Urals loading analysis

Statistical analysis of Urals loadings for forecasting

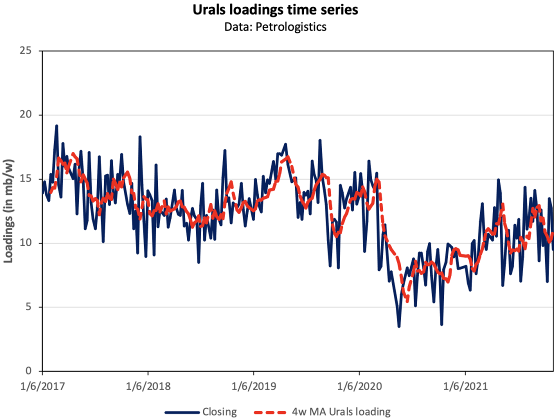
Youssef Louraoui[1]

[1]Summer Analyst
Analysis Desk

October, 2022

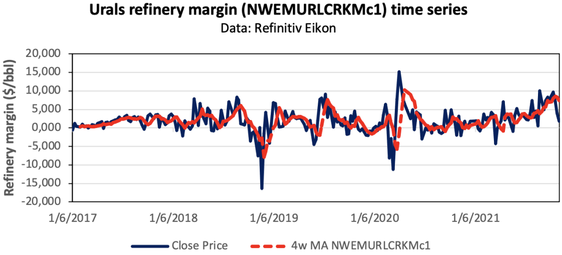
Historical analysis of Ural loading

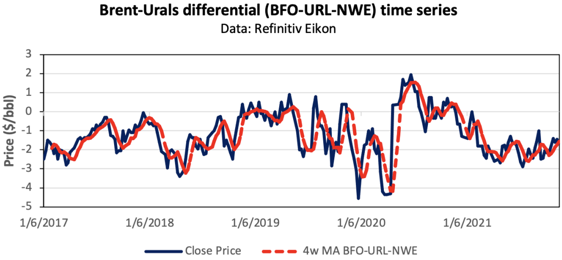
- Over the period spanning 2017-2021, Urals loadings was moving in a range between 3.19mb/d end Feb 2017 to less than 0.58mb/d May 2020.
- Average mean for the whole period covering 2017-2021 for Urals loading reached the equivalent of 2.03mb/d (stdev: 3.07mb/d).



Urals refining margin and Ural-Brent differential analysis

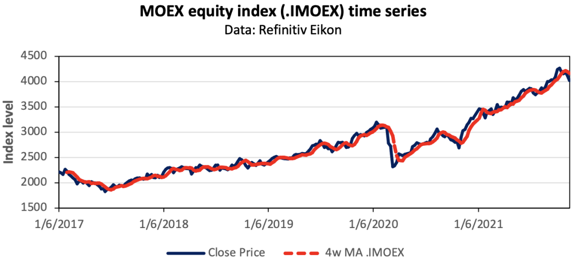
- Urals refining crack was fluctuating around end of 2018Q4, where it reached it lowest point in the sample (-16.34 USD/bbl), and in 2020Q1 where it reached it highest level across the sample analysed (15.1 USD/bbl).
- Brent-Urals differential has been drifting around its long term mean during the period covered in this analysis (-1.24 USD/bbl, stdev 1.18 USD/bbl).

