On the Evaluation of Multivariate Event Probability Predictions in Electricity Price Forecasting

Arne Vogler & Florian Ziel

House of Energy Markets and Finance
University of Duisburg-Essen

12.12.2018

Motivation

- Probabilistic Electricity Price Forecasting (EPF)
  - either full Conditional Multivariate Distribution
  - or Densities, Intervals, Quantiles
- Need to assess Statistical Validity of Distribution Forecasts
- It may often suffice
  - to define statistical event of interest
  - to forecast associated probability
  - to evaluate probability forecast methodology
- Multivariate Events in EPF
- No coherent Framework for Event Probability Forecasts in EPF

Events

Event 1

Probability of n consecutive Hours of negative Prices

Event 2

Probability of peak-base Price Spread being greater than/equal to Threshold a [1]

Event 3

Probability of a pumped-storage Plant Operation being profitable during a given Day

[1] Note that Spread is also calculated for Saturdays and Sundays.

Overview

Models

- Naive Model

$$P_{t,h} = \begin{cases} P_{t-7,h} + \epsilon_{t,h} & , t = Sat, Sun, Mon \\ P_{t-1,h} + \epsilon_{t,h} & , otherwise \end{cases}$$

- Expert Model

$$\begin{aligned} P_{t,h} = \beta_{h,0} + \beta_{h,1} P_{t-1,h} + \beta_{h,2} P_{t-2,h} + \beta_{h,3} P_{t-7,h} \\ + \beta_{h,4} P^{Max}_{t-1,h} + \beta_{h,5} P^{Min}_{t-1,h} + \beta_{h,6} P_{t-1,24} \\ + \sum_{i=1}^{6} \beta_{h,6+i} D_i + \epsilon_{t,h} \end{aligned}$$

Estimation, Simulation and Probability Forecasts

- Sample and Estimation
  - German Day-Ahead Prices 2014-2017
  - Mutivariate approach
  - Rolling window of $N = 730$ days
- Simulation
  - Simulation of German Day-Ahead Prices for 2016-2017
  - Based on Vector of Residuals $\hat{\epsilon}_t = [\hat{\epsilon}_{t,1}, ..., \hat{\epsilon}_{t,24}], t = 1...N$
  - 3 Methods - Bootstrap, Multivariate Normal, Multivariate Student's t
- Probability Forecast
  - Probability Forecasts for 2016-2017
  - Relative Frequency across Ensemble of Simulated Paths

Summary of Specifications

N-Boot	**N-Norm**	**N-t**
Naive - Bootstrap	Naive - Normal	Naive - Student's t
Ex-Boot	**Ex-Norm**	**Ex-t**
Expert OLS Bootstrap	Expert OLS Normal	Expert OLS Student's t
QREx-Boot	**QREx-Norm**	**QREx-t**
Expert QR($\tau = 0.5$) Bootstrap	Expert QR($\tau = 0.5$) Normal	Expert QR($\tau = 0.5$) Student's t

Supervised Classification

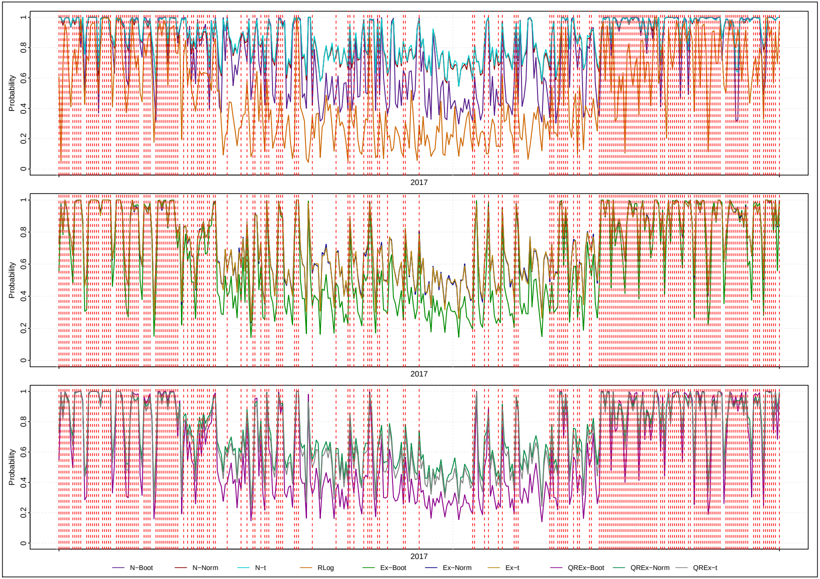
- Classification Approach
  - Historical Price and Event Indicator Series
  - Classification Model for Paths
  - Logistic Regression with Regularization (**RLog**)

