Global Energy Forecasting Competition 2012

Tao Hong, Pierre Pinson and Shu Fan

**Abstract**

Global Energy Forecasting Competition (GEFCom2012) attracted hundreds of participants worldwide, who contributed a lot of novel ideas to the energy forecasting field. This paper introduces both tracks of GEFCom2012, hierarchical load forecasting and wind power forecasting, from the aspects of the problem, data and summary of the methods used by selected top entries. We also discuss lessons learned from this competition from the organizers’ perspective. Along with this paper, we publish the complete data set including the solution data as an effort of establishing a benchmark data pool for the community.

**1. Background**

Energy forecasting, in a broad sense, covers a wide range of forecasting problems in the utility industry, such as generation forecasting, load forecasting, price forecasting, demand response forecasting, etc. While the deployment of smart grid technologies offers the utility industry higher granular data than ever before, it also presents the challenge to drive business value out of the big data. As a result, energy forecasting, one of the most fundamental and classical problems, finds its new life in today's utility industry.

Although there has been a significant amount of literature devoted to energy forecasting, most are still on the theoretical level with little practical value. No formal benchmarking process and data pool have been established in the field. New publications rarely reproduce and compare with the results from the past work done by other research groups. Few academic programs in electrical engineering, statistics or economics offer courses concentrated on energy forecasting. To improve the forecasting practices of the utility industry, to bring together the state-of-the-art techniques for energy forecasting, to bridge the gap between academic research and industry practice, to promote analytics in power and energy education, to prepare the industry to overcome the forecasting challenges in the smart grid world, IEEE Working Group on Energy Forecasting (WGEF) organized the Global Energy Forecasting Competition 2012 (GEFCom2012). The competition included two tracks, hierarchical load forecasting and wind power forecasting. In this paper, we introduce GEFCom2012 in detail. Meanwhile, we publish the complete competition data as an effort of establishing the benchmarking data pool for energy forecasting.

We started planning the competition in late 2011, which mainly included identifying field interest, seeking sponsorships, and setting up rules and schedule. Most of the previous forecasting competitions used a centralized communication approach, where the participants could communicate with the administrators but not 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 et al,. 2011) took a different approach by using Kaggle's platform, where 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 automatically calculated and displayed to this team. If the score is the best one from this team, the public leader board is refreshed to reflect the changes. Due to these key features, GEFCom2012 selected Kaggle as the competition platform and became the second forecasting competition hosted by Kaggle. The first Call for Participants was issued in May 2012. Prior to the launching date, we received registrations from around 120 people across 30 countries. The competition was active on Kaggle for two months, from 8/31/2012 to 10/31/2012. At the end of the competition, the data on each track was downloaded by over 600 unique users.

This paper is organized as follows: Section 1 presents the background, including the motivation, platform, and an overview of the participants. Sections 2 and 3 introduce the two tracks respectively in terms of the problem, data, and a brief summary of the methods and results. Section 4 discusses the issues and lessons learned from this competition. The paper is concluded in Section 5 with an outlook of the potential future work. We also acknowledge the key contributors at the end.

**2. Hierarchical Load Forecasting**

Problem Description

Short term load forecasting (STLF) provides load forecasts in hourly or sub hourly interval for the next one day to two weeks. The forecasts are used by all sectors of the utility industry, from generation and transmission to distribution and retail. The business needs of short term load forecasts include unit commitment, T&D (transmission and distribution) operations and maintenance, and energy market activities. A lot of statistical and artificial intelligence techniques have been applied to STLF over the past three decades, such as multiple linear regression, Box-Jenkins approach, and Artificial Neutral Networks, etc. A comprehensive review of the literature is by Hong (2010).

In the hierarchical load forecasting track, the participants are required to backcast and forecast hourly loads (in kW) for a US utility with 20 zones at both zonal level (20 series) and system level (sum of the 20 zonal level series), totally 21 series. We provide the participants with 4.5 years of hourly load and temperature history with 8 non-consecutive weeks of load data taken out. The backcasting task is to predict the loads of these 8 weeks in the history given actual temperatures, in which the participants are permitted to use the entire history to backcast the loads. The forecasting task is to predict the loads of the week right after the 4.5 years of history without actual temperatures or temperature forecasts given. This track is designed to mimic a short term load forecasting job, in which the forecaster first build a model with historical data, and then develop the forecasts for the next few days. Traditionally, most STLF jobs are conducted with the system level data only. In this competition, we provide the zonal level data to further mimic a STLF job in the smart grid era, where the forecasters have access to the smart meter information.

