Rio Oil & Gas Expo and Conference 2022

ISSN 2525-7579

Conference Proceedings homepage: https://biblioteca.ibp.org.br/riooilegas

Technical Paper

Oil production and pressure multimodal forecasting integrating high-frequency production data

Leopoldo André Dutra Lusquino Filho [1]
Rafael de Oliveira Werneck [2]
Pedro Ribeiro Mendes Júnior [3]
Manuel Castro [4]
Eduardo dos Santos Pereira Eduardo Pereira [5]
Renato Moura [6]
Vitor Hugo de Sousa Ferreira [7]
Alexandre Mello Ferreira [8]
Alessandra Davólio Gomes [9]
Anderson de Rezende Rocha [10].

1. UNICAMP, IC , RECOD.AI. CAMPINAS - SP - BRASIL, leopoldolusquino@gmail.com
2. UNICAMP, IC, RECOD.AI. CAMPINAS - SP - BRASIL, rafael.werneck@ic.unicamp.br
3. UNICAMP, IC, RECOD.AI. CAMPINAS - SP - BRASIL, pedro.junior@ic.unicamp.br
4. UNICAMP, IC, RECOD.AI. CAMPINAS - SP - BRASIL, castroavila@ic.unicamp.br
5. UNICAMP, IC, RECOD.AI. CAMPINAS - SP - BRASIL, eduardo.santos@inpe.br
6. UNICAMP, IC, RECOD.AI. CAMPINAS - SP - BRASIL, rlmoura@unicamp.br
7. UNICAMP, IC, RECOD.AI. CAMPINAS - SP - BRASIL, vhsf@unicamp.br
8. UNICAMP, IC, RECOD.AI. CAMPINAS - SP - BRASIL, melloferreira@ic.unicamp.br
9. UNICAMP, IC, RECOD.AI. CAMPINAS - SP - BRASIL, davolio@unicamp.br
10. UNICAMP, IC, RECOD.AI. CAMPINAS - SP - BRASIL, anderson.rocha@ic.unicamp.br

Abstract

Production forecasts (oil, gas, water) are important for the management of petroleum reservoirs. Previous results in the literature indicate the effectiveness of using machine learning models in short-term predictive analysis; however, such models are dependent on a large amount of input data. For example, rates are measured daily in the Brazilian Pre-salt fields. Some measures are high-frequency data (wellhead pressure and temperature, for instance), thus requiring additional treatment so that such time series can be used together by the same forecasting model. However, combining the daily series with an interpolation of the high-frequency series can prevent ML models from learning relevant patterns. This work proposes a multimodal approach in which high-frequency and daily series are treated separately. These series can be in different granularities and the model can take advantage of their particularities. This strategy was validated for a 30-day forward forecast in real Brazilian Pre-salt wells, and the results obtained indicated that the use of multimodal learning brought an improvement in oil production and pressure forecasting.

Keywords: Multimodal forecasting. high-frequency data. machine learning. oil production. pre-salt

Received: October 05, 2021 | **Accepted:** August 25, 2022 | **Available online:** September 26, 2022
Article nº: 308
Cite as: Proceedings of the Rio Oil & Gas Expo and Conference, Rio de Janeiro, RJ, Brazil, 2022.
DOI: https://doi.org/10.48072/2525-7579.rog.2022.308

1. Introduction

Forecasting production accurately is essential for managing hydrocarbon fields effectively over their lifetime. Reservoir simulation models are traditionally used for long-term forecasting, but the work is very time consuming, especially in large and complex reservoirs, as demonstrated by Ertekin and Sun (2019). A different and valuable approach for establishing short-term forecast projections for reservoirs is to leverage machine learning techniques, as explored by Liu, Liu, and Gu (2020) and Sun, Ma, and Kazi (2018). Previous investigations performed by Werneck et al. (2022) showed that Recurrent Neural Networks (RNNs) with a sliding window mechanism is an approach with potential for forecasting oil production and Bottom-Hole Pressure (BHP). As these solutions are data-driven, it is crucial to look for methods that make the most of all the data collected by the different sensors. However, specific sensors operate at distinct sampling rates: some measurements are taken daily (such as oil, water, and gas fluid productions, and BHP) and others are taken on the minute scale (such as temperature and choked pressure).

To address this issue, we investigate a multimodal approach for handling each high-frequency (HF) series (time series with a periodicity less than daily) separately and report validation findings in an actual pre-salt reservoir. Outside the oil industry, this approach has been validated in different deep-learning architectures based on Long Short Term Memory (LSTM) networks, introduced by Hochreiter and Schmidhuber (1997), and Gated Recurrent Unit (GRU) networks, introduced by Cho, van Merrienboer, Bahdanau and Bengio (2014), expressly indicated to capture temporal information.

Some of the benefits of using multimodal learning in this task are: (i) it allows different attention mechanisms for each HF series; (ii) it enables independent HF measurements, leading to models that are more robust to noise in some of the sensors; (iii) it allows the selective treatment of HF data, rendering less weight to weak modalities or ignoring them altogether.

This paper is organized as follows: Section 2 describes related work on mixed-frequency time series forecasting, multimodal machine learning, and multi-output forecast validation setup; Section 3 presents the multimodal architectures being proposed in this work; Section 4 contains a description of the experiments we carried out, as well as their analyses; and Section 5 states the conclusions and future work.

2. Related work

This section presents the related work on mixed-frequency time series, multimodal learning, and how to approach multi-output forecasting.