$$P(I_t^E|X) = \frac{1}{1 + e^{-X\beta}}$$

$$X\beta = \beta_0 + \sum_{h=1}^{24} \sum_{k\in\{1,2,7\}} \beta_{h,k} P_{t-k,h} + \sum_{i=1}^{6} \beta_i D_i + \epsilon_t$$

- Probability Forecast
  - Probability Forecasts for 2016-2017
  - Model directly provides Forecast

Event III - Probability Time Series



Overview

Quadratic Probability Score

Quadratic Probability Score

$$QPS(f, x) = \frac{1}{T} \sum_{t=1}^{T} (f_t - x_t)^2$$

$\Rightarrow$ **Diebold Mariano Test** can be used with QPS

Murphy Decomposition of QPS

Murphy Decomposition

$$QPS(f,x) = \underbrace{\bar{x}(1-\bar{x})}_{\text{UNC}} + \underbrace{\frac{1}{T}\sum_{j=1}^{J} T_j(\bar{f}_j - \bar{x}_j)^2}_{\text{REL}} - \underbrace{\frac{1}{T}\sum_{j=1}^{J} T_j(\bar{x}_j - \bar{x})^2}_{\text{RES}}$$
$$+ \underbrace{\frac{1}{T}\sum_{j=1}^{J}\sum_{t=1}^{T_j}(f_{tj} - \bar{f}_j)^2}_{\text{WBV}} - \underbrace{\frac{1}{T}\sum_{j=1}^{J}\sum_{t=1}^{T_j}(x_{tj} - \bar{x}_j)(f_{tj} - \bar{f}_j)}_{\text{WBC}}$$

ROC Curve and AUROC

- Contingency Table for Discrete Classifier

	$x_t = 1$	$x_t = 0$
$f_t = 1$	True P	False P
$f_t = 0$	False N	True N
	P	N

- ROC Curve
  - Define $FPR = \frac{FP}{N}$ and $TPR = \frac{TP}{P}$
  - ROC Space: Two-dimensional Graph of FPR against TPR
  - Discrete Classifier: Point in ROC Space
  - Probabilistic Classifiers with varying Threshold: Curve in ROC Space
- AUROC: Aggregate Measure of Classifier Performance

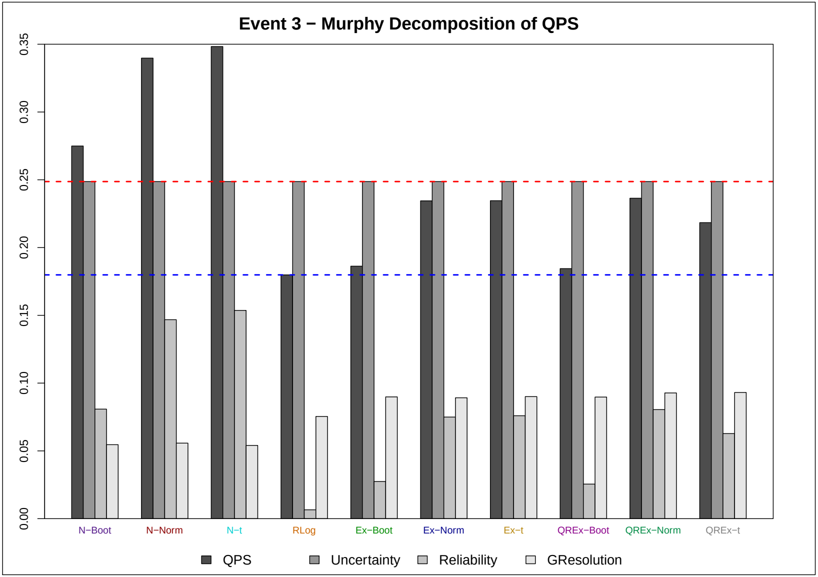
Overview

Quadratic Probability Score

Event	N-Boot	N-Norm	N-t	RLog	Ex-Boot
1	0.0313	0.0321	0.0323	0.0339	0.0156
2	0.0620	0.0611	0.0621	0.1441	0.0561
3	0.2749	0.3398	0.3483	<u>0.1798</u>	0.1863
Event	**Ex-Norm**	**Ex-t**	**QREx-Boot**	**QREx-Norm**	**QREx-t**
1	0.0154	0.0154	0.0141	0.0138	<u>0.0137</u>
2	0.0537	<u>0.0535</u>	0.0572	0.0544	0.0548
3	0.2345	0.2346	0.1845	0.2364	0.2184

