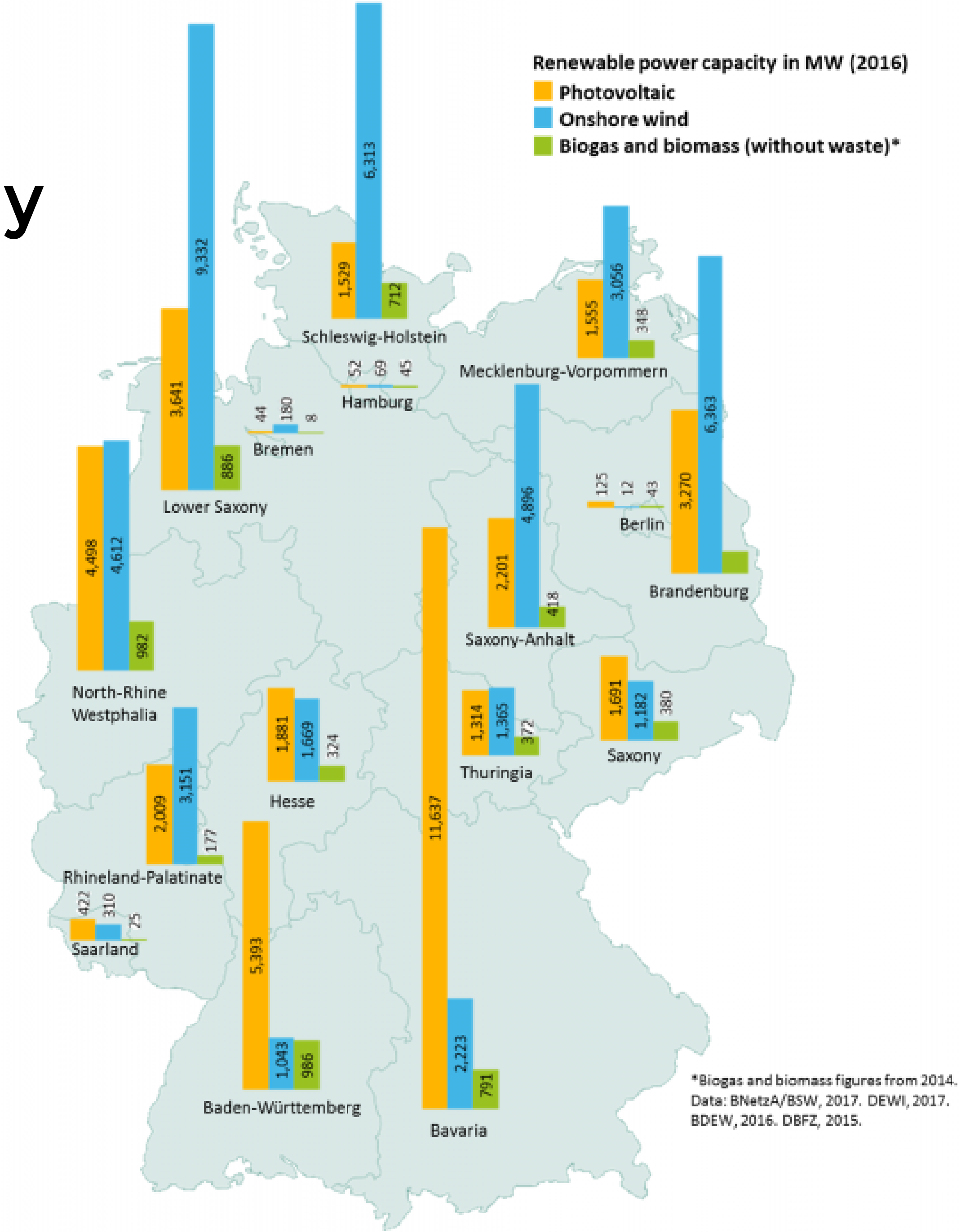
Short Term Wind Power Forecasting

Andrew Nachtigal

Data Science Community Day July 3, 2018

Renewable Power Transition in Germany

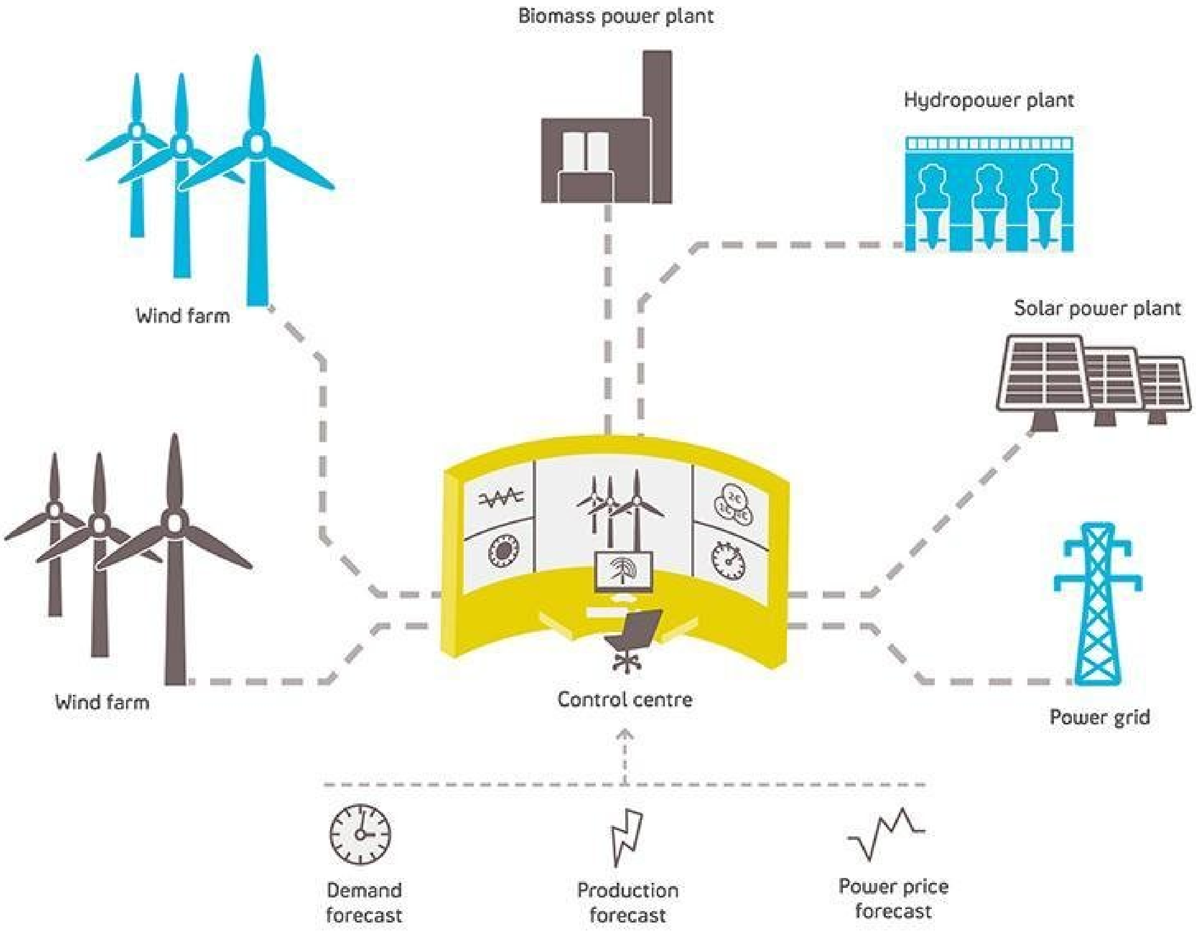
- Renewables now account for 36% of power consumption (2017)
- Wind is the fastest growing renewable power resource
- Changing energy balance: windy north, sunny south



Why forecast power production?

- Grid balancing: match energy Supply & Demand

- Accurate Forecast allow more efficient use of renewable energy



What is short term wind power forecasting?

- Predict electrical power generation from wind resources

- Short term: 1 - 48 hours
  "day-ahead"