2.1. Mixed-frequency time series

Most proposed solutions in the literature flatten all series at the same frequency, typically bringing the HF series to the same granularity as the lower-frequency series, thus discarding potentially important information. However, Ghysels and Valkanov (2006) and Hyung and Granger (2008) showed that it is generally advantageous for forecast accuracy to keep the series with mixed-frequency. Two approaches based on statistical modeling can be found in the literature: the multivariate, continuous, autoregressive, moving average (VARMA) model-based approach, proposed by Zadrosny (1998), and Mixed Data Sampling (MIDAS), proposed by Ghysels, Santa-Clara, and Valkanov (2004).

The first solution assumes that the model operates at the highest frequency of the data and that all data are generated but not necessarily observed at this frequency. The second approach specifies conditional expectations as a distributed lag of regressors at higher sampling frequencies, using a weighting function. Wohlrabe (2009) carried out the first literature review of the joint use of time series of mixed frequencies and performed a thorough comparison of the VARMA-based approach and MIDAS in various scenarios related to German gross domestic product forecasting.

More recently, domain-agnostic machine learning solutions that are less dependent on a priori knowledge than statistical modeling have been proposed. Xu, Zhuo, Jiang, and Liu (2018) propose a MIDAS-based neural network. Xu, Liu, Jiang, and Zhuo (2021) extend this model using a quantile regression neural network. Toda, Moriwaki, and Ota (2021) employ aggregate learning to combine different-frequency series into the same model.

2.2 Multimodal learning

Multimodal learning [Baltrušaitis, Ahuja, and Morency (2018)] is a Machine Learning (ML) technique that mixes data from many sources, frequently combining numerous domains, e.g., pictures, text, or tabular data. All these data have various statistical features and a joint representation of such modalities is required. Multimodal learning can also potentially supply missing modalities in noisy data based on the other modalities, making models that use it more robust.

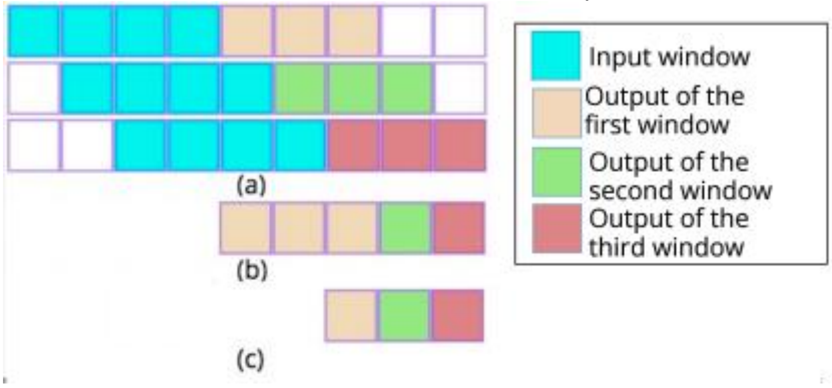
We find two main types of multimodal representations in the literature: joint and coordinated. Joint representations project all inputs onto the same space and some of the approaches to accomplish this include RNNs encoder-decoder frameworks, introduced by Bahdanau, Cho, and Bengio (2014), probabilistic graphical models, as proposed by Bengio, Courville, and Vincent (2013), and autoencoders, as explored by Silberer and Lapata (2014). Coordinated representations, in turn, exist in their own space. However, they are coordinated through a similarity or structure constraint, typically using similarity models, as seen in Weston, Bengio, and Usunier (2010), or a structured, coordinated space, as used by Wang, Shen, Song, and Ji (2014).

One of the main topics of interest in multimodal machine learning is multimodal data fusion, explored in three surveys in this area: D'mello and Kory (2015), Zeng, Pantic, Roisman, and Huan (2009), and Baltrušaitis, Ahuja, and Morency (2017). The three primary forms of multimodal data fusion are early, late, and hybrid fusion. Early fusion creates a joint representation of the multimodal data at some point in the processing prior to the classification or regression itself, usually through simple concatenation or a bag approach, as investigated by Werneck, Dourado, Fadel, Tabbone, and Torres (2018). Late fusion uses systems composed of different models, each of which deals with a single modality, and then merges the inferences from those models using averaging, as seen in Shutova, Kelia, and Maillard (2016), or voting schemes, as performed in Morvant, Habrard, and Ayache (2014), or even using another model for this, as suggested by Glodek (2011). Finally, hybrid fusion combines late fusion mechanisms inferences performed by early fusion models and unimodal models.

2.3. N-th Day

Werneck et al. (2022) propose two validation paths that aim to be more realistic for multi-output forecasting than those traditionally found in the literature, while avoiding mixing different prediction confidences for long forecasts. The first approach is called First Prediction and its forecast considers the entire output of the first predicted window followed by only the last point of the following predictions. The second approach is called N-th Day and uses the same multi-output sliding window as the previous approach but only considers the last point of each output in its forecast for performance evaluation. Figure 1 details the final forecast of the same model using the First Prediction and N-th-Day setups, with a sliding window of four days of input to predict the three following days.

