# EVALUATION OF RECENT ADVANCEMENTS IN MACHINE LEARNING METHODS IN VERY SHORT-TERM TIME SERIES FORECASTS OF WIND POWER PRODUCTION

John W Zack, Ph.D.
UL Renewables
463 New Karner Rd
Albany, NY
jzack@awstruepower.com

# OBJECTIVE AND AGENDA

**Objective: Assess the impact of advanced machine learning tools on the performance of wind power forecasts**

- How much value over traditional statistical methods do they provide for the wind power forecasting applications?
- Are there specific aspects of wind power forecasting where they provide the most value?

- Background
- Experimental Design
- Application of Machine Learning to NWP Output ( a.k.a Model Output Statistics (MOS)
- Application of Machine Learning real-time measurement-based time series forecasts
- Summary & Conclusions

# TYPICAL MODES OF APPLICATION OF STATISTICAL METHODS IN WIND POWER FORECASTING

- **NWP Paradigm: Adjusting or Combining (Blending/Mixing) Output from NWP Models**
  - Optimal prediction of a target met variable from output of a single NWP
    - Often referred to as Model Output Statistics (MOS)
  - Optimal combinations of the forecasts from multiple NWP systems
  - Integrated combination of both of the above items
- **TS Paradigm: Development of Short-term Prediction Models from Time Series Data**
  - Time series models using near-real-time data from facilities (or facility aggregates)
  - Time series models with facility (or facility aggregate) data and off-facility data
- **Power production model (conversion of meteorological variable to power output)**

**This presentation will focus on deterministic forecasts but there are analogous approaches and issues for probabilistic forecasts**

Background

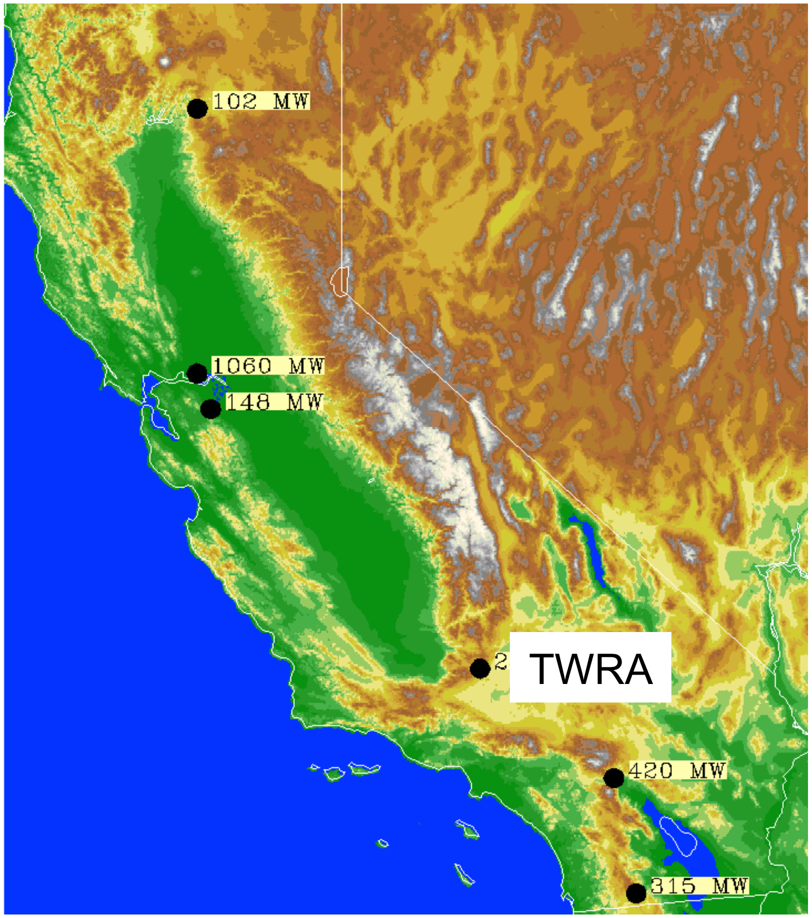
# WHAT IS MACHINE LEARNING?

