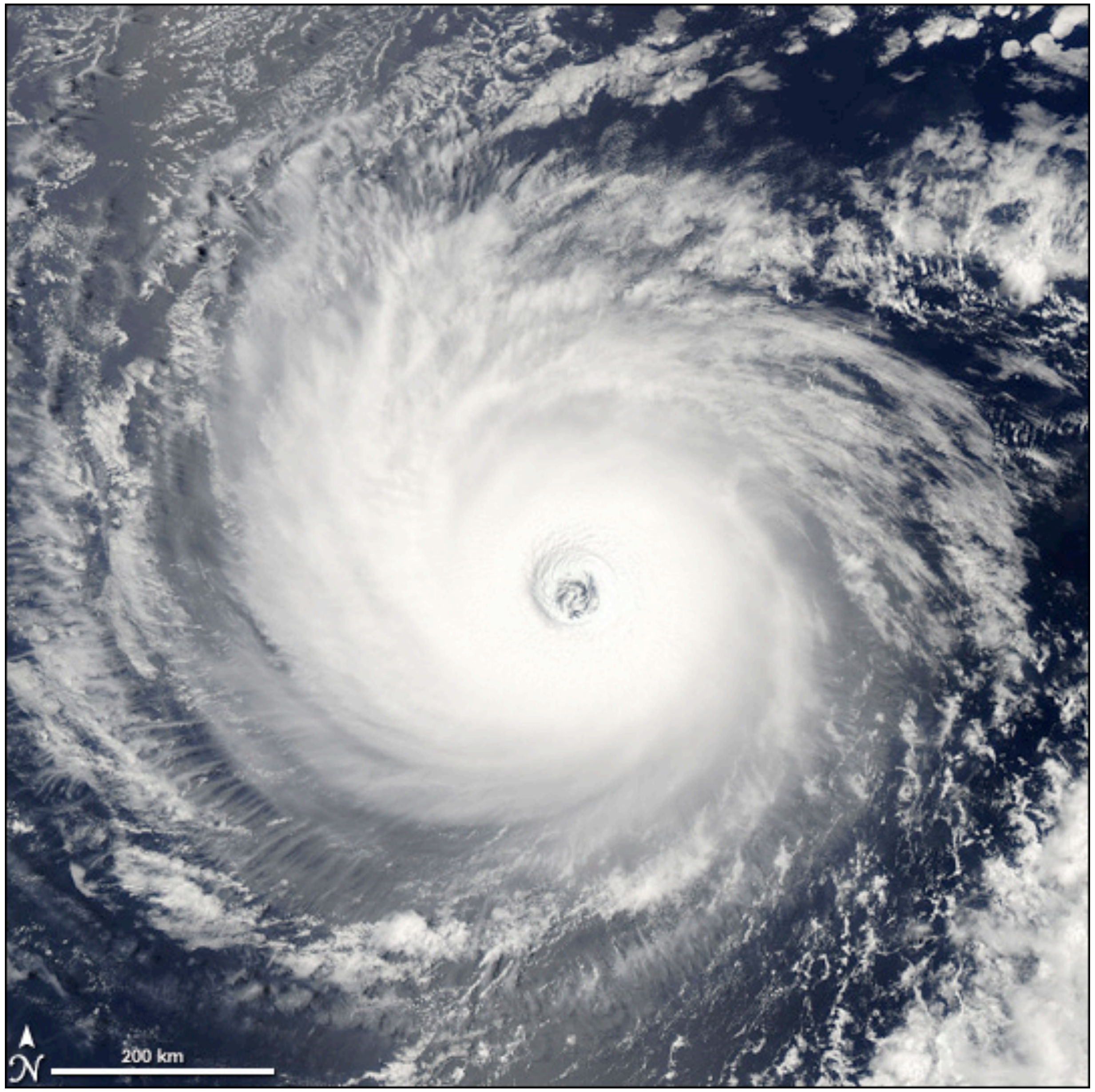
Hurricane Forecasting Using Robust Locally Weighted Regression

