What’s New with Probabilistic Forecast Development

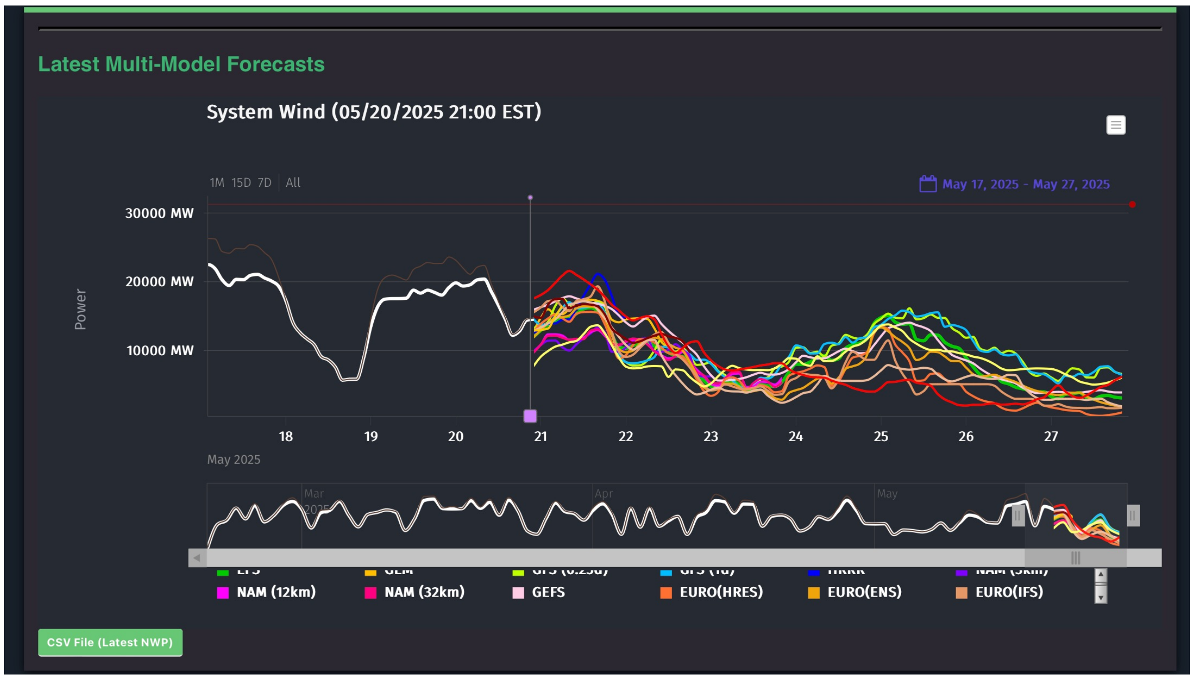
Craig Collier
ESIG Forecasting and Markets Workshop
Nashville, Tennessee
June 25, 2025

Energy Forecasting Solutions

Needs for Probabilistic Renewables Forecasts

- Classical predictions of variable power are deterministic, still widely applied in operations.
- Several specific power systems management applications benefit from forecast **uncertainty** information:
  - Reserves management
  - Trading strategies
  - Unit commitment
  - Storage sizing for storage systems

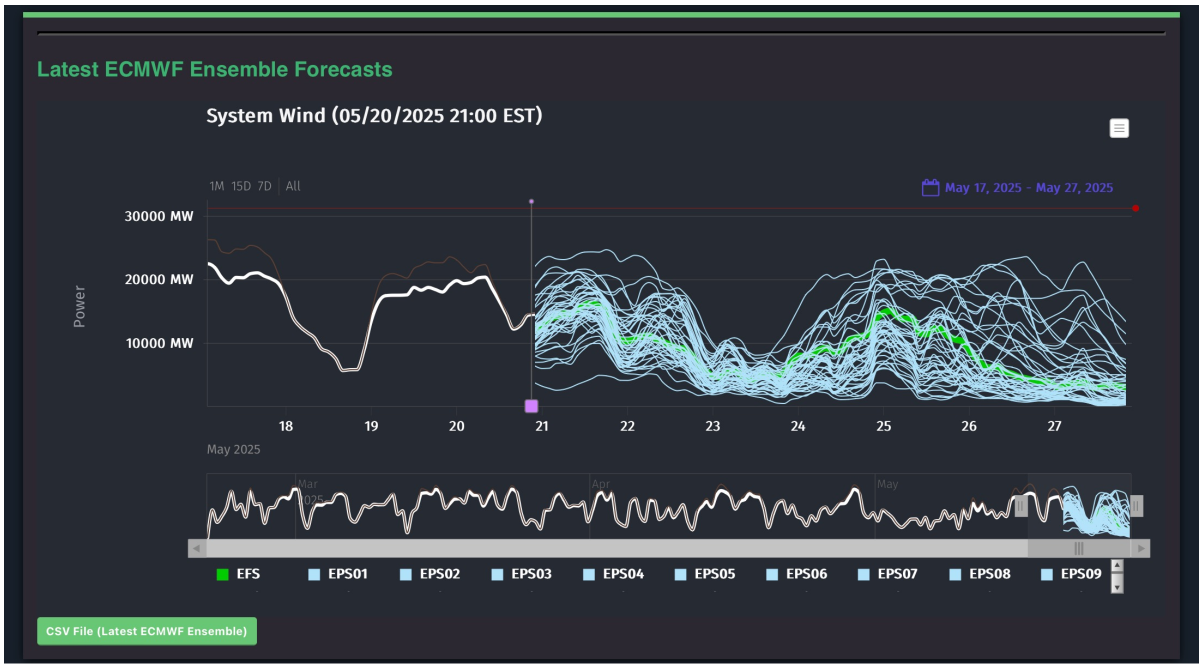
How Do We Convey Uncertainty?



Multi-model forecasts are differentiated by physics and dynamics cores, horizontal and temporal resolution, parameterization.

Multi model scenarios not well-suited for automated risk management and reserves estimation – but rather for situational awareness.

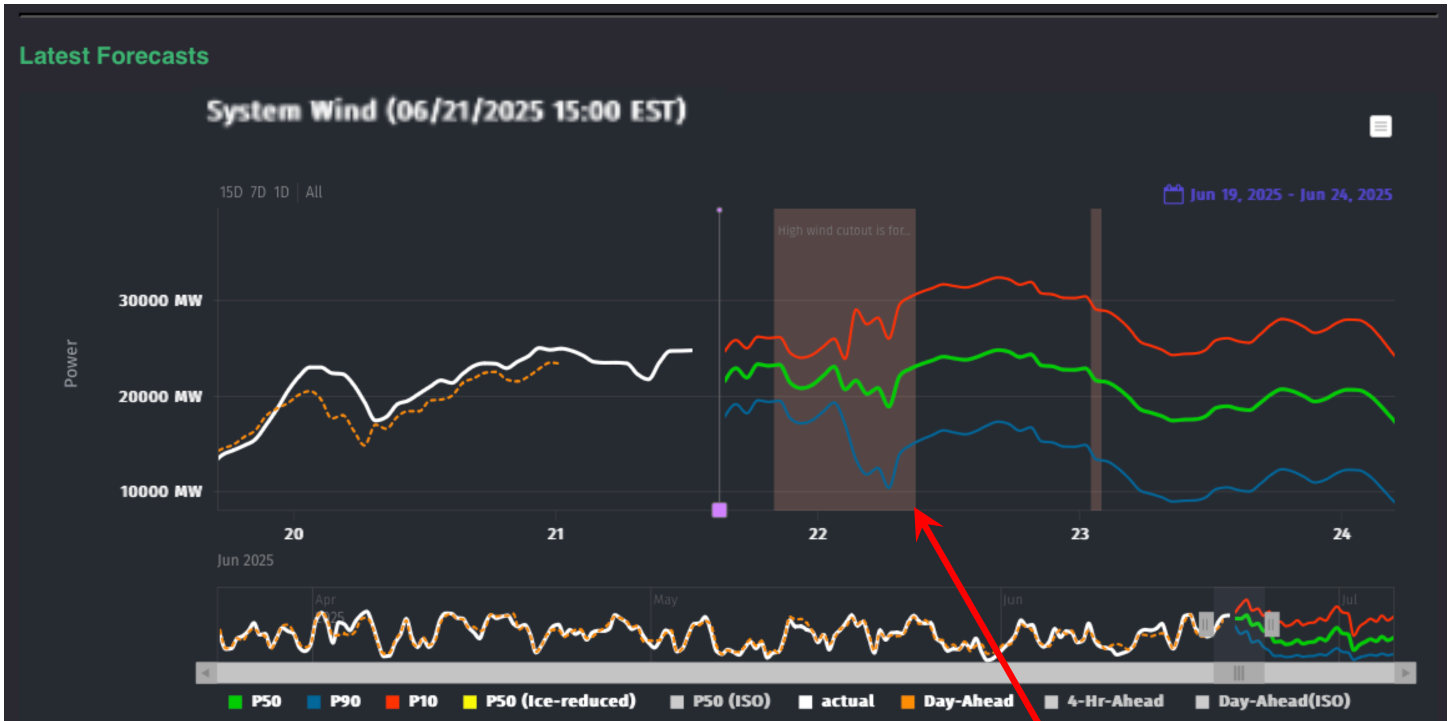
How Do We Convey Uncertainty?