Wind Farm Data

Wind Farm Data

- Power generation and weather forecasts for 45 wind farms

- 2 consecutive years of hourly measurements

  Power Generation

  Wind Speed

  Wind Direction

  Air Pressure

  Humidity

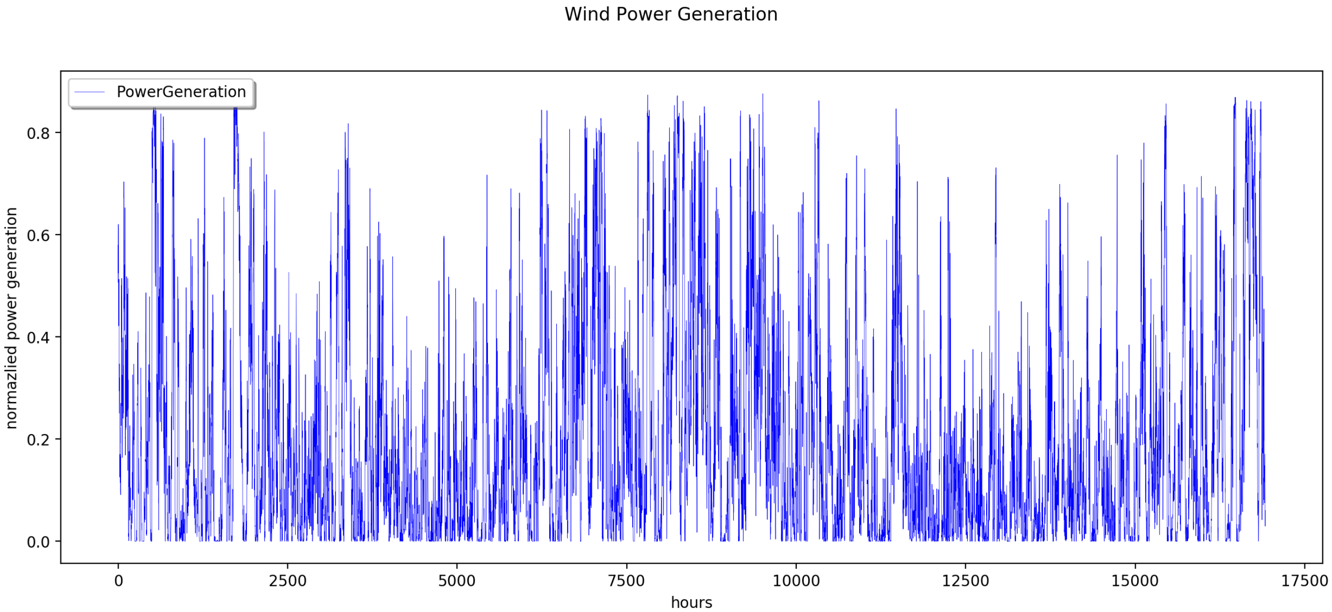
  Temperature

Weather Forecast Data: European Center for Medium Range Weather Forecasts

Wind Power

