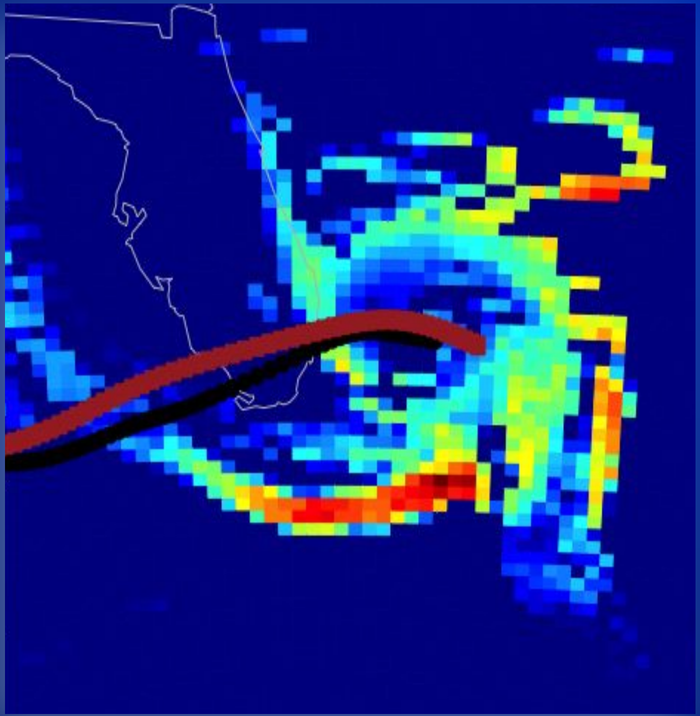
Statistical Learning for Stochastic Tropical Cyclone Simulation



Laurette Hamlin

Introduction

The basis of my culminating experience is a research paper entitled

Stochastic Tropical Cyclone Precipitation Field Generation
(William Kleiber et al. 2020)

which successfully modeled precipitation patterns during tropical cyclones in Houston, Texas.

I tested the success of the research by replicating and validating the process on new storm data from

Miami, Florida

Introduction

Inland Flooding

Flooding from heavy rains is the second leading cause of fatalities from landfalling tropical cyclones. Widespread torrential rains associated with these storms often cause flooding hundreds of miles inland. This flooding can persist for several days after a storm has dissipated.

A total of six persons are known to have died directly as a result of the effects of winds and water in south Florida. Three of the deaths, all in Broward County, were

More Precipitation

A hurricane's ability to produce rain is affected by the temperature of the air and ocean water. Warm air can hold more moisture; more moisture often leads to more rain. That's how climate change causes wetter storms. Researchers studying Hurricane Harvey found that human-induced climate change made extreme rainfall more likely. In general, models show hurricane rainfall increasing by 10 to 15 percent on average by the end of the century. That means that we may see more storms like Harvey.

hurricane strength in the southeast Gulf of Mexico. The distribution of winds and rain near the center of Katrina while passing across the south Florida peninsula was asymmetrical with the strongest winds and heaviest rain located to the south and east of the center.

Research Goal

Simulate tropical cyclones using only a handful of features from standard track databases

- ➢ Radius of Maximal Winds
- ➢ Pressure Deficit at Sea Level
- ➢ Wind Direction
- ➢ Storm Center

Why Statistical Learning?

- ➔ Typical physics-based models require lots of computing power and time
    - ◆ We want **fast, localized desktop estimations**
- ➔ Tropical cyclones are relatively **rare meteorological events**
- ➔ Instruments that take measurements are strategically **spaced apart**

Statistical Techniques Used

- **Smoothing Splines** - piecewise step functions used here for interpolation
- **Random Forests** - decision tree-based method for regression analysis
- **Spatial Smoothing** - multi-dimensional interpolation
- **Principal Component Analysis (PCA)** - unsupervised learning to reduce feature set.
- **Empirical Orthogonal Functions (EOFs)** - for estimating orthonormal basis from a data set

Environment

R

- ➢ **fields** - tools for spatial statistics
- ➢ **LatticeKrig** - multi-resolution kriging tool based on Markov random fields
- ➢ **randomForest** - random forest generator for classification and regression

Data Description

- **WRF** (Weather Research & Forecasting model)
  - Mesoscale numerical weather prediction system
  - used for all of our model data
  - 79 variables modeled over time for each TC event, kept 10

- **IBTrACS** (International Best Track Archive for Climate Stewardship)
  - 21 actual measurements for each TC event
  - Kept storm centers in latitude/longitude, tracked over time

