Resources Policy 52 (2017) 427–434

Contents lists available at ScienceDirect

Resources Policy

journal homepage: www.elsevier.com/locate/resourpol

Forecasting copper prices by decision tree learning

Chang Liu[a,*], Zhenhua Hu[a], Yan Li[b], Shaojun Liu[b]

[a] School of Business, Central South University, Changsha, Hunan, 410083, China
[b] College of Mechanical & Electrical Engineering, Central South University, Changsha, Hunan, 410083, China

ARTICLE INFO

Keywords:
Machine learning
Copper prices forecasting
Decision tree
Correlation
Prediction

ABSTRACT

Forecasting the prices of metals is important in many aspects of economics. Metal prices are also vital variables in financial models for revenue evaluation, which forms the basis of an effective payment regime using by resource policymakers. In this article, we utilized a machine-learning algorithm based on decision tree to predict future copper prices. We showed that our method is capable of accurately and reliably predicting copper prices in both short-term (days) and long-term (years), with mean absolute percentage errors below 4%. We also demonstrated that the current method is assumption-free, robust, and not prone to human bias. This method is easily and readily applicable to predicting the prices of other metals and other commodities, and we expect that such method could be useful in a broad range of fields.

1. Introduction

Copper price plays vital roles in various aspects in today's economies Li and Li (2015); Sánchez Lasheras et al. (2015); Watkins and McAleer (2004); Gargano and Timmermann (2014). On one hand, copper is strongly associated with many industries, such as electrical wiring, construction, and equipment manufacturing; and therefore, copper price has become a significant impact factor on the performance of related companies and economies Sánchez Lasheras et al. (2015). On the other hand, for certain developing countries whose economy relies extensively on copper production, such as Chile and Zambia, fluctuations in copper price has been placed on a national strategic position as it will impact the national income Sánchez Lasheras et al. (2015). Furthermore, copper price provides critical information to financial market participants such as banks and investment funds as these speculators are active participants in metal futures markets Watkins and McAleer (2004).

Forecasting copper price is particularly important for policymakers and participants in the market. First, budgeting is crucial for mining companies as the payback periods are typically long and therefore it requires a steady and reliable financial model to adequately evaluate and forecast the future revenue and cost to maintain a healthy cash flow along the project phase Runge (1998). In addition, an effective financial model is necessary to the authorities who issue mining permits to assess the project revenue and to determine the appropriate payment or loyalty from contractors, as payment regimes are formulated based on the models Garnaut (2010). As the price of copper is directly coupled to the project revenue, and is one of the most important parameters in financial models, accurate forecasting of copper price in the long term is essential for policymakers to make decisions Wets and Rios (2015).

In the past decades, the forecast of copper pricing has attracted a wide academic interest Gargano and Timmermann (2014). Copper prices are highly volatile and depend on many external factors Bresnahan and Suslow (1985); Wets and Rios (2015). Much effort has been made to predict future copper prices using various modeling approaches Gargano and Timmermann (2014); Li and Li (2015); Sánchez Lasheras et al. (2015); Buncic and Moretto (2015). For example, factor-augmented methods based on standard principal components analysis or partial least squares regression have been attempted to extract dynamic factors from large data sets for forecasts of commodity prices Groen and Pesenti (2011). Multivariate regression has also been utilized recently Gargano and Timmermann (2014). More recently, dynamic averaging and selection models, as well as neural networks and ARIMA models, have also been exploited to forecast copper prices Buncic and Moretto (2015); Sánchez Lasheras et al. (2015). However, the literature on modeling copper prices has found mixed results regarding predictability Buncic and Moretto (2015).

Recently, we have been exploring the possibility of utilizing a machine-learning algorithm based on decision tree for the prediction of copper prices. Decision tree learning is a method commonly used in data mining Breiman et al. (1984), and has been exploited for forecasting house prices, oil prices, and stocks Fan et al. (2006); Tiwari et al., ; Charlot and Marimoutou (2014); Bhar et al. (2015), as well as measuring firm performances Delen et al. (2013). A decision

* Corresponding author.
E-mail address: laviniacliu@gmail.com (C. Liu).

http://dx.doi.org/10.1016/j.resourpol.2017.05.007
Received 14 August 2016; Received in revised form 14 May 2017; Accepted 15 May 2017