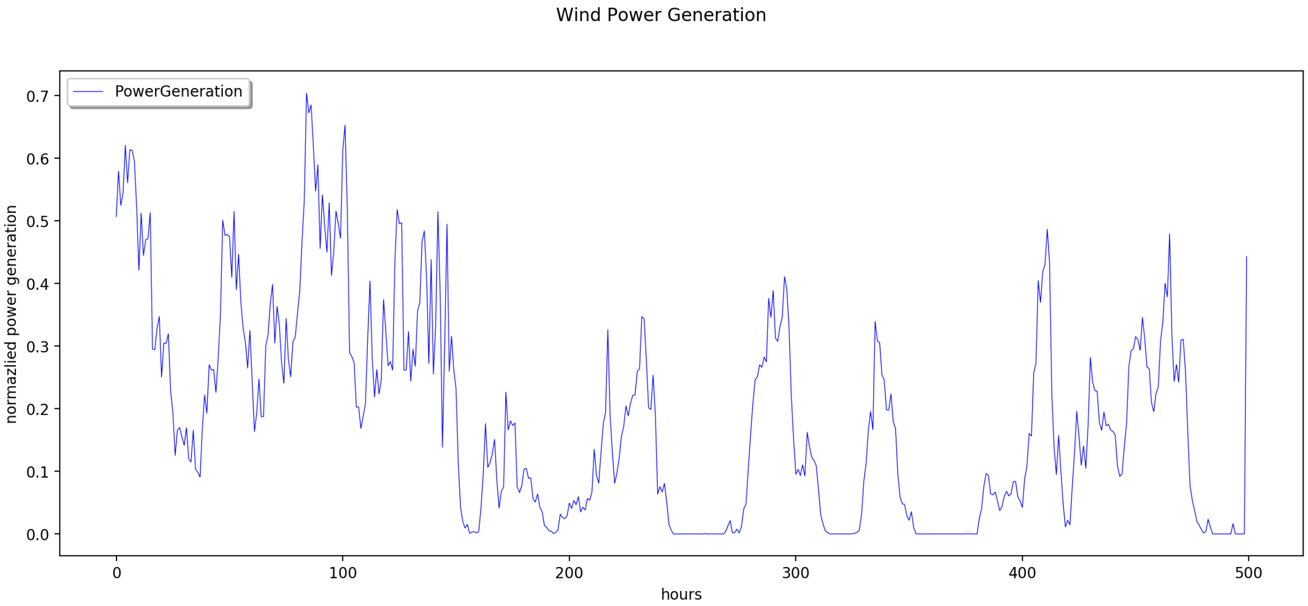
Time Series

All Hours



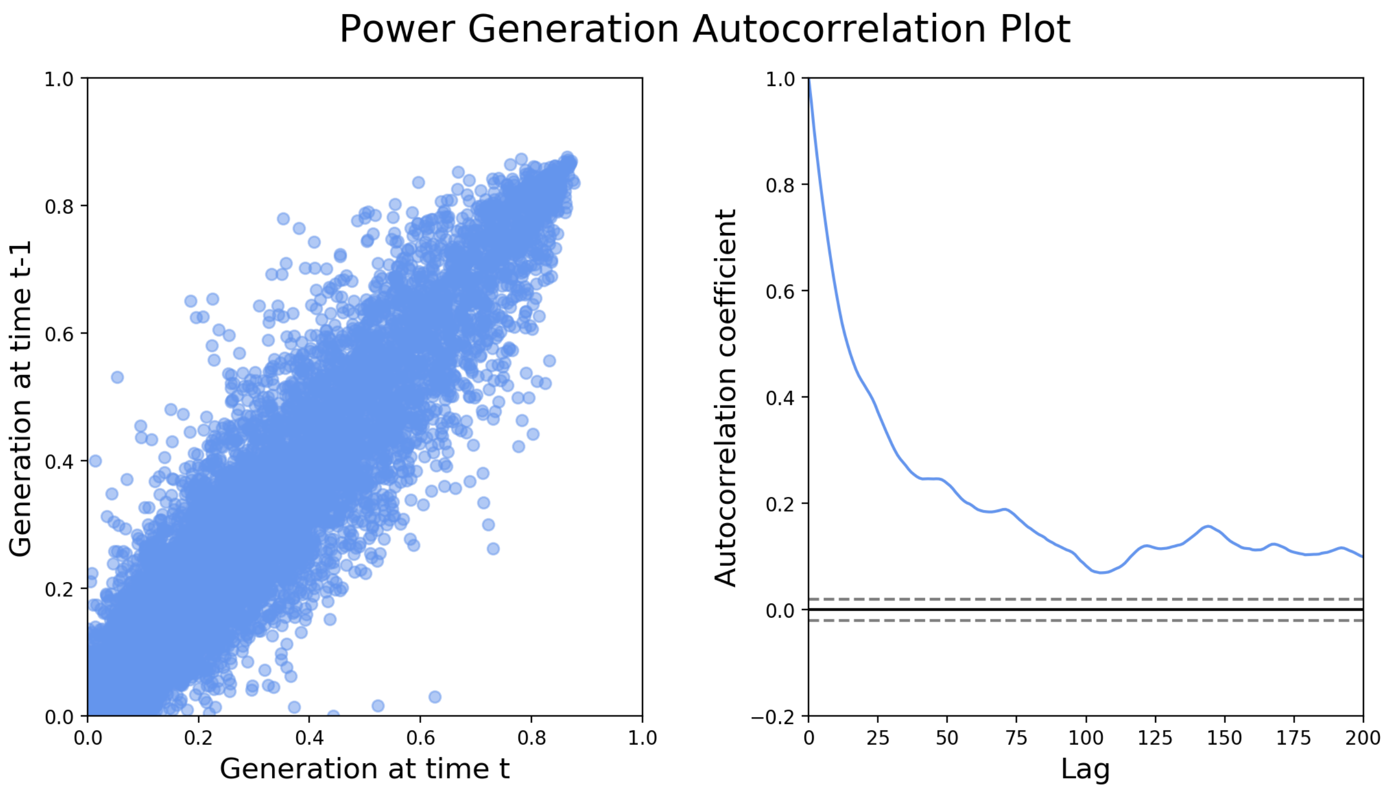
Time Series

Hours 0 - 500



Wind Power

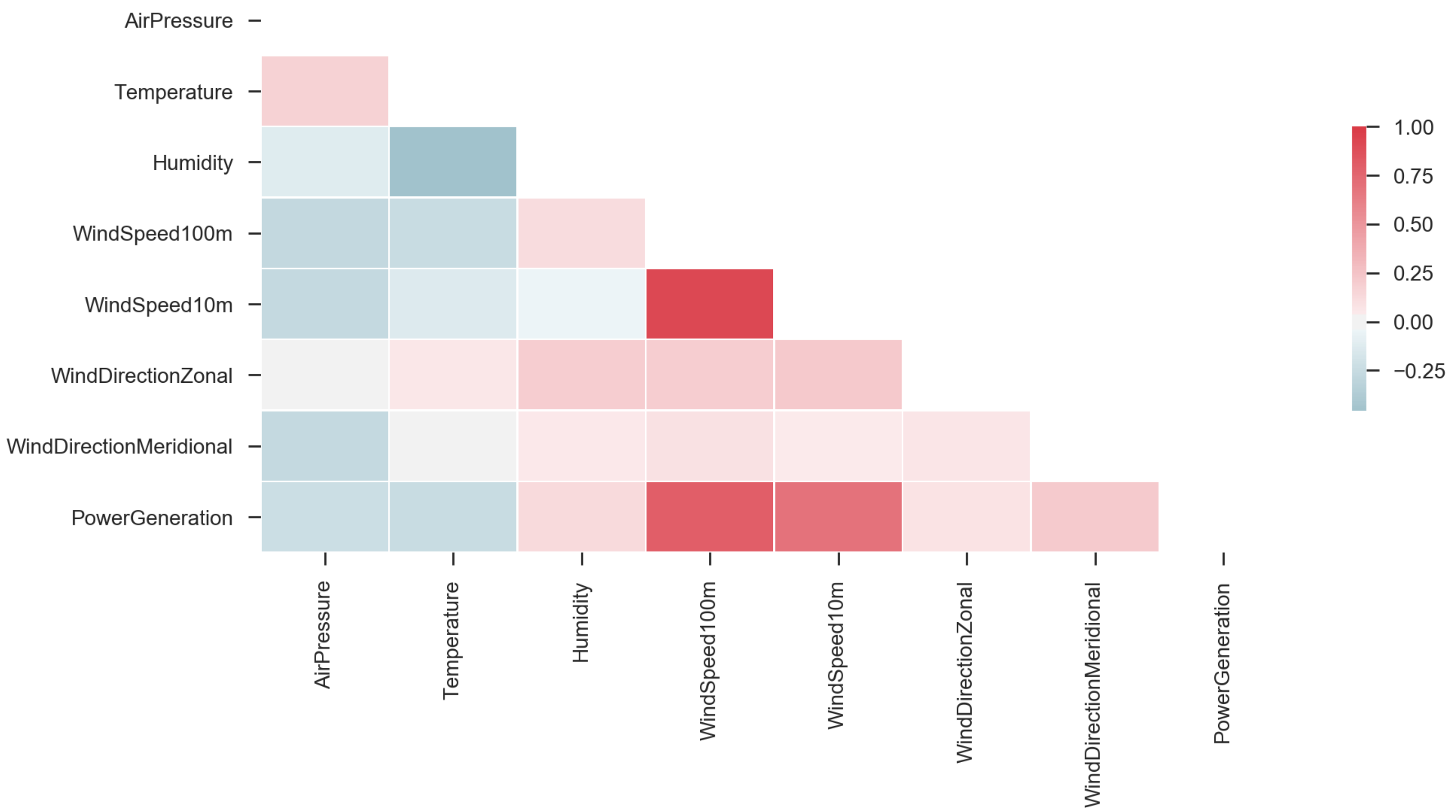
Wind Power
Temporal
Dependency



Wind Power Weather Forecast Variables

Target Variable: Observed Wind Power Generation

Prediction Variables: Wind Speed,Wind Direction, Air Pressure, Humidity, Temperature