- **Broadest definition: algorithms that find relationships in a dataset between inputs (X) and outputs (Y)….**
  - Linear Regression?
- **Typically reserved for advanced algorithms**
  - Families of Artificial Neural Network (ANN) algorithms
  - Families of Decision Tree algorithms
    - Random Forests
    - **Gradient Boosted Machine (GBM) including the XGBoost version**
  - Support Vector Regression (SVR)
  - Enhanced regression schemes
- **Could also include "data pre-conditioning" methods**
  - PCA
  - Clustering Algorithms (k-means, etc.)

# EXPERIMENTAL VENUE AND DATA

- **Venue: Tehachapi Wind Resource Area (TWRA) in California (USA)**
  - East side of a pass through the Tehachapi Mountains
  - Variations in channeled flow
- **Three Dominant Weather Regimes**
  - Winter storms from the Pacific Ocean
  - Warm season diurnal ocean-desert circulation
  - Late warm season monsoon
- **Wind Power Specifications**
  - Total of slightly over 3000 MW
  - Facilities with capacity of 2319 MW used in this project

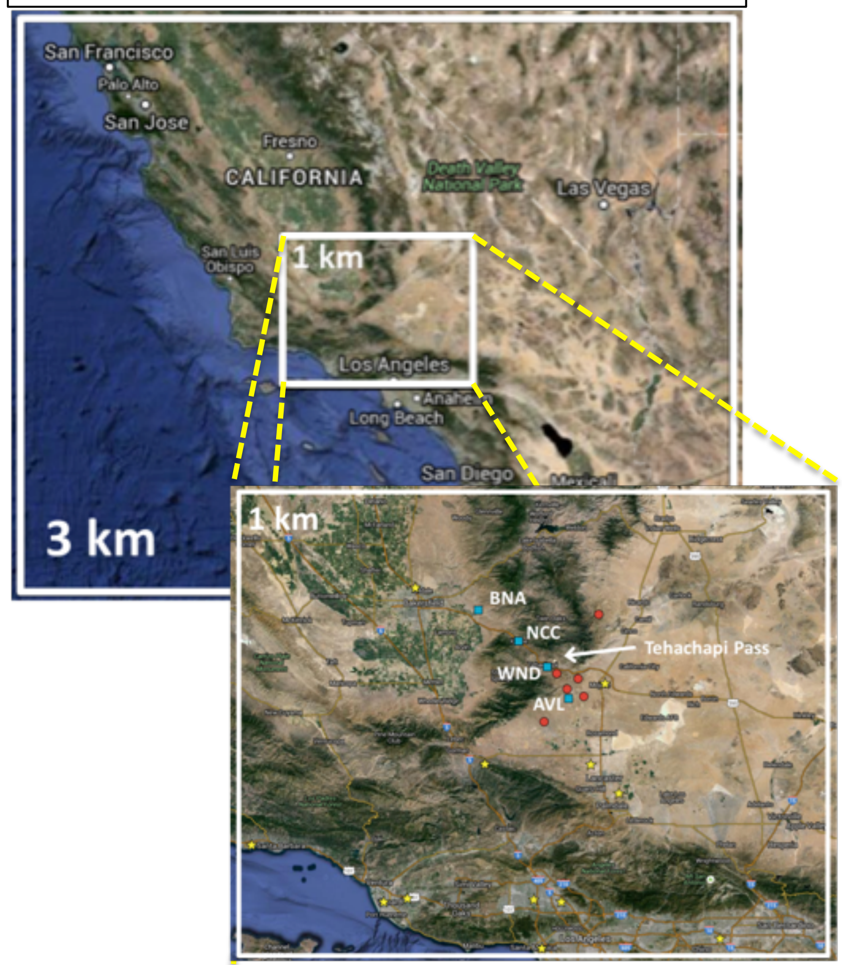
Location of Experimental Venue



# NWP EXPERIMENT SPECIFICATIONS: MODELS

- **Forecast Cycles: every 6 hours**
- **Forecast Length:15 hours**
- **NWP Models**
  - High Resolution Rapid Refresh (HRRR)
    - US national rapid update model
    - 3 km resolution
  - Custom 1-km WRF
    - In-house UL-AWST model
    - Physics optimized for wind prediction in the TWRA
- **Model Initialization**
  - HRRR: standard HRRR data assimilation
  - WRF: background state from HRRR 1-hr forecast with additional assimilation of local sensor data

