Forecasting Commodity Prices using Futures:

The case of copper.

December 2023

GONZALO CORTAZAR
Pontificia Universidad Católica de Chile
gcortaza@uc.cl

MARIAVICTORIA ENBERG
Pontificia Universidad Católica de Chile
mdenberg@uc.cl

HECTOR ORTEGA
Pontificia Universidad Católica de Chile
hiortega@uc.cl

Keywords: Forecasting, Copper prices; Futures; Expected Prices; Pricing Models

Forecasting Commodity Prices using Futures:

The case of copper.

Abstract

This paper analyzes three alternatives for forecasting commodity spot prices and applies them to predicting copper prices. The first is using futures prices from either LME or COMEX. Second, using analysts' expectations consensus, reported by Bloomberg. The third is jointly using futures and analysts' expectations as input to a stochastic model that smooths its data by applying the Kalman Filter. All three alternatives are compared with the well-known no-change forecast benchmark and among themselves. The results show that when futures prices are relatively higher than spot prices, the model presented is the best alternative for forecasting copper prices at any horizon up until 24 months. Also, when prices are relatively lower than spot prices, the model is the best alternative for long-term forecasts and the LME futures prices for 1 to 12 months.

Keywords: Forecasting, Copper prices; Futures; Expected Prices; Pricing Models

1. Introduction

Commodities have become increasingly relevant for the real economy and provide financial assets related to many economic sectors like energy, industrial metals, precious metals, and agriculture. Exposure to their risks can be managed through trades on the spot and derivative markets, on individual commodities, or an index representing an aggregate portfolio (Tang & Xiong, 2012; Boyd et al., 2018).

Forecasting the price of a given commodity is of the foremost interest to many economic agents. That is why many industry analysts regularly provide their price estimations for each commodity at different horizons. In addition to spot prices, futures contracts at different maturities are regularly traded.

This paper analyzes different alternatives for using these information sources to obtain the best possible copper price forecasts for up to 24 months. We start by providing an overview of copper's importance, relevant spot and derivative markets, and where it is traded. We then present some forecasting models found in the literature to conclude this introduction with an overview of our proposed forecasting method, which will be described in more detail in the following sections.

Copper, the chosen commodity to be analyzed in this paper, is essential nowadays due to its wide use in different industries. It is the world's third most used metal (following iron and aluminum), playing an indispensable role in several businesses in the financial environment (Wang et al., 2019).

Forecasting copper prices is relevant for different reasons. For instance, movements in copper prices can be seen as an early indicator of global economic performance, given the importance of copper in various industries such as transportation, telecommunications, and construction (Buncic

& Moretto, 2015). Also, some countries like Chile[1] and Zambia have become strongly dependent on copper prices (Sánchez Lasheras et al., 2015).

Moreover, the role of copper has evolved from being a commodity that is only used as a primary input in the production process of final goods to a financial asset that is also held and traded for speculative purposes (Buncic & Moretto, 2015). Thus, copper prices are becoming challenging to forecast, given the number and diversity of market participants such as producers, consumers, investors, and governments (García & Kristjanpoller, 2019).

Multiple models have been proposed in the literature to forecast copper prices. Different data have been used to input these models, including combinations of past spot prices, futures prices, and fundamental and non-fundamental variables. A wide variety of techniques and methods have been used. The simplest models use only past spot values to predict prices. For instance, the no-change forecast model, in which prices are assumed to follow a random walk with no drift, makes the current spot price the best forecast (Alquist et al., 2013). A more complex model is the wavelet-ARIMA model Kriechbaumer et al. (2014).

Cortazar et al. (2015) generate copper price forecasts, adding an estimation of the risk premium obtained from the CAPM model to the futures price.

Buncic and Moretto (2015) use a dynamic model averaging and selection approach to forecast copper prices. This method selects the predictor variables for a model chosen from three different groups: (i) fundamentals, (ii) financialization, and (iii) exchange rates and stock prices.

Sánchez Lasheras et al. (2015) propose two neural networks (multilayer perceptron neural network and Elman neural network). Chen et al. (2016) use a grey wave forecasting method to predict metal prices. Liu et al. (2017) predict copper prices using a machine learning algorithm. This method uses variables correlated with copper prices, such as gold, silver, crude oil, natural

[1] For example, in Chile this metal represents about half of Chilean exports and nearly 45% of Foreign Direct Investment (Brown & Hardy, 2019).

gas, lean hogs, coffee, the Dow Jones Index, and past copper prices. Dehghani & Bogdanovic (2018) propose a bat algorithm, and Dehghani (2018) uses an artificial neural network called gene expression programming.

Alameer et al. (2019) propose ten input variables as predictors for copper price fluctuations using a hybrid model. The model employs a genetic algorithm to adjust the adaptive neuro-fuzzy inference system (ANFIS) parameters. Wang et al. (2019) predict copper prices with a hybrid predictive technique combining complex and artificial neural network techniques.

This paper explores the forecasting performance of the Cifuentes et al. (2020) model. This model integrates analysts' forecasts and futures prices by proposing a three-factor stochastic model to estimate futures prices, expected prices, and the term structure of risk premiums for copper. The model uses both futures prices and analysts' expectations obtained from Bloomberg. Initially developed for estimating risk premiums, this model will now be studied in its forecasting ability for copper prices. The research hypothesis is that including futures price data outperforms using only analysts' forecasts of short and medium-term[2] copper prices.

The reference price to be forecasted is the London Metal Exchange (LME) copper price since this exchange is the primary international market for copper and provides appropriately located storage facilities to enable market participants to take or make physical deliveries (Dooley & Lenihan, 2005; Watkins & McAleer, 2004). Besides, it is the biggest futures exchange[3] for copper, handling more than half of world trades and a world reference for copper prices (Ciner et al., 2020; Li & Li, 2015). Futures prices are also obtained from the New York Commodity Exchange (COMEX) and analysts' expectations from Bloomberg (Cortazar et al., 2021).

[2] A 24-month horizon.
[3] The 3 biggest futures exchange are the London Metal Exchange (LME), the New York Commodity Exchange (COMEX) and the Shanghai Futures Exchange (SHFE).

Different performance metrics are used to analyze how good the proposed joint model is at forecasting copper prices, compared to using futures or analysts' expectations individually and to the no-change benchmark.

The paper is organized as follows. Section 2 presents the forecasting alternatives that will be compared. Section 3 describes the metrics under which model performance will be measured. Section 4 shows the data. Section 5 summarizes the results of the forecasting alternatives under each performance metric. Section 6 discusses the best forecasting alternatives under different price scenarios. Finally, Section 7 concludes.

2. Forecasting Models

In what follows, we present the forecasting alternatives for copper spot prices that will be compared.

2.1 No-change

The simplest benchmark for measuring the forecast performance of a given model is to compare it with the no-change forecast. This assumes that prices follow a no-drift, random walk, in which the best forecast is the current spot price.

$$\hat{S}_{t+h|t} = S_t$$

where $\hat{S}_{t+h|t}$ is the prediction of the spot price in h periods and S_t is the current spot price.

