a functional PDG. This surface data typically includes measurements of oil, gas, and water flow rates, choke bean aperture size, pressure, and temperature. For offshore wells where the wellhead is on the seabed, additional pressure and temperature gauges are usually present. However, modeling flow dynamics can be challenging due to variations in the relationships among different fluids in multiphase flow, which are influenced by pressure and temperature gradients along the path from the reservoir to the surface, as well as the consequent phase changes.

Empirical equations are traditionally employed for these estimates, though they are often calibrated for specific datasets in each field and may not be suitable for reproduction in other plays where fluid and flow characteristics differ. Some of the more notable works in this area are [3], [4], and [5], which continue to be widely used in the industry decades after their publication..

II. Related Work

With the advent of data-driven, artificial intelligence, and deep learning methods, this area of research headed to a new direction. Developments on analytical and numerical solutions, nonetheless, are still useful for studies of the multiphase flow, especially in cases where we must calculate pressure losses in tubing and pipelines [6]. In specific cases, merging equation-based, numerical, and artificial intelligence methods can yield positive results [7].

1946-0759/24/$31.00 ©2024 IEEE
DOI 10.1109/ICMLA61862.2024.00236

The use of data-driven methods to estimate well variables is, however, a reality, being accepted in the industry as a replacement for physical meters in monitoring actions and even for operational control, as highlighted by [8]. This same work also provides an extensive analysis of practical aspects of this replacement (focused on estimating BHP), including variable selection, the choice of suitable models, feature engineering, data conditioning, and challenges in the real-world application of these so-called soft sensors.

The complexity of applying data-driven methods in real cases is also highlighted in [2]. The authors show that, for a simulated model, the reproduction of BHP is possible even using only one input variable, the choke bean size. For real noisy data, obtained from wells operating with gas lift and under slug flow, a far more complex framework was required. Nevertheless, they achieved good results using Echo State Networks.

Depending on the produced fluids and artificial lift methods, the approach and relevant variables to estimate BHP can differ significantly, as we can observe in the recent literature. The work in [9] focused on coal seam gas wells, gathering surface and subsurface data from five wells over periods of up to 18 months. They classified these wells into groups based on flow characteristics and added a categorical variable to represent these groups. Testing Neural Networks alongside a set of regression methods, they found that the Neural Networks outperformed the regression methods, concluding that their ability to handle non-linear relationships was essential.

The study in [10] used Neural Networks to estimate the BHP, using as inputs surface measurements, but also a downhole variable, in this case, the temperature. Their case study used data from the Volve offshore oilfield, an open dataset frequently used as a benchmark in petroleum engineering publications.

In their approach to estimate flowing BHP in a tight sand gas condensate field, [11] used pressure and temperature at the wellhead, depth, fluid specific gravity, chloride content, water cut, and fluid rates as inputs. They emphasized that most classical correlations are empirical and developed for specific conditions due to the complexity of systems and multiphase flow. Their results showed that Neural Networks could outperform these correlations. Although they implemented an LSTM model, its performance was not the best.

As an example of different approach while dealing with artificial lift, [12] estimated BHP in oil wells using Electrical Submersible Pumps (ESP), using a combination of Response Surface Modelling (RSM) and Neural Networks. In this study, they selected variables related to the pumping system, as inlet and outlet pressures and temperatures, electric current, and rpm.

The work by [13] calculate BHP for an oil well using empirical correlations in the form of a system of linear equations, which are selected through a Quasi-Monte Carlo method and having as input variables oil, gas, water rates, and flowing WHP.

We consider, however, that the literature has a gap addressing flowing BHP estimation in oilfields where concept drift is relevant, meaning that production parameters vary significantly from well to well and over time. Additionally, most frameworks are not suitable for practical, real-time applications. So, our intention is to go further and explore the possibilities of a new deep learning approach, thought to be validated over time for diverse production conditions.

III. Objectives and Contributions

The objective of this study is to estimate the flowing bottom-hole pressure of producing oil wells in regular operation, using a data-driven approach based on machine learning (deep learning) methods . In particular, we implement and evaluate several configurations of Long Short-Term Memory (LSTM) and Multi-Layer Perceptron (MLP) Neural Networks, as well as Ridge Regression to accomplish this task, extracting knowledge and "learning" the fluid flow dynamics in well tubing and production flowlines in order to estimate BHP from surface measurements and wellhead gauge data, without relying on analytical models, empirical equations, or the properties of produced fluids. We selected this ensemble of methods to encompass both simpler, widely adopted techniques and some of the most advanced methods prominent in recent deep learning applications.

