HUNGARIAN CENTRAL STATISTICAL OFFICE

Nowcasting European Production Prices in Industry

Gergely Attila Kiss, gergely.kiss@ksh.hu

GASP, June 14, 2023

Agenda

1. Introduction
    1. European Statistics Awards
    2. Goal of the project
2. The method of nowcasting
    1. Data
    2. Preprocessing
    3. Hyperparameter tuning
    4. Estimation
3. Summary

Introduction of the Project

European Statistics Awards for Nowcasting

- *„Divided into 3 separate competitions, each focusing on the nowcasting of one economic indicator (PPI, PVI, Tourism), aimed at discovering promising methodologies and potential external data sources that could, now or in the future, be used to improve the timeliness of key EU economic indicators."*
- The competitions were 8 months long from Sep. 01, 2022- Apr. 30, 2023.
- There is two separate section in the competition for each of the three indicators. An accuracy and a reproducibility award.
  - The accuracy award has 3 rewards for giving the best 3 teams in terms of giving the most accurate awards:
    - for at least 5 countries,
    - for 6 consecutive months,
    - for each months 5 available prediction for each country.
  - The reproducibility award is targeted at the top 25% of the accuracy competitors rewarding the most generalizable and best documented approach.
- The algorithm I will present made predictions for PPI for 6 months and PVI for the last few.

Goal of the Project

- We focused on creating a general method that could predict any time series with the appropriate auxiliary data.
- Therefore, the aim was to give predictions for all the participating countries (26).
- Unfortunately, only competed since November, as it took some time to develop the method appropriately.
- It was also, a goal to create a method that is accurate enough so it might be implemented later on as experimental statistics for timlier estimates in the HCSO.
- The current results for PPI are very promising in terms of measuring results in stabilty and accuracy compared to other competitors.
- The continuation of this project currently is applying for a European Grant to implement timlier estimates on three indicators for Hungary (PPI, PVI and SPPI)

The method of nowcasting

The method

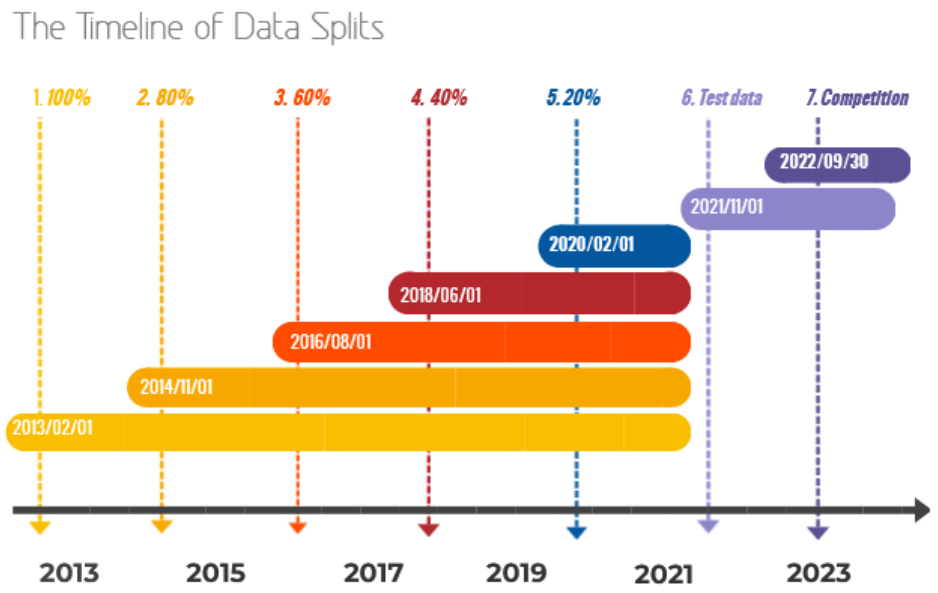
- In our opinion the most important part of any ML model is the data.
  - Our data consists of world market prices of commodities
  - Inputs for production <- correlation with the Producers' prices
- Since the time series structure of the data there were several issues with the standard processes in the ML part that had to be reconsidered:
  - Feature selection as there are a great many of available variables, mainly due to differencing and lagging
  - Preprocessing: Splitting the time series
  - Cross validation: has to preserve the order of the sequence
- The method itself is a two step approach as the first explorative analyses show some simultaneity across Europe.
  - We chose Germany to be estimated in the first step and then use the extended version of the German series as an extra explanatory variable for other countries.

