NREL
Transforming ENERGY

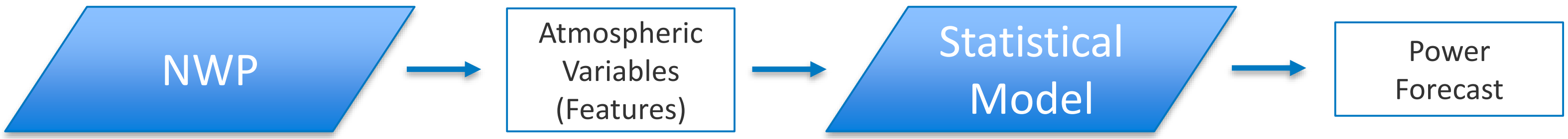
# *Statistical Modelling of Wind Plant Power Production: Contrasting Different Machine Learning Algorithms and Atmospheric Features*

Mike Optis, Senior Scientist
Jordan Perr-Sauer, Data Scientist

National Renewable Energy Laboratory
AMS 2019
January 10, 2018

# State of Literature

- Machine learning (ML) is now a standard step in wind power forecasting




- Extensive literature with a large range of datasets, algorithms, input variables, timescales, forecast horizons, and error metrics

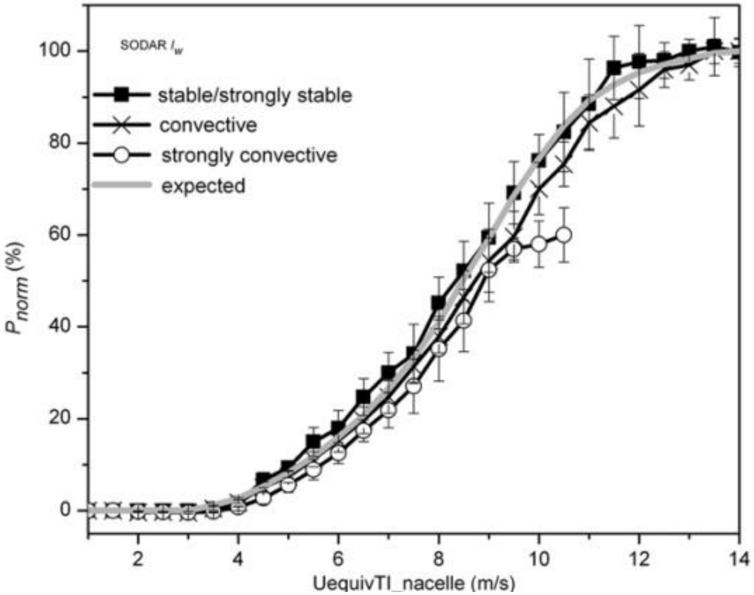
- **No clear consensus on best practices, no standardized methods to assess**

| Algorithms | Counts |
|---|---|
| Neural network | 14 |
| Neuro-fuzzy inference system | 6 |
| K-means clustering | 5 |
| Support vector machine | 4 |
| Regression tree approaches | 4 |
| Gaussian processes | 2 |
| Particle swarm optimization | 2 |
| Principle component analysis | 1 |
| Markov chain approaches | 1 |
| Bayesian learning | 1 |

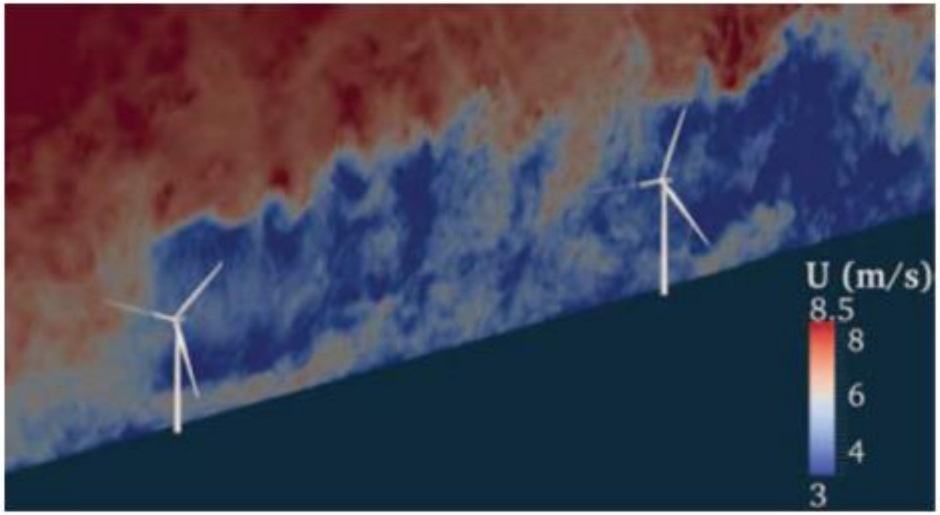
| Atmospheric Input | Counts |
|---|---|
| Wind speed | 22 |
| Wind direction | 10 |
| Temperature | 8 |
| Air pressure | 4 |
| Humidity | 3 |
| Air density | 1 |
| Potential temperature gradient | 1 |