Our motivation is to obtain a virtual meter capable of estimating BHP reliably enough to be used in well and reservoir monitoring actions, including real-time possibilities, being able to be deployed when hard bottom hole measures are unavailable.

The proposed framework is evaluated using datasets from an offshore oilfield in the Brazilian pre-salt. In this field, as a project assumption all wells have one or more PDGs installed at completion, but we observed that 15.2% (19/125) of these have failed since then. We also noted that failures in downhole gauges are 220% more frequent (and their repair is harder and more expensive) than in wellhead gauges, what results in a relevant number of wells to which our framework can be applied. We highlight that this methodology is intended to be general enough to be applied to other fields and situations where similar input variables (as we will detail later) are available, including wells for which installing a PDG is not economically or technically viable.

As main contributions of this work, we list the features that distinguish our framework:

- Approach using time-related data with deep learning;
- Suitability for a broad range of reservoir and flow conditions, as we verify the validity of the models over time for wells under different production conditions;
- Potential application in real cases as a virtual meter in wells without available PDGs, including real-time monitoring.

IV. Methods

In this section we describe how the dataset was obtained, discuss its characteristics, introduce the machine learning methods chosen to estimate the BHP in the case study, and present the proposed framework to solve this problem.

A. Database

We used as a case study data collected over the last six years from five oil wells attached to one of the platforms producing from an offshore reservoir in the Brazilian coast. The variables

we assumed as relevant to estimate BHP and that are readily available in the operator company's databases are the following:

- wellhead pressure and wellhead temperature (gauges on seabed);
- choke inlet pressure, choke inlet temperature, and choke aperture (gauges at topside, attached to well control valves);
- oil, gas, and water production rates (measured daily at the platform).

An important characteristic of this particular dataset is the diversity of reservoir and fluid conditions both spatially and temporally, what is a consequence of a superposition of factors: (1) spatial compositional variations of the original fluids due to the large volume of the reservoir; (2) evolution of drainage at different regions of the reservoir; (3) rock heterogeneities; and (4) recovery strategies used over time, including Water Alternating Gas (WAG) injection. These factors, plus differences in perforation depths, well geometries and flowline lengths, lead to a large variety of flow conditions, resulting in a challenging problem as we want our machine learning to estimate BHP with good accuracy in this plethora of conditions. To illustrate this, we present in Table 1 the list of wells that make our dataset, accompanied by their ranges of values for some of the selected variables.

Although the dataset includes wells from a single platform, it is comprehensive enough to validate the method before involving other platforms

For production rates, we have daily data derived from the total production of the platform, prorated to individual wells according to their most recent production test, which imposes additional uncertainties. All other variables, including the BHP itself, come from the Plant Information (PI) real-time monitoring system, where they are registered continuously.

It is noteworthy that we have access to bottom-hole temperature (BHT) data; however, as our objective is to estimate BHP in the absence of a PDG signal, we decided not to use information from another variable measured at the bottom hole, as the most common failure mode is losing both PDG signals.

B. Data Conditioning and Feature Engineering

Data conditioning is a fundamental step to support good feature extraction and "feed" the machine learning methods with information in an adequate format, with quantity and quality to perform well. In this step, we started from the database described in the previous section, loaded as a data frame in a Python environment, to which we performed the following sequence of actions:

- Adjusting all the variables to the same time base, by calculating daily averages for those recorded in the PI system.
- Calculating gas-oil rates (GOR) and water cuts from the production rates, to replace gas and water rates in our dataset with variables that may have reduced ranges of values. The equations are as follows:

$$GOR = Q_g / Q_o \tag{1}$$

$$W_{cut} = Q_w / (Q_o + Q_w) \tag{2}$$

  where Q_o, Q_g, and Q_w represent oil, gas, and water production rates, respectively.
- Excluding from the dataset samples that have null values for at least one of the variables.
- Removing from the dataset samples in the periods of shut-ins, for which the choke aperture and the production rates are zero, since we are dealing only with flowing BHP (static pressures would require a completely different approach, due to stabilization periods).
- Removing samples with on-time less than 2 hours per day and with choke apertures close to zero (<5%) in the daily averages to avoid biases of transients and unusual production conditions.
- Separating data from one of the wells (P-46) to be used as our test dataset, then proceeding with additional data conditioning with the remaining data (that will be used as training and validation datasets). This is required to prevent the learning procedure from having any a priori information from the test dataset.
- Checking for outliers, initially observing the interquartile ranges (*IQR*) and then removing the entire sample *i* if one of the variables x_i is considered an outlier, using:

$$x_i < Q_1 - 1.5\ IQR, \text{ or} \tag{3}$$

$$x_i > Q_3 + 1.5\ IQR \tag{4}$$

  where Q_1 and Q_3 are the values for which the Cumulative Distribution Function for x equals to 25% and 75%, respectively, and $IQR = Q_3 - Q_1$. Exceptions are made if deviations slightly larger than the threshold make physical sense, e.g. oil rates in wells operating temporarily with restrictions.