Among thousands of papers in the load forecasting literature, most were devoted to various modeling techniques, while a lot of practical issues have not received enough attentions. When designing the competition problem, we would like to highlight a few challenges with the aim to bring new ideas from the following aspects:

1. Data cleansing. The competition data is real-world data, in which there are significant data quality issues due to the outages, load transfers and various other data errors. An effective data cleansing method is expected to enhance the forecasting accuracy. This challenge is also applicable to the wind forecasting track.
2. Hierarchical forecasting. To fully utilize the hierarchical information, the participants may choose a bottom-up, middle-out or top-down approach. In addition, to avoid the situation that some participants use external data, we did not specify the locations of the zones and weather stations. Therefore, another challenge is to decide which weather station(s) should be associated with which delivery point. In practice, although the forecasters have access to the geographical information, they still need to decide which weather station(s) should be used for each zone and how to use them.
3. Special days forecasting. The loads of holidays and the surrounding days are usually less predictable than the regular days due to the limited sample size and variability of the pattern over time. When selecting the weeks to be backcasted and forecasted, we include holidays in some of the weeks.
4. Temperature forecasting. In operations environment, some utilities purchase commercial weather forecast, while some have their meteorologists to develop in-house weather forecast. In this competition, we did not release the temperature forecasts for the forecasted week. If the participants decide to use temperature variables, they have to develop the temperature forecast for the forecasted week.
5. Combining forecasting. The participants are not restricted to any techniques or tools used for this competition. We would like to see applications of combining forecasting methods in both tracks of GEFCom2012.
6. Integration. A load forecasting job covers a few tasks including the ones above. Integration of these tasks is another important task. For instance, a decent temperature forecast may or may not result in a decent load forecast. A good integration strategy may further enhance the load forecasting accuracy. However, among the reports we received from both tracks, none of them discuss the integration strategy.

Other than the standard Kaggle rules, we set up the following two rules:

1. The participants are not allowed to use more weather, load and economy data than what has been provided.
2. At each hour, the sum of the zonal level loads should be equal to the system level load.

The error score in the hierarchical load forecasting track is Weighted Root Mean Square Error (WRMSE) as described below:

,

where *Ai* and *Pi* are the actual and predicted values of observation *i*, while the weight for this observation is denoted as *Pi* and specified in Table 1.

Table 1. Weight assignment

|  |  |
| --- | --- |
| Week(s) | Weight |
| Forecasted week at system level | 160 |
| Forecasted week at zonal level | 8 |
| Backcasted week at system level | 20 |
| Backcasted week at zonal level | 1 |

Data Description

The complete dataset can be roughly divided by two parts based on different purposes of usage: a training set for model identification and parameter estimation, and an evaluation set for calculating scores. Kaggle randomly selects 25% of the evaluation data as the validation set for calculating public scores, and the rest 75% as the test set for calculating private scores. The public scores can be seen by all participants and competition administrators during the competition, while the private scores are being published at the end of the competition. During the competition, the validation and test data is not released to the participants. Along with this paper, we publish the complete dataset including five spreadsheets in Comma-Separated Values (CSV) format for the hierarchical load forecasting track:

1. Load\_history. Houly load history of 20 zones ranges from the 1st hour of 2004/1/1 to the 6th hour of 2008/6/30 with the following 8 weeks set to be missing for backcasting purpose: 2005/3/6 - 2005/3/12, 2005/6/20 - 2005/6/26, 2005/9/10 - 2005/9/16, 2005/12/25 - 2005/12/31, 2006/2/13 - 2006/2/19, 2006/5/25 - 2006/5/31, 2006/8/2 - 2006/8/8, and 2006/11/22 - 2006/11/28.
2. Temperature\_history. Hourly temperature history of 11 weather stations ranges from the 1st hour of 2004/1/1 to the 6th hour of 2008/6/30.
3. Holiday\_list. A list of US Federal holidays from 2004/1/1 to 2008/7/7.
4. Load\_benchmark. Predicted hourly loads from 2008/7/1 to 2008/7/7. The weight column shows the weights are assigned to different weeks and levels,
5. Load\_solution. Actual hourly loads from 2008/7/1 to 2008/7/7. The format is similar to “Load\_benchmark”. The indicator column shows how we split the solution data to calculate the scores for public and private leaderboards.