Data Example

Tropical Storm Bonnie

July, 2010

- Variable: Sea Level Pressure

- Simulated by WRF at
  - Spatial field: 59 x 59 locations near Miami
  - Temporal vector: every hour for 3.5 days

- 59 x 59 x 84 = **292,404 total data points**

Data Processing
Time Synchronization

- WRF Data
  - 1 hour intervals 2010-07-24 18:00:00 GMT
  - Actually the integration of 17:00-18:00
  - Switched to midpoints (i.e. 17:30)

- IBTrACS Data
  - 6 hour intervals
  - Interpolated to 1 hour increments,
  - Reduced an additional 30 minutes for midpoint calculation
  - Storm centers interpolated using a natural spline function

Data Transforms

- Pressure Deficit at Sea Level
  - Interpolated using Spatial Smoothing
    - "fields" package, normal kernel smoother (λ = 10)

- Distance to Coast

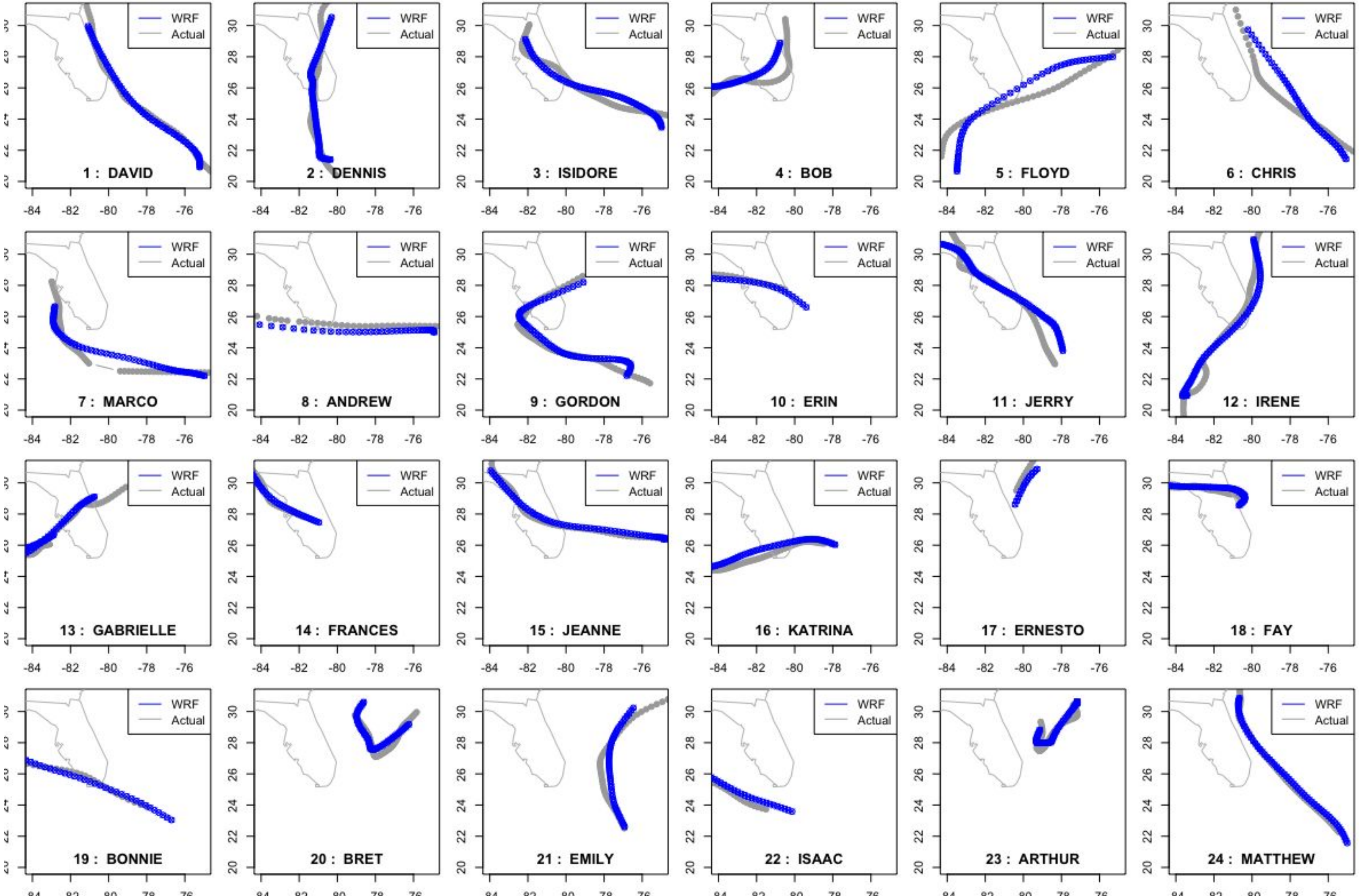
- Radius of Maximal Winds
  - Interpolated using a smoothing spline (λ = 0.75)

