NYU | TANDON

Commodities Time Series Prediction using Relevance Filtering and CEnet

NYU Capstone Project hosted by Bank of America

Andong Zhou (az1418), Anurag Rathore (ar6634), Jiayu Shi (js11603)

05.17.22

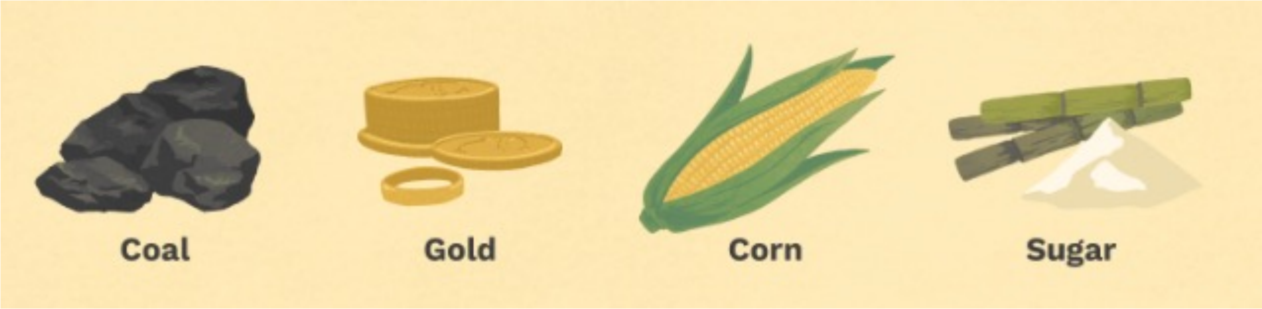
PART 01

Introduction & Motivation

Drawbacks of traditional time series analysis

Introduction - Commodity

- Commodities: grains, gold, beef, oil, and natural gas, etc.
    - producers and consumers: hedge price risk
    - Investors: diversify portfolios, hedge inflation and volatility



Motivation - Traditional Approach Drawbacks

1. **Observations Processing**

    - Not considering the relevance between historical observations and the most recent one.

    → relevance filtering (Czasonis et al., 2021)

2. **Feature Selection and Forecast Combination**

    - Multi-variate models in linear settings

    → C-Enet (Rapach and Zhou, 2020)