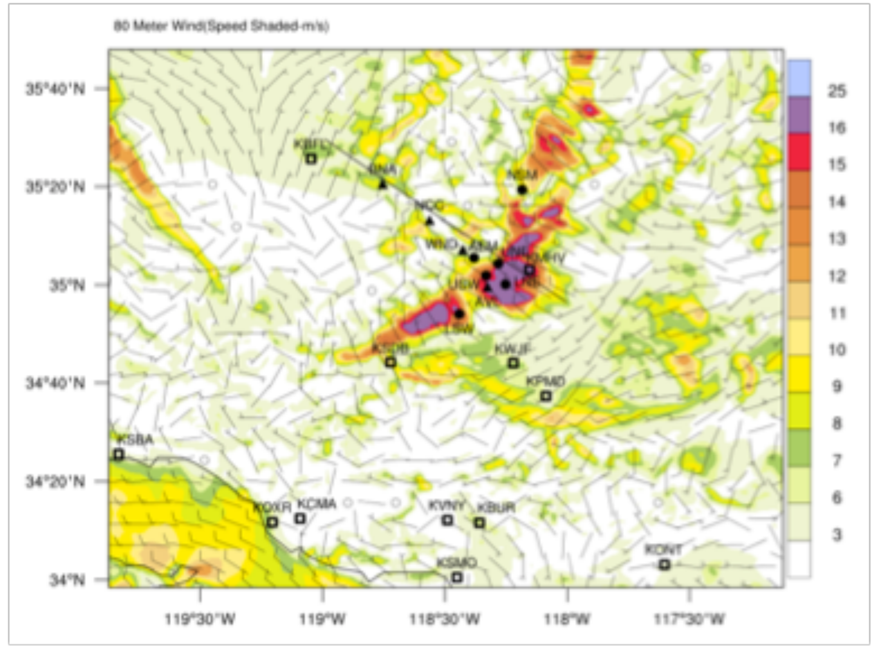
Domain of Custom WRF Model



# NWP EXPERIMENT SPECIFICATIONS: MOS

- **Statistical Methods**
  - 26 algorithms available in the Python-based Scikit-learn package
  - Employed default values for all algorithm parameters
- **Predictors**
  - Same 9 NWP variables used for all statistical methods
- **Predictand (target variable)**
  - 15-min average power production for each of 17 facilities (total of 2319 MW)
- **Statistical Model Training**
  - Rolling 11-month sample excluding forecast month
  - All look-ahead intervals (0-15 hrs) trained in one sample

