Comparison Of Nearshore Modeled Wave Heights Utilizing Meteorological Vs. Offshore Wave Conditions

Elizabeth Livermont,
Jon Miller, Thomas Herrington

WAVES 2011, Kohala Coast, Hawaii's Big Island

STEVENS Institute of Technology

The Center for Maritime Systems

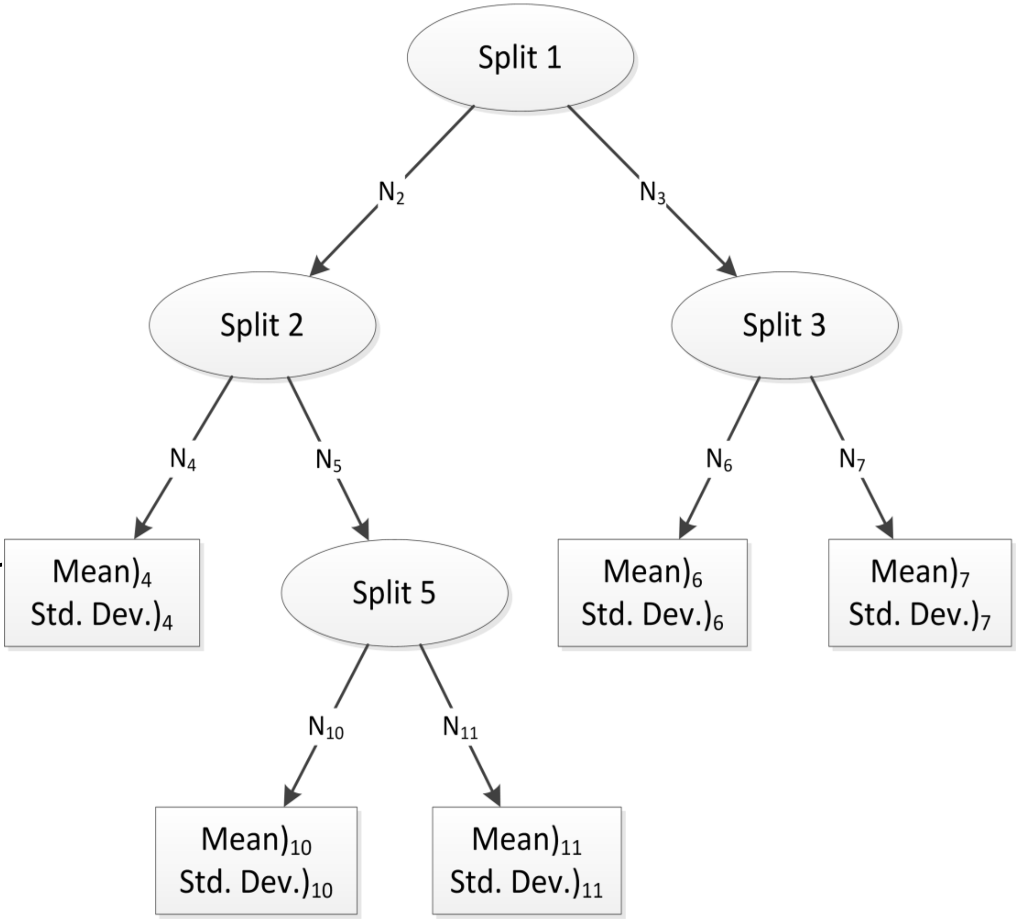
Stevens Institute of Technology Davidson Laboratory CPTAS Coastal Protection Technical Assistance Service

Summary of Motivation

- Coastal engineers need accurate high-resolution wave information:
  - Wave height,
  - Wave period,
  - Wave direction, etc.
- Physical processes are less important than accurate predictions
- Nearshore gauges are…
  - Expensive
  - Prone to failure
  - Relatively rare
- Numerical models require tradeoffs between computational speed and a complete physical description

Regression Tree Methodology

- Decision trees are a non-parametric technique
  - Formed by recursive partitioning of the data
  - Aim is to minimize mean square error of the dependent variable
- Utilize offshore buoy data (NDBC buoys) and onshore weather data for independent data



Summary & Conclusions

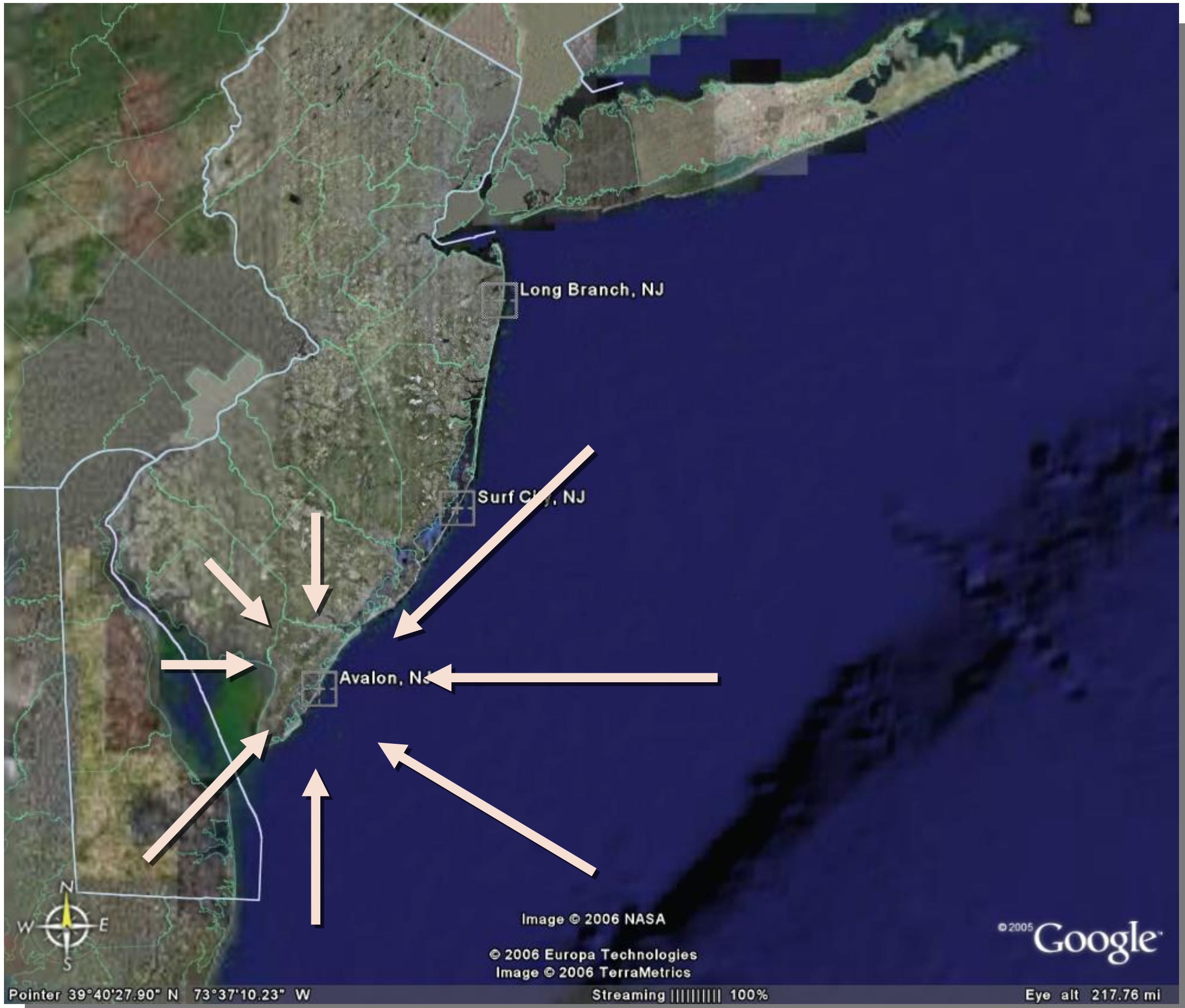
- Concurrent weather observations capture 45% of variability (according to R^2)
- Lagged weather observations capture 2.5% of variability
- Combining the two captures 75% of the variability
- Still does not match 93.5% based on offshore buoy data (44009 and 44025)
- Correlation does not imply causation

Overview

- Introduction
  - Motivation, Methodology and Conclusions
- Overview of Avalon, NJ Data
- Decision Tree Overview & “Best” Fit Model
- Local Weather Models
- Summary and Conclusions
- Future Work

Factors Influencing Wave Climate

- Oriented from northeast to southeast
- Broad continental shelf extending 150 km
- Southern extent surrounded by water
- Most waves approach perpendicular to the shoreline
- Highly refractive wave field, requiring nonlinear wave theories