Figure 1 –- Forecast setup: (a) multi-output forecast with sliding window using four points to predict three ahead;
(b) First Prediction; (c) N-th Day



Source: Authors

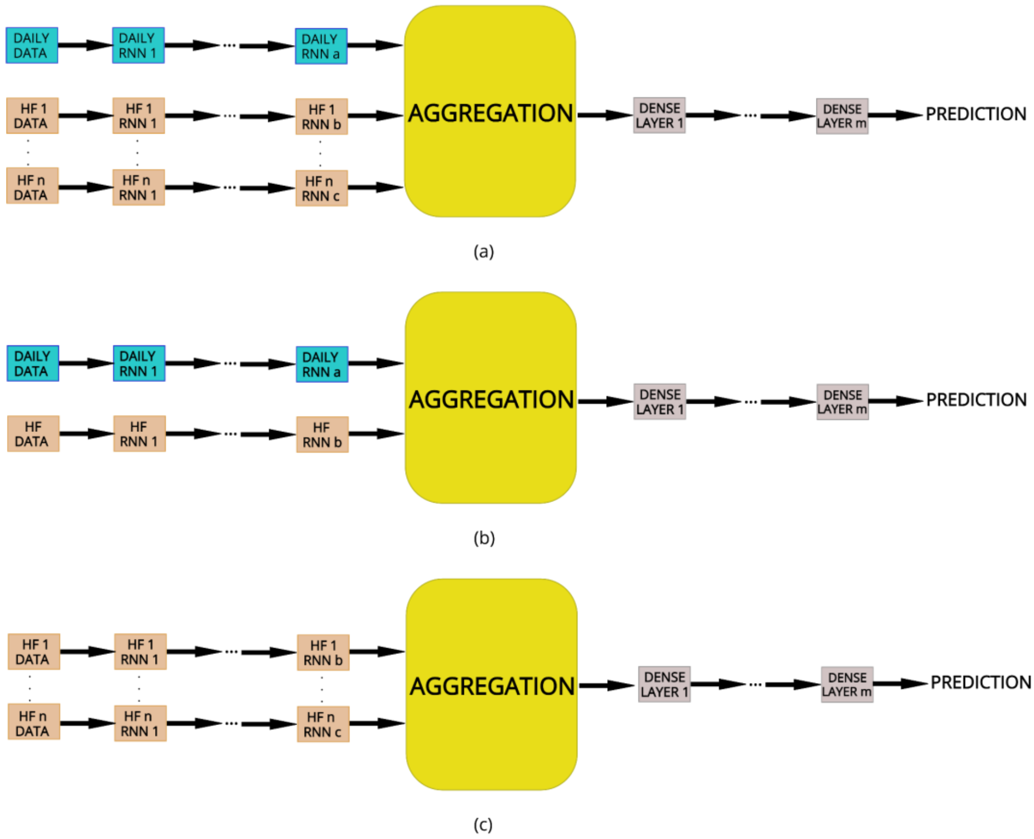
3. Proposed approaches

The problem being addressed in this research is to create an approach to forecasting where the HF data can be used alongside daily data without the need to change the granularity of the data; that is, we seek an ML-underpinned solution that uses mixed-frequency time series for forecasting. The proposed method is based on the notion that mixed-frequencies time series have different distributions, each of which can be viewed as a modality. We propose deep neural networks with early fusion through a combination layer to perform this multimodal learning. As each HF series can have a different periodicity, each one is considered a separate modality, while all features that make up the daily series form a single modality. Figure 2.a depicts this multimodal approach.

Each HF series and the daily series enter the deep network through a separate initial layer and run through a series of stacked recurrent neural networks. The RNNs' outputs are then concatenated and sent through a series of dense layers, which performs forecasting. Given that this model is used here with supervised training, in the training stage, the labels are associated with the input data of the daily series, and in the concatenation step, the data of each HF series are distributed equally, according to the sequential order of daily data. Different granularities may be applied in each HF pipeline without affecting the training samples' correlation and labels. Also, the same aggregation ratio between the HF data and the daily data applied in training is preserved in the test stage.

Another architecture is proposed for cases in which all HF data are on the same frequency (Figure 2.b), with the model having only two modalities: one for daily data and another for HF data. This architecture can be used whenever the amount of high periodicity HF data compromises the computational cost of training and makes the model time consuming. Therefore, the HF data can be interpolated to a lower granularity, though still higher than daily.

Figure 2 –- Multimodal forecasting model for high-frequency data: (a) using HF data separately and daily data; (b) using HF together in the same modality and daily data; (c) using HF data separately.



Source: Authors

Figure 2.c shows a third variation of the architecture, in which only the HF data are used, each one being a separate modality, which may have different granularities. The target used to train this architecture is still associated with the daily data and their quantity is used to guide the process of merging the data in the concatenation layer. This approach can be interesting in specific domains where the daily series features lack much influence or are even noisy compared to the target to be predicted.

4. Experiments and results

This section describes the dataset used, the performed experiments, the validation used, and an analysis of the results obtained.

4.1. Dataset

The dataset used in this research is production data from a Brazilian pre-salt oil field with 1,269 days of historical data. Four wells (P1–P4) were selected for this validation; all wells share the same features as the daily series: oil, water, and gas fluid productions, BHP, water cut, gas-oil ratio, and gas-liquid ratio values. HF data vary from well to well in type and quantity but are related to temperature and pressure. For example, P3 has 10 HF features, whereas the others have 18. Both daily oil production and daily BHP were the prediction targets for all wells. Table 1 provides more details on the wells.

Table 1 – Wells used in the validation with range, variance, and standard deviation of the training and test data.

