An innovative model to mitigate the impact of oil and steel price dynamics on the oil & gas sector projects

Aguinaldo Júnio Flor[1], **Luis França**[1]

[1]CIn - Universidade Federal de Pernambuco (UFPE)
Av. Jornalista Aníbal Fernandes, s/n, Cidade Universitária, Recife, PE, Brasil, 50.740-560

{ajf2,lvlf}@cin.ufpe.br

Abstract. *This paper addresses the development and application of an innovative model to analyze the historical price series of commodities, significantly impacting the profitability of Brazil's oil and gas projects. The experiment focuses on six historical price series of commodities critical to significant oil and gas exploration companies. It highlights the volatility of steel prices in the Brazilian and international markets and their direct impact on the key suppliers and explorers in the sector. The research introduces an advanced model, employing Deep Learning techniques with automated hyperparameters to optimize the selection of the most effective model for each dataset. This selection is based on a score of seven distinct metrics, ensuring the choice of the most suitable model to predict market trends relevant to the Oil and Gas sector.*

1. Introduction

Brazil's Oil and Gas sector offers significant market-value projects. Despite this, the Brazilian Oil & Gas (O&G) sector is relatively archaic in adopting Artificial Intelligence (AI), Data Mining, and Machine Learning techniques for project financial risk analysis. Such tools can provide reassurance to critical decision-makers concerning financial risk mitigation. Brent crude oil is globally recognized as a strategic product, thus significantly influencing the price dynamics of various other products. Price volatility significantly increases, contributes to inflationary pressures, and dictates the main challenges for economic growth in countries. Given the impact of oil on various supply chains, significant suppliers must understand the dynamics of oil price fluctuations along with the fluctuations and interdependence of other raw materials. A barrel of oil's current and future price guides investment volumes in oil exploration in the coming years.

The production of subsea products has an average timeline ranging from 12 to 18 months. Depending on the contract structure and the composition of these product packages, the total contract period can reach between 30 and 40 months of supply. In some instances, a contract involves several exploration wells, thereby determining the number of orders each contract will generate. In other words, the first link in the supply chain can be financially exposed to foreign exchange variations, commodity price changes, or inflationary indexes throughout the contractual period.

Companies producing subsea oil extraction systems must develop medium- to long-term strategies for market forecasting, raw material inventory management, and cash flow optimization to maintain financial stability, particularly in the Brazilian Oil & Gas sector, where fluctuations in oil barrel prices, such as WTI and Brent, are key indicators for future investment decisions [Foroutan and Lahmiri 2024, Mohsin and Jamaani 2023a]. These fluctuations also affect the volatility of commodities like metals and steel, essential for subsea systems production, while the predictability of oil prices is increasingly analyzed using hybrid time series models, combining classical statistical methods with deep learning techniques, as seen in recent studies focused on price curve dynamics and global price benchmarks, including OPEC [Cihan 2024, Maiti et al. 2024].

Deep Learning in financial risk analysis of projects applied to the O&G sector impacts companies' cash flow, risk, and return in this field. In this article, we use a product composed of approximately 38% carbon steel and stainless steel, which is highly exposed to the price variation of these commodities. Considering the variations in carbon and stainless-steel prices between 2020 and 2022, there was an impact of 7.22% on the raw material's total value.

The findings of this research underscore that, in sectors characterized by significant financial risk and complexity, artificial intelligence (AI) offers a powerful tool to support strategic and financial decision-making. In the Oil & Gas industry, Deep Learning has proven to be an effective and reliable method for addressing high-risk scenarios that involve potential loss of shareholder value. However, there is not an extensive body of literature focused on applying Deep Learning models to financial risk management in large industries, highlighting the need for further research in this area. Strengthening the connection between academia and industry in this sector could be a pivotal step toward modernizing how financial aspects are monitored and managed.

2. Methodology

We employed the DarTS library to develop our model, enhancing hyperparameter optimization using the Optuna tool. This was applied to time series data of Brent oil prices and Brazilian and imported steel prices, which are pivotal in the production systems of subsea products for oil and gas exploration. Furthermore, the project is designed to optimize the best model selection. This approach aims to identify the most effective technique for each dataset under analysis. This initiative stems from the goal of constructing a financial risk index based on steel price volatility relative to the contract values offered by an oil exploration company in bids for subsea products. A significant issue in Brazilian Oil & Gas (O&G) sector projects is the discrepancy between the procurement prices of raw materials, determined approximately six months to a year before contract signing, and the updated value of the contract post-signing. At certain market junctures, this price fluctuation between the initial procurement estimates and the acquisition of raw materials is not reflected in the updated contract value between the O&G operator and its supplier.

