Using Machine Learning Methods to find a Representative and Conservative Set of Conditions for Fatigue Analysis of Offshore Wind Turbines

Sam Kanner PhD, R&D Lead

Bingbin Yu PhD, Global Performance Lead

PRINCIPLE POWER

EERA Deepwind '19 Conference
Trondheim, Norway
16 Jan 2019

skanner@principlepowerinc.com

Outline

- **Motivation**
- **Algorithm**
  - LASSO
  - Gradient Descent
  - Clustering
- **Metocean Data**
- **Results**
- **Conclusion**

Independent "Dimensions" of Fatigue

Aerodynamics: 2-D+
(Wind speed, wind direction, [turbulence])

Sea-surface: 3-D+
(Per spectrum: Wave height, wave period, wave direction)

Subsurface: 2-D
(Current speed, current direction)

Anything else?

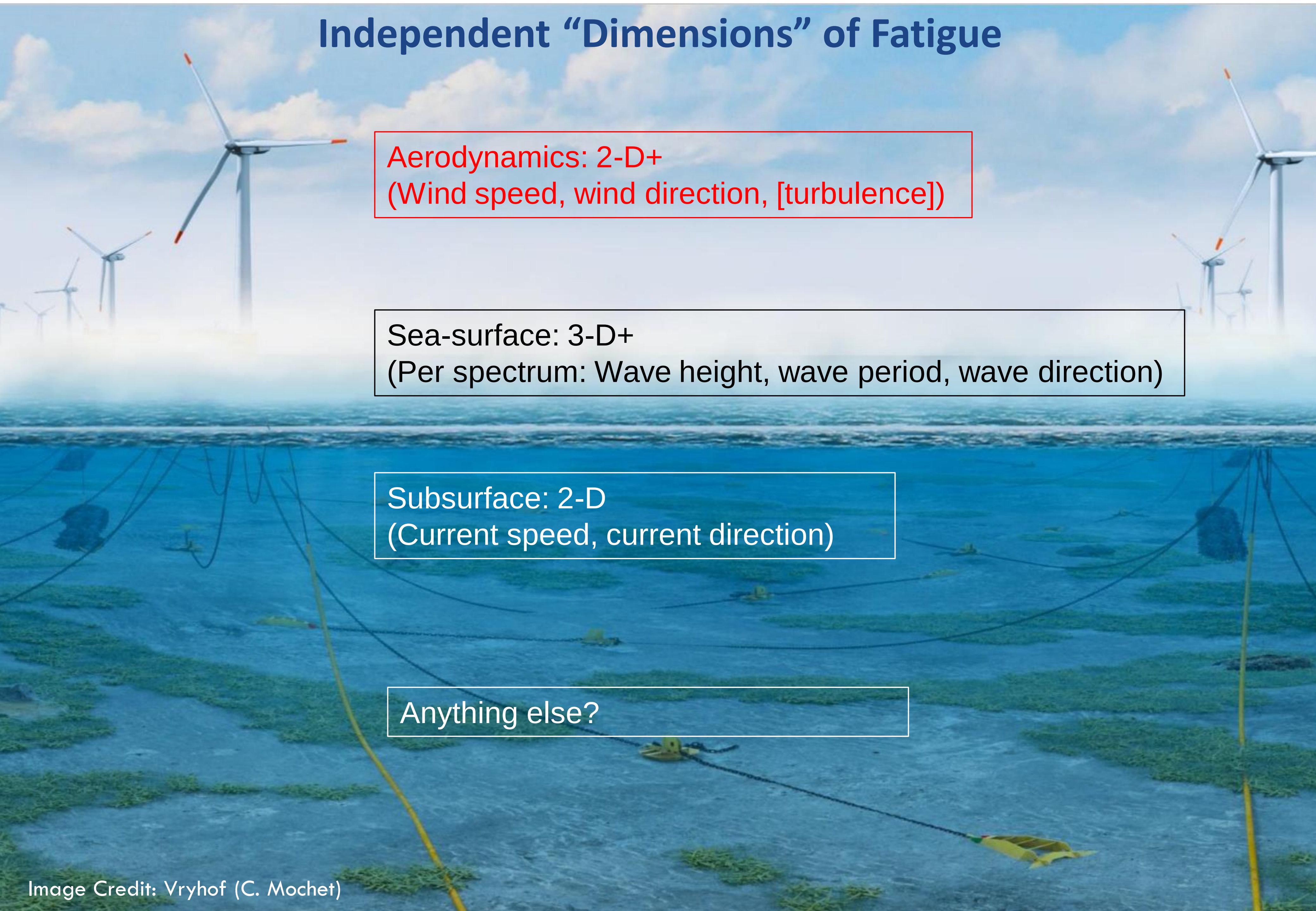


Image Credit: Vryhof (C. Mochet)

