

Article

Nickel and Cobalt Price Volatility Forecasting Using a Self-Attention-Based Transformer Model

Shivam Swarup *  and Gyaneshwar Singh Kushwaha

Department of Management Studies, Maulana Azad National Institute of Technology, Bhopal 462016, India

* Correspondence: ss.193121001@manit.ac.in

Abstract: Both Nickel and Cobalt have been extensively used in cutting-edge technologies, such as electric vehicle battery manufacturing, stainless steel, and special alloys production. As governments focus on greener solutions for areas such as transportation and energy generation, both metals are increasingly used for energy storage purposes. However, their price uncertainty makes for an interesting case in the modern economy. This study focuses on the price volatility forecasting of Nickel and Cobalt using ANN (Artificial Neural Network) built on a special class of Transformer models used for multi-step ahead forecasts. Our results suggest that the given model is only slightly better in predictive accuracy compared to traditional sequential deep learning models such as BiLSTM (Bidirectional Long Short-Term Memory) and GRUs (gated recurrent units). Moreover, our findings also show that, like conventional approaches, in-sample behavior does not guarantee out-of-sample behavior. The given study could be utilized by industry participants for an inquiry into new and efficient ways to forecast and identify temporal-based structural patterns in commodity-based time series.

Keywords: commodity prices; self attention; volatility forecasting; deep learning approaches; transformer models



Citation: Swarup, S.; Kushwaha, G.S. Nickel and Cobalt Price Volatility Forecasting Using a Self-Attention-Based Transformer Model. *Appl. Sci.* **2023**, *13*, 5072. <https://doi.org/10.3390/app13085072>

Academic Editors: Xiaoping Ma, Difei Zhao and Qinxia Wang

Received: 15 March 2023

Revised: 11 April 2023

Accepted: 13 April 2023

Published: 18 April 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent times, there has been a focus by governments as well as non-profit organizations to cut down on fossil-fuel-based energy sources and increase the consumption of environmentally friendly and renewable sources of energy. To cater to this, private sector-based industries are focusing on battery-based energy storage solutions, which extensively use metals such as Nickel and Cobalt for their regular operations. This results in a kind of uncertainty in the prices of these metallic commodities, which ultimately results in the volatile nature of their pricing behavior. Numerous studies have been conducted on the connection between resources and economic growth to determine the relationship between resource readiness and how it helps a country in the expansion of its economy. Since natural resources are a necessary component of activities such as manufacturing, it would be reasonable to assert that having access to these basic materials stimulates the economy. However, this is not the case with every country or region. For example, studies such as [1,2] reveal how oil-rich countries such as Venezuela and Nigeria lag behind oil-scarce countries of East Asia.

Additionally, a significant amount of greenhouse gases has been released into the atmosphere because of world economic progress, indirectly resulting in a significant rise in global temperatures and consequent climate change. To curb this phenomenon, it is necessary to put more focus on those energy mechanisms consuming fewer GHGs, such as battery-dominated power sources for automobile and various other industries. Nonferrous base metals such as Lithium, Nickel, and Cobalt are extensively used in new-age battery technologies and, hence, have witnessed frequent price variations, especially in the last decade.

Due to the limited size of the Lithium market, there is a lack of active trading in lithium futures, resulting in a limited level of market liquidity to sustain a healthy lithium futures market. Moreover, the majority of the global lithium supply is generated by only a handful of nations and a limited number of corporations. This situation could create a concentration risk and restrict the ability of market participants to mitigate their market risk.

On the other hand, both Nickel and Cobalt have an active financial trading market since they are widely employed in a wide range of industries, including the production of computer parts, home appliances, automobiles, aircraft, military applications, and coinage; see [3]. One of Nickel's most significant benefits over other metals is its high resistance to oxidation and corrosion. Because of its economical price and superior energy density in comparison to similar products, nickel is a critical metal in the production of EV batteries [4,5]. Between 2017 and 2025, there will be a 39% yearly rise in the application of Nickel in batteries for electric vehicles. According to estimates, battery manufacturing will account for 37% more of the world's Nickel usage by 2030 than it did in 2017 [5].

Similarly, Cobalt is also one of the essential elements of the clean energy generation supply chain. Composite materials, electric cars and mobile phone batteries, and the move to renewable energy all depend on Cobalt, a vital element that makes it possible to store renewable energy. Additionally, due to rising prices and product performance losses, Cobalt replacement is essentially impossible. In that regard, the use of Cobalt in known technologies is inevitable. Cobalt is also produced as a by-product of other significant base metals such as Nickel, thus making its prices more in sync with it; see [6].

If these metal price predictions are not accurate, they can distort economic analyses and lead to significant financial losses on investments. Additionally, there is a significant need for companies to assess the scope and financial viability of mining decisions beforehand. As a result, there is a substantial need to evaluate the price fluctuations of these two metallic commodities, and also price estimates should be carried out using a very precise method to minimize error. The emergence of deep learning methodology as a new and advanced form of AI has provided us with opportunities to create improved prediction models that can more accurately forecast the volatility of prices for Nickel and Cobalt.

The paper is structured as follows: Section 2 provides a review of the literature, while Section 3 describes the dataset and forecasting methods. Section 4 presents the forecast results from the TFT method and its comparison to other deep learning models, along with a discussion of these results. The paper concludes with Section 5, which offers some final remarks and some future recommendations.

2. Literature Review

In their 1995 study on the volatility of metal prices, ref. [7] created a model that connects metal volatility to metal balance.

Plourde and Watkins [8] were the first to examine the fluctuation in prices of crude oil and a variety of industrial metals, including Nickel. They found that the volatility of crude oil for the period of 1985–1994 was comparable to that of Nickel.

Behmiri and Manera [9] used the generalized autoregressive conditional heteroskedastic and Glosten–Jagannathan–Runkle models to investigate the effect of outliers and oil price shocks on the volatility of metal prices. It was discovered that outliers skewed estimate models, and that the price volatility of metals responds asymmetrically and differentially to changes in the price of oil. Chen et al. [10] used monthly Nickel and Zinc prices from the year 2015 to assess the effectiveness of the modified grey wave forecast model.

The findings of these two tests showed that, in terms of prediction accuracy and computing efficiency, the modified grey wave approach outperformed the original grey wave method.

Shao et al. [11] recommended an approach for forecasting Nickel prices that employed an improved particle swarm optimization (PSO) technique combined with LSTM networks. They compared the performance of classic LSTM networks and ARIMA with the upgraded PSO-LSTM model. The results revealed that the enhanced PSO method could notably

enhance the predictive accuracy of the LSTM networks and had a quicker convergence rate. For anticipating changes in the price of gold, Alameer et al. [12] suggested a hybrid model combining the whale optimization algorithm (WOA) with multilayer perceptron neural networks (NN).

