A Crowdsourcing Approach to Energy Forecasting

Tao Hong, PhD

Director of Big Data Energy Analytics Laboratory (BigDEAL)

Energy Production and Infrastructure Center (EPIC)

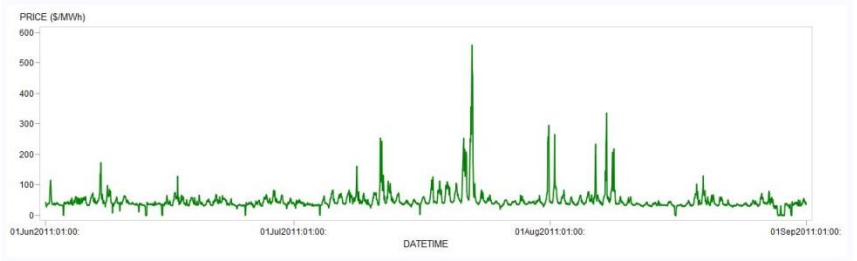
University of North Carolina at Charlotte

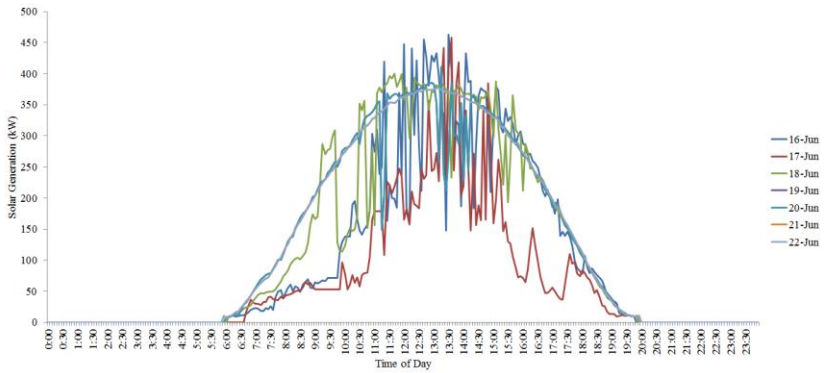
UNC CHARLOTTE

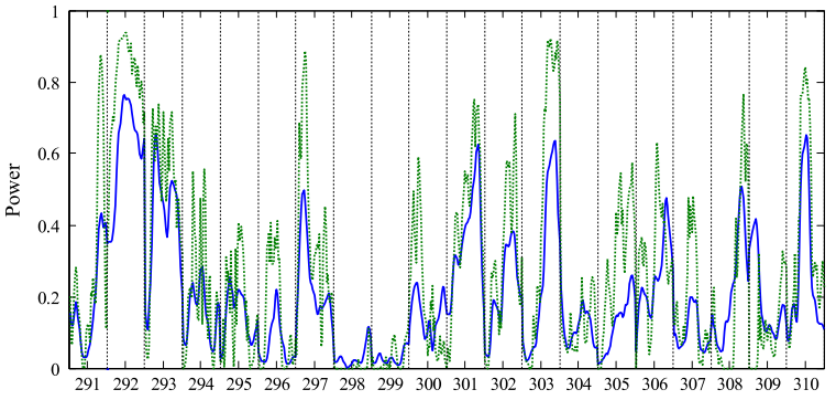
Agenda

- Why forecasting?
- Why crowdsourcing?
- Global Energy Forecasting Competitions (GEFCom)

Why Forecasting?





