Jeff Patra - Data Scientist

Solar Energy Prediction

Capstone Project

"We have this handy fusion reactor in the sky called the sun, you don't have to do anything, it just works. It shows up every day."
– Elon Musk, CEO Tesla Motors

Solar Energy

Advantages

❖ Renewable

❖ Low maintenance

❖ Environmentally friendly

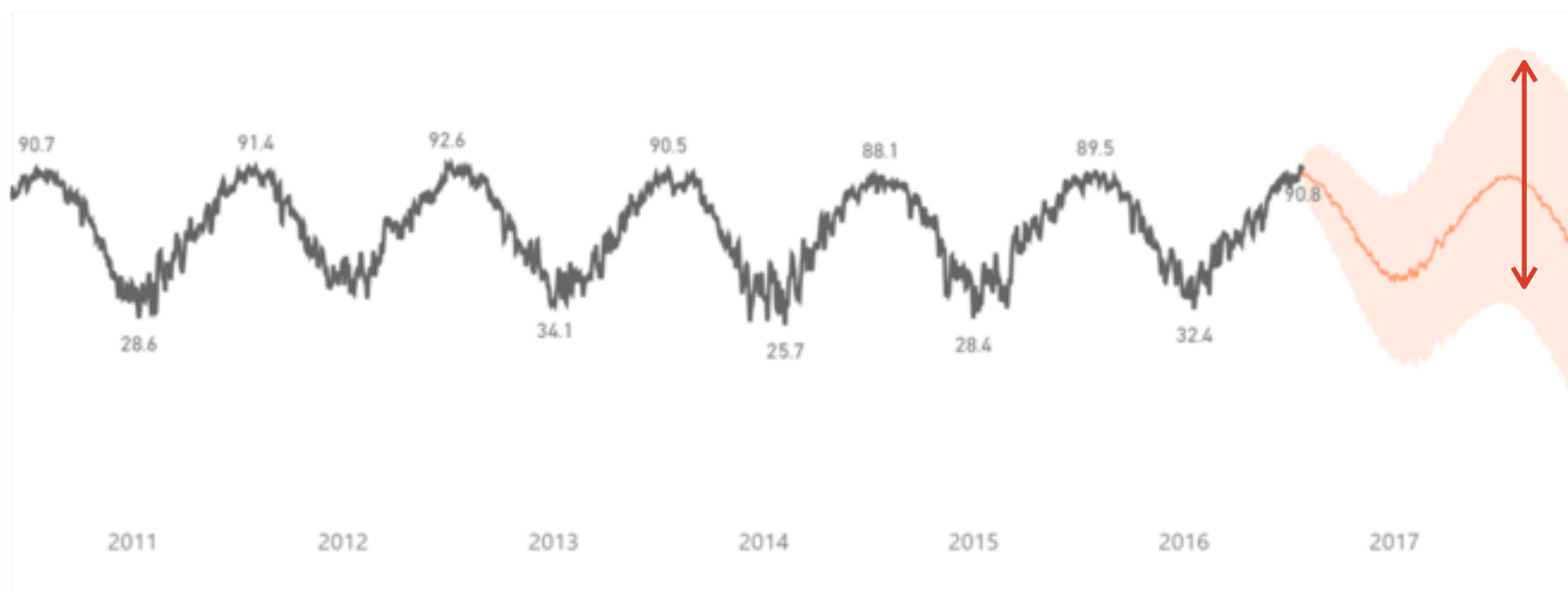## Disadvantages

❖ Expensive upfront cost

❖ Inefficient panel conversion

❖ Inconsistent power generation

Key Stakeholders For Solar Prediction

Solar added to power generation portfolio

Large forecast errors can result in excess expenditure

My Solar Prediction
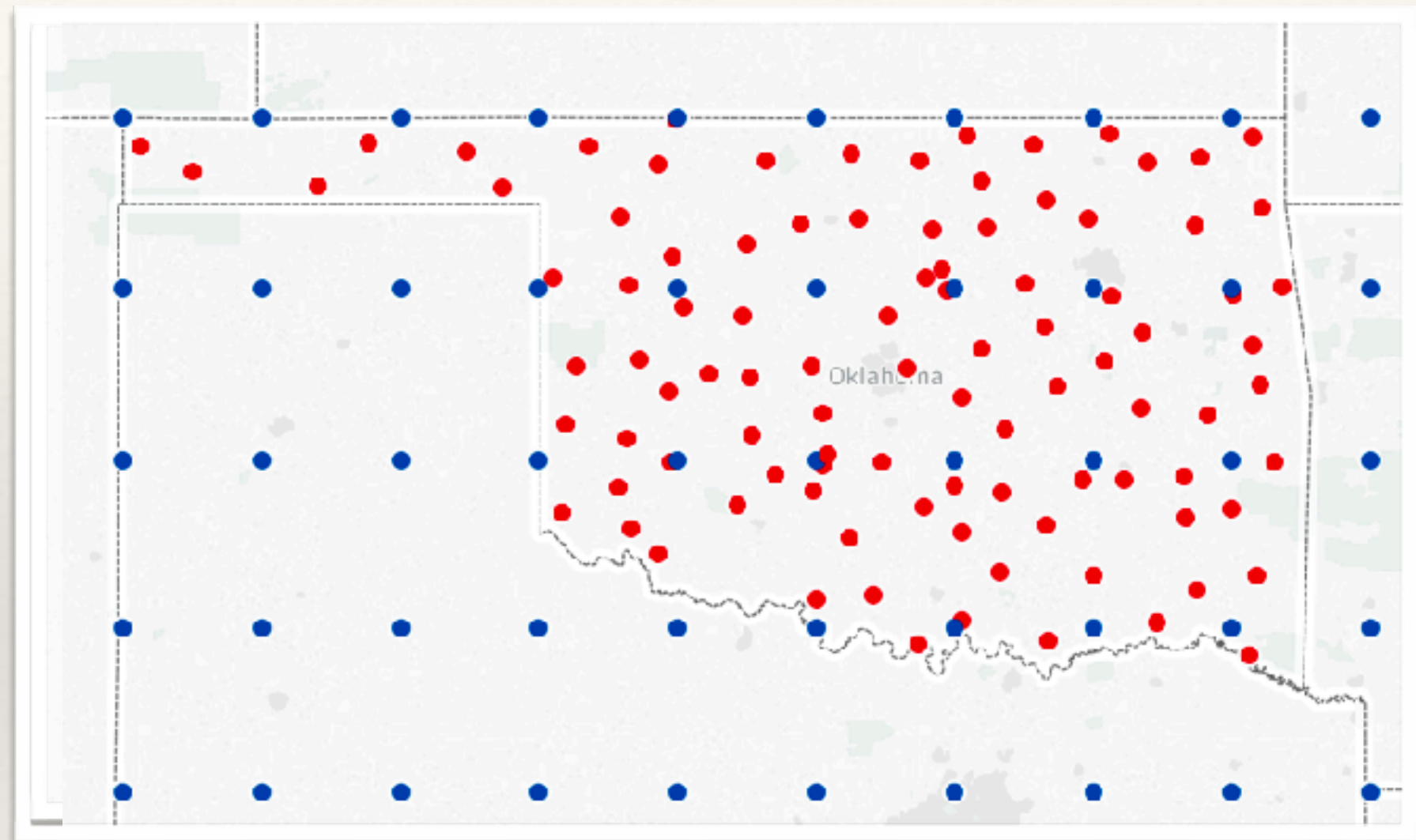
Where?

Oklahoma State

What?