2.2 Futures

Another forecasting alternative is to use the futures price as an unbiased forecast (Cortazar et al., 2015). This assumes there are no relevant risk premiums. We will test two alternative implementations, using futures from the London Metal Exchange (LME) or the New York Commodity Exchange (COMEX).

2.3 The Analysts' Expectations-Consensus

Analysts from different institutions provide expectations for some quarters and years ahead. Bloomberg delivers each of these predictions. Given that various analysts provide forecasts, they are very volatile. To summarize these predictions, Bloomberg offers the median of the analysts' expectations for each horizon in what they call the Consensus.

2.4 The proposed Model (Joint Futures and Analysts' Expectations)

The proposed model is a joint futures and analysts' expectations model used to estimate the term structure of risk premiums for copper in Cifuentes et al. (2020) and for oil in Cortazar et al. (2022) and to forecast oil prices in Cortazar et al. (2021).

We now propose to use this model to analyze its copper price forecasting performance compared to the other alternatives presented before that use similar input data but take each one individually.

2.4.1 Model definition

Let S_t be the spot price at time t, then:

$$\ln S_t = Y_t = h'x_t$$

$$dx_t = \left(-Ax_t + \begin{bmatrix} b_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \right) dt + dw_t$$

where h is an $n\ x\ 1$ vector of constants, x_t is an $n\ x\ 1$ vector of state variables, b_1 is a scalar, A is an $n\ x\ n$ upper triangular matrix with its first diagonal element being zero and the remaining elements all different and strictly positive. Let dw_t be a $n\ x\ 1$ vector of uncorrelated Brownian motions such that:

$$dw_t dw_t' = I\ dt$$

where I is an $n\ x\ n$ identity matrix.

Let Π_t be the commodity risk premium at time t. We assume that:

$$\Pi_t = \lambda + \Lambda\, x_t$$

Hence, the risk-adjusted version of the model is:

$$Y_t = h' x_t$$

$$dx_t = \left(-(A+\Lambda)x_t + \begin{bmatrix} b_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} - \lambda \right) dt + dw_t^Q$$

where dw_t^Q is a Brownian motion under the risk-neutral measure Q, λ is an $n\ x$ 1 vector and Λ is an $n\ x\ n$ matrix that needs no additional condition.

The expected price under the risk-adjusted (futures) and under the historical process are:

$$F_t(\mathrm{T}) = E_t^{\mathrm{Q}}(x_T) = e^{-(A+\Lambda)(T-t)} x_t + \left(\int_0^{T-t} e^{-(A+\Lambda)\tau} d\tau \right) (b - \lambda)$$

$$E_t(x_T) = e^{-A(T-t)} x_t + \left(\int_0^{T-\mathrm{t}} e^{-A\tau} d\tau \right) b$$

The implicit volatilities of futures σ_F and expected prices σ_E are:

$$\sigma_F = \sqrt{h' e^{-(A+\Lambda)(T-t)} e^{-(A+\Lambda)(T-t)'} h}$$

$$\sigma_E = \sqrt{h' e^{-A(T-t)} e^{-A(T-t)'} h}$$

2.4.2 Model estimation

The state variables and the model's parameters are estimated using the Kalman filter (Kalman, 1960). This method uses all the available data in each iteration to estimate the state variables' optimal value, defined by the measurement and the transition equations.

The measurement equation indicates the relationship between the observable variable vector z_t and the state variable vector x_t, as follows:

$$z_t = \mathrm{H_t} x_t + d_t + v_t \qquad v_t \sim N(0, R_t)$$

where z_t is an $m_t \, x \, 1$ vector that contains logarithm of price observations (futures and expected spot prices) at time t. H_t is an $m_t \, x \, n$ matrix, x_t is an $n \, x \, 1$ vector, d_t is an $m_t \, x \, 1$ vector, and, v_t is a measurement error vector of $m_t \, x \, 1$ dimension with zero mean and covariance R_t. In the model, m_t depends on the number of observations at each time. Thus, the dimension of $z_t, H_t, d_t, v_t,\ y\ R_t$ can vary in each iteration.

The expected spot prices, proxied by the analysts' expectations, are nosier than futures prices, so there will be two measurement errors, and the matrix R_t is defined by:

$$R_t = \begin{bmatrix} {\sigma_f}^2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & {\sigma_f}^2 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & {\sigma_e}^2 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & {\sigma_e}^2 \end{bmatrix}$$

The second equation is the transition equation, which describes the stochastic process that the state variables follow:

$$x_{t+1} = \bar{A}x_t + \bar{c} + w_t \qquad w_t \sim N(0, Q)$$

where $\bar{A}$ is an $n \, x \, n$ matrix, and $\bar{c}$ is an $n \, x \, 1$ vector. $\bar{A}$ and $\bar{c}$ represent the discretization of the process. In the above expression w_t is a vector of random variables with zero mean and an $n \, x \, n$ covariance matrix Q.

The parameters of the model are estimated by maximum likelihood.

3. Performance metrics

The above forecasting alternatives will be compared to the no-change forecast, which assumes prices follow a random walk. Thus, each of the other three models will be first compared with this no-change benchmark and later between them.

We use three performance metrics to measure each model's forecasting accuracy. These are Root Mean Squared Error, Relative Mean Squared Prediction Error, and Dstat. We now provide a brief description of each one.

3.1 Root Mean Squared Error

The first performance metric is to compute the root of the squared errors. This metric is used in many studies on copper forecasting (Hesam et al., 2018; Kriechbaumer et al., 2014; Wets & Rios, 2015). This metric is calculated as follows:

$$RMSE^h = \sqrt{\frac{1}{N}\sum_{t}^{N}\left(S_{t+h|t} - \hat{S}_{i,t+h|t}\right)^2}$$

3.2 Relative Mean Squared Prediction Error

A second metric is the relative mean squared prediction error, which divides the mean squared prediction error of the model by that of the no-change forecast error.

The Relative Mean Squared Prediction Error (Watson & Stock, 2004) is defined as:

$$Relative\ MSPE_i = \frac{\sum_t^N\left(\mathrm{S_{t+h|t}} - \hat{S}_{i,t+h|t}\right)^2}{\sum_t^N\left(\mathrm{S_{t+h|t}} - \hat{S}_{0,t+h|t}\right)^2}$$

where i is the forecasting model analyzed, and i=0 refers to the no-change benchmark.

3.3 Dstat

Finally, the most straightforward forecasting metric is the directional prediction, i.e., if prices will go up or down at a given horizon.

The Directional Change statistic (Yao & Tan, 2000) is calculated as follows:

$$D_{stat}^h = \frac{1}{N}\sum_{t\in T} a_{t,h}$$

where:

$$a_{t,h} = \begin{cases} 1 & if\ (S_{t+h} - S_t)(\hat{S}_{t+h} - S_t) > 0 \\ 0 & otherwise \end{cases}$$

A D_{stat}^{h} greater than 0.5 means that the obtained prediction is better than the no-change model, which is expected to have a D_{stat}^{h} 0.5.