PART 02

Data

PART 02

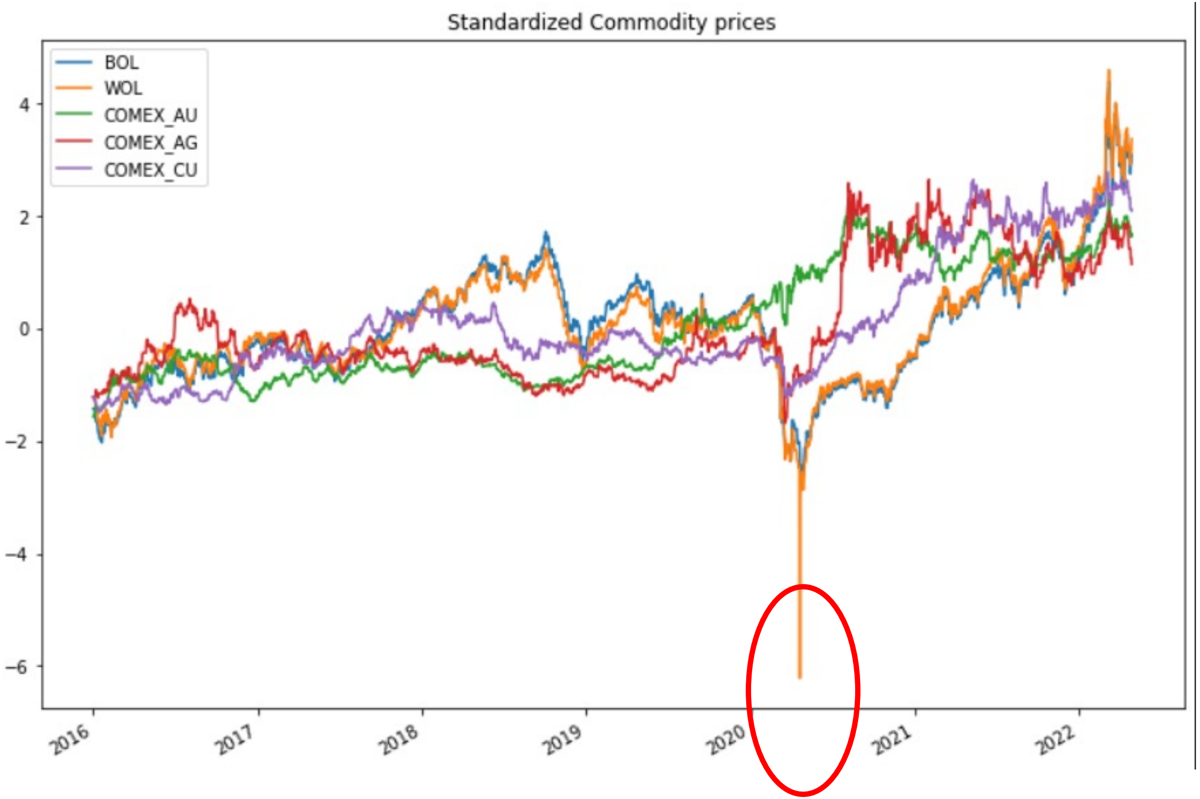
Data

- Source

Yahoo Finance API (yfinance)

- Time range

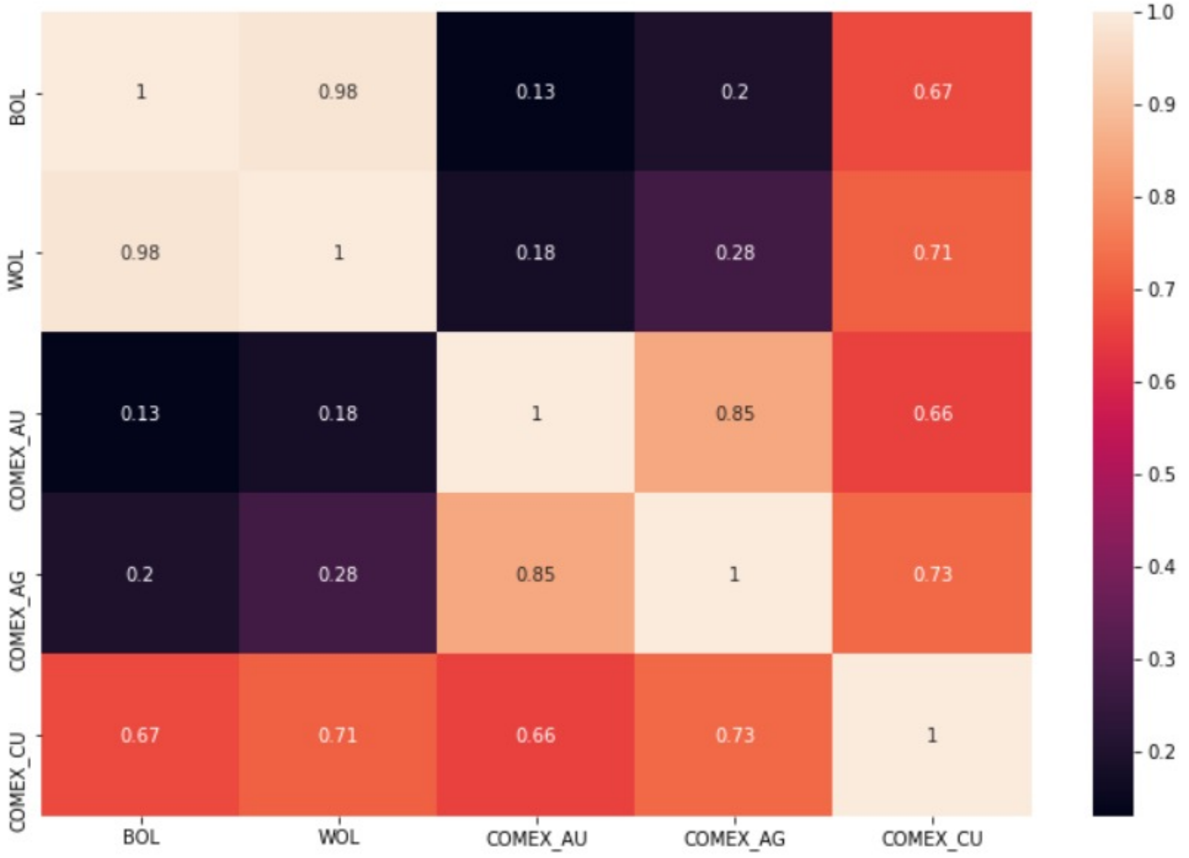
01/01/2016 ~ 04/30/2022



NYU | TANDON

PART 02

Data



Commodity Correlation Matrix

PART 03

Methodologies

Relevance Filtering & CEnet

Methodology - Relevance Filtering

For any pair of observations x_i and x_j,

$$sim_{ij} = sim(x_i, x_j) = -\frac{1}{2}(x_i - x_j)\Omega^{-1}(x_i - x_j)'$$

$$info_i = info(x_i) = \frac{1}{2}(x_i - \overline{x})\Omega^{-1}(x_i - \overline{x})'$$

Relevance is defined as

$$r_{ij} = r(x_i, x_j) = sim_{ij} + info_i + info_j$$

Methodology - Relevance Filtering

The generalized equation for relevance can be given by:

$$r_{it} = (x_t - \overline{x})\Omega^{-1}(x_i - \overline{x})'$$

In our modelling, when predicting point $t + 1$, we consider only the observations that satisfy:

$$r_{it} > 0$$

for any i in the sample of interest.