# Turbulence and Stability

- Significant impacts on both turbine and plant power
  - Laminar vs. turbulent flow
  - Effect of wind shear
  - Turbine wake propagation and dissipation
- Almost never considered in ML models for wind power forecasting




Wharton and Lundquist (2012). Env. Res. Let., **7** (1)

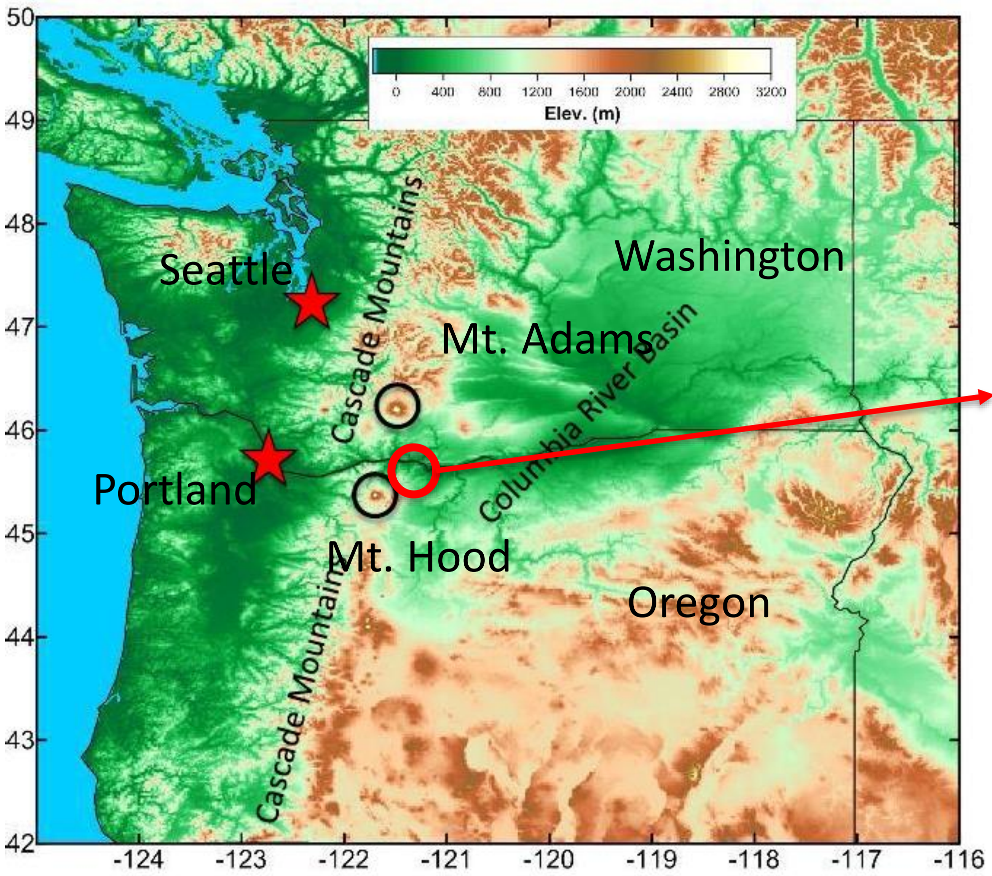



Churchfield et. al (2012). Journal of Turbulence, **13**

# Goal and Experiment Setup

**Can the use of turbulence and stability inputs improve ML predictions?**

**What is uncertainty associated with different model choices?**




| | |
|---|---|
| **Atmospheric Data** | WFIP2 Physics Site |
| **Wind power data** | Nearby wind farm |
| **Time period** | 2016-07 to 2017-03 |
| **Time resolution** | Hourly |