The method: Data

- The most important factor in creating a relevant and accurate method.
- For estimating PPI in Europe we found the World Bank Commodity Price data as a basis (64 commodities).
  - Although it is published a few days after the submission deadline for the competition, so we had to replace it for the competition.
- Replacement data is from the Eikon Refinitive workspace, where we looked up the referenced time series from the World Bank data (48 indicators).
- Next we transformed all the series to be consistent and took the logged, lagged and differenced versions for them (~3200 indicators).
  - Also, have to take car of the Ragged edge problem.
- Then we did some feature selection to filter out the most essential variables for the target. (e.g.: Gas prices, Oil and fat prices, sugar prices, etc.)

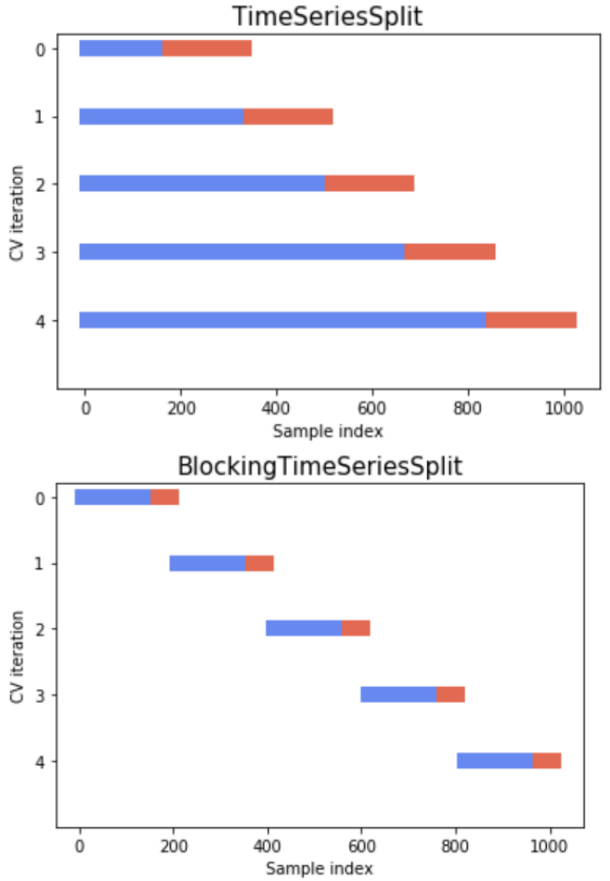
The method: Preprocessing

- Due to the time series nowcasting task at hand we split up our data.
  - The assumption is that the more recent an observation the more relevant it should be to the prediction.
- The result is 5 different training samples.
  - Each containing 20% less of the original sample.
  - Each dropping the latest observations
- Test sample is always up to the current month or one period before as they are not published yet.



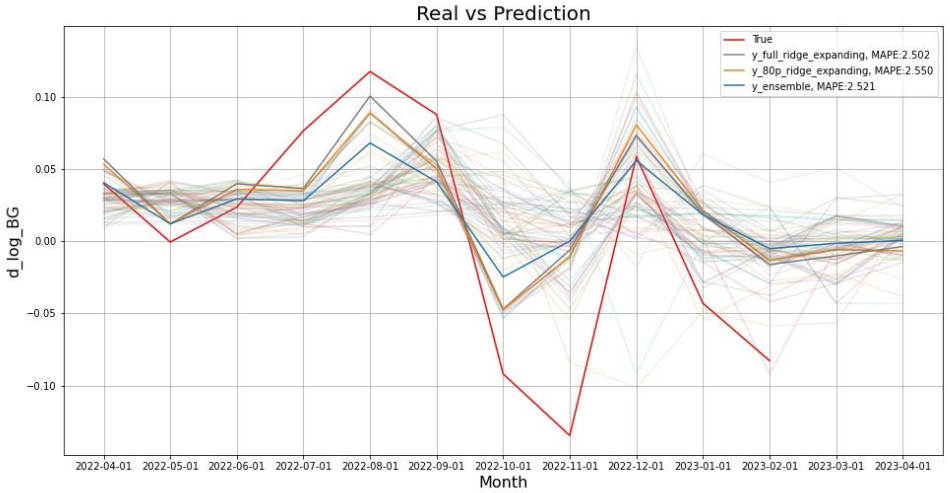
The method: Hyperparameter tuning

- We used 7 different type of ML models. Best performers seem to be boosted random forest regressions.
- We created our own Cross Validation method to match the goal of precision.
  - Day Forward-Chaining validation: we split the training sample by months and calculate the RMSE for each hyperparameter.
  - Always creating one-step ahead forecasts.
  - Two different approach to estimation: expanding window and sliding window
- Small hyperparameter grid and it is searched exhaustively.
- Hyperparameters optimized once and not re-tuned until there is a structural break in the series.