Summary of methods and results

The benchmark is created based on a multiple linear regression model with an intercept and the following effects as discussed by Hong (2010):

1. main effects: *Trend* (an increasing normal number assigned to each observation in the chronological order), *T* (temperature of the current hour), *T2*, *T3*, *Month* (a class variable with 12 levels representing 12 months of a year), *Weekday* (a class variable with 7 levels representing 7 days of a week) and *Hour* (a class variable with 24 levels representing 24 hours of a day).
2. cross effects (interactions): *Hour\*Weekday*, *T\*Month*, *T2\*Month*, *T3\*Month*, *T\*Hour*, *T2\*Hour* and *T3\*Hour*.

The parameters are estimated with the 4.5 years of history less the 8 backcasted weeks. For each zone, we build 11 models, one per weather station. The weather station with the best fit is being assigned to the corresponding zone. We predict the 8 weeks of loads using the same model with actual temperatures from the selected weather station. We forecast the last week of loads using the same model with forecasted temperatures, of which temperature forecast at each hour is the average of the same date and hour over the past four years.

Table 1 summarizes the methods used by selected entries based on their reports. We also calculate the WRMSE of the 8 backcasted weeks, 7/1/2008 and the forecasted week as shown in Table 2.

Table 2. Summary of methods in hierarchical load forecasting track

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Kaggle ID | Techniques | Data Cleansing | Weather Station Selection | Holiday Effect | Temperature Forecast | Combining Forecasting |
| CountingLab | MLR, SVD, combining | Yes | Eleven models corresponding to the eleven weather stations were built | Yes | Using the average temperature of the same hour from similar days in the previous years. | Combine forecasts from the 5-best fitted models. |
| James Lloyd | Gradient boosting machines, Gaussian process regression, MLR, combining | Not discussed | Temperatures from all stations were used | No | Estimating the smooth trend and daily periodicity of temperature separately. | Combine forecasts from three models. |
| Tololo (EDF) | Semi-parametric regression, with B-splines or cubic regression splines as smooth function | Not discussed | A stepwise procedure was used for each zone to select the station that minimizes forecasting error on a test set. | Yes | Not discussed | No |
| TinTin | Nonparametric additive models with P-spline, component-wise gradient boosting | Yes | A testing week (the last week of the available  data) was used to determine the station for each zone. | Yes | Using the average temperatures at the same period across the previous years. | No |
| Quadrivio | Multiple linear regression (MLR) | Yes | Load was fitted to temperature at each station separately, and the best three were used for each zone. | No | Averaging the temperatures during the same days from previous years | No |
| Chaotic Experiments | Random Forest, GMB models, combining | Not discussed | Not discussed | Yes | Not discussed | Combine forecasts from three models. |
| Andrew L | Generalized additive model, spline, PCA | Not discussed | The first component of PCA was used temperature variable for each hour | No | Using a generalized additive model | No |
| NHH | Wavelet decomposition, Mutual Information, Neural Networks | Not discussed | Temperatures from all stations were considered as input candidate | No | Not discussed | No |
| TheJellyTeam | Neural Networks | Not discussed | Temperatures from all stations were considered. | Yes | Using the mean of the same period from the previous years. | No |
| Shooters Touch | Regression models and neural network | No | Weighted average of up to 3 stations, selected based on the fitted result for each station | Yes | Not discussed | No |
| Tao’s Vanilla Benchmark | MLR | No | Best fit from the 11 weather stations | No | Average of the same date/time of the past four years | No |

Table 3. Error statistics (WRMSE) of selected entries in the wind power forecasting track.

|  |  |  |  |
| --- | --- | --- | --- |
| Kaggle ID | Backcast Accuracy | 1-day ahead | 1-week ahead |
| CountingLab | 61890 | 72504 | 73900 |
| James Lloyd | 58406 | 59273 | 82346 |
| Tololo (EDF) | 46756 | 52136 | 82776 |
| TinTin | 50926 | 112410 | 86590 |
| Quadrivio | 71663 | 63186 | 81645 |
| Chaotic Experiments | 78238 | 50967 | 89783 |
| Andrew L | 68638 | 133005 | 106272 |
| NHH | 65360 | 121818 | 109850 |
| TheJellyTeam | 72197 | 120752 | 101066 |
| Tao’s Vanilla Benchmark | 69557 | 148352 | 123758 |