In Table 2, we present a statistical summary of the variables after the steps described so far.

TABLE I. GENERAL DATA FROM WELLS USED IN THE CASE STUDY

Well	Depth PDG (m)	Oil Potential range (m^3/d)	Water Cut range (%)	GOR range (adim.)	BHP range (kgf/cm^2)	Temp. surf. range (°C)
P-2	4831	403.7 to 3094.6	7.7 to 45.1	183.6 to 244.0	333.6 to 488.6	25.0 to 39.5
P-8	4746	218.9 to 1666.5	0 to 12.2	269.8 to 574.5	233.9 to 403.0	0.3 to 22.6
P-9	4853	662.0 to 4423.5	0 to 61.5	173.1 to 619.0	286.7 to 489.4	13.6 to 42.5
P-46	4670	900.6 to 4094.4	0 to 56.0	177.6 to 609.6	268.8 to 439.6	17.1 to 42.0
P-86	4834	567.7 to 2585.7	0 to 42.5	192.1 to 809.7	216.1 to 440.3	0.4 to 27.5

TABLE II. STATISTICAL SUMMARY OF THE DATASET

Stats.	Depth PDG (m)	Oil rate (m³/d)	Water cut (%)	GOR (adim.)	Choke Open. (%)	Press. Choke in. (kgf/cm²)	Temp. Choke in. (°C)	WHP (kgf/cm²)	WHT (°C)	BHP (kgf/cm²)
count	4710	4710	4710	4710	4710	4710	4710	4710	4710	4710
mean	4833.6	1803.8	21.3	310.6	79.2	29.9	25.9	168.0	49.5	362.1
std	29.7	1021.2	20.6	126.2	33.1	17.1	12.2	43.0	5.9	69.4
min	4746.0	209.4	0.0	173.1	10.0	13.6	0.3	80.3	15.0	216.1
max	4853.0	4423.5	61.5	809.7	103.3	143.9	42.5	292.0	57.0	489.4

After the data conditioning steps, we analyzed the correlations between the eight variables available in the database that could be used to estimate BHP. Based on the data from the wells, excluding the one used for the tests, we obtained the correlation heatmap shown in Fig. 1, consisting of values calculated using Pearson's coefficient [14]:

$$r = \frac{\sum(x_i - \bar{x})\cdot(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \cdot (y_i - \bar{y})^2}} \quad (5)$$

where r is the correlation coefficient, x and y the samples of two different sets of variables.

From this plot, we observed that the choke inlet pressure has a correlation with BHP that is much weaker than that of the other variables. Therefore, we decided to eliminate this variable to reduce complexity without losing significant information. We extended this step to the test dataset.

To prepare our dataset for k-fold cross-validation [15], we created four versions of training-validation subsets, removing from the original one well at each time to be used as validation while training the machine learning methods with the remaining data, as illustrated in Fig. 2.

As the final step in data preparation, we performed Gaussian normalization. It is important to note that the mean and standard deviation are calculated using only the training data to prevent the learning procedure from having any prior information from the validation and test sets. We save the transformation parameters to be used later with these datasets. We emphasize that the only output variable to be estimated is the bottom-hole pressure (BHP), while the other variables serve as inputs for the machine learning methods.

Fig. 1. Correlation heatmap for the variables of the problem.

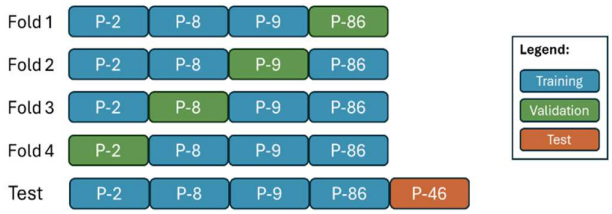

Fig. 2. Dataset preparation for training and 4-fold cross-validation.

C. *Aspects of the Machine Learning Methods*

In this work, we evaluated three machine learning methods to address the proposed task. These methods are briefly described in the following subsections, and detailed information can be found in the cited references.