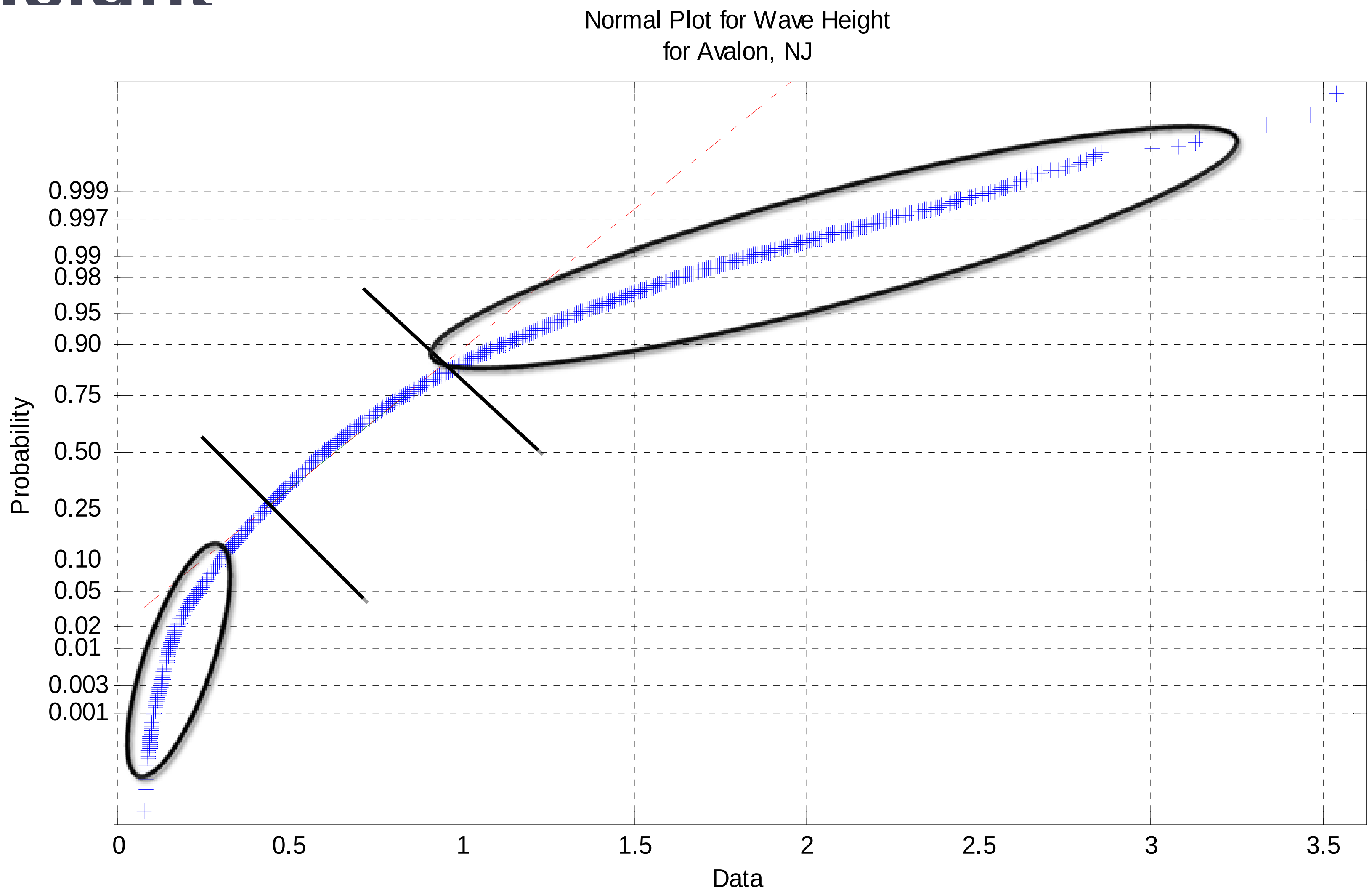


Volume of Observations

Month	Year									Grand Total
	1999	2000	2001	2002	2003	2004	2005	2006	2007	
1		569		742	575			705	744	3335
2		634		663	546			672	625	3140
3		676		699	579			738	611	3303
4		1		1	1			718	450	1171
5	118	715		349	716	608		744	725	3975
6	623	718	269	684	715	714		720	652	5095
7	553	735	736	596	742	91		744	675	4872
8	604	738	556	433	429		120	741	599	4220
9	645	705	717	544			720	703		4034
10	651	735	738	723			738	670		4255
11	658	716	715	605			720	719		4133
12	653	260	741	212			742	719		3327
Grand Total	4505	7202	4472	6251	4303	1413	3040	8593	5081	**44860**

Data from nearshore wave gauge in Avalon, NJ (operated by Stevens)

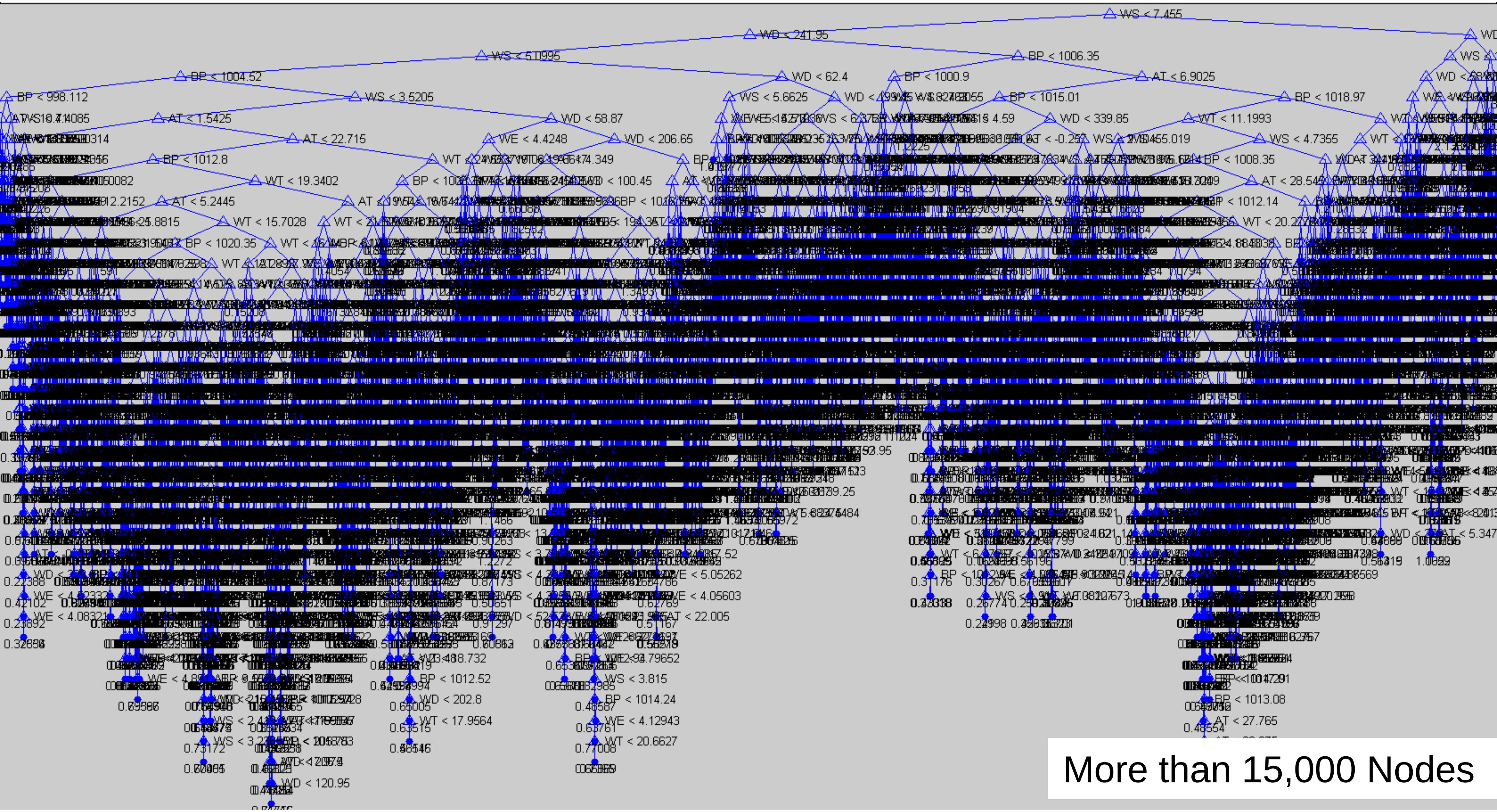
Deviation from Normality for Wave Height