**3. Wind Power Forecasting**

Problem description

With the ever increasing deployment of wind power capacities as a viable renewable energy solution to the electricity mix, a number of decision-making problems related to power system operations and participation in electricity markets require some form of forecasts as input. The development of methods for wind power forecasting can be traced back to the work of Brown et al. (1984), who used simple time-series models for wind forecasting at a site of interest, then converted to electric power generation by passing wind forecasts through a theoretical manufacturer's power curve. Since then, three decades of research and development have led to the proposal of a wide range of approaches, with a clear intensification of these efforts since the beginning of the new millennium, as wind power capacities are spreading more globally (while they were mainly concentrated in the European region before that). A set of reviews of the state of the art in wind power forecasting exists, to which the readers are referred to for an exhaustive coverage of the alternative approaches. Certainly the most complete of these reviews are that by Giebel et al. (2011) and Monteiro et al. (2009).

In the wind power forecasting track, the participants are required to forecast the hourly wind power generation for 7 wind farms. We provide 3 years of historical data including both wind power generation and wind forecast. The error score for the wind power forecasting track is Root Mean Square Error (RMSE). Similar to the hierarchical load forecasting track, in addition to new techniques, we also anticipate some novel ideas on data cleansing, combining forecasting and integration.

Data description

In the wind power forecasting track, we use about 3 years of data of 7 wind farms from the same region of the world as a basis to design the competition problem. This data consists of historical power measurements for these wind farms, as well as meteorological forecasts of wind components at the level of these wind farms.

Historical power measurements have an hourly temporal resolution with a high level of availability over that period and for all the wind farms. They were normalized by the respective nominal capacities of the wind farms, so as to obtain normalized power values between 0 and 1, hence permitting to mask the original characteristics of the wind farms. It also enables a scale-free comparison of forecasting results for the various wind farms.

Meteorological forecasts were gathered for the zonal (u) and meridional (v) components of surface winds at 10 m above ground level. They were extracted from the archive of the European Centre for Medium-range Weather Forecasts (ECMWF). ECMWF issues high-resolution deterministic forecasts twice a day at 00UTC and 12UTC, with a temporal resolution of 3 hours out to 10 days ahead. In order to match the hourly resolution of the power measurements, also required by most forecast applications, the forecasts were interpolated using cubic splines so as to have hourly resolution. Only the first 48 hours of each forecast series were collated in the dataset. Note that these meteorological predictions were also given in the form of wind speed and direction for those who preferred to use them in such a format.

A number of 48-hour periods with missing power observations are defined for validation and testing purposes. The first one is from 1 January 2011 at 01:00 to 3 January 2011 at 00:00. The second one is from 4 January 2011 at 13:00 to 6 January 2011 at 12:00. Note that to be consistent, only the meteorological forecasts relevant for periods with missing power data, which would be available in practice, were given. Each of these two periods then repeats itself every 7 days until the end of the dataset. In between periods with missing data, power observations are available for updating the models if necessary.

Along with this paper, we publish the complete data including 11 spreadsheets (in comma-separated values CSV format) for the wind power forecasting track:

1. WindPower\_train. Houly wind power observations for the 7 wind farms from 2009/7/1 to 2010/12/31 (i.e., the training set), without any holes except potentially due to data quality issues.
2. WindPower\_eval. Houly wind power observations for the 7 wind farms from 2011/1/1 to 2012/6/28 (i.e., the evaluation set), with holes for the periods of which the forecasts are expected to be produced as mentioned above.
3. WindForecasts\_wf1, …, WindForecasts\_wf7. Wind forecasts for the 7 wind farms and for the same period as for the measurements. Forecasts are issued every 12 hours with a forecast horizon of 48 hours and an hourly temporal resolution.
4. WindPower\_benchmark. Predicted hourly wind power at the 7 wind farms for the holes in the evaluation set.
5. WindPower\_solution. Actual wind power measurements for the holes defined in the evaluation set. The format is similar to “WindPower\_benchmark”. The indicator column shows how we split the solution data to calculate the scores for public and private leaderboards.