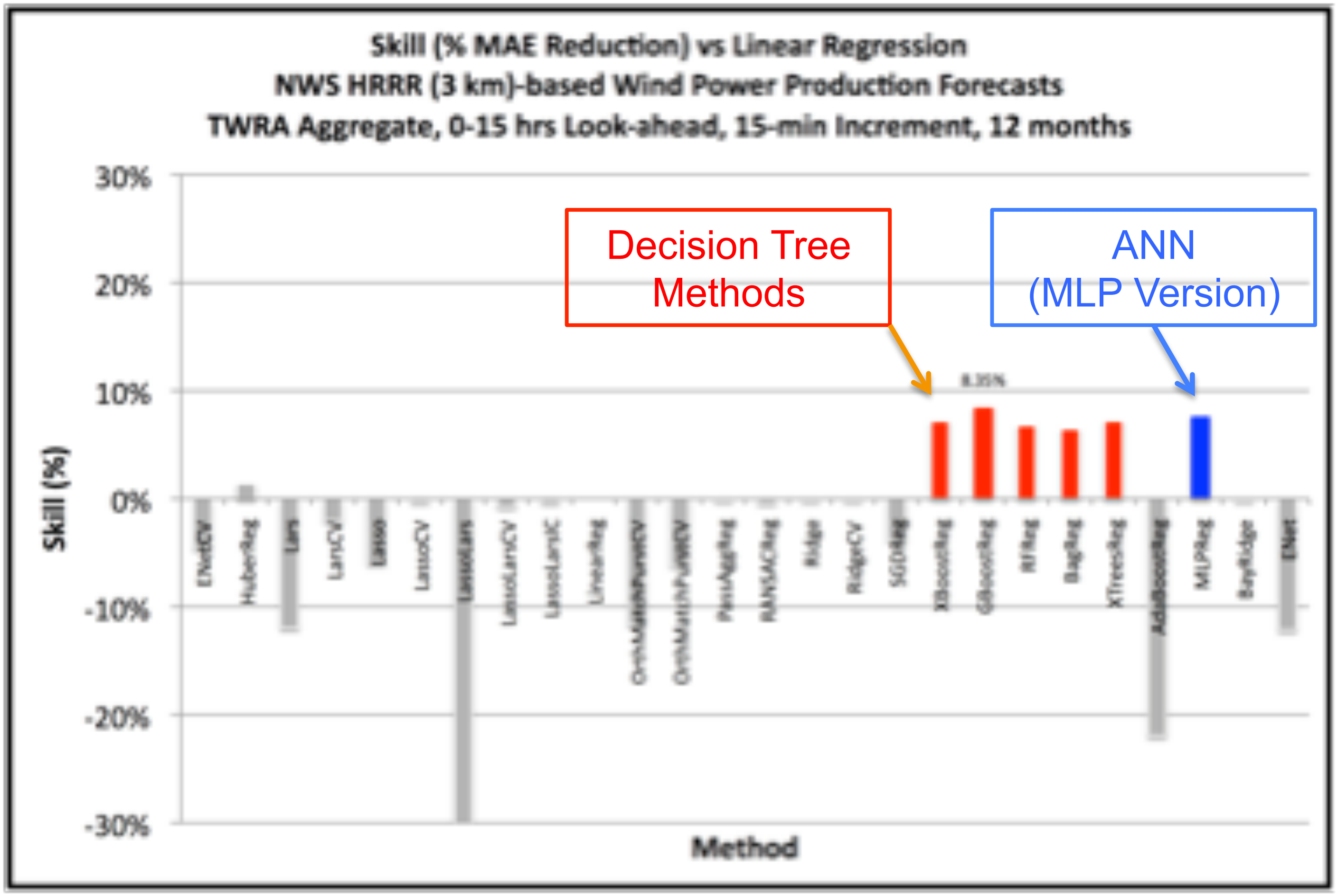
Example of a 80-m wind speed forecast from the custom 1-km WRF model



NWP Paradigm

# RELATIVE PERFORMANCE OF STATISTICAL METHODS: HRRR

0-15 Hour Wind Power Forecasts for TWRA Aggregate
Evaluation Period: Oct 2015-Sept 2016