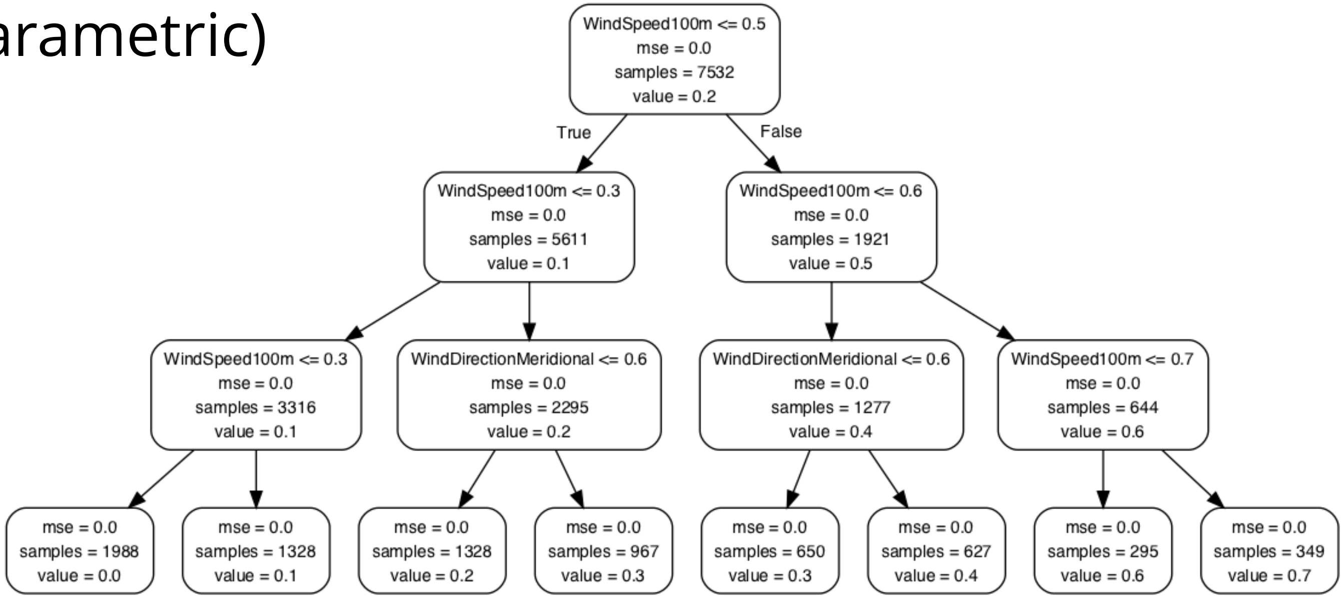
Forecasting Approach

- Random Forest Regression
- Long Short Term Memory Network

Random Forest Regression

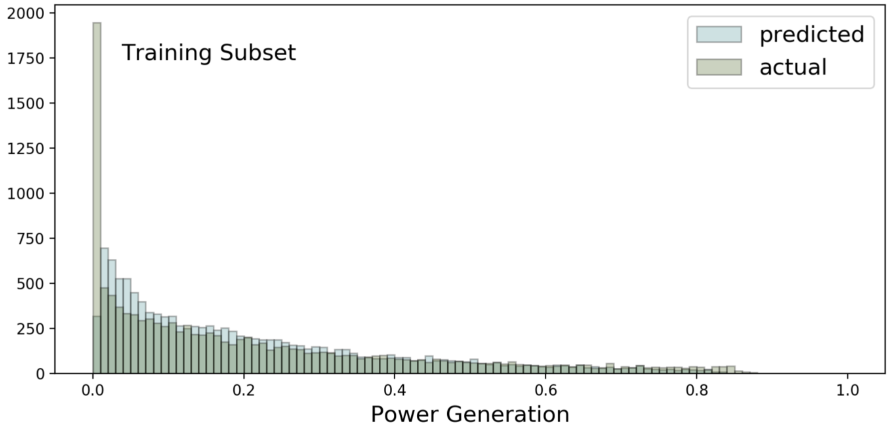
Why use a Random Forest Regression?

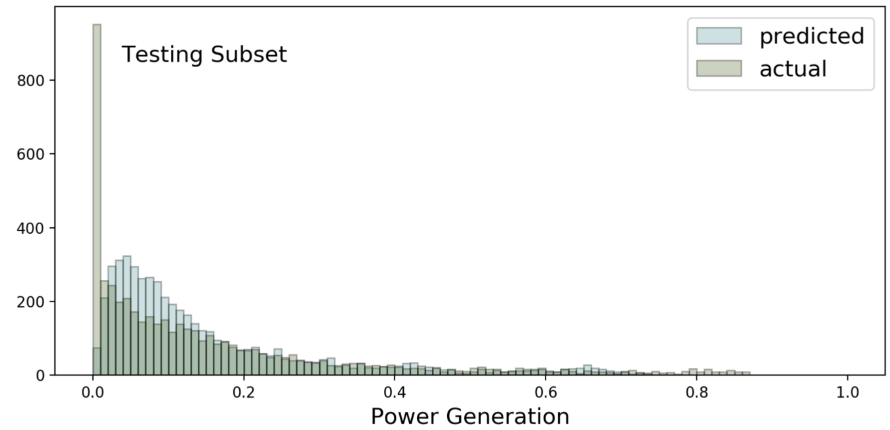
- Random Forest Regression -> Hourly Wind Power Forecast
- Avoids overfitting data