- Estimation of Storm Centers
  - Calculated from vector curl
  - Interpolated using a cubic smoothing spline (λ = 0.75)

Data Transform
Verification of Storm Centers

- 9 storms were removed from original 35
  - spatial edge cases
  - transforms/estimations performed poorly

- Study went forward with **24 storm events**
  - from 1979 - 2016



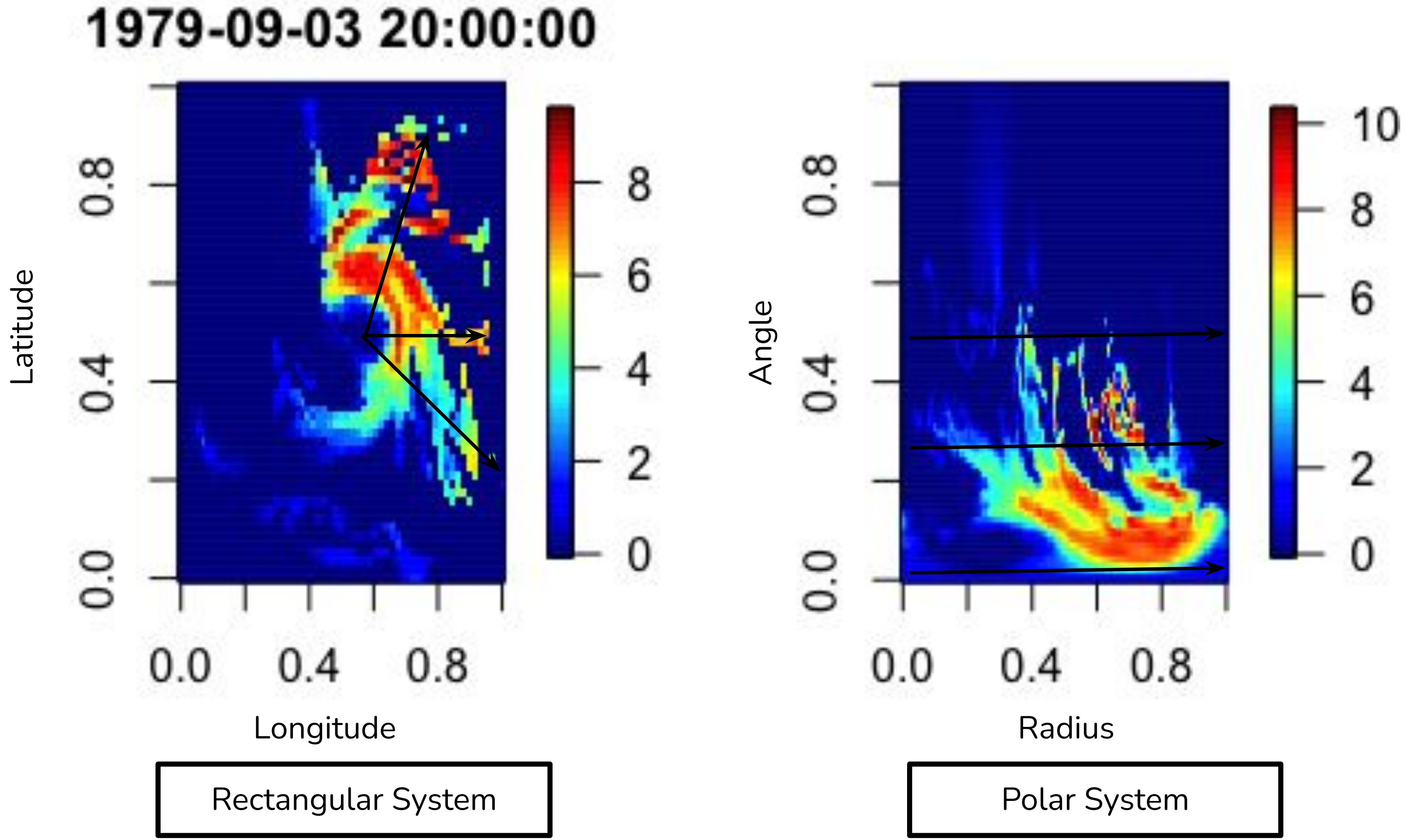
WRF
Actual
1 : DAVID
2 : DENNIS
3 : ISIDORE
4 : BOB
5 : FLOYD
6 : CHRIS
7 : MARCO
8 : ANDREW
9 : GORDON
10 : ERIN
11 : JERRY
12 : IRENE
13 : GABRIELLE
14 : FRANCES
15 : JEANNE
16 : KATRINA
17 : ERNESTO
18 : FAY
19 : BONNIE
20 : BRET
21 : EMILY
22 : ISAAC
23 : ARTHUR
24 : MATTHEW