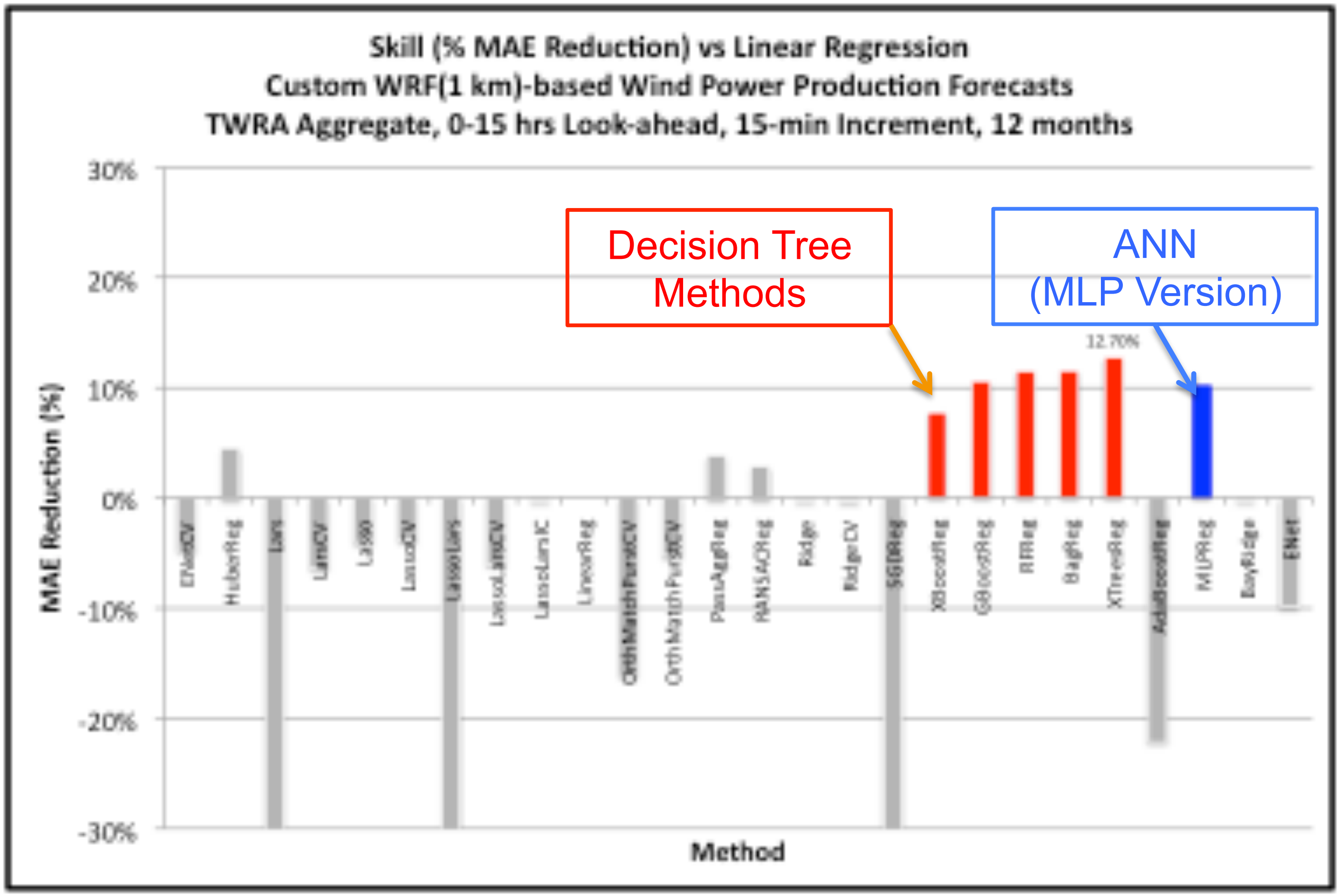



Tables, graphs and charts should be designed with neutral colors using accent colors sparingly. Use UL colors whenever possible.
Table borders should be 1pt.

NWP Paradigm

# RELATIVE PERFORMANCE OF STATISTICAL METHODS: WRF

0-15 Hour Wind Power Forecasts in 15-min Increments
Evaluation Period: Oct 2015-Sept 2016





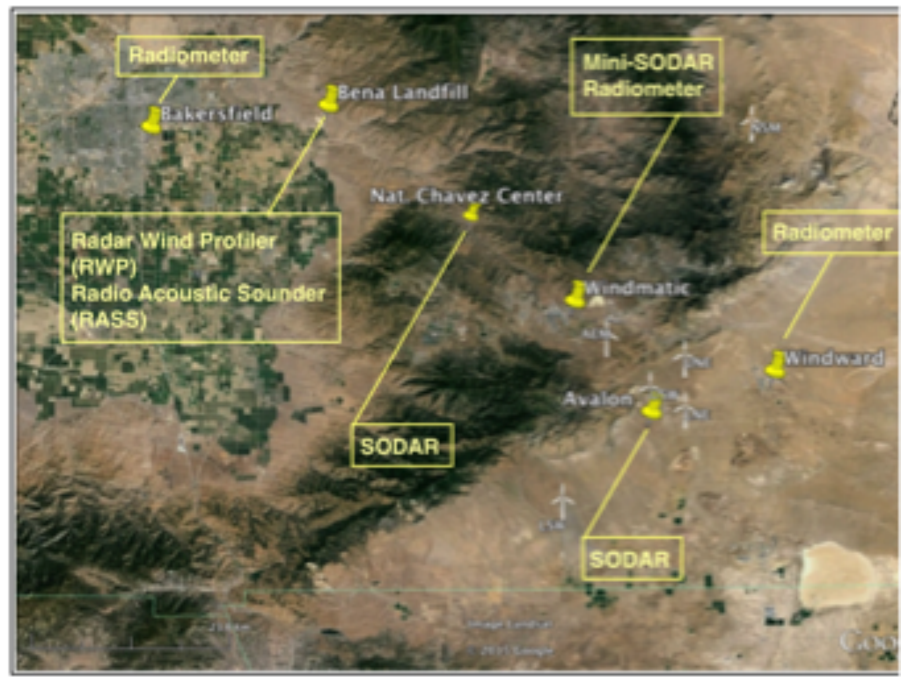



Tables, graphs and charts should be designed with neutral colors using accent colors sparingly. Use UL colors whenever possible.
Table borders should be 1pt.