1) Multivariable Linear Regression: Linear Regression is perhaps the simplest and most used regression method. It assumes that the outputs can be expressed as a weighted sum of the input variables [16]. The training of such a model consists of optimizing those weights to minimize a loss function applied to the training dataset [17]. Due to its ease of implementation and interpretation, this method is often the first benchmark for any regression problem. Some variants of the linear regression include regularization terms, which act as penalties for the regression coefficients. This approach helps prevent overfitting by reducing the influence of input variables that may capture noise or outliers, thereby hindering the generalization capabilities of the regressor.

Ridge regression adds a term in the equation of loss function (that will be minimized in the training phase), as shown below:

$$f_{ridge} = \sum_{i=1}^{n}[y_i - (w \cdot x_i + b)]^2 + \alpha \cdot \sum_{j=1}^{p} w_j^2 \quad (6)$$

where y is the predicted value, x is the input value, w and b the coefficients to be adjusted. The last term, responsible for the regularization, is weighted by a factor α, which can be tuned to control the penalties of the regression coefficients. Setting α=0 reduces the equation to the simple linear regression, while large values may exaggeratedly shrink the coefficients [18]. We use the regression based on Eq. 6, with α as a hyperparameter adjusted during the training and validation of the model.

2) Neural Networks: Developed from an approach known as connectionist, Artificial Neural Networks are based on units called neurons (or nodes), which individually can take an input data vector and calculate an output by performing a weighted sum followed by the application of an activation function [17].

The combination of a number of these basic units in different structures forms Neural Networks with distinct characteristics. Their training consists of adjusting the weights of the connections (and eventually biases) to reproduce the relations between inputs and outputs of a training dataset.

In this work, we adopted the Multilayer Perceptron (MLP) model, a traditional feed-forward structure with neurons arranged in sequential layers, capable of performing regression for non-linear problems [17].

3) Long Short-Term Memory: The current availability of high-capacity CPUs and GPUs allows, however, that more complex Neural Network models become universally accessible [16]. This includes a plethora of recently developed methods that are already popular in applications involving all areas of knowledge, reffered to as Deep Learning. Among these, one of the most prominent branches is that of recurrent networks, designed to operate with internal states in the neurons that can retain previous states, what make them more suited for dealing with sequential data, as time series [16]. This capability is well explored by LSTM, whose long-term and short-term memory units can retain information from distant sequences, using dedicated gates to decide when to keep or override this information, and also to decide when to access it [19].

One remarkable difference between LSTM and the previous methods is how we construct the input samples. Since we are dealing with time-related events, each sample carries values for a number of previous time steps, what requires an additional procedure in data preparation. Fig. 3 exemplifies this procedure, considering n input variables and p previous time steps feeding LSTM to estimate BHP at a time t. It is noteworthy that while p is a hyperparameter for LSTM, the other benchmark methods are not time-related, using p=0.

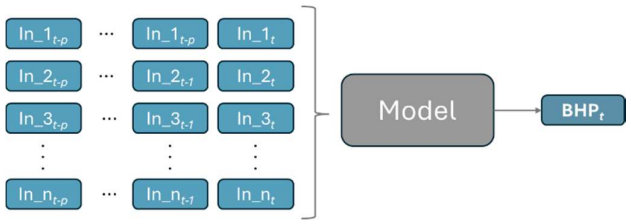

Fig. 3. Input samples preparation.

D. Evaluation metrics

We place special attention on the proper performance evaluation of predictors/estimators, especially when working with datasets in the form of time series. In the literature, the metrics most often used are the Mean Absolute Percentage Error (MAPE) and the Root Mean Squared Error (RMSE). However, these metrics have limitations, such as a greater penalty for models that tend to diverge towards higher values—since the error can be unlimited in this direction but is limited when it approaches lower values. Therefore, we also include the Symmetric Mean Absolute Percentage Error (SMAPE) in our evaluations, which addresses the issue described, despite showing a slight tendency in the opposite direction [20]. These metrics are defined by the following equations:

$$MAPE = \frac{1}{n}\sum_{j=1}^{n}\left|\frac{y_j-\hat{y}_j}{y_j}\right| \quad (7)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(y_j-\hat{y}_j\right)^2} \quad (8)$$

$$SMAPE = \frac{1}{n}\sum_{j=1}^{n}\frac{|y_j-\hat{y}_j|}{\frac{(y_j+\hat{y}_j)}{2}} \quad (9)$$

where y_j are the predicted values and $\hat{y}_j$ the actual values of the dependent variable [21].