	Well	Training set			Test set		
		Range	**Var**	**Std**	**Range**	**Var**	**Std**
Oil	**P1**	0 - 5996.86	1322260.22	1149.90	0 - 4935.68	518923.74	720.36
	P2	0 - 5842.70	1136987.36	1066.30	0 - 5568.66	830724.06	911.44
	P3	0 - 5864.88	1468909.94	1221.99	0 - 4378.30	843377.17	918.36
	P4	0 - 5246.40	4927390.62	2219.77	0 - 5012.36	748220.22	865.00
BHP	**P1**	0 - 526.79	6486.96	80.54	0 - 502.69	3712.67	60.93
	P2	0 - 533.42	7967.28	89.26	0 - 509.09	1988.76	44.60
	P3	0 - 531.24	6477.82	80.48	0 - 492.72	672.11	25.93
	P4	0 - 495.62	54396.66	233.23	0 - 486.05	606.2	24.62

Source: Authors

4.2. Validation

To provide a more precise model assessment in short-term forecasting, a multi-output window with sliding steps from the past 85 days (input) was used to predict the following 30 days (output) with the N-th Day setup. These parameters took Werneck et al. (2022) as a reference. Finally, the predictions were validated using the SMAPE (Symmetric Mean Absolute Percentage Error) metric, which measures the percentage error of the predicted values according to Equation 1.

$$SMAPE(y,h) = \frac{100}{n}\sum_{i=1}^{n} \frac{|h(x_i) - y_i|}{\frac{(|y_i| + |h(x_i)|)}{2}} \quad (1),$$

where $h(x)$ is the forecast value, y is the actual value, and n is the total number of observations made.

As the periodicity of these HF data is low, they were interpolated through the median for the hourly frequency, despite the multimodal architecture proposed in Section 3, which can deal with a series of any frequencies. This option was adopted to prevent models from becoming time consuming. Data were also interpolated through the median and autoencoder (AE) for daily frequency for testing purposes. The configuration used for the AE was: one decoding and one qualifying, the loss being the Mean Absolute Error (MAE), using the Adam optimizer and ten epochs in training. The data frame had its dimensionality reduced through AE, and its transposition was used as input for the HF modes of the network.

This approach seems counter-intuitive since the HF data could be applied in a unimodal model when interpolated to daily frequency. However, an advantage of applying them in the multimodal model is that each series will be processed by a different pipeline, making the model more robust to noise coming from a specific HF series. All architectures proposed in Section 3 were used in this validation.

A GRU-based architecture (GRU2) that only uses the daily series was applied as a baseline. This architecture was introduced by Werneck et al. (2022), which was validated in the same daily dataset of this work (i.e., no HF data) in another historical period and other forecast datasets. In addition, multimodal architectures use LSTM, GRU, or both, and convolutional layers preceding the RNNs, and attention mechanisms were also tested. All experiments were performed with 100 epochs, using an early stopping approach after ten epochs without improving the validation loss.

4.3. Results and analysis

Table 2 shows the benchmark between the results of GRU2 and the best models for each multimodal approach. The results are obtained from an average of 10 runs. Figures 3–5 show forecasting plots for the best setup with and without HF data. Due to space limitations, we only show some results here. In these images, the purple vertical band marks the beginning of the test data, the blue dots are ground-truth data, the red dots are the predicted values, and the green region represents the range of predicted values considering the ten runs to understand the forecast variability of that model for the target well represented in the figure.

The results indicate that the best approach depends on the well and the target but that, in general, the use of HF interpolated data for daily granularity with autoencoder was the best in most scenarios and was competitive even in scenarios where it was not the best. For the oil production forecast, the multimodal solution was superior in SMAPE in wells P2 and P4 and tied in well P3. Concerning BHP, the multimodal approach had an advantage in wells P3 and P4, with a tie in P1. The unimodal solution was consistently better when the variance and standard deviation of the test dataset were smaller, as is shown in Table 1. In the oil production forecast of well P1, the unimodal solution outperformed architectures that use HF in SMAPE terms, but HF Bimodal approached the curve better, with more sensitivity to its oscillations.

The hourly granularity approach in the HF data (HF Hourly) was competitive with the daily granularity (HF Daily Median and HF Daily AE) in the oil production forecast, with the best performance in well P2. However, it was worse in the BHP forecast, where both ranges of the data of training and testing were smaller. The Only HF approach proved better in scenarios with high data variance. Though the HF Bimodal approach did not have the best performance in any target, it proved a competitive alternative in almost all scenarios. The only exception is the forecast of oil production in well P3. The bimodal network also proved the most stable approach and is an interesting alternative for wells whose behavior is not fully known.

Regarding the architectures themselves, LSTM-based models benefited more from attention mechanisms and median interpolation and were more efficient for forecasting oil production. Meanwhile, the GRU-based models had a more favorable impact from convolutional layers and AE interpolation. They had better performance in the BHP forecast. The use of batch normalization after the concatenation layer generally had a positive impact, mainly when hourly granularity was used.

Table 2 – SMAPE benchmarking of oil production and bottom-hole pressure forecasting for the unimodal baseline (GRU2) and the best multimodal architectures.

