

# ***WIND AND SOLAR FORECASTING TRIALS EXPERIENCE: DO'S AND DON'TS***

## ***PART 2: INTRODUCTION TO THE IEA WIND TASK 36 GUIDELINE FOR EVALUATION OF FORECASTING APPROACHES AND SELECTION***

John W Zack, Ph.D.  
Vice President – Grid Solutions  
jzack@awstruepower.com



# OVERVIEW AND CONNECTION TO IEA TASK 36 WP 2

- AWS Truepower Intro
- Trial planning & Setup
- Evaluation Data
- Representativeness of Sample
- Performance Metrics
- Communication of Results to Forecasters



## IEA Task 36: Forecasting for Wind Energy

2016 - 2018

### Task Objective is to encourage improvements in:

- 1) weather prediction
- 2) power conversion
- 3) use of forecasts

### Task Organisation is to encourage international collaboration between:

- Research organisations and projects
- Forecast providers
- Policy Makers
- End-users and stakeholders

### Task Work is divided into 3 work packages:

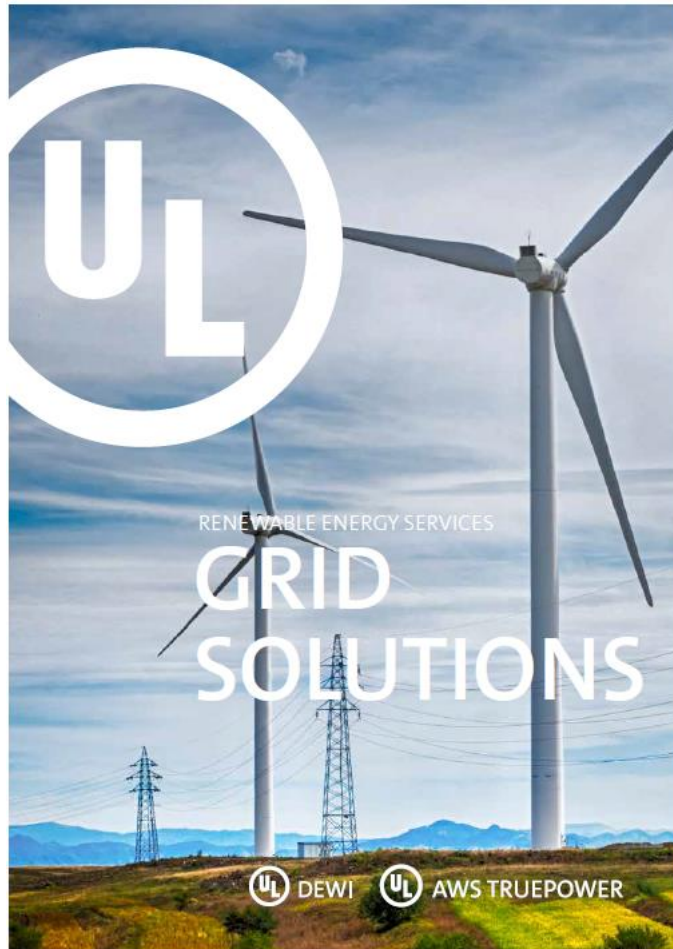
- WP1: Weather Prediction Improvements inclusive data assimilation
- WP2: Development of a benchmarking platform & best practice guidelines
- WP3: Communication of best practice in the use of wind power forecasts



Follow us on our webpage: [www.ieawindforecasting.dk](http://www.ieawindforecasting.dk)

# AWS TRUEPOWER, A UL COMPANY

## GRID SOLUTIONS BRIEF



- Founded in 1983 in Albany, NY
- Acquired by Underwriters Laboratory in 2016
- Short-term and seasonal forecasting for renewable energy generation, utility electric loads, and other weather-sensitive industries
- **Began renewable energy forecasting in 1998**
- Atmospheric modeling and applied research
- Grid integration and curtailment studies
- Emerging smart grid applications related to transmission management, distributed generation, storage management, and others
- Climate change assessment and impact mitigation