Projection of Precipitation to Polar Coordinates - Reasoning

- ➢ Precipitation in a hurricane does not fall uniformly from the sky
- ➢ It forms and moves in accordance with the circular, spiral character of the tropical cyclone.
- ➢ It makes things much easier computationally to transform
  - from: fixed, rectangular coordinates (latitude, longitude)
  - to: polar coordinates radiating out from the storm center (r, θ)

Projection of Precipitation to Polar Coordinates - Methodology

- Fine-Grained for localized predictions
  - 100 grid points for each parameter
- Calculated using fields::rdist.earth
  - geographic distance matrices
- Required spatial smoothing
  - both variances and distances within random field are used
  - field is polar
  - calculated using LatticeKrig package
    - "a variation of Kriging with fixed basis functions that uses a compactly supported covariance to create a regular set of basis functions on a grid. The coefficients of these basis functions are modeled as a Gaussian Markov random field (GMRF)" (Nychka et al. 2016)



1979-09-03 20:00:00
Latitude
0.0
0.4
0.8
Longitude
0.0
0.4
0.8
0
2
4
6
8
Rectangular System
Angle
0.0
0.4
0.8
Radius
0.0
0.4
0.8
0
2
4
6
8
10
Polar System

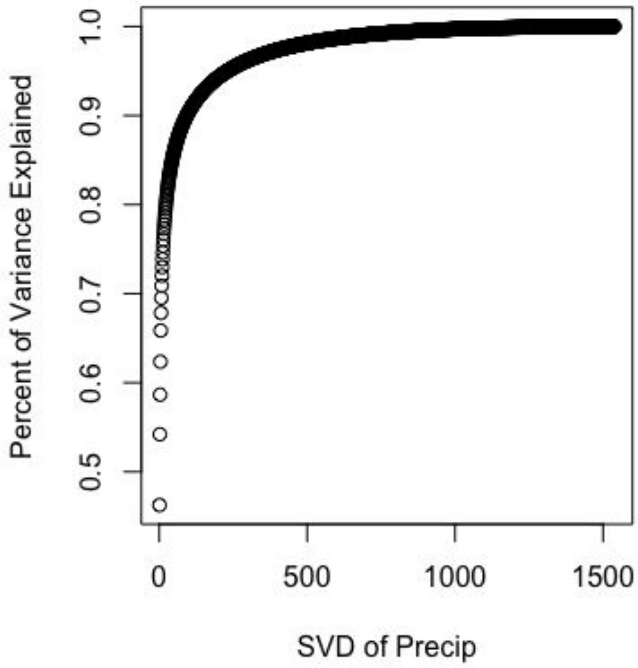
Transform of Precipitation
Singular Value Decomposition