Estimation of Fatigue Life of an Offshore Structure & Mooring

DNV-OS-J103, DNV-OS-E301

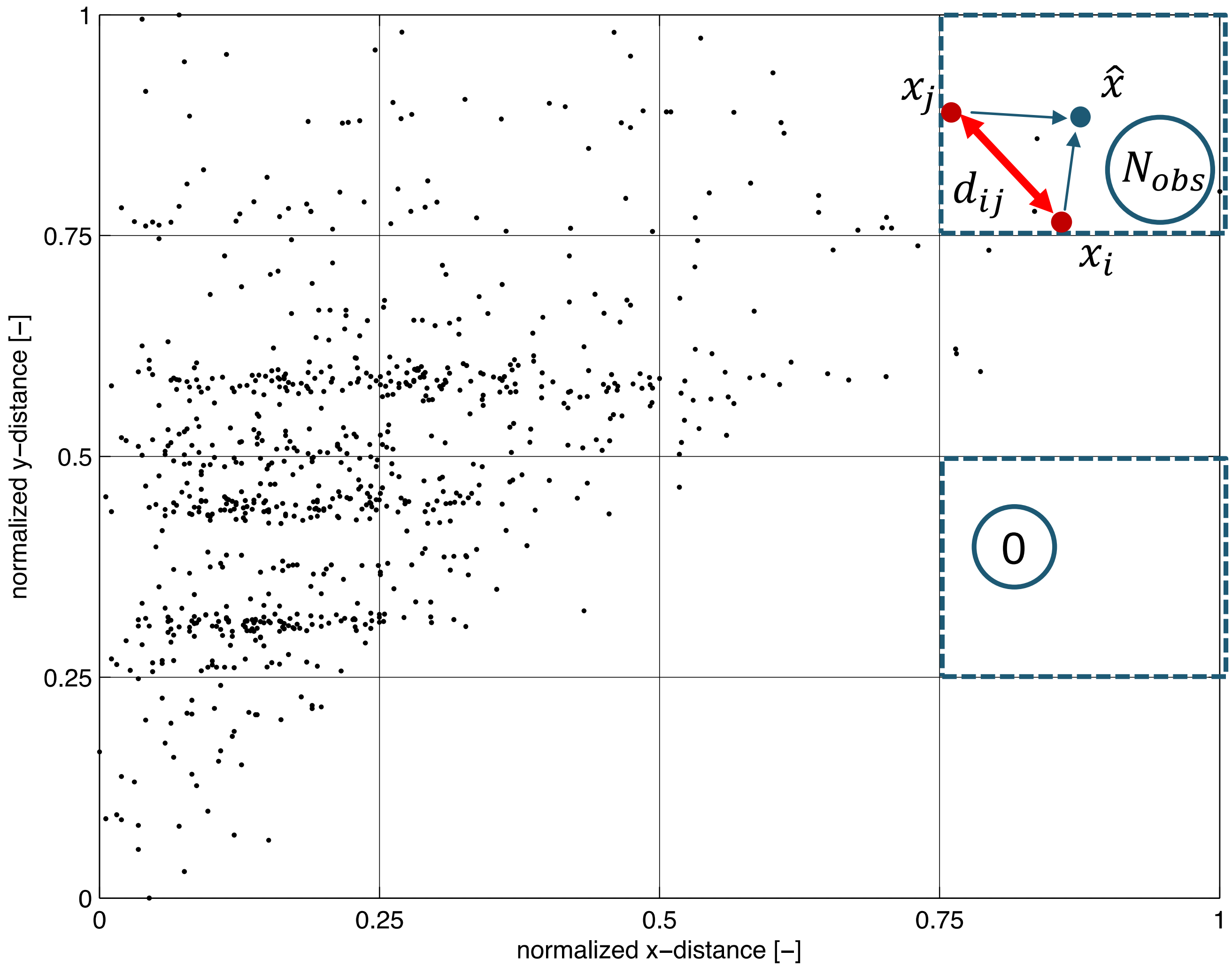
(most accurate and computationally intensive procedure)

1. **Numerous specific environmental conditions (load cases)**
    1. Wave direction: 8-12 bins
    2. Wave height/period: 10-50 bins
    3. Wind speed/direction: ? Bins
    4. Current speed/direction: ? bins
2. **Time-domain modelling tool**
3. **Rainflow counting method to assess range of "sensor" (e.g., tension in mooring line, principal stress at specific location)**
4. **Estimate damage from each load case using properties of material (e.g., S-N, T-N curve)**
5. **Estimate fatigue life from sum of damage, taking into account the probability of occurrence of each load case during design life of structure**

Dowling SD, Socie DF. *Simple rainflow counting algorithms.* Int J Fatigue 1982;4:31–40.

B. Yeter, Y. Garbatov, C. Guedes Soares, *Evaluation of fatigue damage model predictions for fixed offshore wind turbine support structures,* Intl J. Fatigue, 2016; 87:71-80

Traditional clustering method (visualized in 2D)



$$P_{bin} = \frac{N_{obs}}{N_{tot}}$$

Conservative?