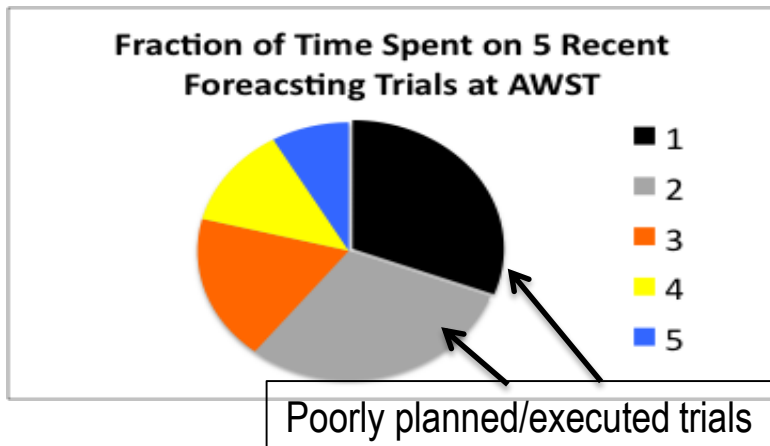
# PLANNING FOR TRIAL/BENCHMARK PROJECT

## • Considerations

- **PREPARE DETAILED TRIAL/BENCHMARK PLAN:** before the trial setup begins evaluator should prepare a detailed trial plan
- **PROVIDE TRIAL PLAN TO EACH PROVIDER**
- **ALLOW TIME FOR PROVIDER TO PROVIDE FEEDBACK**
- **DO NOT CHANGE PLAN DURING TRIAL WITHOUT NOTIFICATION**

## • Impact

- Misunderstandings in trial setup often waste the time of evaluators and providers and can negatively impact representativeness of results



- **Trial/Benchmark Plan should include:**
  - Accurate locations of forecast sites
  - Content and format of data to be provided
  - Mechanism and frequency of providing data
  - Precise definition of forecast target variables
  - Mechanism and frequency of forecast delivery
  - Specify expected outcomes (selection criteria etc.)

# DATA FOR FORECAST EVALUATION

## • Considerations

- **PROVIDE EVALUATION DATASET:** evaluator should either:
  - Provide documentation of exactly how the raw evaluation data will be quality-controlled so that the providers can perform the identical QC, **OR**
  - Provide the exact QC'd dataset that will be used to evaluate the forecasts

## • Impact

- Having the exact data that will be used for evaluation enables
  - the provider to routinely compute their own performance metrics
  - the provider to know exactly the nature of the forecast target variable (for example how outages and curtailments are identified and handled)
- Impact of differences in QC procedures can often be on the order of the differences in performance among providers

# REPRESENTATIVENESS OF SAMPLE

- **Considerations**

- **SIZE:** should be large enough to produce statistically meaningful results.
  - Adjacent forecast cases are often highly correlated
  - Differences in forecast performance may be variable and noisy
  - 3 months may be adequate under ideal circumstances
- **REPRESENTATIVENESS:** should include all of the important modes of variability for the forecast parameter that are relevant to the user
  - Trial timing (winter, summer etc.) & duration should be chosen carefully

- **Impact**

- Long trials are a burden to the evaluator and the providers but unrepresentative ones may be useless

# PERFORMANCE METRICS

## • Considerations

- **VALUE FOR USER'S APPLICATION:** ideally metrics should measure the sensitivity of the user's application to forecast error
  - MAE/RMSE are popular, but do they measure what the user should want to know?
  - Worthwhile reading: DOE SUNSHOT report/papers on forecast metrics
- **REPRODUCIBLE:** method to calculate metrics should be well documented and able to be independently calculated by evaluator and providers
- **ASSESSMENT OF DELIVERY RELIABILITY:** If a real-time trial, a metric for the missed forecast rate should be a part of the trial
- **APPROPRIATE TREATMENT FOR MISSING FORECASTS:** Evaluation sample should be the same for all providers.
  - Eliminate times missed by any provider for ALL providers **OR**
  - Fill-in the missing forecasts with a reference forecast (e.g. persistence, climatology)

## • Impact

- Inappropriate metrics fail to provide optimal information for the evaluator's decision-making process (business case etc.)

# COMMUNICATION WITH FORECASTER PROVIDERS

- **Considerations**

- **RECONCILIATION OF PERFORMANCE RESULTS:** Entity conducting the trial should periodically reconcile performance results with each provider
  - Find reasons for any differences
- **PERIODIC COMPETITIVE PERFORMANCE UPDATES:** provide each provider with anonymous competitive performance data with respect to other trial participants and/or the user's reference benchmark

- **Impact**

- Lack of reconciliation can result in persistence of flaws in the execution of the performance analysis and invalidate results
- Feedback on competitive standing provides forecaster with value for their effort (especially important in free trials) and can also provides added incentive for forecast optimization