		GRU2 (baseline)	HF Daily Median	HF Daily AE	HF Hourly	Only HF	HF Bimodal
	P1	**8.7 ± 0.10**	10.9 ± 0.27	10.3 ± 0.16	10.7 ± 0.10	19.0 ± 0.62	9.7 ± 0.17

Oil	**P2**	20.4 ± 0.35	27.2 ± 2.50	28.2 ± 2.54	**15.2 ± 0.47**	17.2 ± 1.21	17.0 ± 1.19
	P3	**27.9 ± 1.44**	28.2 ± 0.66	**27.1 ± 1.57**	30.4 ± 0.53	31.9 ± 0.49	36.3 ± 1.75
	P4	14.4 ± 0.25	12.1 ± 0.32	**11.7 ± 0.18**	12.0 ± 0.11	13.0 ± 0.20	11.9 ± 0.15
BHP	**P1**	**0.8 ± 0.02**	**0.8 ± 0.03**	1.2 ± 0.10	2.5 ± 0.08	2.4 ± 0.95	1.3 ± 0.05
	P2	**0.4 ± 0.01**	0.7 ± 0.02	0.9 ± 0.07	3.5 ± 0.21	4.8 ± 0.13	0.7 ± 0.03
	P3	3.6 ± 0.07	**2.6 ± 0.06**	3.3 ± 0.36	7.7 ± 0.05	8.0 ± 0.29	3.0 ± 0.10
	P4	0.9 ± 0.06	1.3 ± 0.16	**0.7 ± 0.10**	1.2 ± 0.09	1.4 ± 0.02	0.9 ± 0.04

Source: Authors

5. Conclusion

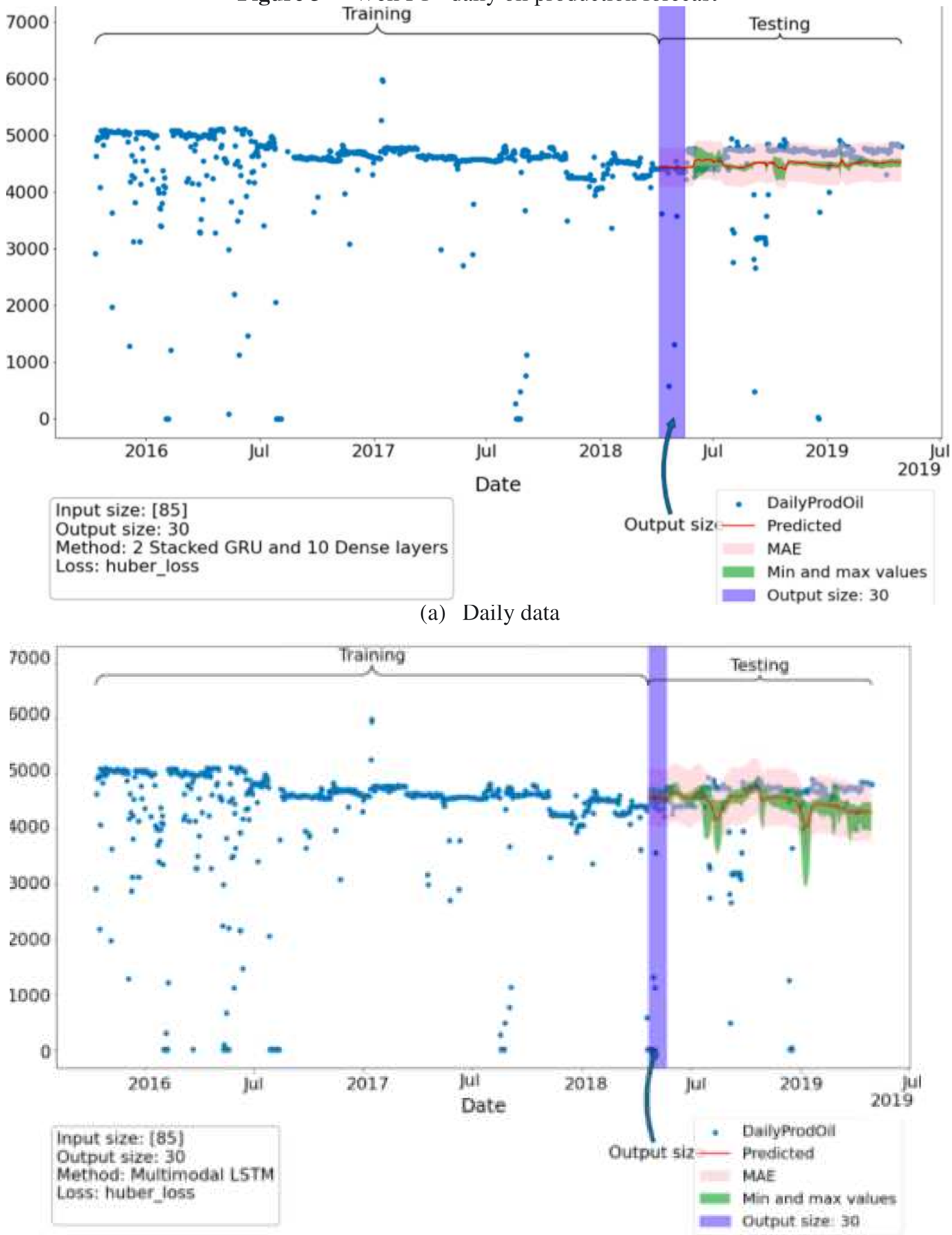
In this paper, we propose using deep multimodal learning to integrate the HF production data with the daily production data. The main contribution is investigating HF data in oil production and pressure forecasting solutions using ML techniques, which has received little attention in the literature to date. Its originality is the use of multimodal learning in forecasting with HF data, which can be used in various fields. Different multimodal architectures were tested in four pre-salt wells and the use of HF data brought gains for the accuracy of the production forecast of half of the validated scenarios.