Single-model ensemble members derive from the same underlying forecasting system but typically use uniquely-perturbed initial conditions.

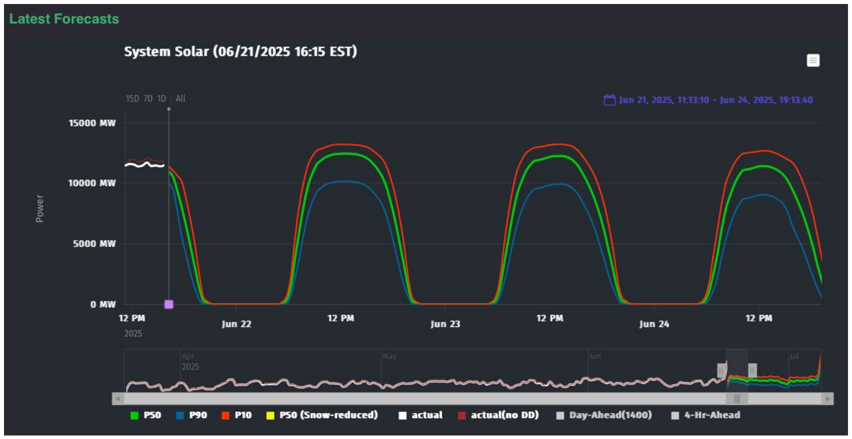
Single-model ensembles may be used in confidence interval determination but, on their own, are primarily for situational awareness for event-based risk.

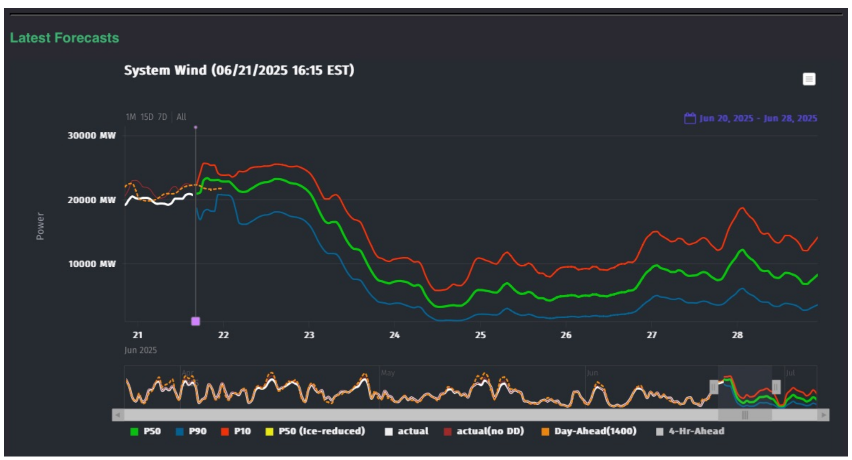
How Do We Convey Uncertainty?



Forecast Warnings and Alerts generally are only visual and focused on discrete event types only.

How Do We Convey Uncertainty?





- Most actionable probabilistic forecast is typically continuous and distribution-based
  - Confidence interval defines a risk tolerance
  - Probability-of-Exceedance (POE) → so-called “P Levels”
- Primary objective of the Forecaster is to determine estimate the probability of something occurring based on X , or X + Y , or X + Y + Z, or . . .
- Derives from Conditional Probability Distribution Functions

Two Families of Probabilistic Forecasting Techniques

A. Parametric

- simple and most widely used
- assume what we predict has an easy-to-fit distribution
- forecasters use different methods to estimate distribution parameters, including Bayesian Theory and Regression

Probabilistic Renewables Forecasts

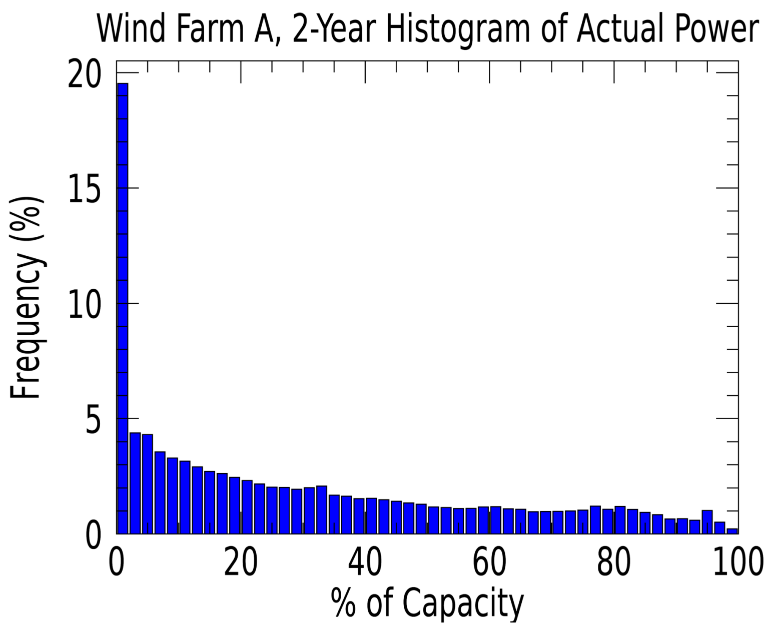
Energy ForecastingSolutions

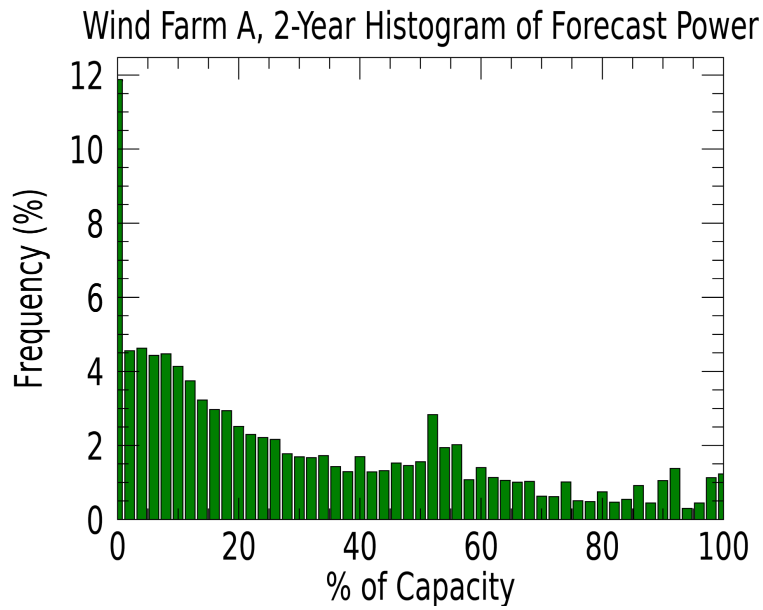
Two Families of Probabilistic Forecasting Techniques

