Nina Bury 255758
Sandra Stanowska 256566

Predictive analytics

Final report

„On the importance of the long-term seasonal component in day-ahead electricity price forecasting with NARX neural networks"

1. Introduction

The paper written by Grzegorz Marcjasz, Bartosz Uniejewski and Rafał Weron discusses the importance of seasonality in electricity price forecasting (EPF) and explores different modeling approaches to improve the accuracy of price predictions. Electricity price forecasting is a crucial aspect of energy markets and has various applications, including optimizing energy production, managing energy consumption, and making informed investment decisions. In EPF, seasonality refers to the regular and predictable patterns in electricity prices that occur over different time scales, such as daily, weekly, and long-term seasonal components. These patterns can be influenced by factors like weather, economic conditions, and market dynamics. Understanding and effectively modeling these seasonal patterns are essential for making accurate price forecasts. *Seasonality* is a crucial factor in EPF, and its appropriate treatment is essential for accurate predictions. While long-term seasonal components are considered relevant for medium-term horizons (weeks or months), daily and weekly seasonality is important for short-term forecasting. Traditionally, long-term seasonality has been ignored to avoid complexity in modeling. *The Seasonal Component AutoRegressive (SCAR)* modeling framework is introduced as a method to decompose electricity price series into a long-term seasonal component and a stochastic component. These components are modeled independently and then combined to improve forecast accuracy. SCAR is based on linear models estimated using ordinary least squares (OLS). Authors of discussed paper have used non-linear autoregressive (NARX) neural network-type models to compare outcomes of forecasting between above-mentioned models. They have been considering long-term seasonality in case of day-ahead electricity forecasting.

In the article study there were two datasets used. The source of the first one is Global Energy Forecasting Competition 2014. From this dataset authors have chosen to analyse locational marginal prices (LMP) and day-ahead predictions of system loads. The data ranges between 1.1.2011–17.12.2013. Second dataset comes from Nord Pool, which is pan-European power exchange. The data considers hourly system prices and hourly consumption prognosis values for Denmark, Finland, Norway and Sweden. It covers the dates between 1.1.2013–26.12.2015. All missing data was handled the same way as in exemplary paper, meaning the blanks were filled with arithmetic average of neighbours.

In the study it was assumed that there is rolling calibration window and the modeling is implemented separately across hours, moreover the forecast for next day is done simultaneously for all 24 hours. The methods used in the research were: naïve benchmark, expert benchmark, seasonal component autoregressive (SCAR) models, artificial neural network (ANN) models, committee machines of ANN networks, seasonal component artificial neural network (SCANN) models, committee machines of SCANN networks.

The results of forecasts were evaluated in terms of weekly-weighted mean absolute error (WMAE) loss function and shown in a table. From the results authors have draw few conclusions such as: seasonal component approach used in SCARX and SCARX5 models could better the day-ahead forecasts. Another statement regarding committee machines implies that increasing the number of averaged neural networks improves the forecasts. The best linear SCANN-5 models are outperforming the best SCARX models, but SCARX is outperforming SCANN-1 models. Considering deseasonalizing the exogenous variable (being either load or consumption forecasts) the results for the SCARX models are in line with a study by Uniejewski and show an average improvement of 0.711 for GEFCom2014 and 0.152 for Nord Pool to the original study by Nowotarski and Weron (2016), where load and consumption forecasts were not deseasonalized. The analyse in this paper suggests that models with HP filter-based LTSCs are not always outperformed by those with wavelet-based LTSCs in contrast to Nowotarski and Weron (2016) findings. The models were compared using Diebold–Mariano (DM) tests, which presents if the two forecast are significantly different. In the GEFCom2014 data the best models were SCARX and SCANN5, since they significantly outperform others. In case of Nord Pool there were four statistically significant better models: SCANN5-S10, SCANN5-HP5·109 , SCARX-S9 and ANN5. The number of hidden neurons has been set to 5, since it was performing quite well with most of the models. The hidden layer was tested with 2 to 10 neurons. The committee machine performance was tested with different numbers of averaged networks between 1 and 10. Considering performance of models, but also computational cost of enlarging the network, the optimal networks to be averaged in committee machine was 5.