Ali Yahya

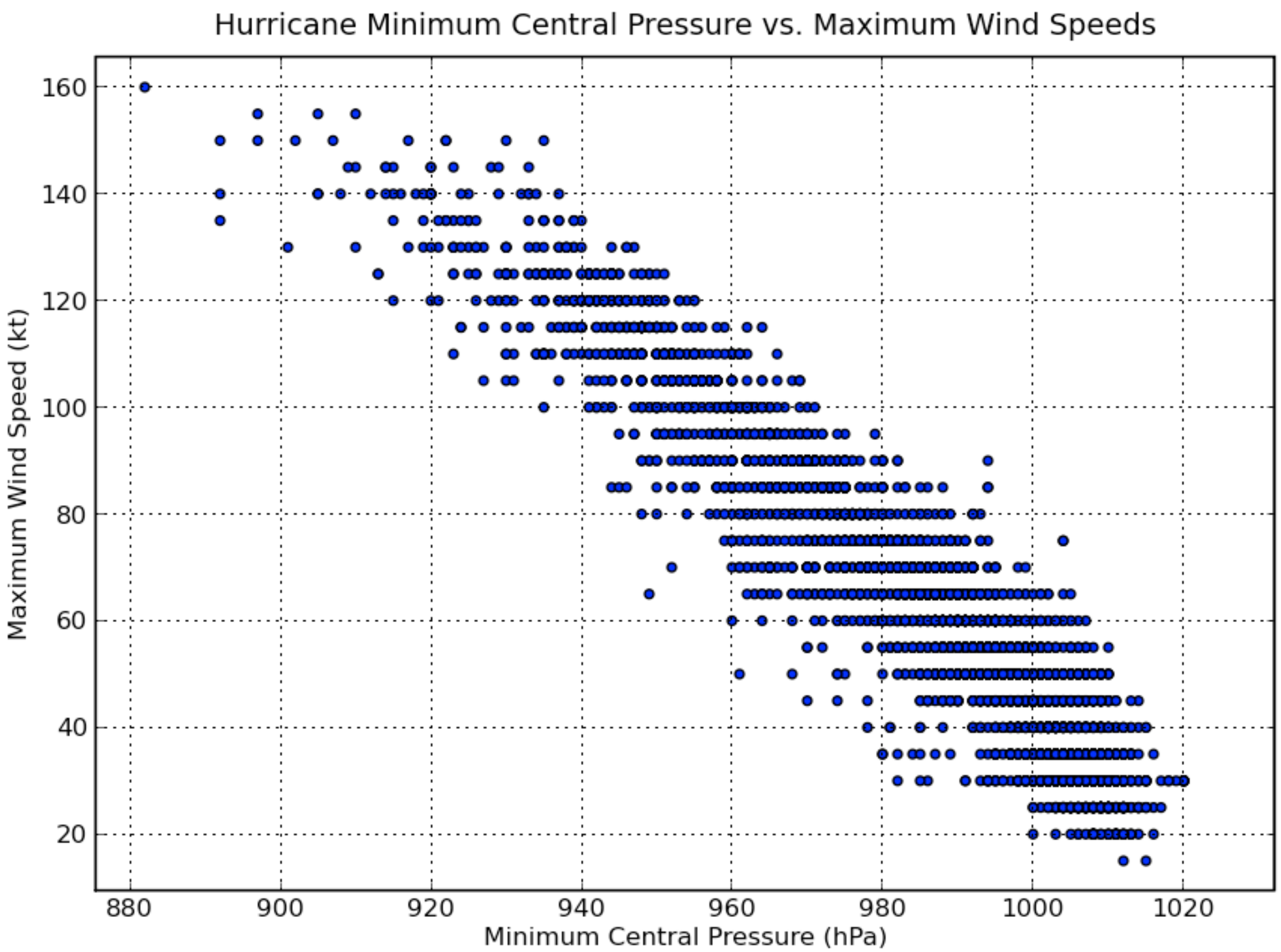
Project View at 1,000,000 ft.



