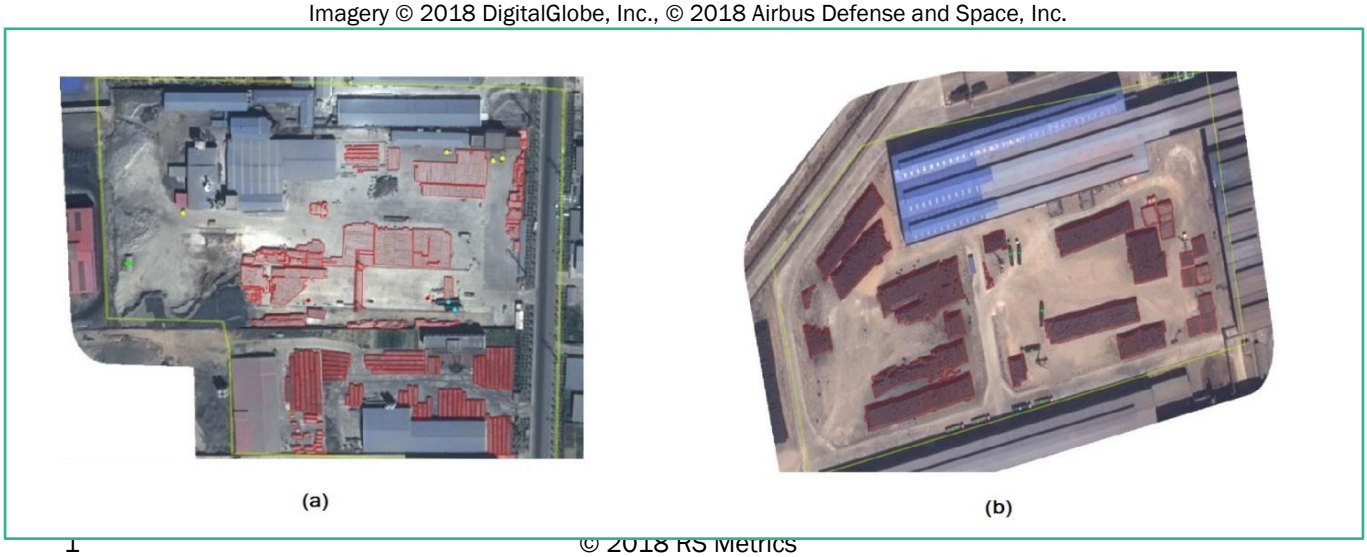
RS Metrics CME Group Copper Futures Price Predictive Analysis Explained

Disclaimer

Summary - Directional Change Prediction

RS Metrics monitors and tracks the estimated amount of metals stored outside many global smelters and storage facilities including Copper (Figure 1 shows images of two such locations and the estimated metal stocks we measure are highlighted in red). The data collected is then used to identify periods of significant change that RS Metrics finds predictive of the CME Group futures copper price. RS Metrics presents below a couple of different methodologies we use (K-Nearest Neighbors and Reversed Moving Average Crossover) that have been experimented to generate 1 to 3 months ahead directional change signals for CME Group Copper futures prices. RS Metrics uses the estimated lagged total metal stocks, piles and concentrates in the KNN model, but use only the total estimated metal stocks in the moving average crossover method to keep it simple. Based on historical data from March 2013 to June 2018, RS Metrics use of the KNN method predicts 1 month, 2 months and 3 months ahead directional changes in CME Group Copper futures prices with 78%, 76% and 81% accuracy levels respectively. Though not as accurate and complex as the KNN method, RS Metrics use of the Reversed Moving Average Crossover method is easier to understand and generates fairly good results for the above directional changes (65%, 68% and 72% accuracy levels respectively). This white paper explains RS Metrics belief in the power of the estimated metal signals all by itself, but when using the RS Metrics data, one should include other data such as consumption, demand for even better results.

Figure 1: (a) Yunnan Copper Chifeng Jinfeng – Copper Smelter (China), (b) Baotou Huading – Copper Smelter (China).



K-Nearest Neighbors (KNN)

K-Nearest Neighbors classification is a non-parametric method that uses a set of features and class labels of a target variable from training data as input to predict the class label of an unknown data point with known feature values. The unknown class label will be determined based on a distance measure between the unknown data point and k-neighboring points with known class labels. Fitting a KNN model involves the selection of two key values; (i) the size of the training sample (n) and (ii) the number of nearest neighbors (k). In our analysis, we use the following variables and inputs to predict directional changes in the CME Group High-Grade (HG) Copper futures' price and use a training sample of size n = N x 80% (rounded up to the nearest integer), where N is the total number of complete data points available. The value of k is determined out of a range of odd integers so that the historical hit rate is maximized (Refer section 'KNN Model Selection' for further details). Only odd integers are used in the selection procedure in order to avoid having ties for the two category (positive and negative) class labels.