The conclusions from the paper:

- The NARX-type non-linear models, similarly to linear models with OLS, enhance performance when using seasonal component approach in one-day-ahead EPF.
- Seasonal component approach used in non-linear models improves accuracy in more advantageous way than in linear model.
- The SCAR-type models can be outperformed by seasonal non-linear models.

2. Results

The first step of analysing and further developing the discussed article begins with replicating the results of the initial document.

2.1. The naïve and expert benchmarks

In this part of the project, we aim to implement and evaluate two benchmark models for electricity price forecasting (EPF) data from the GEFCOM2014 and Nord Pool datasets. The first benchmark, referred to as the Naïve method, was based on a similar-day technique introduced by Nogales, Contreras, Conejo, and Espinola (2002). This method assumes that the price forecast for a specific hour on a given day is equal to the price for the same hour on the previous week's corresponding day. We implemented this benchmark for both datasets, visualizing its performance against the actual electricity prices.

The second benchmark, termed the Expert model, was inspired by the work of Misiorek et al. (2006) and incorporates an autoregressive structure with various components. The model utilizes the natural logarithm of electricity spot prices, incorporating autoregressive effects, non-linear effects, load, and weekday dummies. We implemented and evaluated this model, demonstrating its ability to capture more complex patterns in electricity prices compared to the Naïve method.

The created code successfully loads and preprocesses the data, implements both benchmark models, and visualizes the forecasted prices against the actual prices for the GEFCOM and Nord Pool datasets. These benchmarks provide a foundation for further model development and evaluation in the domain of electricity price forecasting.

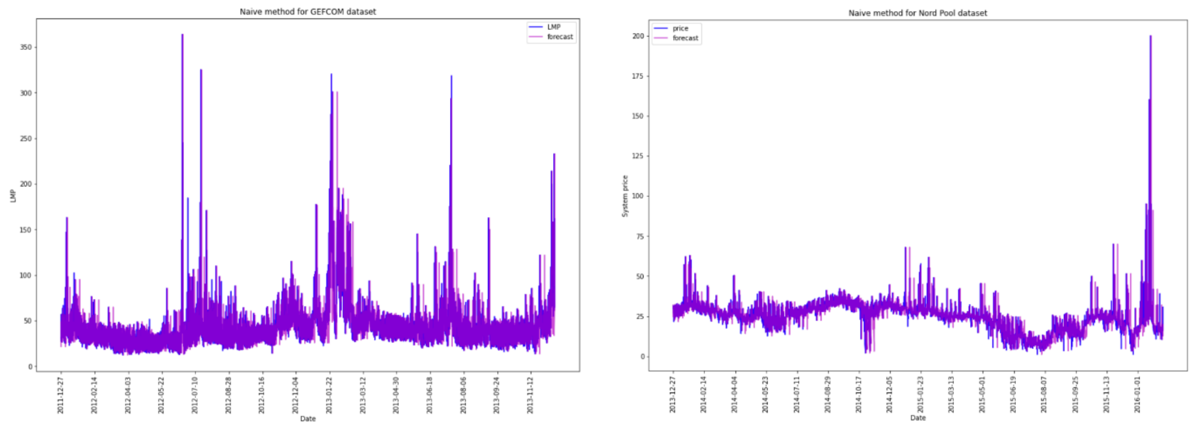

Figure 1. This figure illustrates the performance of the Naïve method in forecasting hourly electricity prices (LMP) for the GEFCOM dataset and system prices for the Nord Pool dataset. The x-axis represents the date, while the y-axis depicts the electricity spot prices. The blue line represents the actual hourly prices, while the magenta line corresponds to the Naïve method forecast.