# TIME SERIES EXPERIMENT SPECIFICATIONS

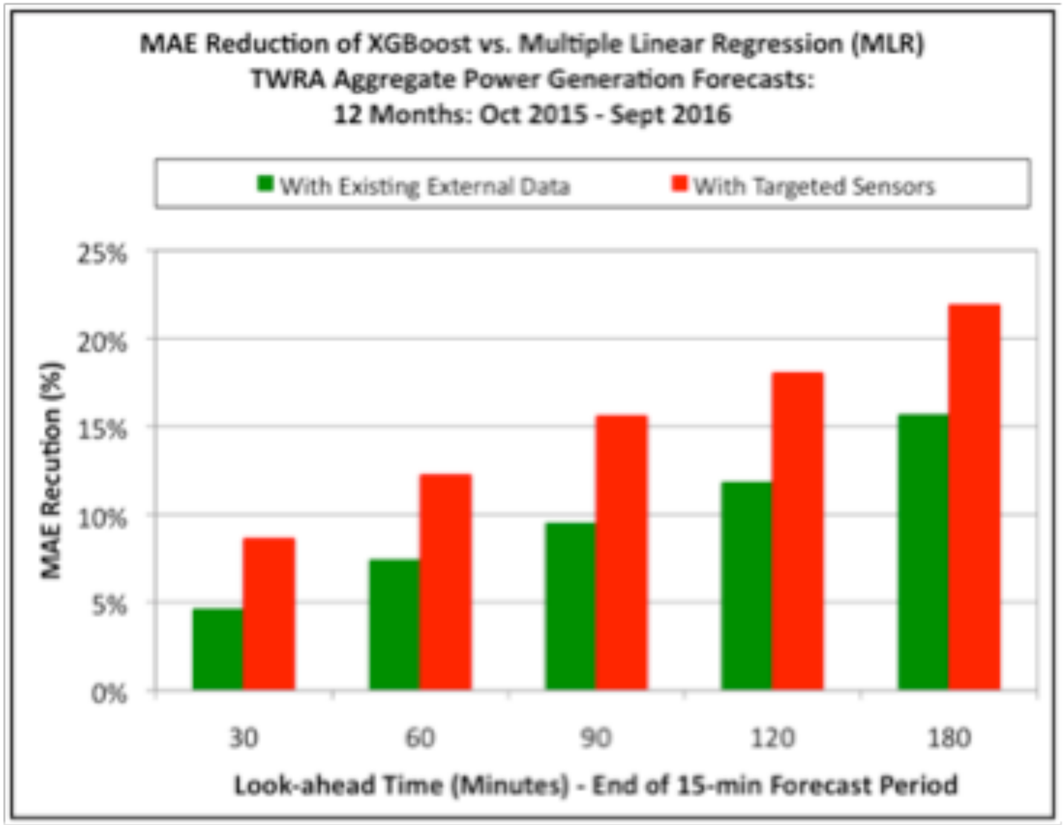
- **Statistical Methods**
  - Two algorithms: XGBoost (Decision Tree method) and Linear Regression
  - Default values for algorithm parameters
- **Predictors (time-lagged)**
  - 122 variables from 3 source classes: (1) facility data, (2) pre- existing off-facility data and (3) targeted sensors
- **Predictand (target variable)**
  - 15-min average power production for <u>each</u> of 17 facilities (total of 2319 MW)
- **Statistical Model Training**
  - Rolling 11-month sample excluding forecast month
  - Each 15-minute look-ahead period screened and trained separately

Network of targeted sensors configured from the results of a forecast sensitivity analysis and deployed during a 2.5-year experimental project