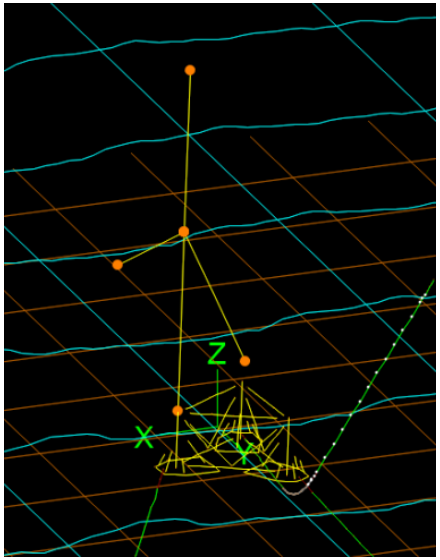
Representative?

PRINCIPLE POWER

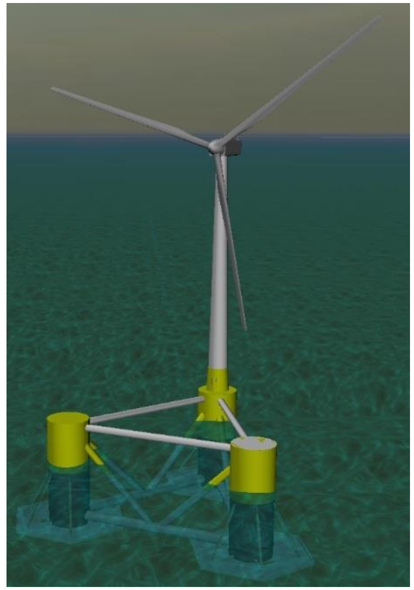
BEGIN:
ALGORITHM

Proposed (Machine Learning-based) Algorithm

1. **Load *p*-dimensional set of multi-decadal environmental conditions**
   i. Normalize data to all lie in [0,1].
2. **Initialize with a "representative" set of *M* clusters (or bins)**
   i. Modified Maximum Dissimilarity Algorithm (MDA-based) clustering method to associate all observations with closest cluster
3. **Run time-domain simulations to estimate fatigue damage**
   i. OrcaFAST coupled aero-hydro-mooring simulations
   ii. OrcaFlex: Time domain solver including first and second-order hydrodynamics (from WAMIT) and instantaneous mooring force
   iii. FAST: Open-source BEM tool with linearized structural dynamics
   iv. In-house rainflow counting algorithm



Representative



Kanner, S., Yu, B., Aubault, A., Peiffer, A., 2018. Maximum Dissimilarity-Based Algorithm for Discretization of Metocean Data into Clusters of Arbitrary Size and Dimension OMAE2018-77977

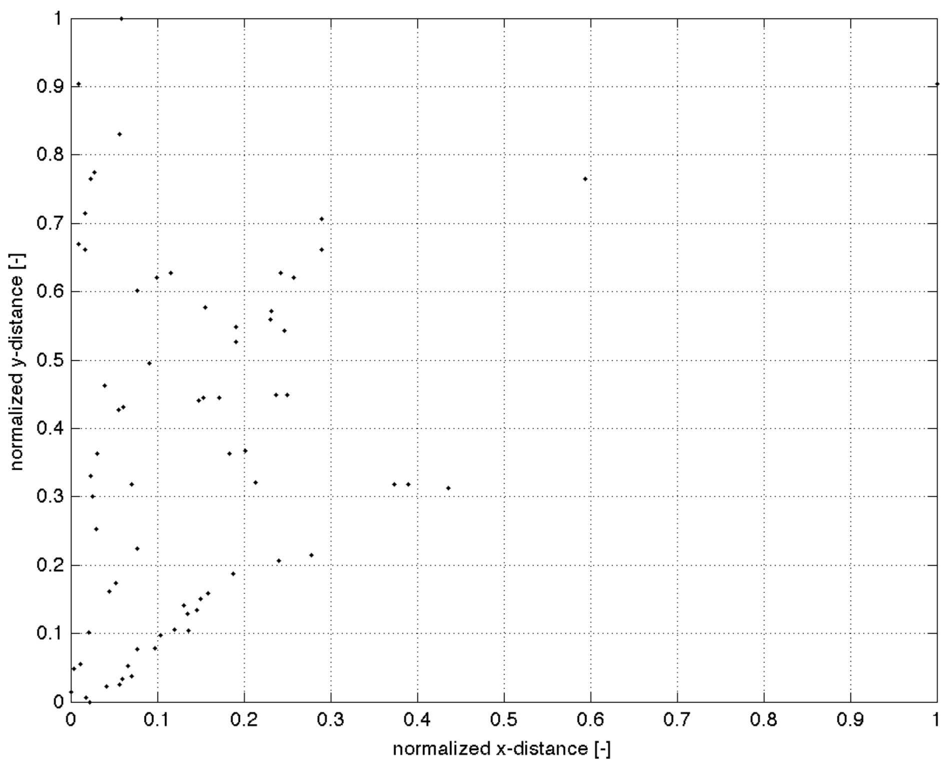
Proposed (Machine Learning-based) Algorithm (cont.)