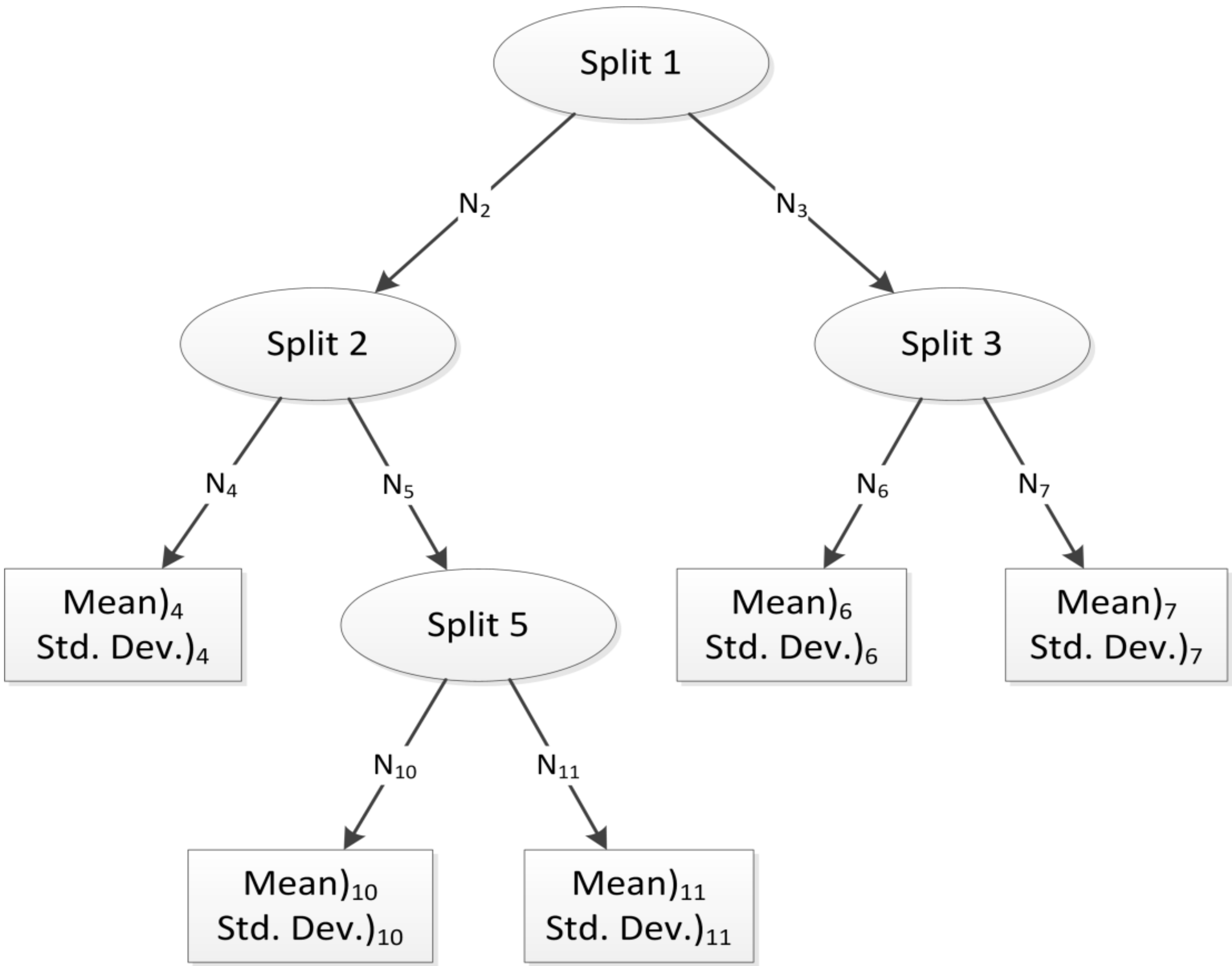
Data from nearshore wave gauge in Avalon, NJ (operated by Stevens)

DECISION TREE MODELS

Full Tree


WS < 7.455
WD < 241.95
WS < 5.1995
BP < 1004.52
BP < 1006.35
WD < 62.4
AT < 6.9025
BP < 1000.9
BP < 998.112
WS < 3.5205
WS < 5.6625
BP < 1015.01
BP < 1018.97
AT < 1.5425
WD < 58.87
WD < 339.85
WT < 11.1993
AT < 22.715
WE < 4.4248
WD < 206.65
WS < 4.7355
BP < 1012.8
BP < 1008.35
WT < 19.3402
AT < 28.545
AT < 5.2445
BP < 1012.14
WT < 15.7028
WT < 20.2278
BP < 1020.35
More than 15,000 Nodes

Decision Tree Theory



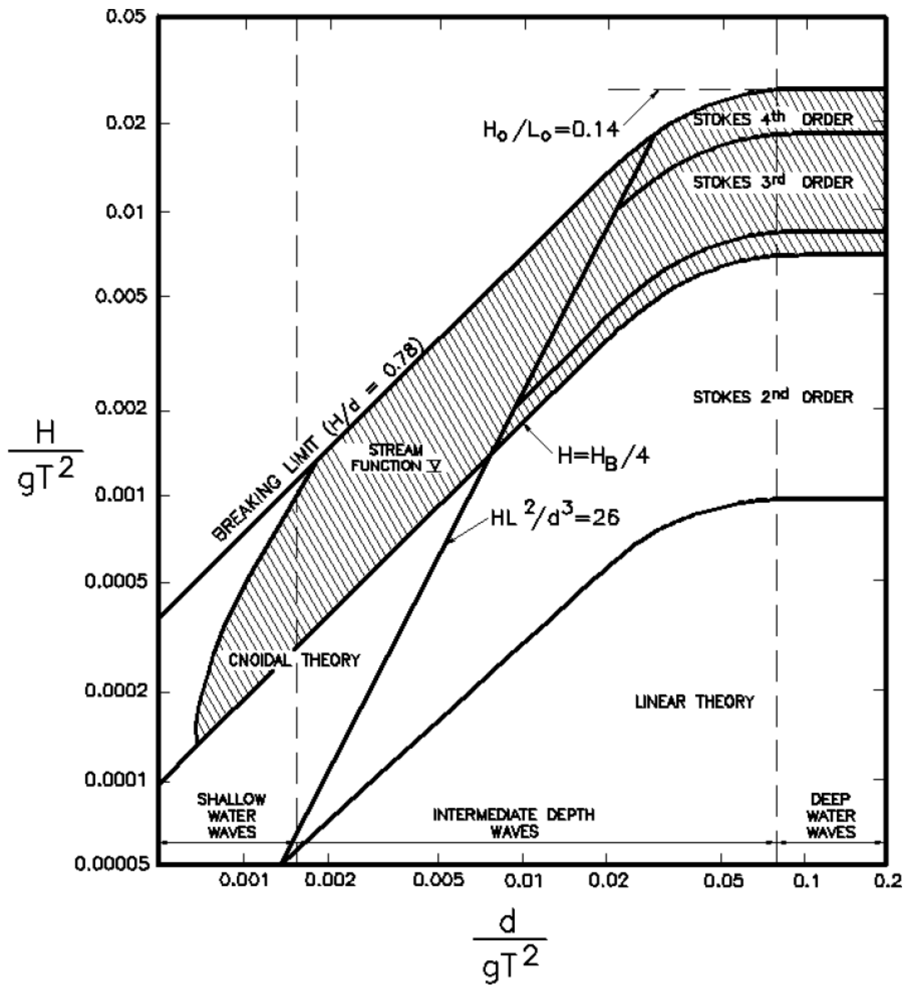
Advantages and Disadvantages

Advantages:

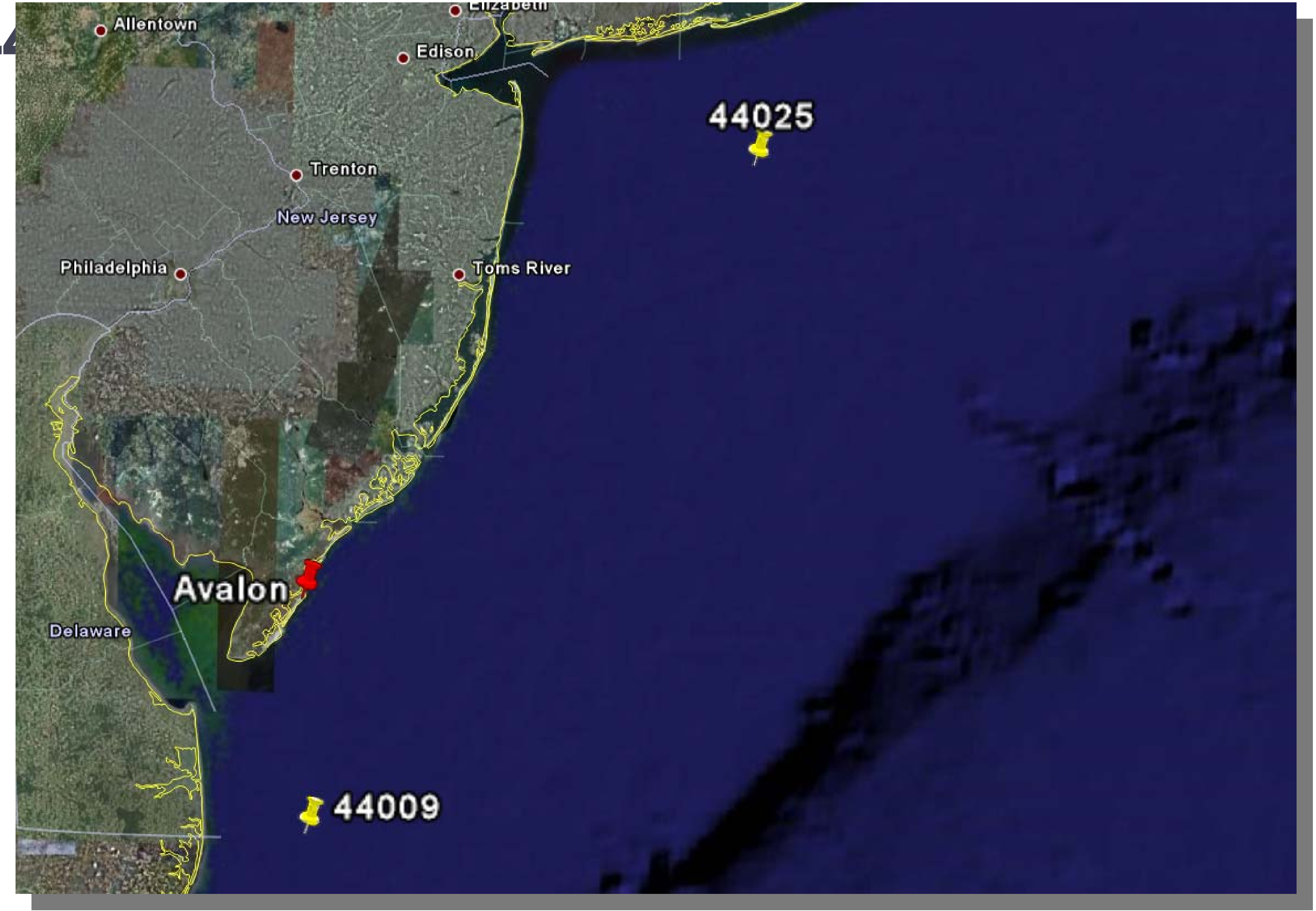
- Data-driven
- No assumption of normality
- Bridges existing wave theories