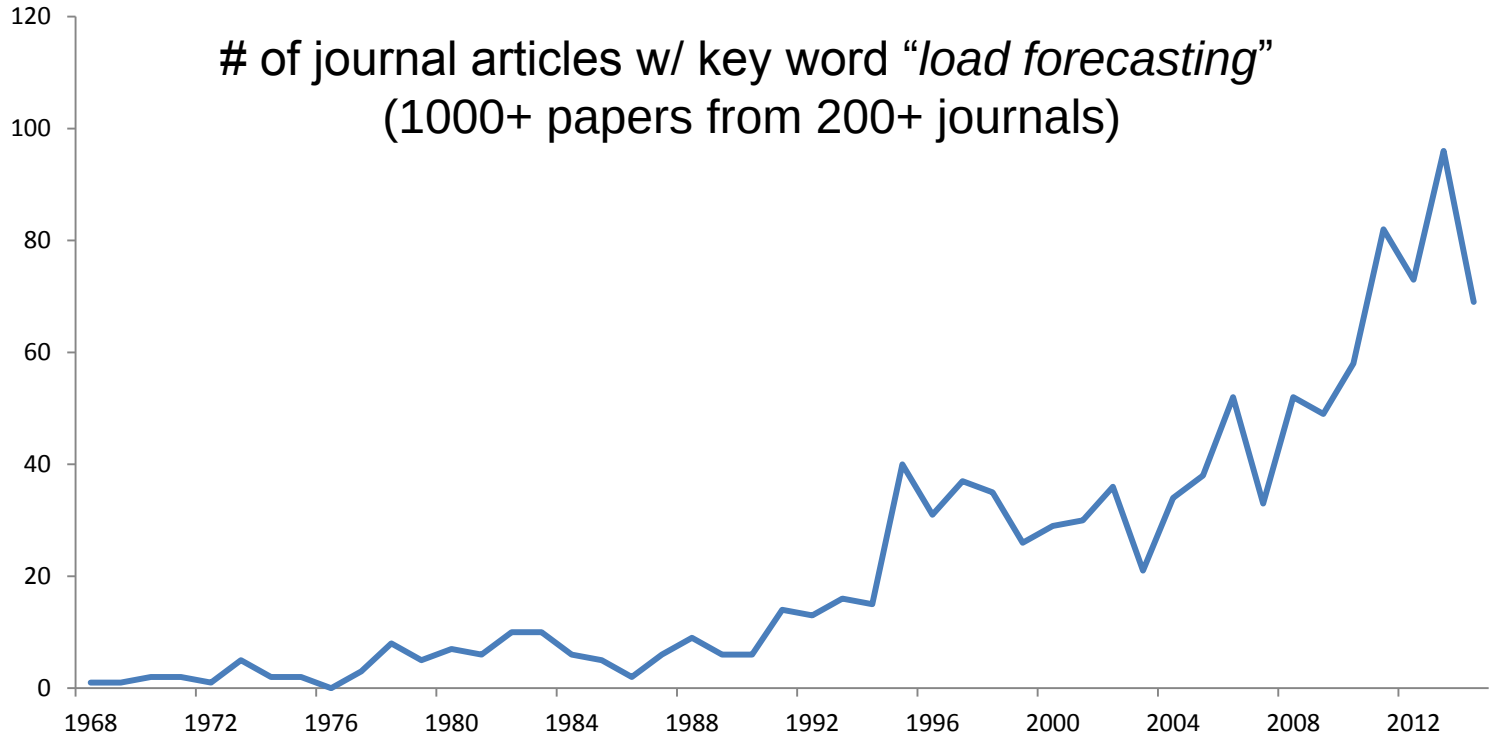


galvin center for electricity innovation at ILLINOIS INSTITUTE OF TECHNOLOGY

IEEE PES Power & Energy Society®

Why Crowdsourcing?

- What's happening in ivory tower



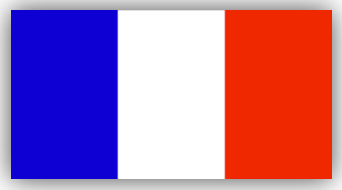
GEFCom2012

GEFCom2012
Load Forecasting

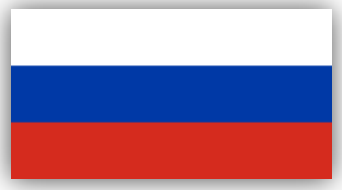
GEFCom2012
Wind Forecasting

GEFCom2012

- 8 winning teams from 8 countries





Hierarchical Load Forecasting Track:

#1. Colin Singleton and Nathaniel Charlton from Counting Lab (UK);
#2. James Robert Lloyd from University of Cambridge (UK);
#3. Raphael Nedellec from EDF R&D (France), Jairo Cugliari from INRIA (France) and Yannig Goude from EDF R&D (France);
#4. Souhaib Ben Taieb from Université Libre de Bruxelles (Belgium) and Rob J Hyndman from Monash University (Australia).

Wind Power Forecasting Track:

#1. Lucas Eustáquio Gomes da Silva from DTI Sistemas (Brazil);
#2. Ekaterina Mangalova from Siberian State Aerospace University and Evgeny Agafonov from Siberian Federal University (Russia);
#3. Gabor I. Nagy from Budapset University of Technology and Economics (Hungary);
#4. Duehee Lee from University of Texas at Austin (USA).

GEFCom2012

- Wind power forecasting

Kaggle ID	WF1	WF2	WF3	WF4	WF5	WF6	WF7	Validation	Test	All	Submissions
Leustagos	0.145	0.138	0.168	0.144	0.158	0.133	0.140	0.146	0.146	0.146	37
DuckTile	0.143	0.145	0.172	0.145	0.165	0.137	0.146	0.149	0.147	0.148	82
MZ	0.141	0.151	0.174	0.145	0.167	0.141	0.145	0.148	0.149	0.149	19
Propeller	0.144	0.153	0.177	0.147	0.175	0.141	0.147	0.148	0.153	0.152	64
Duehee Lee	0.157	0.144	0.176	0.160	0.169	0.154	0.148	0.155	0.155	0.155	10
MTU EE5260 forecast team	0.161	0.172	0.193	0.162	0.192	0.156	0.160	0.166	0.169	0.168	20
SunWind	0.174	0.177	0.193	0.176	0.179	0.157	0.162	0.173	0.171	0.172	26
ymzsmsd	0.163	0.186	0.200	0.164	0.192	0.162	0.167	0.173	0.174	0.174	24
4138 Kalchas	0.180	0.179	0.197	0.175	0.200	0.160	0.165	0.179	0.176	0.177	3
Benchmark	0.302	0.338	0.373	0.364	0.388	0.341	0.361	0.361	0.353	0.355	1

Over 60% error reduction!