B. Non-parametric

- Derive a model of the forecasts and the actuals with no assumption of underlying distribution
- Span techniques from Kernel Density Estimation to artificial neural networks and generative AI
- Most computationally expensive but more versatile than parametric techniques

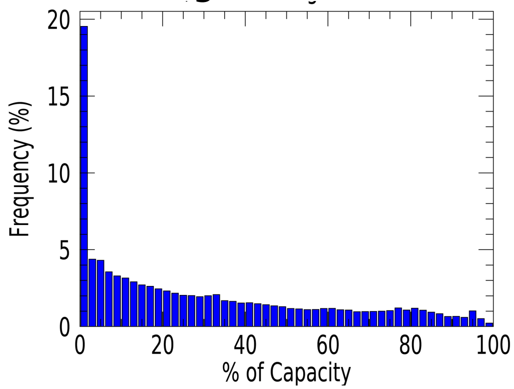
Distributions for a Single Wind Farm



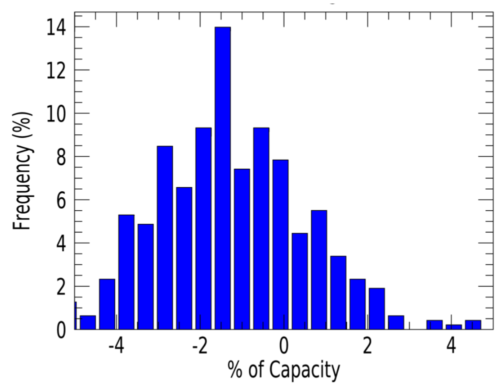


Distributions for a Single Wind Farm

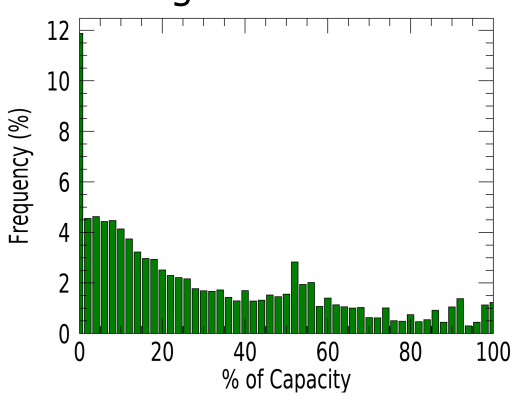
Histogram: Actuals



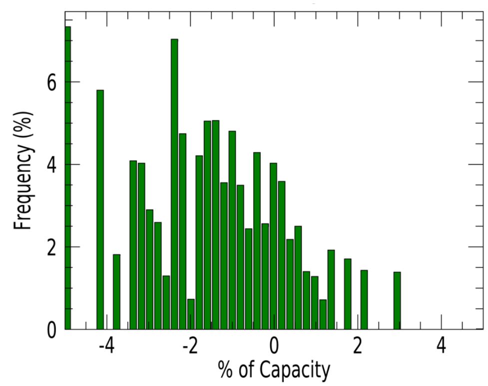
Histogram: Trans. Actuals



Histogram: Forecasts



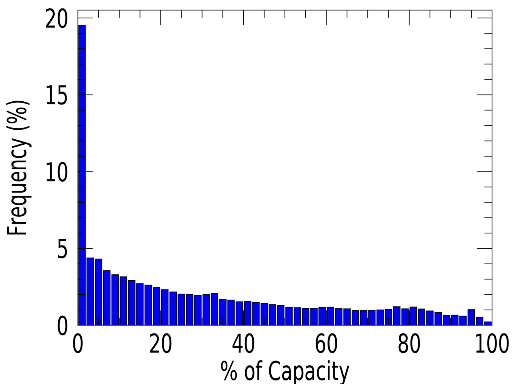
Histogram: Trans. Forecasts



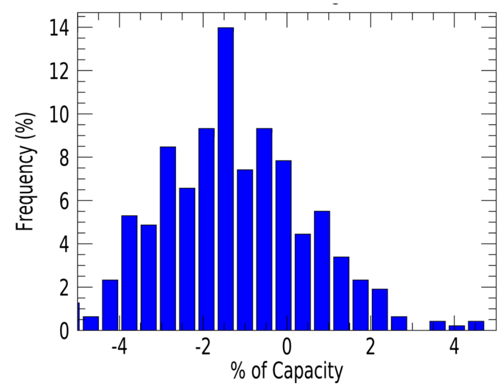
- Can be transformed into logarithmic space (Mauch et al. 2013; Shen et al, 2025, and others).
- Resembles a Gaussian – use Bayesian learning to estimate the parameters of the posterior.
- Simple, cost effective but makes a huge assumption on the underlying predictive data.