Numerous methods have been applied previously for understanding the forecasting accuracy of metallic commodity prices. For example (support vector machine) with the genetic algorithm to predict gold future prices. Likewise, ref. [13] used SVR for foretelling the prices of copper futures. Dehghani et al. [14] used a bat algorithm for predicting copper prices. Kriechbaumer et al. [15], on the other hand, used wavelet-based ARIMA (auto regressive integrated moving average) to forecast metallic prices. Similarly, experts such as [12,16,17] used hybrid models including neuro-fuzzy inference systems and genetic algorithms for predicting the forecasting accuracy of various metallic prices. Similarly, ref. [18] advanced the Heterogeneous AR (HAR) model by [19] for increasing the accuracy of price volatilities of commodities. Liu et al. [20] utilized a hybrid neural network that incorporated Bayesian optimization and wavelet transform techniques to predict copper prices over the short- and long-term. In addition, they employed LSTM and GRU networks to analyze the data and predict future copper prices. Based on their research, both LSTM and GRU networks demonstrated accuracy in predicting copper prices over various timeframes.

There have been other relevant studies conducted using various deep-learning-based neural networks models. For example, ref. [21] forecasted Nickel prices based on LSTM and GRU and found that both the methods provided significant improvements over previous ones in terms of their accuracy as well as the computation rate.

This paper primarily deals with the impact of an energy commodity, namely oil prices on precious metals which are used mainly for clean energy purposes. The reason for the interest is primarily that the oil market is particularly susceptible to fluctuations in market volatility and economic growth, as stated by [22]. This study takes a step further and utilizes more advanced Recurrent Neural Networks (RNNs) that employ self-attention mechanisms such as TFTs (Temporal Fusion Transformers) to analyze their performance in terms of both speed and accuracy.

Temporal Fusion Transformers (TFTs) are a variant of the transformer design of [23], a class of neural networks initially developed for natural language processing applications. When used with time series data, the attention mechanisms of the transformer architecture allow the model to focus on the most important data, enabling it to capture long-term relationships between time steps accurately.

TFTs are particularly well-suited for applications that require the projection of multiple time steps into the future, and they can handle large amounts of datasets. They have been shown to achieve state-of-the-art performances on a variety of time series forecasting tasks, including univariate and multivariate forecasting.

In addition to forecasting, TFTs can be utilized for various time series assessment tasks such as anomaly detection and the interpolation of missing values. They are versatile, and can process many types of time series data including both typical and atypical time series. Recent studies that have utilized TFTs include those by [24–26] which have applied them to a variety of applications ranging from wind speed prediction to tourism demand.

In this study, we focus on forecasting the volatility of the two most critical metals used for fossil-free energy, namely Nickel and Cobalt.

3. Data and Methodology

Depending on the type of data being studied and the company or financial scenario, several forecasting models may be employed. For instance, one may utilize statistical models, machine learning techniques, or professional judgment. As substitutes for traditional methods, Recurrent Neural Networks (RNNs) and their variations, such as Long Short-Term Memory Networks (LSTMs), have been suggested for modeling intricate sequential data such as natural language, audio waves, and video frames. Studies conducted recently have demonstrated the ability of LSTMs to capture complex temporal patterns in dynamic

time series and apply them to forecasting tasks [27,28]. By using an LSTM encoder-decoder model to map the past of a time series to its future, sequence-to-sequence learning can naturally formulate multi-step forecasting.

Multi-horizon forecasting is a common practice that involves making forecasts at various time horizons, from short-term predictions for the coming days or weeks to long-term predictions for several years. One of its many benefits is that, over a certain time frame, it offers directions for scheduling resources and making decisions [29]. The objective of multi-horizon forecasting is to give businesses and investors a thorough understanding of potential trends and changes, so that they can make wise judgments. In volatile or fast-changing contexts, the neural network architecture called Temporal Fusion Transformers (TFT), developed by [30], was applied in the present study. The novelty of this neural network structure is its ability to integrate other network architectures such as LSTM layers and attention heads. Three major blocks underlie the architecture of this neural network, namely:

- A multiheaded temporal attention block whose purpose is to identify long-range patterns present in the time series;
- Sequence-by-sequence LSTM encoder-decoders to account for the short-range patterns present in the time series;
- Residual network blocks (GRN) whose function is to eliminate unimportant and unused inputs to the neural network.

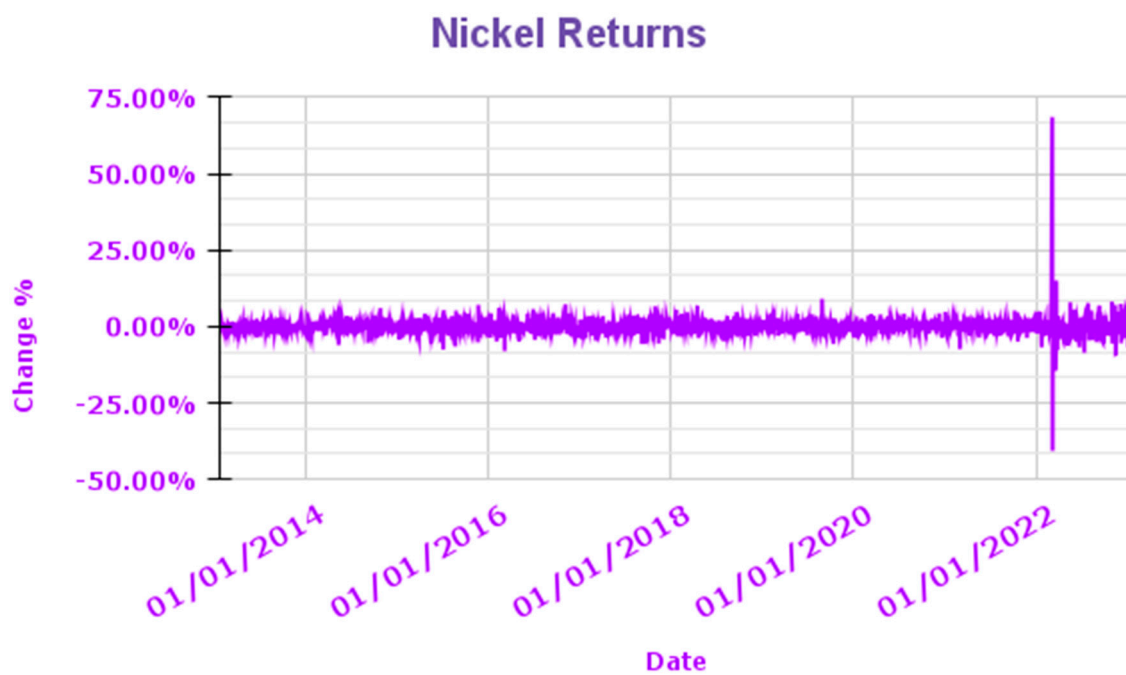
Another important feature of TFT is the possibility of determining the weight or contribution of explanatory variables to network performance, similar to the determination of statistical significance in econometric regression models, but with the difference that no distributional assumption is required.