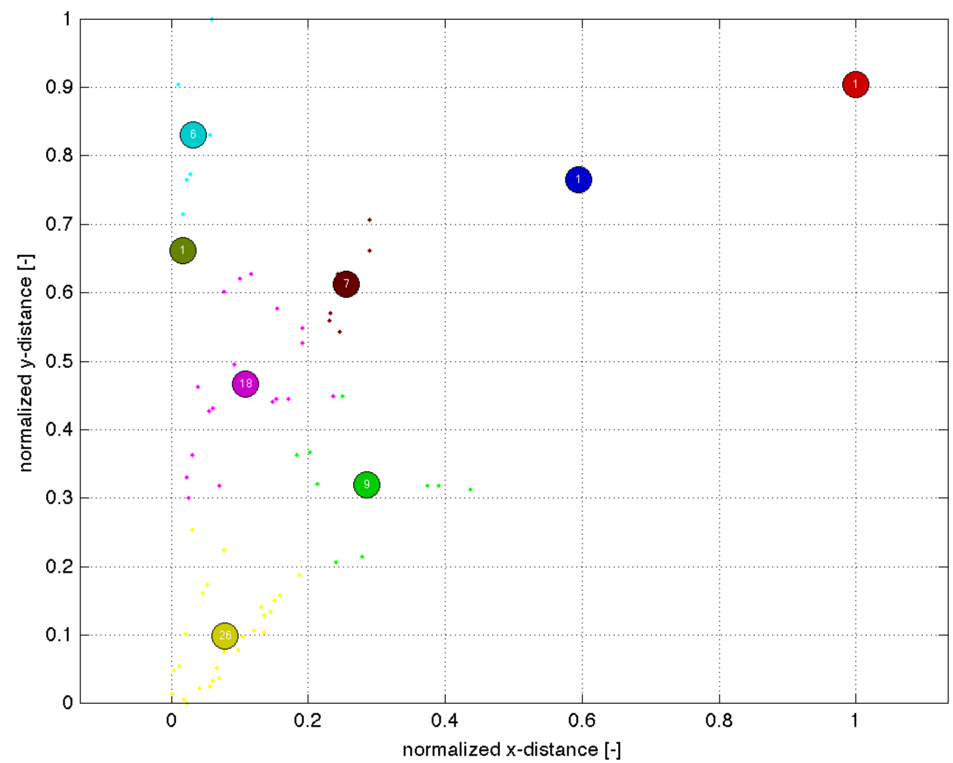
4. **Choose a set of predictors to estimate how environmental conditions effect fatigue damage**
   i. Damage $= H_S + H_S^2 + H_S^3 + T_P + T_P^2 + T_P^3, \ldots, + H_S \cdot T_P$?
5. **Run regularized linear regression analysis: Least Absolute Shrinkage and Selection Operator (LASSO)**
   i. Come up with a 'constrained' model on how fatigue damage depends on predictors
6. **Use gradient ascent algorithm to determine direction of maximum damage**
   i. Pick step-size to determine speed of approach to maxima — **Conservative**
   ii. Select clusters that are in 'high-damage' areas and spawn *N* new clusters that may be more damaging — **Conservative**
   iii. Keep number of clusters *M* constant by creating (*M*-*N*) new "representative" clusters using MDA-based method. — **Representative**
7. ***Re-cluster all observational data using M new clusters.***
8. **Iterate (steps 3-7) to try and find a conservative value of fatigue damage**

PRINCIPLE POWER

Step 1-2: Modified Maximum Dissimilarity Based Algorithm

$$d_{tol} = 0.17$$





$$M = 8$$

Representative

PRINCIPLE POWER

Step 4-5: Least Absolute Shrinkage and Selection Operator (LASSO)

1. **Try to find the best fit:** $\hat{y}_i = \sum_{i=1}^{M} \sum_{j=1}^{N_x} x_{i,j}\beta_j = \sum_{j=1}^{N_x} x_i^T \beta$ $\quad x_i = H_{s,i}, H_{s,i}^2, H_{s,i}^3, \theta_{w,i}, \theta_{w,i}^2, \theta_{w,i}^3, \ldots$

2. **For a given λ (λ > 0), LASSO algorithm attempts to solve the problem**

$$\min_{\beta,\beta_0} \left(\frac{1}{2N} \sum_{i=1}^{M} (y_i - \beta_0 - x_i^T \beta)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \right)$$

Calculated (weighted) damage at i^{th} observation

Regression coefficients

Regularization parameter to penalize 'overfitting'

i^{th} observation, with N_x parameters (using some linear combination of the p parameters)

3. **Use coordinate descent algorithm to determine "relevant" predictors**

Steps 6: Gradient Ascent & Selection Criteria

1. **Move in the direction of a local maximum:**

$$\hat{x}^{n+1} = \hat{x}^{n} + \gamma \nabla \hat{y}(\hat{x}^{n})$$

Proposed next location of cluster

Previous location of cluster

Learning rate

Gradient of best fit

2. **Selection criteria**

   i. If *weighted* damage from cluster is in top quintile and calculated damage is greater than estimated, then set $\gamma = 0$ and keep it as a good candidate.