# PERFORMANCE OF XGBOOST VS LINEAR REGRESSION

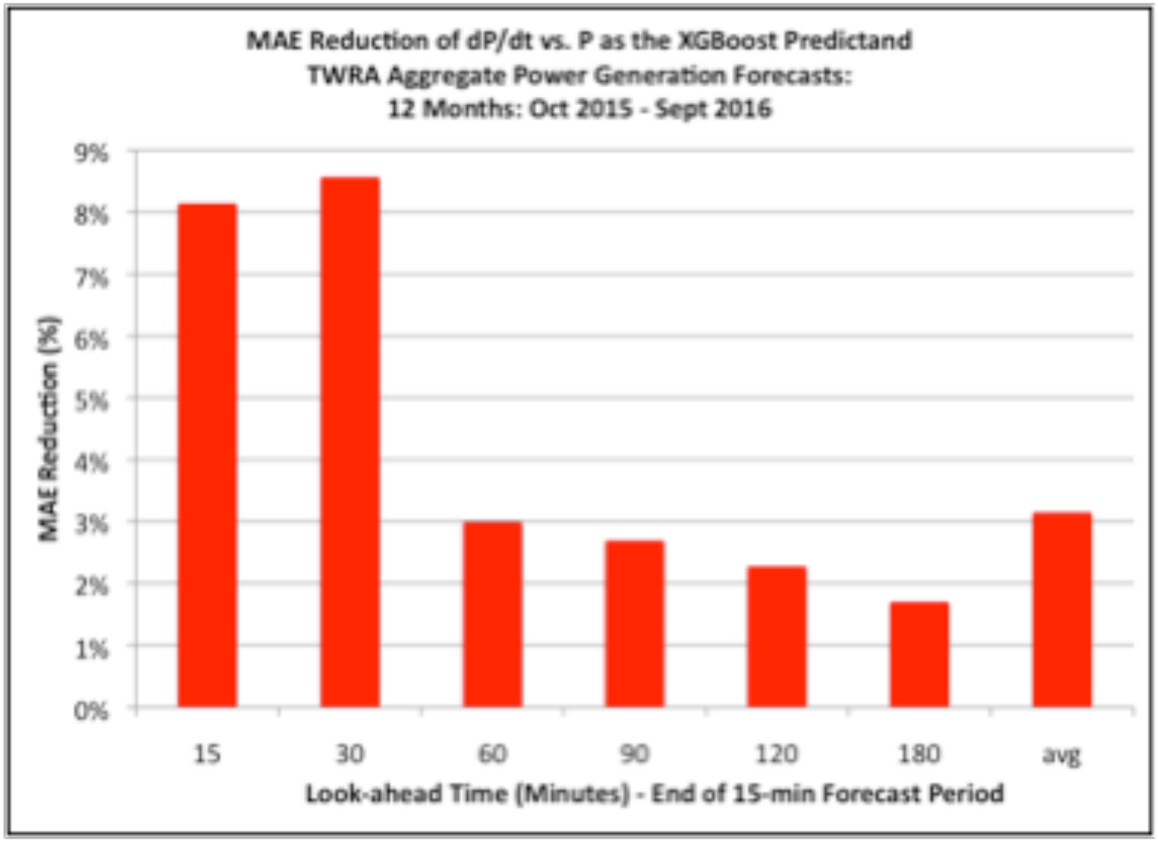
0-3 Hour Time series Wind Power Forecasts
2 Input Data Scenarios




- XGBoost (advanced decision tree algorithm) outperforms Linear Regression for all look-ahead time periods for both input datasets
- Use of targeted local area data reduces error with XGBoost or Linear Regression
- **XGBoost yields a greater improvement over Linear Regression when using the larger and more complex input dataset**

# RELATIVE PERFORMANCE OF PREDICTAND FORMULATION

0-3 Hour Time series Wind Power Forecasts
Evaluation Period: Oct 2015-Sept 2016




- Two forms of the forecast target variable were used in XGBoost:
    - Power (P) (Direct prediction)
    - Change in power from initial time (dP/dt) (Indirect prediction)
- **Use of the change in power as the XGBoost predictand resulted in lower MAE for all 0-3 hour look-ahead times**
    - Largest impact for the 0-30 minute time frame