Our analysis, based on the results obtained from Deep Learning models for Brent oil and the prices of Brazilian and international carbon steel and stainless steel, along with a brief historical review of the impact of geopolitical factors on the steel time series, will enable us to propose a risk factor to be considered in the project's Profit & Loss (P&L) statement. Equation 1 represents the financial risk index, a weighted sum of each component multiplied by a specific weight. The formula adapts the parametric Value at Risk (VaR) framework [Jorion 2007].

$$FinancialRiskIndex = l_1 * w_1 + w_2 * l_2 + w_3 * l_3 \quad (1)$$

Where, l_1 will represent the risk of steel in Brazil; l_2 will represent the risk of international steel; l_3 will be the risk of Brent oil combined with the risk of price loss due to geopolitical factors. The weights are determined by w_x based on the assessment of the time series of oil and steel prices, as well as a historical evaluation of potential geopolitical impacts on the behavior of the curve.

3. Related Work

Machine Learning and Artificial Intelligence are technologies emerging in Corporate Finance and Financial Risk Analysis studies. To improve the efficacy of new studies in the financial field, Deep Learning models have been used for predicting financial disasters [Knuth 1984] and evaluating financial risk in various sectors of the economy: Electric Energy [Oreshkin et al. 2020] using models capable of projecting consumption, including the use of the DarTS library [Thomas and K.V. 2023], [Kazmi et al. 2023]; Medicine [Salehin et al. 2024], [Panja et al. 2023], [Orji and Ukwandu 2024], tourism, economy, retail, demography, among others [Athanasopoulos et al. 2024].

The importance of oil for the economic development of a nation generates a search for coverage against the uncertainty in the price of the commodity. In addition to the effect of the volatility of the price itself, there is also the high variation over time of the raw materials used by suppliers who are part of the first link in the chain of significant oil operators. From more traditional research, techniques in price prediction studies for oil include ARIMA, vector autoregressive models, Monte Carlo Simulation, among others. In the context of nonlinear scenarios, these models tend to perform poorly despite having good efficacy for handling linear and stationary time series.

Contemporary literature has shown a shift in studies on oil pricing from traditional econometric and statistical models to more advanced, nonlinear models with machine learning and Deep Learning techniques to capture the high volatility in oil price curves. This moment of change has brought about some studies by researchers [Ali Salamai 2023], [Yang et al. 2024], [Fang et al. 2023] who are now using artificial intelligence and deep learning techniques for oil price projection. This sector is relatively archaic but has been updated through new technologies. This includes explaining price fluctuations during COVID-19 [Xu et al. 2024] or discussing regional prices, such as in China [Guo et al. 2023].

Recent research using Deep Learning models, such as Convolutional Neural Networks - CNN [Mohsin and Jamaani 2023b], Temporal Fusion Transformer - TFT [He et al. 2023], and Recurrent Neural Networks - RNNs using LSTM and GRU [Sen and Dutta Choudhury 2024], [Wang et al. 2023], have datasets with daily closing values of the oil market in China and the United States. Although published between 2023 and 2024, the datasets are from periods up to 2020 - 2021, which may offer distortions with applications carried out between 2022 and 2024, a post-pandemic period with a cooling in the barrel price.

In addition to presenting an updated dataset, with the final date in December 2023, the correlation of carbon and stainless steel to the fluctuation curve of oil price linked to the geopolitical risk index GPR, our model covers the gap of research done with data from the Brazilian scenario, such as the price of carbon and stainless steel. Additionally, the novel experimentation with the Python DarTS library offers a variety of models, from classics like ARIMA to deep neural networks [Herzen et al. 2022].

4. Dataset

All the time series applied in the model are listed in Table 1. Each series is derived from real-world data, calculated monthly, and indexed to a base of 100.

The datasets underwent an individual preprocessing phase, accounting for observed seasonality and trends as well as the analysis of missing data or any additional irregularities.

After structuring the time series, the Dickey-Fuller (ADF) test was applied to check the stationarity of each variable. It was necessary first to transform the data using logarithms, which showed a high p-value. In the second phase, the technique of log differentiation by a twelve-month moving average was applied to make the series stationary. From this point, the experimental phase began.