# Algorithms and Inputs

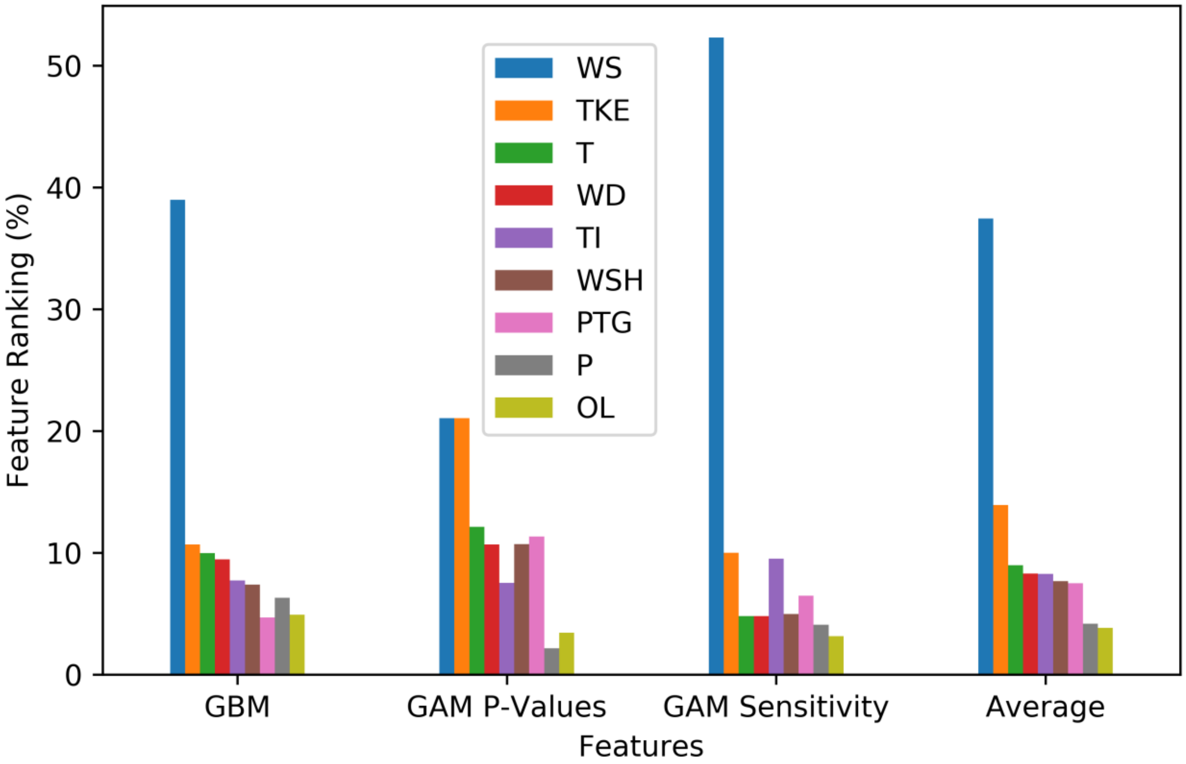
| Learning Algorithms | Atmospheric Features |
| --- | --- |
| Neural network (ANN)*<br>Generalized additive model (GAM)<br>Gradient boosting model (GBM)<br>Extra trees regression (ERT)<br>Support vector machine (SVM) | Wind speed (80m)<br>Wind direction (80m)<br>Air density (80m)<br>Temperature (80m)<br>Turbulence Intensity (80m)<br>Wind speed shear (80-50m)<br>Potential temperature gradient (80-17m)<br>Turbulent kinetic energy (80m)<br>Obukhov length (80m) |

- Hyperparameters for each algorithm optimized through cross-validated grid search over parameter space
- K-fold cross-validation (95% train, 5% test)
  - ~1 week test periods
  - Common time interval for curtailment loss estimation

