# [Daily Crude Oil Price Analysis and Forecasting Based on the Sequential Majorization Method]

[Jian Shen, PhD Candidate, +44 07599072812, J.Shen@soton.ac.uk]
[Xiaohang Ren, PhD Candidate, +44 07761618538, X.Ren@soton.ac.uk]

## Overview

Crude oil prices analysis and forecasting attracts a lot of attention from both academic researchers and energy industrial participators due to its importance. In this paper we will discuss singular spectrum analysis (SSA) and iterative Cadzow method that focused on Hankel structured low-rank approximation (HSLRA) problems as well as sequential majorization method (SMM) which aims to obtain a finite rank series by solving weighted HSLRA problem. We employ these methods to forecast the future daily contract prices of crude oil and some main petroleum products while ARIMA and GARCH are also involved in the forecasting experiments as benchmark models. Numerical results provide empirical evidences that compared with all other models, SMM generates the most accurate forecasting result especially in the middle-long term predictions.

## Methods

Hankel structured low-rank approximation (HSLRA) methods are considered to be effective in time series analysis and forecasting. In this research we mainly discuss three representative algorithms: singular spectrum analysis[1], iterative Cadzow method[2] and sequential majorization method[3]. The HSLRA problem is described as

$$\min f(x) \coloneqq \frac{1}{2}\,||X - Y\,||^2$$
$$\text{s.t. } X \in \mathbf{H} \cap \mathbf{M}_{\boldsymbol{r}}$$

where $\mathbf{H}$ is the set of Hankel matrices and $\mathbf{M}_{\boldsymbol{r}}$ is the set of matrices with rank equal or smaller than r. The singular spectrum analysis (SSA) method tries to approximate this problem through alternating projection methods, that is calculating the projection of giving matric on low rank subspace and Hankel spaces alternately. The iterative Cadzow methods keep doing above alternating projection until some stopping conditions are met..

To deal with the convergence issues as well as the weighted HSLRA problem as following, the sequential majorization method (SMM) is proposed to approximating the low rank Hankel matrix.

$$\min f(x) \coloneqq \frac{1}{2}\,||W \circ (X - Y)\,||^2$$
$$\text{s.t. } X \in \mathbf{H} \cap \mathbf{M}_{\boldsymbol{r}}$$

It is proved [3] that SMM can generate the convergent results since the sequence of function value is non-increasing. Theoretically, SMM outperforms other two methods because of the introduced weighted matrix to reflect the importance of data and the convergent framework. Hence in this paper, above three methods are implemented to forecast the future values of giving series. ARIMA and GARCH models are also introduced as the benchmark model so that we can compare the forecasting result.

## Results

In this study we mainly focusing on forecasting the future prices of crude oil as well as other four kinds of petroleum products. RMSE of forecasting result by SMM algorithms are listed in **Table 1** as well as the RMSE ratio of SMM to other four candidates models. In general, SMM significantly outperforms all other four methods in 8 cases over 15 as indicated by the result of Diebold-Mariano test at 5% significant level. There are three situations that only ARIMA or iterative Cadzow method can provide a slightly better forecasting results compared with SMM which are not significantly, and only 1 case that there exists other models generating more accurate forecasting values significantly compared with SMM. While in the most of remaining cases, SMM provides the most accurate forecasting result by achieving the smallest RMSE.

Table 1 RSME comparison of estimation result by five different models

| Price Series | Forecasting Steps | $RMSE\ of\ SMM$ | $\frac{SMM}{ARIMA}$ | $\frac{SMM}{GARCH}$ | $\frac{SMM}{SSA}$ | $\frac{SMM}{Cadzow}$ |
|---|---|---|---|---|---|---|
| Crude Oil | h = 10 | 2.4177 | 1.01 | 0.99 | 0.60** | 0.63** |
| | h = 15 | 2.5399 | 0.83** | 0.82** | 0.45** | 0.90** |
| | h = 20 | 2.4600 | 0.72** | 0.70** | 0.35** | 0.86** |
| Conventional Gasoline | h = 10 | 0.0999 | 1.55 | 1.53 | 0.89** | 0.66** |
| | h = 15 | 0.0853 | 0.97 | 0.95 | 0.48** | 1.03 |
| | h = 20 | 0.0947 | 0.92** | 0.85** | 0.51** | 1.00 |
| No.2 Heating Oil | h = 10 | 0.0631 | 1.03 | 1.00 | 0.59** | 0.86** |
| | h = 15 | 0.0666 | 0.87* | 0.84* | 0.51** | 0.87** |
| | h = 20 | 0.724 | 0.86* | 0.82** | 0.51** | 0.97** |
| Ultra-Low-Sulfur No.2 Diesel Fuel | h = 10 | 0.496 | 0.87* | 0.87 | 0.40** | 0.78** |
| | h = 15 | 0.529 | 0.69** | 0.68** | 0.33** | 0.75** |
| | h = 20 | 0.0612 | 0.72** | 0.69** | 0.33** | 0.82** |
| Kerosene-Type Jet Fuel | h = 10 | 0.0637 | 0.95 | 0.92 | 0.40** | 0.75** |
| | h = 15 | 0.0682 | 0.80* | 0.77** | 0.38** | 0.75** |
| | h = 20 | 0.0757 | 0.79** | 0.76** | 0.41** | 0.78** |

Notes: **/* represents the results of Diebold-Mariano test. ** indicates SMM provides more accurate results than another model at 5% significant level. * indicates SMM provides more accurate results than another model at 10% significant level.

We also notice from **Table 1** that with forecasting steps being longer, RMSE ratios of SMM to ARIMA and GARCH models become smaller except the case of ultra-low-sulfur No.2 diesel fuel price series forecasting. On the other hand, although iterative Cadzow method failed to generate better predictions than SMM, its forecasting result did not get worse when the forecasting steps extended from 10 steps to 20 steps. For all five daily price series considered in this paper, RMSE ratio of SMM to Cadzow raised from 10 steps forecasting to 20 steps forecasting. As a result, we may conclude that iterative Cadzow method and SMM can provide more stable estimation results in the middle-long term period forecasting, while SMM still outperforms iterative Cadzow method due to its ability in capturing the underlying periodicity components in a time series better

## Conclusions

In this study we mainly introduced three novel and powerful time series dimensional reduction methods dealing with HSLRA and weighted HSLRA problems as SSA, iterative Cadzow method and SMM. These methods were implemented to analyse and forecast daily price series of crude oil as well as several kinds of main petroleum products including conventional gasoline, heating oil, diesel fuel and jet fuel. In the core part of this paper, forecasting performance of three HSLRA methods were analysed and compared based on different forecasting periods. Using ARIMA and GARCH as benchmark model, we found numerical evidences that in most cases, SMM can generate most accurate forecasting result with smaller RMSE while there is only 1 cases out of 15 that other candidate models can generate predictions significantly better. We further discussed the stable forecasting performance of iterative HSLRA methods in the middle-long term compared with parametric models. Therefore, we concluded that SMM is a suitable method in predicting the daily price series of crude oil and petroleum products.

## References

[1]Golyandina, N., & Zhigljavsky, A. (2013). *Singular Spectrum Analysis for time series*. Springer Science & Business Media.

[2]Qi, H. D., Shen, J., & Xiu, N. (2017). A sequential majorization method for approximating weighted time series of finite rank. *Statistics and Its Interface*.

[3]Zvonarev, N., & Golyandina, N. (2017). Iterative algorithms for weighted and unweighted finite-rank time-series approximations. *Statistics and Its Interface*, *10*(1), 5-18.