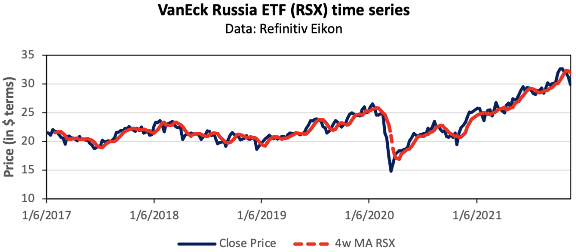




Equity index analysis

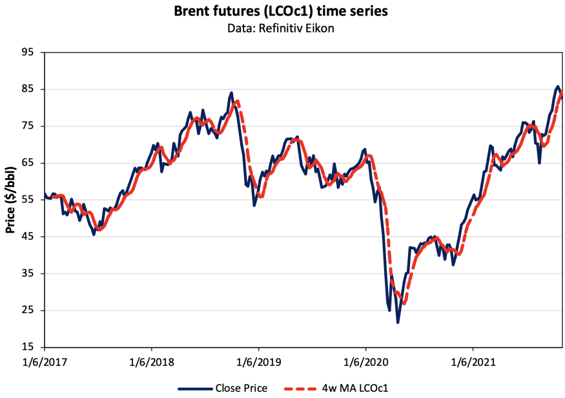
- During the period covering 2021, MOEX performed above its average for the period (avg. 2021: 3710.74 pts vs avg. sample: 2708.96 pts).
- With a correlation of 0.8782 (87.82% between the two), the VanEck tracker closely resembles the Russian equity index (MOEX). Due to the holdings' differences from the MOEX index, this divergence from the tracker can be justified.





Brent futures analysis

- Brent futures traded in a range covering (21.8 USD/bbl) at its lowest value during Mar. 2020 to more than 84 USD/bbl end of 2021.
- Across the last year worth of data, Brent futures where trading above their long term mean for the period 2017-2021 (avg 2022: 69.03 USD/bbl vs avg. 2017-2021: 60.7 USD/bbl).



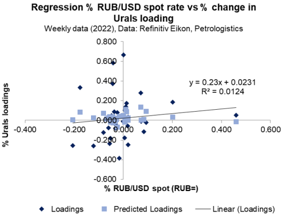
Effects of the variables on change in Urals loading

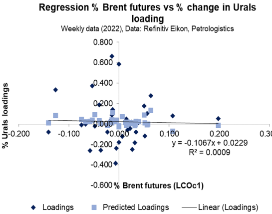
The figure below captures the effect of change in RUB/USD spot rate, Brent futures, and USD index with respect to the change in Urals loadings over 2022, or the equivalent of 40 data points. The regression can be captured mathematically as follows $(*)$:

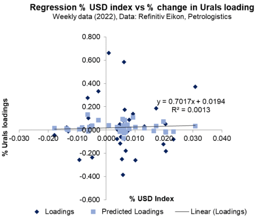
$$y_{\text{Loadings 2022}} = \alpha + \beta_{\%RUB/USD_{spot}} X_{\%RUB/USD_{spot}} + \beta_{\%Brent_{futures}} X_{\%Brent_{futures}} + \beta_{\%USD_{index}} X_{\%USD_{index}} + \epsilon_t$$

- Very poor relationship between the Urals loading rate with respect to the other explanatory variables used.
- Urals loadings seem to be sensitive to changes in Brent futures price ($R^2 = 2.1\%$).
- Overall R^2 for the regression is not conclusive (0.034, 3.4%). In other words, 3.4% of the variations in the Urals loadings can be attributed to the changes in the explanatory variables retained in this analysis.
- This slightly negative relationship can seem intuitive since Urals difference is a sound gauge for Russian crude attractiveness in the crude oil market.
- This analysis offers an additional view on possible variables that can be used to better predict Russian Ural loading.

Augmented econometric modeling analysis







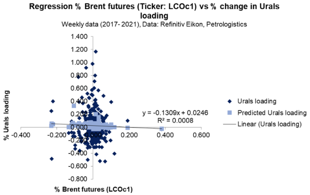
Augmented econometric modeling analysis

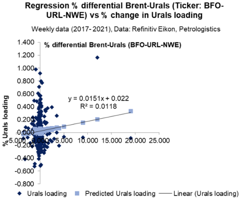
The figure below captures the effect of change in Brent futures, Brent-Urals differential, MOEX equity index, VanEck Russia ETF and Urals refinery crack with respect to the change in Urals loadings over the equivalent of 235 data points. The regression can be captured mathematically as follows (*):