Our experimental findings corroborated investigations carried out by Ghysels and Valkanov (2006) and Hyung and Granger (2008) in other domains, demonstrating that the use of mixed-frequency time series is advantageous in many short-term forecasting scenarios, as it is better able to approximate the curve oscillations of the ground-truth values even in some cases where the use of daily data only obtains a better predictive accuracy. The results also indicate the potential for using multimodal ML as an approach to forecasting mixed-frequency time series, so we encourage its use in other domains. Because the multimodal architectures proposed here allow handling different time series separately, providing robustness in cases of noise present in the original HF data or generated by its interpolation, this approach is also useful in scenarios where the HF data must be interpolated to a higher granularity because its large quantity makes using all of the data unfeasible in terms of time consumption. However, using solely HF data was found to be disadvantageous in almost all of the circumstances investigated in this study.

Our results showed that the appropriate granularity for the HF data and the best architecture for the multimodal networks depend on the dataset and the target, but interpolation for daily granularity with auto-encoders performed better in most cases. The main strength of the multimodal approach was in cases where there was greater variance in the test dataset. Although multimodal forecasting with HF data has shown to be a promising approach, we need to better understand its idiosyncrasies and how this solution relates to others found in the literature. For this, one must explore it in other domains, such as a comparison with the statistical models VARMA and MIDAS.

The investigations demonstrate the potential for forecasting using HF data and how multimodal architectures can efficiently deal with mixed-frequency time series. Future work includes using early and hybrid fusion-based approaches, using LSTM and GRU-based architectures simultaneously, testing the use of data augmentation on daily series data when used in conjunction with the HF data, and testing the use of the HF data in its original granularity even at the cost of much longer processing time.