Table: Quadratic Probability Score

Event III - Murphy Decomposition



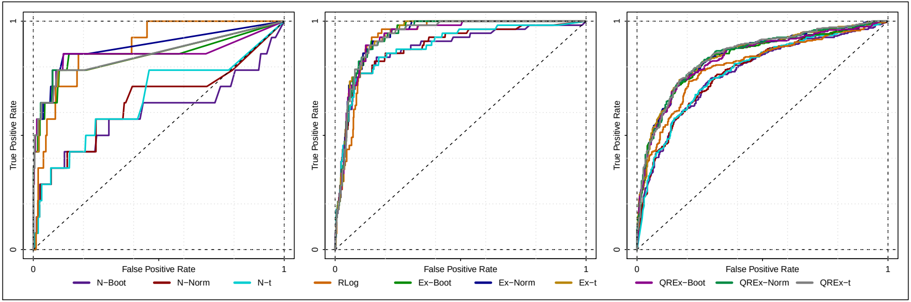
Diebold Mariano Test

	N−Boot	N−Norm	N−t	RLog	Ex−Boot	Ex−Norm	Ex−t	QREx−Boot	QREx−Norm	QREx−t
N−Boot	NA	0.96	0.98	0.65	0.00	0.00	0.00	0.00	0.00	0.00
N−Norm	0.04	NA	0.95	0.61	0.00	0.00	0.00	0.00	0.00	0.00
N−t	0.02	0.05	NA	0.59	0.00	0.00	0.00	0.00	0.00	0.00
RLog	0.35	0.39	0.41	NA	0.00	0.00	0.00	0.00	0.00	0.00
Ex−Boot	1.00	1.00	1.00	1.00	NA	0.23	0.24	0.06	0.05	0.05
Ex−Norm	1.00	1.00	1.00	1.00	0.77	NA	0.70	0.08	0.05	0.05
Ex−t	1.00	1.00	1.00	1.00	0.76	0.30	NA	0.07	0.04	0.04
QREx−Boot	1.00	1.00	1.00	1.00	0.94	0.92	0.93	NA	0.24	0.20
QREx−Norm	1.00	1.00	1.00	1.00	0.95	0.95	0.96	0.76	NA	0.26
QREx−t	1.00	1.00	1.00	1.00	0.95	0.95	0.96	0.80	0.74	NA

	N−Boot	N−Norm	N−t	RLog	Ex−Boot	Ex−Norm	Ex−t	QREx−Boot	QREx−Norm	QREx−t
N−Boot	NA	0.11	0.55	1.00	0.14	0.06	0.05	0.19	0.07	0.09
N−Norm	0.89	NA	1.00	1.00	0.17	0.07	0.07	0.23	0.09	0.11
N−t	0.45	0.00	NA	1.00	0.14	0.05	0.05	0.19	0.07	0.08
RLog	0.00	0.00	0.00	NA	0.00	0.00	0.00	0.00	0.00	0.00
Ex−Boot	0.86	0.83	0.86	1.00	NA	0.00	0.00	0.73	0.16	0.22
Ex−Norm	0.94	0.93	0.95	1.00	1.00	NA	0.11	0.97	0.69	0.75
Ex−t	0.95	0.93	0.95	1.00	1.00	0.89	NA	0.98	0.75	0.79
QREx−Boot	0.81	0.77	0.81	1.00	0.27	0.03	0.02	NA	0.00	0.00
QREx−Norm	0.93	0.91	0.93	1.00	0.84	0.31	0.25	1.00	NA	0.86
QREx−t	0.91	0.89	0.92	1.00	0.78	0.25	0.21	1.00	0.14	NA