   Conservative

   ii. For all other observations, find the distance between the closest observation and the proposed (more-damaging) location

   iii. If the distance is less than a tolerance AND is in the "right" direction, then it is a good candidate.

   Conservative

   iv. Count how many observations are good candidates.

   v. Randomly select bins from lower (1^{st}-4^{th}) quintiles to remove from candidacy so that at least 20% of bins are removed from each iteration.

3. **Re-run MDA algorithm to 'top-up' set (keeping number of bins constant)**

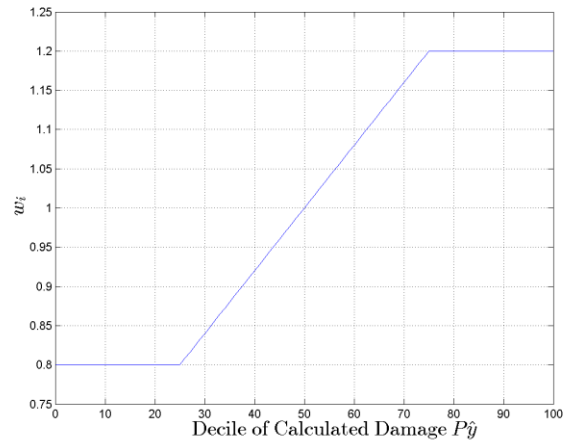
Representative

Step 7. Use weights to associate observations with damaging clusters

1. **Euclidian distance of k^{th} observation to i^{th} cluster:**

$$d_k = w_i \sqrt{\sum_{j=1}^{p} (\hat{x}_{j,i} - x_{j,k})^2}$$

2. **Add in weight function, based upon calculated damage:**



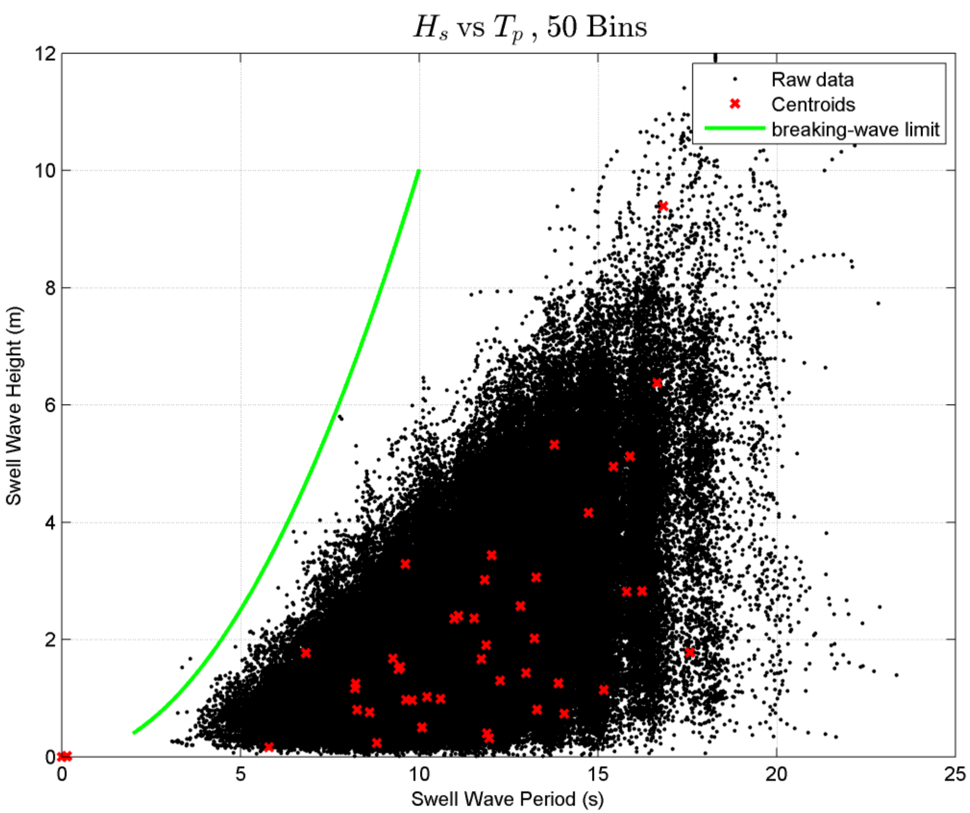
Conservative

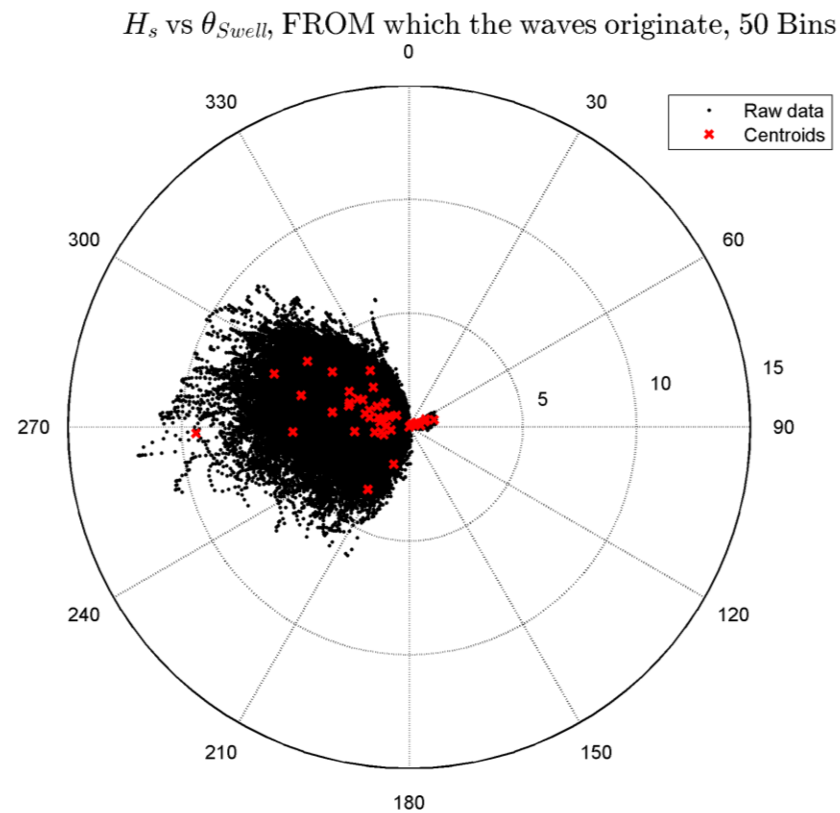
3. **Re-cluster observations to associate observations with "nearest" cluster**