- Forecast model parameters:
  - maximum wind speed
  - minimum central pressure
  - radius of maximum wind speed
  - radius of outer closed isobar
- Project goal:
  - device a predictive model for any one feature given the remaining features
  - develop an understanding of the intricate relationships between features
- Motivation:
  - incomplete/inconsistent data
  - better understanding of each parameter's side-effects

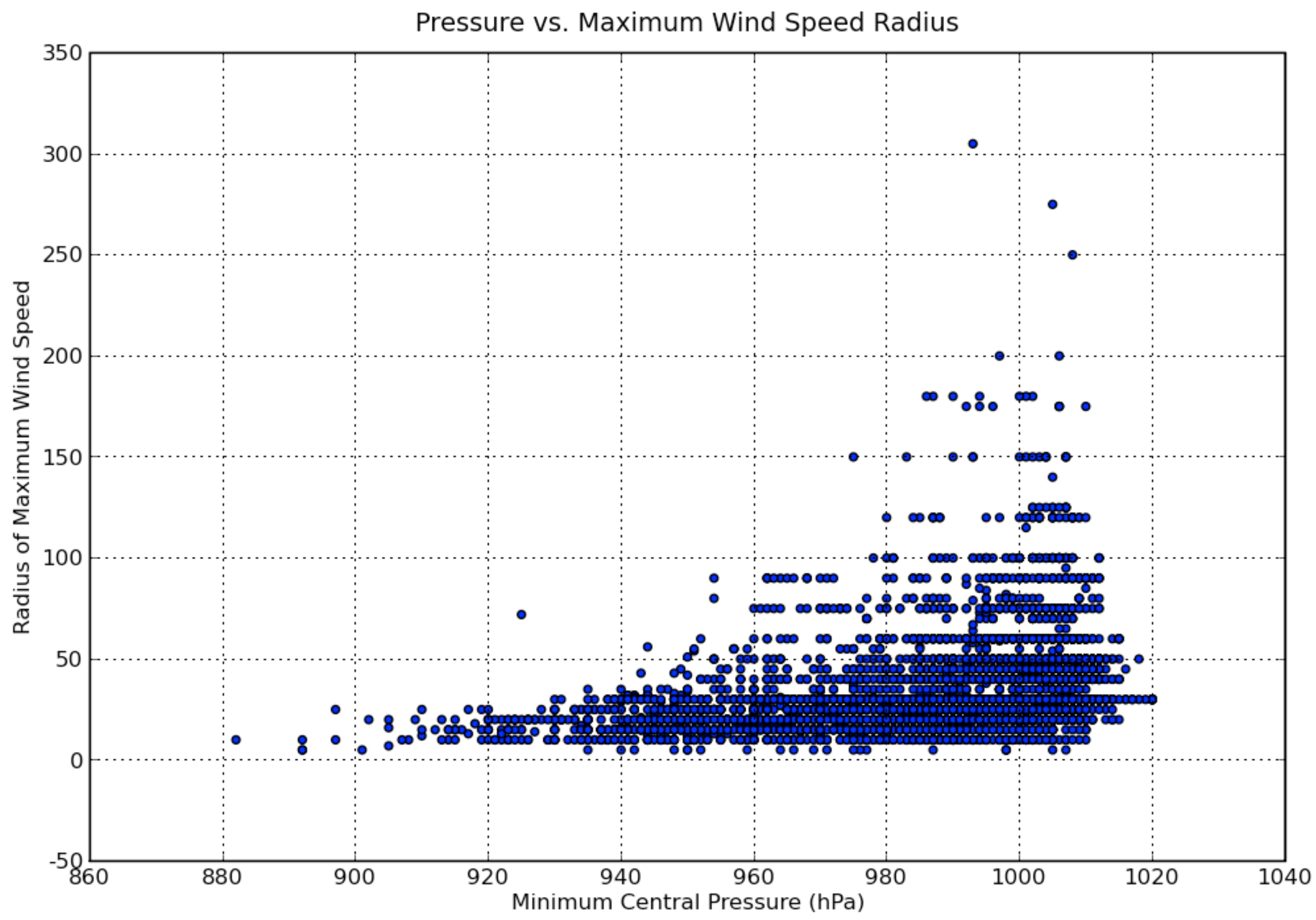
Looking for Patterns

- Data collection:
  - processed data from ~10,000 hurricanes
  - payed attention to correlations between model features
- Correlations Observed:
  - Min. Pressure vs. Max. Speed
  - Pressure vs. Max. Speed Radius
  - Max. Speed vs. Max. Speed Radius
  - Longitude vs. Max. Speed