4. Data

In this section, we describe the data that will be used to compute the performance of the alternative forecasts. It consists of spots, futures, analysts' expectations, and Bloomberg's consensus expectations from January 2010 to December 2020.

4.1 Spot prices

We will use the London Metal Exchange-LME cash prices as the spot prices to be forecasted. The LME is the largest copper trading market and a well-recognized world reference for prices of this commodity (Ciner et al., 2020; Park & Lim, 2018; Li & Li, 2015; Dooley & Lenihan, 2005; Watkins & McAleer, 2004). Forecasts will be done yearly, out-of-sample, for the following 24 months.

To illustrate how variable these prices are, Figure 4.1 plots them from January 2010 to December 2020.

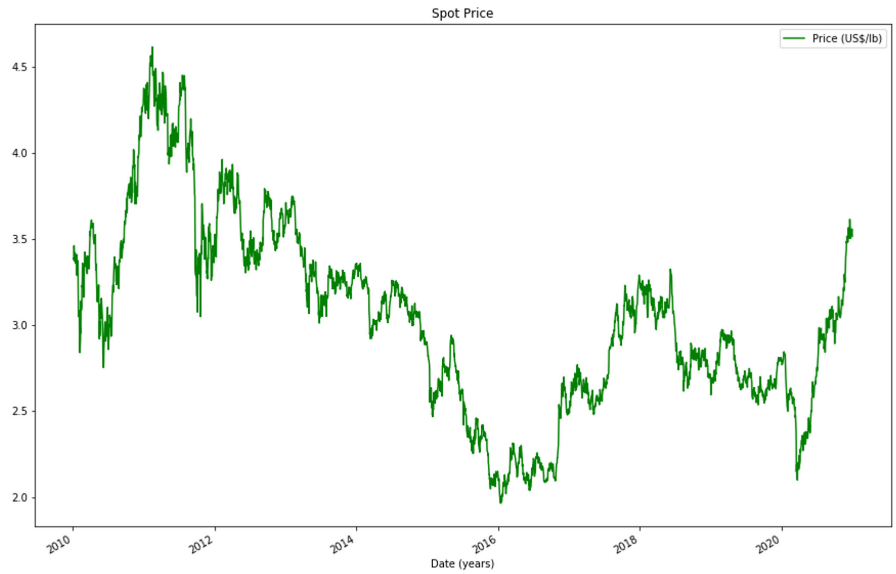

Figure 4.1: Copper spot prices from January 2010 to December 2020.

4.2 Futures prices

The two primary sources of copper futures prices are LME in the UK and COMEX in the USA. In the LME exchange, futures expire at the current month and for the following 123 months. In the COMEX exchange, futures expire at the current month, the next 23 months, and any March, May, July, September, and December within 60 months.

We will use futures prices for two purposes: first, as one of the inputs to the proposed Model (joint futures and analysts' expectations) described before. Following Cifuentes et al. (2020), we use LME weekly futures prices for the contract closest to its maturity and those maturing every six months. Figure 4.2 shows the LME weekly copper futures prices from January 2010 to December 2020, and Table 4.1 summarizes the LME data used as input to the proposed model.

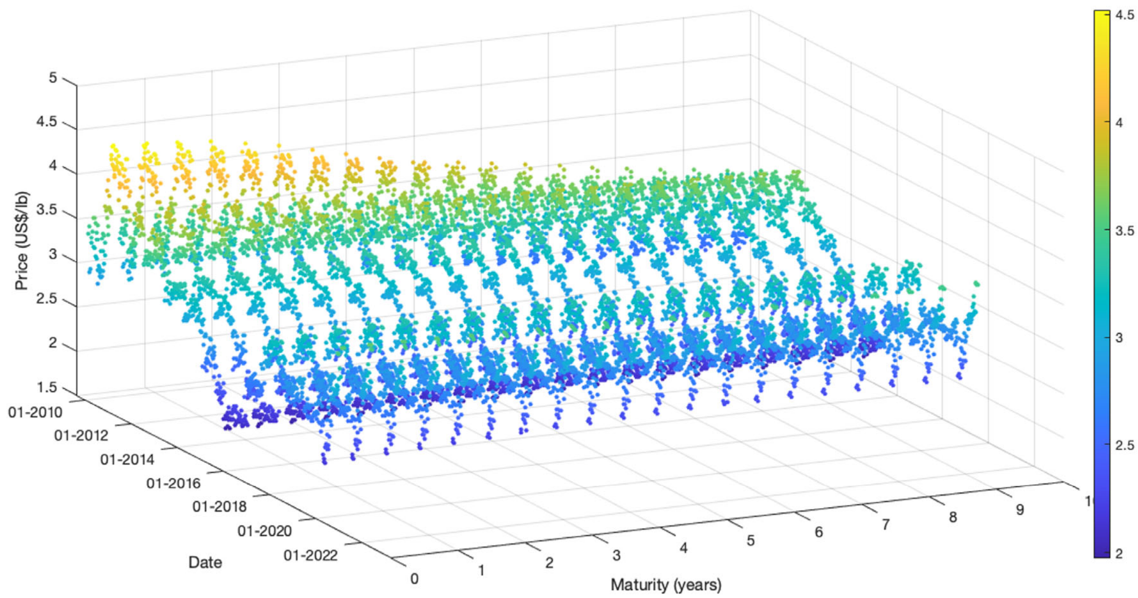

Figure 4.2: LME weekly copper futures prices from January 2010 to December 2020

Table 4.1: LME weekly Copper Futures Prices used as input to the model (in sample), grouped by year.

Year	Amount of data
2010	1040
2011	1040
2012	1040
2013	1040
2014	1060
2015	1040
2016	1040
2017	1040
2018	1040
2019	1040
Average	1042

A second way of using futures data is as the forecast for the next 24 months of the spot prices. The following section compares forecasting performance using LME or COMEX futures contracts. The LME futures used as the spot price forecast include weekly data for contracts with maturities up to 24 months. The following figure and table describe this data.

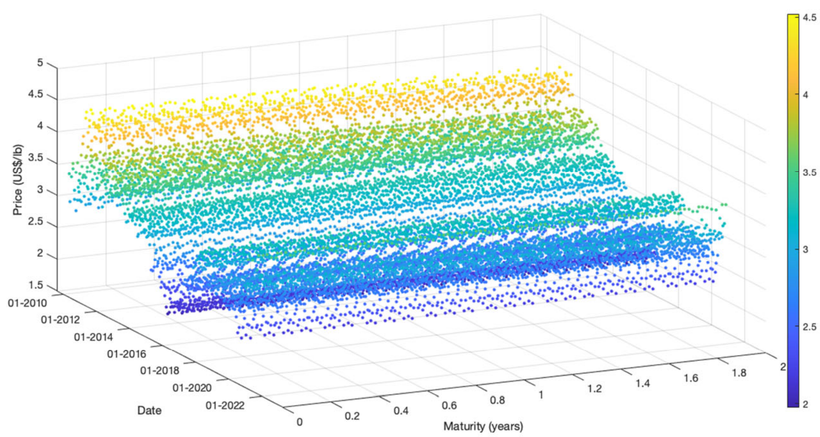

Figure 4.3: LME weekly copper futures prices from January 2010 to December 2020

Table 4.2: LME weekly Futures Prices up to 24 months (in sample) by year.