Weather forecast data

98 Weather stations

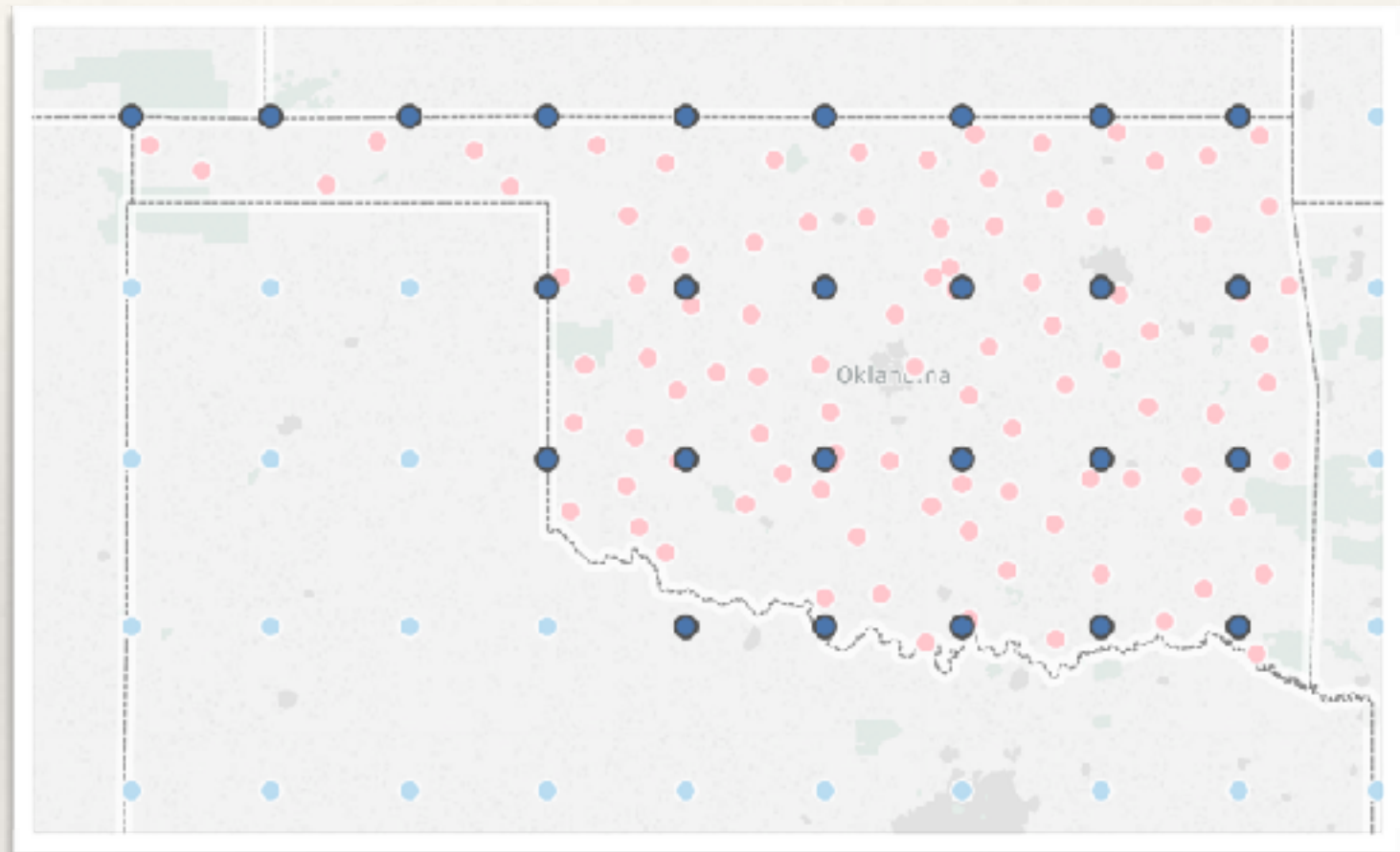## When?

1994 - 2007 (14 years)

The Data
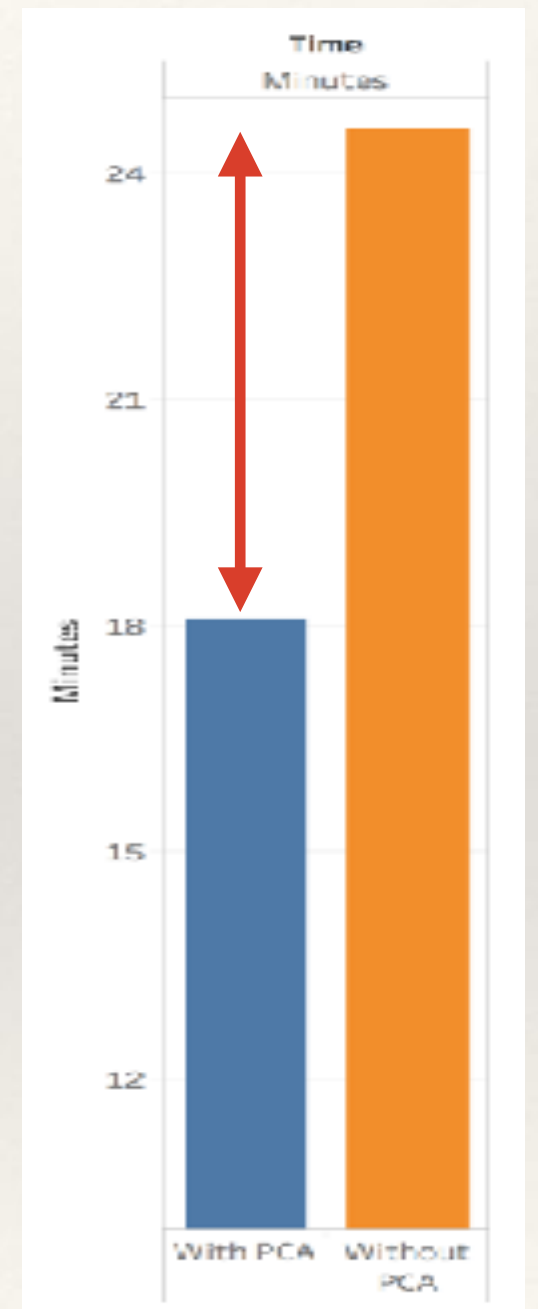
- 15 features in netCDF4 format
    - Temperature
    - Pressure
    - Humidity, etc
- Multi-dimensional array in space and time
- 1.5 GB of data