GEFCom2014

235

153

139

147

Further Readings

Tao Hong, Pierre Pinson and Shu Fan (2014), "Global Energy Forecasting Competition 2012", International Journal of Forecasting, vol.30, no.2., pp 357-363

International Journal of Forecasting 30 (2014) 357–363

Contents lists available at ScienceDirect

International Journal of Forecasting

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

Global Energy Forecasting Competition 2012

Tao Hong[a,*], Pierre Pinson[b], Shu Fan[c]

[a] SAS Institute Inc, United States
[b] Technical University of Denmark, Denmark
[c] Monash University, Australia

ABSTRACT

The Global Energy Forecasting Competition (GEFCom2012) attracts numerous participants worldwide, who contributes many novel ideas to the energy forecasting field. This paper introduces both tracks of GEFCom2012, namely load forecasting and wind power forecasting, with details on the aspects of the problems, the data, and a summary of the numerous ways by which top entries fit. We also discuss the lessons learned from this competition from the organizers' perspective. The complete data set, including the solution data, is published along with this paper, in an effort to establish a benchmark data pool for the field.

© 2013 International Institute of Forecasters. Published by Elsevier B.V. All rights reserved.

1. Background

In a broad sense, energy forecasting covers a wide range of forecasting problems in the utility industry, such as generation forecasting, load forecasting, price forecasting, demand response forecasting, and so on. While the deployments of smart grid technologies drives the utility industry data of a higher granularity than ever before, it also presents the challenge of obtaining business value from big datasets. As a result, energy forecasting, one of the most fundamental and classical problems, has found a new life in today's utility industry.

Although a significant amount of the literature has been devoted to energy forecasting, most such studies are still at the theoretical level, having little practical value. No formal benchmarking process or data pool has been established in the field, and new publications rarely reproduce the results from past work done by other research groups for a comparison. Few academic programs in electrical engineering, statistics or economics offer courses which concentrate on energy forecasting. Given these facts, the IEEE Working Group on Energy Forecasting (WGEF) organized the Global Energy Forecasting Competition 2012 (GEFCom2012) in order to: (i) improve the forecasting practices of the utility industry, (ii) bring together state-of-the-art techniques for energy forecasting, (iii) bridge the gap between academic research and industry practice, (iv) promote analytics in power and energy education, and (v) prepare the industry to overcome the forecasting challenges posed by the smart grid world. The competition included two tracks, hierarchical load forecasting and wind power forecasting. In this paper, we introduce GEFCom2012 in detail, as well as publishing the complete competition dataset in an attempt to establish a benchmarking data pool for energy forecasting.

We started planning the competition in late 2011; this mainly involved identifying field interns, seeking sponsorships, and setting up the rules and schedule. Most previous forecasting competitions have used a centralized communication approach, where the participants were able communicate with the administrators but not with each other. As a result, the participants did not know the scores and ranks until the administrators calculated them after the competition. The tourism competition (Athanasopoulos, Hyndman, Song, & Wu, 2011) took a different approach, by using Kaggle's platform, where both the participants and administrators can share questions, ideas and findings with each other on Kaggle's forum. As soon as a team submits its entry, the score is calculated and displayed to the team automatically. If the score is the best one presented by this team, the public leaderboard is refreshed to reflect the changes. Based on these key features, GEFCom2012 selected Kaggle as the competition platform, becoming the second forecasting competition to be hosted

* Corresponding author.
E-mail addresses: hongtao01@gmail.com (T. Hong), ppin@elektro.dtu.dk (P. Pinson), shu.fan@monash.edu (S. Fan).

0169-2070/$ - see front matter © 2013 International Institute of Forecasters. Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ijforecast.2013.07.001

"GEFCom2012 is the largest known energy forecasting competition to date. Not only does it bring together many new ideas to the energy-forecasting field from data scientists in many different industries but the competition data has already been used by scholars for benchmarking purposes."

-- IEEE Power and Energy Society, 9/30/2013

Questions?

For more information, please contact

Dr. Tao Hong
hong@uncc.edu
www.drhongtao.com
blog.drhongtao.com