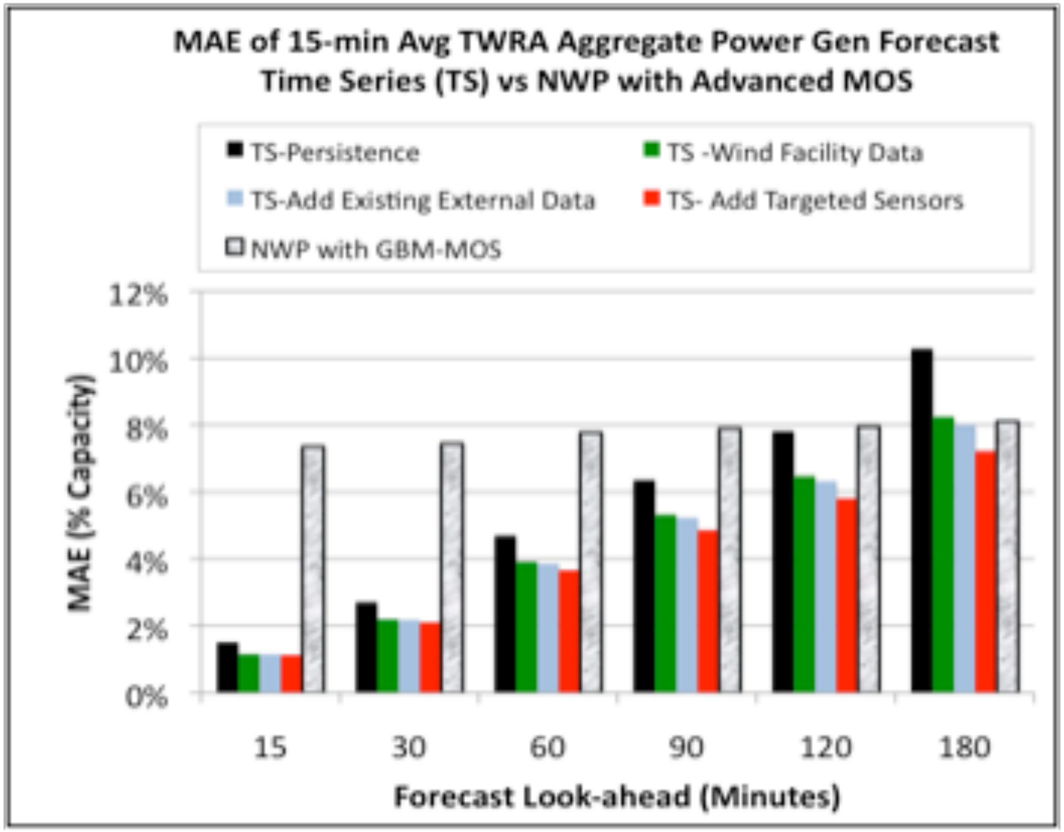




Tables, graphs and charts should be designed with neutral colors using accent colors sparingly. Use UL colors whenever possible.
Table borders should be 1pt.

# SUMMARY: RELATIVE PERFORMANCE OF NWP+MOS VS TS

0-3 Hour Wind Power Forecasts
Evaluation Period: Oct 2015-Sept 2016




- XGBoost-based time series methods & persistence outperform a high resolution NWP ensemble + advanced MOS for 0-2 hour look-ahead periods
  - Note: No time series info used in NWP-MOS
- **After 2-hours the relative performance depends on which off-site data is used in the time series model**
  - NWP+MOS performs better than persistence after 2 hrs
  - TS with targeted off-site data better than NWP+MOS at 3 hrs





Tables, graphs and charts should be designed with neutral colors using accent colors sparingly. Use UL colors whenever possible.
Table borders should be 1pt.

# CONCLUSIONS

- **Application of Machine Learning to NWP-MOS (0-15 hour look-ahead):**
  - Decision Tree and Neural Network algorithms were the best performers
  - MAE reduction by advanced methods over Linear Regression was 8% to 13% depending on specific statistical method and the underlying NWP model
  - MAE reduction by advanced methods was greater for custom 1-km WRF model vs 3-km HRRR (more structure in the data? – amount of input data was the same)
- **Application of Machine Learning to Time Series Prediction (0-3 hour look-ahead)**
  - Advanced machine learning method (XGBoost) outperformed based on (MAE reduction) Linear Regression for all look-ahead periods and predictor datasets
  - XGBoost achieved a substantially greater MAE reduction over Linear Regression with more predictor variables from additional sources (a bigger and more diverse dataset)
- **Comparison of NWP-MOS to Time Series Performance (0-3 hour look-ahead)**
  - All TS methods and persistence outperformed NWP+MOS for 0-2 hrs
  - For 2-3 hrs, TS vs NWP+MOS winner depended on input dataset and method for TS