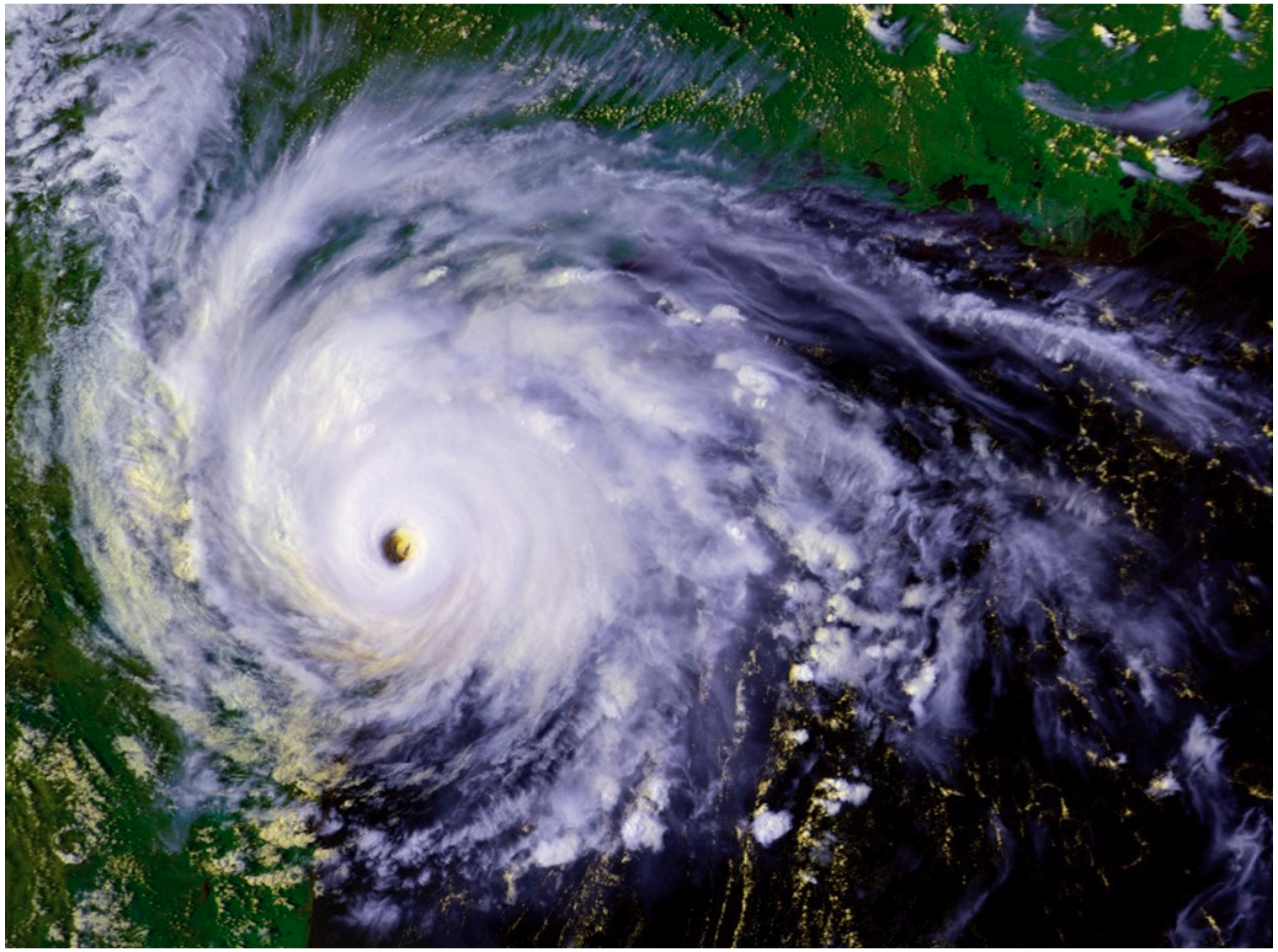
Looking for Patterns

- Data collection:
    - processed data from ~10,000 hurricanes
    - payed attention to correlations between model features
- Correlations Observed:
    - Min. Pressure vs. Max. Speed
    - Pressure vs. Max. Speed Radius
    - Max. Speed vs. Max. Speed Radius
    - Longitude vs. Max. Speed



Looking for Patterns

- Data collection:
  - processed data from ~10,000 hurricanes
  - payed attention to correlations between model features
- Correlations Observed:
  - Min. Pressure vs. Max. Speed
  - Pressure vs. Max. Speed Radius
  - Max. Speed vs. Max. Speed Radius
  - Longitude vs. Max. Speed

