SAND2018-2865C

Application of Statistical Clustering to Characterize Ocean Sites for Wave Energy Converter Placement

PRESENTED BY

Bibiana E. Seng

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Abstract

The reliability of a Wave Ocean Converter (WEC) design is heavily reliant on the extreme ocean conditions at a chosen deployment site. Current analysis of ocean sites uses statistical contours to determine expected values for the significant wave height and energy period for a given time interval, known as a return period. The statistical contour method is currently run on a site-by-site basis; as such it can be difficult to detect overall patterns and expected values in large regions. My presentation will cover the application of a statistical clustering method that uses the Mean Shift algorithm to group together similar sites, and identify trends in the clusters. I will also discuss how these results were inferred. From there, I will discuss how this method can assist in choosing the deployment site for a WEC design, and how this could be further refined and researched in the future.

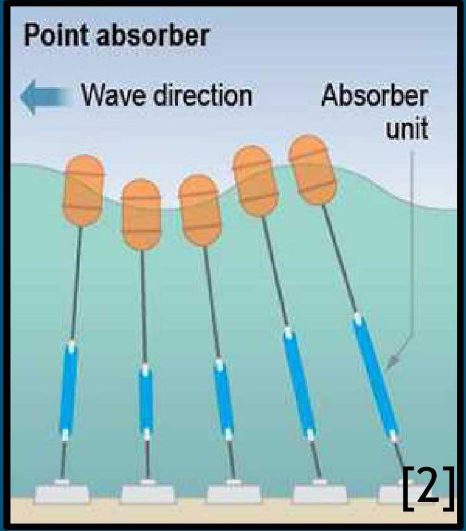
Outline

1. Characterizing Ocean States Using Extreme Condition Modeling (ECM) and Contours
2. Mean Shift Algorithm and Clustering
3. Preliminary Clustering Results
4. Discussion on Results and Their Application to Site Analysis
5. Future Work

Predicting Wave Conditions

- Waves need to be characterized for Wave Energy Converter (WEC) placement.
- Key Factors: Significant Wave Height (Hs) and Energy Period (Te)
  - Data collected from NDBC buoys
- **If we place a WEC in a suboptimal area, it could get damaged or destroyed by waves.**

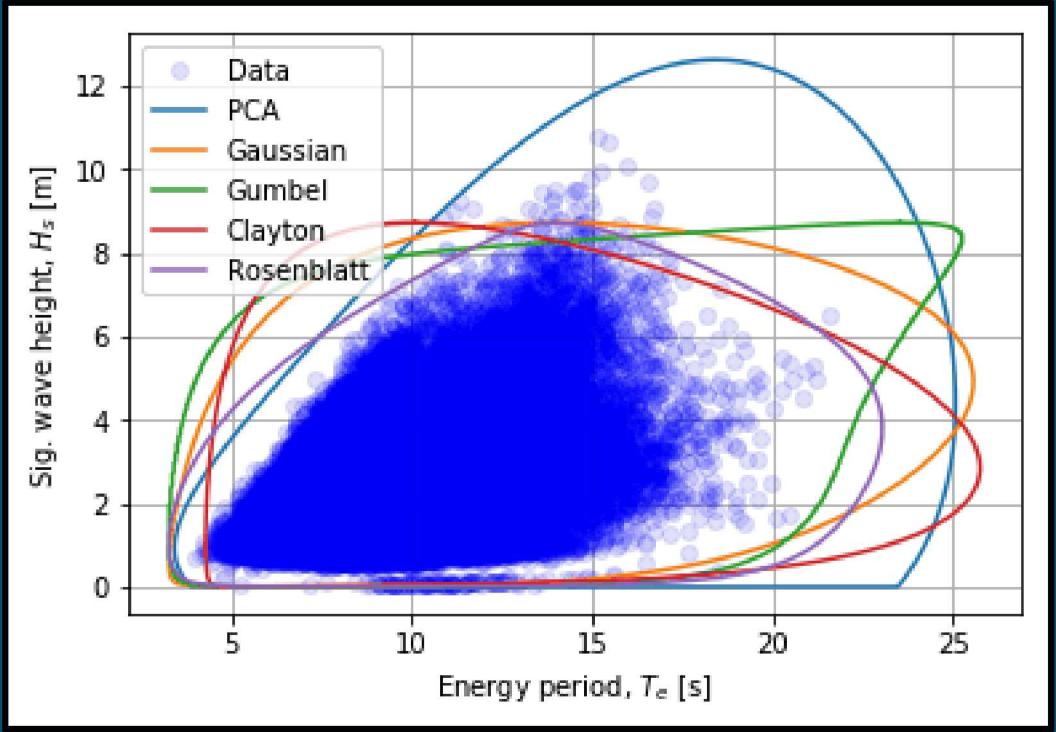
[1]