Data and Variables

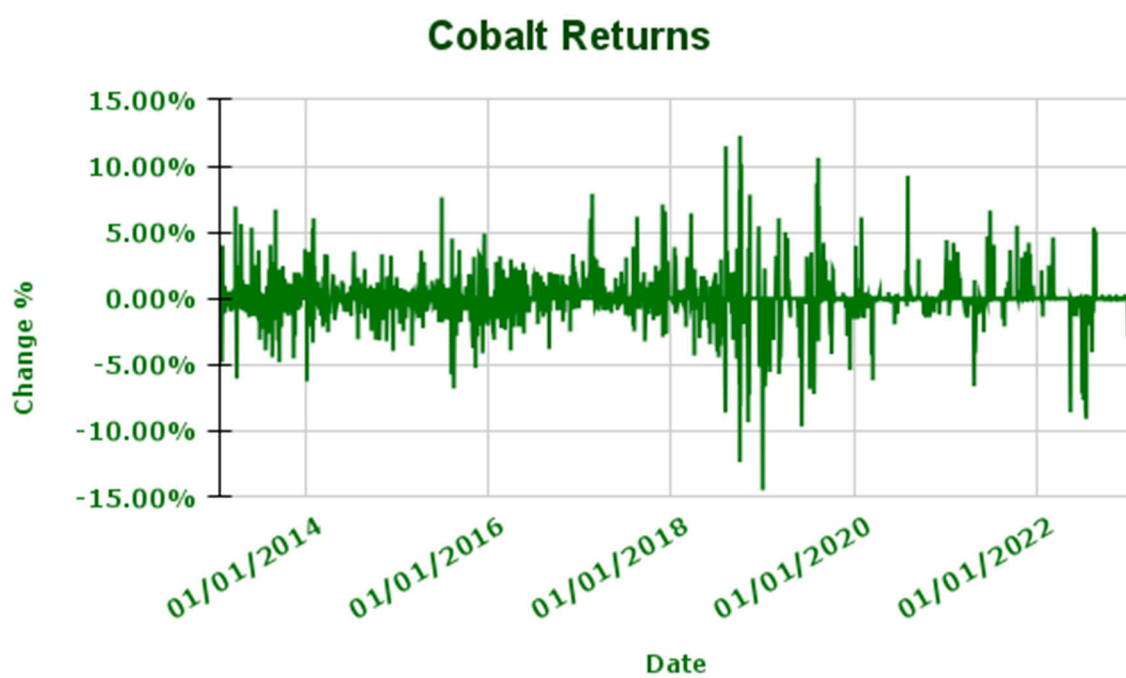
The variables used in this work were Cobalt futures prices, Nickel futures prices, and WTI Crude Oil Futures prices, taken from Bloomberg data services. The intention for using the futures rather spot prices was due to the increased practical viability of the study. The data were considered on a daily basis from a historical period spanning from 7 March 2013 to 12 December 2022. From this basic data, the returns of Cobalt and Nickel future prices were generated and, finally, a time series was obtained describing the evolution of the volatility of Cobalt and Nickel returns, the main variables of the study. Volatilities were estimated through GARCH (1, 1) models.

As can be seen from Figure 1a, there was surge in the prices of Nickel, especially after the onset of the Russia–Ukraine War in early 2021, which was followed by a huge spike. They have remained highly volatile thereafter. Cobalt prices, however, remained quite stable throughout the duration of study (Figure 1b).

When it came to crude oil prices, high fluctuations could be observed during the COVID-19 pandemic, with a spike, and they have stabilized in recent years with production increases from various OPEC members.



(a)



(b)

Figure 1. Cont.

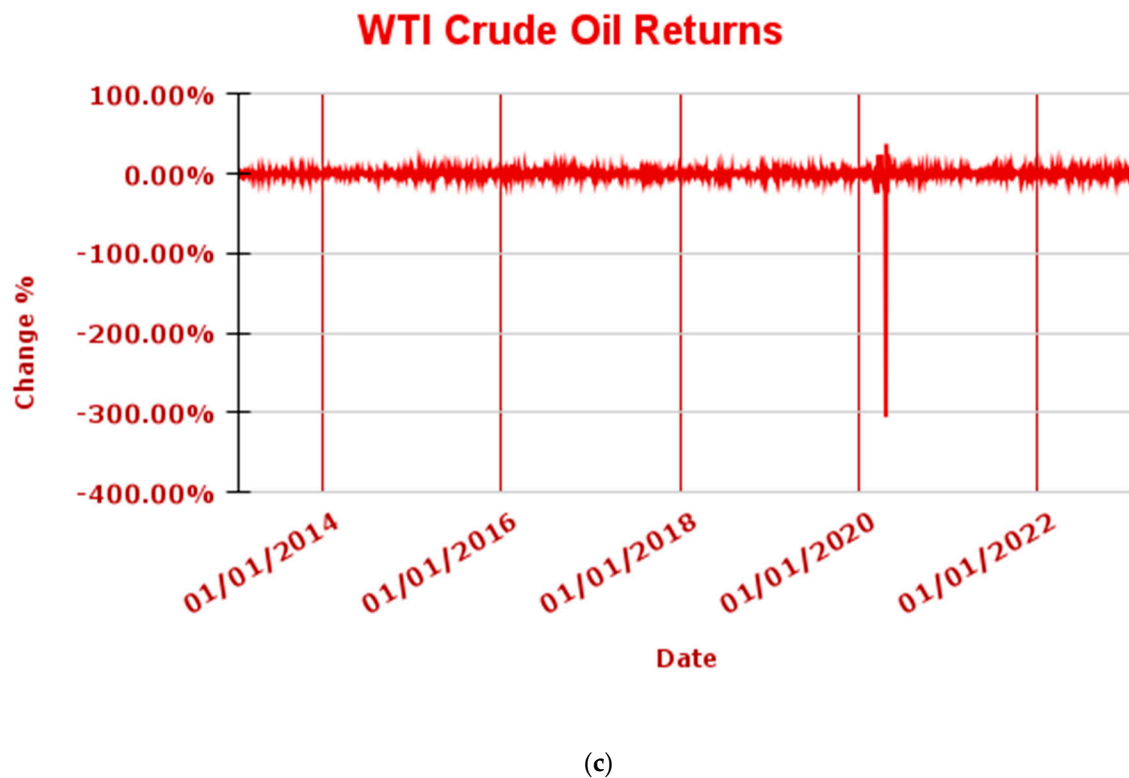


Figure 1. Shows the daily returns of (a) Nickel, (b) Cobalt and (c) WTI Crude Oil for the Period under study.