Work Flow and Processing
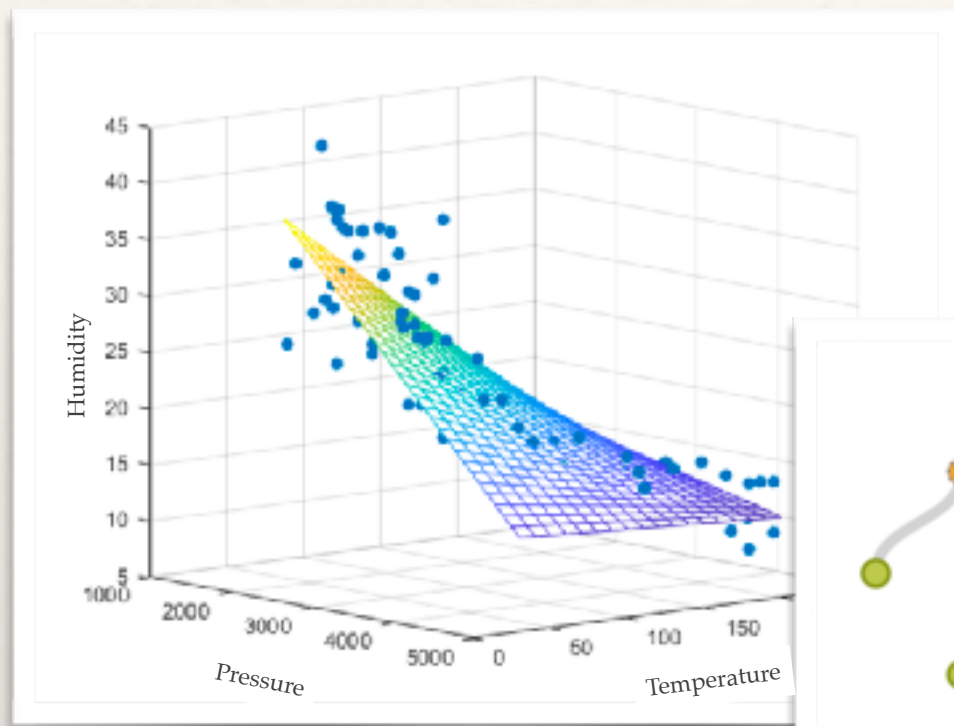
- Load features

- Standardise

- Principal Component Analysis (PCA)