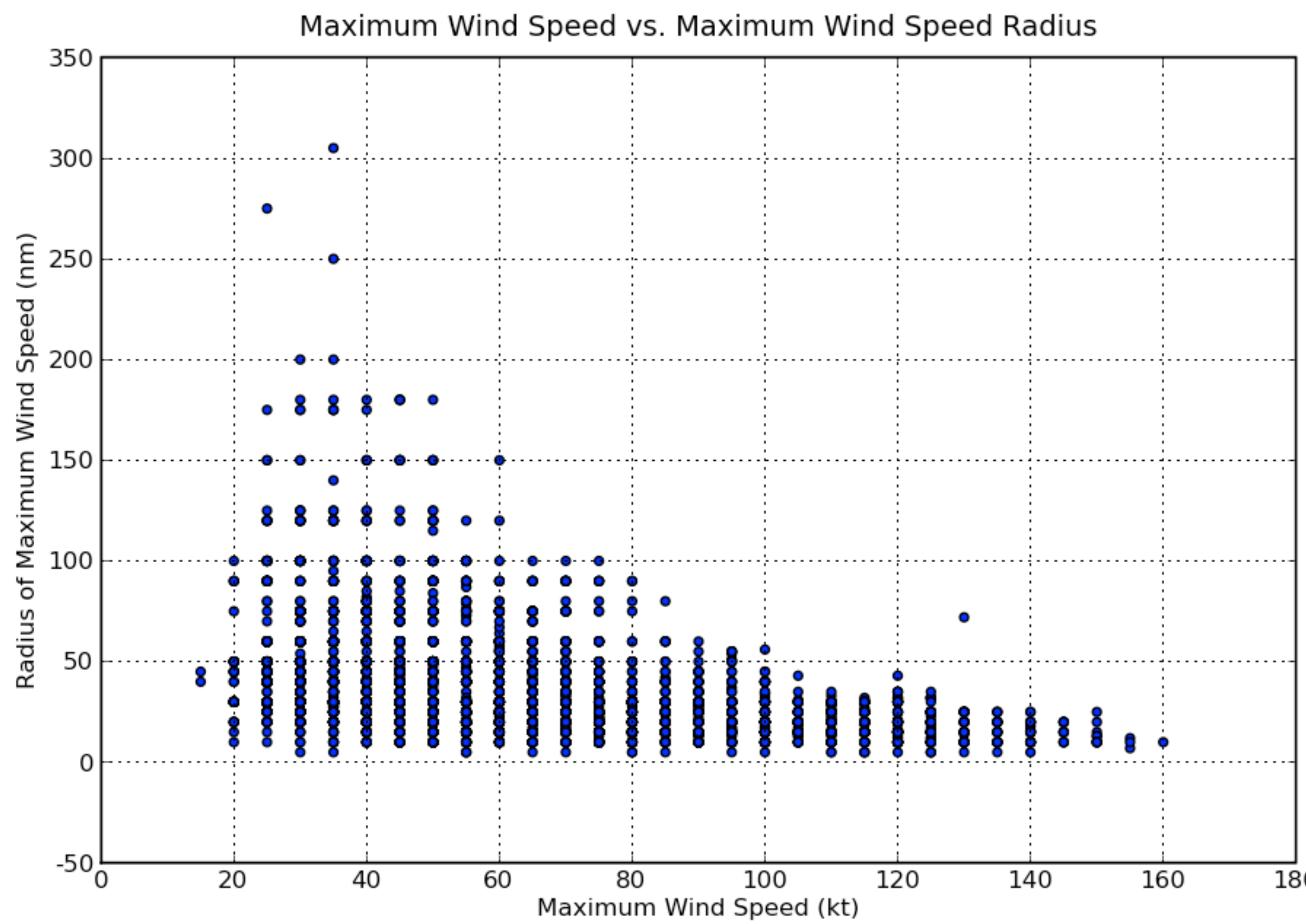


Hurricane Bret, 1999

Maximum Wind Speed Radius: 305 nm

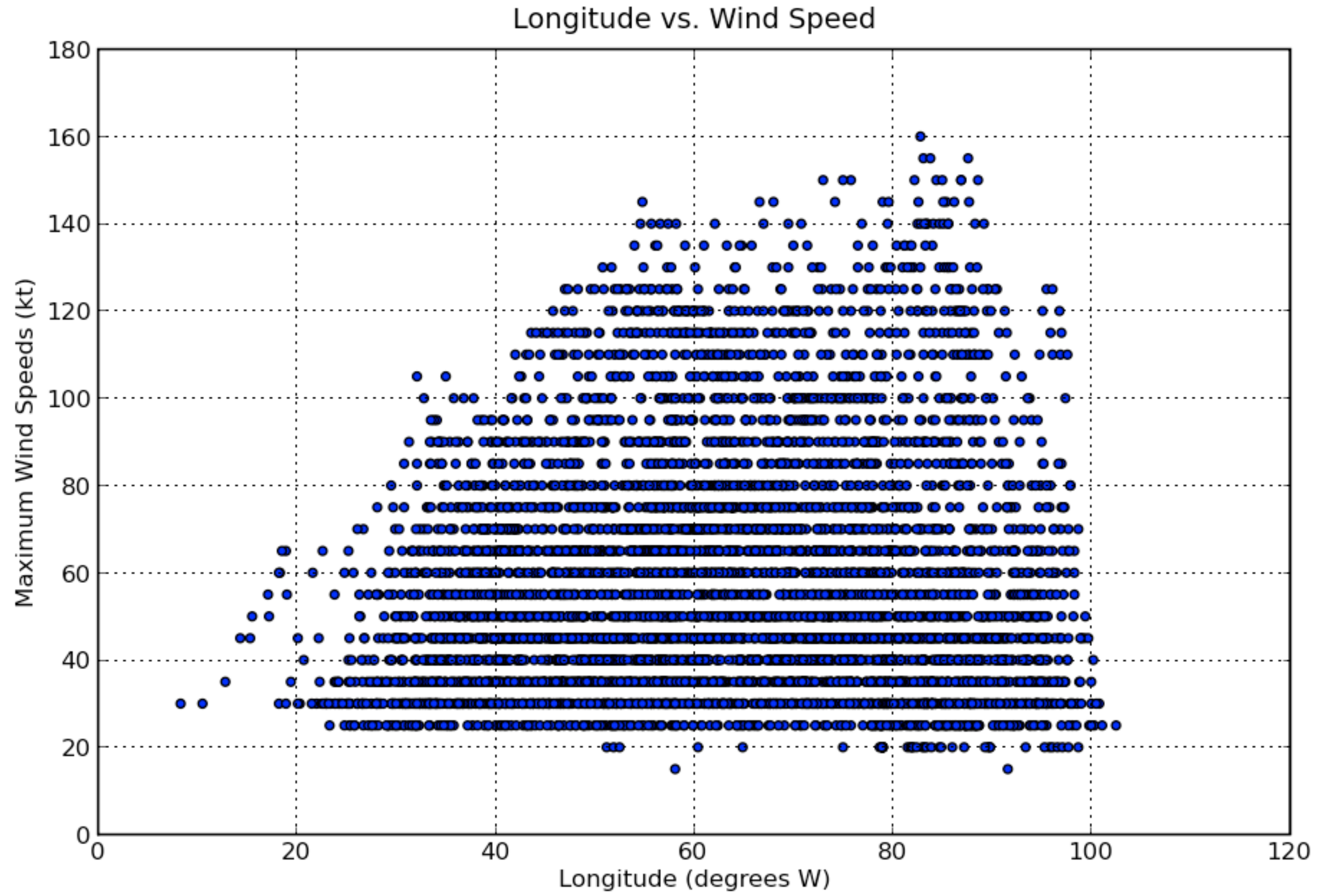
Looking for Patterns

- Data collection:
  - processed data from ~10,000 hurricanes
  - payed attention to correlations between model features
- Correlations Observed:
  - Min. Pressure vs. Max. Speed
  - Pressure vs. Max. Speed Radius
  - Max. Speed vs. Max. Speed Radius
  - Longitude vs. Max. Speed



Looking for Patterns

- Data collection:
  - processed data from ~10,000 hurricanes
  - payed attention to correlations between model features
- Correlations Observed:
  - Min. Pressure vs. Max. Speed
  - Pressure vs. Max. Speed Radius
  - Max. Speed vs. Max. Speed Radius
  - Longitude vs. Max. Speed



Formalized Observations

- Implemented LOESS,

  A machine learning algorithm based on Locally Weighted Regression