4. Empirical Results

The results are compiled based on two parts. In the first, the results of the accuracy measures of TFTs are considered with respect to other deep learning models such as BiLSTM, GRU, and the No change forecasting model. In the second, the application of the Temporal Fusion Transformer model (TFT) is considered based upon 6 months and 12 months of historical data.

The total dataset was divided into training and test sets, with approximate proportions of 97% and 3%, respectively. The TFT model was estimated with the last 365 observations of the training set, which corresponds to a one-year history. A No change forecast model was also developed, with an assumption that the volatilities would remain constant for future periods. Additionally, the daily return values of WTI crude oil prices were added as an external variable. The given model accuracy was compared with other comparative deep learning models, namely, BiLSTM and GRU, given that the TFT model was estimated with the same number of observations as the training set, which corresponds to about a preceding year of history.

We could observe from Table 1 that the volatilities of Cobalt and Nickel future prices presented similar mean values and standard deviations of around 1.3 and 0.4, respectively. The minimum and maximum volatilities reached by Cobalt returns were 0.54 and 2.41, respectively, while for Nickel returns the minimum and maximum volatilities were 0.75 and 3.10, respectively. On the other hand, WTI crude oil prices averaged around USD 65 per barrel during the study period, reaching a minimum price of around USD 19 per barrel and a maximum price of approximately USD 123 per barrel after adjusting for the outliers. All the variables presented a positive asymmetry, and distributions were more flattened than for a normal distribution (platykurtic) since the Kurtosis coefficient was less than 3 for all the variables; this was confirmed by the Jarque Bera test, with a p -value very close to zero with respect to three variables, indicating the rejection of the null hypothesis that establishes the distributional normality of the variables. This also suggested that the

right branch of the curve was more elongated than the left one, indicating the presence of very high volatility values and WTI prices that could possibly act as outliers.

Table 1. The descriptive statistics of the variables of interest in the study over the total dataset.

	Volatility		Prices
	Cobalt	Nickel	WTI
Mean	1.31	1.3	65.05
SD	0.46	0.47	22.16
Min	0.54	0.75	19.78
Max	2.41	3.1	123.7
Skewness	0.13	1.43	0.54
Kurtosis	−1.13	1.46	−0.75
Jarque-Bera	54.07	413.87	69.32
N	974	974	974

4.1. Hyperparameter Tuning

The optimal parameters' prior determination is important in developing models with neural network architectures. This involves establishing a grid of feasible values for each parameter and evaluating each combination of parameters on an objective function until the optimal combination has been established. This process usually requires considerable computational effort, given the highly nonlinear nature of neural network architectures. In the case of the Temporal Fusion Transformer model, the parameter grid considered in both scenarios was as follows.

The values in the last two columns of Table 2 contain the optimal parameters used in the estimation of the TFT model for each respective scenario.

Table 2. Parameters for the optimal model.

Parameters	Optimum Values		
	Range	Scenario 1	Scenario 2
No. of Trials	200	200	200
Maximum No. of Epochs	50	50	50
Gradient Clip Value Range	(0.01, 1.0)	0.150672	0.011087
Hidden Size Units	(8, 128)	23	119
Hidden Continuous Size Units	(8, 128)	11	14
Attention Head Size	(1, 4)	3	2
Learning Rate	(0.001, 0.1)	0.001071	0.002029
Dropout	(0.1, 0.3)	0.115563	0.130221

4.2. Forecasts Generated from the Test Set

For the first part, three set-ups were considered from the perspective of the historical period: (i) 30-day forecasts and training history of one year (this scenario would reflect the medium-long term), (ii) 20-day forecast and training history of one year, and finally (iii) 15-day forecast and training history of one year, scenarios that would reflect the short term trend, respectively.

Additionally, other sequential models such as BiLSTM (Bidirectional Long Short-term memory) and GRU (Gated Recurrent Units) were also tested for the future 15, 20 and 30 days.

The forecasts of the TFT model on the test set for 15 days, as can be seen from Table 3, did not show a good fit; the mean absolute percentage error (MAPE) with respect to the volatility of Cobalt returns was 26.91%, while with respect to the volatility of Nickel returns the MAPE was 26.98%. As the horizon period increased to 20 days, the error rate decreased to a drastic level of 3.47% for Cobalt and 2.817% for Nickel using the TFT model. Similar settings were noticeable in no change forecasts, where there was a drastic change in the error values. This was somehow the result of internal designs of TFTs, which favor only the key features to be taken to gauge future values.

Table 3. MAPE values of forecasting accuracy of the used models.

Cobalt				
MAPE	TFT	BLSTM	GRU	NCF
15	26.91	35.463	31.7975	27.6
20	3.4785	15.394	12.021	4.1105
30	9.05	16.882	16.459	5.0178
Nickel				
MAPE	TFT	BLSTM	GRU	NCF
15	26.98	26.39	28.92	28.22
20	2.817	2.95	4.336	5.9525
30	3.659	4.11	5.17	6.178

In contrast the error values for BLSTM and GRU forecasts, a major change can be seen for the Nickel forecasts but not for the Cobalt forecast. A familiar leaning was shown for the 30-day-ahead forecasts. All in all, the TFT forecasts were the most accurate in comparison to all other models, and this was evidenced by the Diebold–Mariano test in Table 4. In the case of Cobalt (Table 4), TFTs were significantly better at predicting pricing volatilities for different time horizons, namely, 15, 20 and 30 days in comparison to No change and GRU forecasts. However, with respect to BiLSTMs, the accuracy results were not found to be statistically significant for 20- and 30-days-ahead forecasts, and were less accurate for 15-day-ahead forecasts. Likewise, for the case of Nickel, the accuracy of TFTs was found to be greater with comparison to the No change forecasts as well as GRUs; however, this was not statistically significant for the 15-days-ahead forecasts. In comparison to the BiLSTMs forecasts, the results were not much different. Additionally, the empirical practice considered a MAPE less than or equal to a 5% good fit, which was the case with accuracy measures of TFT forecasts. Given that models considered a one-year history, the results seemed to suggest that from information representing the medium-long term it was not possible to predict a month of daily observations in an acceptable way with the TFT model; likewise, it provided indirect evidence of the poor predictive power of WTI crude oil futures prices.