Table 1. Statistical description of the data

#	Column	Period	count	mean	std	min	25%	50%	75%	max
0	oil	jan/82-nov/23	503	130.3	88.8	26.6	53.8	85.6	191.0	372.6
1	carbon_br	jan/96-dec/23	336	149.1	137.2	18.8	53.2	124.5	162.2	576.7
2	carbon_us	jun/82-dec/23	499	151.2	59.9	92.5	101.1	113.7	193.5	340.7
3	carbon_cru	apr/82-dec/23	357	152.2	57.3	68.9	100.9	154.7	177.2	355.4

5. Metrics

The accuracy of each model is measured using seven statistical indices calculated by the library itself: Mean Absolute Error (MAE), Mean Absolute Scaled Error (MASE), Mean Absolute Percentage Error (MAPE), Mean Squared Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Ranged Relative Error (MARRE), and Root Mean Squared Logarithmic Error (RMSLE). Additionally, we have developed a classification approach that considers the combination of all these metrics through a score, allowing us to automatically determine the best model among all those tested in the DarTS library. This automation in the developed model brings more clarity to the interpretation of results, offering a more concise structure to assist decision-makers based on the obtained outcomes. Having a model that demonstrates high levels of accuracy in predicting steel prices helps mitigate the imminent risk of financial loss in projects with a high cost of raw materials, with a concentration of 34% in a single commodity.

6. Experiment and Analysis

DarTS is a library for forecasting and anomaly detection in time series, comprising a wide array of models ranging from classic ones like ARIMA through regression models to neural networks for Deep Learning. The diversity of models makes the DarTS library a versatile tool that is applicable in various fields, such as the oil and gas sector. Despite being a recent library, its comprehensive documentation facilitates understanding the functioning of different models and the implementation of time series forecasting projects. This makes using this innovative library for forecasting and anomaly detection in time series particularly relevant. The importance of using DarTS is evident in this project, specifically in forecasting steel prices, where the library becomes a good choice due to its available models, which are crucial for predicting steel prices. This is evident from the results of experiments showing that the library's neural network-based models performed satisfactorily. The training set, which corresponds to 75% of the data, is used to fit the model, meaning the model "learns" relevant patterns from historical data. The testing set, representing the remaining 25%, is used to evaluate the model's ability to generalize to new, unseen data during training.

The initial models were those referenced in published articles mentioned in the library's documentation: RNN (Recurrent Neural Networks) as per article [Herzen et al. 2022] with Vanilla structure, LSTM (Long Short-Term Memory), and GRU (Gated Recurrent Unit), which are architectures of recurrent neural networks; NBeatsModel (Neural Basis Expansion Analysis for Time Series) as per article [Oreshkin et al. 2019], and finally, the Transformer model. Thus, we can understand the performance of PyTorch-based models compared to other classics for constructing a financial risk index in the procurement of Carbon and Stainless Steel for Oil and Gas industry projects.

According to article [Flunkert et al. 2017], DeepAR is a probabilistic forecasting method using autoregressive recurrent neural networks (RNNs) to learn a global model from historical data of all-time series. This enables DeepAR to capture complex relationships between time series and make more accurate predictions. DeepAR also provides quantile forecasts, which can estimate the likelihood of different future outcomes. DeepAR is suitable for this case as it can capture complex relationships between time series, which is crucial for steel price forecasting. Experiments applied vanilla, LSTM, and GRU RNNs.

On the other hand, article [Flunkert et al. 2017] discusses a forecasting method that uses neural basis functions to represent the underlying structure of a time series. This makes N-BEATS more interpretable than other forecasting methods, allowing users to see which factors are most important for predictions. N-BEATS can also handle long-term dependencies in time series data. It's useful for this case as it's more interpretable than other forecasting methods and crucial for understanding the causes of variations in steel prices. This article also provides insights into how the Transformer model can be used to learn relationships between different time series, like steel and oil prices.

Continuing with experiments, we implemented the TCNModel (Temporal Convolutional Network). The convolutional neural network is also used in the DarTS library for time series forecasting [Bai et al. 2018],[Chen et al. 2020]. This method aims to learn more complex patterns, such as seasonality and holiday effects, both within and in multivariate series. The architecture consists of stacked residual blocks based on dilated causal convolutional networks designed to capture the temporal dependencies of the series under analysis.

Another model used was N-HiTS (Neural Hierarchical Interpolation for Time Series) [Challu et al. 2022], which combines hierarchical interpolation techniques with multi-rate data sampling. Predictions are sequentially structured, emphasizing components presenting different frequencies and scales, decomposing the input signal, and synthesizing the forecast. The paper argues that N-HiTS achieves an average accuracy improvement of nearly 20% over Transformer architectures, with a 50-fold reduction in time.

Finally, the TiDE (Time-series Dense Encoder) was tested on the datasets. This model is an encoder-decoder based on a Multi-Layer Perceptron (MLP). Authors of the article [Das et al. 2024] argue that in more recent works, simple linear models can outperform Transformer-based approaches for long-term time series. They continue to discuss that TiDE can match or exceed previous approaches in popular benchmarks, being 5 to 10 times faster than Transformer-based models.