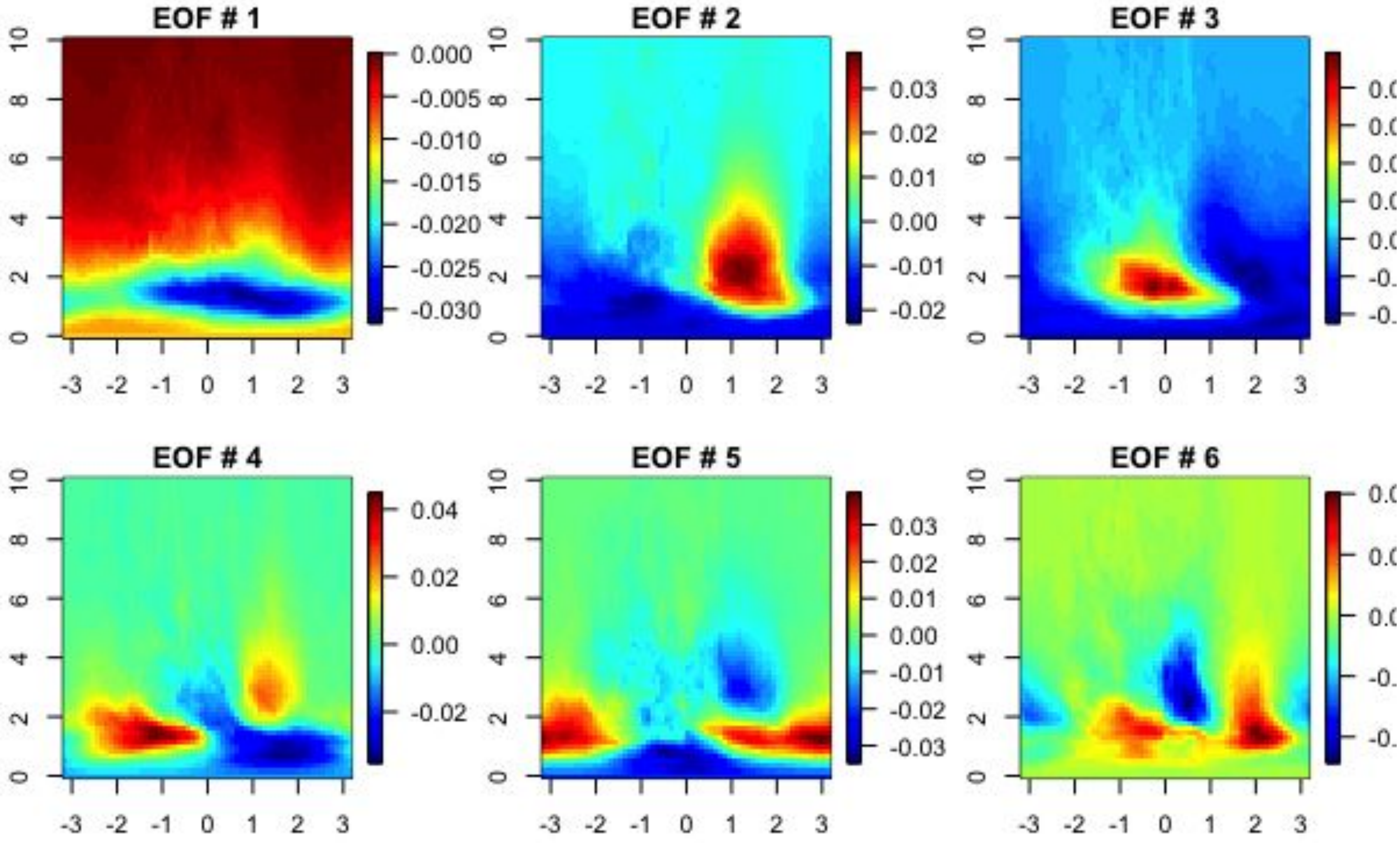
- Input Variable(s):
  - polar spatial grid matrix (100 x 100)
  - smoothed precipitation measures at each hourly time interval

