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

Grzegorz Marcjasz[1,2], Bartosz Uniejewski[1,2] & Rafał Weron[1]

[1]Department of Operations Research

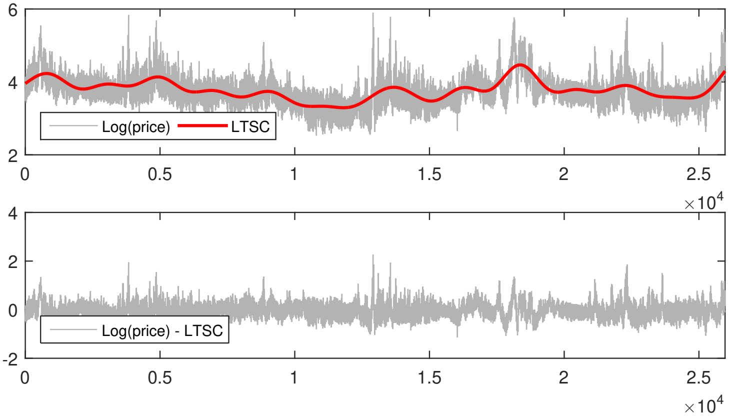
and

[2]Faculty of Pure and Applied Mathematics
Wrocław University of Science and Technology, Poland

*Partially supported by the National Science Center (NCN, PL) through grant no. 2015/17/B/HS4/00334

LTSC and short-term price forecasting

- Does removing the long-term seasonal component (LTSC) improve short-term (day-ahead) electricity price forecasts?



Yes, it does!

Energy Economics 57 (2016) 228–235

Contents lists available at ScienceDirect

Energy Economics

journal homepage: www.elsevier.com/locate/eneco

On the importance of the long-term seasonal component in day-ahead electricity price forecasting

Jakub Nowotarski, Rafał Weron*

Department of Operations Research, Wrocław University of Technology, Wrocław, Poland

ARTICLE INFO

Article history:
Received 16 March 2016
Received in revised form 21 May 2016
Accepted 25 May 2016
Available online 2 June 2016

JEL classification:

ABSTRACT

In day-ahead *electricity price forecasting* (EPF) the daily and weekly seasonalities are always taken into account, but the long-term seasonal component (LTSC) is believed to add unnecessary complexity to the already parameter-rich models and is generally ignored. Conducting an extensive empirical study involving state-of-the-art time series models we show that (i) decomposing a series of electricity prices into a LTSC and a stochastic component, (ii) modeling them independently and (iii) combining their forecasts can bring – contrary to a common belief – an accuracy gain compared to an approach in which a given time series model is calibrated to the prices themselves.

Yes, it does!

Percentage (no.) of hours for which a SCARX model significantly outperforms the ARX benchmark (#better) and vice versa (#worse):

	GEFCom2014	Nord Pool
	SCARX-S_{12}	SCARX-S_9
#better	46% (11)	29% (7)
#worse	0% (0)	17% (4)
	mSCARX-S_{12}	mSCARX-S_{10}
#better	92% (22)	42% (10)
#worse	0% (0)	0% (0)

- Only Seasonal Component ARX vs. ARX models tested
- **But is this phenomenon more general? E.g., for ANNs?**

This study is based on a forthcoming paper:

International Journal of Forecasting (2018)

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

ELSEVIER

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

Grzegorz Marcjasz[a,b], Bartosz Uniejewski[a,b], Rafał Weron[a]

