PIRP Workshop

Hosted by CA ISO: March 23, 2005

Wind Power Production Forecasting for CA ISO PIRP

John W. Zack
AWS Truewind LLC
Albany, New York
jzack@awstruewind.com

Overview

- PIRP forecasting requirements
- The forecasting challenge
- Meeting the challenge: the forecast tools
- PIRP forecast system architecture
- Forecast performance: examples & factors
- Forecasting the future of forecasting

PIRP Forecast Specifications

Delivery

Performance

PIRP Forecast Requirements: Delivery

- Next Hour
  - Production (MW) for each of the next 7 hours
  - Delivered by 15 minutes after each hour
- Next Day
  - Production (MW) for each hour of next calendar day
  - Delivered by 5:30 AM Pacific Prevailing Time (PPT)
- Extended Forecasts
  - Production (MW) for each hour of days 2, 3 and 4 after delivery day
  - Delivered by 5:30 AM PPT on Thursdays, Fridays and on selected days before scheduling holidays

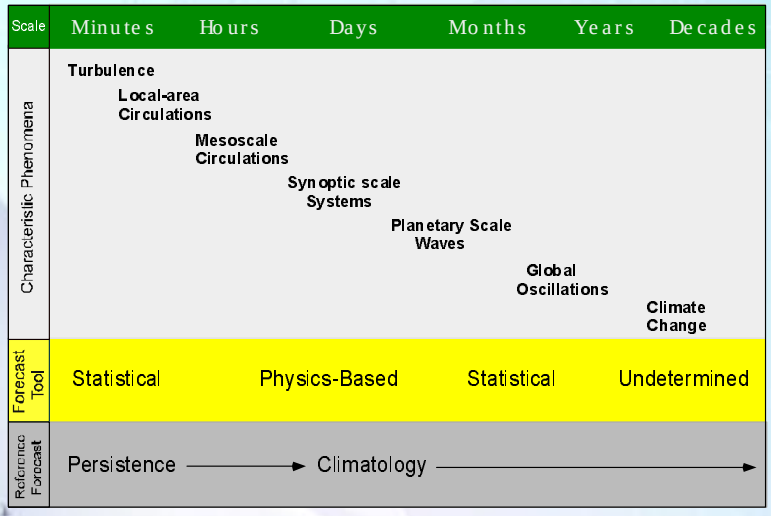
PIRP Forecast Requirements: Performance

- Next Operating Hour
  - Definition: Hour starting 2 hr 45 min after forecast delivery
  - Penalty
    - Monthly MAE > 12% of installed capacity
    - Monthly Bias > 0.6% of monthly production
  - Bonus
    - Monthly MAE < 10% of installed capacity
    - Monthly Bias < 0.1% of monthly production
- Next Day
  - No performance criteria

The Forecasting Challenge

Space and Time Scales

Wind Energy Forecasting Time Scales



Forecast Tools

Physics-Based Models

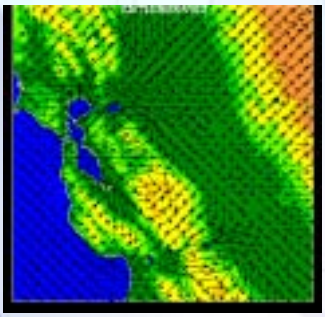
Statistical Models

Meteorological Data

Wind Plant Data

The physics-based model approach

Differential equations for basic physical principles are solved on a 3-D grid

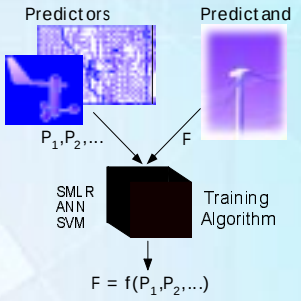


Performance Factors

- Grid has finite resolution - some processes are at the "sub-grid" scale
- Sub-grid scale processes must be modeled since they affect grid scale
- **Initial values for all variables must be specified for all grid cells.**
- Boundary values must be specified for all boundary cells (usually from another model with a larger domain)
- Numerical methods must be used to solve the equations - therefore solutions are not exact