Disadvantages:

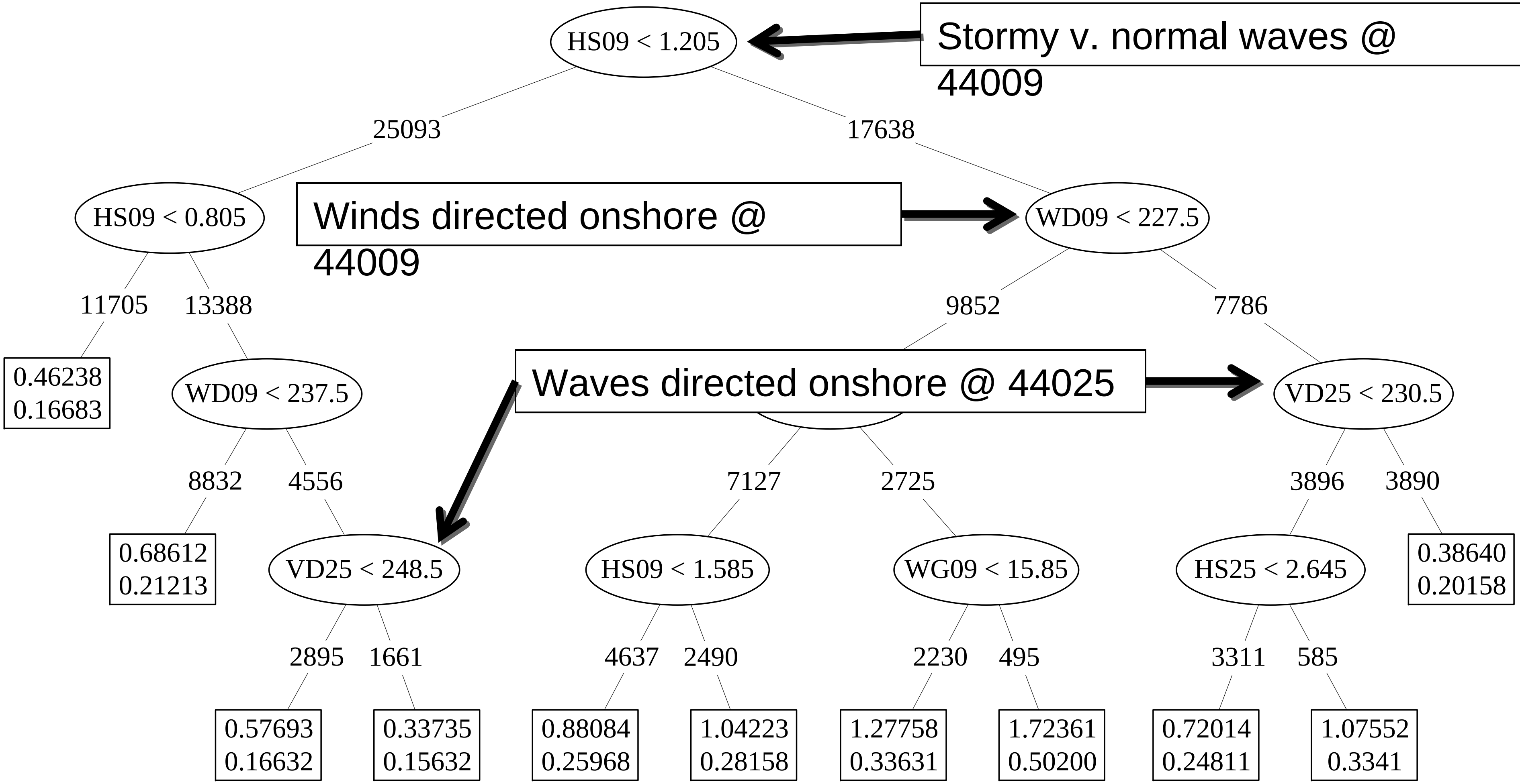
- No explicit link to physical processes
- Data- intensive
- Cannot predict outcomes which have not previously occurred



Model Utilizing Buoy 44009 and 44



Buoy 44009 and 44025



Buoy 44009 and 44025

Goodness-of-Fit Criteria	Value	Target
R^2	0.9366	1
Bias	0.0100	0
SI	0.0994	0
RMSE	0.0662	0
U	1.3212	>1

$$R^2 = 1 - \frac{\sum(x_i - y_i)^2}{\sum(x_i - \bar{x})^2}$$

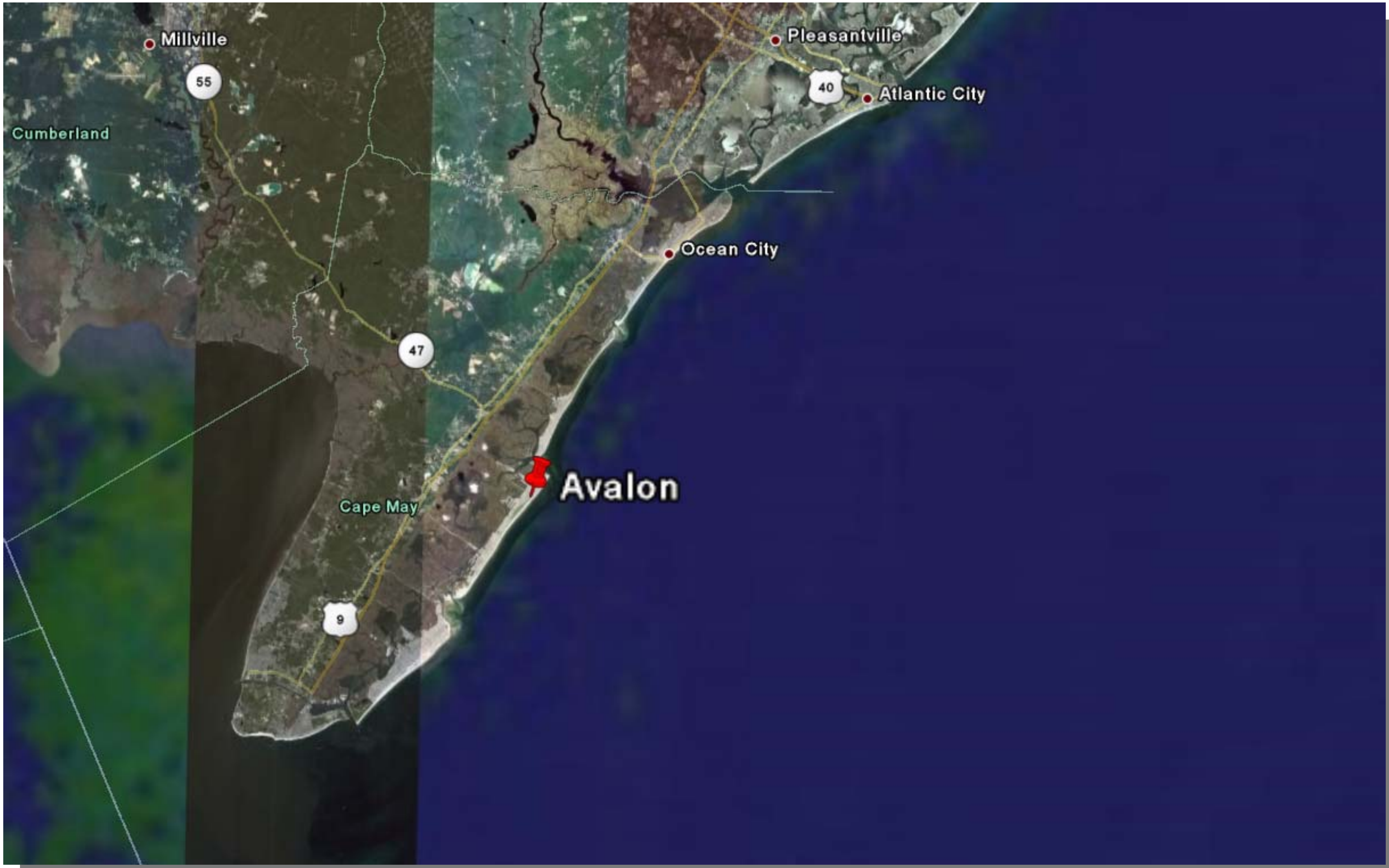
$$Bias = \bar{y} - \bar{x}$$

$$SI = \frac{1}{\bar{x}}\sqrt{\frac{1}{n}\sum[(y_i - \bar{y}) - (x_i - \bar{x})]^2}$$