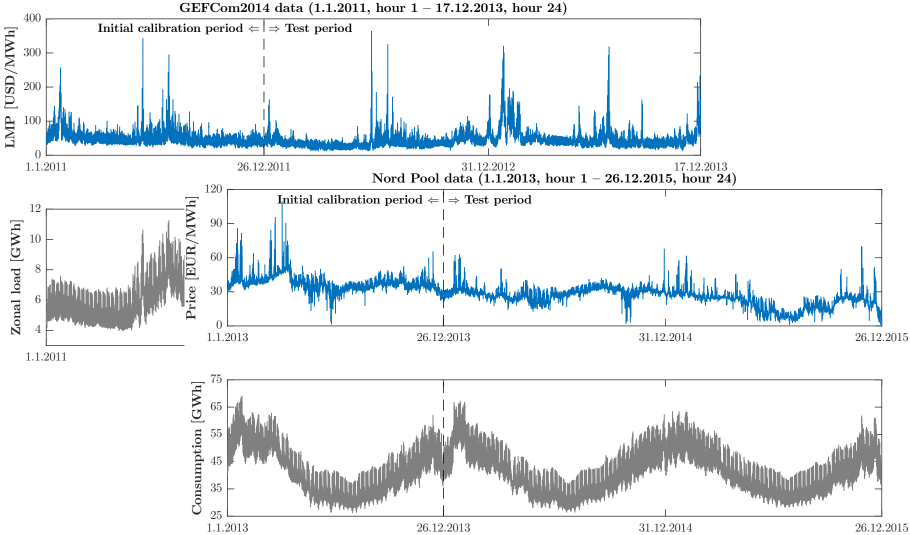
[a] *Department of Operations Research, Wrocław University of Science and Technology, Wrocław, Poland*
[b] *Faculty of Pure and Applied Mathematics, Wrocław University of Science and Technology, Wrocław, Poland*

Abstract

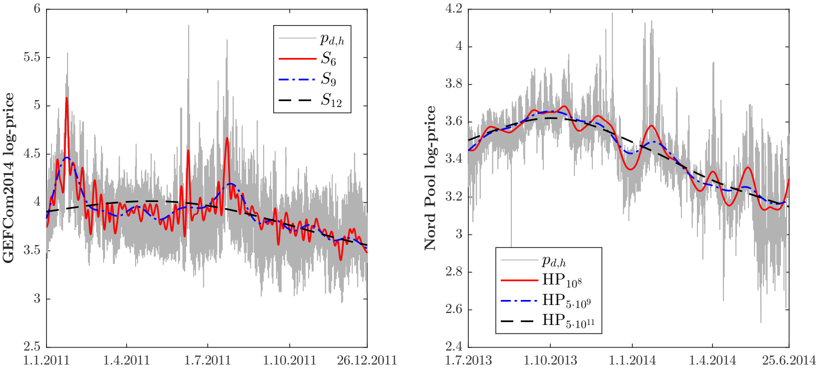
In day-ahead *electricity price forecasting* the daily and weekly seasonalities are always taken into account, but the long-term seasonal component was believed to add unnecessary complexity and in most studies ignored. The recent introduction of the *Seasonal Component AutoRegressive* (SCAR) modeling framework has changed this viewpoint. However, the latter is based on linear models estimated using Ordinary Least Squares. Here we show that considering non-linear autoregressive (NARX) neural network-type models with the same inputs as the corresponding SCAR-type model

Study setup

The same as in Nowotarski & Weron (2016)



18 wavelet and HP-filter based LTSCs



- **Wavelet filters (-S$_J$)**: $S_5, S_6, \ldots, S_{14}$, ranging from 'daily' smoothing ($S_5 \to 2^5$ hours) up to 'biannual' ($S_{14} \to 2^{14}$ hours)
- **HP-filters (-HP$_\lambda$)**: with $\lambda = 10^8, 5\cdot10^8, 10^9, \ldots, 5\cdot10^{11}$

The **ARX** model

For the log-price, i.e., $p_{d,h} = log(P_{d,h})$, the model is given by:

$$p_{d,h} = \underbrace{\beta_{h,1} p_{d-1,h} + \beta_{h,2} p_{d-2,h} + \beta_{h,3} p_{d-7,h}}_{\text{autoregressive effects}} + \underbrace{\beta_{h,4} p_{d-1,\text{min}}}_{\text{non-linear effect}}$$
$$+ \underbrace{\beta_{h,5} z_t}_{\text{load forecast}} + \underbrace{\sum_{i=1}^{3} \beta_{h,i+5} D_i}_{\text{Mon, Sat, Sun dummies}} + \varepsilon_{d,h} \qquad (1)$$

- $p_{d-1,min}$ is yesterday's minimum hourly price
- z_t is the logarithm of system load/consumption
- Dummy variables D_1, D_2 and D_3 refer to Monday, Saturday and Sunday, respectively