Summary of methods and results

The persistence method, as one of the simplest approach to issue wind power forecasts for these wind farms, is used here as a benchmark. This forecasting approach is based on a random walk model, defining the forecasted value as the most recent available observation. The methods used by 9 selected teams together are summarized in Tables 4. We also show the error statistics of these 9 teams and the persistence benchmark in Table 5, where the error statistics (in RMSE) are broken down to each wind farm.

Table 4. Summary of methods in the wind power forecasting track.

|  |  |  |  |
| --- | --- | --- | --- |
| Kaggle id | Technique | Data Cleansing | Combining Forecasting |
| Leustagos | Linear combination of 9 models (regression from meteorological forecasts to power, inter-wind farms dependencies, autoregressive components, with different model structures) | No | Yes |
| DuckTile | Data cleaning, and then local linear regression with wind forecasts, day and time of the year as inputs | Yes | No |
| MZ | Linear models estimated with regularized least squares with radial basis functions spanning the space of wind forecasts, and autoregressive features | No | No |
| propeller | Linear regression from wind forecasts to power measurements, then nonlinear correction with gradient boosting machines (with optimal inputs identified through cross-validation) | Yes | No |
| Duehee Lee | Plain combination of a large number of Neural Networks (52) and Gaussian Process models (5) mapping all input data to power measurements | No | Yes |
| MTU EE5260 | Linear regression and Neural Networks for the conversion of meteorological forecasts to power | No | No |
| SunWind | Plain combination of a power curve model, an autoregressive model, a local linear regression model, and a support vector machine model | No | Yes |
| ymzsmsd | Sparse Bayesian learning with input measurements and forecasts from all wind farms | No | No |
| 4138 Kalchas | Regularized kernel-based regression for the conversion of meteorological forecasts to power | No | No |
| Benchmark | Persistence | No | No |

Table 5. Error statistics (RMSE) of selected entries in the wind power forecasting track.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Kaggle id | WF1 | WF2 | WF3 | WF4 | WF5 | WF6 | WF7 |
| Leustagos | 0.145 | 0.138 | 0.168 | 0.144 | 0.158 | 0.133 | 0.140 |
| DuckTile | 0.143 | 0.145 | 0.172 | 0.145 | 0.165 | 0.137 | 0.146 |
| MZ | 0.141 | 0.151 | 0.174 | 0.145 | 0.167 | 0.141 | 0.145 |
| propeller | 0.144 | 0.153 | 0.177 | 0.147 | 0.175 | 0.141 | 0.147 |
| Duehee Lee | 0.157 | 0.144 | 0.176 | 0.160 | 0.169 | 0.154 | 0.148 |
| MTU EE5260 | 0.161 | 0.172 | 0.193 | 0.162 | 0.192 | 0.156 | 0.160 |
| SunWind | 0.174 | 0.177 | 0.193 | 0.176 | 0.179 | 0.157 | 0.162 |
| ymzsmsd | 0.163 | 0.186 | 0.200 | 0.164 | 0.192 | 0.162 | 0.167 |
| 4138 Kalchas | 0.180 | 0.179 | 0.197 | 0.175 | 0.200 | 0.160 | 0.165 |
| Benchmark | 0.302 | 0.338 | 0.373 | 0.364 | 0.388 | 0.341 | 0.361 |

**4. Discussion**

Figure 1 shows the cumulative number of unique IDs that downloaded the data from each track since the beginning of the competition. The vertical dash dot line indicates the end of the competition, when there were about 600 unique IDs from each track. After the competition, the data are still being downloaded by the Kaggle users. Through Kaggle’s platform, the competition attracted much more participants than expected, of which many are very experienced data scientists outside the utility industry. While the diverse background of the participants brings a lot of new ideas to the energy forecasting field, some participants are not interested in joining the post-competition activities, such as submitting reports, presenting the work at conferences and writing scientific papers.