* Simple multi-layer perceptron model with back propagation

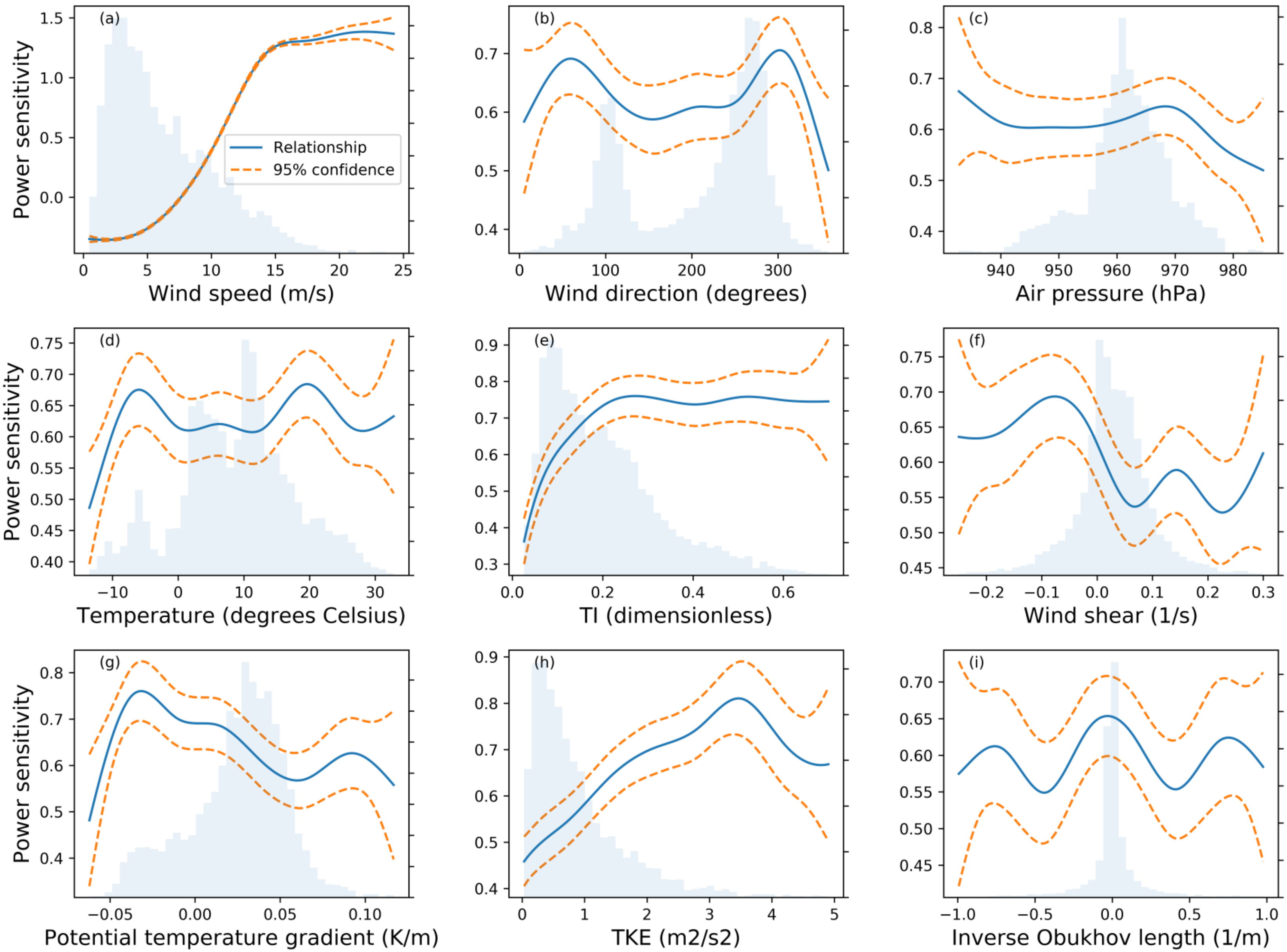
# Feature Importance

- 3 methods to assess feature importance
- Turbulence and stability all statistically relevant
- **TKE is second most significant after wind speed**