To achieve better results, hyperparameters were optimized using Optuna, an automatic hyperparameter optimization library for Python. This optimization was crucial for making the models more efficient, showing improved performance for each model. A total of 20 iterations were performed for each model, optimizing key parameters.

Thus, the experiment made it possible to evaluate the performance of different time series forecasting models for constructing a financial risk index in procuring carbon and stainless steel for the Oil and Gas industry. Analyzing the results, it is understood that, as expected, models based on deep neural networks showed superior performance to classic models, considering that the Transformer model had the best overall performance, and RNN-based models also performed well. On the other hand, the TCN and TiDE models showed inferior and below-expected performance.

In the comparative analysis of machine learning models on Table 2 for oil price forecasting, as indicated in our results table, the Transformer model demonstrates superior performance across multiple accuracy metrics. Specifically, it achieved the highest aggregated score of 54, reflecting its robustness in predicting oil prices. The evaluation metrics used include MAE, MSE, MAPE, RMSE, MASE, MARRE and RMSLE. The Transformer model consistently registered the lowest error values in MAE, MSE, and RMSE, among others, and secured the top or near-top ranking across the individual metrics. While other models such as Long Short-Term Memory (LSTM), TIDE, N-BEATS, Gated Recurrent Unit (RNN GRU), Temporal Convolutional Network (TCN), and Vanilla Recurrent Neural Network (RNN) were included in the study, none outperformed the Transformer in overall score. However, LSTM and TIDE showed competitive results. This finding suggests that the Transformer model, known for capturing complex temporal dependencies, is particularly effective for this forecasting task and may offer substantial predictive capabilities for stakeholders in the oil market sector.

Table 2. Comparison of different models based on various metrics and scores - oil.

oil	MAE	MSE	MAPE	RMSE	MASE	MARRE	RMSLE	Score
Transformer	**0.113 -8**	**0.022 -8**	**0.148 -8**	29.490 -7	**2.544 -8**	**13.860 -8**	0.098 -7	54
RNN LSTM	0.116 -3	0.022 -7	0.149 -7	**28.676 -8**	2.614 -3	14.241 -3	**0.098 -8**	39
NHiTS	0.113 -6	0.022 -5	0.150 -5	29.574 -4	2.560 -6	13.948 -6	0.099 -5	37
TiDE	0.113 -7	0.022 -4	0.150 -4	29.764 -2	2.556 -7	13.923 -7	0.099 -4	35
RNN GRU	0.116 -2	0.022 -6	0.149 -6	29.534 -5	2.622 -2	14.285 -2	0.098 -6	29
NBEATS	0.114 -5	0.023 -2	0.151 -2	29.506 -6	2.575 -5	14.027 -5	0.100 -2	27
TCN	0.115 -4	0.023 -3	0.150 -3	29.576 -3	2.593 -4	14.130 -4	0.099 -3	24
RNN Vanilla	0.118 -1	0.024 -1	0.154 -1	32.513 -1	2.664 -1	14.513 -1	0.103 -1	7

Regarding the Carbon US dataset, displayed in Table 3, the Transformer achieved the best overall performance with the highest total score of 46, indicating consistently high rankings across all metrics and locations. The RNN GRU had the weakest performance, with the lowest total score of 7. Generally, the models appear to have varied performances, with some (like the Transformer and RNN Vanilla) demonstrating robustness across multiple metrics and others (such as TiDE and RNN GRU) showing limitations in their predictive performance.

Table 3. Comparison of different models based on various metrics and scores - carbon_us.

carbon_cru	MAE	MSE	MAPE	RMSE	MASE	MARRE	RMSLE	Score
Transformer	**0.115 -8**	0.022 -5	0.149 -5	21.634 -7	**5.476 -8**	**17.163 -8**	0.093 -5	46
RNN Vanilla	0.116 -7	0.022 -6	0.148 -6	21.898 -6	5.525 -7	17.317 -7	0.093 -6	45
NBEATS	0.117 -6	**0.021 -8**	**0.144 -8**	24.284 -2	5.564 -6	17.44 -6	**0.092 -8**	44
RNN LSTM	0.119 -3	0.022 -7	0.147 -7	22.585 -4	5.635 -3	17.663 -3	0.092 -7	34
NHiTS	0.118 -4	0.023 -4	0.150 -4	22.217 -5	5.588 -4	17.513 -4	0.094 -4	29
TCN	0.117 -5	0.024 -2	0.155 -2	**20.862 -8**	5.568 -5	17.452 -5	0.096 -2	29
TiDE	0.119 -2	0.023 -3	0.150 -3	23.222 -3	5.678 -2	17.796 -2	0.095 -3	18
RNN GRU	0.132 -1	0.024 -1	0.156 -1	26.53 -1	6.254 -1	19.601 -1	0.099 -1	7