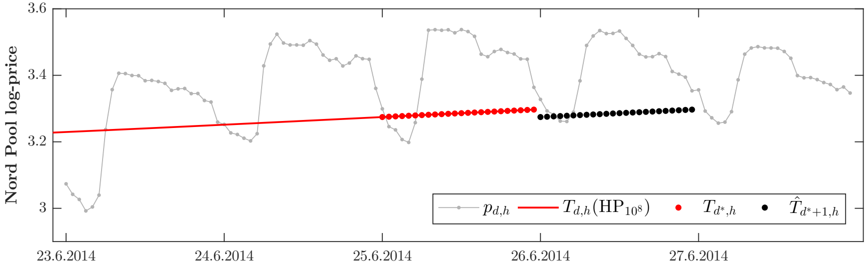
The SCAR modeling framework

(Nowotarski & Weron, 2016, ENEECO; Uniejewski, Marcjasz & Weron, 2017, WP)

The *Seasonal Component AutoRegressive* (SCAR) modeling framework consists of the following steps:

1. (a) Decompose the log-price in the calibration window into the LTSC $T_{d,h}$ and the stochastic component $q_{d,h}$
   (b) Decompose the exogenous series in the calibration window using the same type of LTSC as for prices
2. Calibrate the **ARX** model to q_t and compute forecasts for the 24 hours of the next day (24 separate series)

The SCAR modeling framework cont.

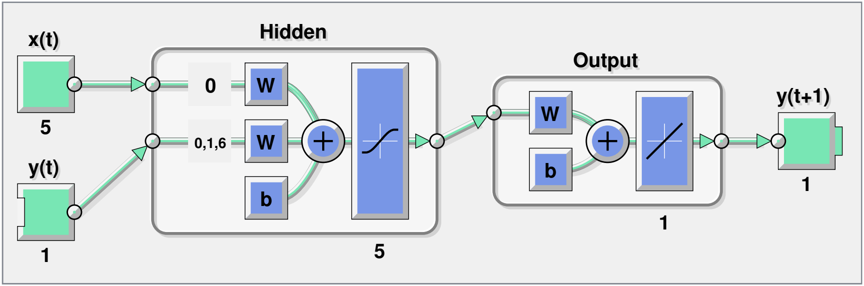


3. Add stochastic component forecasts $\hat{q}_{d+1,h}$ to persistent forecasts $\hat{T}_{d+1,h}$ of the LTSC to yield log-price forecasts $\hat{p}_{d+1,h}$
4. Convert them into price forecasts of the **SCARX** model, i.e., $\hat{P}_{d+1,h} = \exp\left(\hat{p}_{d+1,h}\right)$

ANNs in other EPF studies

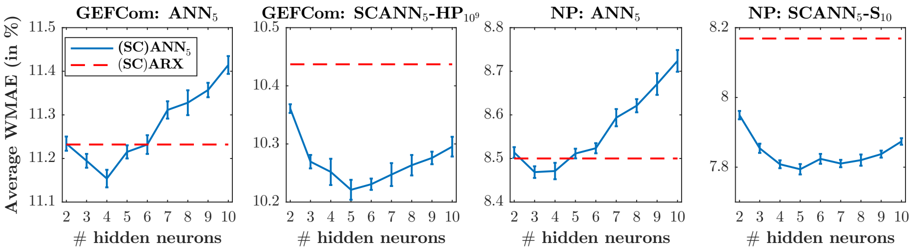
- A variety of ANN implementations
- Different datasets and inputs → impossible to compare with published studies that use regression models
- A few papers acknowledge the need for deseasonalizing data before fitting neural network models:
  - Andrawis et al. (2011)
  - Zhang and Qi (2005)
  - Keles et al. (2016), the only one in the context of EPF

ANN: Based on Matlab's NARXnet



- One hidden layer with 5 neurons and sigmoid activation functions
- Inputs identical as in the **ARX** model
- Trained using Matlab's `trainlm` function, utilizing the Levenberg-Marquardt algorithm for supervised learning