Distributions for a Single Wind Farm

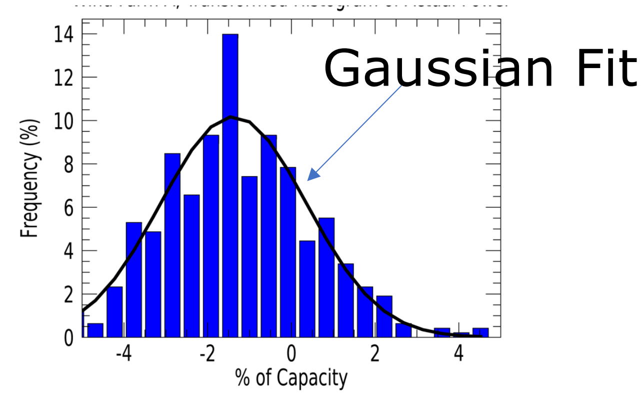
Histogram: Actuals



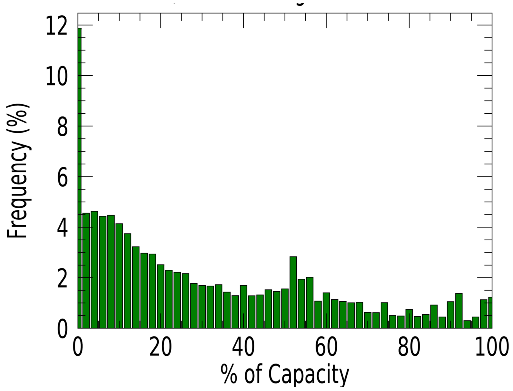
Histogram: Trans. Actuals



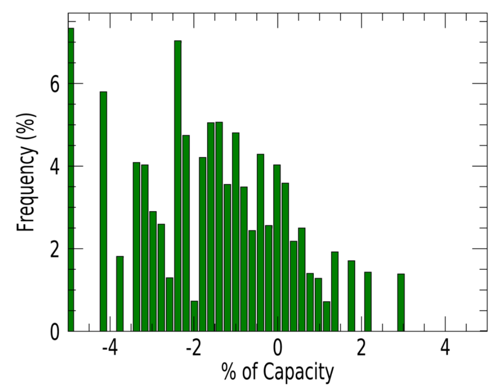
Histogram: Trans. Actuals



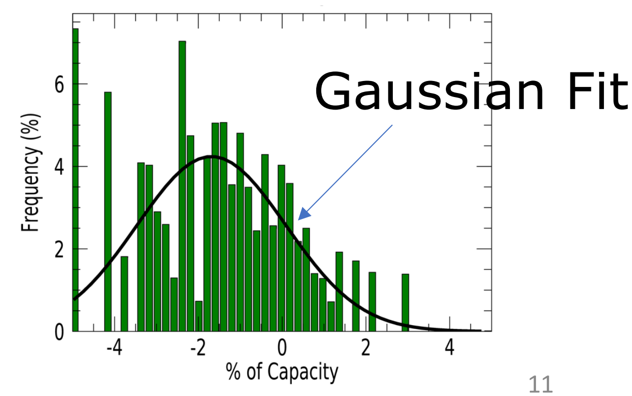
Histogram: Forecasts



Histogram: Trans. Forecasts

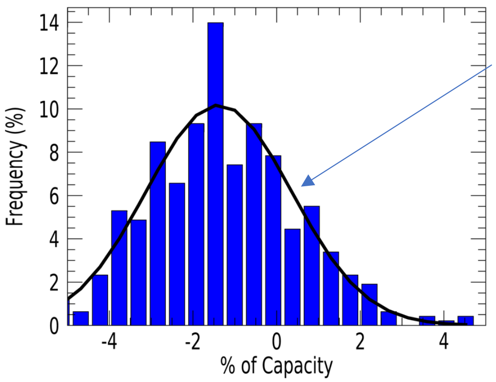


Histogram: Trans. Forecasts



Distributions Jointly Form a Conditional PDF/CDF

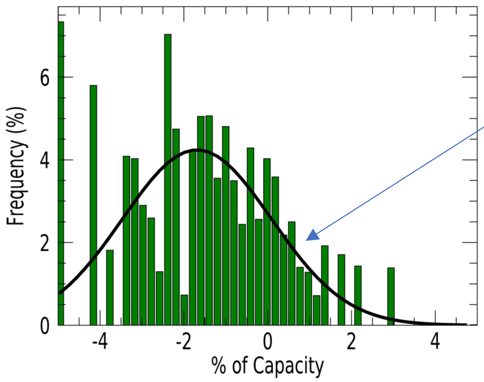
Histogram: Trans. Actuals



Gaussian Fit

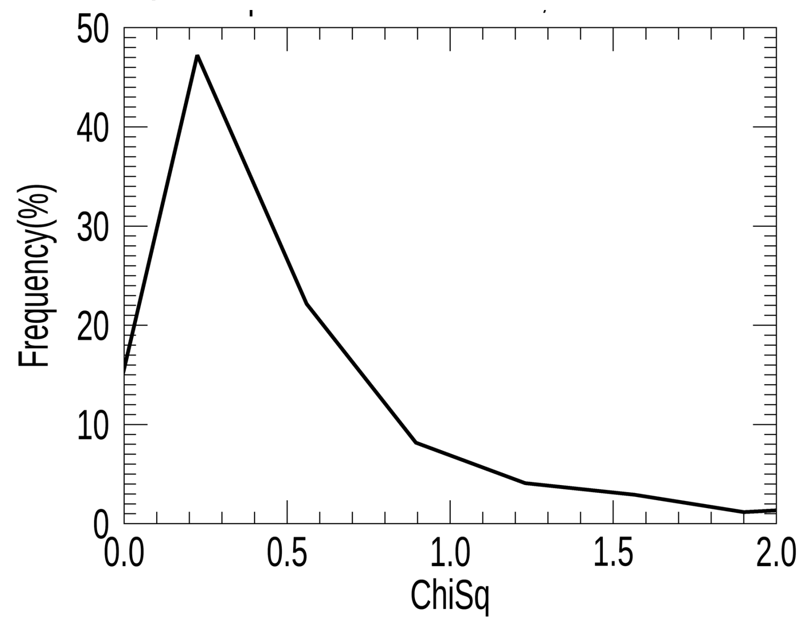
$$\longrightarrow \quad P(Y|X)$$

Histogram: Trans. Forecasts



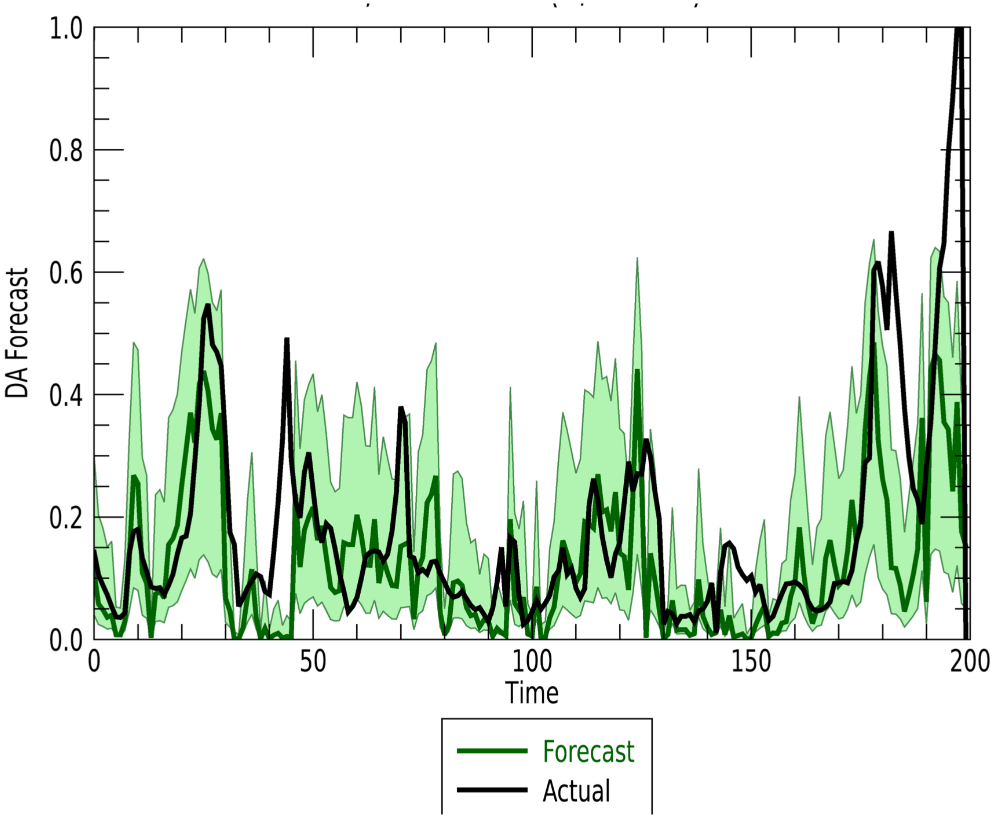
Gaussian Fit

Chi Square of Posterior Dist., Forecasts



The CDF gives a (1-∝)% Confidence Interval

Wind Forecast w/ 80% Confidence Interval



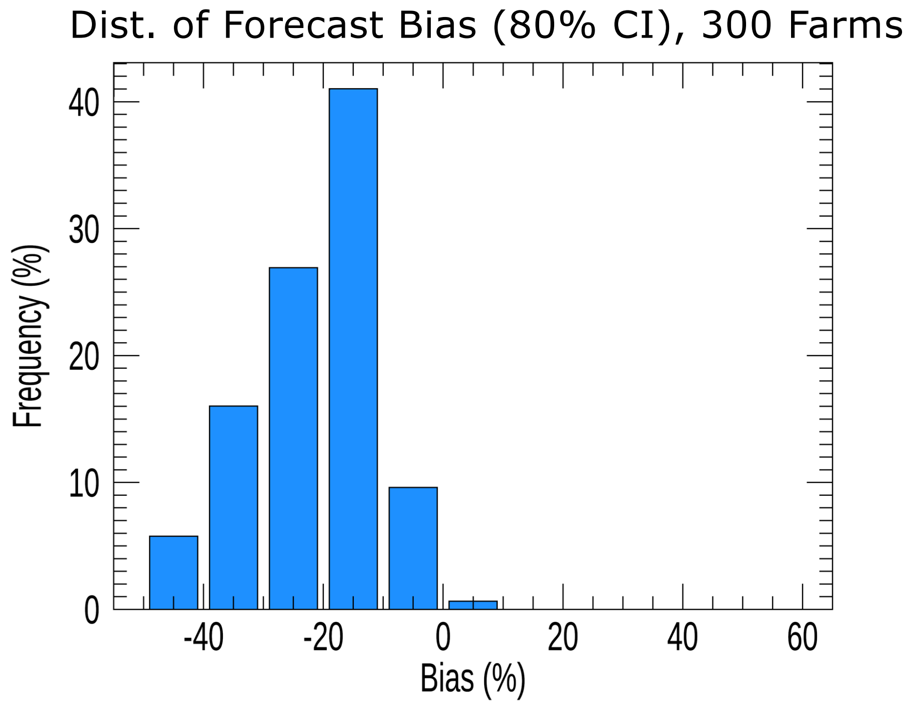
How do we assess it?