Year	Amount of data
2010	1243
2011	1243
2012	1245
2013	1245
2014	1266
2015	1243
2016	1244
2017	1244
2018	1245
2019	1243
Average	1246.1

As mentioned, COMEX futures prices may also be used as the spot price forecast. Figure 4.4 and Table 4.3 present this data.

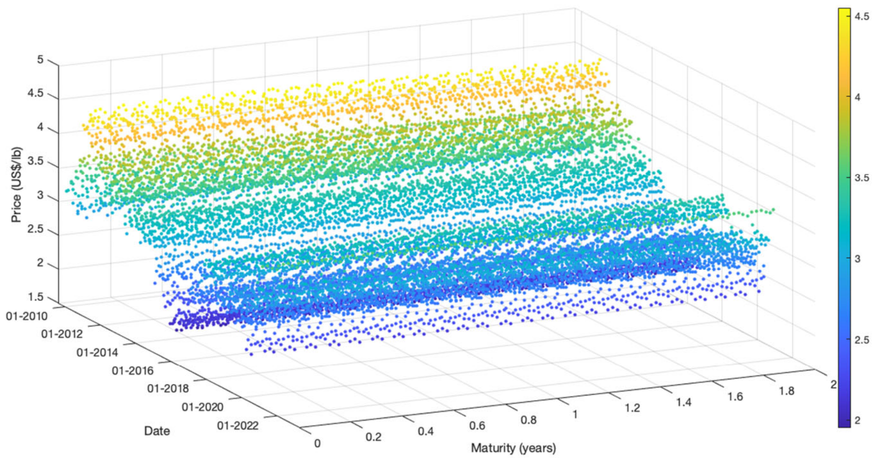

Figure 4.4: COMEX weekly copper futures prices from January 2010 to December 2020

Table 4.3: COMEX weekly Futures Prices up to 24 months (in sample) by year.

Year	Amount of data
2010	1250
2011	1249
2012	1236
2013	1231
2014	1258
2015	1248
2016	1251
2017	1251
2018	1273
2019	1286
Average	1253.3

4.3 The Analysts' Expectations-Consensus

Bloomberg reports the forecasts made by various analysts from different financial institutions. There are two types of forecasts: quarterly and annual. These predictions are made for the average price each quarter or year. Following Cifuentes et al. (2020), they represent the price in the middle of their period.

Quarterly forecasts are available for the current quarter and for the following five quarters. Annual forecasts are valid for the year the forecast is made and for the next four years. These forecasts are not available on a previously defined schedule. Analysts can forecast some, all, or none of these horizons at any given date. All forecasts in the same week and for the same horizon are averaged to obtain weekly analysts' expectation data.

Figure 4.5 shows all weekly analysts' expectations data, while Table 4.4 summarizes the analysts' weekly expectations data for up to 24 months.

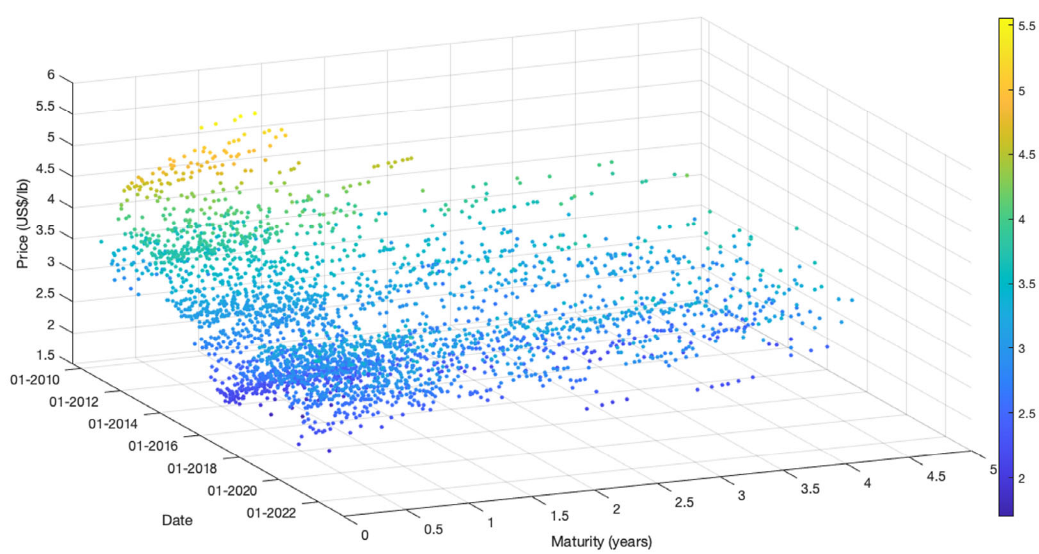

Figure 4.5: Analysts' Expectations Weekly data from January 2010 to December 2020

Table 4.4: Analysts' Expectations up to 24 months (in sample), by year

Year	Amount of data
2010	240
2011	278
2012	344
2013	621
2014	711
2015	740
2016	783
2017	746
2018	561
2019	331
Average	535.5

Given the variety of analysts, horizons, and dates, data is particularly volatile and difficult to use directly. Figure 4.6 shows how volatile the analysts' data is during the week.

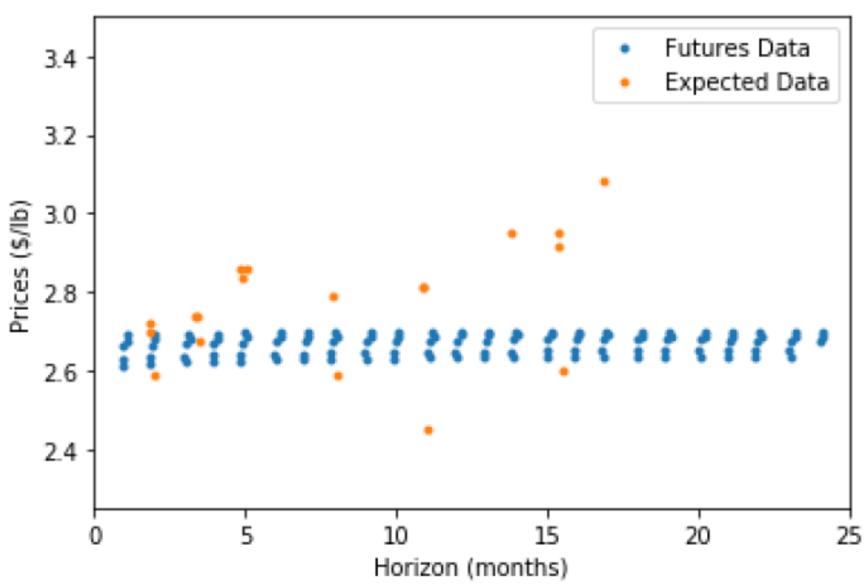

Figure 4.6: Futures and analysts' expected price data, third week, March 2017

It becomes clear that to use this data for forecasting effectively, some smoothing must be done.