Figure 3 — Well P1 - daily oil production forecast

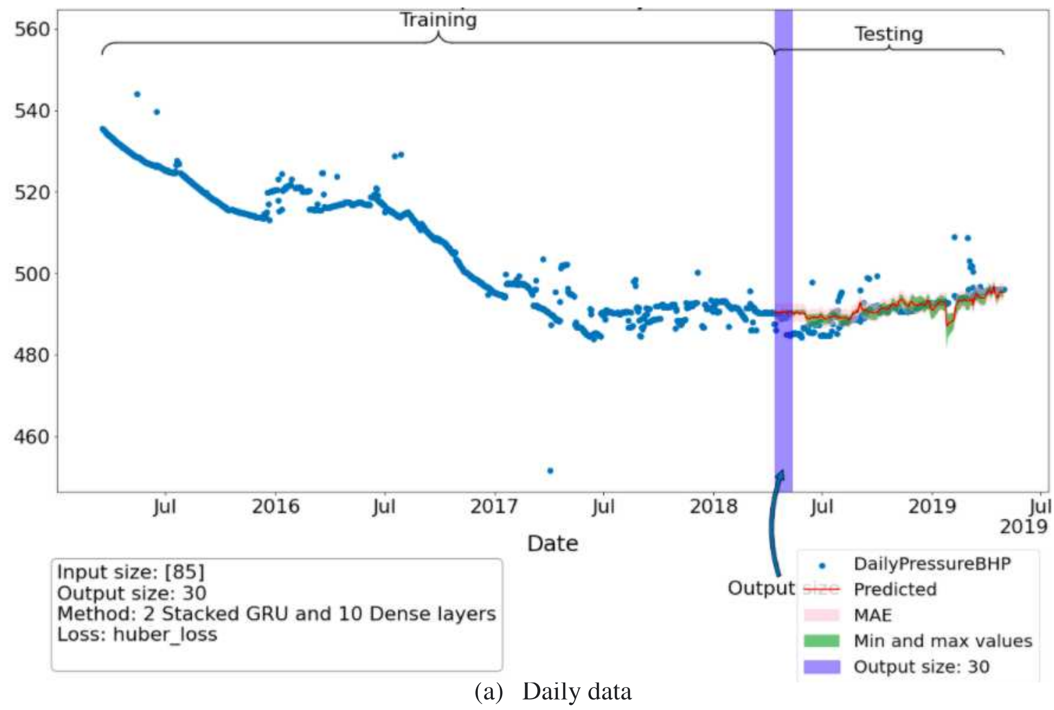


(a) Daily data

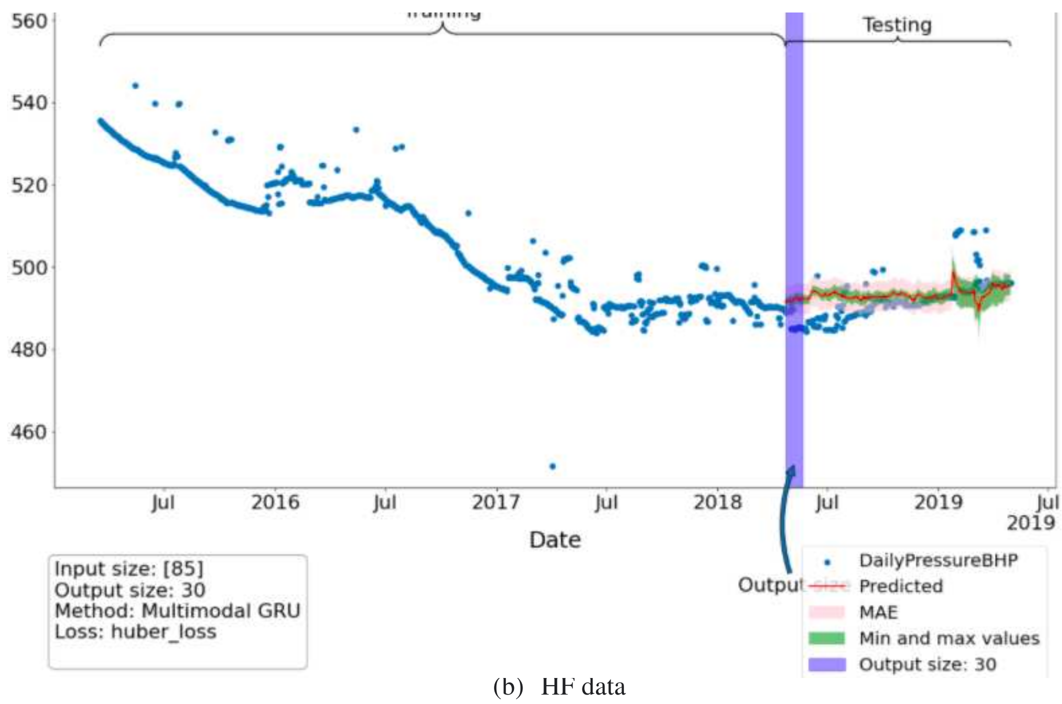
(b) HF data

Source: Authors

Figure 4 — Well P2 - daily BHP forecast



(a) Daily data



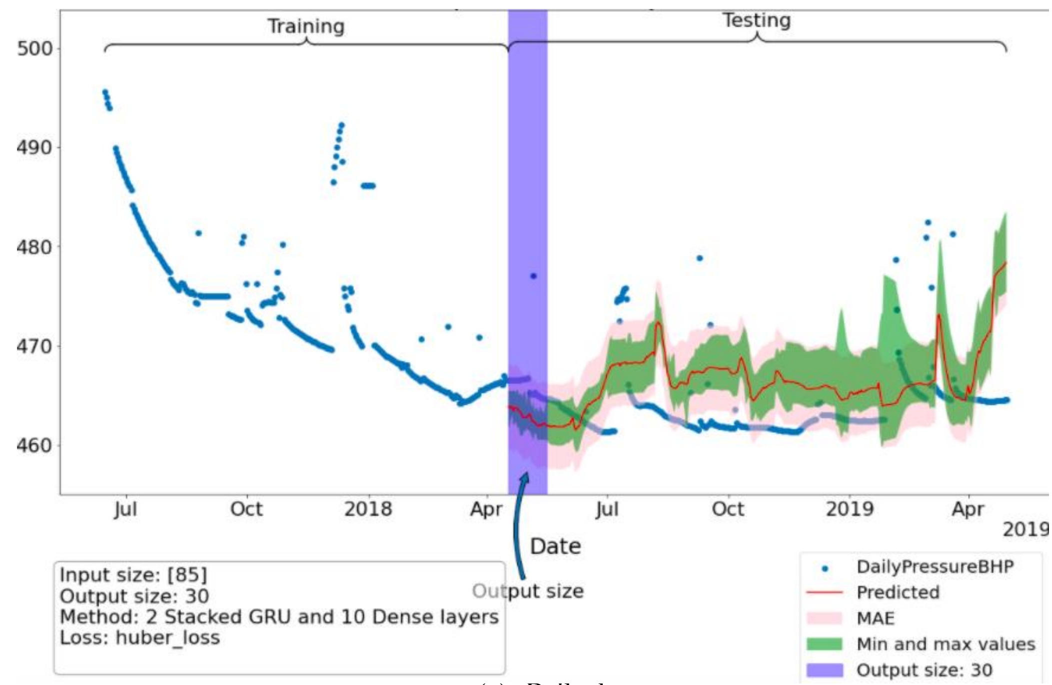
(b) HF data

Source: Authors

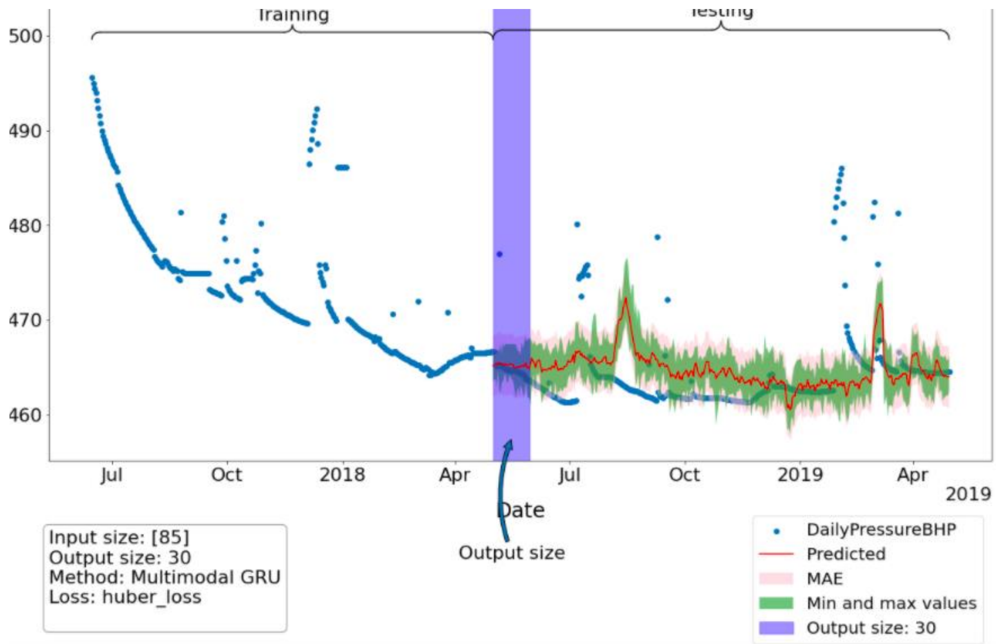
Acknowledgments

This research was conducted in association with the ongoing Project registered under ANP number 21373-6 as "Desenvolvimento de Técnicas de Aprendizado de Máquina para Análise de Dados Complexos de Produção de um Campo do Pre-Sal" (UNICAMP/Shell Brazil/ANP) - "Machine-Learning Development for Analysis of Complex Production Data in a Pre-Salt Carbonate Field" - funded by Shell Brazil, under the ANP R&D Levy as "Compromisso de Investimentos com Pesquisa e Desenvolvimento". The authors thank Schlumberger and CMG for software licenses.