Target Variable:

Directional change in CME Group futures Price (1-month, 2-months or 3-months ahead)

Feature Set:

Set of normalized YOY series with specified lags for total estimated metal stocks, piles and concentrates

Methodology

Steps

1. Obtain monthly estimates of average metal stored outside global smelter and storage locations
2. Impute each series using 'Kalman Smoothing' method to adjust for missing values
5. Aggregate series location-wise by taking a weighted average
4. Convert each aggregated series to YOY (Year-Over-Year) change in order to remove seasonal effects
3. Perform KNN method for RS Metrics to predict CME Group futures price and inventory directional changes
    a. Normalize the features to remove influences due to measurement units
    b. Obtain target price and volume signals
    c. Obtain the best model to do the prediction using the most recent complete set of data
    d. Predict the unknown signal using the best model obtained from current data

KNN Model Selection

As we continue to add more locations and more estimated metal data each month, the KNN model we use to make forecasts is adjusted each month in order to take into account the improved dataset and to maximize the hit rate. The following procedure is followed in order to select a better KNN model to make forecasts each month.

Steps

1. Let initial best guess of k be $k_0 = \sqrt{n}$ (take next odd integer, if k_0 isn't odd), where n is the training window size
2. Define a range of odd integers to test for best k as {3, 5, … , k_0, … , $k_0 + 5$}
3. For each combination of k and the feature set;
    a. Use a moving window of size n (first window starts at Mar-13 data point) and fit the KNN model
    b. Predict the $(n+1)^{th}$ signal based on the fitted model
    c. Move the window and repeat step (b) till we predict the last known directional change(s)
    d. Calculate the hit rate
    e. Select the best k (k_{best}) by the value of k that provides the highest historical hit rate
4. Use the selected best model to predict the next unknown directional change(s)

Price Hit Rate

RS Metrics has found that the KNN approach predicts all three directional changes (1-month, 2-months and 3-months ahead) in CME Group copper futures price with an accuracy level of more than 75%. Currently, the 3-months ahead forecasts seem to be the most accurate carrying a hit rate of 81% based on historical data (see Figure 2). The green color represents positive directional forecasts while red color represents negative directional forecasts in figure 2.

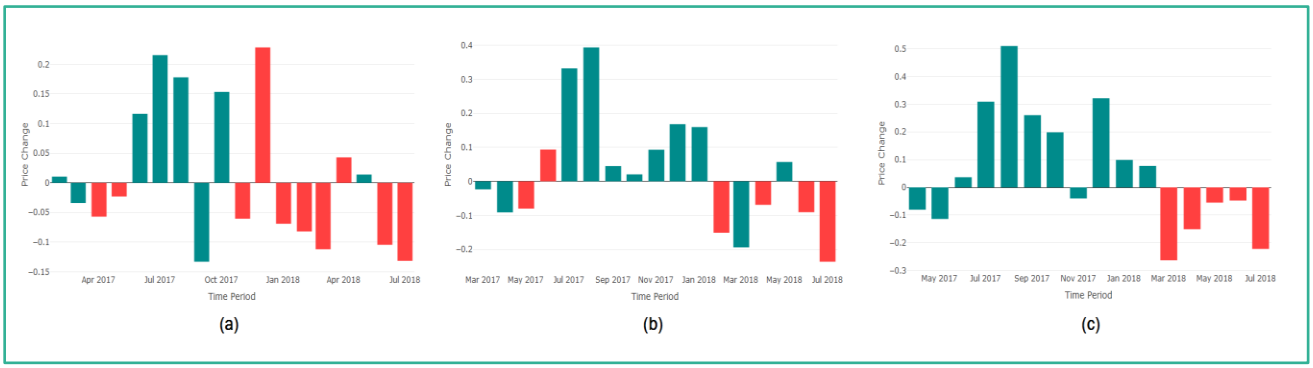


Figure 2: Directional change hits or misses for CME copper futures price (a) 1-month (b) 2-months and (c) 3-months ahead forecasts

Reversed Moving Average Crossover

RS Metrics experimented with a directional change signal generation methodology based on the common moving average crossover trading strategy that involves two moving average series, one with a shorter rolling period, and another with a longer rolling period. When moving average crossover method is used in general stock trading, it is assumed that a positive signal (bullish signal) is indicated if the short term moving average rises above the long term moving average of the relevant stock price. In our analysis, the short term and long term moving averages of estimated metal stocks stored outside Copper smelter locations are used to predict the directional changes in CME Group copper futures price. RS Metrics observed that a reversed moving average crossover methodology (i.e. a negative signal is indicated if the short term moving average rises above the long term moving average) works well to achieve this goal.