We consider two ways of processing this data. The first one uses what Bloomberg reports as the consensus, the median of the available analyst forecasts for each horizon on a given week. Figure 4.7 and Table 4.5 present this data.

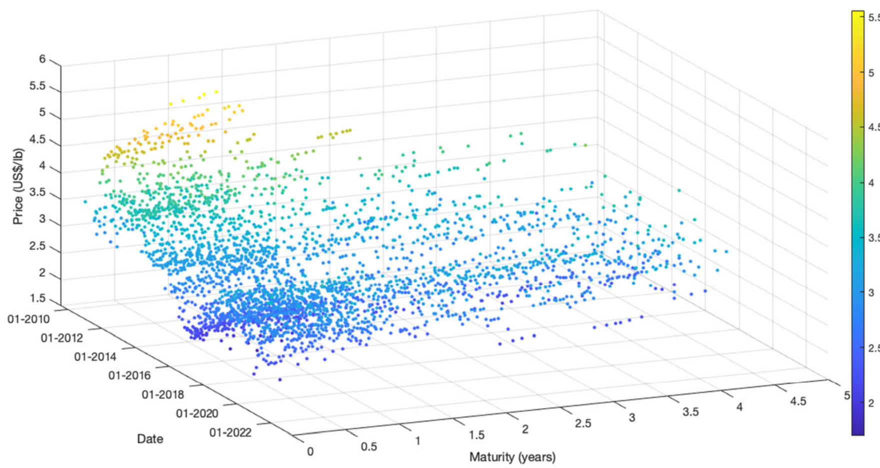

Figure 4.7: Bloomberg's weekly Consensus expectations

Table 4.5: Bloomberg's weekly Consensus expectations (in sample) by year

Year	Amount of data
2010	411
2011	460
2012	473
2013	490
2014	458
2015	453
2016	466
2017	425
2018	391
2019	407
Average	443.4

The second way of smoothing this volatile data is by applying the Kalman Filter to all data shown in Figure 4.5 when calibrating our proposed model.

5. Forecasting Results

In what follows, we summarize the out-of-sample results of weekly applying the alternative forecasts to 1-to-24-month horizons from 2014 to 2020[4]. Each result is ranked against the standard no-change forecast benchmark, boldfacing the best.

5.1 Forecasting Results Using Futures

The forecast for any given horizon is calculated as the weighted average of the prices of the two futures with maturities closest to the horizon. Tables 5.1, 5.2, and 5.3 present the RMSE, Relative MSPE, and Dstat metrics using futures when implemented with the LME or the COMEX data.

Table 5.1 RMSE Performance Metric using Futures from 2014 to 2020, implemented with LME, COMEX. The "BEST" column shows the best model for each horizon.

Horizon (months)	LME	COMEX	**BEST**
1	**0.147**	0.150	**LME**
3	**0.262**	**0.265**	**LME**
6	**0.352**	**0.354**	**LME**
9	**0.407**	**0.410**	**LME**
12	**0.494**	**0.498**	**LME**
15	**0.551**	**0.557**	**LME**
18	**0.605**	**0.613**	**LME**
21	**0.642**	0.651	**LME**
24	**0.664**	0.671	**LME**
Horizons up to 12 months	**0.354**	**0.3577**	**LME**
Horizons between 13 and 24 months	**0.604**	0.611	**LME**
Horizons up to 24 months	**0.485**	**0.490**	**LME**

The boldface indicates that the model performs better than the no-change benchmark.

[4] For the year 2020 forecasting errors are computed only for horizons up to 12 months

Table 5.2 Relative MSPE Performance Metric using Futures from 2014 to 2020 implemented with LME, COMEX. The "BEST" column shows the best model for each horizon.

Horizon (months)	LME	COMEX	**BEST**
1	**0.987**	1.018	**LME**
3	**0.978**	**0.997**	**LME**
6	**0.972**	**0.984**	**LME**
9	**0.965**	**0.975**	**LME**
12	**0.968**	**0.985**	**LME**
15	**0.971**	**0.992**	**LME**
18	**0.971**	**0.997**	**LME**
21	**0.983**	1.009	**LME**
24	**0.988**	1.010	**LME**
Horizons up to 12 months	**0.970**	**0.983**	**LME**
Horizons between 13 and 24 months	**0.976**	1.000	**LME**
Horizons up to 24 months	**0.974**	**0.995**	**LME**

The boldface indicates that the model performs better than the no-change benchmark.

Table 5.3 Dstat Performance Metric using Futures 2014-2020 implemented with LME, COMEX. The "BEST" column shows the best model for each horizon.

Horizon (months)	LME	COMEX	**BEST**
1	**0.587**	**0.507**	**LME**
3	**0.571**	**0.503**	**LME**
6	**0.644**	**0.559**	**LME**
9	**0.606**	**0.554**	**LME**
12	**0.589**	**0.510**	**LME**
15	**0.570**	**0.563**	**LME**
18	**0.564**	**0.610**	**COMEX**
21	**0.533**	**0.628**	**COMEX**
24	**0.544**	**0.659**	**COMEX**
Horizons up to 12 months	**0.599**	**0.526**	**LME**
Horizons between 13 and 24 months	**0.558**	**0.600**	**COMEX**
Horizons up to 24 months	**0.580**	**0.560**	**LME**

The boldface indicates that the model performs better than the no-change benchmark.

Several conclusions may be obtained from the above tables. First, futures prices provide a better forecast for most metrics and horizons than the no-change benchmark. Second, LME futures give better forecasts than the no-change benchmark for all metrics and horizons. Lastly, LME futures provide better forecasts than COMEX futures for almost all horizon and performance metrics. This result is not surprising, given that the spot price to be forecasted is the cash price from the LME exchange.

If, for any reason, LME futures are not available, COMEX futures provide better forecasts than those of the no-change benchmark for some metrics, especially for horizons from 3 to 18 months.

5.2 Forecasting Results Using Analysts' Expectations-Consensus

As discussed, analysts' expectations are very volatile, so some smoothing is required. In this section, we analyze the performance using the consensus of the analysts' expectations, defined by the median of the analysts' predictions, as reported by Bloomberg.

The forecast for any given horizon is calculated as the weighted average of the two analyst consensus forecasts closest to the horizon.

Table 5.4 shows that consensus forecasts are better than the no-change benchmark for all performance metrics for only 21- and 24-month horizons. The Dstat metric, which indicates if prices are going up or down, provides good results, especially for longer horizons.

In summary, using analysts' forecasts, represented by the consensus provided by Bloomberg, gives mixed results, thus shedding some doubts on the value of considering analysts' expectations as a valuable data source. This preliminary conclusion, however, will be revised in the next section.