- Built-in cross validation
- Flexible (non parametric)



Random Forest Model Comparison

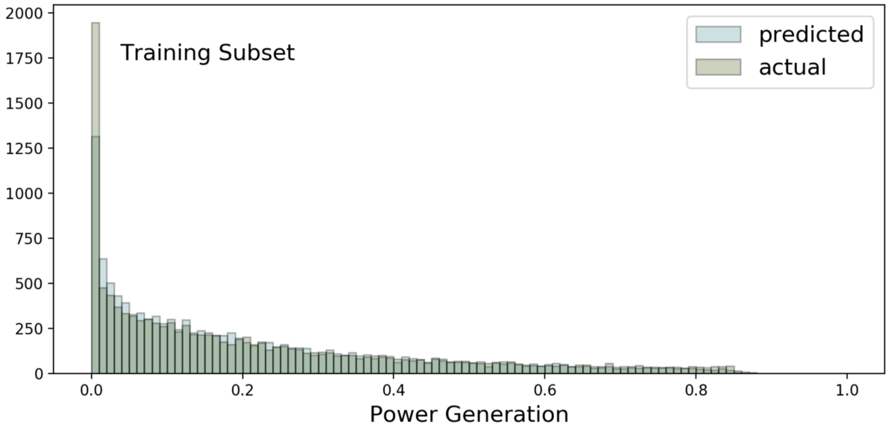
Model 1: Weather Forecast Variables

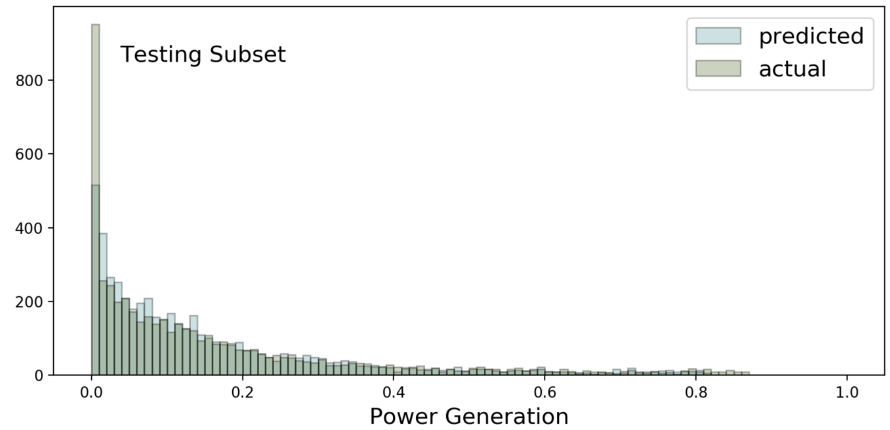




M1 Performance Evaluation - Forecasting Accuracy			
		MAE	*RMSE*
	M1 train data	*0.046*	*0.066*
	M1 test data	*0.071*	*0.105*