- Log Loss
- Brier Skill Score
- Continuous Ranked Probability Score (CRPS)

- . . . among others

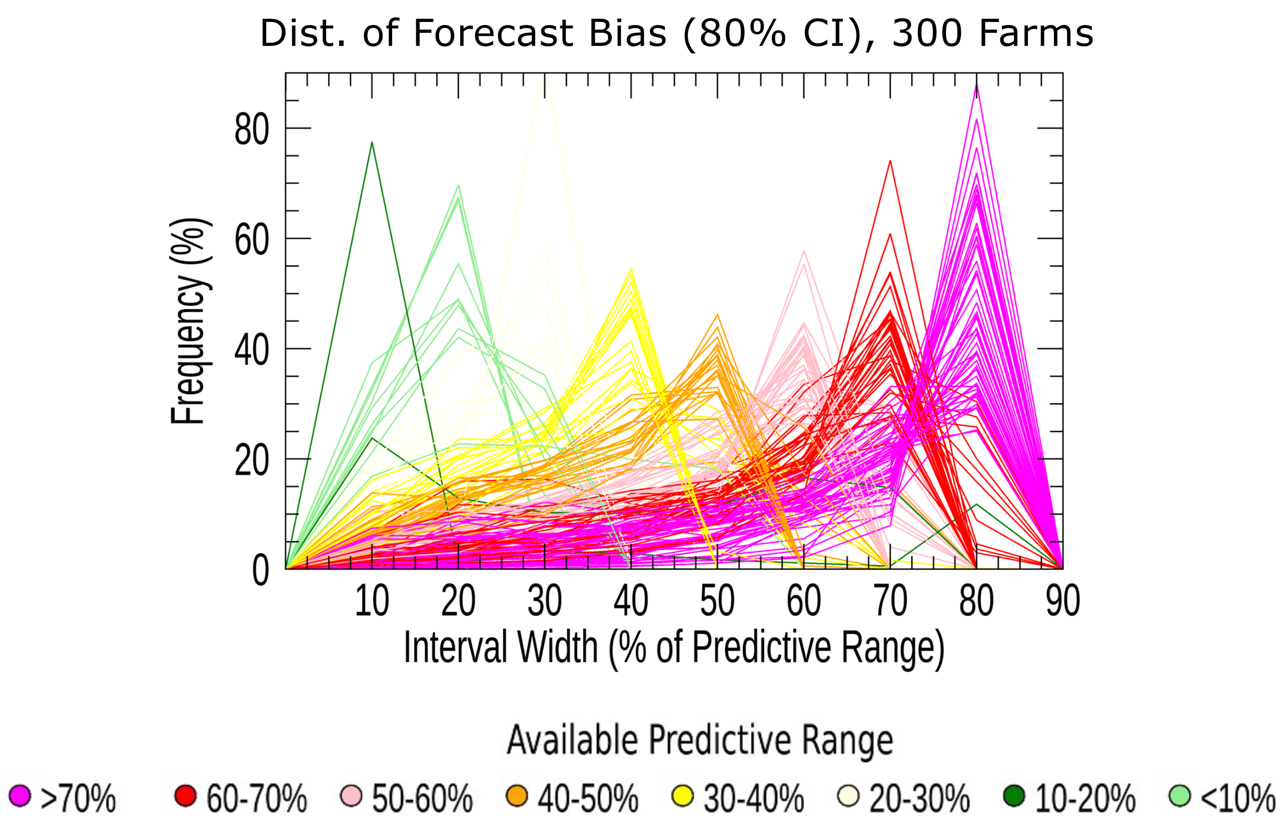
Two Key Features: Reliability & Sharpness

Reliability:
Compare nominal confidence interval to the empirical confidence interval
→ what is the bias?



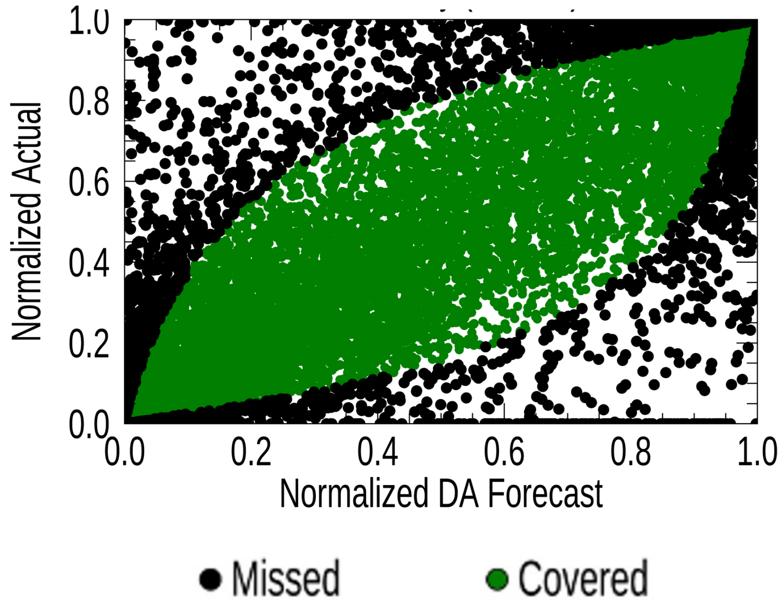
Two Key Features: Reliability & Sharpness

Sharpness: Is the interval precise enough?