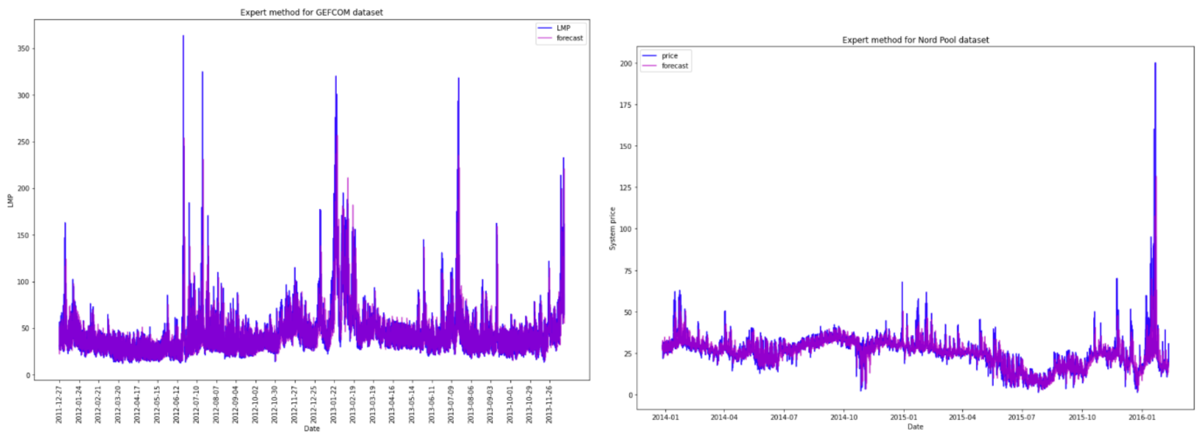

Figure 2, This figure serves as a comparative analysis to Figure 1, presenting the performance of the Expert method in forecasting hourly electricity prices (LMP) for the GEFCOM dataset and system prices for the Nord Pool dataset. Similar to the Naïve method, the x-axis corresponds to the date, and the y-axis represents the electricity spot prices. The blue line represents the actual hourly prices, while the magenta line illustrates the Expert method forecast.

2.2 The seasonal component autoregressive (SCAR) benchmarks

The SCAR modeling framework, inspired by standard seasonal decomposition approaches, introduces a nuanced perspective to electricity log-price forecasting. In particular, the framework decomposes the log-prices into a Long-Term Seasonal Component (LTSC) and a stochastic component with short-term periodicities. A critical enhancement, proposed by Nowotarski and Weron (2016), involves deseasonalizing the exogenous series using the same LTSC as prices before incorporating it into the model. This modification demonstrates significant improvements in forecasting performance, as highlighted by Uniejewski, Marcjasz, and Weron. The SCAR algorithm unfolds in four steps: decomposing log-prices, decomposing the exogenous series, calibrating an Autoregressive model (ARX) on the stochastic component, and combining ARX forecasts with persistent LTSC forecasts to yield log-price forecasts. The final step involves converting log-price forecasts into price forecasts, completing the SCARX model. Notably, the framework explores 18 SCARX models, varying in the choice of LTSC filters, including wavelet and Hodrick & Prescott filters. This diverse approach accommodates different smoothers, contributing to the adaptability of the SCARX model to the unique characteristics of electricity markets.