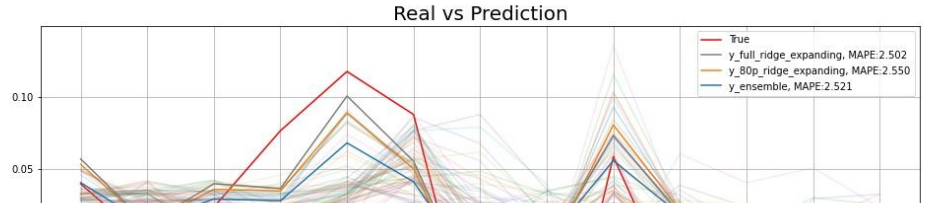
The method: Results

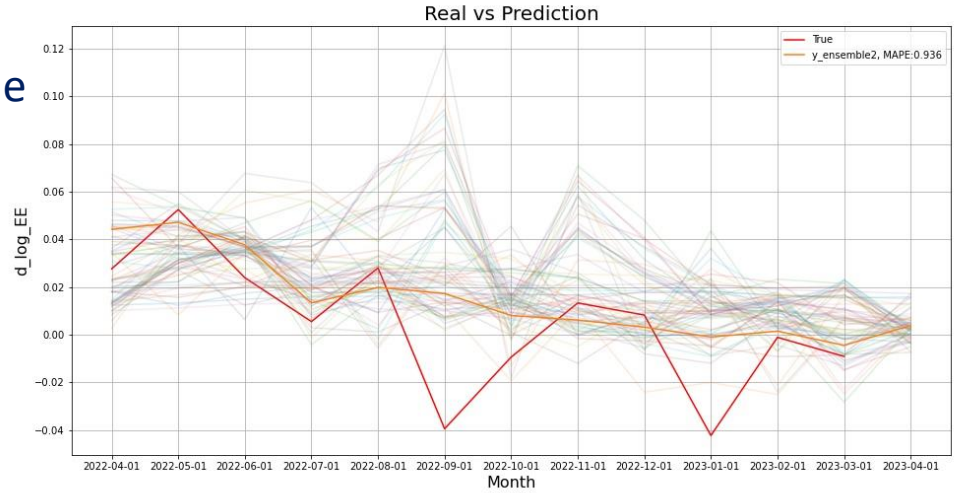
- Model selection: after hyperparameter tuning the next step is to select the best performing models out of the estimated 70 (5 time splits x 2 time windows x 7 models)
- Model selection is done on the test set, by calculating the MAPE of the one-step ahead predicitons
- The 3 and 5 best models then averaged for an ensemble predictor



The method: Results

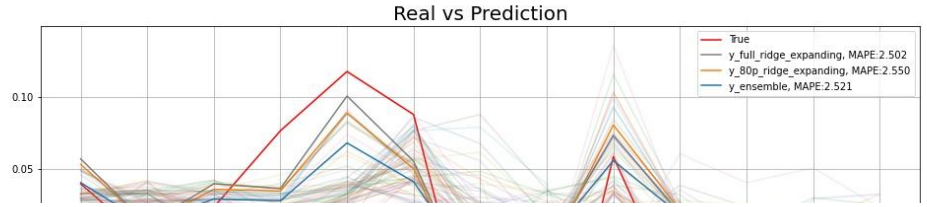
- Model selection: after hyperparameter tuning the next step is to select the best performing models out of the estimated 70 (5 time splits x 2 time windows x 7 models)
- Model selection is done on the test set, by calculating the MAPE of the one-step ahead predicitons
- The 3 and 5 best models then averaged for an ensemble predictor