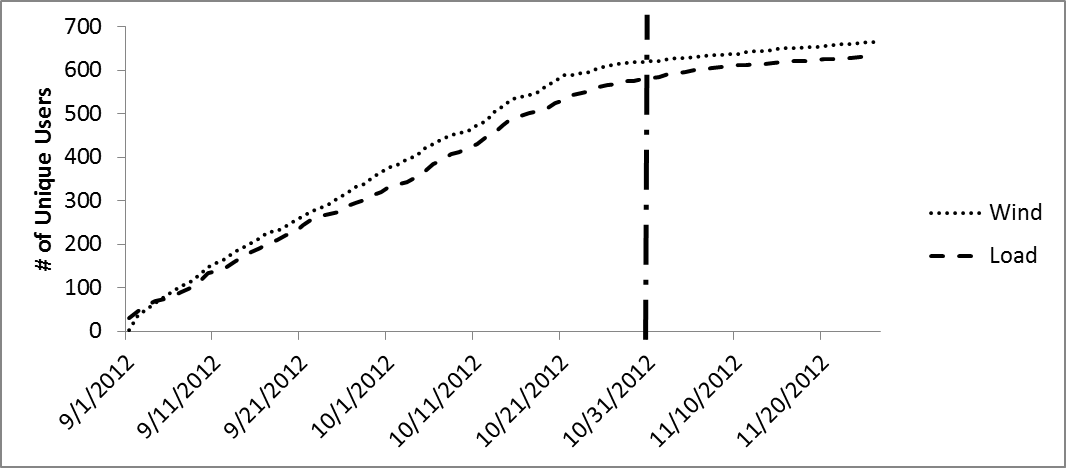


Figure 1. Number of unique users who have downloaded competition data from both tracks during the first 3 months.

Kaggle provides a forum for participants and competition administrators to post questions, answers and findings. This feature allows the participants to help each other in the public domain. It also allows the administrators to address issues as soon as they are raised. As the competition goes on, there is rich information in the forum, which requires the new participants to review the old posts. Some of the participants, who are new Kaggle users, may not review the previous posts, which leads to violation of some competition rules. To avoid similar situations in the future, we would recommend the competition administrators to increase the awareness of the participants about the important posts in the forum discussion.

In the hierarchical load forecasting track, to maintain the load level of each zone, we give the actual loads instead of standardized values, which open the possibility that some participants may be able to guess the location of the utility and use external information to win the competition. To avoid this situation, we require the teams to submit reports and codes, which will be evaluated by the award committee of GEFCom2012. Two teams were disqualified at the end due to this reason. In the wind power forecasting track, the data has been standardized, so that the participants can hardly find the solution by guessing where the wind farm is.

In real world short term load or wind forecasting jobs, the forecasters have to develop the forecasts on daily basis with newly available information. In other words, the forecast origin moves every day. To implement this feature in a competition, we have to host multiple phases with new data released in each phase. While each phase may take a couple of weeks to complete, the entire competition will take much longer than two months. Implementing this feature will also require the participants to be fully engaged throughout the competition. This is more achievable as an in-class competition than an inaugural international competition. Therefore, we did not set up this feature when designing GEFCom2012. As an amendment, we leave a few missing periods in the history for prediction. Since we can hardly find out whether the participants are using data after a missing period when predicting this missing period, we did not restrict the participants to use the data prior to each missing period being predicted. This set up may offer regression or some other data mining techniques some advantage over some time series forecasting techniques such as ARIMA, which may contribute to the fact that we did not receive any reports using Box-Jenkins approach in the hierarchical load forecasting track.

Forecasting by nature is a stochastic problem. In the utility industry, some applications in some utilities require probabilistic forecasts in the form of predictive densities or scenarios as inputs, such as annual peak demand forecasting for system planning (Hyndman and Fan 2009), systems reserve quantification (Matos and Bessa et al., 2011), unit commitment (Tuohy et al. 2009), as well as trading of wind power generation (Pinson et al., 2007). On the other hand, a lot of decision making processes are set to take point forecasts only. Majority of the energy forecasting literature has been on point forecasts. In GEFCom2012, to keep the competition problem and error score straightforward, we let the participants develop point forecasts instead of the probabilistic ones.

**5. Conclusion**

GEFCom2012 includes two tracks: hierarchical load forecasting and wind power forecasting. The competition attracted hundreds of participants worldwide. In this paper, we introduced GEFCom2012 from several aspects, including the background, problem, data, methods, results, and lessons learned. We also published the complete data of both tracks as an effort of establishing a data pool for energy forecasting. In the future, we would like to expand the competition by adding more tracks, such as long term load forecasting, price forecasting and solar generation forecasting. We would also like to explore other features, such as rolling forecast origin, comprehensive error scores, and probabilistic forecasts.