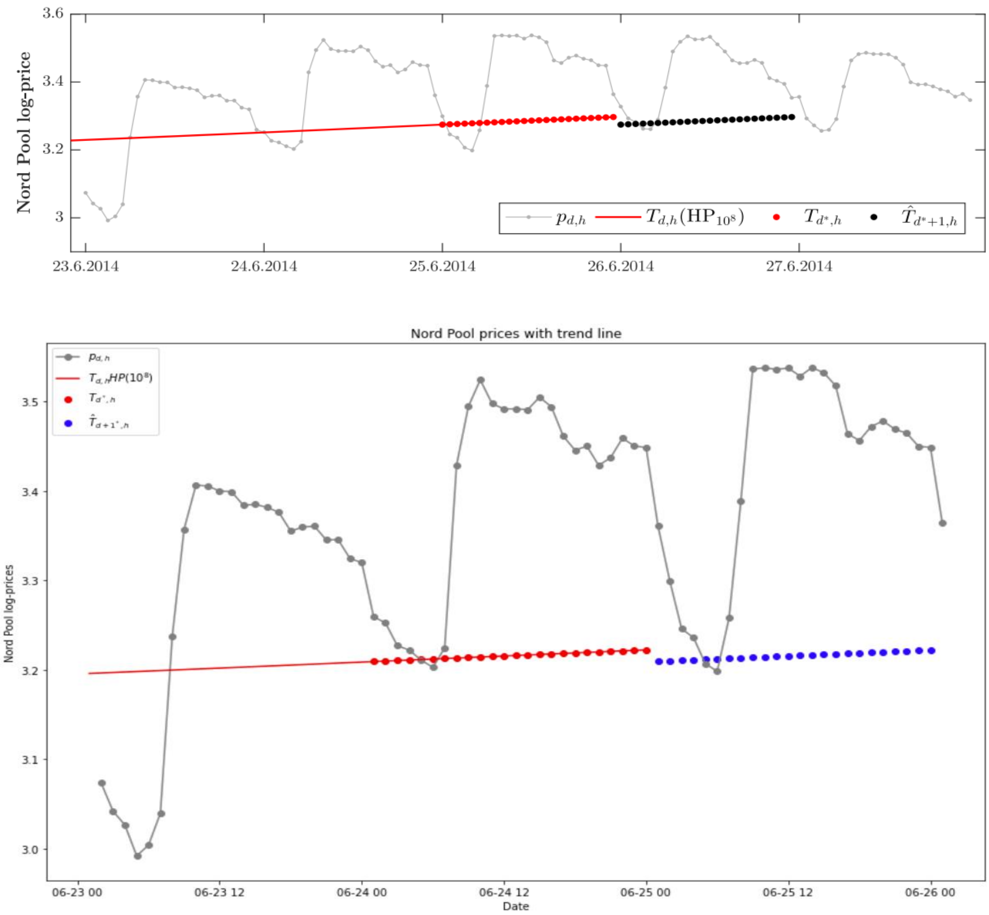

Figure 3. This figure presents a side-by-side comparison of our generated results (bottom plot) and the corresponding illustration from the article, capturing the Hodrick-Prescott (HP) filter-based Long-Term Seasonal Component (LTSC). The black circles on the reference figure depict our persistent forecasts for the 24 hours of 25.06.2014, while the red circles represent the LTSC values on the last day (d∗) in the 360-day calibration window (25.06.2014).

The code successfully decomposes log-prices into a long-term seasonal component (LTSC) and a stochastic component using the Hodrick-Prescott (HP) filter. The red circles showcased on both plots are aligned with each other confirming correct replication of the original result. Having the decomposition it is possible to carry on with further forecast procedure. Following the algorithm given in the article the SCAR benchmark was implemented and the results can be viewed in Figure 4. There is visible improvement of the forecasts comparing to naïve method and very probable enhancement in case of expert benchmark.

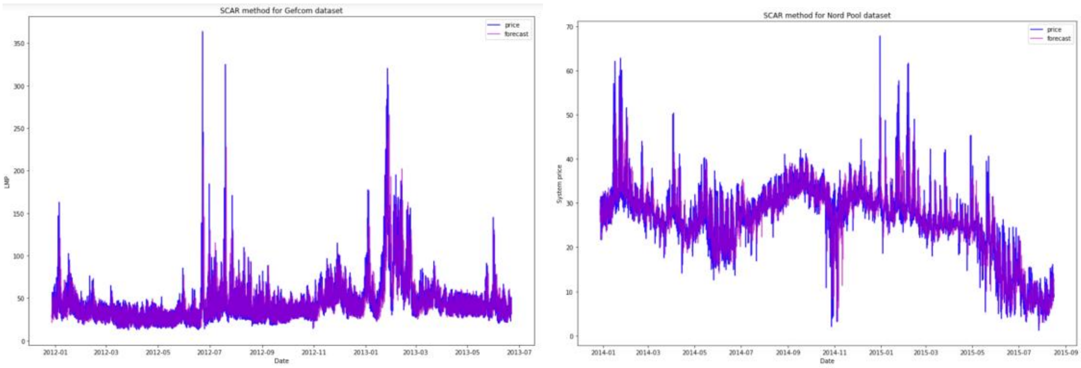

Figure 4. The above figure depicts usage of SCAR benchmark on different datasets. The plot on the left presents the comparison between the prices of electricity from GEFCom2014 dataset and forecast made with SCAR method based on Expert model. On the right plot the data comes from Nord Pool dataset and the analyzed variable is system price and the forecast was also made with SCAR benchmark. The blue lines are actual values of prices, meanwhile the magenta lines are forecasts. The x-axis represents the date and y-axis value of price.