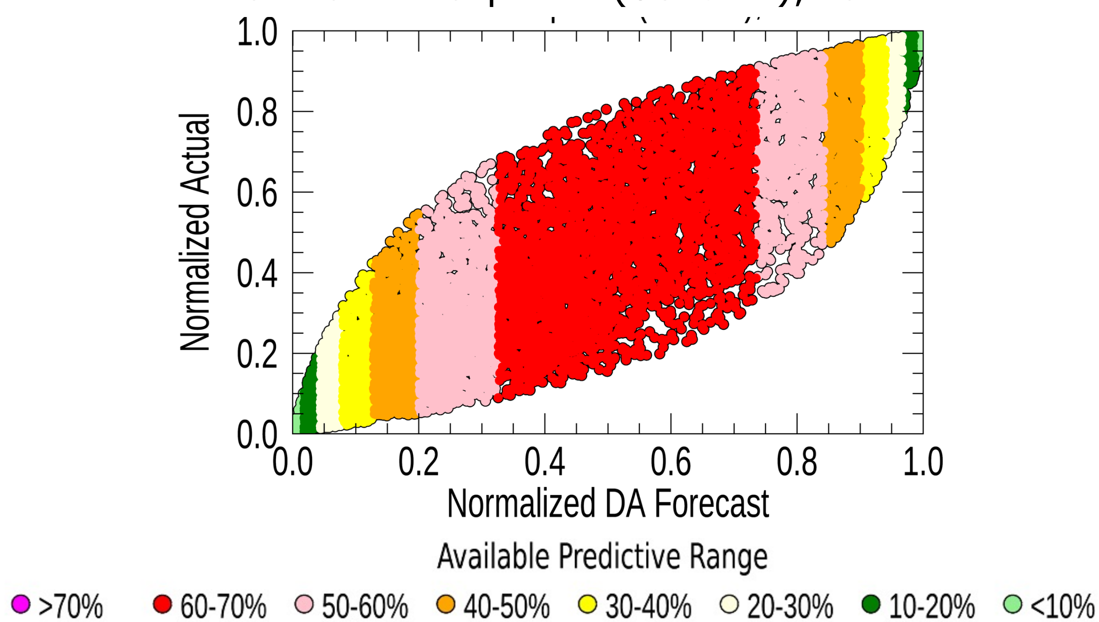


Quality of the Confidence Interval: Reliability & Sharpness

Forecast Reliability (80% CI), Farm X



Forecast Sharpness (80% CI), Farm X

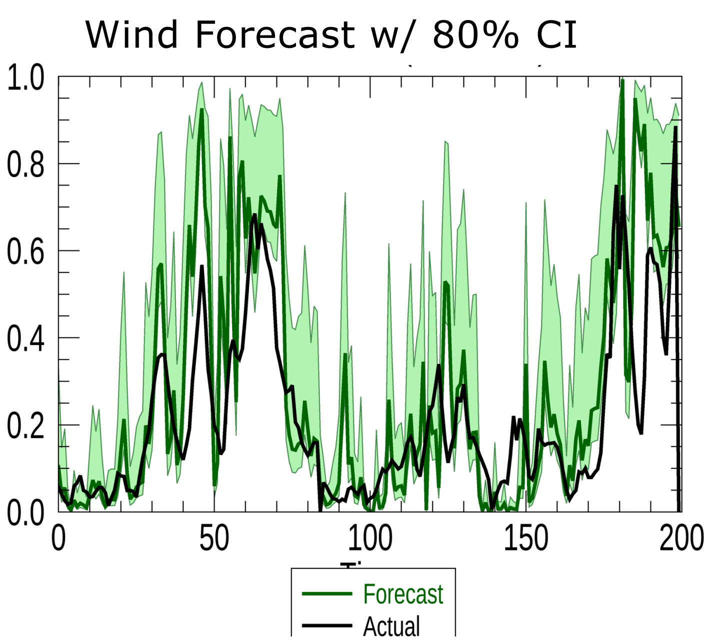


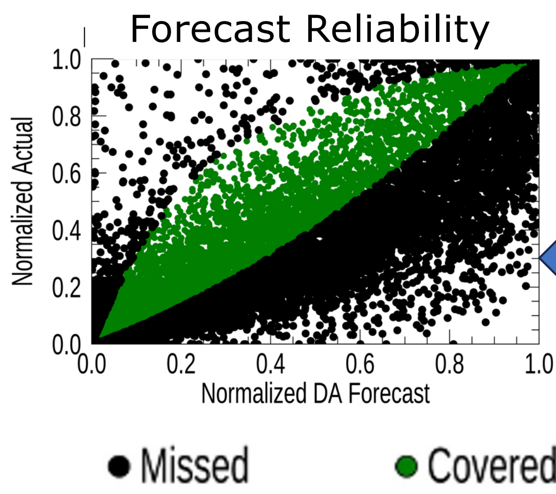
Fewer black dots → better

Warmer colors → fatter intervals (less sharp)

Probabilistic Forecast with Downside Risk

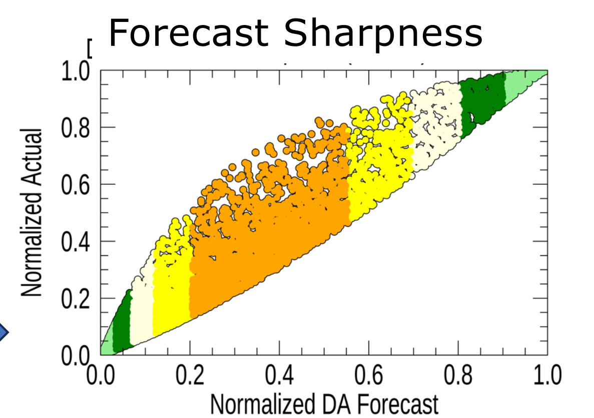
Energy ForecastingSolutions





Interval is noticeably top-heavy.

Frequently consumes between 30 and 60% of the predictive range.

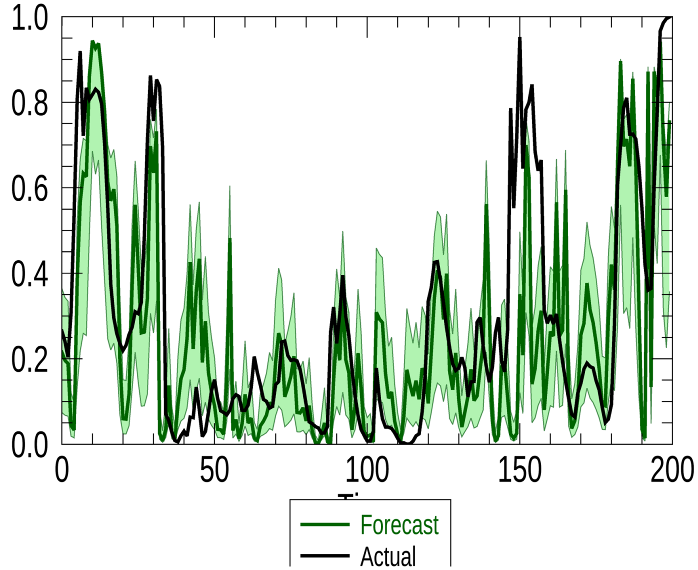


Available Predictive Range

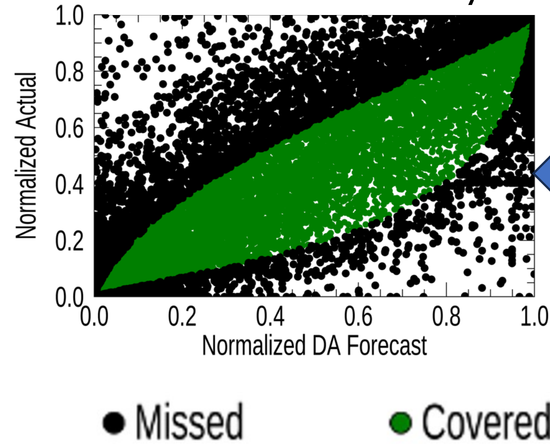
>70% 60-70% 50-60% 40-50% 30-40% 20-30% 10-20% <10%