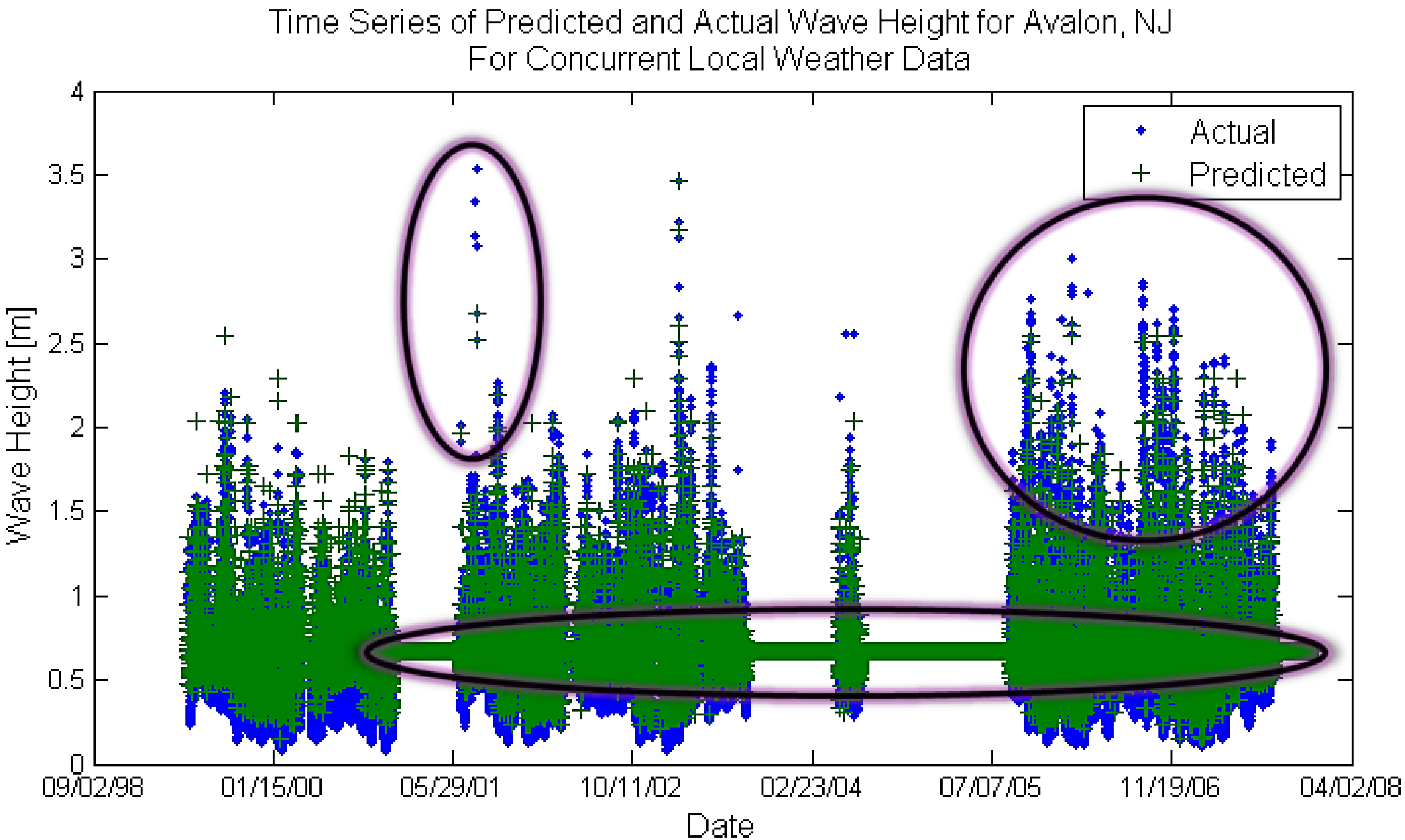
$$RMSE = \sqrt{\frac{1}{n}\sum(y_i - x_i)^2}$$