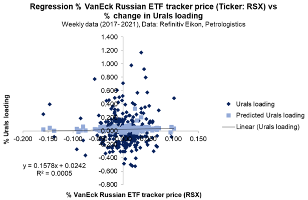
$$\begin{aligned} y_{\text{Loadings 2017–2021}} = \alpha &+ \beta_{\%\text{Brent futures}} X_{\%\text{Brent futures}} + \beta_{\%\text{BFO–URL NWE}} X_{\%\text{BFO–URL NWE}} \\ &+ \beta_{\%\text{MOEX index}} X_{\%\text{MOEX index}} \\ &+ \beta_{\%\text{RSX}} X_{\%\text{RSX}} + \beta_{\%\text{Urals ref–crack}} X_{\%\text{Urals ref–crack}} + \varepsilon_t \end{aligned}$$

- Again, very poor relationship between the Urals loading rate with respect to the other explanatory variables used in this multivariate regression.
- Even after expanding the original regression with the addition of other variables (five factors), we still yield poor regression outputs.
- We can capture an interesting pattern; Urals loadings seems to be sensitive to change in the differential Brent-Urals ($R^2 = 1.18\%$)
- Overall R^2 for the regression is not conclusive (0.0136, 1.36%). In other words, 1.36% of the variations of the Urals loadings can be attributed to the changes in the explanatory variables retained in this analysis.

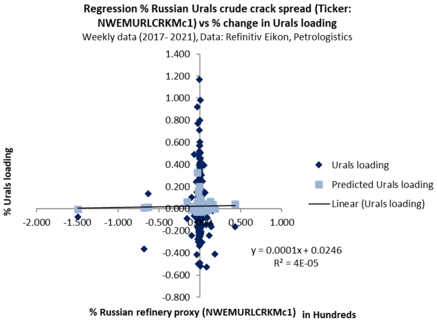
Augmented econometric modeling on Urals loading

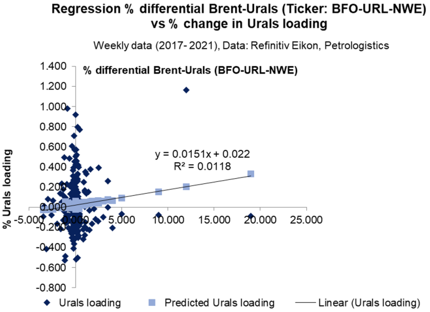






Augmented econometric modeling on Urals loading



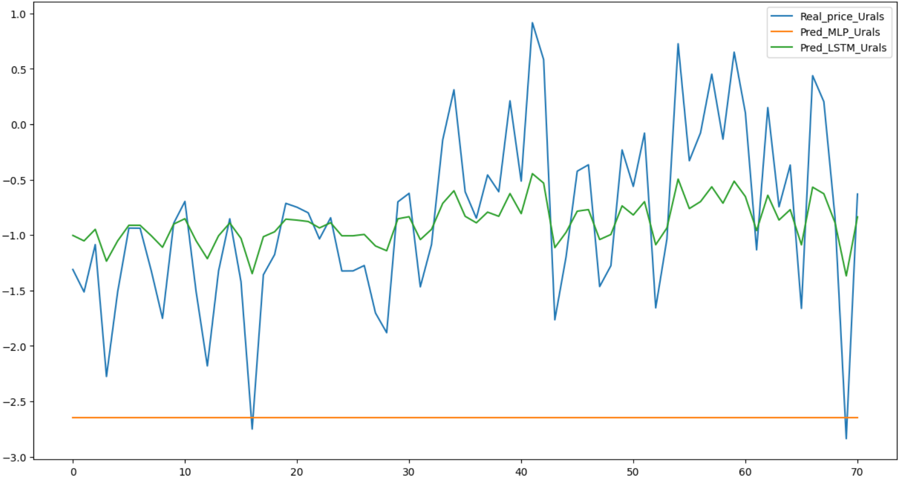


Implementation of MLP and LSTM models

The results obtained from the multiple regression are not very satisfying, the results cannot help much in predicting Urals loading patterns. In this sense, we decide to implement alternative machine learning techniques into the dataset and assess their predictive power(*):

