What drives the accuracy of PV output forecasts?

Topic: (3) Integrierte Netze der Zukunft - Integration Erneuerbarer in das Energiesystem
Thi Ngoc NGUYEN[1(1)], Felix MÜSGENS[2(1)]
[(1)] Brandenburgische Technische Universität Cottbus - Senftenberg

Motivation and central question

Accurate photovoltaic (PV) forecasts are increasingly important to the integration of PV into grid, attracting a consistently high interest from grid operators, investors, politicians, and forecasters from both industry and academia. This leads to such a vast number of literatures focusing on enhancing PV forecast accuracy that it requires a scientific knowledge systemization.

While there are already some survey papers (we found thirteen) summarizing findings from the literature, our work is the first to do statistical analysis of the individual errors reported in PV output forecast papers to systematically answer the question "What drives the accuracy of PV output forecasts?".

Methodologies

Our paper analyses PV output forecasts' performance from the literature focusing on the models' errors, controlling for the effects of various factors such as the evaluation metric, the forecast horizon, the length of the testing data set, etc.

To do that, we build a data base of models' forecast errors from the existing literature. First, we collect papers on PV output forecasts using historical review papers and Google Scholar and get 180 papers in total. Next, we carry out preliminary examination and remove the papers forecasting solar irradiance (we focus exclusively on PV output (electricity) forecasts), and the papers of insufficient information, incorrect approach of PV forecasting, or incomparable error metric or forecast horizon. This leaves 66 papers for final data collection and analysis. After processing the data, we have 1,136 observations with 18 key features, covering a variety of models, regions, training and testing data sets etc. Our data base is large enough to control for various factors and produce robust, statistically significant results.

We use an OLS regression to quantify the contribution of different factors and methodologies to the forecast quality. In addition, we use boxplots to further examine the data and to visualize the results.

Results and Conclusions

1. Inter-methodology comparison:
   1.1. Hybrid models are robustly superior to the others and outperform the conventional methods by 0.84-2.68 percentage point (pp).
   1.2. Machine Learning performs much worse than expected. For example, its average NRMSE (normalized by averaged value) for day-ahead forecasts is 35% – compared with hybrid methods (15-17%) and conventional methods (19-20%).
   1.3. The performance gap between state-of-the-art and conventional methods is much less impressive than expected. The complexity-accuracy trade-off therefore favours the conventional models in the short and medium run. However, the complex models show much higher potential to enhance forecasts' quality in the long run thanks to the development of new optimization techniques.
2. Besides the model, other factors driving PV output forecasts' accuracy include the forecast horizon, the test set length, the time publishing the models, and the data processing techniques.
   2.1. Longer forecast horizons have higher errors. For example, intra-day forecasts have errors being 1.34-5.80 pp higher than intra-hour forecasts.
   2.2. Models report lower errors for shorter test sets: an additional day in the test set increases the error by 0.008-0.018 pp.
   2.3. Data processing techniques significantly lower forecast errors by 0.85-4.45 pp.
   2.4. Models published one year later have the errors 0.14-3.07 pp lower.

Literature

Due to the word limit, kindly find the list of the literature *here*.

Others

Further tables and figures are provided *here*.

[1] Ad: Siemens-Halske-Ring 13, 03046 Cottbus; T: +49 (0) 355 69 3227; E: nguyen@b-tu.de
[2] Ad: Siemens-Halske-Ring 13, 03046 Cottbus; T: +49 (0) 355 69 4504; E: felix.muesgens@b-tu.de