The statistical model approach

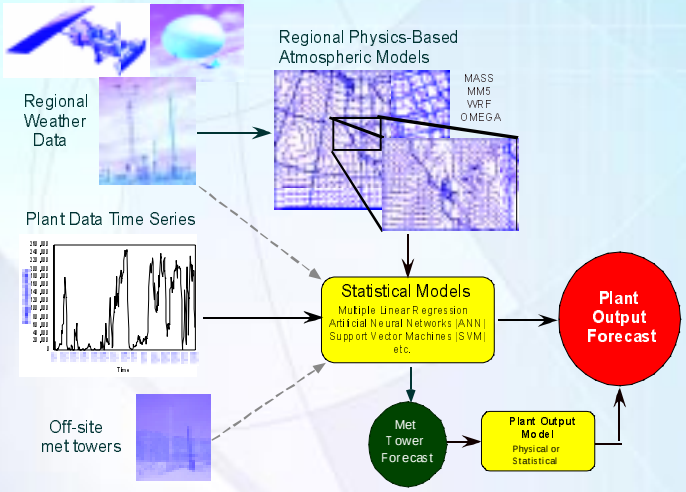
- Empirical equations are derived from historical predictor and predictand data ("training sample")
- Current predictor data and empirical equations are used to make forecasts



Performance factors

- Type & configuration of the statistical model and training algorithm
- Size, quality and representativeness of the training sample
- Issue: difficult to understand the reasons for observed performance

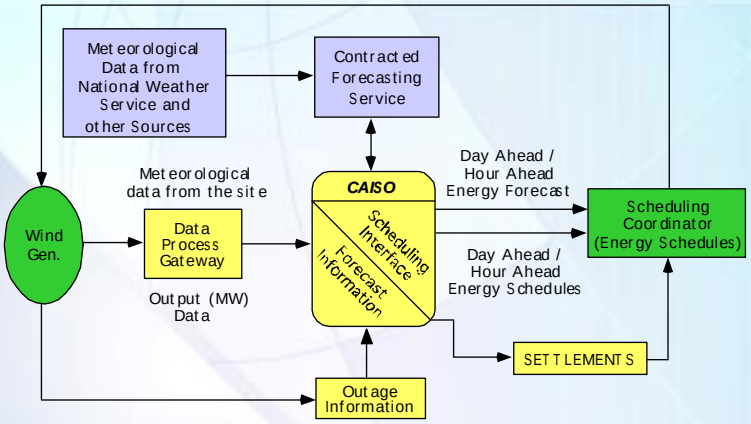
Integrated System: eWind



PIRP Forecast System Architecture

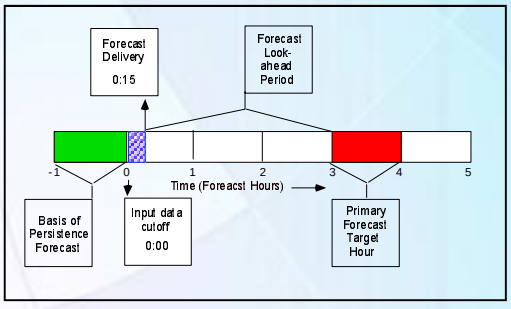
- PIRP data flow
- Forecast System Configuration

Data Flow for CA ISO PIRP Wind Generation Scheduling Process



eWind Configuration for Next Operating Hour PIRP Forecast

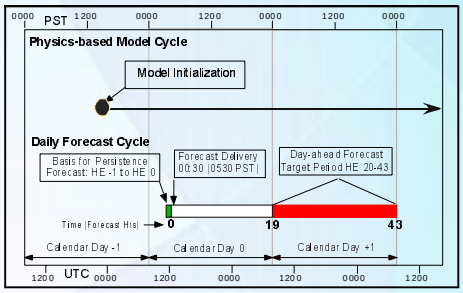
- Forecasts of hourly output and met variables for next 1–7 hrs
- 10-min data from PIR (via ISO) and output from regional physics-based models used as input
- Ensemble of statistical models with rolling 30-day training sample
- Bias correction algorithm is employed



Input data cut-off: on the hour
Typical forecast delivery: ~ hour +15 minutes

eWind Configuration for Next Day PIRP Forecast

- Forecasts of hourly output & met variables for next calendar day
- Ingests NWS raw and model output data
- Uses physics-based model run from previous 0000 UTC (4 PM PST) data time
- Model Output Statistics (MOS) with rolling 30-day training sample
- Can be extended to multiple days