**Acknowledgement**

The authors gratefully acknowledge the following organizations that contributed to the success of GEFCom2012: IEEE Power& Energy Society, IEEE Power Systems Planning and System Implementation Committee, IEEE Power and Energy Education Committee, IEEE Working Group on Energy Forecasting, Kaggle, WeatherBank Inc., International Journal of Forecasting, and IEEE Transactions on Smart Grid. The authors also appreciate the support on organizing GEFCom2012 from the following individuals: David Hamilton, Eric Wang, Hamidreza Zareipour, ML Chan, Rob J Hyndman, Wei-Jen Lee, Fangxing Li, Shanshan Liu, Anil Pahwa, Mohammad Shahidehpour, and Kumar Venayagamoorthy.

**References**

Ath­anasopoulos G., Hyndman R., Song H., & Wu D. (2011), [The tourism forecasting competition](http://robjhyndman.com/papers/the-tourism-forecasting-competition/). [Inter­na­tional Journal of Fore­cast­ing](http://www.forecasters.org/ijf), 27(3), 822–844.

Brown, B.G., R.W. Katz & Murphy, A.H. (1984). Time series models to simulate and forecast wind speed and wind power. Journal of Climate and Applied Meteorology, 23, 1184-1195.

Fan, S., Hyndman, R. (2012). Short-term load forecasting based on a semi-parametric additive model. IEEE Transactions on Power Systems, 27, 134-141

Hong, T. (2010), Short term electric load forecasting. North Carolina State University

Giebel, G., Brownsword, R., Kariniotakis, G., Denhard, M. & Draxl, C. (2011). The state-of-the-art in short-term prediction of wind power: A literature overview, 2nd edition. Tech. Rep. EU project ANEMOS.plus, available at: <http://orbit.dtu.dk>.

Matos, M.A. & Bessa, R.J. (2011) Setting the operating reserve using probabilistic wind power forecasts. IEEE Transactions on Power Systems, 26, 594-603.

Monteiro, C., Bessa, R., Miranda, V., Botterud, A., Wang, J. & Conzelmann, G. (2009) Wind power forecasting: State of the art 2009. Tech. Rep., Argonne National Laboratory, Decision and Information Sciences Division, ANL/DIS-10-1.

Pinson, P., Chevallier, C. & Kariniotakis, G.N. (2007). Trading wind generation from short-term probabilistic forecasts of wind power. IEEE Transactions on Power Systems, 22, 1148-1156.

Tuohy, A., Meibom, P., Denny, E. & O'Malley, M. (2009). Unit commitment for systems with significant wind penetration. IEEE Transactions on Power Systems, 24, 592-601.

**Bio**

Tao Hong received his B.Eng. in Automation from Tsinghua University, Beijing, an M.S. in Electrical Engineering, an M.S. in Operations Research and Industrial Engineering, and a Ph.D. in Operations Research and Electrical Engineering from North Carolina State University. He is the Head of Energy Forecasting at SAS Institute Inc., Chair of IEEE Working Group on Energy Forecasting, General Chair of Global Energy Forecasting Competition, and Guest Editor-in-Chief of IEEE Transactions on Smart Grid Special Issue on Analytics for Energy Forecasting with Applications to Smart Grid. His research interests are in analytics and the applications to the utility applications.

Pierre Pinson holds a M.Sc. in Applied Mathematics, as well as a Ph.D. in Energy from Ecole des Mines de Paris, France. He is the Associate Professor in Stochastic Energy Systems at the Technical University of Denmark, with a particular interest in probabilistic forecasting for energy and electricity markets. He is involved as principal scientist and workpackage leader in a number of Danish and European projects related to these topics. He acts as editor for the journals Wind Energy and for IEEE Transactions on Power Systems.

Shu Fan received the B.S., M.S., and Ph.D. degrees in the Department of Electrical Engineering from China’s Huazhong University of Science and Technology (HUST), Wuhan, China, in 1995, 2000, and 2004, respectively. He conducted postdoctoral research sponsored by the Japanese Government in Osaka Sangyo University from 2004 to 2006. He was a Visiting Assistant Professor at the Energy Systems Research Center at the University of Texas at Arlington from 2006 to 2007. Presently, he is a Senior Research Fellow at Monash University, Clayton, Australia. His research interests include energy system forecasting, power system control, and high-power power electronics.