- We use a dataset of weekly data capturing 234 data points. We made sure to eliminate any missing data and applied a z_score to eliminate data above three standard deviation threshold.
- We decompose the dataset into two samples: A training period with the first 120 observations and a testing set of the remaining observations (114 observations).
- We implement at first Multiple Layer Perceptron (MLP) and Long-Short Term Memory (LSTM) algorithms. The (1) and (2) drops 25% of the data to preserve the model from overfitting.
- Both models will run on the training dataset ten times (epoch=10) with a batch of sixteen, meaning the model will be updated after sixteen points are processed.

Machine learning analysis: MLP and LSTM

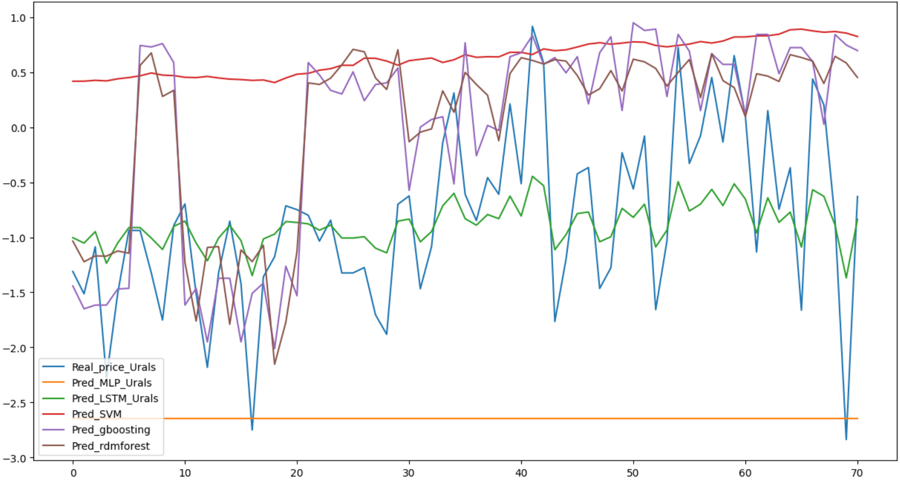


Implementation of alternative models (GBoost, SVM and Random Forest)

For this particular forecasting exercise, the LSTM model shows the best fit. Even though the Random Forest model isn't as precise as the LSTM, it still has good predictive power. Despite employing strong algorithms, both the Gradient Boosting and MLP models imply that careful adjustment is necessary to improve their performance. Finally, the less accurate findings of the SVM highlight the difficulties of using kernel-based techniques in complex time-series forecasting without considerable feature engineering and tweaking (*):

- Exhibits the highest accuracy among all models with the lowest MAE, MSE, and RMSE scores. This suggests a strong capacity for capturing temporal dependencies and sequential patterns within the data.
- Shows competitive accuracy, with relatively low error metrics, particularly the MAE being the second-best. This ensemble model's robustness against overfitting may have contributed to its solid performance.
- Registers the highest error metrics across all three indicators, indicating a lower predictive accuracy for this dataset. The kernel and hyperparameter selection are critical for SVM performance and might require optimization.

Implementation of alternative models (GBoost, SVM and Random Forest)



Conclusion

- After running the regression for each time series, there is a very poor relationship between the Urals loading rate with respect to the other explanatory variables used in this multivariate regression.
- Some interesting relationships* with Urals loadings with Brent futures and Urals loadings with Brent-Urals differential.
- Modelling the relationship of Urals loading via marchine learning techniques gives mixed results. More fine tuning of the parameters and the underlying models is needed to obtain more robust forecasts.
- Modelling the relationship of Urals loading is complicated. The results serve as a basis for further analysis in order to obtain better results.