The Use of Analog Ensembles to Improve Short-Term Wind Forecasting

Steven H. Young and John W. Zack

AWS Truepower, LLC, MECO, Inc.

The Hawaii Wind Forecast Problem

- Isolated island electric grid with relatively significant wind power penetration (wind capacity is close to 100% of night-time load on Maui, 30% on Oahu).
- Wind farms tend to be on the upwind side of the island with few upwind observatories.
- Wind speeds can change rapidly due to small-scale, poorly observed features such as rain showers or the interaction of terrain, changing stability and the upwind wind speed/direction.
- High resolution, frequently updated NWP (FNL) models can help, but a paucity of upwind observations can limit the effectiveness of this approach.
- WEC and DOE have supported the targeted placement of sensor systems in an effort to improve forecast performance.
- In addition to assimilation into FNL simulations, an analog ensemble approach is being developed to use these observations to improve the 0-3 hour forecast.

Analog Ensemble Method

The analog ensemble method selects a historical sample of similarly observed cases by picking those cases that most resemble the current situation.

1. Choose a set of observed or simulated “case-matching” variables.
2. Compute case-matching score components. A case-matching score component is the difference between a case-matching variable for the forecast case and the same variable from a historical case.
3. Scale case-matching score components by their standard deviation and compute a final case-matching score that measures the “distance” between the forecast case and the historical case in case-matching variable space.
4. The historical cases with the smallest case-matching scores are selected as ensemble members.
5. The method can be combined with regime-based forecasting by allowing only ensemble members that are classified in the same regime in the current forecast case.

The analog ensemble can be used to generate:

- A probabilistic forecast of wind speed or wind power from each of the distribution of wind speeds among ensemble members.
- A probabilistic forecast of up wind speed or power change (ramp rate) from the distribution of wind speed change among ensemble members.

Morning Ramp up at the Apollo Wind Farm

- Apollo wind farm extends from 6 to 7.7 km north of the southern tip of the big island of Hawaii at an elevation of 190-260 m.
- A SuDAR is located at South Point. - 4 km NNE of the wind farm at 60 m.
- Northeast low-level jet is kept offshore by:
  - Nontropical drainage flow/land breeze from the higher terrain.
  - Increased blocking of the stable night-time flow by the terrain.
- As daytime heating returns, upslope/valley sea breeze develops and blocking is reduced. This allows the jet to shift inland.
- Regime Identification:
  - Wind speed at South Point is greater than wind speed at Apollo.
  - Wind speed at South Point has recently increased.
- Time is between 0400 and 1100 HST.
- Wind is from the north/northwest quadrant.

Performance: Probabilistic Forecasts

Scoring Method: Ranked Probability Score

\[ \text{RPS} = -\ln \sum_{k=1}^{K} \left( CDF_{f,k} - CDF_{o,k} \right) \]

Where:
- \( K \) is the number of outcome bins,
- \( CDF_{f,k} \) is the cumulative distribution function for forecast probabilities \( f \) e.g. the sum of all probabilities from 0 to \( k \) and
- \( CDF_{o,k} \) is either the observation falls in bins 1 to \( k \) and if the observation falls in bins \( k+1 \) to \( K \).
- RPS rewards forecasts that assign a high probability to the observed outcome bin and a low probability to other bins.
- RPS \( \leq 0 \) is perfect forecast, RPS \( \geq 0 \) is the worst possible forecast.
- RPS can be averaged over many forecasts to measure probabilistic forecast skill.

Performance: Deterministic Forecasts

MAE For Deterministic Wind Speed Forecast

- Regime average performs only slightly better than persistence.
- Using only cases with similar initial speed reductions most especially at longer lead times, the weighted ensemble average reduces error significantly more, especially at longer lead times.

Case Studies

Case 1: Large Upward Ramp
- A large upward ramp occurred 20 to 70 minutes after forecast time.
- Many members predicted the ramp, but most underestimated the amplitude.
- Ensemble weighted mean showed a small upward ramp about 20 minutes earlier than observed and then a slow increase after that.

Case 2: Slow Downward Ramp
- General observed decrease in wind speed with slight down ramp in first 20 minutes.
- Regime is designed to select upward ramp cases, so downward ramp forecasts are a challenge. However, this case was reasonably well predicted by the ensemble.

Case 3: Small Downward Ramp Then Large Upward Ramp
- Initial moderate downward ramp followed by large upward ramp.
- Some ensemble members show one or both ramps to some degree.
- Weighted mean shows a slow, moderate upward ramp.

Summary, Issues and Future Plans

- An analog ensemble forecast method was tested as a wind speed ramp forecasting tool for the morning ramp up regime at the Apollo Wind Farm near south tip of the island of Hawaii.
- The analog ensemble was compared to baseline forecasts based on climatology and an ensemble of all cases that fall within the morning ramp up regime.
- Most of the probabilistic ramp forecasting benefit over climatology was obtained with the regime based method. The analog ensemble provided some additional benefit at most lead times.
- Deterministic wind speed forecasts showed significant improvement over the regime and climatology at lead times of 30 minutes and longer.
- For upward ramps of at least 1.5 m/s, the analog ensemble provided some additional benefit at most lead times.
- Limited data from the wind farm and a few SuDARs near the wind farm were available to generate case matching variables.
- Next: identify additional data sources or innovative ways of using existing data to generate an ensemble with more predictive value. Possibilities include:
  - Radial velocity data from a S-1500 rotor blade for the wind farm
  - Data from an NWP model that is well initialized with or near available remotely sensed data.
- For more information, contact Steven Young at steve@meeco.com.