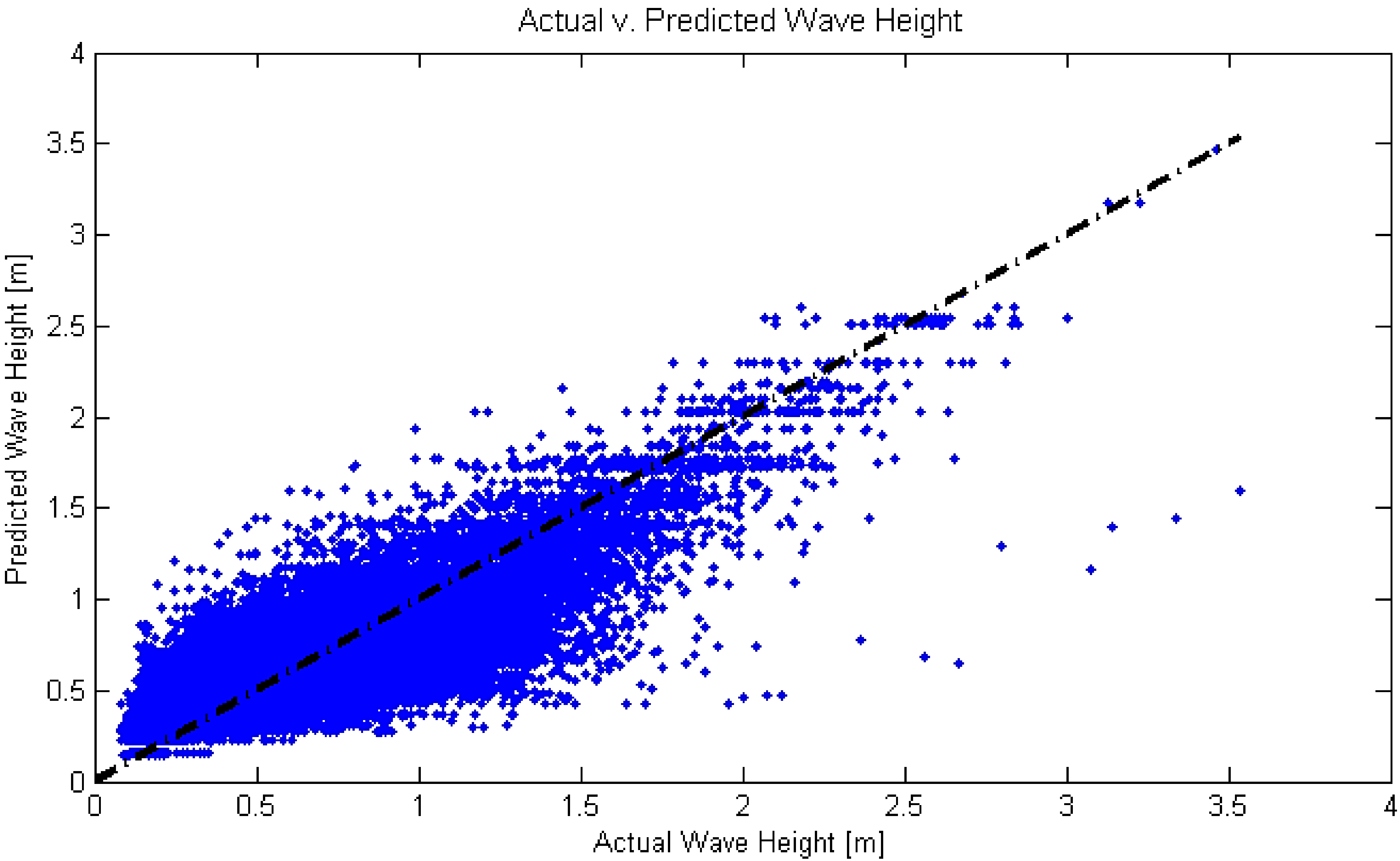

LOCAL WAVE MODELS

Local Weather Variables



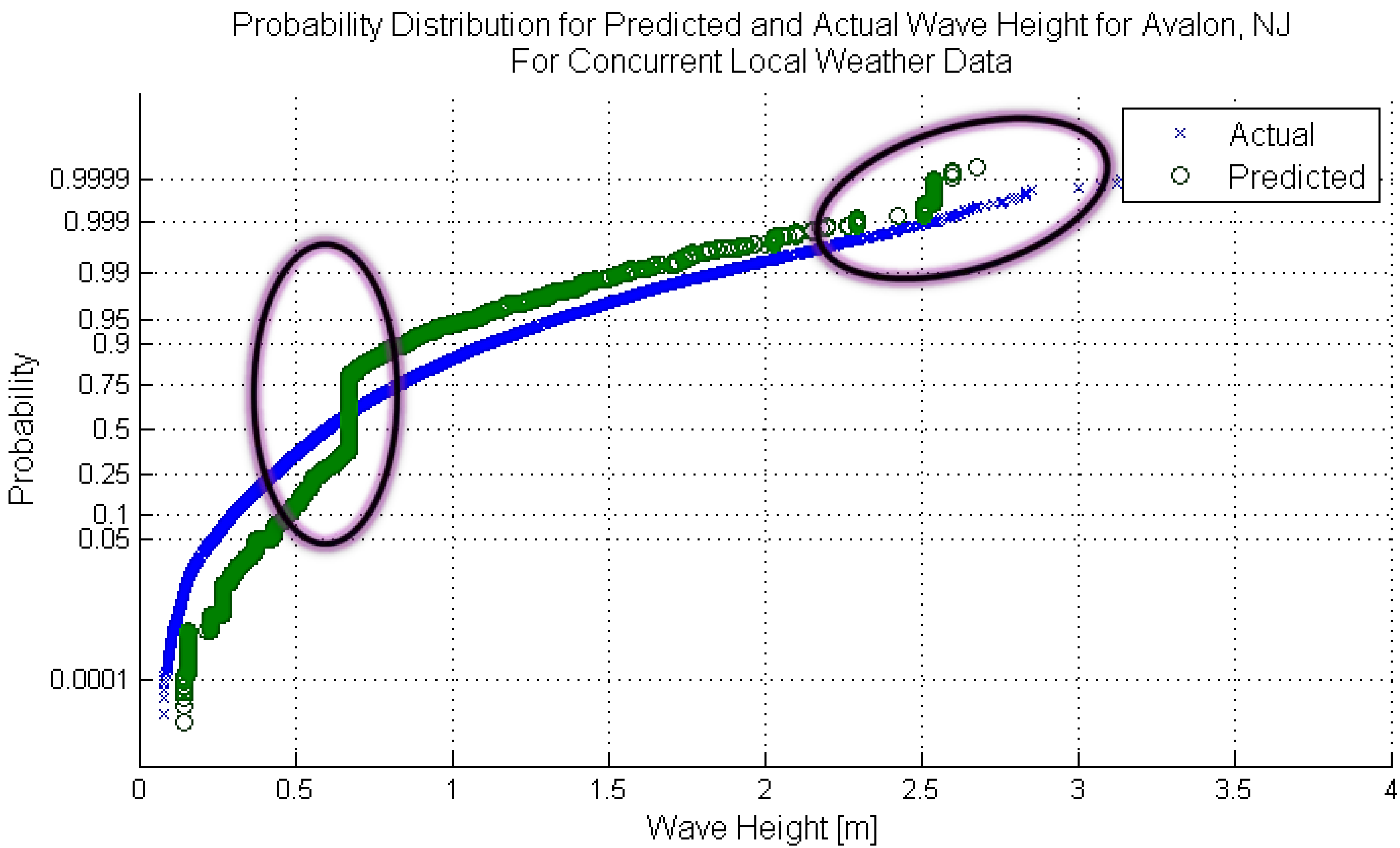
Predictions based on Concurrent Local Data





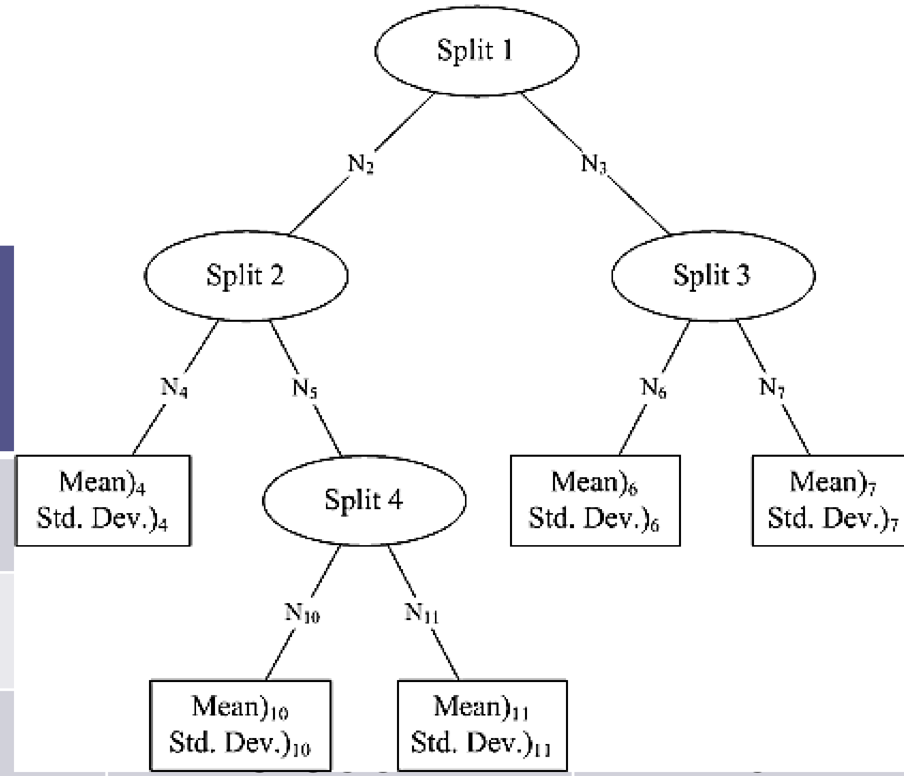
Actual v. Predicted Wave Height
Predicted Wave Height [m]
Actual Wave Height [m]
0
0.5
1
1.5
2
2.5
3
3.5
4