In Table 4, we find the results for the Carbon CRU dataset, where the Transformer model continues to lead the ranking, like previous datasets, with the highest total score of 50, reflecting a consistently strong performance across all metrics. On the other hand, the RNN GRU has the lowest total score of 16, suggesting it is the least effective model among those listed for this specific dataset. The results for the other models vary, with the RNN Vanilla and TCN positioning between the Transformer and RNN GRU, having total scores of 44 and 36, respectively. The N-HiTS, N-BEATS, and TiDE exhibit intermediate performances with total scores of 28, 23, and 21, respectively. These results indicate that, for the Carbon_CRU dataset, the Transformer can be considered the most robust model, while the RNN GRU shows significant limitations in its predictive capacity. The other models display a spectrum of effectiveness, with some offering reasonable performance and others showing potential for improvement.

Table 4. Comparison of different models based on various metrics and scores - carbon_cru.

carbon_cru	MAE	MSE	MAPE	RMSE	MASE	MARRE	RMSLE	Score
Transformer	**0,116 -8**	0,022 -6	0,149 -6	21,503 -7	**5,493 -8**	**17,216 -8**	0,093 -7	50
RNN Vanilla	0,117 -6	0,022 -7	0,149 -7	21,863 -6	5,539 -6	17,361 -6	0,093 -6	44
TCN	0,116 -7	0,023 -2	0,152 -2	**21,310 -8**	5,495 -7	17,222 -7	0,095 -3	36
RNN LSTM	0,119 -2	**0,022 -8**	**0,148 -8**	22,861 -4	5,664 -2	17,753 -2	**0,093 -8**	34
NHiTS	0,118 -4	0,023 -4	0,15 -4	22,950 -3	5,632 -4	17,652 -4	0,094 -5	28
NBEATS	0,118 -5	0,023 -1	0,153 -1	22,536 -5	5,62 -5	17,616 -5	0,096 -1	23
TiDE	0,119 -3	0,023 -3	0,151 -3	23,026 -2	5,652 -3	17,714 -3	0,095 -4	21
RNN GRU	0,123 -1	0,023 -5	0,150 -5	24,097 -1	5,841 -1	18,306 -1	0,095 -2	16

It is important to note that in addition to performing best on the crude oil price dataset, the Transformer model also excelled on carbon steel price datasets in the United

States and Global price, with Recurrent Neural Networks, whether vanilla or LSTM, being the second choice. This leads us to conclude that the short-term memories of LSTMs may be sufficient to capture the dynamics of Carbon curves. In the Brazilian market, however, the Transformer did not maintain its dominance in the first place, suggesting that the pricing dynamics in Brazil have a different structure.

Unlike American and global prices, Brazilian carbon steel had the NHiTS as the best model, as shown in Table 5, with RNN models again in second place. The dynamics of the Brazilian price curve are better accommodated by the convolutional network architecture of NHiTS than by the attention mechanisms of the Transformers. In any case, the Transformer model comes in second place, trailing only behind the RNN LSTM model.

Table 5. Comparison of different models based on various metrics and scores - carbon_br.

carbon_br	MAE	MSE	MAPE	RMSE	MASE	MARRE	RMSLE	Score
NHiTS	**0,149 -8**	0,048 -7	0,220 -7	35,265 -6	**3,935 -8**	**16,784 -8**	**0,141 -8**	52
RNN Vanilla	0,157 -7	0,064 -4	0,252 -4	**29,351 -8**	4,161 -7	17,75 -7	0,163 -4	41
Transformer	0,172 -4	0,048 -6	0,220 -6	46,763 -4	4,569 -4	19,488 -4	0,144 -7	35
RNN LSTM	0,163 -6	0,069 -3	0,262 -3	29,599 -7	4,324 -6	18,443 -6	0,171 -3	34
TiDE	0,175 -3	**0,048 -8**	**0,218 -8**	49,264 -3	4,638 -3	19,784 -3	0,144 -6	34
NBEATS	0,172 -5	0,049 -5	0,220 -5	49,614 -2	4,560 -5	19,448 -5	0,145 -5	32
RNN GRU	0,211 -2	0,091 -2	0,301 -2	40,628 -5	5,583 -2	23,812 -2	0,205 -2	17
TCN	0,322 -1	0,151 -1	0,388 -1	69,722 -1	8,537 -1	36,413 -1	0,283 -1	7

After running all the comparative models, the Friedman test yielded a p-value significantly lower than 0.05. This indicates the presence of statistically significant differences between the models in the metrics analyzed, suggesting that their performances are not equivalent. Therefore, it is essential to explore further which models stand out for each time series being evaluated. This result reinforces the validity of mapping through the scores, demonstrating a significant difference between the models' performances.