Number of hidden neurons

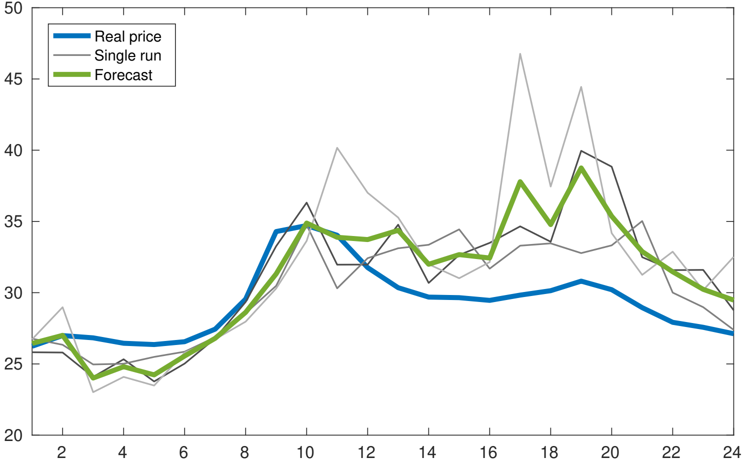


There is no universally optimal number, but the errors are smallest for 4 to 6 neurons in the hidden layer

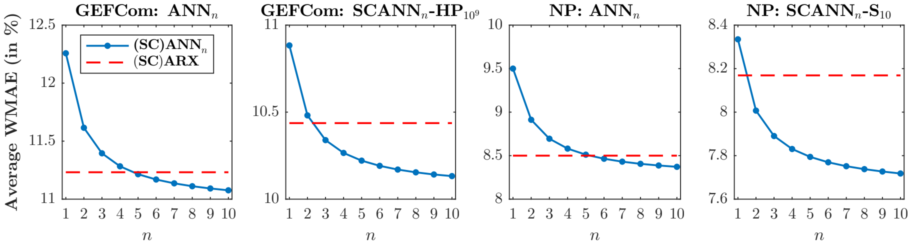
Committee machines of **(SC)ANN** networks

- The training function finds only local minima and initial weights are random
- Every forecast yields slightly different results $\Rightarrow$ two 'models' are considered:
  - $\overline{\textbf{ANN}}_1$ – the 'expected' result for a single **ANN** network, an average of error scores across separate runs
  - $\textbf{ANN}_5$ – a forecast average of **five** runs (hour-by-hour) with identical parameters, a so-called **committee machine**

Committee machines of **(SC)ANN** networks



Sample gains from using committee machines



- Forecast errors roughly scale as a power-law function of the number of networks in a committee machine
- We should use as large committee machines as we can ...

Sample gains cont.

- ... however, the time needed may be substantial, e.g., for generating forecasts for the next 24 hours:

Model	ARX	SCARX-HP$_{10^8}$	SCARX-S$_9$	ANN$_1$	ANN$_5$
Time	8.6ms	13.5ms	37.3ms	7.6s	38.2s

- SCANN times are omitted here, because LTSC computation is negligible compared to training the ANN

Results

Weekly-weighted Mean Absolute Error (WMAE)

- Following Conejo et al. (2005), Weron & Misiorek (2008) and Nowotarski et al. (2014), among others, we use:

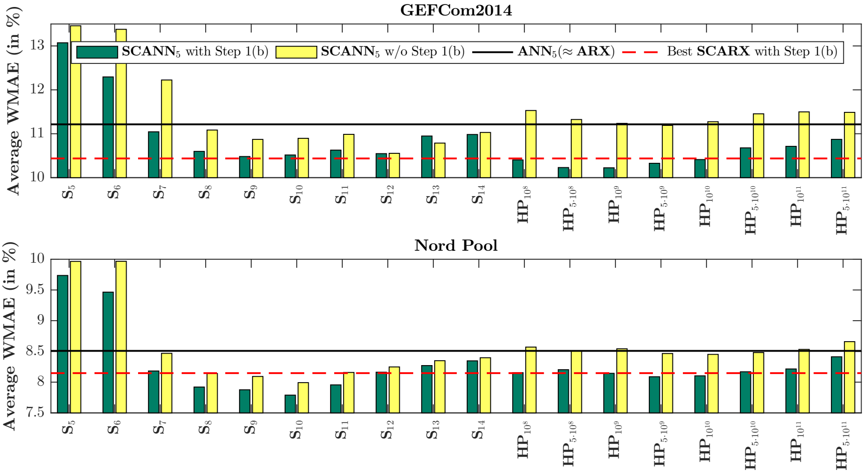
$$\mathrm{WMAE}_w = \frac{1}{\bar{P}_{168}}\mathrm{MAE}_w = \frac{1}{168 \cdot \bar{P}_{168}} \sum_{d=Mon}^{Sun} \sum_{h=1}^{24} \left| P_{d,h} - \hat{P}_{d,h} \right|$$

- where $\bar{P}_{168} = \frac{1}{168} \sum_{d=Mon}^{Sun} \sum_{h=1}^{24} P_{d,h}$

$$\overline{\mathrm{WMAE}} = \frac{1}{w_{max}} \sum_{w=1}^{w_{max}} \mathrm{WMAE}_w$$

- where $w_{max} = 103$ for GEFCom and 104 for Nord Pool

Aggregate results of SCANN performance



Note: Step 1(b) is important (green vs. yellow)!