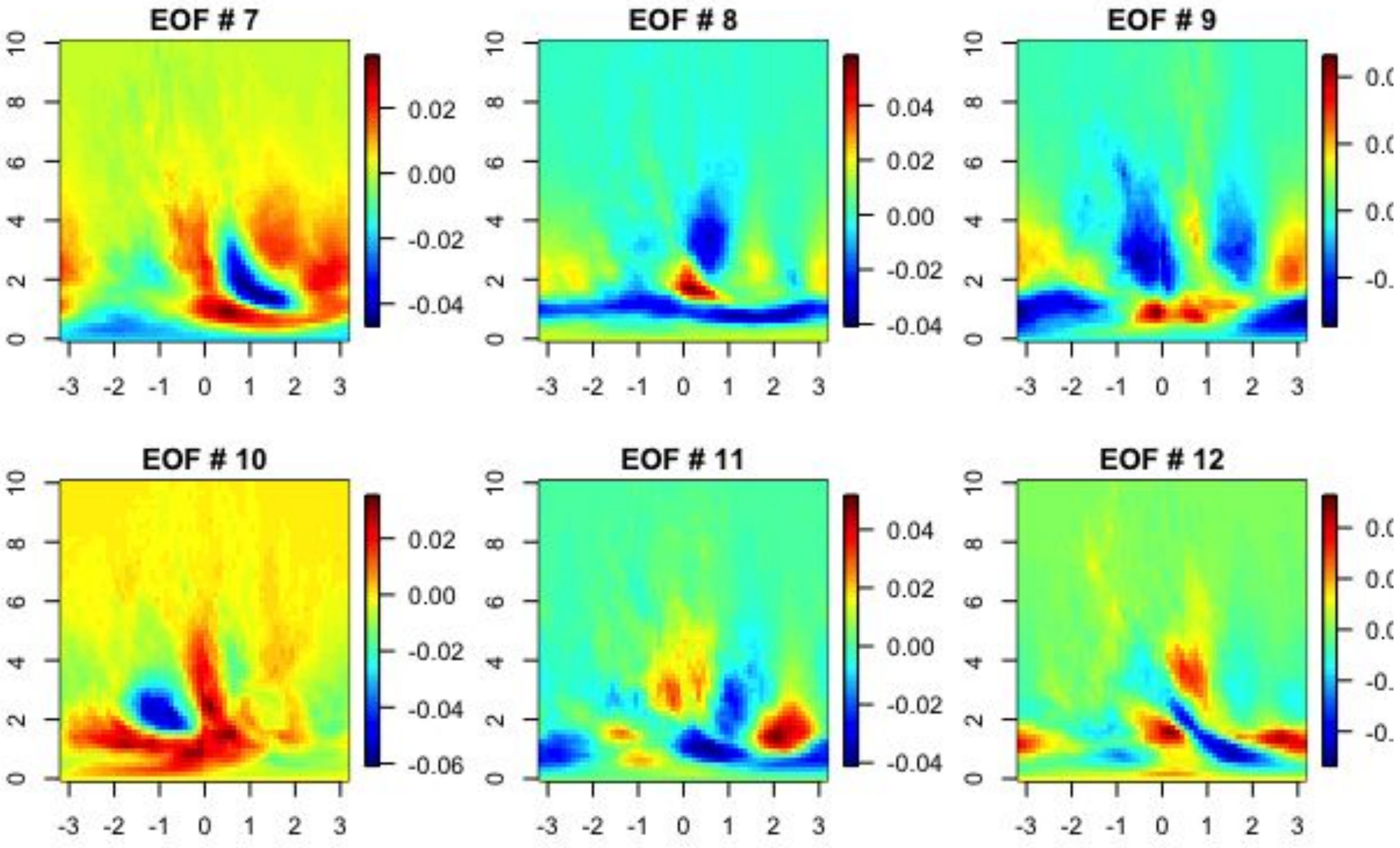
- Goal: calculate $\mathbf{Z} = \mathbf{UDV}^T$
  - Created a matrix by taking their cross product
  - svd package
  - **D** is the diagonal matrix of singular values that represent the square roots of the eigenvalues

Transform of Precipitation
Empirical Orthogonal Functions

Number of EOFs	Percent of Variance Explained
12	75%
23	80%
94	90%
231	95%




EOF # 1
EOF # 2
EOF # 3
EOF # 4
EOF # 5
EOF # 6
0.000
-0.005
-0.010
-0.015
-0.020
-0.025
-0.030
0.03
0.02
0.01
0.00
-0.01
-0.02
0.04
0.02
0.00
-0.02
0.03
0.02
0.01
0.00
-0.01
-0.02
-0.03


EOF # 7
EOF # 8
EOF # 9
EOF # 10
EOF # 11
EOF # 12

Fitting the Statistical Model - Random Forest

- Response
  - Principal Components

- Predictors
  - Radius of Maximal Winds
  - Pressure Deficit at Sea Level
  - Wind Direction (Lat/Long)
  - Storm Center (Lat/Long)
  - Distance to Coast

Fitting the Statistical Model - Random Forest

Predictor	Importance
Pressure Deficit at Sea Level	2814298.06
Storm Center - Polar Radius	1265974.08
Radius of Maximal Wind	1024219.61
Storm Center - Polar Angle	1000045.48
Wind Direction - Polar Angle	800550.80
Wind Direction - Polar Radius	757230.31
Distance to Coast	699380.21
White Noise	95819.21

Summarize

- ➢ Spatial Statistics is a fascinating, active area of research
- ➢ Careful consideration has to be given at each step of the analysis
- ➢ Statistical Learning is excellent for fitting fine-grained predictions

Thank You!!!

🌎 ⛈️ Dr. Kleiber 🌀 🌤️

📚 📊 Dr. Zaharatos 📈 ∑

Ⓡ👍 Dr. Julien 💯 ♕

😎 🙆 Laura Gooch & APPM Staff 😉 ⌛

🌟 🎓 Class of '21 PMD Cohort 🎉 👏