Methodology - Combination Elastic Net

Conventional multiple predictive regression takes the form

$$r_t = \alpha + \sum_{j=1}^{J} \beta_j x_{j,t-1} + \varepsilon_t$$

$$\hat{r}_{t+1|t}^{OLS} = \hat{\alpha}_{1:t}^{OLS} + \sum_{j=1}^{J} \hat{\beta}_{j,1:t}^{OLS} x_{j,t}$$

Rapach and Zhou consider a combination forecast that takes the form of a simple average of the univariate predictive regression forecasts based on $x_{j,t}$ for $j = 1, \ldots, J$

$$r_t = \eta + \sum_{j=1}^{J} \theta_j \hat{r}_{t|t-1}^{(j)} + \varepsilon_t$$

NYU | TANDON

Methodology - Combination Elastic Net

Rapach and Zhou then consider the Granger and Ramanathan regression:

$$r_s = \eta + \sum_{j=1}^{J} \theta_j \hat{r}_{s|s-1}^{(j)} + \varepsilon_t$$

For $s = t_1 + 1, \dots, t$.

Let $\mathcal{J}_t \subseteq \{1, \dots, J\}$ denote the index set of individual univariate predictive regression forecasts selected by the ENet, i.e., $\theta_j > 0$.

$$\hat{r}_{t+1|t}^{CEnet} = \frac{1}{|\mathcal{J}_t|} \sum_{j \in \mathcal{J}_t} \hat{r}_{t+1|t}^{(j)}$$

Where $|\mathcal{J}_t|$ is the cardinality of $\mathcal{J}_t$

Methodology – Relevance Filtering + C-Enet

Step 1 For each predictor variable group, we compute rolling partial-sample multivariate predictive regression forecasts with relevance-filtered observations and L2 norm (Ridge) over the holdout out-of-sample period:

$$\hat{r}_{s|s-1}^{(c)} = \hat{\alpha}_{1:s-1}^{c} + \sum_{j=1}^{J_c} \hat{\beta}_{1:s-1}^{(c,j)} x_{c,j,s-1}$$

For $c = 1, \dots, C$, $s = t_1 + 1, \dots, t$ and $j = 1, \dots, J_c$.

Step 2 We estimate the Granger and Ramanathan regression via the ENet over the holdout out-of-sample period:

$$r_s = \eta + \sum_{c=1}^{C} \theta_c \hat{r}_{s|s-1}^{(c)} + \varepsilon_s$$

For $s = t_1 + 1, \dots, t$. Let $\mathcal{C}_t \subseteq \{1, \dots, C\}$ denote the index set of categorical multivariate predictive regression forecasts selected by the ENet.

Step 3 We compute the C-ENet forecast as

$$\hat{r}_{t+1|t}^{CEnet} = \frac{1}{|\mathcal{C}_t|} \sum_{c \in \mathcal{C}_t} \hat{r}_{t+1|t}^{(c)}$$

NYU | TANDON

PART 04

Model

Model - Variables

- Dependent variable (Y)
  - Oil group: Brent Crude Oil Price, WTI Oil price
  - Metal group: COMEX Gold, Silver, and Copper price

- Independent variable (X), total 79
  - Stock market indices (10)
  - CBOE volatility index (VIX) (7)
  - Treasuries (5)
  - Transport fees (6)
  - Commodity prices and indices (7)
  - Fundamentals of commodity (33)
  - Simple combination of features (11)

Model

1. derive respective relevance filtered daily predictions (Observations Selection)
   - 126-day (half-a-year) rolling basis
2. train C-Enet on predictions of all groups (Feature Selection and Forecast Combination)
   - rotational 3-year in-sample
   - 3-month holdout period

Model - Benchmarks

1. Relevance-filtered regression
2. Ridge regression
    - 126-day (half-a-year) rolling basis

Model - Metrics

1. R^2_{OS} statistic

$$R^2_{OS} = 100 * \left(1 - \frac{MSFE_{PREP}}{MSFE_{HM}}\right) = 100 * \left[1 - \frac{\sum_{t=t_0}^{T-1} (r_{t+1} - \widehat{r_{t+1}})^2}{\sum_{t=t_0}^{T-1} (r_{t+1} - \overline{r_t})^2}\right]$$

