ENTSO-E Data: Gaps and Inaccuracies

The ENTSO-E Transparency Platform (ETP) has systematically collected and uploaded operational data from European electricity markets since 5th January 2015. Thus, 2016 was the first complete year, and I have recommended users of my own collection of time series on <u>http://pfbach.dk/</u> to use time series from ETP.

It is a main advantage of the ETP data that they are available quite soon after the hour concerned. On the other hand, most ETP time series have the disadvantages that they are incomplete and that they are less accurate, particularly for wind and solar power.

The transmission system operators collect metered data for accounting and statistics. These data are more accurate than the data collected and published immediately, but are not available until weeks or months later. Comparing the ETP data with the later published time series offers an opportunity to evaluate the accuracy of the ETP data.

Missing data

The 2016 data at <u>http://pfbach.dk/</u> are mainly based on ETP data. Each sheet shows the data source in the top. Missing data have been replaced by estimates.

2016	DE	ES	FR	GB	IE	NL	NO	SE
Source	ENTSO-E	ENTSO-E	ENTSO-E	Gridwatch	ENTSO-E	ENTSO-E	ENTSO-E	ENTSO-E
Load	0.7%	0.1%	0.4%	0.4%	9.6%	0.2%	0.2%	0.7%
Wind	2.0%	0.1%	0.4%	0.4%	30.0%	1.1%	0.0%	0.0%
Offshore						1.1%		
Onshore						1.9%		
PV	1.2%	0.1%	0.4%					
Hydro							0.0%	0.3%
Nuclear		0.1%	0.4%	0.4%				0.3%
Spot price		0.0%		0.5%		0.7%		

Table 1 - Share of missing data

The French data have later been replaced by data from RTE. The Swedish data has been replaced by data from Svenska Kraftnät. The gaps in the Irish data are extensive and an estimation of complete time series was dropped.

Data Accuracy

A comparison between ETP data and later data from the system operators was made for four French time series and for four Swedish time series.

Fig. 1 shows a comparison as an example. Two sources, RTE (French system operator) and ETP have published time series for French wind power in 2015. See all eight charts in annex 1.





Paul-Frederik Bach

The difference between these two time series has been calculated and arranged as a duration curve.

The difference in energy content of the two time series is 7.1%, being reflected as different areas above and below the X-axis. The standard deviation of the difference is 11.2%, but the correlation between the two time series is as high as 99.7%.

In spite of a high standard deviation, the quality of these ETP data will be acceptable for most purposes, but not for estimating the energy content of the wind power.

SE 2016 Load. So	urces: Sven	ska Kraftna	ät & ENTSO-	Ε.		
Svenska Kraftnät	133967	GWh	Difference	2,7%	Correlation	99,1%
ENTSO-E	137638	GWh	St. dev.	3,2%		
SE 2016 Wind. So	urces: Sver	nska Kraftn	ät & ENTSO-	Ε.		
Svenska Kraftnät	15483	GWh	Difference	0,0%	Correlation	99,7%
ENTSO-E	15478	GWh	St. dev.	5,3%		
SE 2016 Nuclear.	Sources: Sv	venska Kraf	ftnät & ENTS	ю-е.		
Svenska Kraftnät	60458	GWh	Difference	0,1%	Correlation	99,8%
ENTSO-E	60384	GWh	St. dev.	1,2%		
SE 2016 Hydro. So	ources: Sve	nska Kraftr	nät & ENTSO	-E.		
Svenska Kraftnät	62021	GWh	Difference	0,6%	Correlation	97,3%
ENTSO-E	61669	GWh	St. dev.	9,1%		
FR 2015 Load. So	urces: RTE	& ENTSO-E				
RTE	473153	GWh	Difference	0,6%	Correlation	99,7%
ENTSO-E	470083	GWh	St. dev.	1,8%		
FR 2015 Wind. So	ources: RTE	& ENTSO-I	E.			
RTE	21094	GWh	Difference	7,1%	Correlation	99,7%
ENTSO-E	19601	GWh	St. dev.	11,2%		
FR 2015 PV. Sour	ces: RTE &	ENTSO-E.				
RTE	7286	GWh	Difference	1,5%	Correlation	99,1%
ENTSO-E	7175	GWh	St. dev.	20,9%		
FR 2015 Nuclear.	Sources: R	TE & ENTS	о-е.			
RTE	416454	GWh	Difference	0,2%	Correlation	99,9%
ENTSO-E	415432	GWh	St. dev.	0,5%		

Table 2 - Main results from the eight comparisons. "St. dev." (standarddeviation) refers to the difference between the two time series, but as% of the average of one of the two time series.

Table 2 shows different properties of the eight time series. The sample may be too small for robust conclusions, but the standard deviations of the differences are significantly higher for wind and solar power than for the other time series. The correlation coefficients are high for all cases. This means that the ETP time series reflect shapes of the variations in each of the time series.

Conclusion

For practical use, the gaps are more disturbing than the inaccuracies. Manual filling of gaps can be time-consuming. For instance, the creation of a complete time series for German wind power in 2016 takes the estimation of 363 quarters for offshore wind and 280 quarters for onshore wind. Inaccuracies are unavoidable in fast data, but it must be possible for the data providers to fill the gaps, if necessary with some delay.

Paul-Frederik Bach

Annex 1

Charts for the eight comparisons

Sweden 2016









France 2015







The PV difference is practically zero for the dark hours.