PRINCIPLE POWER

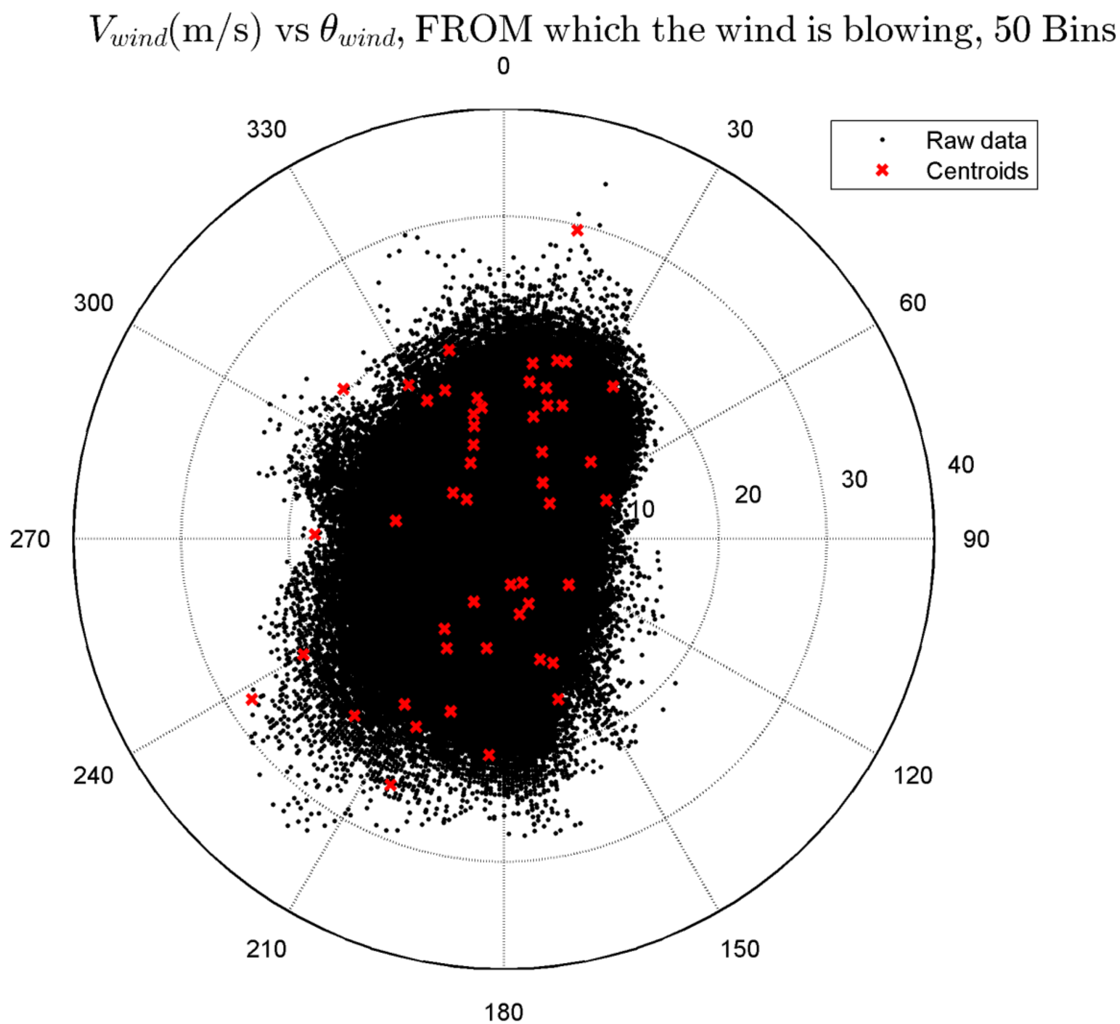
END:
ALGORITHM

Metocean Data: Swell Waves



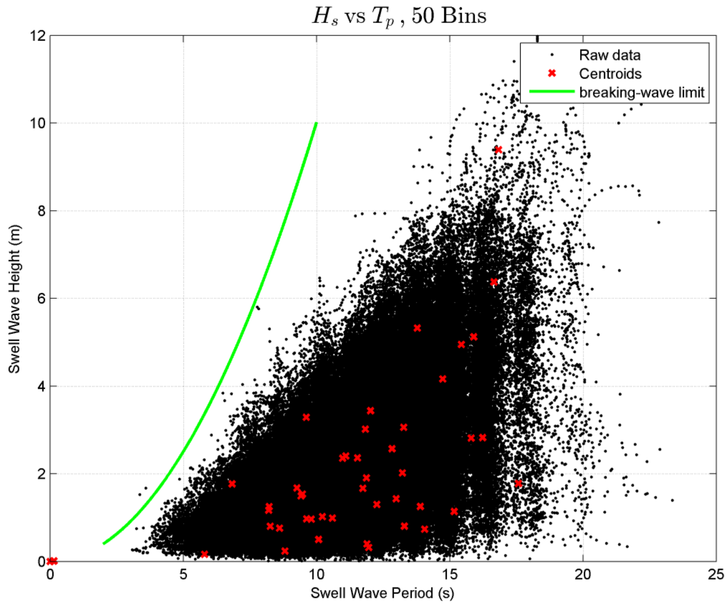