7. Geopolitics and the impact on oil prices

During the post-war period, the price of oil remained stable throughout the '50s and '60s, staying below 2 dollars. In the early '70s, following a coup in Libya, the establishment of a tax structure on the value of oil, and international conflicts in Egypt, Syria, the Sinai Peninsula, and the Gaza Strip, the First Oil Shock began in 1973, raising the price of this commodity to previously unseen levels.

Since then, as shown in Figure 1, geopolitical factors have started to influence price volatility, in addition to demand, which, with the economic development during the post-war period, led to a scarcity of this product during the '80s and '90s. Consequently, this developed the sector for increased production and the discovery of new wells. In 1979, the Iranian revolution and the Afghanistan war with the Soviet Union caused a sharp increase in the price of oil, reaching historically unprecedented levels, thus leading to the well-known second oil shock. Since then, the volatility of this commodity has become highly significant in the global market, impacting various other sectors. Upon reaching 1990, the Gulf War between the United States and Iraq initiated the third oil shock. Prices only began to cool down in 2009, but with the death of Osama bin Laden in 2011, values rose again, peaking above 100 dollars a barrel. The time series remained unstable, with no apparent trends until the arrival of COVID-19, which again raised barrel prices above 100 dollars.

It is significant to monitor the price of oil, both for explorers and for first-link suppliers in the chain, as the current barrel value will dictate the volume of investments in the sector for the exploration of new wells, as well as the production of new products for the maintenance of existing wells.

Figure 1 displays the fluctuating trends of the GPR index [25] on the left axis and the Brent Oil Price on the right axis during this research period. As observed in the chart, the GPR exhibits significant volatility, indicating an unstable global political environment, like the behavior of oil prices. Oil is susceptible to external information, showing movements like the GPR with some delay. Intuitively, major geopolitical events may explain the characteristic movements in the oil market. This same comparison of the GPR and oil prices was presented in [26]; however, the researchers used the daily price of a barrel of oil as the database instead of the monthly index.

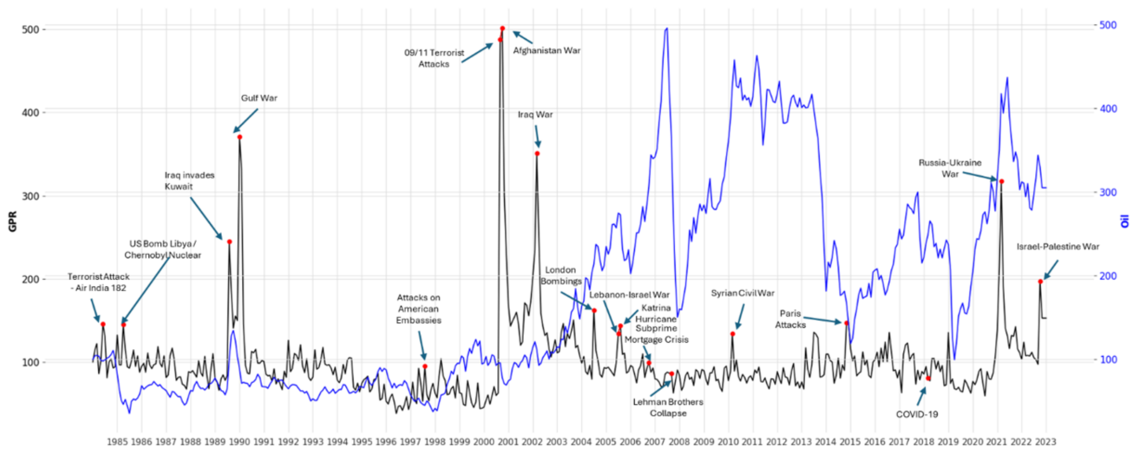

Figure 1. Placement of geopolitical events during the historical series of the global crude oil price.

8. Risk index for steel in the oil and gas sector

Creating a financial risk index for oil and gas projects in Brazil is crucial due to the complexity and challenges inherent to this sector. Projects generally involve significant investments in exploration, production, and infrastructure. In the case of this research, the evaluated project is linked to infrastructure. Assessing and managing financial risks helps protect invested capital and ensure the economic sustainability of the project and all stakeholders involved in the chain.