Comparison of Probability Distributions



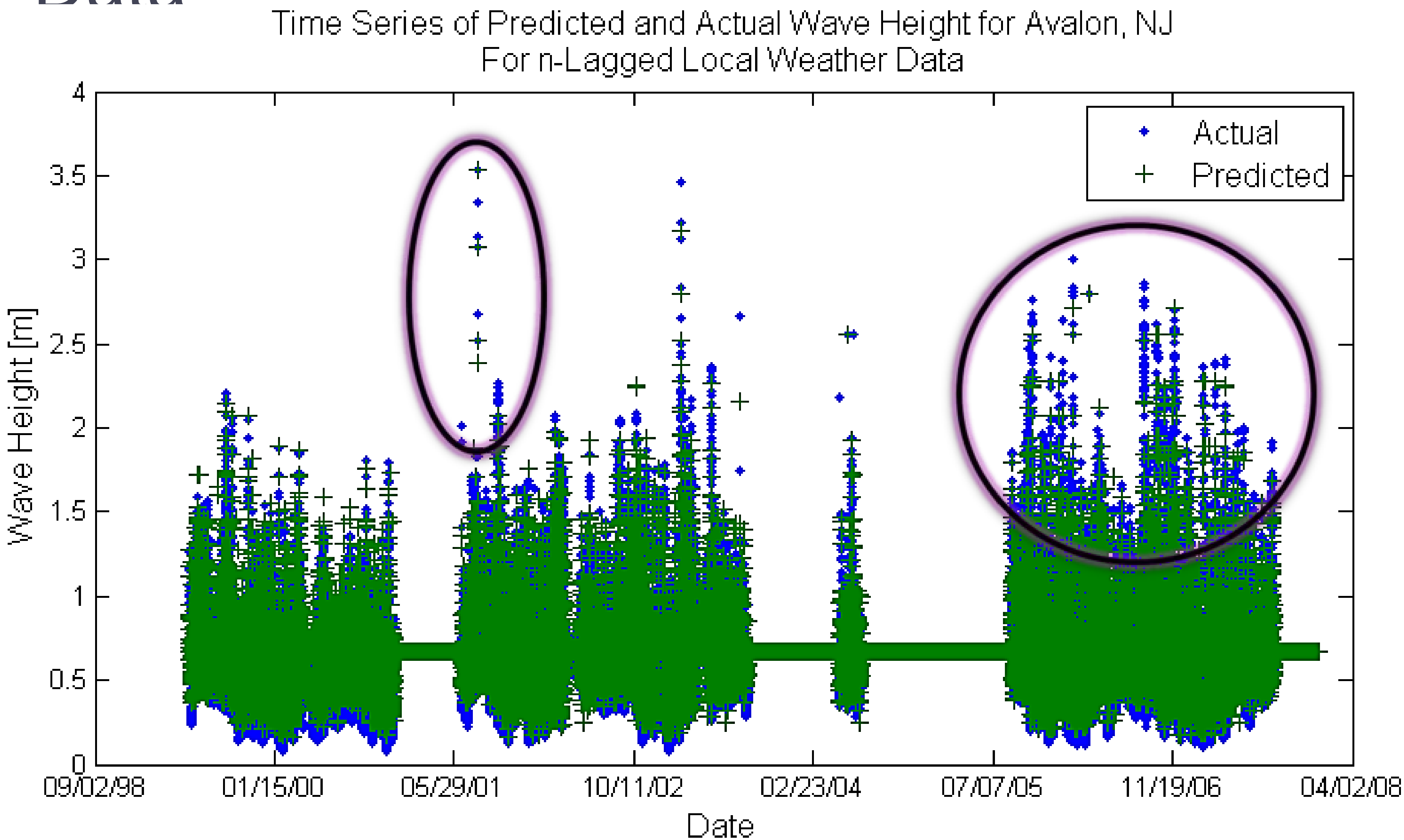
Comparison of Quantitative Measures

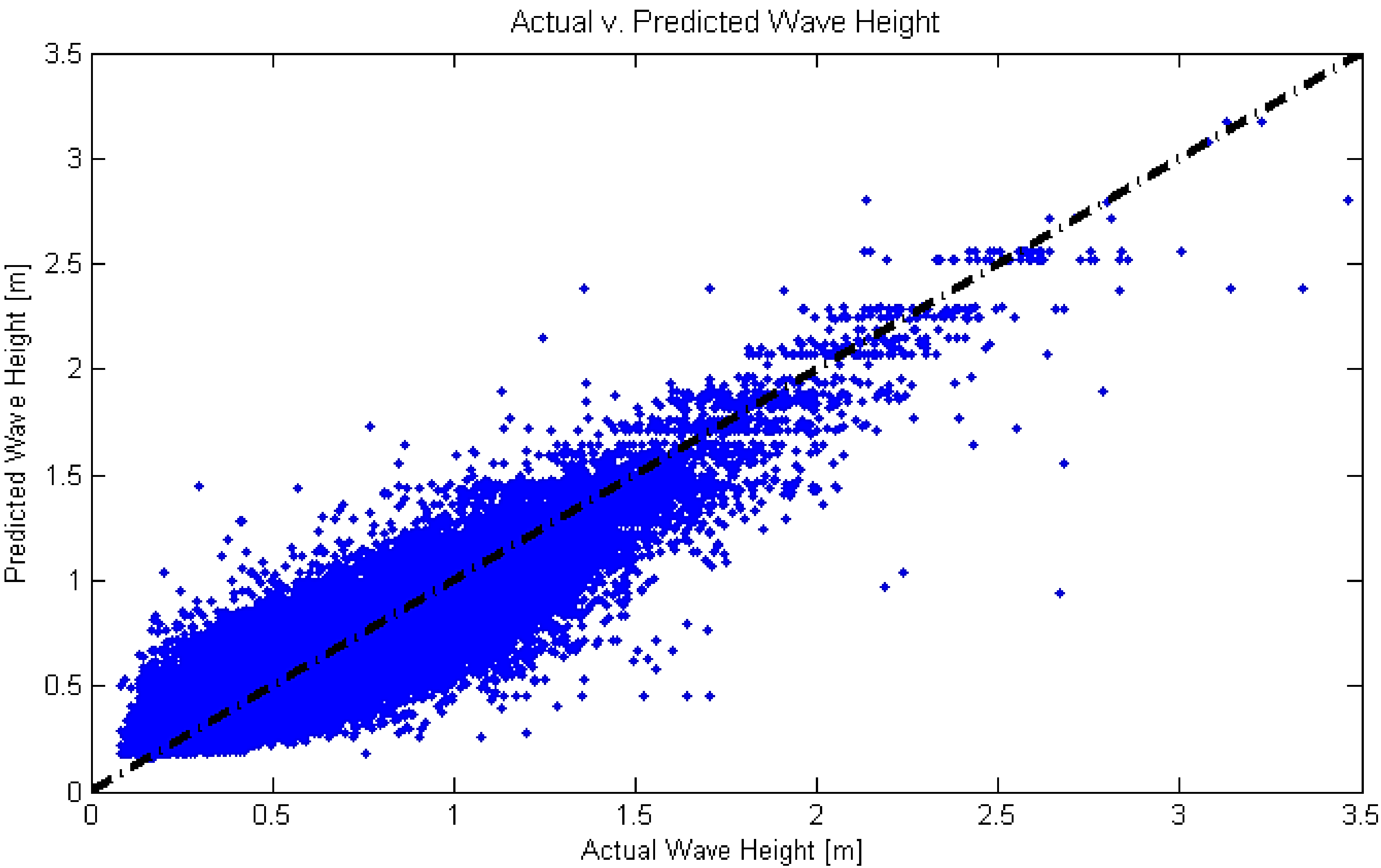
- Variable Importance:
  - WS, WD, BP, AT



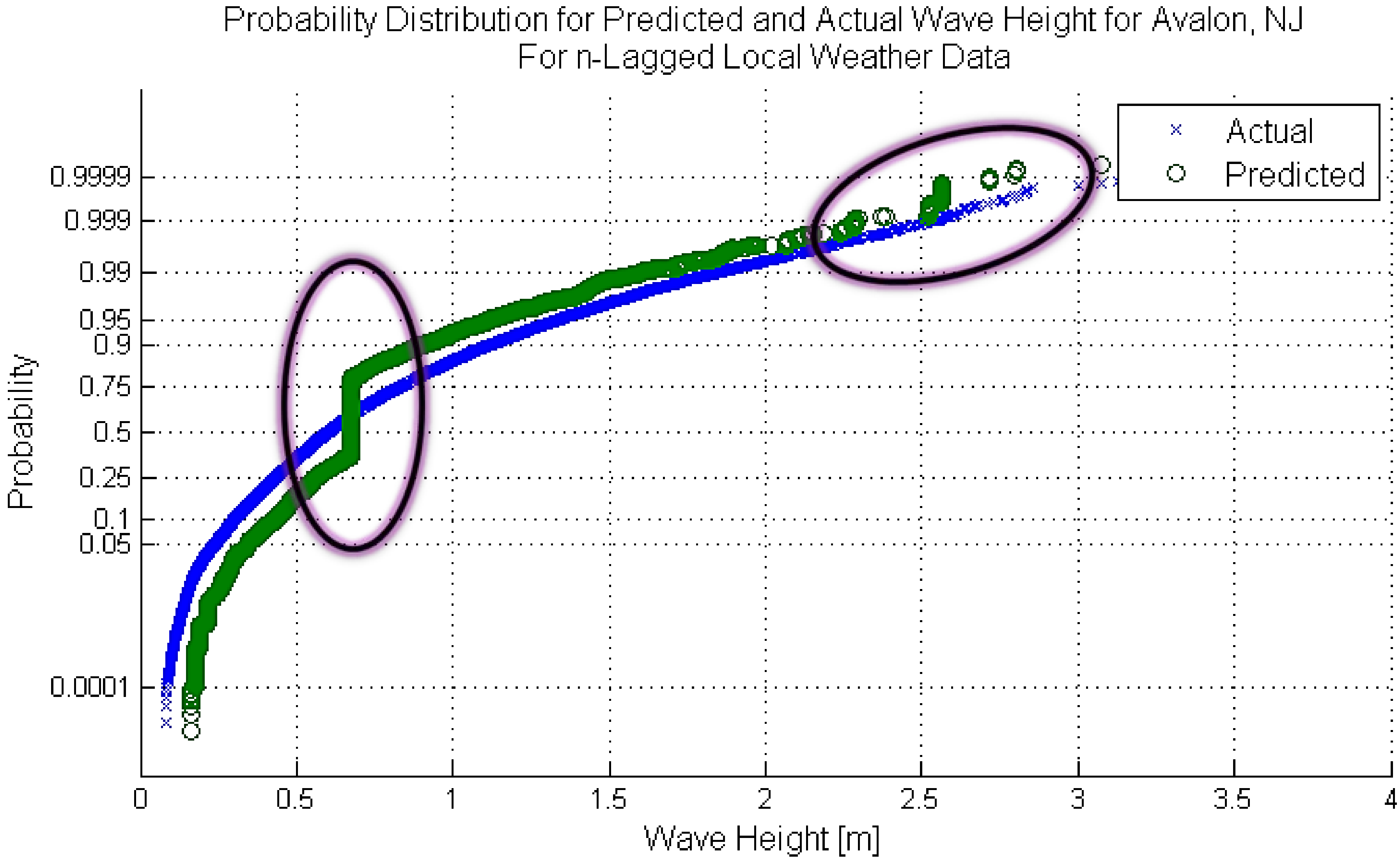
GOF Criteria	Value		
R^2	0.4518		
Bias	0.0012		
SI	0.2464		
RMSE	0.1652	0.0662	0
U	3.9987	1.3212	>1

Predictions based on Local Weather Data




Actual v. Predicted Wave Height
Predicted Wave Height [m]
Actual Wave Height [m]
0
0.5
1
1.5
2
2.5
3
3.5

Comparison of Probability Distributions



Comparison of Quantitative Measures

- Variable Importance:
  - WS_0, WD_0, BP_0, AT_0
  - WD_3 (13th Node)