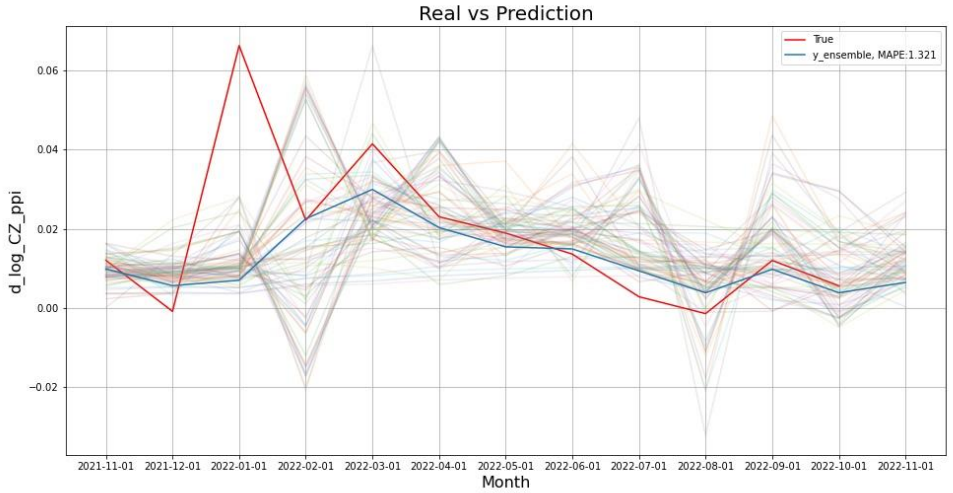




The method: Results

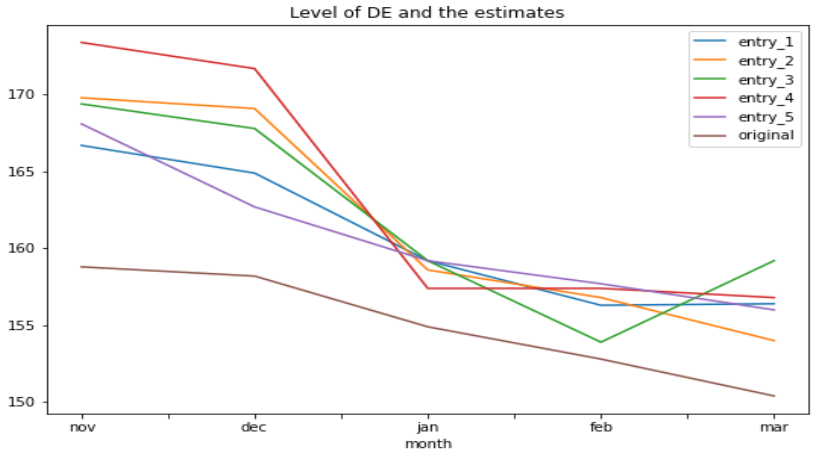
- Model selection: after hyperparameter tuning the next step is to select the best performing models out of the estimated 70 (5 time splits x 2 time windows x 7 models)
- Model selection is done on the test set, by calculating the MAPE of the one-step ahead predicitons
- The 3 and 5 best models then averaged for an ensemble predictor
- The Errors of our nowcasted values:
  - Average value of indices: 165.2
  - Average RMSE of all 5 entries: 13.9 (8.4%)
  - Weighted Average RMSE: 14.9 (9%)
  - Average MSRE of all 5 entries: 0.078
  - Weighted Average MSRE: 0.084
  - The best 5 predictions' average MSRE: 0.000054

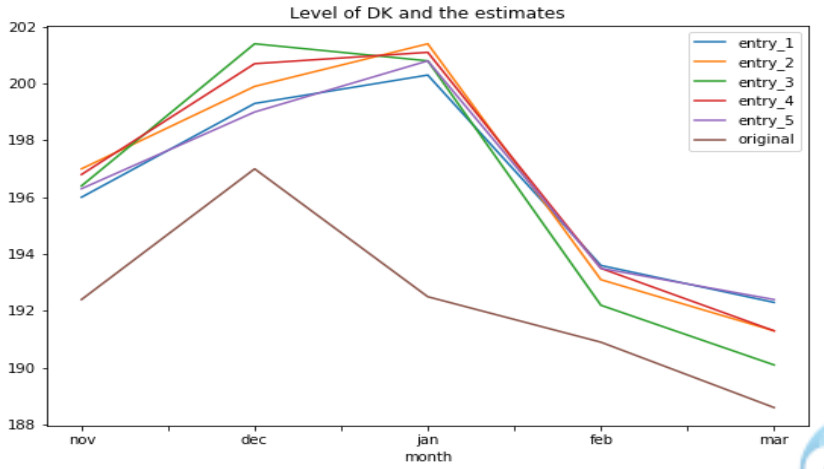




The method: Results

- Model selection: after hyperparameter tuning the next step is to select the best performing models out of the estimated 70 (5 time splits x 2 time windows x 7 models)
- Model selection is done on the test set, by calculating the MAPE of the one-step ahead predicitons
- The 3 and 5 best models then averaged for an ensemble predictor
- The Errors of our nowcasted values:
  - Average value of indices: 165.2
  - Average RMSE of all 5 entries: 13.9 (8.4%)
  - Weighted Average RMSE: 14.9 (9%)
  - Average MSRE of all 5 entries: 0.078
  - Weighted Average MSRE: 0.084
  - The best 5 predictions' average MSRE: 0.000054





Summary

- This method is still under development.
- The results for several countries suggest us that the direction is good, but we need some further tweaks to reach a generally well behaving estimator.
- Comparing our top 5 predictions between November and April we could provide very consistent results. Our placements are 2, 4, 2, 4, 4 and x.
- We plan to generalize further the method to become a time series nowcasting tool for general purpose.
  - Also, thinking on a web application that could help in providing estimates.
- Any suggestions are welcome!

Thank you for your attention!

Gergely Attila Kiss, gergely.kiss@ksh.hu