[2]

Characterizing Ocean Sites

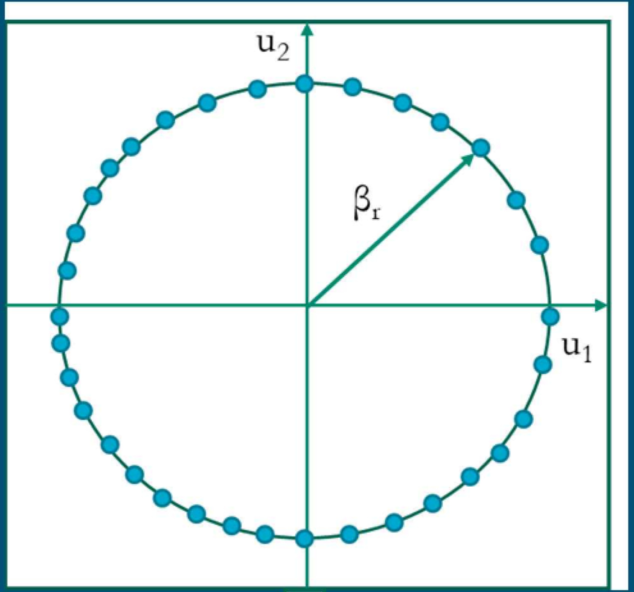
- Extreme Condition Modelling (ECM) calculates predicted extreme values in H_s and T_e for ocean sites in the future (10 to 100 years) by using statistical contours.
- Statistical contours [7] visualize how many points are contained within a given threshold.
- Useful for determining likelihoods of certain events.



Constructing a Statistical Contour

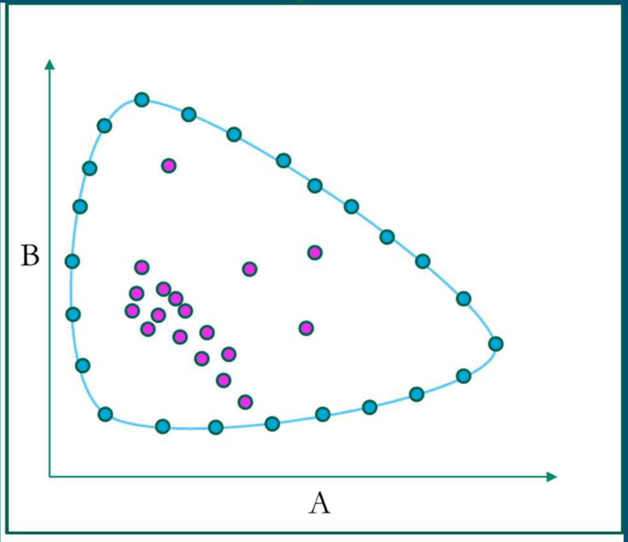
1. Clean and order and the data
2. Apply a chosen transformation to the data
   - Examples: PCA, Rosenblatt, etc.
3. Identify a *n*-sphere with radius of the reliability index β_r
   - The reliability index is calculated by evaluating the cdf at the reliability of the event we want to see.
4. Transform points along the *n*-sphere back into original space.