GOF Criteria	Value	"Best Fit" Values	Target
R^2	0.7525	0.9366	1
Bias	9.89×10^{-4}	0.0100	0
SI	0.1829	0.0994	0
RMSE	0.1226	0.0662	0
U	2.9669	1.3212	>1


SUMMARY AND CONCLUSIONS

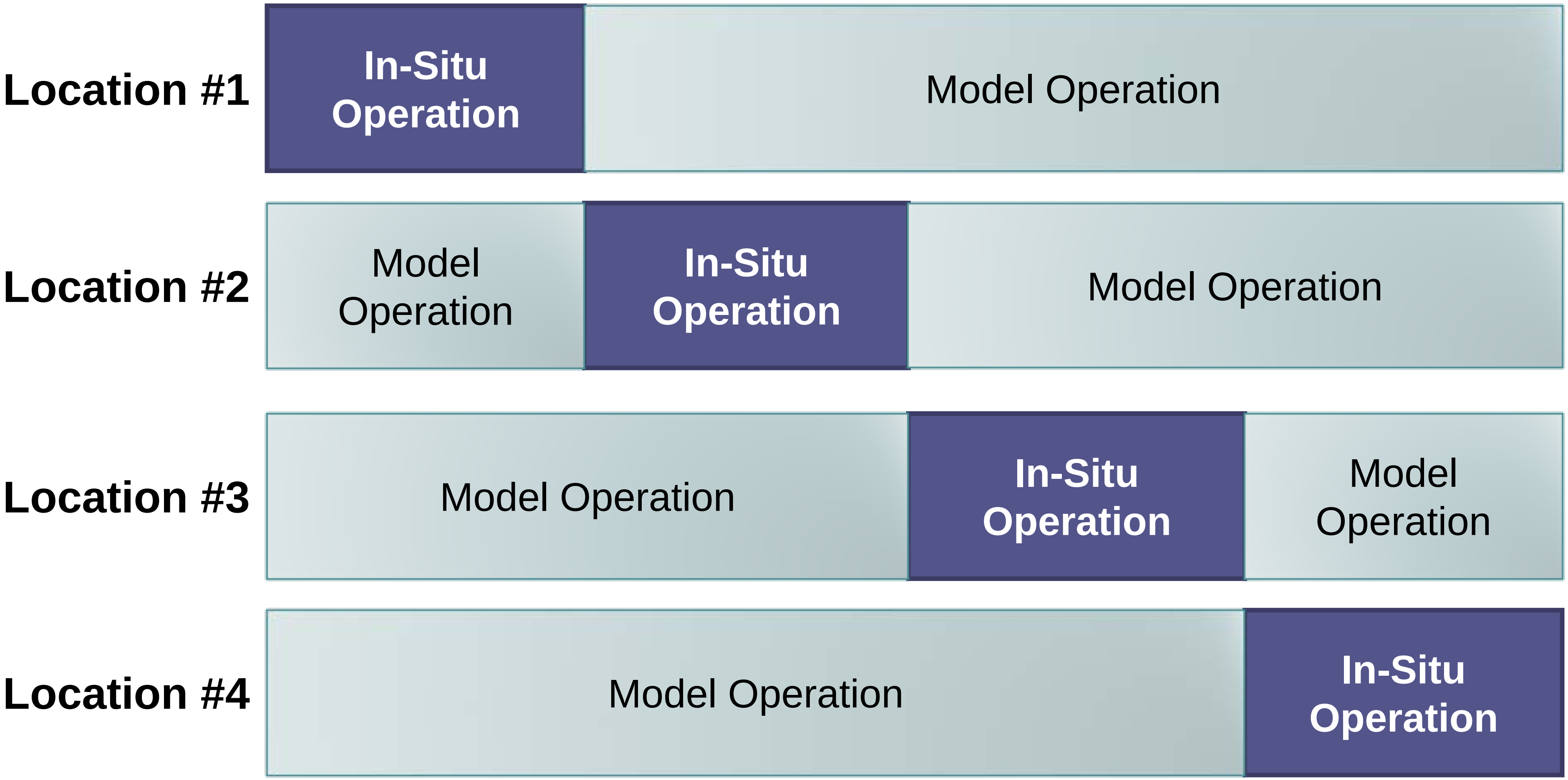
Summary and Conclusions

- Concurrent weather observations capture 45% of variability (according to R^2)
- Lagged weather observations capture 2.5% of variability
- Combining the two captures 75% of the variability
- Still does not match 93.5% based on offshore buoy data (44009 and 44025)
- Correlation does not imply causation

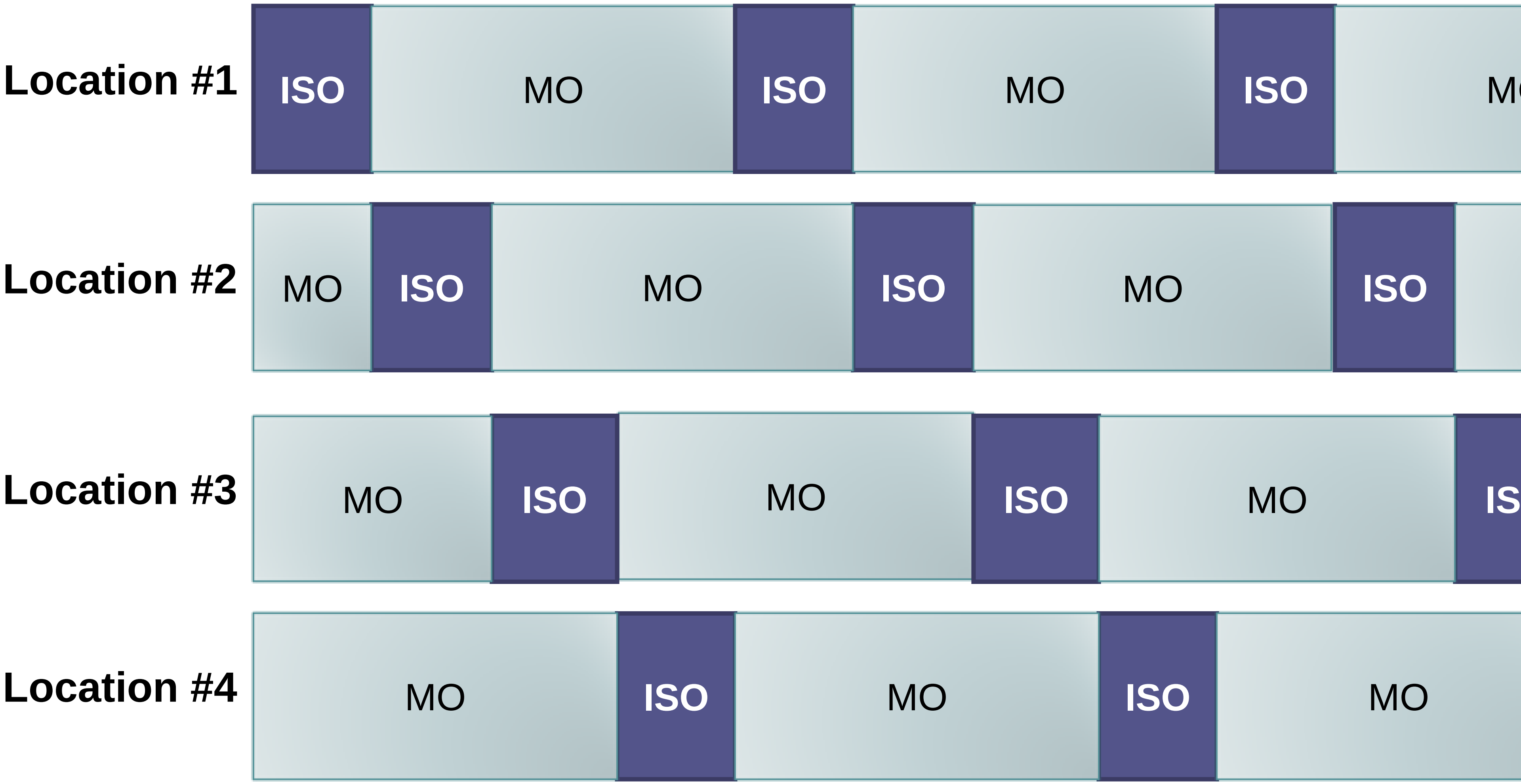
Future Work

- Operationalize decision trees
- Investigate use of climatic indices (NAO, ENSO, etc.)
- Incorporate dynamic/ numerical model data
- Explore alternate data-driven methods:
  - SOM, Neural Networks, etc.
- Apply to:
  - CODAR
  - Rip Tides
  - SEI (Storm Erosion Indices)
  - Shoreline Change
  - (Baseball)

Potential Future Experiment



Potential Future Experiment (Pt. 2)



All Lagged Wind Data

Goodness-of-Fit Criteria	Value
R^2	0.3327
Bias	0.0016
SI	0.2621
RMSE	0.1757
U	4.6790

Lagged Data without Concurrent Observation

Goodness-of-Fit Criteria	Value
R^2	0.0247
Bias	0.0014
SI	0.2916
RMSE	0.1955
U	4.7632