Our evaluation consists of comparing the estimated values for BHP with the actual measured values. Additionally, we give importance to a qualitative evaluation, observing the estimators' ability of reproducing trends in pressure behavior, for example, when wells are experiencing restrictions, decline or reacting to the effects of an injection.

V. Results

We present the results in two phases: in the first we tune hyperparameters for the LSTM and the benchmark NN and ridge regression models using validation dataset; in the second we evaluate performance for a test dataset, simulating a real-life situation for an oil well without available PDG data.

A. Hyperparameters tuning and model selection

In this step, we submitted the pre-processed and segmented datasets to the machine learning models for training and validation, adjusting the hyperparameters of these models to maximize performance using grid search. As we shown previously in Figure 2, this step is based on a 4-fold cross validation. This means we calculate performance metrics four times, each of them segregating data from one of the wells for validation while training our models with the remaining three.

The multivariable linear (ridge) regression requires only one parameter to be adjusted, that is the regularization coefficient α, as shown in Eq. 6. The MLP network requires a definition of its architecture, including the number of layers and neurons in each layer, in addition to the activation functions of its nodes; in all cases we used the Adam optimizer algorithm for training. For the LSTM we must choose the number of elements in the network, the number of timesteps to be considered in each of the input vectors, and the activation function. For the latter, we also included tests with different numbers of input variables, evaluating if a reduced number (and consequent reduced complexity) would bring similar performance. We did that by removing preferentially the variables with lowest correlation to BHP, according to the coefficients shown in Fig. 1, and tested a total of five combinations. In Table 3, we describe the sets and ranges of hyperparameters evaluated for each method.

In this process, we used the principle of looking for the simplest possible model that meets the quality criteria for the estimates. Thus, we started the empirical adjustment of hyperparameters with simpler models and evaluated the performance gain as their complexity increased [22]. We also monitored the inflection point in the performance curves, which showed us when increased complexity no longer represented an improvement, causing overfitting

TABLE III. SETS OF HYPERPARAMETERS EVALUATED WITH GRID SEARCH

Method	Hyperparameters
Ridge Regression	α [0.1-1.0]
MLP Networks	layers [1-2]; neurons [10-100]; activation [linear, relu, tanh]
LSTM Networks	elements [10-200]; timesteps [3-15]; activation [linear, relu, tanh]; datasets [1-5]

We present the relevant results of the model tuning in Table 4, showcasing the combinations with the best performances for ridge regression and the MLP network, as well as three selected combinations for the LSTM networks, computing SMAPE, MAPE, and normalized RMSE. During this search, we ran each combination multiple times to minimize the inherent randomness in the training process. In the first column, we indicate the parameters for each of the models. Notably, LSTM 1 and LSTM 2 achieved the best performances using six input variables, excluding choke opening and water cut—both of which have the first and second lowest correlations to BHP. We also included LSTM 3, which is the best using all input variables, for the sake of comparison. From the results, we highlight the following general observations:

- Errors are low, indicating good overall performance.
- In the LSTM models, nRMSE values are closer to the MAPE values, indicating small error variance and less influence from outliers [21].
- The estimators with best performance under all conditions were the LSTM models using six inputs.
- Reducing the number of input variables in the MLP and ridge regression models resulted in poorer performance.

TABLE IV. PERFORMANCE OF THE BEST MODELS AFTER HYPERPARAMETER OPTIMIZATION

Model	SMAPE	MAPE	nRMSE
Ridge Regression [$\alpha = 0.2$]	0.0378	0.0364	0.0521
MLP Network [1 layer, 50 el., relu]	0.0316	0.0323	0.0510
LSTM 1 [150 el., 6 steps, relu, 6 inputs]	0.0274	0.0274	**0.0364**
LSTM 2 [120 el., 5 steps, relu, 6 inputs]	**0.0266**	**0.0265**	0.0372
LSTM 3 [100 el., 15 steps, relu, 8 inputs]	0.0504	0.0483	0.0653

Next, in Fig. 4 we compare the predicted BHP values with the real values during a relevant production period for well P-2, included in the 4-fold validation dataset. We also show some of the input variables: oil rate, GOR, and choke aperture.

B. Evaluating the test dataset

For the final evaluation of this work, we utilized data from well P-46, which was initially set aside as the test dataset. Our best LSTM model was trained using data from the other four wells, as previously illustrated in Fig. 2. The results of the BHP estimates are depicted in Fig. 5, that also shows the input variables oil rate, GOR, and choke aperture.