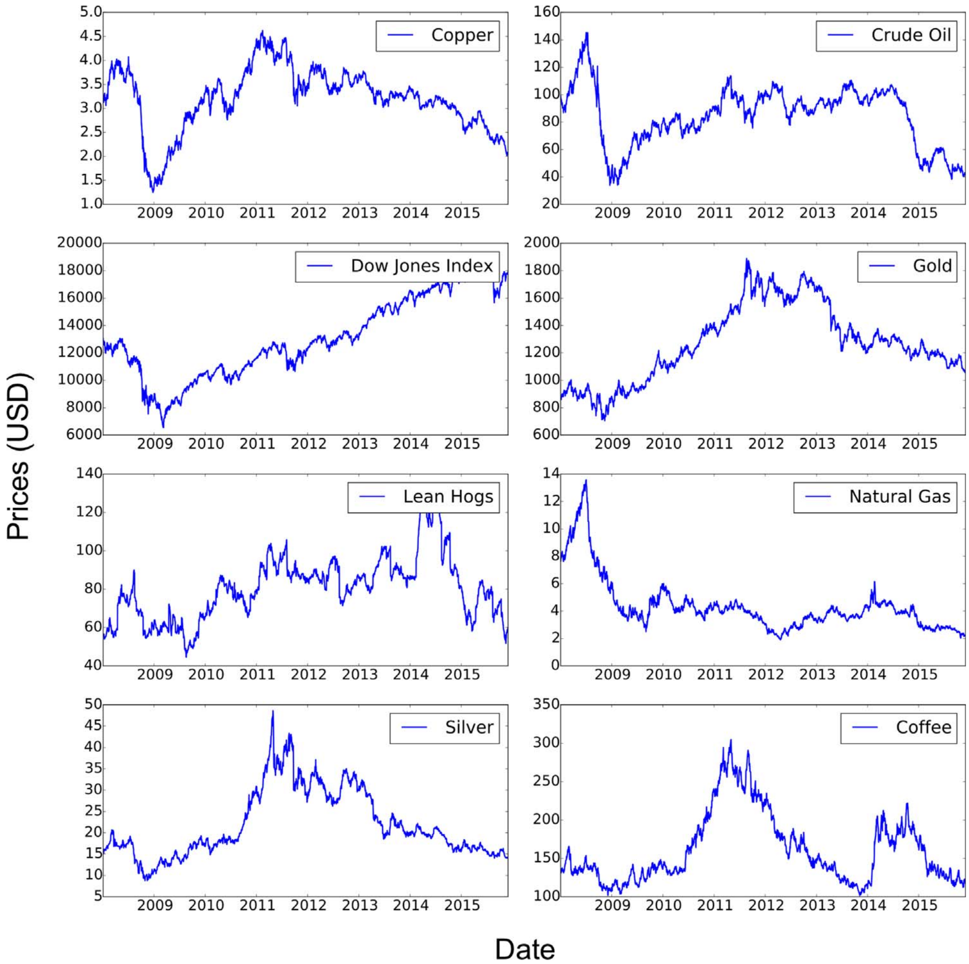

Fig. 1. The actual values of copper price and affecting factors (crude oil, natural gas, gold, silver, lean hogs, coffee, and Dow Jones index) from Jan 2008 to Dec 2015. All the data were available to public and were downloaded from http://www.investing.com.

tree algorithm works by splitting a data set in order to train a model through a recursive partitioning process, and then the model is used to predict the value of a target variable based on the independent variables Breiman et al. (1984).

In the current study, we showed that our method based on decision tree learning is capable of predicting future copper prices accurately, with an error below 5%. In addition, we found that the predictions from our method is precise over a wide range of time scales, from days to months to several years; and the accuracy does not drop for long-term predictions. Due to its simplicity and accuracy, we expect that our method will be useful in various applications for forecasting metal prices and commodity prices.

2. Results and discussions

2.1. Data collection and preparation

Many factors may contribute to the fluctuation of copper prices. For example, the copper price on a future date is likely to be impacted by historic prices of copper. In addition, the price of copper may relate to other metal prices in the market, as alternatives. For example, when other metal price fall down, the copper price is likely to follow the same trend Charlot and Marimoutou (2014). Furthermore, the copper price is likely to depends on energy cost, such as the prices of crude oil and natural gas, due to the connection to the production of copper Joseph and Kundig (1999). Moreover, the copper price could be related to the demand and consumption in industries, which turns to be affected by the general economy environment Gargano and Timmermann (2014); therefore, the copper price might be a function of generic economic indicators such as Dow Jones indices. Lastly, the price of copper might depend on other variables in an unanticipated way. For instance, the prices of lean hogs and coffee might correlate with the copper price unexpectedly.

Based on the reasoning above, and subjected to the data availability, we collected the prices of gold, silver, copper, crude oil, natural gas, lean hogs, and coffee, as well as the Dow Jones indices from Jan 03, 2008 to Dec 01, 2015. All the data were downloaded from http://www.investing.com. Note that, as data for weekends and holidays were missing in the raw data, we excluded weekends from this study. However, because holidays are irregular and sparse, the "missing" prices/values for holidays were estimated by averaging the values from one day before and one day after,

$$P_i = \frac{1}{2}(P_{i-1} + P_{i+1}) \quad (1)$$

if the i-th day is a holiday. At the end, the processed data set contained prices from 2062 days, shown in Fig. 1.

2.2. Correlation between the copper prices and the selected variables

We first confirmed that the copper price is indeed correlated to all the selected other variables. As shown in Fig. 2, where the price of copper and that of the other variables on the same day were ploted as scattered dots (four examples were shown). We observed that the price of copper increases when the prices of crude oil and silver inflate (Fig. 2A–B), which is expected. Interestingly and somewhat surpris-