Table 5.4 Performance Metrics using Analysts' Expectations-Consensus from 2014 to 2020

Horizon (months)	RMSE	Relative MSPE	Dstat
1	0.230	2.392	**0.565**
3	0.321	1.466	**0.545**
6	0.409	1.317	0.462
9	0.477	1.322	0.450
12	0.553	1.215	0.494
15	0.600	1.151	**0.533**
18	0.623	1.031	**0.592**
21	**0.642**	**0.981**	**0.620**
24	**0.611**	**0.839**	**0.663**
Horizons up to 12 months	0.416	1.340	0.493
Horizons between 13 and 24 months	0.617	1.019	**0.586**
Horizons up to 24 months	0.518	1.113	**0.536**

The boldface indicates that the model performs better than the no-change benchmark.

5.3 Forecasting Results for the Model (Joint Futures and Analysts' Expectations)

In this section, we explore the value of using analysts' expectations, jointly with futures prices, in a forecasting model calibrated using the Kalman Filter.

5.3.1 Model Fit

The model must be calibrated several times to provide out-of-sample spot copper forecasts. The first data set used to calibrate model parameters includes prices from 2010 to 2013, which are then used to forecast prices during 2014 for the next 24 months. Then, one year is added to the calibration data set, parameters are estimated, and forecasts are done during 2015. This process continues until the last data set covers prices from 2010 to 2019 to make forecasts for 2020. The model uses all the available futures and analysts' expectation data to jointly estimate the expected and the futures curves.

Figure 5.1 presents the expected and futures curves, and data, for the third week of March of 2017. It can be seen that the expected curve does not perfectly fit the available data because of its volatility. However, the Kalman Filter considers the data for that week and all the past data, providing a smooth expectation curve for each date. On the other hand, the futures curve fits the data much better because it is less volatile.

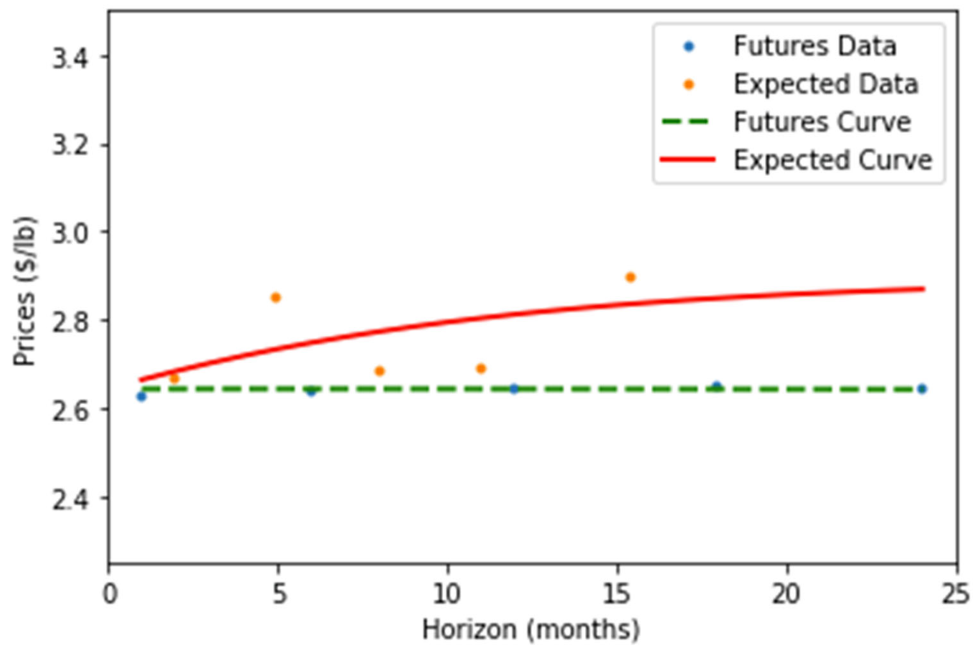

Figure 5.1: Expected and futures curves and data, third week, March 2017

Tables 5.5 and 5.6 compute the Mean Absolute Percentage Error (MAPE) for the in-sample and out-of-sample data. It can be seen that the futures curves have a better fit, as expected. Also, the average MAPE is similar between the in-sample and the out-of-sample calibrations.

Table 5.5: Mean Absolute Percentage Error of expected and futures curves (in-sample)

Calibration Years	MAPE (%) between	
	Curve and Futures Prices Data	Curve and Analysts Expected Prices Data
2010-2013	0.21%	6.77%
2010-2014	0.21%	6.62%
2010-2015	0.19%	6.59%
2010-2016	0.18%	6.71%
2010-2017	0.18%	6.86%
2010-2018	0.18%	6.88%
2010-2019	0.17%	6.85%
Average	0.19%	6.76%
Standard Deviation	0.01%	0.12%

Table 5.6: Mean Absolute Percentage Error of expected and futures curves (out-of-sample). The "Year" column shows the out-of-sample year for each Calibration Year.

Calibration Years	Year	MAPE (%) between	
		Curve and Futures Prices Data	Curve and Analysts Expected Prices Data
2010-2013	2014	0.29%	5.36%
2010-2014	2015	0.36%	6.25%
2010-2015	2016	0.11%	11.51%
2010-2016	2017	0.18%	8.13%
2010-2017	2018	0.20%	5.67%
2010-2018	2019	0.11%	7.14%
2010-2019	2020	0.13%	5.10%
Average		0.19%	7.02%
Standard Deviation		0.10%	2.25%

5.3.2 Model Results

Table 5.7 shows that using both futures and analysts' expectations as inputs in the proposed model provides better forecasts than the no-change benchmark under all metrics for horizons of 1 month and 12 to 24 months. In addition, this holds for the Dstat metric for all horizons.

Table 5.7 Performance Metrics for the Model from 2014 to 2020

Horizon (months)	RMSE	Relative MSPE	Dstat
1	**0.147**	**0.978**	**0.579**
3	0.265	1.001	**0.542**
6	0.362	1.028	**0.529**
9	0.415	1.003	**0.523**
12	**0.472**	**0.885**	**0.513**
15	**0.494**	**0.780**	**0.567**
18	**0.509**	**0.688**	**0.620**
21	**0.516**	**0.634**	**0.675**
24	**0.507**	**0.576**	**0.751**
Horizons up to 12 months	**0.356**	**0.981**	**0.537**
Horizons between 13 and 24 months	**0.505**	**0.684**	**0.629**
Horizons up to 24 months	**0.431**	**0.770**	**0.579**

The boldface indicates that the model performs better than the no-change benchmark.

6. Comparing Forecasting Alternatives

In this section, we analyze the forecasting performance of the three alternatives using futures prices, consensus expectations, and the model. As a robustness check, we split the out-of-sample data into two parts.

Given that it is well-known that, depending on market and inventory conditions, futures prices behave differently, sometimes in contango and others in backwardation, we consider this when dividing the data. We must consider, however, that these two price regimes occur with different frequencies.

In order to generate the two files, we split the out-of-sample futures depending on whether the difference between futures and spot prices is high or low compared to the median of the corresponding in-sample data. Thus, the out-of-sample data is divided into two parts: when "futures are high", which means that futures are in a relative contango, and when "futures are low" or in a relative backwardation.