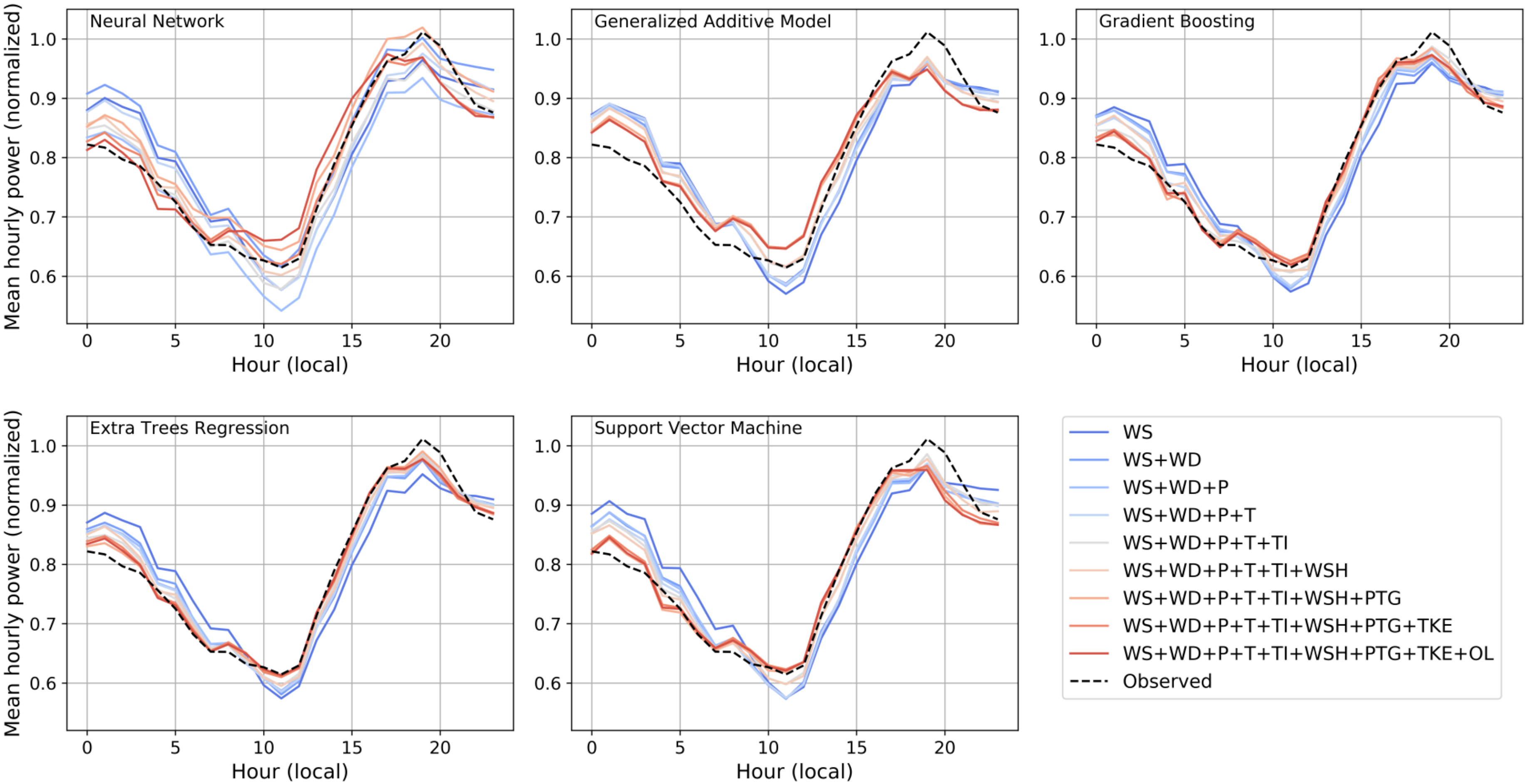



WS - wind speed
TKE - turbulent kinetic energy
T - temperature
WD - wind direction
TI - turbulence intensity
WSH - wind shear
PTG - potential temperature gradient
P - pressure
OL - Obukhov length