4.3. Application of TFTs Based on Training History of Five Months and One Year

The study using the Temporal Fusion Transformer model (TFT) examined two different scenarios based on the historical period. The first scenario involved a 30-day forecast with a training history of one year, which reflects the medium-to-long-term trend. The second scenario involved a 30-day forecast with a training history of approximately five months, which reflects the short-term trend. To evaluate the accuracy of the forecasts, the total dataset was divided into training and test sets, with the training set consisting of approximately 97% of the data and the test set consisting of approximately 3%. In Scenario 1, the TFT model was estimated using the last 365 observations of the training set, corresponding to a one-year history, while in Scenario 2 the model was estimated using the last 142 observations of the training set, corresponding to a five-month history.

Table 4. The levels of statistical significance as shown by Diebold–Mariano Test, indicating a 5% level of significance, based on the p -value and DM statistics (inside the brackets).

Cobalt			
TFT	NCF	GRU	BLSTM
15	0.02(2.193)	0.04(1.069)	1.87(−1.53939)
20	0.01(2.47)	0.039(2.069)	0.54(0.60129)
30	0.003(2.92)	0.000217(3.69)	0.98(0.023)
Nickel			
TFT	NCF	GRU	BLSTM
15	0.12(1.5364)	0.145(1.457)	0.0841(−1.539)
20	0.015(2.4158)	0.148(1.1446)	0.00414(2.867)
30	0.04(1.345)	0.000217(3.69)	0.023(1.98)

The TFT model forecasts on the test set with other models such as; similar to the first scenario, the forecasts did not show a good fit. The mean absolute percentage error concerning the volatility of Cobalt returns was 9.7%, while concerning the volatility of Nickel returns, the MAPE was 13.8%; both values were well above the benchmark value for a good fit of 5%. Since Scenario 2 considered a history of about five months, the results suggest that from information representing the short term it was not possible to predict a month of daily observations in an acceptable manner with the TFT model on the test set, and this continues to provide evidence of the poor predictive power of WTI crude oil prices as an explanatory variable (as shown in Section 3).

Since the sample covered the period from 7 March 2013 to 12 December 2022, the Out-of-Sample forecasts for one month of daily observations covered the period 1 January 2023 to 30 January 2023. Figures 2 and 3 show the Out-of-Sample forecasts under Scenario 1.

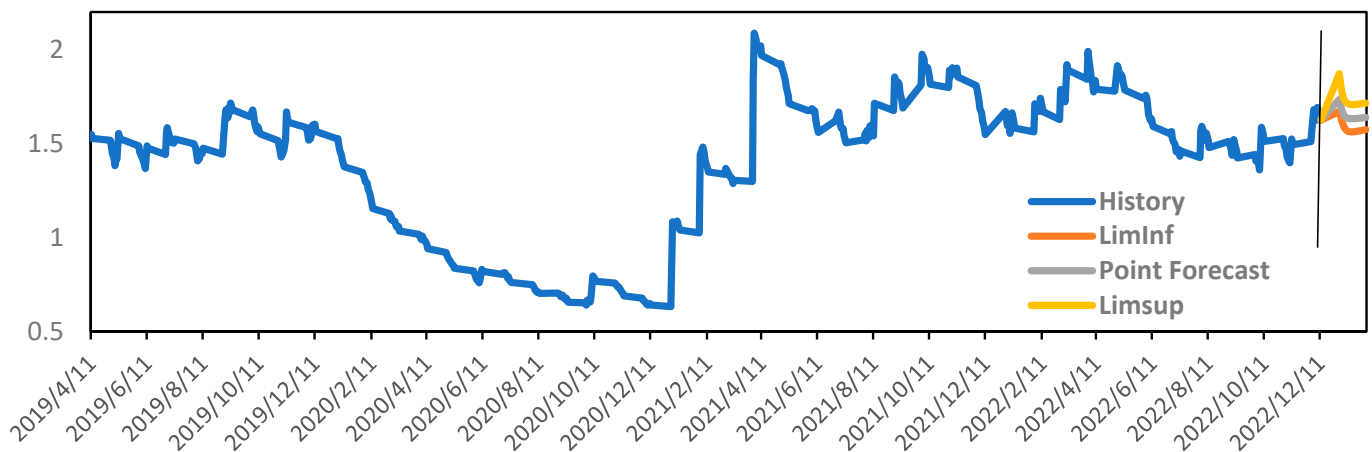


Figure 2. Out-of-Sample Forecast: Scenario 1, Volatility: Cobalt/Confidence Intervals 95% (where LimInf and Limsup are lower and upper confidence limits).

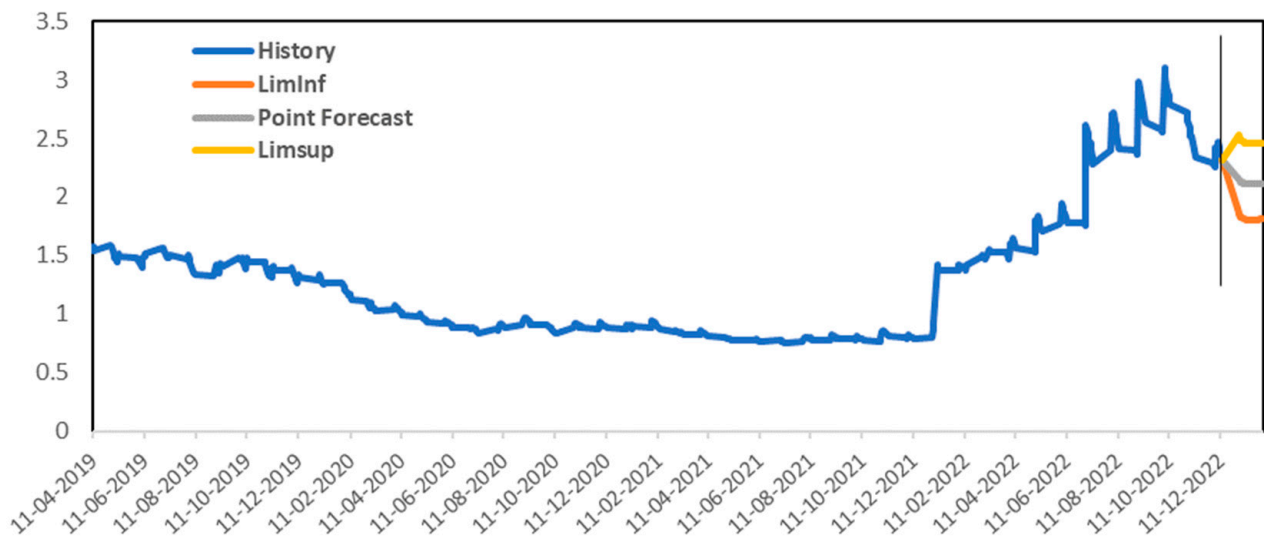


Figure 3. Out-of-Sample Forecast: Scenario 1, Volatility: Nickel/Confidence Intervals 95% (where LimInf and LimSup are lower and upper confidence limits).