Testing for significance: Diebold-Mariano

- We define the error function as

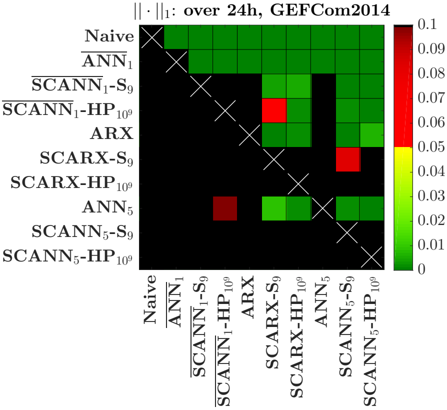
$$L(\varepsilon_d) = ||\varepsilon_d||_1 = \sum_{h=1}^{24} |P_{d,h} - \hat{P}_{d,h}|$$

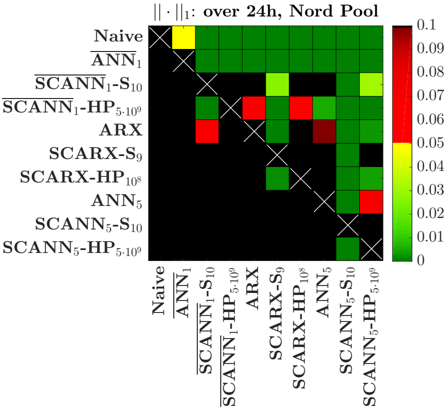
- For each pair of models we compute the loss differential

$$D_d = L(\varepsilon_d^{model_X}) - L(\varepsilon_d^{model_Y})$$

- H_0:$E(D_d) \leq 0$, forecasts of $model_X$ outperform those of $model_Y$
- H_0^R:$E(D_d) \geq 0$, i.e., the reverse hypothesis

Diebold-Mariano test: *p*-values



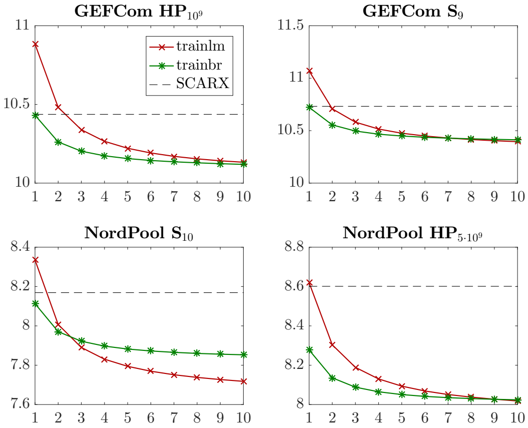


The closer are the *p*-values to zero (dark green) the more significant is the difference between the forecasts of a model on the X-axis ($\rightarrow$ better) and those of a model on the Y-axis ($\rightarrow$ worse)

Alternative networks and training methods

Alternative training methods: MATLAB's `trainbr`

- Better results for a single run than with `trainlm`: $\overline{\textbf{(SC)ANN}}_1$
- Lower gains from averaging: $\textbf{(SC)ANN}_5$