Probabilistic Forecast with Upside Risk

Wind Forecast w/ 80% CI



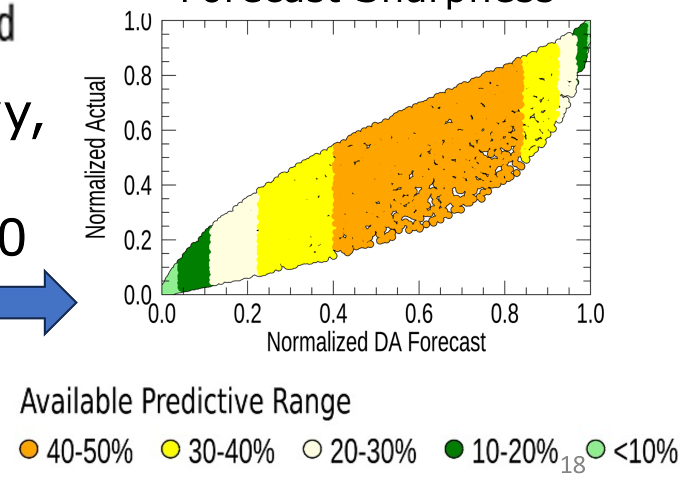
Forecast Reliability



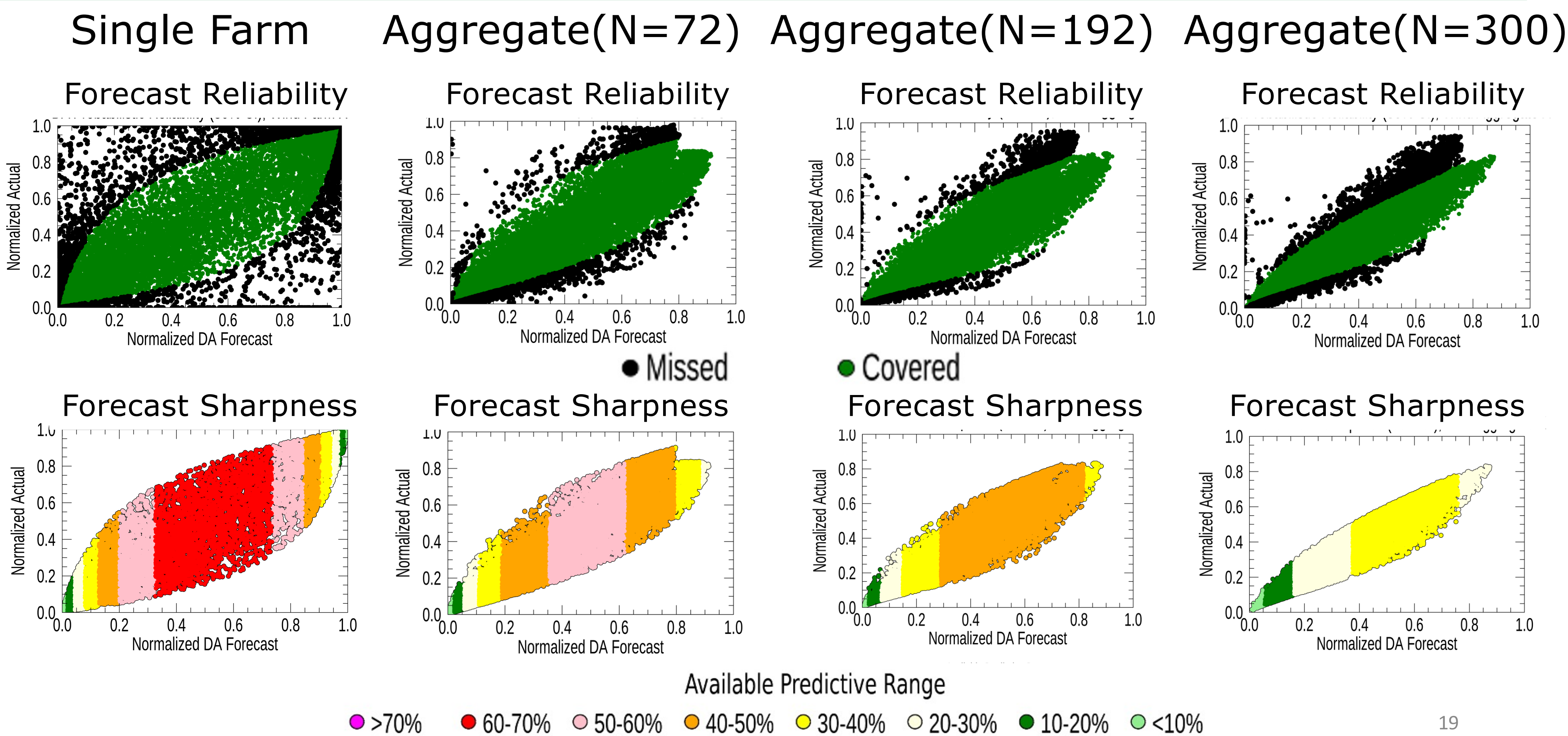
This probabilistic forecast has a noteworthy deficit: it tends to miss over-performance events.

Interval bottom-heavy, and most frequently consumes between 40 and 50% of the predictive range.