Well P-46 experienced a real data unavailability issue in its history, from the beginning of 2018 until late 2020. After this period, the signal was restored. One of the motivations for this work was to estimate BHP values for these periods without signal. Quantitatively, our best model yielded SMAPE = 0.0296, MAPE = 0.0301, and nRMSE = 0.0328, measured during the periods where we had actual BHP values for comparison.

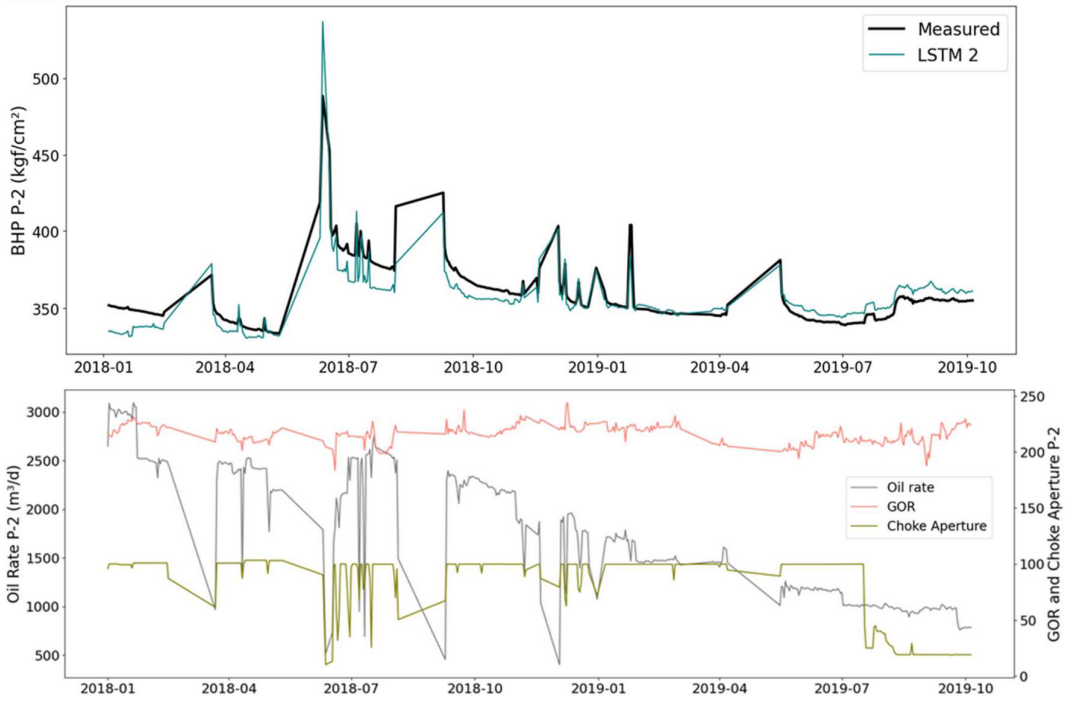

Fig. 4. Predicted BHP (best model) versus real BHP for well P-2, and corresponding inputs (oil rate, GOR, and choke aperture).