Given the volatility of oil prices linked to geopolitical factors, changes in global supply and demand, and unforeseen events, a financial risk index can help assess how these fluctuations may impact project profitability. The initial purpose of verifying the impact of the price of oil on one of the primary raw materials for product development for the oil exploration activity; there are no significant impacts on the steel industry, and there is a low correlation of the price of crude oil and the types of steel used.

Given future investments in the Oil and Gas supply chain in Brazil, the future price of the commodity should be considered a financial risk, as there may be a shortage of proposals to produce submarine exploration systems in the Brazilian market if the price of a barrel of oil were to fall.

In the paper [27], the authors present an approach used to estimate the influence of geopolitical risks on the oil and stock markets. They focus on the asymptotic form of the tail of the distribution rather than modeling the entire distribution. Absorbing part of the methodology used, for this research, we used the whole time series of GPR and calculated the maximum loss by calculating Value at Risk (VaR) on the oil price. The result of the VaR is added to the MAPE, which is the percentage error of the oil price, thus negatively impacting the proposed risk index.

$$FinancialRiskIndex = 20.74\% * w_1 + 13.80\% * w_2 + 14.85\% * (1 + 27.38\%) * w_3 \quad (2)$$

Where, w_1 and w_2 – refer to the purchase of Brazilian and international steel, respectively, and will have their weights distributed depending on the purchase volumes at each location for a given project; w_3 – determines the weight of the Brent oil price impact to be considered by the current price at which the index is being calculated versus the future price, referencing the minimum barrel price considered viable for a project; for example, in Brazil, for oil exploration activities, the barrel should be around 85 dollars per barrel; l_1 – average of the MAPE metric for the model that achieved the most consistent result for the steel price in Brazil: Carbon and Stainless; l_2 – average of the MAPE metric for the model that achieved the most consistent result for the price of international steel: Carbon and Stainless; l_3 – average of the MAPE metric for the model that achieved the most consistent result for Brent oil price plus the maximum loss of the GPR series: $MAPE * (1 + VaRGPR)$.

This index is a preliminary step towards developing a series of studies and improving models for the oil and gas sector, where steel price volatility is a significant concern. The research introduces an innovative approach to addressing this gap, providing a decision-support tool to assess and mitigate risks in such projects. The study establishes a robust method grounded in automated machine learning by comparing various libraries' results to determine the best fit for steel price series and considering significant covariates. This approach automatically selects the most suitable models and hyperparameters tailored to the specific dynamics of commodity price movements.

9. Limitation and future research

In comparison to other literature, this study highlights several gaps that guide us toward future research directions. For instance, it suggests examining the correlation between steel prices and daily crude oil prices, hypothesizing that forecasting daily commodity prices could be more accurate than monthly predictions. Geopolitical factors significantly influence future price trends, and accurately forecasting these prices remains a complex task that requires further investigation. Moreover, hyperparameter optimization could be more effective with additional iterations; however, we were constrained to 20 iterations due to hardware limitations. We plan to continue exploring alternative tools, techniques, and artificial intelligence models to enhance our analysis of oil and steel price fluctuations.

An upcoming area of focus for this research's risk index is to gain a deeper understanding of the steel price dynamics in Brazil, considering covariates from sectors that heavily utilize steel, such as the automotive and civil industries.

References

Ali Salamai, A. (2023). Deep learning framework for predictive modeling of crude oil price for sustainable management in oil markets. *Expert Systems with Applications*, 211:118658.

Athanasopoulos, G., Hyndman, R. J., Kourentzes, N., and Panagiotelis, A. (2024). Forecast reconciliation: A review. *International Journal of Forecasting*, 40(2):430–456.

Bai, S., Kolter, J. Z., and Koltun, V. (2018). An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *CoRR*, abs/1803.01271.

Challu, C., Olivares, K. G., Oreshkin, B. N., Garza, F., Mergenthaler-Canseco, M., and Dubrawski, A. (2022). N-hits: Neural hierarchical interpolation for time series forecasting.

Chen, Y., Kang, Y., Chen, Y., and Wang, Z. (2020). Probabilistic forecasting with temporal convolutional neural network.

Cihan, P. (2024). Comparative performance analysis of deep learning, classical, and hybrid time series models in ecological footprint forecasting. *Applied Sciences*, 14(4):1479.

Das, A., Kong, W., Leach, A., Mathur, S., Sen, R., and Yu, R. (2024). Long-term forecasting with tide: Time-series dense encoder.

Fang, Y., Wang, W., Wu, P., and Zhao, Y. (2023). A sentiment-enhanced hybrid model for crude oil price forecasting. *Expert Systems with Applications*, 215:119329.