	N−Boot	N−Norm	N−t	RLog	Ex−Boot	Ex−Norm	Ex−t	QREx−Boot	QREx−Norm	QREx−t
N−Boot	NA	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
N−Norm	0.00	NA	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
N−t	0.00	0.00	NA	0.00	0.00	0.00	0.00	0.00	0.00	0.00
RLog	1.00	1.00	1.00	NA	0.83	1.00	1.00	0.76	1.00	1.00
Ex−Boot	1.00	1.00	1.00	0.17	NA	1.00	1.00	0.17	1.00	1.00
Ex−Norm	1.00	1.00	1.00	0.00	0.00	NA	0.60	0.00	0.91	0.00
Ex−t	1.00	1.00	1.00	0.00	0.00	0.40	NA	0.00	0.89	0.00
QREx−Boot	1.00	1.00	1.00	0.24	0.83	1.00	1.00	NA	1.00	1.00
QREx−Norm	1.00	1.00	1.00	0.00	0.00	0.09	0.11	0.00	NA	0.00
QREx−t	1.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00	1.00	NA

ROC Curve and AUROC



N-Boot	N-Norm	N-t	RLog	Ex-Boot
0.775	0.782	0.782	0.805	0.839
Ex-Norm	**Ex-t**	**QREx-Boot**	**QREx-Norm**	**QREx-t**
0.847	0.847	0.838	0.849	0.849

Table: Area under ROC Curve

Conclusion

- Calculate Probability Forecasts from well-established Expert Models
- Present Evaluation Framework for Probability Forecasts
- Framework
  - fits and extends existing EPF Evaluation Framework
  - uses MSE-equivalent QPS and Diebold Mariano Test
  - allows for further Insights about Forecast Deficiency
  - allows for Evaluation of Rare Event Forecasts
  - can be extended by ML Performance Measures
- Demonstrate Applicability using three illustrative Events
- Further Research
  - Robustness to Simulation Size
  - Skill Score and Decompositions
  - Additional ML Performance Measures

References I

Bradley, P. S., Bennett, K. P., & Demiriz, A. (2000).
Constrained k-means clustering.
Microsoft Research, Redmond, 1-8.

Diebold, F. X., & Mariano, R. S. (1995).
Comparing predictive accuracy.
Journal of Business & Economic Statistics, 13(6), 253-263.

Fawcett, T. (2006).
An introduction to ROC analysis.
Pattern Recognition Letters, 27(8), 861-874.

Hand, D. J. (2009).
Measuring classifier performance: a coherent alternative to the area under the ROC curve.
Machine Learning, 77(1), 103-123.

References II

Harvey, D., Leybourne, S., & Newbold, P. (1997).
Testing the equality of prediction mean squared errors.
International Journal of Forecasting, 13(2), 281-291.

Lahiri, K., & Wang, J. G. (2013).
Evaluating probability forecasts for GDP declines using alternative methodologies.
International Journal of Forecasting, 29(1), 175-190.

Murphy, A. H. (1993).
What is a good forecast? An essay on the nature of goodness in weather forecasting.
Weather and Forecasting, 8(2), 281-293.

Seillier-Moiseiwitsch, F., & Dawid, A. P. (1993).
On testing the validity of sequential probability forecasts.
Journal of the American Statistical Association, 88(421), 355-359.

References III

Steffen, B., & Weber, C. (2016).
Optimal operation of pumped-hydro storage plants with continuous time-varying power prices.
European Journal of Operational Research, 252(1), 308-321.

Stephenson, D. B., Coelho, C. A., & Jolliffe, I. T. (2008).
Two extra components in the Brier score decomposition.
Weather and Forecasting, 23(4), 752-757.

Yates, J. F. (1982).
External correspondence: Decompositions of the mean probability score.
Organizational Behavior and Human Performance, 30,132-156.

Pumped-Hydro Storage Plant Optimisation [2]

$$\min_{T_t, S_t, X_t} \quad \sum_{t=1}^{H} P_t (T_t \Delta t - S_t \Delta t)$$

$$\text{s.t.} \quad X_t - X_{t-1} = -T_t \Delta t + \eta S_t \Delta t$$

$$0 \leq T_t \leq K_s$$

$$0 \leq S_t \leq K_s$$

$$0 \leq X_t \leq K_x$$

$$X_0 = X^0$$

$$X_H \geq X_0$$

[2] Following Steffen & Weber (2016)