- Run model prediction

- Cross validate

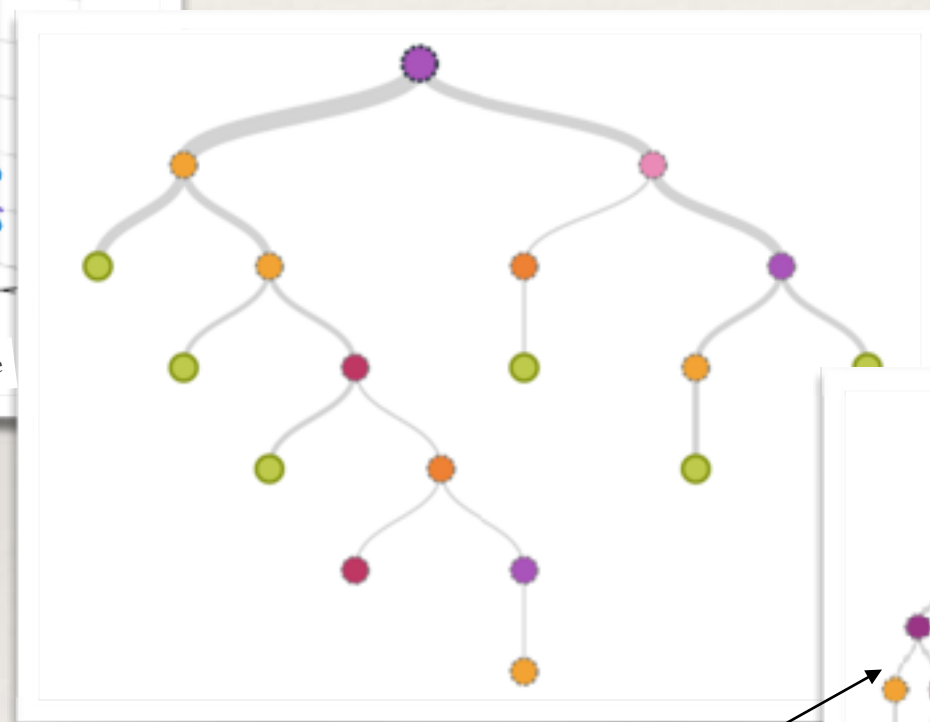- Evaluate results

26% reduction in processing time using PCA
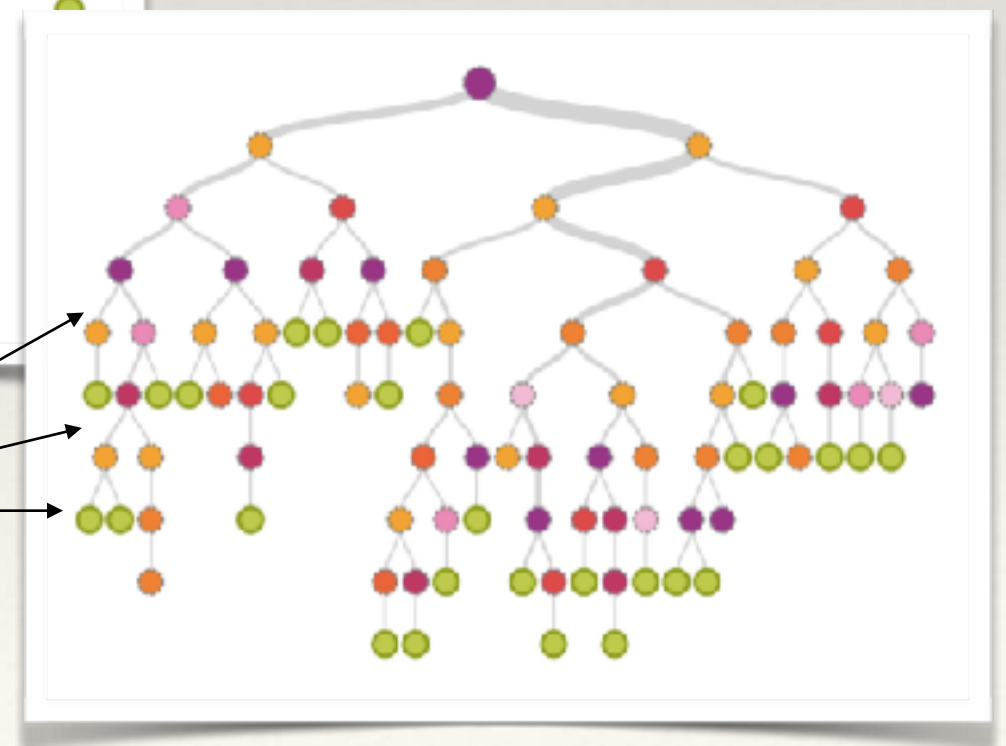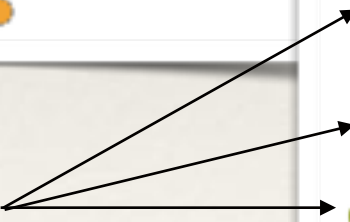
Machine Learning Techniques
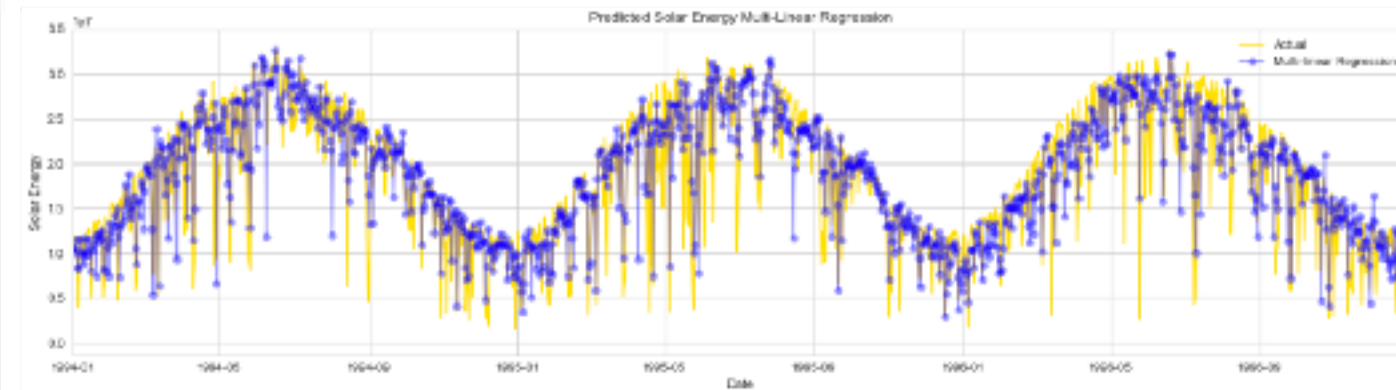
Multi-linear regression