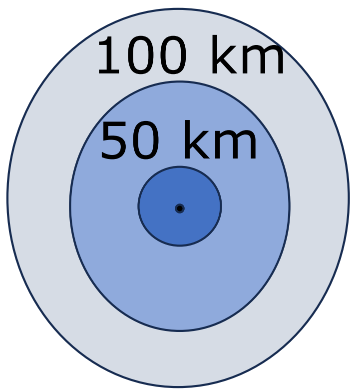
Forecast Sharpness

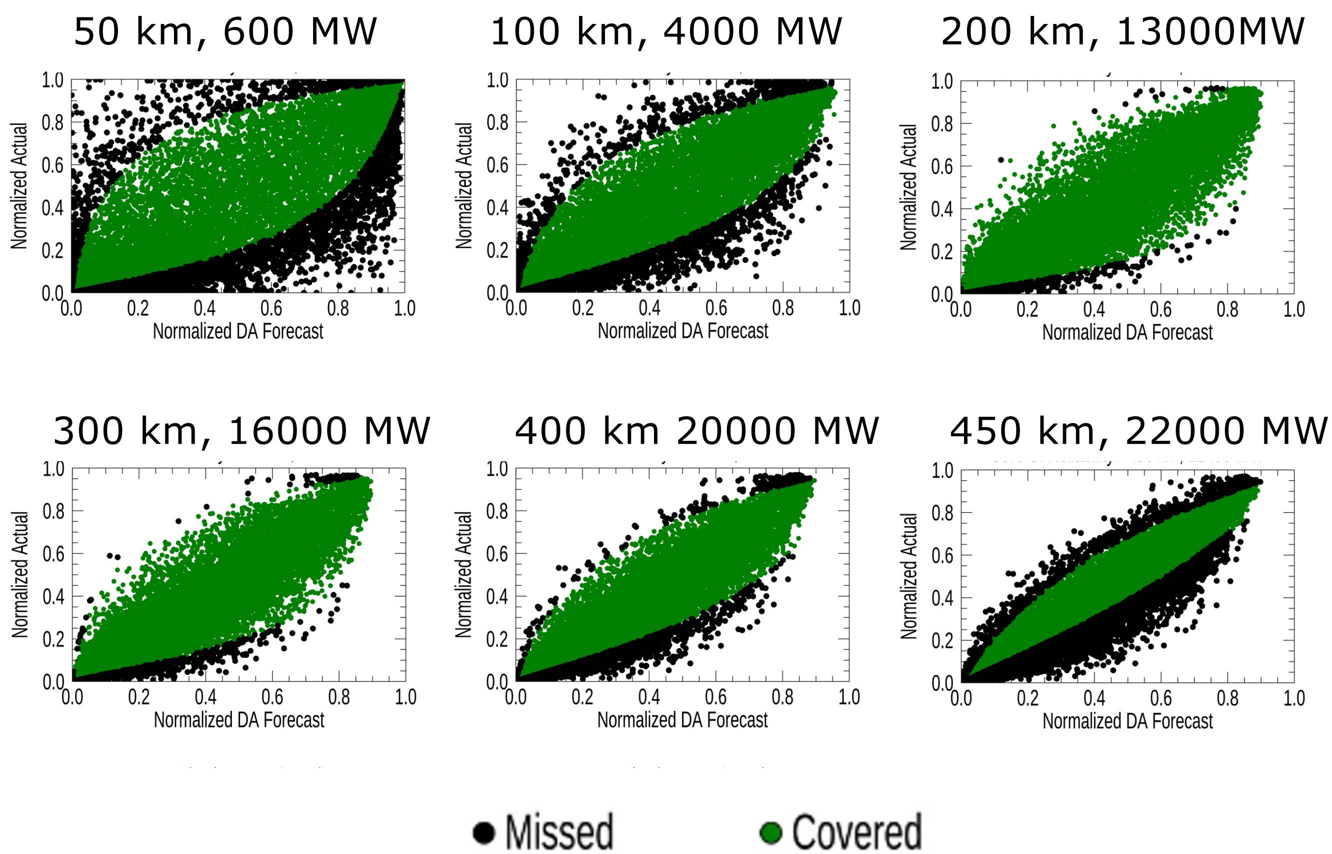


Effects of Aggregation on Reliability and Sharpness

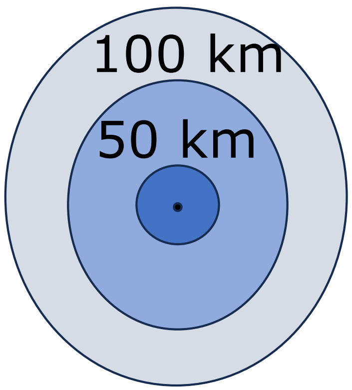


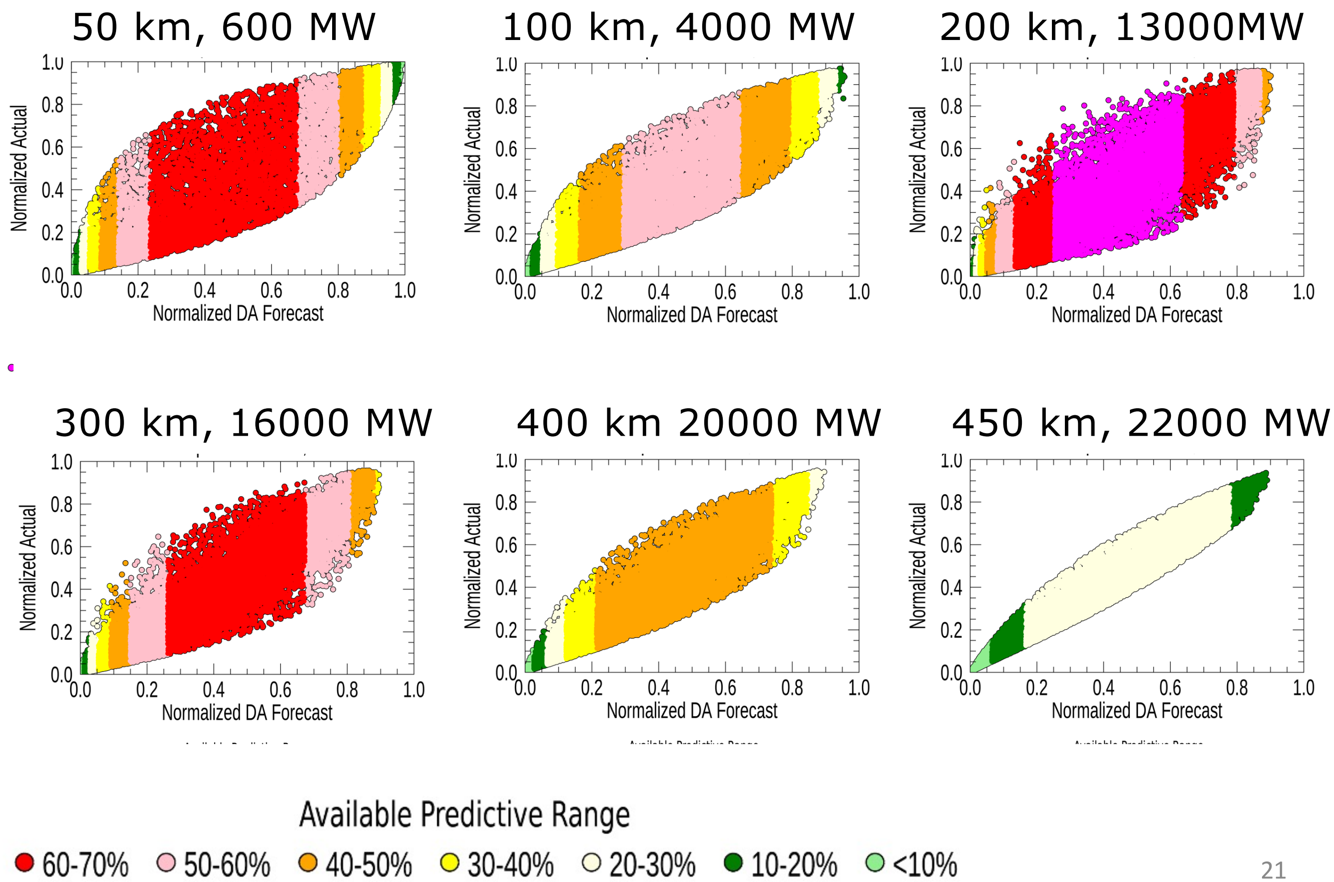
Effects of Distance on Probabilistic Reliability



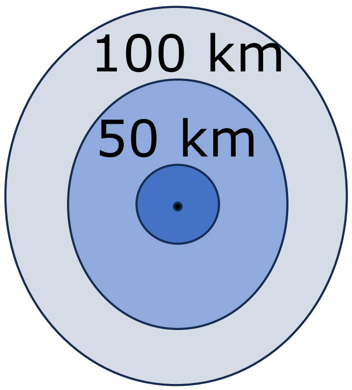


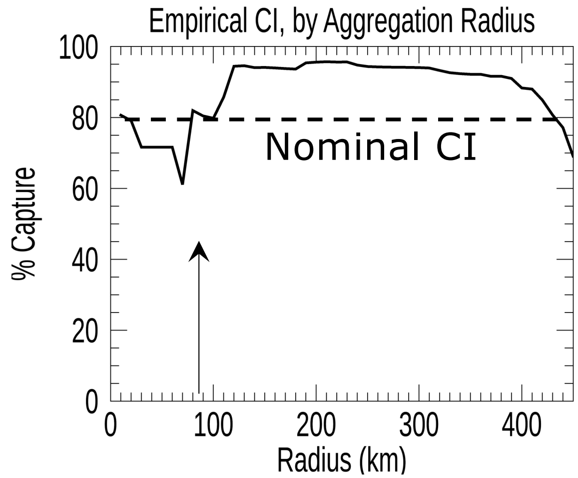
Effects of Distance on Probabilistic Forecast Sharpness

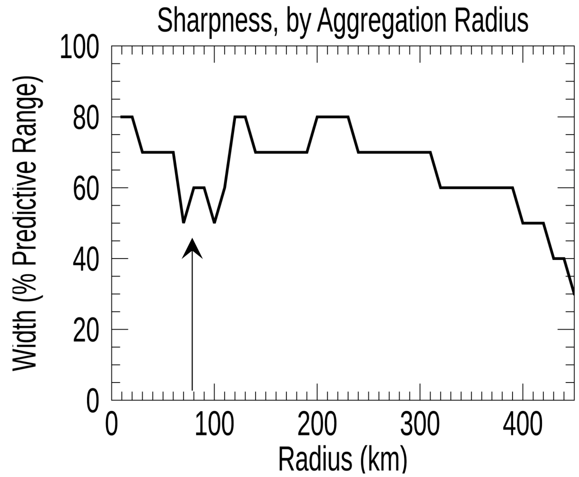




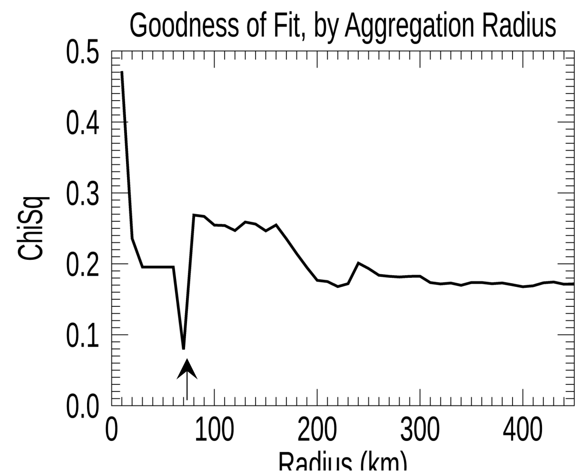
Aggregation Effects on Probabilistic Forecast Value







Over multiple simulations from multiple central wind farm nodes in the central U.S., the confidence interval forecast demonstrates peak effectiveness just under 100 km radius.



Pioneering Non-Parametric Techniques

Kernel Density Estimation (KDE)

Zhang & Wang(2014), Lin et al. (2015), Shi & Chen (2019)

The Law of Total Probability used to estimate the joint PDF of forecast and actual. What's the right kernel function?

Gibbs Sampling of Wx Scenarios

Sun et al. (2020)

Lots of historical NWP data and actuals needed → generate thousands of weather scenarios iteratively → conditional distribution.

Quantile Regression Averaging

Nielsen et al. (2006), Haque et al. (2014), Wan et al. (2017)

Independent NWP variables → conditional quantile functions. CIs can be too wide.

Analog Ensemble

Alessandrini et al. (2015)

Lots of historical NWP data and actuals needed → Monte Carlo approximations of the PDF.

Pioneering Non-Parametric Techniques

AI Methods (LUBE)	Generative AI Methods
Khosravi et al (2011), Quan et al. (2014), Wu et al. 2018) Intervals determined using feed-forward neural networks and minimization of a cost function.	Song et al. (2021), Miraki (2025) Graph-based Denoising Diffusion Probabilistic Models (DDPMs) Uses NNs to add noise to a signal, learn how it is added in order to remove it → clean signal; slow but potentially most accurate

Thank You

craig.collier@energyforecastingsolutions.com