Typical forecast delivery: ~ 5 AM each morning

Forecast Performance

Examples

Factors

Forecast Evaluation Issues

- Many options for performance statistics
  - Mean error (bias),
  - Mean absolute error (MAE)
  - Root mean square error (RMSE)
  - Skill score (% improvement over a reference forecast)
  - Correlation coefficient
  - Full error distributions
  - Many others .....
- Can tune system for a specific statistic (e.g. bias)
- What statistics are relevant to the forecast user?
  - *PIRP: Bias (to participant) and MAE (to CA ISO and some participants)*
  - Others: depends on user's "cost function"

Bias Correction Procedure

- "External" correction procedure is used
- Net Deviation (ND) is calculated from start of month:

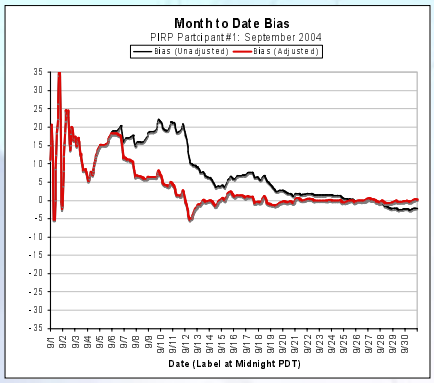
$$ND = \sum_{i=\text{first hr of month}}^{\text{current hour}} (F_i - O_i)$$

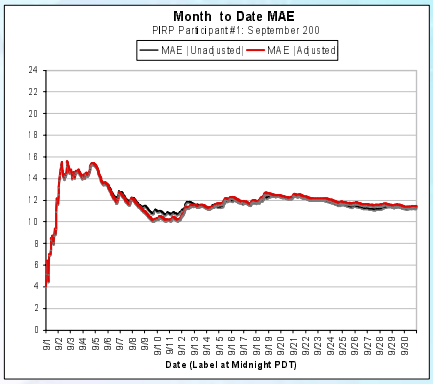
- Bias adjustment is calculated from ND for each forecast hour:

$$F_{biasadj} = F_0 - C * ND$$

- Adjustment phased in between 6th and 10th of month
  - C = 0 from 1st to 5th of month
  - C linearly increases to max value from 6th to 10th
  - C remains at max value from 11th to end of month
- C can be statistically determined from a rolling sample
  - Value that minimizes Bias and MAE

Typical Impact of Bias Correction Procedure



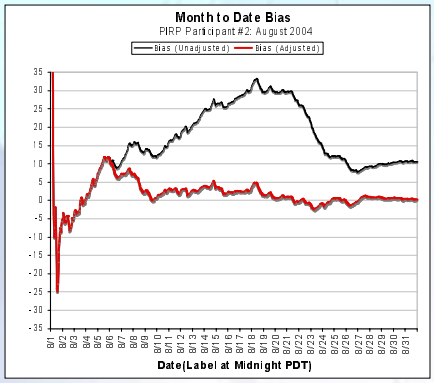


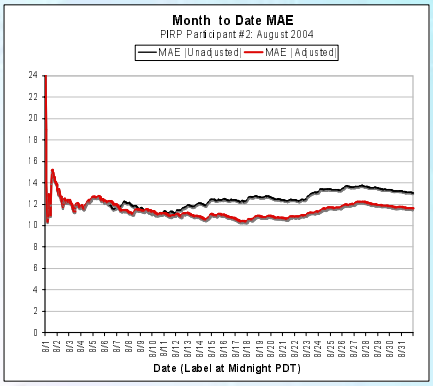
Unadjusted Bias: -2.12 %
Adjusted Bias: 0.23 %

Unadjusted MAE: 11.25%
Adjusted MAE: 11.42%

Bias is reduced and impact on MAE is small

Atypical Impact of Bias Correction Procedure



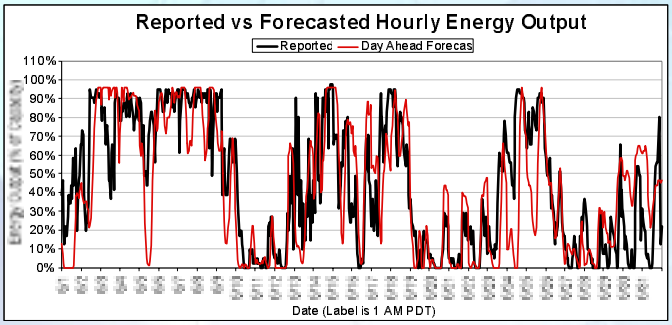


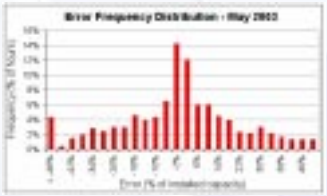
Unadjusted Bias: +10.61 %
Adjusted Bias: + 0.23 %