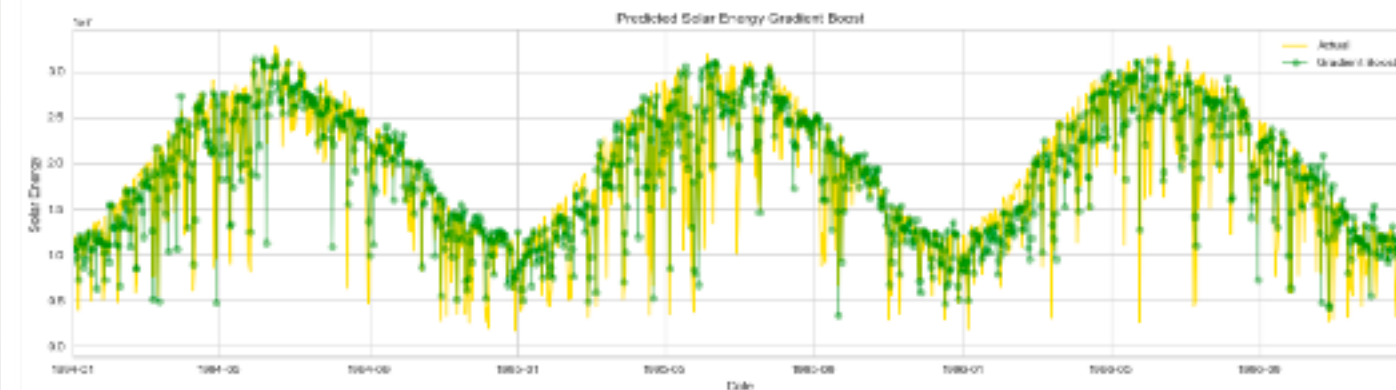Decision Tree
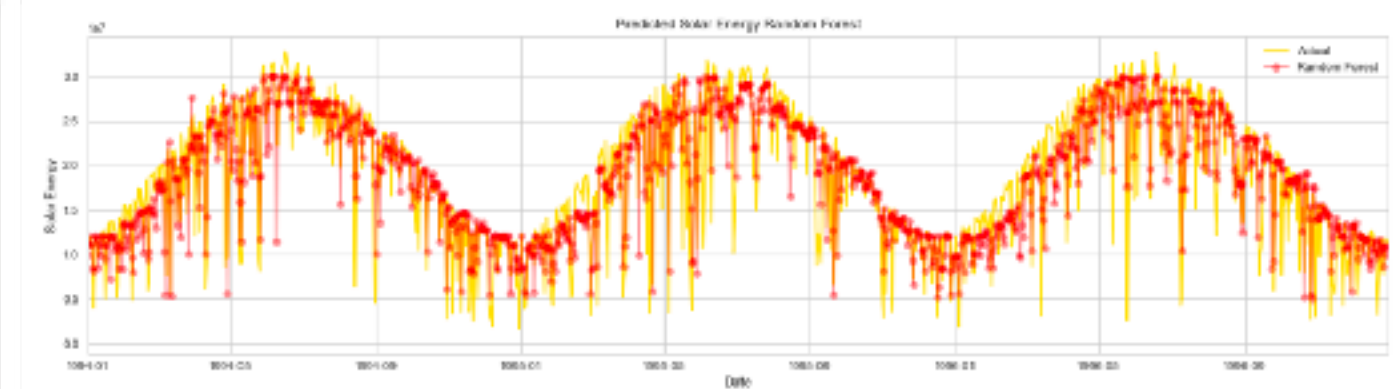
Random Forest