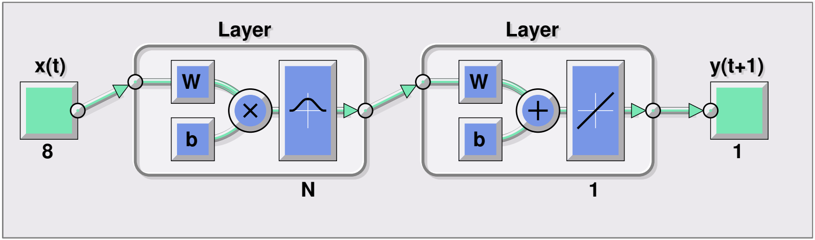


Alternative training methods: `PythonFANN`

- Python interface to the Fast Artificial Neural Network Library

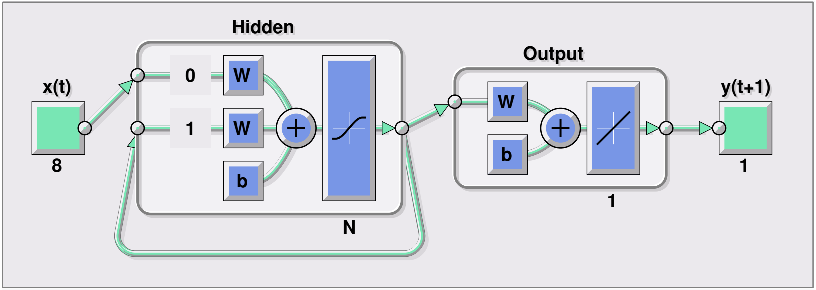
- Unmatched performance for some data periods ...
- ... while on average similar to MATLAB's `trainlm`
- Training algorithms and execution time
- Default training parameters $\Rightarrow$ terrible forecasting performance

Radial Basis Function (RBF) networks



- Potentially good interpolation with many radial basis functions
- Iteratively adds hidden neurons ⇒ high computational cost

(Layer) Recurrent Neural Networks (RNNs)



- Dynamic temporal behavior ⇒ prediction largely based on the input sequence
- Various types, e.g., Long Short-Term Memory (LSTM) networks

Conclusions

Conclusions

- Using Seasonal Component ANN (SCANN) models can yield statistically significant improvement over the ANN benchmark
  - $\mathbf{SCANN}_5$ returns 0.72–0.99% lower WMAE than $\mathbf{ANN}_5$
- The accuracy gains from using LTSC are greater in ANN models than in regression models
  - **SCARX** models yield only a 0.35–0.80% improvement in WMAE vs. the **ARX** benchmark

Conclusions cont.

- Forecast averaging is crucial in outperforming the **SCARX** model
  - $\mathbf{SCANN}_5$ yields 0.21–0.36% lower WMAE than corresponding **SCARX** models ...
  - whereas $\overline{\mathbf{SCANN}}_1$ returns 0.22–1.02% higher WMAE than **SCARX**

International Journal of Forecasting (2018)

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

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

Grzegorz Marcjasz[a,b], Bartosz Uniejewski[a,b], Rafał Weron[a]

[a] *Department of Operations Research, Wrocław University of Science and Technology, Wrocław, Poland*
[b] *Faculty of Pure and Applied Mathematics, Wrocław University of Science and Technology, Wrocław, Poland*

Abstract

In day-ahead *electricity price forecasting* the daily and weekly seasonalities are always taken into account, but the long-term seasonal component was believed to add unnecessary complexity and in most studies ignored. The recent introduction of the *Seasonal Component AutoRegressive* (SCAR) modeling framework has changed this viewpoint. However, the latter is based on linear models estimated using Ordinary Least Squares. Here we show that considering non-linear autoregressive (NARX) neural network-type models with the same inputs as the corresponding SCAR-type model

Bibliography

1. G. Marcjasz, B. Uniejewski, R. Weron (2018) *Importance of the long-term seasonal component in day-ahead electricity price forecasting with NARX neural networks*, International Journal of Forecasting, *forthcoming*. Earlier working paper version available from RePEc: https://ideas.repec.org/p/wuu/wpaper/hsc1703.html

2. J. Nowotarski, R. Weron (2016) *On the importance of the long-term seasonal component in day-ahead electricity price forecasting*, Energy Economics 57, 228-235
3. B. Uniejewski, G. Marcjasz, R. Weron (2017) *On the importance of the long-term seasonal component in day-ahead electricity price forecasting. Part II - Probabilistic forecasting*, submitted. Available from RePEc: https://ideas.repec.org/p/wuu/wpaper/hsc1702.html
4. R. Weron (2014) *Electricity price forecasting: A review of the state-of-the-art with a look into the future*, International Journal of Forecasting 30(4), 1030-1081
5. F. Ziel, R. Weron (2018) *Day-ahead electricity price forecasting with high-dimensional structures: Univariate vs. multivariate modeling frameworks*, Energy Economics, *forthcoming*. Available from RePEc: https://ideas.repec.org/p/wuu/wpaper/hsc1608.html