- Forecast model is a hyperplane in *n*-dimensional space, where *n* is the number of parameters

Allows the following:

```
from cyclone import cylone_predict

prediction = cyclone_predict(x_query, X, alpha)

```

Where:

- X is a matrix containing historical hurricane data,
- x_query is a vector containing info about a new hurricane; info about one of the hurricane's features will be missing
- the return value is a prediction for the missing features

Algorithm Details

- The algorithm minimizes the following cost function with respect to Theta :

$$J(\theta) = \frac{1}{2}\sum_{i=1}^{m} w^{(i)} \left(\theta^T x^{(i)} - y^{(i)}\right)^2 = \frac{1}{2}(X\theta - \vec{y})^T W (X\theta - \vec{y})$$

- Differentiating the cost function, setting equal to zero, and solving for theta yields a closed form solution to the minimization problem:

$$\theta = (X^T W X)^{-1} X^T W \vec{y}$$

- Theta can then be used to compute a prediction for any given query *x*:

$$h_\theta(x) = \theta^T x$$

- RESULT: Average percent error for hypotheses across ~2,000 data points:

8.4%

The End

- Project is Open Source:

https://github.com/ali01/loess.py

- References:

- Cleveland, W.S. (1979). "Robust Locally Weighted Regression and Smoothing Scatterplots". *Journal of the American Statistical Association* **74** (368): 829–836. doi:10.2307/2286407. MR0556476. JSTOR 2286407.
- "Worldwide Tropical Cyclone Names.". National Hurricane Center.National Oceanic and Atmospheric Administration. 2009. Retrieved 2009-05-07.
- NHC Hurricane Research Division (2006-02-17). "Atlantic hurricane best track ("HURDAT")". NOAA. Retrieved 2007-02-22.