As can be seen, while the in-sample fit was not very accurate, the Out-of-Sample forecasts under Scenario 1 appeared to capture the short-term trend. For example, with respect to the volatility of Cobalt price returns, the forecasts appeared to reflect the trend at the end of the volatility series (more precisely, that of the 7 December 2021 to 12 December 2022 period), with the confidence intervals capturing the variability of the earlier period. With respect to the volatility of Nickel price returns, the forecasts also appeared to capture the trend of the final period of the series; that is, the period from 7 March 2022 to 12 December 2022. Figures 4 and 5 show the Out-of-Sample forecasts under Scenario 2.

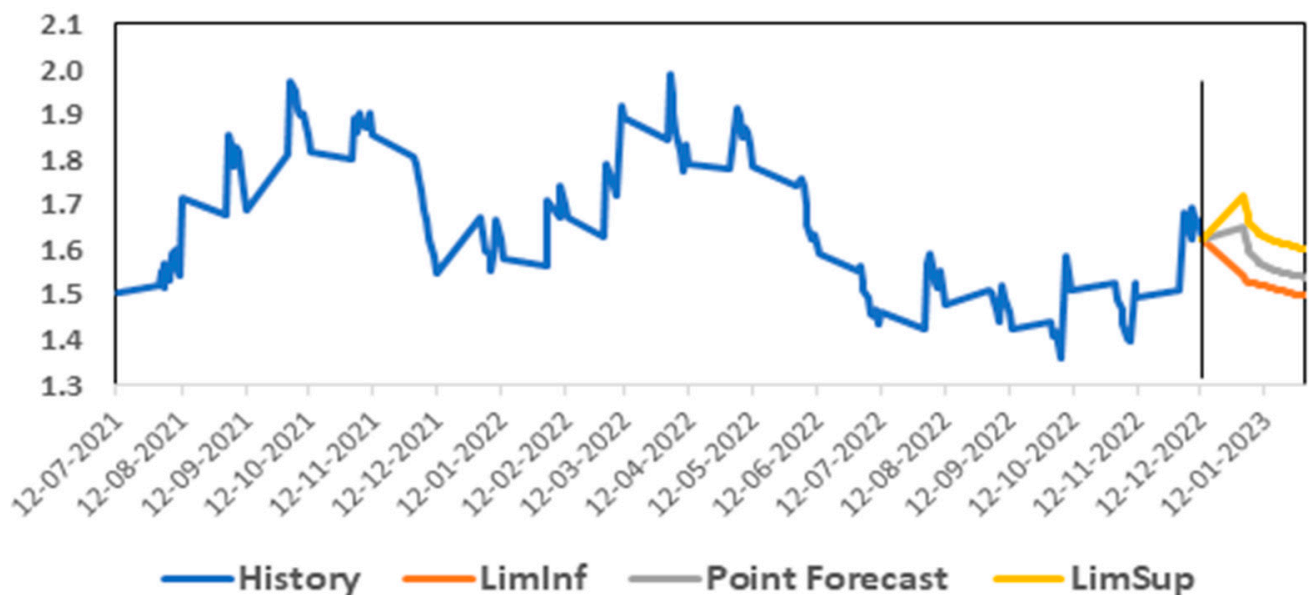


Figure 4. Out-of-Sample Forecast: Scenario 2, Volatility: Cobalt/Confidence Intervals 95% (where LimInf and LimSup are lower and upper confidence limits).

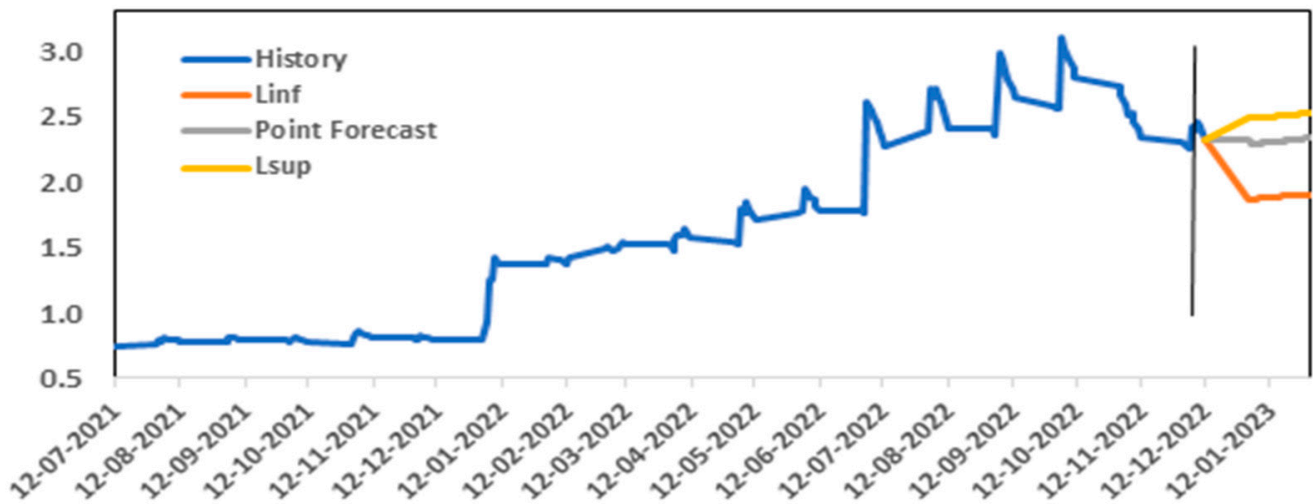


Figure 5. Out-of-Sample Forecast: Scenario 2, Volatility: Nickel/ Confidence Intervals 95% (where Linf and Lsup are lower and upper confidence limits).