Model 2: Weather Variables + Rolling Average of Power





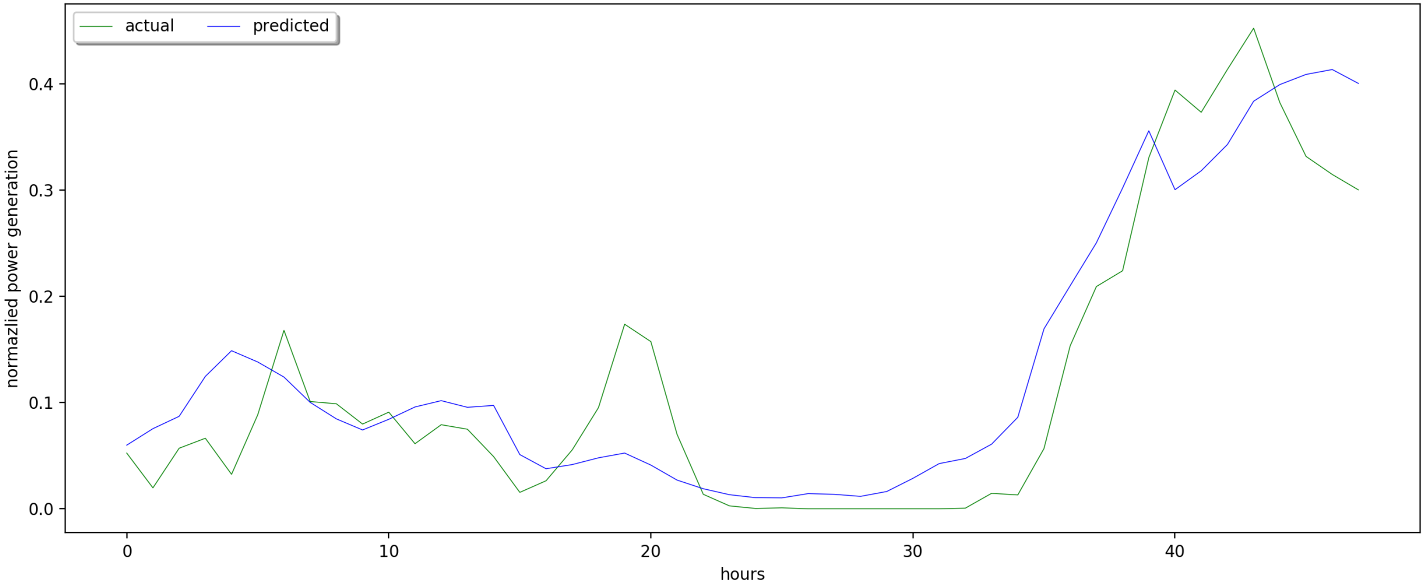
M2 Performance Evaluation - Forecasting Accuracy			
		MAE	*RMSE*
	M2 train data	*0.026*	*0.039*
	M2 test data	*0.042*	*0.064*

Mean Absolute Error (MAE) and Root mean squared error (RMSE) are common metrics used to measure accuracy of continuous variables.

Random Forest Model Comparison

Model 1:

Weather Forecast Variables

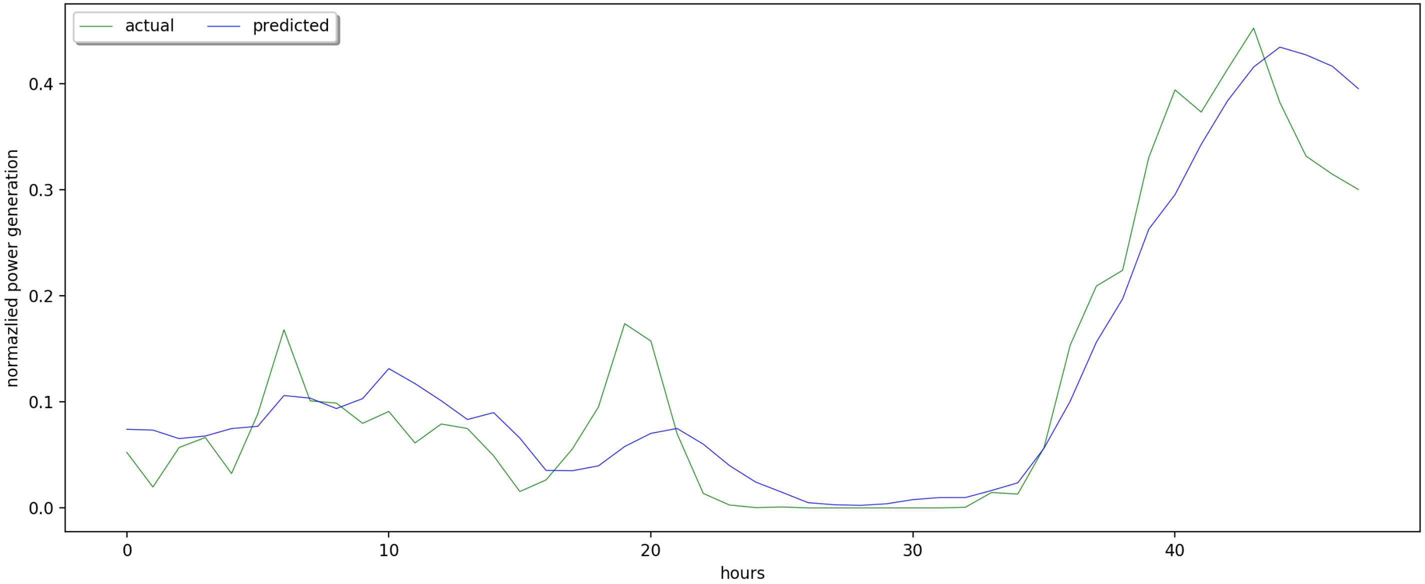


Model 2:

Weather Variables +

Rolling Average of Power

Generation



Forecasting Wind Power with RNN LSTM

LSTM Networks

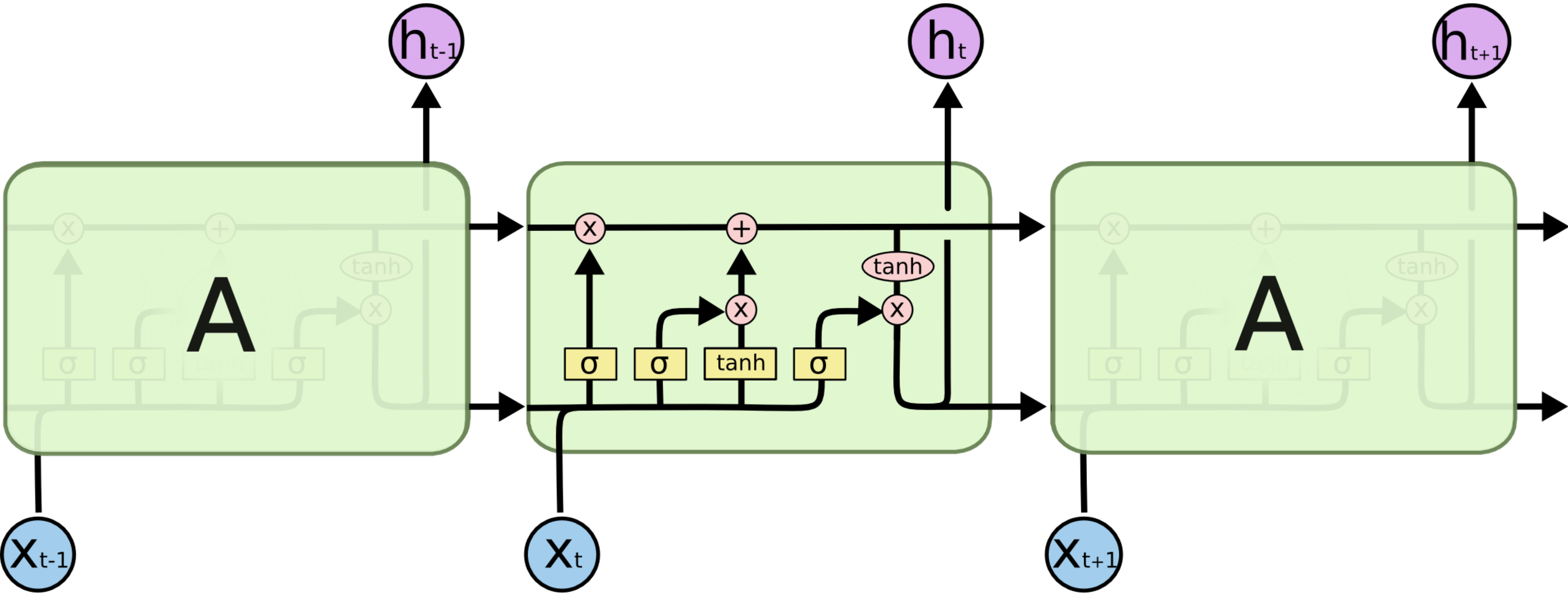
Why use an RNN LSTM Network?

1. Recurrent neural networks can learn temporal features of sequential data, such as time series.

2. LSTM networks can capture long-term dependencies and overcome the vanishing gradient problem in standard RNN models.

LSTM Network

LSTM cells can store, write, and read information



LSTM Cell Information Update Process

Forget Gate - controls what information from previous cell state to forget

Input Gate - controls what new information to store in cell state

Output Gate - controls the extent to which the value in the cell is used to compute the output activation

LSTM Setup

Model

- Stacked LSTM Sequence Model
- Dropout between each layer to reduce overfitting
- Hidden layer size: 20/15/20
- 50 Epochs