	Dependent variable: average error value					
	All error metrics (1)	All error metrics using test sets at least 1 year (2)	NRMSE normalized by peak power or installed capacity (3)	NMAE normalized by peak power or installed capacity (4)	MAPE normalized by averaged value (5)	NRMSE normalized by averaged value (6)
Classical$^{(1)}$	-1.563	-1.683	3.853	-0.320		-0.445
	(2.043)	(2.167)	(4.218)	(0.807)		(3.909)
Ensemble$^{(1)}$	2.116	0.365	-0.473	-2.575		0.267
	(2.631)	(2.114)	(5.198)	(1.682)		(3.716)
Hybrid$^{(1)}$	-3.321**	-4.318***	-2.635	-2.469***	5.232	-5.883*
	(1.646)	(1.653)	(2.970)	(0.679)	(4.731)	(3.492)
Hybrid-Ensemble$^{(1)}$	-0.757	0.406	5.333		-3.548	-4.622
	(3.059)	(2.840)	(8.697)		(7.804)	(3.820)
ML$^{(1)}$	-0.392	0.132	4.298	-2.120***	7.666*	-1.957
	(1.585)	(1.722)	(2.866)	(0.670)	(4.345)	(3.436)
Persistence$^{(1)}$	0.903	-1.883	6.569**	0.294	10.476**	1.005
	(1.849)	(2.040)	(3.310)	(0.744)	(5.257)	(4.175)
Physical$^{(1)}$	6.620**	-0.835	5.908	-0.548		0.848
	(3.297)	(2.833)	(8.607)	(1.109)		(4.665)
Test set length (days)	0.008***	0.012***	0.018***	0.007***	-0.005	0.012***
	(0.001)	(0.002)	(0.003)	(0.001)	(0.005)	(0.002)
intra-day$^{(2)}$	1.344*	2.360***	-1.266	2.733***	-0.148	5.799***
	(0.780)	(0.844)	(1.898)	(0.227)	(3.222)	(1.626)
day-ahead$^{(2)}$	0.131	6.648***	-5.239***	0.831***	-5.530*	14.373***
	(0.677)	(0.927)	(1.625)	(0.275)	(2.978)	(2.126)
Resolution (minutes)	0.008	-0.100***	0.080**	-0.055***	0.292***	-0.266***
	(0.014)	(0.016)	(0.033)	(0.007)	(0.060)	(0.056)
Publishing Year	-0.829***	-1.554***	-0.373	-0.134**	-0.201	-3.068***
(of the papers)	(0.133)	(0.198)	(0.292)	(0.059)	(0.467)	(0.318)
Complexity	-0.398	-0.844**	2.107***	-0.847***	-4.447***	-3.433***
(number of techniques)	(0.257)	(0.379)	(0.524)	(0.108)	(1.357)	(0.960)
Constant	1,680.585***	3,144.080***	752.050	278.803**	409.299	6,205.555***
	(267.513)	(400.352)	(590.081)	(118.351)	(942.393)	(640.404)
Observations	1,121	385	341	328	130	116
R^2	0.158	0.402	0.183	0.866	0.266	0.755
Adjusted R^2	0.149	0.381	0.151	0.861	0.204	0.724
Residual Std. Error	9.065 (df = 1107)	5.715 (df = 371)	11.402 (df = 327)	1.541 (df = 315)	9.566 (df = 119)	6.057 (df = 102)
F Statistic	16.029*** (df = 13; 1107)	19.199*** (df = 13; 371)	5.638*** (df = 13; 327)	170.015*** (df = 12; 315)	4.305*** (df = 10; 119)	24.192*** (df = 13; 102)

Note: [1] Dummies of methodology, baseline: advanced classical models
[2] Dummies of forecast horizon, baseline: intra-hour horizon

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 1: Effects of different factors and methodologies on forecast errors

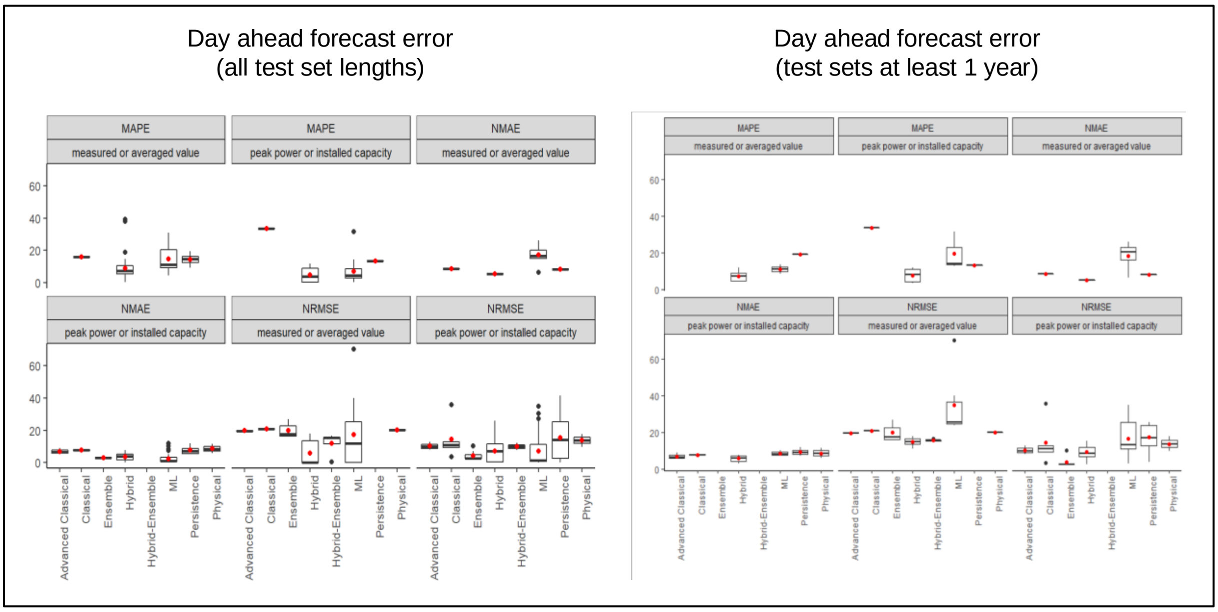

Figure 1: Methodologies' performance in day-ahead PV output forecasting