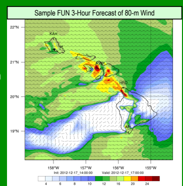
The Use of Analog Ensembles to Improve Short-Term Wind Forecasting

Steven H. Young and John W. Zack
AWS Truepower, LLC, MESO, Inc.

The Hawaii Wind Forecast Problem

- Isolated Island electric grids with relatively significant wind power penetration (wind capacity is close to 100% of night-time load on Maui, 35% on Big Island and 10% on Oahu).
- Wind farms tend to be on the upstream (east to northeast side) of the islands with few upstream observations.
- Wind speeds can change rapidly due to small-scale, often poorly observed features such as rain showers or the interaction of terrain, changing stability and the upstream wind speed/direction.
- High-resolution frequently updated NWP (FUN) models can help, but a paucity of upstream observations can limit the effectiveness of this approach.
- HECO and DOE have supported the targeted placement of sensor systems in an effort to improve forecast performance.
- In addition to assimilation into FUN simulations, an analog ensemble approach is being developed to use these observations to improve the 0-3 hour forecast.



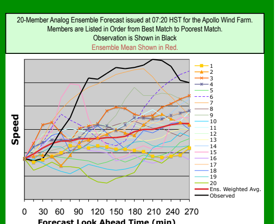
Analog Ensemble Method

The analog ensemble method selects a historical sample of similar situations by picking cases that most closely 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 that same variable from a historical case.
3. Scale case-matching score components by their standard deviation and combine into a 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 as the current forecast case.

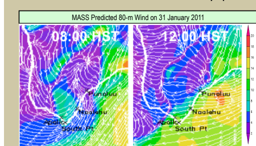
The analog ensemble can be used to generate:

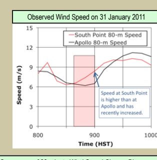
- A deterministic wind speed or power forecast from the ensemble mean.
- A probabilistic forecast of wind speed or wind power from the distribution of wind speed among ensemble members.
- A probabilistic forecast of wind speed or power change (ramp rate) from the distribution of wind speed change among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



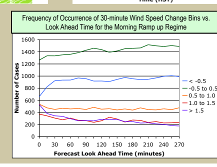
Morning Ramp up at the Apollo Wind Farm

- Apollo wind farm extends from 6 to 7.75 km north of the southern tip of the big island of Hawaii at an elevation of 190-260 m.
- A SoDAR is located at South Point: ~ 4 km SSE of the wind farm at 60 m.
- Northeasterly low-level jet is kept offshore by:
  - Nocturnal drainage flow/ land breeze from the higher terrain.
  - Increased blocking of the stable night-time flow by the terrain.
- As daytime heating returns, upslope flow/ 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 northeasterly quadrant.





- For the first 30 minutes of the forecast, wind speed decreases are less frequent while wind speed increases, especially large ones, are more frequent.
- This suggests the regime is a useful intra-hour forecast tool.



Experimental Design

- **Experimental Sample**: 3360 morning ramp up regime cases from 1 November 2010 to 31 August 2012.
- **Data:** Apollo 15-minute average 80-m wind speed and South Point 10-minute average 80-m wind speed and direction.
- **Ensemble Selection Parameters**: (1) Apollo wind speed, (2) South Point wind speed, (3) South Point wind direction, (4) difference between South Point and Apollo speeds, (5) recent 20-minute speed change at South Point, (6) correlation with the recent 2.5-hour speed time series at Apollo.
- **Case Matching Score**: Square root of the sum of the squares of the six scaled selection parameters.
- **Ensemble Member Selection**: the 25 best matches (i.e. lowest case matching scores) in the historical sample as analog ensemble members.
- **Forecast Production:**
  - Shift wind speeds of ensemble members upward or downward so that the initial member wind speed matches the initial wind speed for the forecast case.
  - Weight ensemble members by the inverse of their case matching score.
  - Create deterministic wind speed forecasts from the weighted ensemble mean.
  - Create probabilistic wind speed and wind speed ramp rate forecasts from the weighted distribution of wind speeds and ramp rates in the analog ensemble.

Performance: Probabilistic Forecasts

Scoring Method: Ranked Probability Score

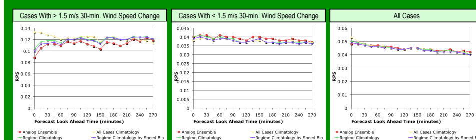
- Used to verify probabilistic forecasts

$$RPS = \frac{1}{K-1}\sum_{k=1}^{K}\left(CDF_{fc,k} - CDF_{obs,k}\right)^2$$

Where:

- K is the number of outcome bins,
- $CDF_{fc,k}$ is the cumulative distribution function for forecast probabilities (i.e. the sum of all probabilities from 1 to k) and
- $CDF_{obs,k}$ is 0 if the observation falls in bins 1 to k and 1 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 = 0 is a perfect forecast, RPS = 1 is the worst possible forecast.
- RPS can be averaged over many forecasts to measure probabilistic forecast skill.

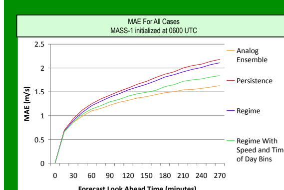
RPS For Probabilistic 30-minute Wind Speed Change Forecasts



- RPS is computed using 14 wind speed change bins: <.-6 m/s, 12 bins of 1 m/s width from -6 to +6 m/s and > +6 m/s.
- Ramp cases (observed wind speed increase > 1.5 m/s, left plot):
  - the morning ramp up regime as a whole performs noticeably better than climatology for the 0-30 minute speed change and slightly better for the 15-45, 30-60 and 45-75 speed changes.
  - Including only cases with similar initial wind speed gives a small improvement.
  - The analog ensemble performs slightly better than either regime method at all lead times up to 210 minutes.
- All other cases (center plot):
  - the regime-based methods and especially the analog ensemble perform very slightly worse than climatology.
- Entire Sample (right plot):
  - The analog ensemble has a very slight performance edge over climatology at very short lead times and a negligible edge over the regime based methods also at very short lead times.

Performance: Deterministic Forecasts

MAE For Deterministic Wind Speed Forecast

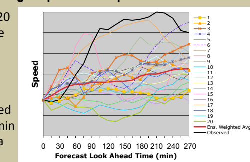


- Regime average performs only slightly better than persistence.
- Using only cases with similar initial speed reduces error, especially at longer lead times.
- Using the weighted ensemble average reduces error significantly more, especially at longer lead times.

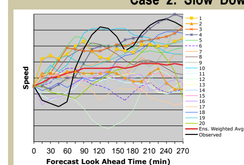
Case Studies

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 min earlier than observed and then a slow increase after that.



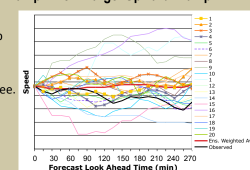
Case 2: Slow Downward Ramp



- 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 the southern tip 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 lead times of 0-30 minutes.
- 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 SoDARs 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 WSR-88D located ~19 km NE of the wind farm
  - Data from an NWP model that is well-initialized with all or most available remotely sensed data.
- For more information, contact Steve Young at steve@meso.com.