Figure 5 — Well P4 - daily BHP forecast



(a) Daily data



(b) HF data

Source: Authors

Referências

Bahdanau, D., Cho, K., & Bengio, Y. (2014). *Neural machine translation by jointly learning to align and translate* arXiV. https://arxiv.org/abs/1409.0473

Baltrušaitis, T., Ahuja, C., & Morency, L. (2018). Multimodal machine learning: A survey and taxonomy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *41*(2), 423–443. https://doi.org/10.1109/TPAMI.2018.2798607

Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(8), 1798–1828. https://doi.org/10.1109/TPAMI.2013.50

Cho, K., Merrienboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). *Learning phrase representations using RNN encoder-decoder for statistical machine translation*. 1724–1734. https://aclanthology.org/D14-1179/

D'Mello, S.K., & Kory, J. (2015). A Review and Meta-Analysis of Multimodal Affect Detection Systems. *ACM Computing Surveys*, *47*(3), 1–36. https://doi.org/10.1145/2682899

Ertekin, T., & Sun, Q. (2019). Artificial Intelligence Applications in Reservoir Engineering: A Status Check. *Energies*, *12*(15), 2897–2919. https://doi.org/10.3390/en12152897

Ghysels, E., Santa-Clara, P., & Valkanov, R. (2004). *The MIDAS touch: MIxed DAta Sampling regression models*. UCLA: Finance. https://escholarship.org/uc/item/9mf223rs

Ghysels, E., & Valkanov, R. (2006). *Linear time series processes with mixed data Sampling and MIDAS regressions models*. mimeo. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=920610

Glodek, M., Tschechne, S., Layher, G., Schels, M., Brosch, T., Scherer, S., & Kächele, M. (2011) *Multiple Classifier Systems for the Classification of Audio-Visual Emotional States*. *6975*, 359–368. http://dblp.uni-trier.de/db/conf/acii/acii2011-2.html#GlodekTLSBSKSNPS11

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Hyung, N., & Granger, C. (2008). Linking series generated at different frequencies. *Journal of Forecasting*, *27*(2), 95–108. https://doi.org/10.1002/for.1042

Liu, J., Liu, J., & Gu, J. (2020). Forecasting oil production using ensemble empirical model decomposition based Long Short-Term Memory neural network. *Journal of Petroleum Science and Engineering*, *189*(2), 1–11. https://doi.org/10.1016/j.petrol.2020.107013

Morvant, E., Habrard, A., & Ayache, S. (2014). *Majority vote of diverse classifiers for late fusion*. *1*, 20–30. https://hal.archives-ouvertes.fr/hal-00985839v2/document

Shutova, E., Kelia, D., & Maillard, J. (2016). *Black Holes and White Rabbits : Metaphor Identification with Visual Features*. *Proceedings of the 2016*, 160–170. https://aclanthology.org/N16-1020/

Silberer, C., & Lapata, M. (2014). *Learning grounded meaning representations with autoencoders*. *1*, 721–732. https://aclanthology.org/P14-1068/

Sun, J., Ma, X., & Kazi, M. (2018). *Comparison of Decline Curve Analysis DCA with Recursive Neural Networks RNN for Production Forecast of Multiple Wells*. *1*, 1–11. https://onepetro.org/SPEWRM/proceedings-abstract/18WRM/4-18WRM/D041S012R009/215403

Toda, T., Moriwaki, D., & Ota, K. (2021). *Aggregate Learning for Mixed Frequency Data*. arXiV. https://arxiv.org/abs/2105.09579

Wang, J., Shen, H.T., Song, J., & Ji, J. (2014). *Hashing for similarity search: A survey*. arXiV. https://arxiv.org/abs/1408.2927

Weston, J., Bengio, S., & Usunier, N. (2010). *Web Scale Image Annotation: Learning to Rank with Joint Word-Image Embeddings*. *81*, 21–35. https://doi.org/10.1007/s10994-010-5198-3

Wohlrabe, K. (2009). *Forecasting with Mixed-frequency Time Series Models* [Doctoral dissertation, Ludwig Maximilian University of Munich]. https://d-nb.info/993272096/34

Xu, Q., Liu, S., Jiang, C., & Zhuo, X. (2021). QRNN-MIDAS: A novel quantile regression neural network for mixed sampling frequency data. *Neurocomputing*, *457*(2021), 84–105. https://doi.org/10.1016/j.neucom.2021.06.006

Zadrozny, P.A. (1988). Gaussian-Likelihood of countinuous-time ARMAX models when data are strocks and flows at different frequencies. *Econometric Theory*, *4*(1), 108–124. https://doi.org/10.1017/S0266466600011890

Zeng, Z., Pantic, M., Roisman, G.I., & Huang, T.S. (2009). A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *31*(1), 39–58. https://doi.org/10.1109/TPAMI.2008.52