Unadjusted MAE: 13.04 %
Adjusted MAE: 11.59 %

Bias is reduced and MAE is significantly improved

Day-Ahead Forecast Example
San Gorgonio Pass, California

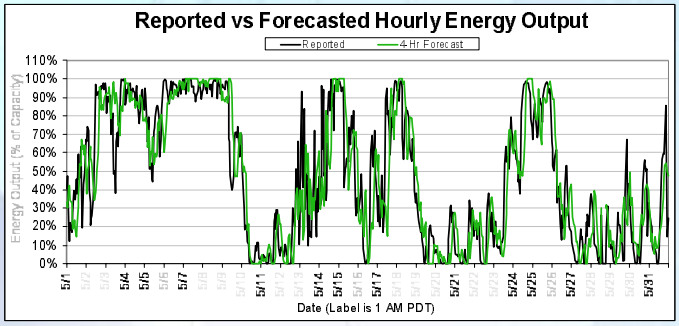


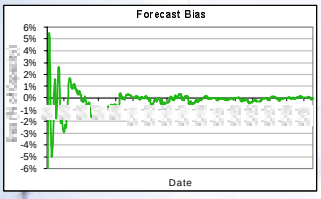


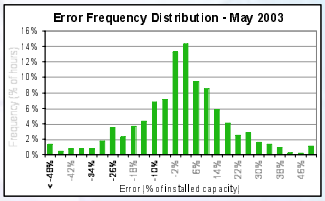
Performance Statistics

Bias	-2.18%
MAE	17.53%
Median AE	12.40%
Correlation	0.755
Skill (vs. Persistence)	53.5%

Next Operating Hour Forecast Example
San Gorgonio Pass, California







Performance Statistics

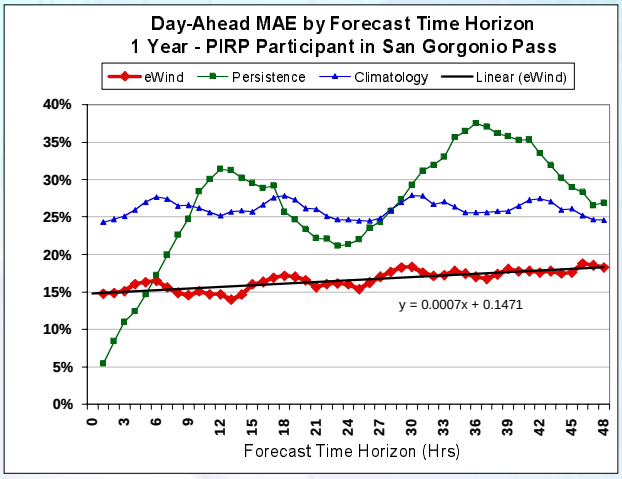
	Shown	Nearby
Bias	-0.04%	+0.06%
MAE	11.92%	10.35%
Median AE	9.02%	7.44%
Correlation	0.869	0.832
Skill (vs. Persistence)	18.3 %	28.3%

Forecast performance factors

- Forecast time horizon (especially for short-term)
- Position on the plant-scale power curve
- Amount of variability in the wind resource
- Quality of generation & met data from the plant
- Meteorological scales of variability (large scale, mesoscale etc.)
- Changes in seasons and weather regimes

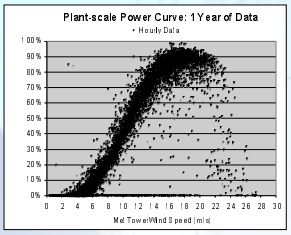
(Of course, methods and data types used by provider are factors but here we look at performance variability for a specific forecast system)