2.3. The artificial neural network (ANN) benchmark

In this part an Artificial Neural Network (ANN) benchmark for forecasting electricity prices was implemented. The model employs a non-linear autoregressive structure (NARX), a recurrent neural network (RNN) architecture. To capture short-term dynamics, feed-forward representation was chosen instead of a Long Short-Term Memory (LSTM) approach. The ANN consists of one hidden layer with five neurons and an output layer with one neuron. Training was conducted using the Levenberg–Marquardt algorithm, with a dataset division into 96% training and 4% validation samples. Despite using load forecasts as exogenous variables, a simplified notation was used by excluding them from the model name.

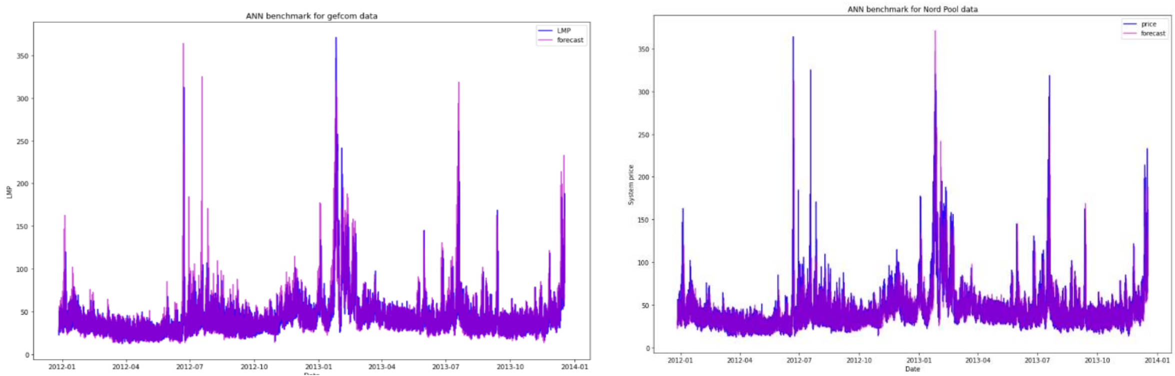

Figure 5. Those plots present ANN benchmark for GEFCom2014 hourly electricity prices on the left side and Nord Pool system prices on the right side. The forecasts are marked with magenta line and the actual values of the LMP/system price are colored with blue. On the x-axis represents the date, while the y-axis exhibits the prices.

3. Performance evaluation of the models

3.1 *Weekly-weighted mean absolute error*

In the article the goodness of fit of the statistical models were measured by weekly-weighted mean absolute error (WMAE) loss function. The formula is shown on Figure 5 and calculated measures for chosen forecasts are shown in Table 1. for Nord Pool data.

$$WMAE = \frac{1}{\bar{P}_{168}} MAE_i = \frac{1}{168 \cdot \bar{P}_{168}} \sum_{d=1}^{7} \sum_{h=1}^{24} |P_{d,h} - \hat{P}^{i}_{d,h}|$$

Figure 5. depicts formula for WMAE.

Model	WMAE

Naive	13.8858
ARX	12.0027
SCAR	12.3938
ANN_1	11.6669

Table 1. shows percentage values of average WMAE for GEFCOM dataset.

WMAE values presented in Table 1. contain information that the error was the smallest in case of ANN_1 benchmark for GEFCom2014 data. The worst performance shows naive method.

Model	WMAE
Naive	9.6610
ARX	8.7896
SCAR	8.6739

Table 2. shows percentage values of average WMAE for Nord Pool dataset.

In the case of Nord Pool data the table of WMAE shows that forecasts made with SCAR model were the closest to the actual values of system price, but also quite close ranks ARX benchmark. Here just as in GEFCom dataset the naive model achieved the worst results.