3.



[8]

4.



[8]

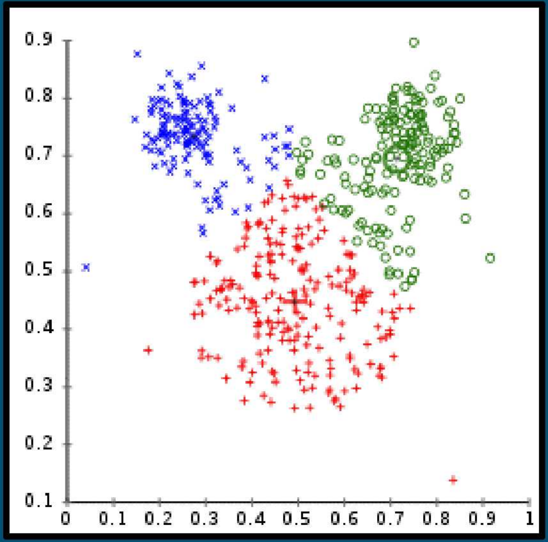
The Task

We use statistical contours to find extreme values at *specific sites*.

Can we use these statistical contours to find *general trends* in sea states?

Clustering Methodology

- Clustering groups similar data points together.
  - Many different types of clustering algorithms.
- Used Mean Shift algorithm [3,4] to cluster data.
  1. Fixes an interval around each point
  2. Finds the mean of the data in the interval
  3. Moves the interval to the calculated mean
- Assumed many outliers, nonspherical data set.



Mean Shift Calculation

1. Denote a kernel function.
   1. Commonly used is the Gaussian Kernel:

$$K(x_i - x) = e^{-c\|x_i - x\|^2}$$

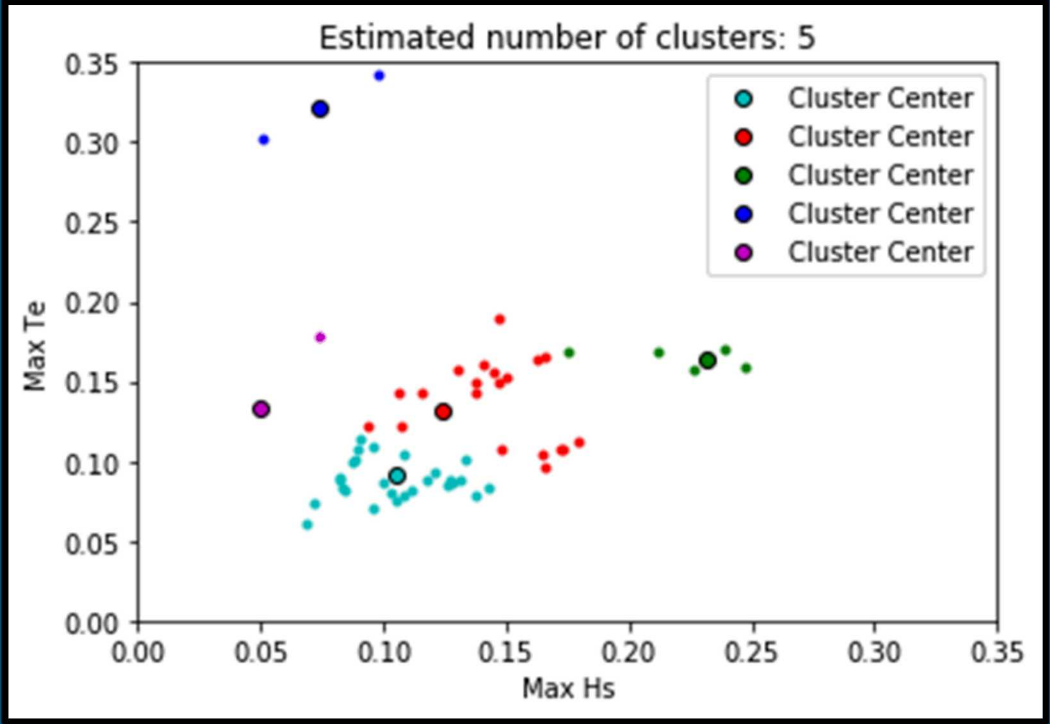
2. Define the weighted mean of the density as:

$$m(x) = \frac{\sum_{x_i \in N(x)} K(x_i - x) x_i}{\sum_{x_i \in N(x)} K(x_i - x)}$$

3. Set $x = m(x)$
4. Repeat until $m(x)$ converges.

Results

- Clustering methodology helps reveal potential trends in NDBC data.
- Input parameters are maximum calculated H_s and T_e from the contours.
- Preliminary results: approximately 3 types of expected extreme ranges.



Cyan: Gulf of Mexico, Great Lakes, East Coast, Hawaii.
Red: East Coast, Hawaii, West Coast
Green: West Coast
Blue: One East Coast, one Great Lakes
Magenta: West Coast

Validation and Discussion

- Determined validity of cluster centers using Silhouette Coefficient. [5]
  - Measurement of how similar data is to its own cluster compared to other clusters.

$$s(i) = \begin{cases} 1 - \frac{a(i)}{b(i)}, & if\ a(i) < b(i) \\ 0, & if\ a(i) = b(i) \\ \frac{b(i)}{a(i)} - 1, & if\ a(i) > b(i) \end{cases}$$

$$-1 \leq s(i) \leq 1$$

- Extreme outliers most likely in the Great Lakes and East Coast regions.
- Cluster centers were validated using the Silhouette Coefficient.
- Method reveals where similar extreme conditions could be, even across geographical boundaries.

Conclusion

- Clustering could demonstrate sea state groupings that could expand site choices considered for WEC placement.
- Clustering could also give insight into where WEC placement would be dangerous, given the clustering outliers.
- Future work will focus on refining algorithms and finding optimal metrics for clustering analysis.

This work was funded by the U.S. Department of Energy's Wind and Water Power Technologies Office. Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Questions?

References

1. National Oceanic and Atmospheric Administration's National Data Buoy Center. "Station 46022" url: http://www.ndbc.noaa.gov/station_page.php?station=46022

2. Office of Energy Efficiency & Renewable Energy. *Marine and Hydrokinetic Technology Glossary*. url: https://energy.gov/eere/water/marine-and-hydrokinetic-technology-glossary

3. D. Comaniciu and P. Meer. "Mean shift: a robust approach toward feature space analysis." *IEEE Transactions on Pattern Analysis and Machine Intelligence, vol 24, pp. 603-619*. (2002). Available: https://courses.csail.mit.edu/6.869/handouts/PAMIMeanshift.pdf

4. K. Fukunaga and L. Hostetler. "The Estimation of the Gradient of a Density Function, with applications in Pattern Recognition." *IEEE Transactions on Information Theory, vol 21*. (1975). Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1055330\&tag=1

5. P. J. Rousseeuw. "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis." *Journal of Computational and Applied Mathematics, vol 20, pp. 53-65*. (1987). Available: http://www.sciencedirect.com/science/article/pii/0377042787901257?via\%3Dihub

6. S. Lloyd. "Least Squares Quantization in PCM" *IEEE Transactions on Information Theory, vol 28, pp. 129 – 137*. (1982). Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1056489

7. Drew, B., Plummer, A., Sahinkaya, M., (2009). Alternative environmental contours for marine structural design: A comparison study. Proceedings of the Institution of Mechanical engineers, Part A: Journal of Power and Energy 223.8, 887–902

8. Eckert-Gallup, A. "Development of Methods for Generating Environmental Contours of Extreme Sea States & for Characterizing Their Uncertainty". (2017). Presented at Sandia National Labs Water Power Brown Bag.