Regarding the volatility of Cobalt price returns under Scenario 2, it can be seen that the Out-of-Sample forecasts appear to be an average between the behavior exhibited before and after the 7 November 2022 date, and likewise the 95% confidence intervals. On the other hand, with respect to the volatility of Nickel price returns under Scenario 2, the forecasts appear to capture the most recent behavior of volatility, i.e., the period from 7 March 2022 to 12 December 2022, similar to the 95% confidence intervals.

4.4. The Importance of the Variables

Multi-horizon forecasting in practical applications typically involves utilizing a range of data sources, such as advance knowledge about future events (also known as the known inputs), external time series, and unchanging metadata, with no prior understanding of how they relate to each other.

The TFT model takes in three types of inputs: unchanging metadata, also known as static covariates, past inputs that change over time, and future inputs that are already known. To select the most important features from these inputs, variable selection is employed to make careful and deliberate choices.

In the given case, the encoder variables used in the prediction process contained features for which previous values were already known. These variables included the features that had been selected previously, along with an index indicating the relative time. As mentioned in Figure 6, percentage-wise variable importance is described. Also, Table 5 shows the weights of the variables included in the study according to the TFT model in both scenarios.

In general, the `time_idx` (time index) variable always tends to have the highest weight in the TFT neural architecture, since it indicates the importance of the time dependence of the series in the construction of the TFT model; a similar argument is valid for the `relative_time_idx` (relative time index) variable, followed by volatility (the target variable). As for the explanatory variable of interest, the `price_wti` (WTI prices), this had greater relevance in the construction of Scenario 2, which reflects the short-term dependence at the level of history, with a weight of 24.2%. In the construction of Scenario 1, it had a weight of only 5.5%. The low weight shown by the variable in relative terms could partly explain the low adjustment capacity shown by the model in the test sets.

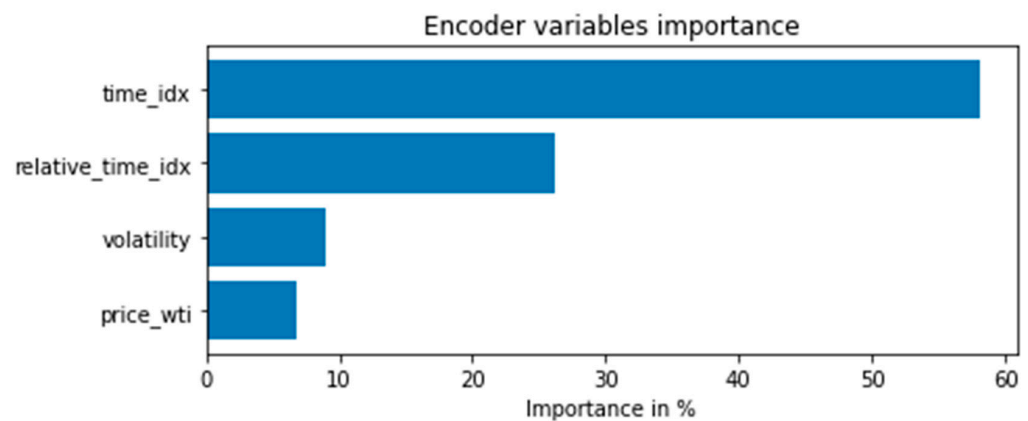


Figure 6. Encoder-based results for Variable Importance.

Table 5. Variables importance.

Scenario 1				
Variables	Time Index	Relative Time Index	WTI Prices	Volatility
Value	1.388912	0.3244441	0.1090103	0.1776332
Weights	69.40%	16.20%	5.50%	8.90%
Scenario 2				
Variables	Time Index	Relative Time Index	WTI Prices	Volatility
Value	0.880674	0.5411383	0.4845534	0.0936345
Weights	44.00%	27.10%	24.20%	4.70%

5. Conclusions and Policy Recommendations

The TFT model was applied in two scenarios based on historical data, one reflecting medium to long-term behavior and the other reflecting short-term behavior. The neural network model did not perform well in terms of fit over the test set in either case. However, the Out-of-Sample forecasts performed better, with the model being able to reflect the average between long and short-term behavior and being able to capture the short-term trend of the volatilities of Cobalt and Nickel price returns in most scenarios.

Our data analysis shows that models such as TFTs are highly nonlinear and computationally intensive, and they do not always outperform traditional methods. This observation is consistent with the phenomenon observed in the Makridakis M4 [31] competitions, where machine learning methods did not perform better than traditional methods. However, these methods are continually developing and have the benefit of enabling the generation of batch forecasts for time series, which can be challenging with traditional methods that rely on a combination of artistic and technical expertise.

A disclaimer regarding the model's performance is that there is wide empirical evidence regarding the difficulty of modeling the volatility of financial variables, even with the most powerful econometric models. Moreover, there is broad evidence that the good fit of an in-sample model does not ensure similar Out-of-Sample behavior, and vice versa, and the TFT model does not seem to escape this dichotomy, as the results of the paper point to better Out-of-Sample performance, but not in-sample.

The study did not find WTI crude oil prices to be very relevant in the present study, leaving an open door for exploring other explanatory variables that could have greater predictive capacity. One weakness of the study is that the variables considered had many gaps, which is a feature that architectures such as the one considered in the study are sensitive to. Most of these neural networks arose for Natural Language Processing (NLP), where the meaning of a text is highly dependent on the previous text. Various sources should be used in the process of constructing the time series to minimize the number of

gaps to a level where interpolation methods can be used with the least possible distortion of the series.

The study suggests using different fossil fuel sources, such as natural gas, and analyzing whether their prices could impact the prices of the metals being used. This research could have important future implications, such as exploring how well the current methodology performs in comparison to other modern deep learning models and how effectively it predicts price volatilities with optimum computations. It could also suggest using other green materials besides Nickel or Cobalt to investigate how the methodology works with different datasets.