Forecast Performance Factors: Forecast "Look-Ahead" Period

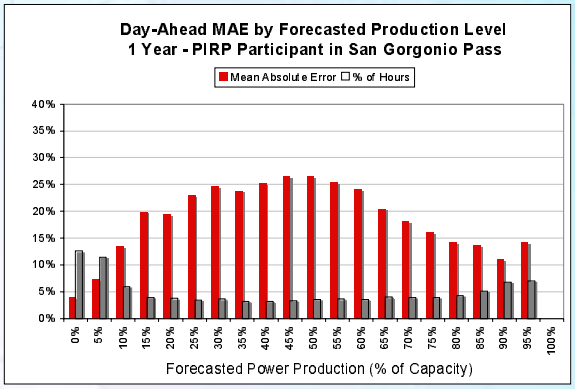


In this case: error growth rate is 0.07% per hr or 1.7% per day

Forecast Performance Factors: Position on the Power Curve

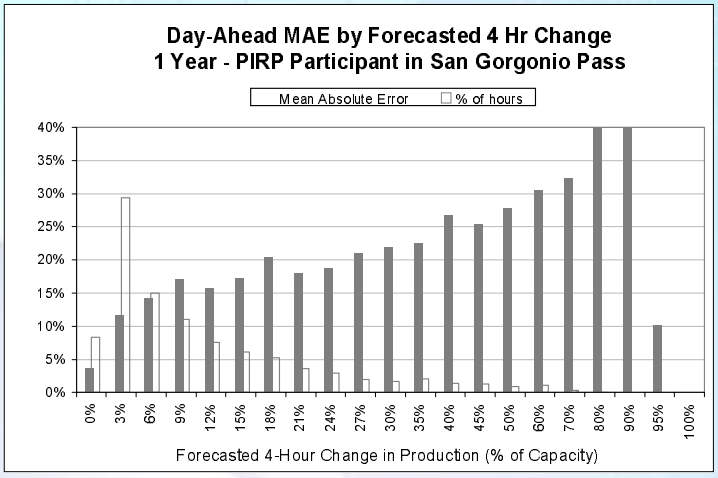


Sensitivity of power production prediction to wind speed forecast error depends on the location on the plant-scale power curve