Gradient Boosting
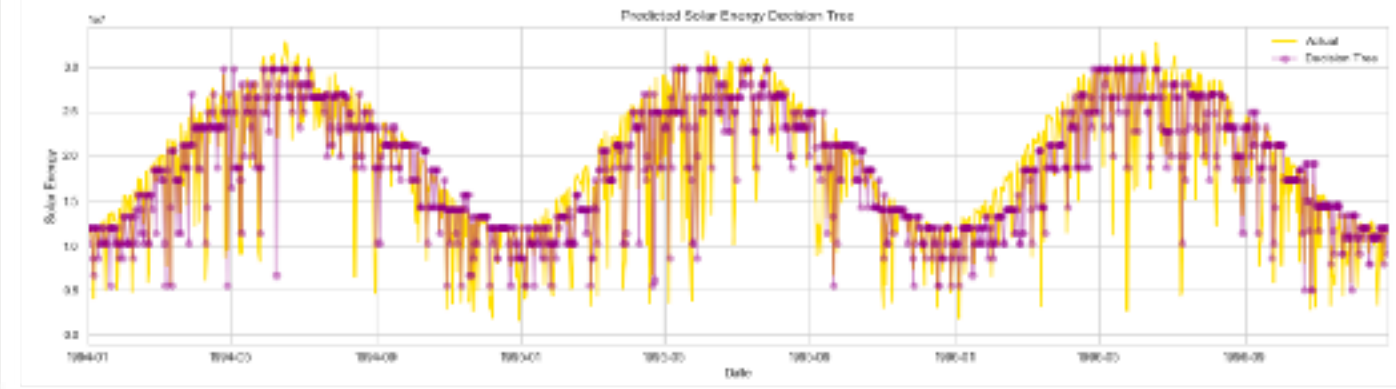
Model Predictions

Multi-linear

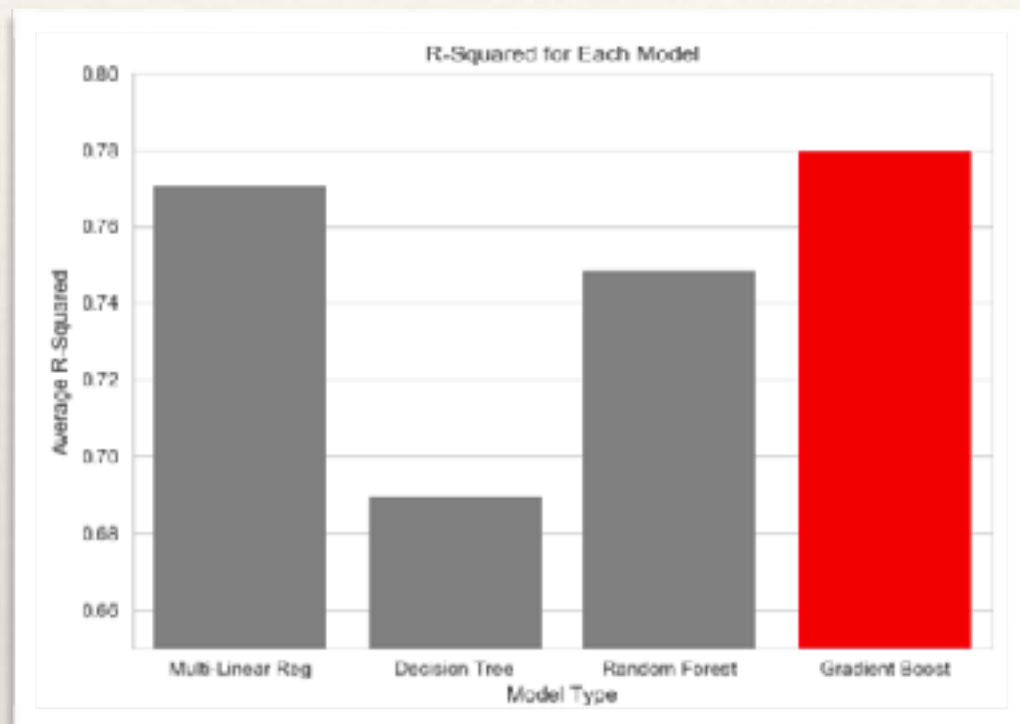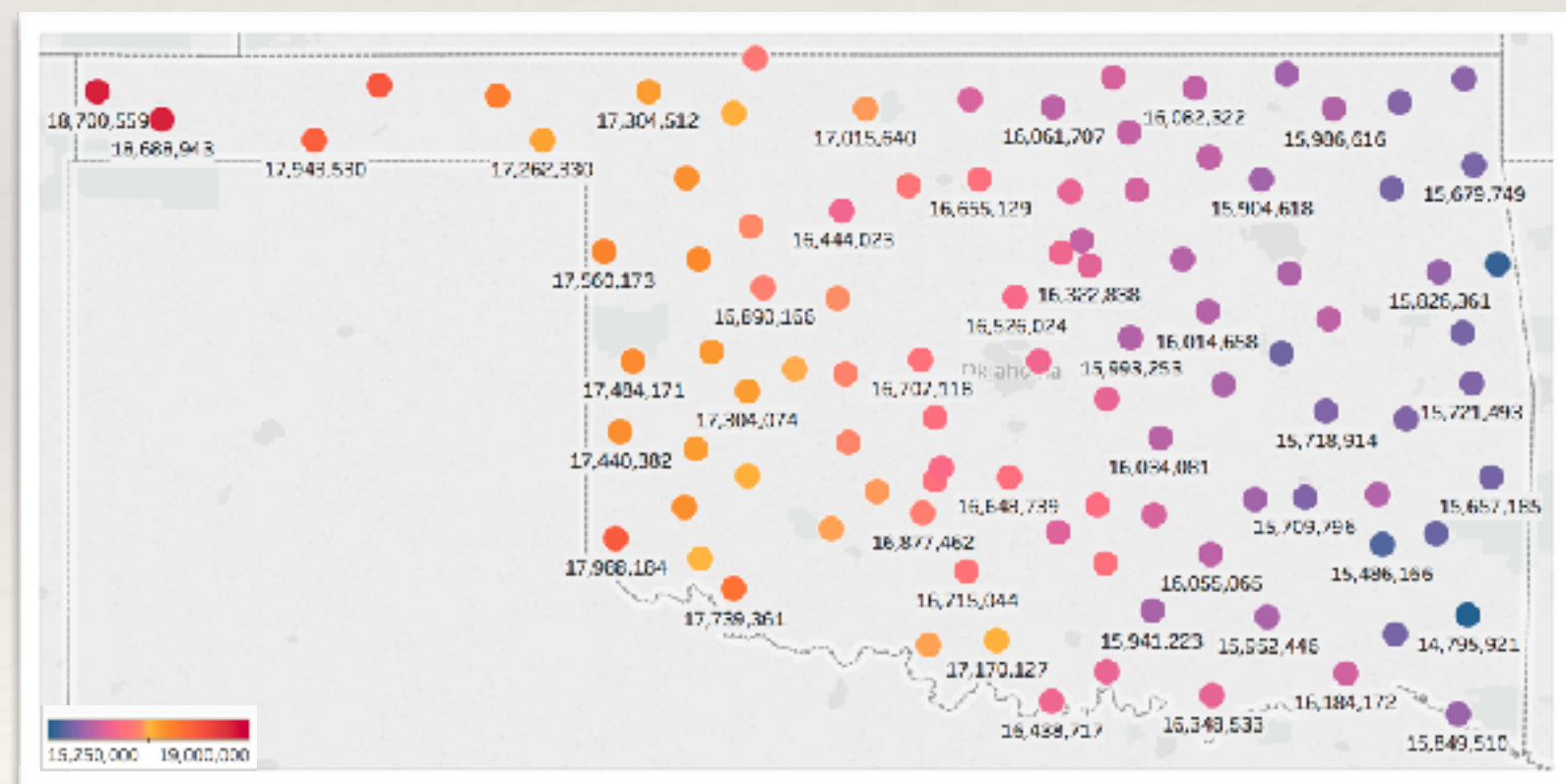Gradient Boost

Random Forest

Decision Tree

Results

Gradient boost wins

Energy production increasing from SE-NW

Results & Recommendations

- Machine learning works!

- Model prediction accounts for severe weather

- Principal component analysis saved 26% in processing time

- Try deep learning neural network