# Power Sensitivity to Inputs

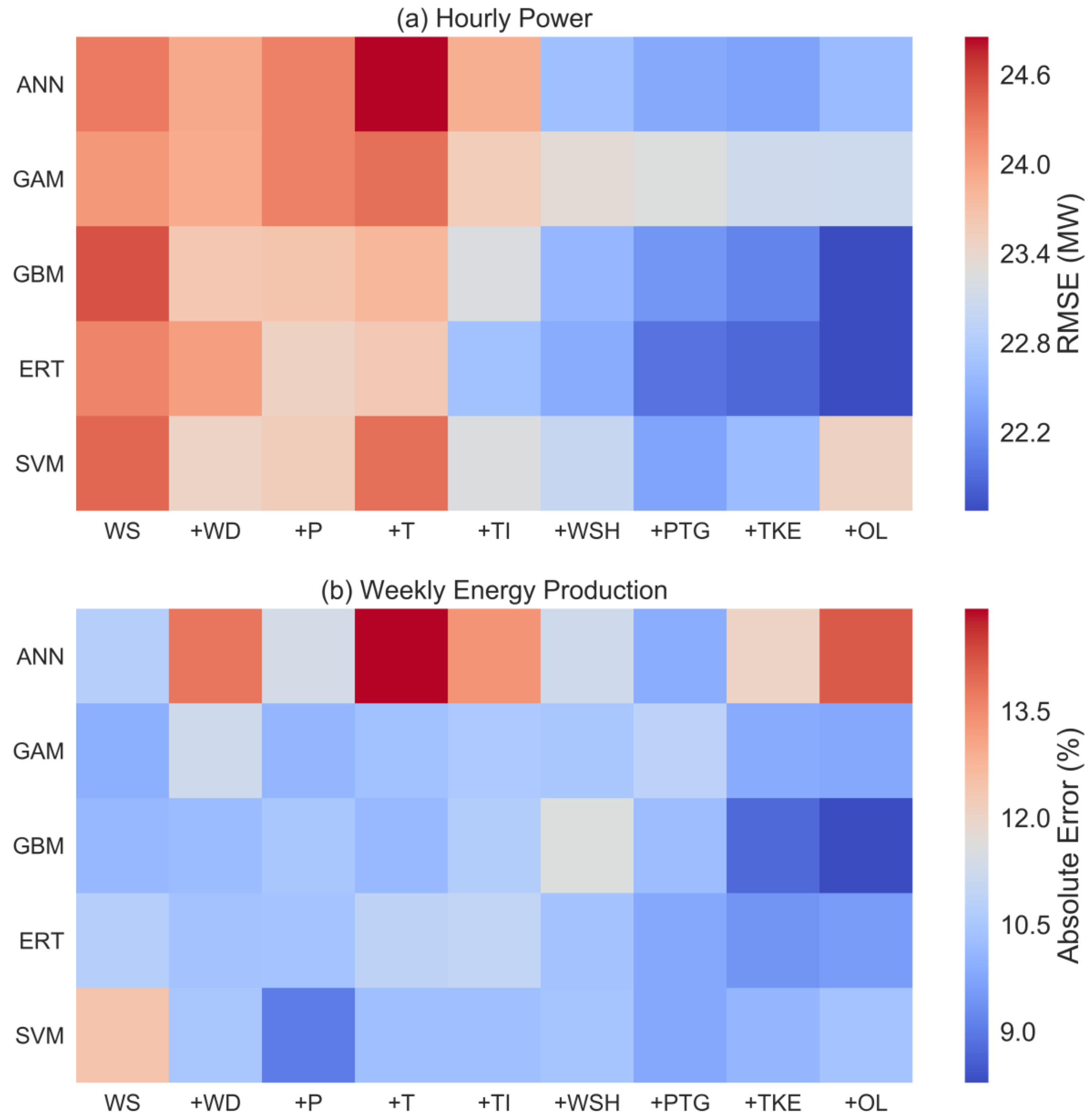


# Diurnal Cycle



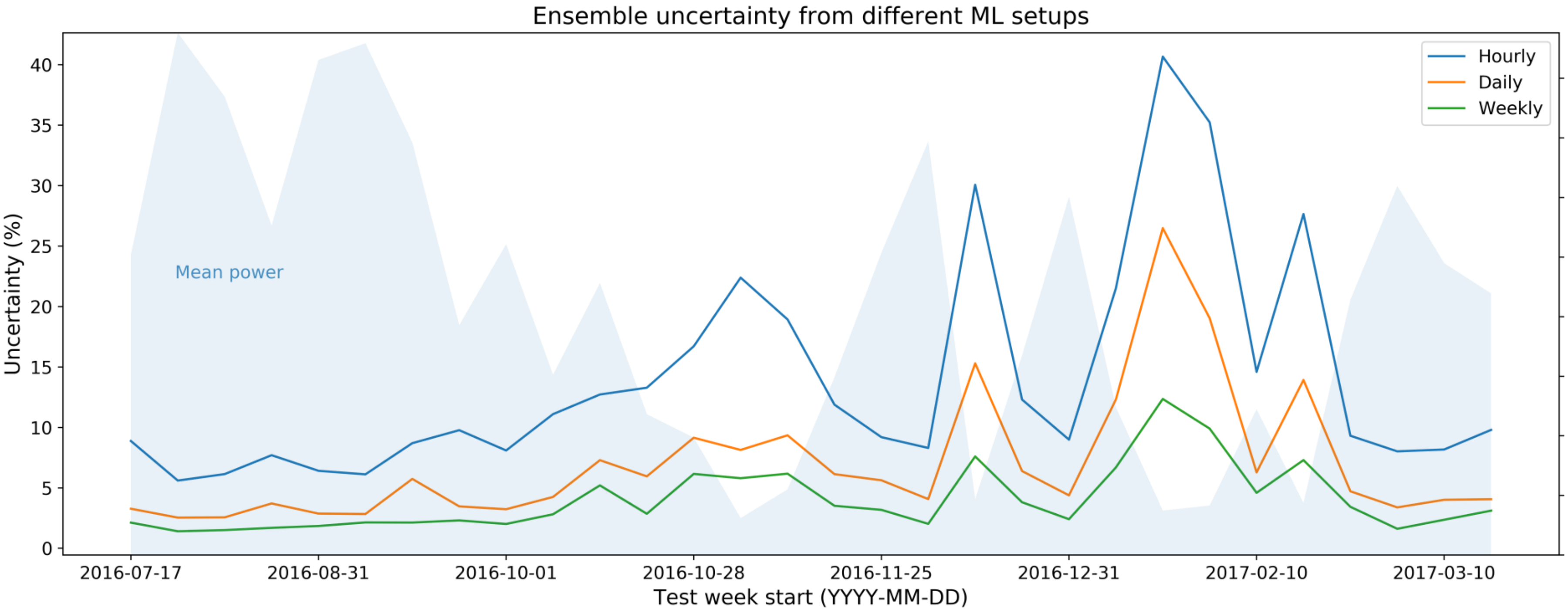

- Better diurnal modeling with more features
- Regression-tree approaches perform relatively well

# Model Accuracy on Different Timescales




- Improvement in hourly performance once turbulence/stratification is considered
- Not much improvement when predicting sum of energy over 1 week
- Model biases tend to average out over longer periods

# Uncertainty from ML Model Choice



# Remarks

- **Atmospheric turbulence and stability features should be considered**
- **Significant uncertainty** depending on inputs and algorithm selected
- No consensus on what model setup is best, although:
  - Regression tree-based approaches worked well here (and are simple)
  - More input variables the better
- Do we need complex neural networks for most wind energy problems?
- **Yet another non-standardized analysis…**

# Next Steps

**How can we standardize ML model assessment for wind power modelling?**

1. Standard datasets
    - Several wind farms in different wind regimes
    - Observed and NWP-modelled atmospheric variables
    - Pre-processed and ready for ingestion by ML frameworks
    - **Challenges with proprietary data**

2. Standard assessment methods
    - Hyperparameter optimization