4. Further development

*4.1. T*he autoregressive integrated moving average (ARIMA) benchmark

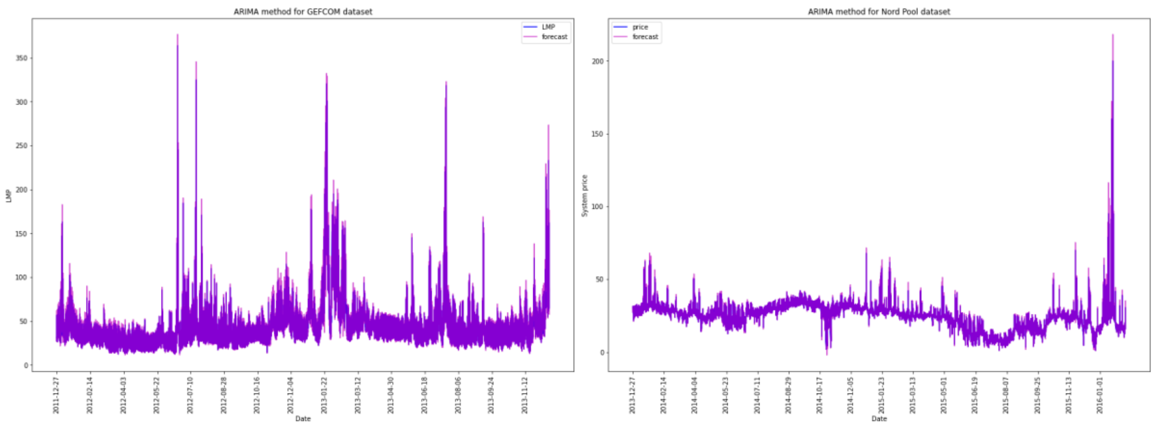

Figure 6. The plots present ARIMA benchmark for GEFCom2014 hourly electricity prices on the left side and Nord Pool system prices on the right side. The forecasts are marked with magenta line and the actual values of the LMP/system price are colored with blue. On the x-axis represents the date, while the y-axis exhibits the prices.

In order to further develop our report, the autoregressive integrated moving average (ARIMA) method was implemented. ARIMA is a time series forecasting technique that combines autoregressive (AR) and moving average (MA) components with differencing to achieve stationarity. The model's three main parameters—p (order of autoregression), d (degree of differencing), and q (order of moving average)—allow flexibility in accommodating various time series structures.). As the ARIMA algorithm is known for its simplicity and effectiveness in capturing linear trends and seasonality, its application provides a benchmark for evaluating forecasting performance. Due to the computational demands of ARIMA, the frequency of our analysis had to be adjusted from hourly to daily. This modification aimed to improve computational efficiency, considering the potential challenges posed by the lengthy computation times observed during hourly forecasting. The results obtained from ARIMA for GEFCOM and Nord Pool datasets align with the trends seen in other forecasting methodologies used in this report.

5. Conclusions

Long term predictions are complicated subject and, in most cases cannot be effective with basic statistical models. The approach presented in the article by Grzegorz Marcjasz, Bartosz Uniejewski and Rafał Weron has been succesfully verifed during our study. The forecasts made based on the 4 benchmarks: Naive, ARX, SCAR, NARX, obtained good results even when tested on 103 (for GEFCom, 104 for Nord Pool) weeks out-of-sample test period. We can conlude that neural networks are time-consuming, but the outcomes of the forecasts carry the least error.

6. Bibliography:

(1) Marcjasz G., Uniejewski B., Weron R., (2019) On the importance of the long-term seasonal component in day-ahead electricity price forecasting with NARX neural networks, *International Journal of Forecasting 35* (1520-1535).

(2) Misiorek, Adam, Trueck, Stefan and Weron, Rafal. (2006) Point and Interval Forecasting of Spot Electricity Prices: Linear vs. Non-Linear Time Series Models, *Studies in Nonlinear Dynamics & Econometrics*, vol. 10, no. 3. https://doi.org/10.2202/1558-3708.1362