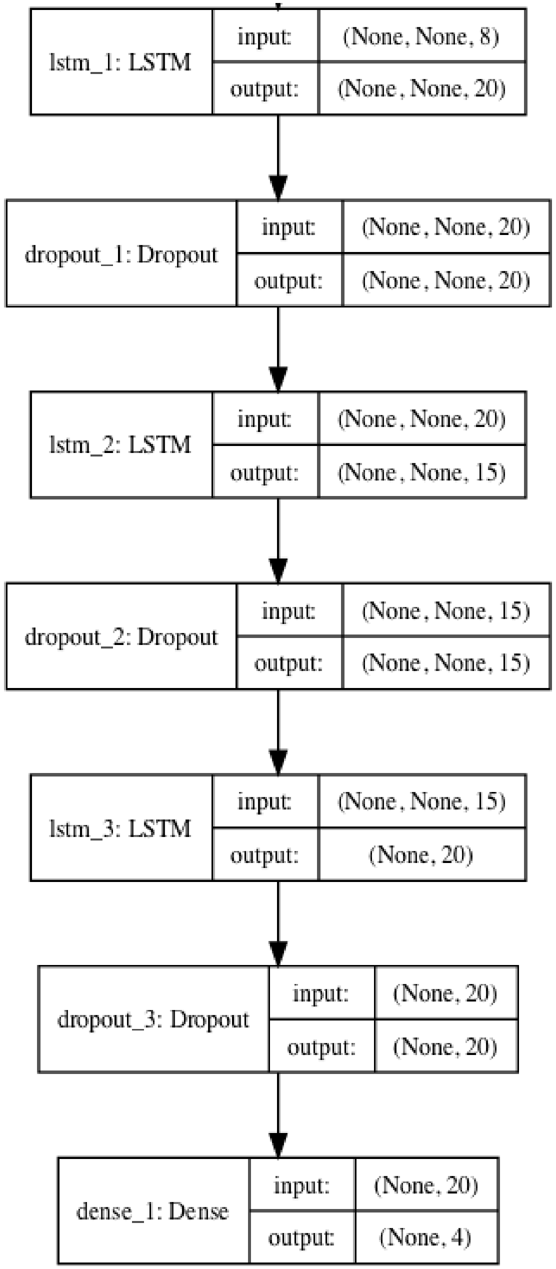
Power Generation Forecast Horizons

- 1, 6, 12, 24 hours

Baseline Comparison - Persistence Method

- "today = tomorrow"

Stacked LSTM

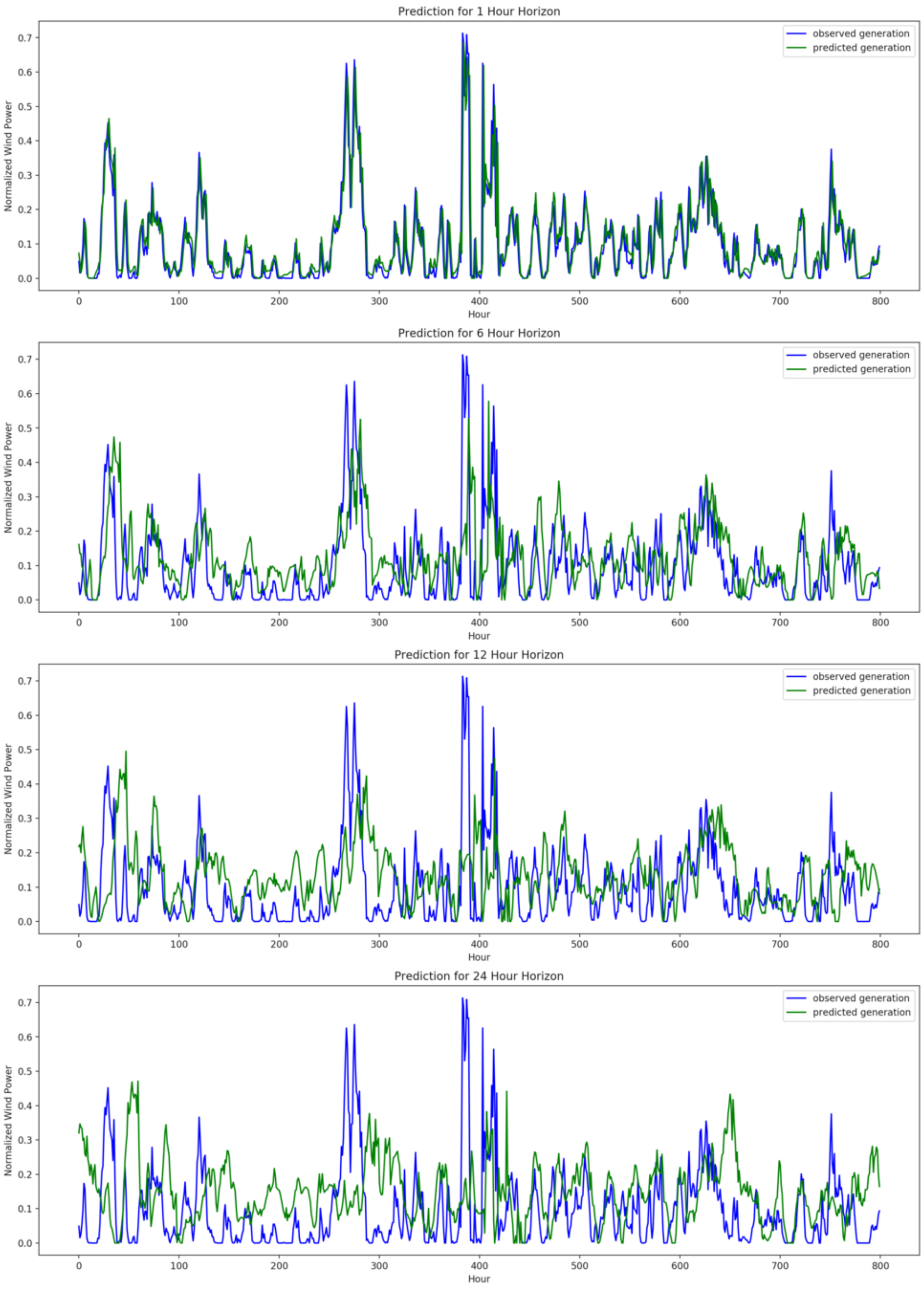


LSTM Model Results

LSTM vs Persistence Method

LSTM Model			
prediction horizon	***MAE***	***MSE***	***RMSE***
1 hour	*0.038*	*0.003*	*0.058*
6 hour	*0.098*	*0.018*	*0.133*
12 hour	*0.134*	*0.032*	*0.178*
24 hour	*0.153*	*0.041*	*0.202*

Persistence Method			
prediction horizon	***MAE***	***MSE***	***RMSE***
1 hour	*0.037*	*0.003*	*0.058*
6 hour	*0.093*	*0.018*	*0.133*
12 hour	*0.123*	*0.032*	*0.178*
24 hour	*0.147*	*0.041*	*0.202*



next steps:

hyperparameter tuning

different data

model architecture

Thanks for Listening!

Special thanks to:

Mentor - Adam Green, Tempus Energy

Data - Dr. Bernhard Sick & André Gensler, University of Kassel

Data Science Retreat - Batch 14 and Staff