Author Contributions: Conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft preparation, writing—review and editing, visualization, S.S.; supervision, G.S.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not Applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data would be made available on request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Auty, R.M. *Resource-Based Industrialization: Sowing the Oil in Eight Developing Countries*; Clarendon Press: Oxford, UK, 1990.
2. Sachs, J.D.; Warner, A. *Natural Resource Abundance and Economic Growth*; Oxford University Press: Oxford, UK, 1995.
3. Henckens, M.L.C.M.; Worrell, E. Reviewing the availability of copper and nickel for future generations. The balance between production growth, sustainability and recycling rates. *J. Clean. Prod.* **2020**, *264*, 121460. [[CrossRef](#)]
4. Nguyen, R.T.; Eggert, R.G.; Severson, M.H.; Anderson, C.G. Global electrification of vehicles and intertwined material supply chains of cobalt, copper and nickel. *Resour. Conserv. Recycl.* **2021**, *167*, 105198. [[CrossRef](#)]
5. Yao, P.; Zhang, X.; Wang, Z.; Long, L.; Han, Y.; Sun, Z.; Wang, J. The role of nickel recycling from nickel-bearing batteries on alleviating demand-supply gap in China's industry of new energy vehicles. *Resour. Conserv. Recycl.* **2021**, *170*, 105612. [[CrossRef](#)]
6. Tisserant, A.; Pauliuk, S. Matching global cobalt demand under different scenarios for co-production and mining attractiveness. *J. Econ. Struct.* **2016**, *5*, 4. [[CrossRef](#)]
7. Brunetti, C.; Gilbert, C.L. Metals price volatility, 1972–1995. *Resour. Policy* **1995**, *21*, 237–254. [[CrossRef](#)]
8. Plourde, A.; Watkins, G.C. Crude oil prices between 1985 and 1994: How volatile in relation to other commodities? *Resour. Energy Econ.* **1998**, *20*, 245–262. [[CrossRef](#)]
9. Behmiri, N.B.; Manera, M. The role of outliers and oil price shocks on volatility of metal prices. *Resour. Policy* **2015**, *46*, 139–150. [[CrossRef](#)]
10. Chen, Y.; He, K.; Zhang, C. A novel grey wave forecasting method for predicting metal prices. *Resour. Policy* **2016**, *49*, 323–331. [[CrossRef](#)]
11. Shao, B.; Li, M.; Zhao, Y.; Bian, G. Nickel price forecast based on the LSTM neural network optimized by the improved PSO algorithm. *Math. Probl. Eng.* **2019**, *2019*, 1934796. [[CrossRef](#)]
12. Alameer, Z.; Abd Elaziz, M.; Ewees, A.A.; Ye, H.; Jianhua, Z. Forecasting gold price fluctuations using improved multilayer perceptron neural network and whale optimization algorithm. *Resour. Policy* **2019**, *61*, 250–260. [[CrossRef](#)]
13. Astudillo, G.; Carrasco, R.; Fernández-Campusano, C.; Chacón, M. Copper price prediction using support vector regression technique. *Appl. Sci.* **2020**, *10*, 6648. [[CrossRef](#)]
14. Dehghani, H.; Bogdanovic, D. Copper price estimation using bat algorithm. *Resour. Policy* **2018**, *55*, 55–61. [[CrossRef](#)]
15. Kriechbaumer, T.; Angus, A.; Parsons, D.; Casado, M.R. An improved wavelet–ARIMA approach for forecasting metal prices. *Resour. Policy* **2014**, *39*, 32–41. [[CrossRef](#)]
16. García, D.; Kristjanpoller, W. An adaptive forecasting approach for copper price volatility through hybrid and non-hybrid models. *Appl. Soft Comput.* **2019**, *74*, 466–478. [[CrossRef](#)]
17. Rubaszek, M.; Karolak, Z.; Kwas, M. Mean-reversion, non-linearities and the dynamics of industrial metal prices. *A forecasting perspective. Resour. Policy* **2020**, *65*, 101538.
18. Zhang, H.; Zhu, X.; Guo, Y.; Liu, H. A separate reduced-form volatility forecasting model for nonferrous metal market: Evidence from copper and aluminum. *J. Forecast.* **2018**, *37*, 754–766. [[CrossRef](#)]
19. Andersen, T.G.; Bollerslev, T.; Huang, X. A reduced form framework for modeling volatility of speculative prices based on realized variation measures. *J. Econom.* **2011**, *160*, 176–189. [[CrossRef](#)]

20. Liu, K.; Cheng, J.; Yi, J. Copper price forecasted by hybrid neural network with Bayesian Optimization and wavelet transform. *Resour. Policy* **2022**, *75*, 102520. [\[CrossRef\]](#)
21. Ozdemir, A.C.; Buluş, K.; Zor, K. Medium-to long-term nickel price forecasting using LSTM and GRU networks. *Resour. Policy* **2022**, *78*, 102906. [\[CrossRef\]](#)
22. Baruník, J.; Kocenda, E.; Vácha, L. Volatility spillovers across petroleum markets. *Energy J.* **2015**, *36*. [\[CrossRef\]](#)
23. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 6000–6010.
24. López Santos, M.; García-Santiago, X.; Echevarría Camarero, F.; Blázquez Gil, G.; Carrasco Ortega, P. Application of Temporal Fusion Transformer for Day-Ahead PV Power Forecasting. *Energies* **2022**, *15*, 5232. [\[CrossRef\]](#)
25. Wu, B.; Wang, L.; Zeng, Y.R. Interpretable wind speed prediction with multivariate time series and temporal fusion transformers. *Energy* **2022**, *252*, 123990. [\[CrossRef\]](#)
26. Zhang, H.; Zou, Y.; Yang, X.; Yang, H. A temporal fusion transformer for short-term freeway traffic speed multistep prediction. *Neurocomputing* **2022**, *500*, 329–340. [\[CrossRef\]](#)
27. Guo, T.; Lin, T.; Antulov-Fantulin, N. Exploring interpretable lstm neural networks over multi-variable data. In Proceedings of the International Conference on Machine Learning, PMLR, Long Beach, CA, USA, 9–15 June 2019; pp. 2494–2504.
28. Song, H.; Rajan, D.; Thiagarajan, J.; Spanias, A. Attend and diagnose: Clinical time series analysis using attention models. In Proceedings of the AAAI Conference on Artificial Intelligence, New Orleans, LA, USA, 2–7 February 2018; Volume 32.
29. Wen, R.; Torkkola, K.; Narayanaswamy, B.; Madeka, D. A multi-horizon quantile recurrent forecaster. *arXiv* **2017**, arXiv:1711.11053.
30. Lim, B.; Arik, S.Ö.; Loeff, N.; Pfister, T. Temporal fusion transformers for interpretable multi-horizon time series forecasting. *Int. J. Forecast.* **2021**, *37*, 1748–1764. [\[CrossRef\]](#)
31. Makridakis, S.; Spiliotis, E.; Assimakopoulos, V. The M4 Competition: 100,000 time series and 61 forecasting methods. *Int. J. Forecast.* **2020**, *36*, 54–74. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.