    - Cross-validation
    - Error metrics

Optis, M. and Perr-Sauer, J. (2018). Renewable and Sustainable Energy Reviews, under review

# Thank You

**www.nrel.gov**

This work was authored by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by the U.S. Department of Energy, within the Atmosphere to Electrons research program. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

NREL Transforming ENERGY

# Annex: Hyperparameter Ranges

| Algorithm | Optimized Hyperparameters | Possible Values |
|---|---|---|
| GAM | Number of splines | 5-40 |
| NN | Activation function | logistic, tanh, relu |
| | Hidden layer size | (100), (100, 100), (100, 50), (100, 100, 100), (100, 100, 50), (100, 100, 25), (100, 50, 25, 9) |
| GBM | Maximum depth | 2, 4, 8, 12, 16, 20 |
| | Maximum features | 1 to number of total features |
| | Minimum sample split | 2-11 |
| | Minimum samples per leaf | 1-11 |
| | Number of estimators | 10 to 400 in increments of 40 |
| ERT | Maximum depth | 4, 8, 12, 16, 20 |
| | Maximum features | 1 to number of total features |
| | Minimum sample split | 2-11 |
| | Minimum samples per leaf | 1-11 |
| | Number of estimators | 10 to 400 in increments of 40 |
| SVM | Penalty parameter of the error term | 0.1, 1, 10, 50, 100 |
| | Kernel coefficient | 0.01, 0.1, 1, 10 |
| | Kernel | Polynomial, RBF, sigmoid |