6.1 Forecasting when "futures are high" (*relative contango*)

Tables 6.1 and 6.2 show that when futures prices are relatively high, compared to spot prices (*relative contango*), the model performs much better than the alternatives for all horizons regarding RMSE and Relative MSPE. This shows that analysts' expected prices, when used as input with futures prices, are valuable for forecasting purposes.

On the other hand, Table 6.3 shows that for forecasting the direction of price movements, using detailed analysts' expectations is not valuable. In this case, it is better to use LME futures prices for short-term horizons and consensus forecasts for long horizons.

Table 6.1: RMSE for the Model, LME, and Consensus from 2014 to 2020-High Futures. The "Best" column shows the best alternative for each horizon.

Horizon (months)	Model	LME	Consensus	Best
1	**0.153**	0.154	0.231	**Model**
3	**0.264**	0.279	0.316	**Model**
6	**0.364**	0.388	0.428	**Model**
9	**0.393**	0.417	0.461	**Model**
12	**0.420**	0.472	0.507	**Model**
15	**0.429**	0.508	0.520	**Model**
18	**0.396**	0.529	0.479	**Model**
21	**0.425**	0.578	0.516	**Model**
24	**0.429**	0.613	0.486	**Model**
Horizons up to 12 months	**0.345**	0.370	0.411	**Model**
Horizons between 13 and 24 months	**0.424**	0.552	0.508	**Model**
Horizons up to 24 months	**0.387**	0.470	0.462	**Model**

The boldface indicates the best-performing alternative.

Table 6.2: Relative MSPE for the Model, LME, and Consensus from 2014 to 2020- High Futures. The "Best" column shows the best alternative for each horizon.

Horizon (months)	Model	LME	Consensus	Best
1	**0.983**	0.989	2.237	**Model**
3	**0.881**	0.984	1.263	**Model**
6	**0.862**	0.982	1.191	**Model**
9	**0.880**	0.989	1.208	**Model**
12	**0.789**	0.997	1.148	**Model**
15	**0.716**	1.005	1.050	**Model**
18	**0.565**	1.011	0.831	**Model**
21	**0.555**	1.029	0.821	**Model**
24	**0.506**	1.032	0.650	**Model**
Horizons up to 12 months	**0.858**	0.988	1.218	**Model**
Horizons between 13 and 24 months	**0.599**	1.016	0.859	**Model**
Horizons up to 24 months	**0.680**	1.007	0.972	**Model**

The boldface indicates the best-performing alternative.

Table 6.3: Dstat for the Model, LME, and Consensus from 2014 to 2020-High Futures. The "Best" column shows the best alternative for each horizon.

Horizon (months)	Model	LME	Consensus	Best
1	0.576	**0.587**	0.565	**LME**
3	0.552	**0.571**	0.545	**LME**
6	0.593	**0.644**	0.462	**LME**
9	0.534	**0.606**	0.450	**LME**
12	0.495	**0.589**	0.494	**LME**
15	0.521	**0.570**	0.533	**LME**
18	0.582	0.564	**0.592**	**Consensus**
21	0.605	0.533	**0.620**	**Consensus**
24	**0.696**	0.544	0.663	**Model**
Horizons up to 12 months	0.549	**0.599**	0.493	**LME**
Horizons between 13 and 24 months	0.585	0.558	**0.586**	**Consensus**
Horizons up to 24 months	0.567	**0.580**	0.536	**LME**

The boldface indicates the best-performing alternative.

6.2 Forecasting when "futures are low" (*relative backwardation*)

Tables 6.4, 6.5, and 6.6 present the performance results of the three alternatives when the futures prices are relatively low compared to spot prices (relative backwardation). Results are very consistent for all three performance metrics, showing that, in this case, it is better to use LME futures prices for short-term horizons (until 12 months) and the model for longer horizons.

Table 6.4: RMSE for the Model, LME, and Consensus from 2014 to 2020- Low Futures. The "Best" column shows the best alternative for each horizon.

Horizon (months)	Model	LME	Consensus	Best
1	**0.140**	0.141	0.228	**Model**
3	0.267	**0.241**	0.327	**LME**
6	0.358	**0.292**	0.382	**LME**
9	0.444	**0.394**	0.499	**LME**
12	0.537	**0.523**	0.613	**LME**
15	**0.586**	0.615	0.714	**Model**
18	**0.689**	0.739	0.849	**Model**
21	**0.691**	0.778	0.877	**Model**
24	**0.674**	0.786	0.866	**Model**
Horizons up to 12 months	0.370	**0.332**	0.424	**LME**
Horizons between 13 and 24 months	**0.642**	0.698	0.797	**Model**
Horizons up to 24 months	**0.494**	0.507	0.597	**Model**

The boldface indicates the best-performing alternative.

Table 6.5: Relative MSPE for the Model, LME, and Consensus from 2014 to 2020-Low Futures. The "Best" column shows the best alternative for each horizon.

Horizon (months)	Model	LME	Consensus	Best
1	**0.972**	0.984	2.583	**Model**
3	1.187	**0.968**	1.781	**LME**
6	1.428	**0.948**	1.622	**LME**
9	1.185	**0.930**	1.490	**LME**
12	0.989	**0.937**	1.287	**LME**
15	**0.848**	0.935	1.258	**Model**
18	**0.808**	0.931	1.227	**Model**
21	**0.732**	0.926	1.178	**Model**
24	**0.680**	0.923	1.120	**Model**
Horizons up to 12 months	1.170	**0.941**	1.531	**LME**
Horizons between 13 and 24 months	**0.786**	0.929	1.210	**Model**
Horizons up to 24 months	**0.886**	0.932	1.293	**Model**

The boldface indicates the best-performing alternative.

Table 6.6: Dstat for the Model, LME, and Consensus from 2014 to 2020- Low Futures.

The "Best" column shows the best alternative for each horizon.

Horizon (months)	Model	LME	Consensus	Best
1	0.582	**0.587**	0.565	**LME**
3	0.531	**0.571**	0.545	**LME**
6	0.440	**0.644**	0.462	**LME**
9	0.507	**0.606**	0.450	**LME**
12	0.538	**0.589**	0.494	**LME**
15	**0.643**	0.570	0.533	**Model**
18	**0.699**	0.564	0.592	**Model**
21	**0.848**	0.533	0.620	**Model**
24	**0.900**	0.544	0.663	**Model**
Horizons up to 12 months	0.522	**0.599**	0.493	**LME**
Horizons between 13 and 24 months	**0.723**	0.558	0.586	**Model**
Horizons up to 24 months	**0.600**	0.580	0.536	**Model**

The boldface indicates the best-performing alternative.

7. Conclusions

We have argued that forecasting copper prices is relevant for many agents, including investors and governments. However, research is still underway to find models and data sources that could be more useful in this endeavor.

This paper presents three alternatives to forecast spot prices for horizons 1 to 24 months. First, futures prices were used, and this alternative was implemented using either LME or COMEX futures. We concluded that, in this case, it was better to use LME futures.