Flunkert, V., Salinas, D., and Gasthaus, J. (2017). Deepar: Probabilistic forecasting with autoregressive recurrent networks. *CoRR*, abs/1704.04110.

Foroutan, P. and Lahmiri, S. (2024). Deep learning-based spatial-temporal graph neural networks for price movement classification in crude oil and precious metal markets. *Machine Learning with Applications*, 16:100552.

Guo, L., Huang, X., Li, Y., and Li, H. (2023). Forecasting crude oil futures price using machine learning methods: Evidence from china. *Energy Economics*, 127:107089.

He, K., Zheng, L., Yang, Q., Wu, C., Yu, Y., and Zou, Y. (2023). Crude oil price prediction using temporal fusion transformer model. *Procedia Computer Science*, 221:927–932. Tenth International Conference on Information Technology and Quantitative Management (ITQM 2023).

Herzen, J., Lässig, F., Piazzetta, S. G., Neuer, T., Tafti, L., Raille, G., Pottelbergh, T. V., Pasieka, M., Skrodzki, A., Huguenin, N., Dumonal, M., Kościsz, J., Bader, D., Gusset, F., Benheddi, M., Williamson, C., Kosinski, M., Petrik, M., and Grosch, G. (2022). Darts: User-friendly modern machine learning for time series.

Jorion, P. (2007). *Value at risk: the new benchmark for managing financial risk*. McGraw-Hill.

Kazmi, H., Fu, C., and Miller, C. (2023). Ten questions concerning data-driven modelling and forecasting of operational energy demand at building and urban scale. *Building and Environment*, 239:110407.

Knuth, D. E. (1984). *The TEX Book*. Addison-Wesley, 15th edition.

Maiti, R., Menon, B. G., and Abraham, A. (2024). Ensemble empirical mode decomposition based deep learning models for forecasting river flow time series. *Expert Systems with Applications*, 255:124550.

Mohsin, M. and Jamaani, F. (2023a). A novel deep-learning technique for forecasting oil price volatility using historical prices of five precious metals in context of green financing–a comparison of deep learning, machine learning, and statistical models. *Resources Policy*, 86:104216.

Mohsin, M. and Jamaani, F. (2023b). A novel deep-learning technique for forecasting oil price volatility using historical prices of five precious metals in context of green financing – a comparison of deep learning, machine learning, and statistical models. *Resources Policy*, 86:104216.

Oreshkin, B. N., Carpov, D., Chapados, N., and Bengio, Y. (2019). N-BEATS: neural basis expansion analysis for interpretable time series forecasting. *CoRR*, abs/1905.10437.

Oreshkin, B. N., Dudek, G., and Pelka, P. (2020). N-BEATS neural network for mid-term electricity load forecasting. *CoRR*, abs/2009.11961.

Orji, U. and Ukwandu, E. (2024). Machine learning for an explainable cost prediction of medical insurance. *Machine Learning with Applications*, 15:100516.

Panja, M., Chakraborty, T., Nadim, S. S., Ghosh, I., Kumar, U., and Liu, N. (2023). An ensemble neural network approach to forecast dengue outbreak based on climatic condition. *Chaos, Solitons Fractals*, 167:113124.

Salehin, I., Islam, M. S., Saha, P., Noman, S., Tuni, A., Hasan, M. M., and Baten, M. A. (2024). Automl: A systematic review on automated machine learning with neural architecture search. *Journal of Information and Intelligence*, 2(1):52–81.

Sen, A. and Dutta Choudhury, K. (2024). Forecasting the crude oil prices for last four decades using deep learning approach. *Resources Policy*, 88:104438.

Thomas, J. B. and K.V., S. (2023). Neural architecture search algorithm to optimize deep transformer model for fault detection in electrical power distribution systems. *Engineering Applications of Artificial Intelligence*, 120:105890.

Wang, J., Zhao, W., Tsai, F.-S., Jin, H., Tan, J., and Su, C. (2023). A study of crude oil futures price volatility based on multi-dimensional data from event-driven and deep learning perspectives. *Applied Soft Computing*, 146:110548.

Xu, Y., Liu, T., and Du, P. (2024). Volatility forecasting of crude oil futures based on bi-lstm-attention model: The dynamic role of the covid-19 pandemic and the russian-ukrainian conflict. *Resources Policy*, 88:104319.

Yang, K., Cheng, Z., Li, M., Wang, S., and Wei, Y. (2024). Fortify the investment performance of crude oil market by integrating sentiment analysis and an interval-based trading strategy. *Applied Energy*, 353:122102.