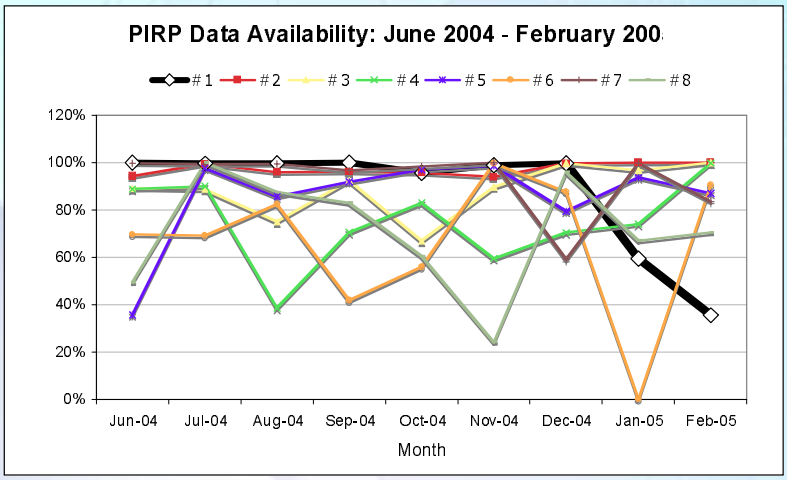
Annual MAE: 16.6%

Forecast Performance Factors: Forecasted Rate of Change

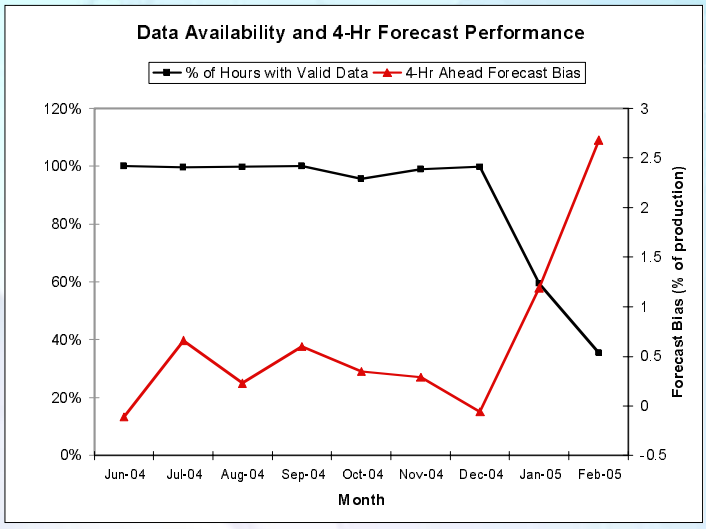


Annual MAE: 16.6%

Forecast Performance Factors: Data Availability



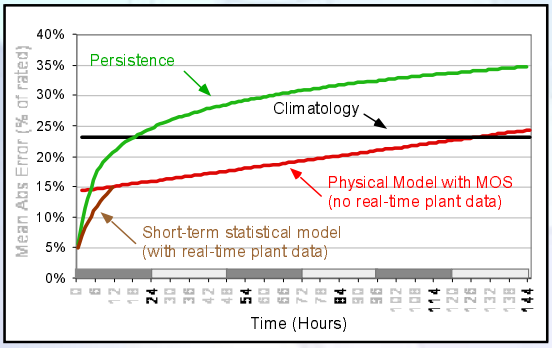
Impact of Data Availability



Missing/bad data has large impact on short-term forecast performance

Forecast Performance Summary

Wind Energy Forecast MAE by Forecast Hour
Composite of sites in California and central US



- Physical model forecasts adjusted by a MOS without real-time plant data outperform persistence and climatology from about 4 to 6 hours to approximately 5 days
- Short-term statistical models with real-time plant data outperform persistence and the "physical model-MOS without real-time plant data" procedure from 1 to 12 hours

Prospects for the Future

More Data
Improved Models

How can forecasts be improved?
(Top Three List)

- (3) Improved physics-based/statistical models
  - Improved physical modeling of sub-grid scale and surface processes
  - Better data assimilation techniques for physical models
  - Advances in learning theory: how to extract more relevant info from data
- (2) More effective use of models
  - Enabled by more computational power
  - Higher resolution, more frequent physics-based model runs
  - Ensemble forecasting
  - Wider use of more advanced statistical models and training methods
- (1) More/better data
  - Enhanced "off-site" data in the vicinity of wind plants
  - A leap in quality/quantity of satellite-based sensor data
    - Issue: how to most effectively use it?

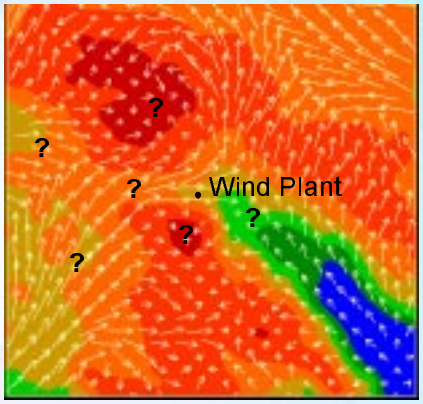
Off-site Data Collection: Met Towers

Popular idea:

- Install met tower "upstream" as an indicator of what is coming
- Can improve performance for **short-term (0-6 hr)** forecasts but…….

Issues

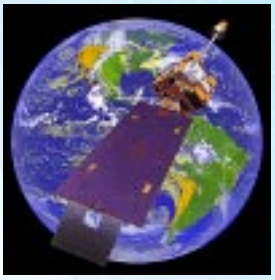
- Where do you put it (them)?
  - Physics-based model studies can help ($$)
  - Can use existing sites if available
- Performance in different regimes?
- What time scales?
- Installation and maintenance $$
- **PIRP ADVANTAGE: Can use data from other participants for PIRP forecasts**



Very Off-Site Data Collection:
Satellite-Based Sensors: The Next Generation

GIFTS: Geosynchronous Imaging Fourier Transform Spectrometer

- Temperature and moisture: vertical profiles
- Winds: tracking cloud and moisture features
- Less than 10 km between profiles
- 6 X vertical resolution of current profiles
- Profiles in cloud-free air only
- Initial version launched within 2 years



ATMS: Advanced Technology Microwave Sounder

- Temperature and moisture profiles through clouds
- Initial version will be deployed in 2006

Satellite-based wind profilers

- Active wind profiling systems have been proposed (LIDAR)
- None scheduled for deployment but discussions continue….

Summary

- PIRP forecasting status
  - Forecasting for "full PIRP" officially began on September 1, 2004
  - Currently forecasting for 9 resources (387 MW)
  - Forecast performance has been erratic
    - Data quality issues are a major factor
    - Limited data history from resources
  - Performance criteria met for many participant-months
- PIRP forecast performance will continue to improve
  - Quantity and quality of PIRP data will (hopefully) improve
  - Will take advantage of cross-use of PIRP data for next-hour forecasts
  - Results of CEC-sponsored research will be implemented
  - Improvements from internal forecast method R&D by AWS Truewind
  - New (non-PIRP) data sources will become available