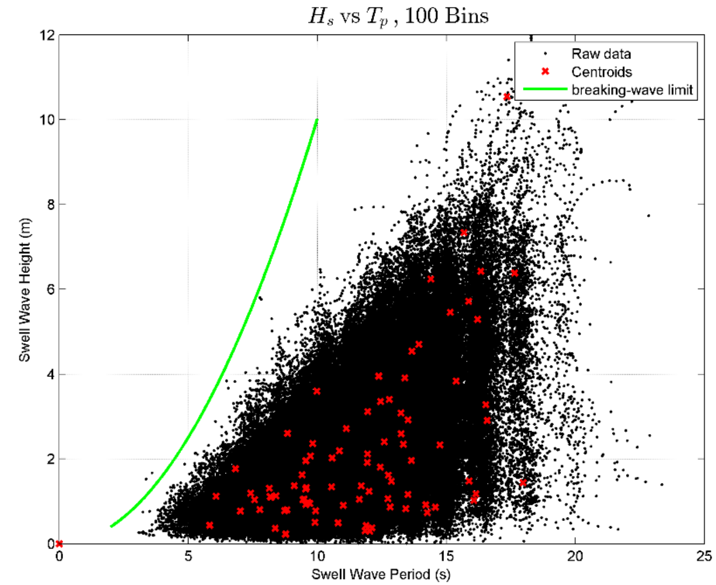
Metocean Data: Wind

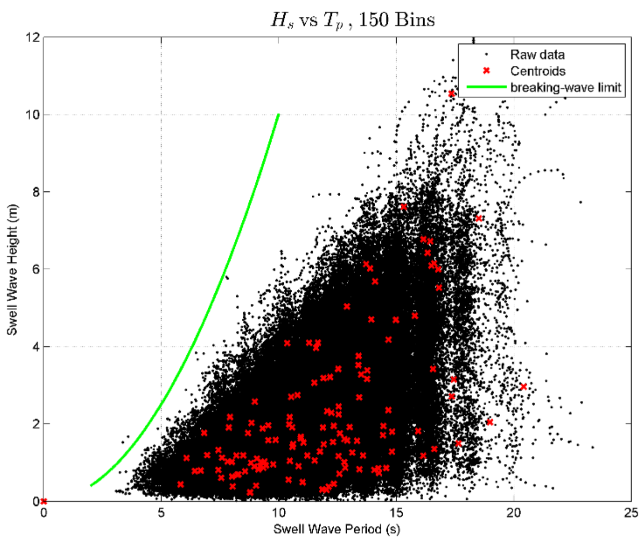