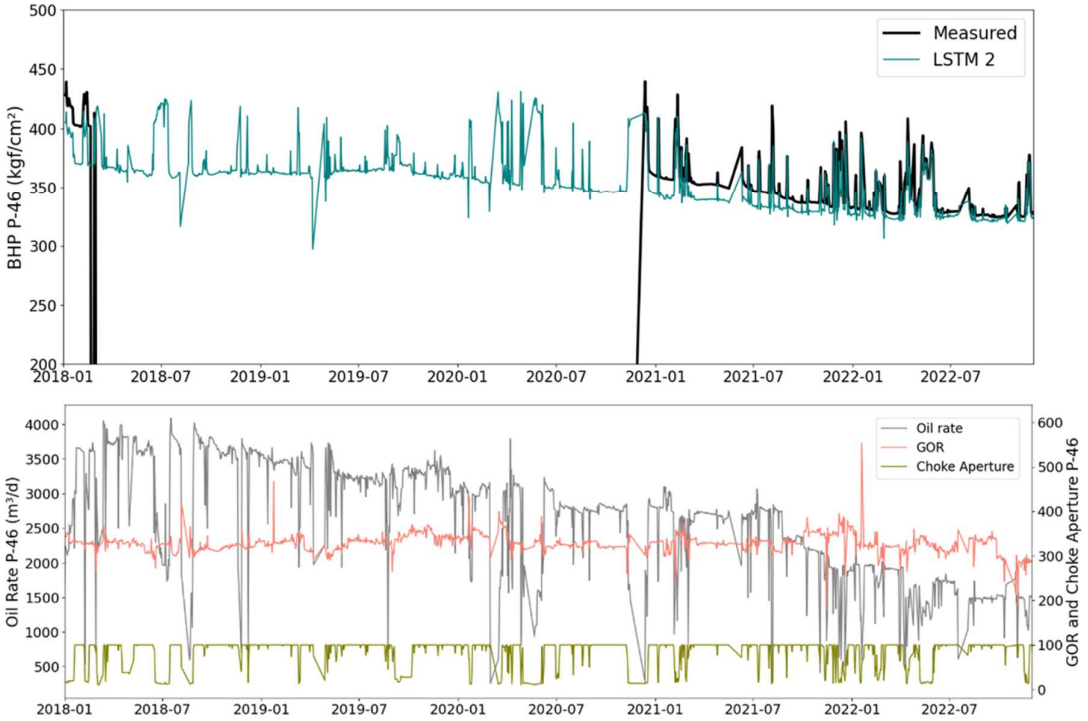

Fig. 5. Predicted BHP (best model) versus measured BHP for well P-46, and corresponding inputs (oil rate, GOR, and choke aperture).

Observing the plots, we can see that the model had good performance both before the first failure and after the signal restoration in late 2020. This gives us confidence to adopt the values estimated by our model during the "blank" periods in studies regarding well and reservoir behavior, as we demonstrated that variations over time in production parameters did not cause degradation in our model's performance.

We highlight that the qualitative analysis in general reinforces the results from the quantitative metrics, given that our best models presented excellent abilities to reproduce the pressure behavior, even in transient situations (peaks) or when production conditions are altered, showing model stability over time. These results can be considered satisfactory, noting that the measured data themselves present an inherent range of uncertainty.

However, to achieve results like these, we need a dataset comprehensive enough to encompass most of the situations we will encounter in the predictions. Cases, for example, of wells producing with combinations of water cut and GOR not seen during training can lead to greater deviations or unexpected outcomes. Therefore, we believe that a larger database would be beneficial to enhance robustness and reduce uncertainties in the machine learning methods.

In terms of processing times, LSTMs have a significantly longer training duration, averaging between 1-2 minutes in the evaluated configurations while running on a cluster with multiple processors. MLP neural networks and regression are trained significantly more rapidly, with training times of 0.4 s and 0.075 s. Once training is completed, all models can estimate BHP within few milliseconds. This latency is deemed suitable for well and reservoir monitoring in practical applications, allowing for seamless integration with PI systems to facilitate continuous, real-time data tracking and analysis.

VI. Conclusions

In this work, we presented the analysis of a data-driven and machine learning-based model to estimate the flowing bottom-hole pressure in oil-producing wells. Models of LSTM Neural Networks are implemented and evaluated as our flagship models, and their results were compared to those obtained using MLP networks and Ridge Linear Regression. These methods have proven capable of operating within a simple framework and providing fast response, using as inputs data that is regularly obtained in well operations, such as the flow rates of the fluids in production, choke valve aperture, pressures and temperatures at wellhead and topside. Importantly, we dismiss the use of correlations, knowledge about flow patterns, or even details about well and flowline's geometries.

We used an oilfield in the Brazilian pre-salt region as a case study, collecting data from five wells. One of these was used as the test dataset, while the others were arranged for a 4-fold training/validation scheme. After adequate pre-processing, feature extraction, and hyperparameter tuning, we were able to achieve consistently MAPE and SMAPE values under 3% with our best LSTM estimators.

When we compare the estimates with real measured data in a time-series plot, it becomes clear that the method is a viable alternative to replace real PDG data in cases where there is sufficient history to construct a robust training dataset. Thus, we can reliably use estimated BHPs when PDGs fail or are unavailable.

ACKNOWLEDGMENT

The authors express their gratitude to the Laboratory of Reservoir Simulation and Management (LASG), University of São Paulo (USP) and Texas A&M University for their support. We also extend our thanks to Petróleo Brasileiro S.A. (PETROBRAS) for granting permission to publish this paper.

REFERENCES

[1] Alain Gringarten, Thomas Von Schroeter, Trond Rolfsvåg, and John Bruner, "Use of downhole permanent pressure gauge data to diagnose production problems in a North Sea horizontal well", SPE Annual Technical Conference and Exhibition, Denver, Colorado, 2003.

[2] Eric A. Antonelo, Eduardo Camponogara, and Bjarne Foss, "Echo State Networks for data-driven downhole pressure estimation in gas-lift oil wells", Neural Networks, vol. 85, pp. 106-117, 2017.