2. cumulative sum of squared error differences (CSSED)

$$CSSED_{it} = \sum_{\tau=1}^{\tau=t} (e^2_{Bmk,i\tau} - e^2_{Pbrk,i\tau})$$

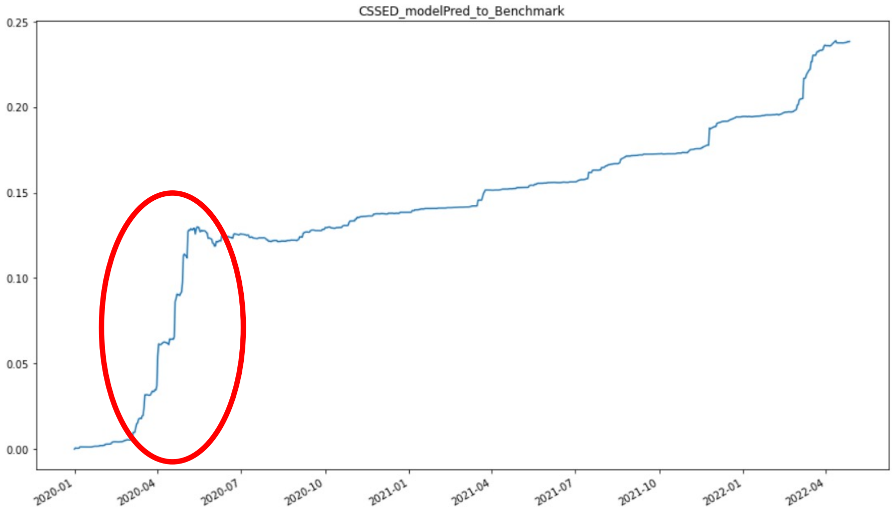
PART 05

Experiment Result

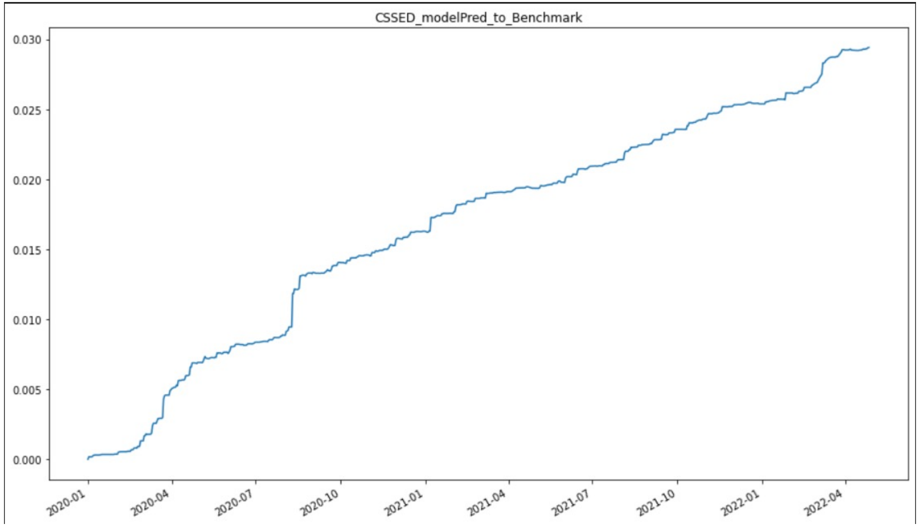
Result - Proportional Reduction

Proportional reduction	Oil group	Metal group
in model prediction error	0.6672	0.7015
in relevance filtered prediction error	0.5103	0.4835
in benchmark prediction error	0.3140	0.2950

Result - Cumulative Sum of Squared Error Differences compared to Benchmark



Oil group



Metal group

PART 06

Reference

Campbell, J. Y. & Thompson, S. B. (2008) 'Predicting excess stock returns out of sample: Can anything beat the historical average?', *The Review of Financial Studies*, 21(4), pp. 1509-1531, JSTOR [Online].

Czasonis, M., Kritzman, M. & Turkington, D. (2021) 'Relevance', *MIT Sloan Research Paper No. 6417-21*, SSRN [Online].

Rapach, D. E., J. K. Strauss & G. Zhou (2010) 'Out-of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy', *Review of Financial Studies*, 23(2), pp. 821–862, JSTOR [Online].

Rapach, D. E. & Zhou, G. (2020). 'Time-series and cross-sectional stock return forecasting: new machine learning methods', *Machine Learning for Asset Management: New Developments and Financial Applications*, SSRN [Online].

Thank you!