Second, we considered using analysts' expectations and discussed how volatile this data is, stating the convenience of using some smoothing process. For this alternative, we initially used the consensus expectations, reported by Bloomberg as the median of the available data for a given horizon.

The third alternative presented was to jointly consider futures and analysts' expectations as input to a model that smooths data using the Kalman Filter. All three alternatives, Futures, Consensus, and Model, were compared with the well-known no-change forecast benchmark and among themselves under different price scenarios.

The main conclusions that can be drawn from these exercises are the following. First, analysts' expectations are a valuable source of data that can be useful for forecasting copper prices. Second, being this data very volatile, smoothing by using Bloomberg´s Consensus data (which provides the median forecast) is not helpful. Third, when futures prices are relatively higher than spot prices (compared with recent history), the presented model is the best alternative for forecasting copper prices at any horizon between 1 and 24 months. Fourth, when prices are relatively lower than spot prices (compared with recent history), the model is the best alternative for long-term forecasts and the LME futures prices for 1 to 12 months.

During the preparation of this work the author(s) used Grammarly in order to improve readability- After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

REFERENCES

Alameer, Z., Elaziz, M. A., Ewees, A. A., Ye, H., & Jianhua, Z. (2019). Forecasting Copper Prices Using Hybrid Adaptive Neuro-Fuzzy Inference System and Genetic Algorithms. *Natural Resources Research*, *28*(4), 1385–1401. https://doi.org/10.1007/s11053-019-09473-w

Alquist, R., Kilian, L., & Vigfusson, R. J. (2013). *Forecasting the price of oil*. *Handbook of Economic Forecasting* (Vol. 2). Elsevier B.V. https://doi.org/10.1016/B978-0-444-53683-9.00008-6

Boyd, N.E., Harris, J.H., Li, B. (2018). An update on speculation and financialization in commodity markets. *Journal of Commodity Markets* 10, 91–104.

Brown, P. P., & Hardy, N. (2019). Forecasting base metal prices with the Chilean exchange rate. *Resources Policy*, *62*(April), 256–281. https://doi.org/10.1016/j.resourpol.2019.02.019

Buncic, D., & Moretto, C. (2015). Forecasting copper prices with dynamic averaging and selection models. *North American Journal of Economics and Finance*, *33*, 1–38. https://doi.org/10.1016/j.najef.2015.03.002

Chen, Y., He, K., & Zhang, C. (2016). A novel grey wave forecasting method for predicting metal prices. *Resources Policy*, *49*, 323–331. https://doi.org/10.1016/j.resourpol.2016.06.012

Cifuentes, S., Cortazar, G., Ortega, H., & Schwartz, E. S. (2020). Expected prices, futures prices and time-varying risk premiums: The case of copper. *Resources Policy*, *69*(July), 1–15. https://doi.org/10.1016/j.resourpol.2020.101825

Ciner, C., Lucey, B., & Yarovaya, L. (2020). Spillovers, integration and causality in LME non-ferrous metal markets. *Journal of Commodity Markets*, *17*, 100079

Cortazar, G., Kovacevic, I., & Schwartz, E. S. (2015). Expected commodity returns and pricing models. *Energy Economics*, *49*, 60–71. https://doi.org/10.1016/j.eneco.2015.01.015

Cortazar, G., Liedtke, P., Ortega, H., and Schwartz, E. (2022). Time-Varying Term Structure of Oil Risk Premiums. *The Energy Journal*, Vol.43, No 5.

Cortazar, G., Ortega, H., & Valencia, C. (2021). How good are analyst forecasts of oil prices ? *Energy Economics,* Vol 102, (October), Article 105500

Dehghani, H. (2018). Forecasting copper price using gene expression programming. *Journal of Mining and Environment*, *9*(2), 349–360. https://doi.org/10.22044/jme.2017.6195.1435

Dehghani, H. & Bogdanovic, D. (2018). Copper price estimation using bat algorithm. *Resources Policy*, *55*(October), 55–61. https://doi.org/10.1016/j.resourpol.2017.10.015

Dooley, G., & Lenihan, H. (2005). An assessment of time series methods in metal price forecasting. *Resources Policy*, *30*(3), 208–217. https://doi.org/10.1016/j.resourpol.2005.08.007

García, D., & Kristjanpoller, W. (2019). An adaptive forecasting approach for copper price volatility through hybrid and non-hybrid models. *Applied Soft Computing Journal*, *74*, 466–478. https://doi.org/10.1016/j.asoc.2018.10.007

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *J. Basic Eng.*, *82*(1), 35–45. https://doi.org/10.4271/490207

Kriechbaumer, T., Angus, A., Parsons, D., & Rivas Casado, M. (2014). An improved wavelet-ARIMA approach for forecasting metal prices. *Resources Policy*, *39*(1), 32–41. https://doi.org/10.1016/j.resourpol.2013.10.005

Li, G., & Li, Y. (2015). Forecasting copper futures volatility under model uncertainty. *Resources Policy*, *46*, 167–176. https://doi.org/10.1016/j.resourpol.2015.09.009

Liu, C., Hu, Z., Li, Y., & Liu, S. (2017). Forecasting copper prices by decision tree learning. *Resources Policy*, *52*(May), 427–434. https://doi.org/10.1016/j.resourpol.2017.05.007

Sánchez Lasheras, F., de Cos Juez, F. J., Suárez Sánchez, A., Krzemień, A., & Riesgo Fernández, P. (2015). Forecasting the COMEX copper spot price by means of neural networks and ARIMA models. *Resources Policy*, *45*, 37–43. https://doi.org/10.1016/j.resourpol.2015.03.004

Tang, K., Xiong, W., (2012). Index investment and the financialization of commodities. *Financial Analysts Journal*. 68 (5), 54–74.

Wang, C., Zhang, X., Wang, M., Lim, M. K., & Ghadimi, P. (2019). Predictive analytics of the copper spot price by utilizing complex network and artificial neural network techniques. *Resources Policy*, *63*(January). https://doi.org/10.1016/j.resourpol.2019.101414

Watkins, C., & McAleer, M. (2004). Non-Ferrous Metal Prices. *Journal of Economic Surveys*, *18*(5), 651–702. https://doi.org/10.1111/j.1467-6419.2004.00233.x

Watson, M. W., & Stock, J. H. (2004). Combination forecasts of output growth in a seven-country data set. *Journal of Forecasting*, *23*(6), 405–430. Retrieved from http://ideas.repec.org/a/jof/jforec/v23y2004i6p405-430.html

Wets, R. J. B., & Rios, I. (2015). Modeling and estimating commodity prices: copper prices. *Mathematics and Financial Economics*, *9*(4), 247–270. https://doi.org/10.1007/s11579-014-0140-2

Yao, J., & Tan, C. L. (2000). A case study on using neural networks to perform technical forecasting of forex. *Neurocomputing*, *34*, 79–98. https://doi.org/10.1016/S0925-2312(00)00300-3