Methodology

Steps
1. Obtain the weekly average estimated metal stocks stored outside global smelter locations
2. Impute the missing values based on 'Kalman Smoothing' method
3. Obtain the relevant current average and long term moving average of estimated total metal stocks for each month
4. Calculate the difference between the two series
5. Assign a positive label if difference (current avg – long term avg) < 0 and a negative label if difference > 0

Price Hit Rate

RS Metrics use of the reversed crossover approach predicts all three directional changes (1-month, 2-months and 3-months ahead) in CME Group copper futures price with an accuracy level of more than 64%. After RS Metrics experimenting with multiple combinations of short term and long term moving averages, it appeared that the following combinations generate fairly good signals for the three types of directional changes RS Metrics is interested in. A 3-week current estimated average and a 14-week long term estimated average resulted in a 65% hit rate for 1-month ahead forecasts. A 6-week current estimated average and an 8-week long term estimated average resulted in a 68% hit rate for 2-months ahead forecasts. A 2-week current estimated average and a 22-week long term estimated average resulted in a 72% hit rate for 3-months ahead forecasts.

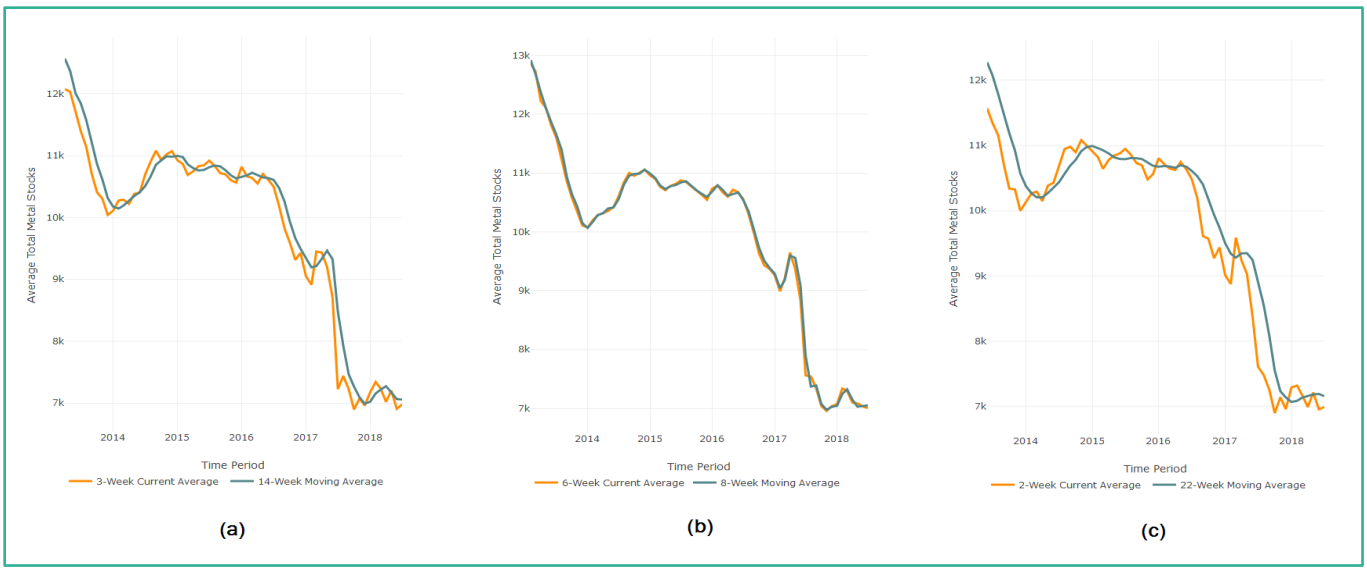


Figure 3: (a) 3-week current average vs. 14-week moving average, (b) 6-week current average vs. 8-week moving average and (c) 2-week current average vs. 22-week moving average

RS Metrics analysis presented above focuses only on predicting CME Group HG Copper futures' price changes using RS Metrics estimated metal data. As previously noted, the current results are purely based on RS Metrics estimated metal data and to obtain better forecasts, it is recommended that you use RS Metrics data along with other data such as consumption and demand.