[3] H. Duns and N. C. J. Ros, "Vertical flow of gas and liquid mixtures in wells", 6th World Petroleum Congress, Frankfurt, Germany, 1963.

[4] A. R. Hagedorn and K. E. Brown, "Experimental study of pressure gradients occurring during continuous two-phase flow in small diameter vertical conduits", Journal of Petroleum Technology, pp. 475-484, 1965.

[5] H. D. Beggs and J. P. Brill, "A study of two-phase flow in inclined pipes", Journal of Petroleum Technology, vol. 25, n. 5, pp. 607-617, 1973.

[6] Clinton Akinseye, Omotara Oluwagbenga, and Abdulwahab Giwa, "Determination of flowing bottom-hole pressure of a well using modified Guo's model", Journal of Environmental Science, Computer Science and Engineering & Technology, vol. 9, pp. 270-286, 2020.

[7] X. Li, J. L. Miskimins, and B. T. Hoffman, "A combined bottom-hole pressure calculation procedure using multiphase correlations and artificial neural network models", SPE Annual Technical Conference and Exhibition, Amsterdam, Netherlands, 2014.

[8] Luis A. Aguirre, Bruno O. S. Teixeira, Bruno H. G. Barbosa, Alex F. Teixeira, Mario C. Campos, and Eduardo M. Mendes, "Development of soft sensors for permanent downhole Gauges in deepwater oil wells", Control Engineering Practice, vol. 65, pp. 83-99, 2017.

[9] Mahshid Firouzi and Suren Rathnayake, "Prediction of the flowing bottom-hole pressure using advanced data analytics", Asia Pacific Unconventional Resources Technology Conference, Brisbane, Australia, 2019.

[10] Oluwatoyin Akinsete and Blessing Adesiji, "Bottom-Hole Pressure Estimation from Wellhead Data Using Artificial Neural Network", Nigeria Annual International Conference and Exhibition, Lagos, Nigeria, 2019.

[11] Fahad H. Al Shehri, Anton Gryzlov, Tayyar Al Tayyar, and Muhammad Arsalan, "Utilizing Machine Learning Methods to Estimate Flowing Bottom-Hole Pressure in Unconventional Gas Condensate Tight Sand Fractured Wells in Saudi Arabia", SPE Russian Petroleum Technology Conference, Virtual, 2020.

[12] Sherif Sanusi, Adenike Omisore, Eyituoyo Blankson, Chinedu Anyanwu, and Eremiokhale Obehi, "Estimation of bottom hole pressure in electrical submersible pump wells using machine learning technique", SPE Nigeria Annual International Conference and Exhibition, Lagos, Nigeria, 2021.

[13] Daniel Opoku, Ahmed Al-Ghamdi, and Albert Osei, "Novel method to estimate bottom hole pressure in multiphase flow using quasi-Monte Carlo method", International Petroleum Technology Conference, Dhahran, Kingdom of Saudi Arabia, 2020.

[14] M. M. Mukaka, "Statistics corner: A guide to appropriate use of correlation coefficient in medical research", Malawi Medical Journal, vol. 24, n. 3, pp. 69-71, 2012.

[15] Jason Brownlee, Machine learning mastery with Python: Understand your data, create accurate models and work projects end-to-end, v1.20 ed., Australia: Jason Brownlee, 2021.

[16] Aston Zhang, Zachary C. Lipton, Mu Li, Alexander J. Smola, Dive into Deep Learning, 2019.

[17] S. Russell and P. Norvig, Artificial Intelligence: a modern approach, 4th ed., Pearson, 2021.

[18] Jugal K. Kalita, Dhruba K. Bhattacharyya, and Swarup Roy, Fundamentals of Data Science. Cambridge: Academic Press, 2024, pp. 69-89.

[19] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, vol. 9, pp. 1735–1780, 1997.

[20] Eryk Lewinson. "Choosing the correct error metric: MAPE vs. sMAPE", Towards data science, 2020. [Online]. Available: https://towardsdatascience.com/choosing-the-correct-error-metric-mape-vs-smape-5328dec53fac

[21] T. Rezende, "RMSE or MAE? How to evaluate my machine learning model?", LinkedIn, 2018. [Online]. Available: https://pt.linkedin.com/pulse/rmse-ou-mae-como-avaliar-meu-modelo-de-machine-learning-rezende

[22] R. A. Rubo, C. C. Carneiro, M. F. Michelon, and R. S. Gioria, "Digital petrography: Mineralogy and porosity identification using machine learning algorithms in petrographic thin section images", Journal of Petroleum Science and Engineering, vol. 183, pp. 106382, 2019.