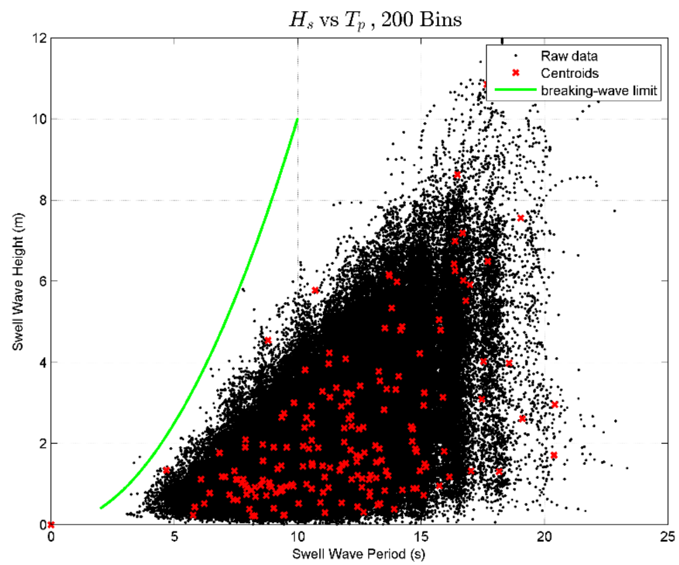
PRINCIPLE
POWER

Metocean Data: Swell Waves, Bin Dependence



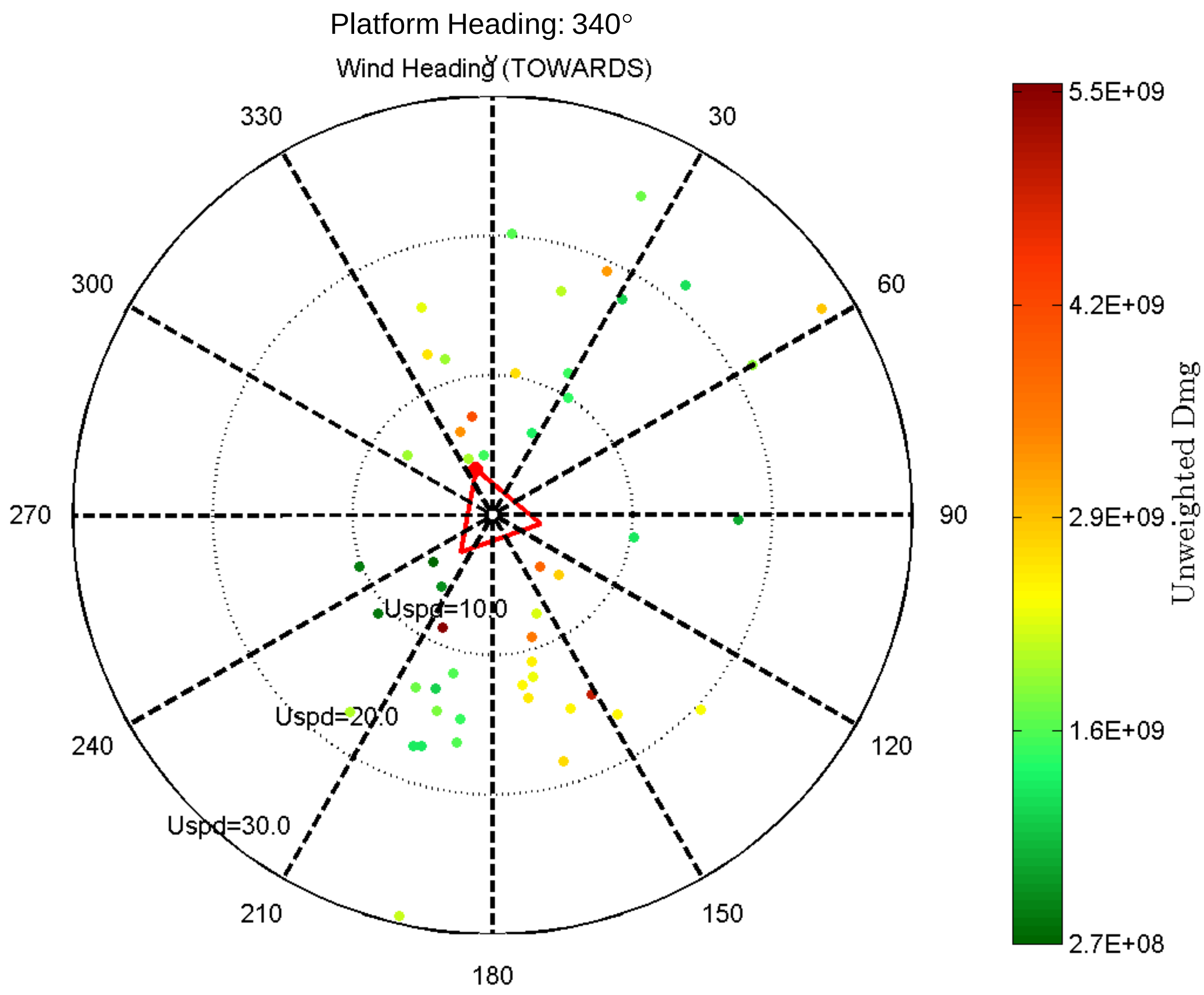






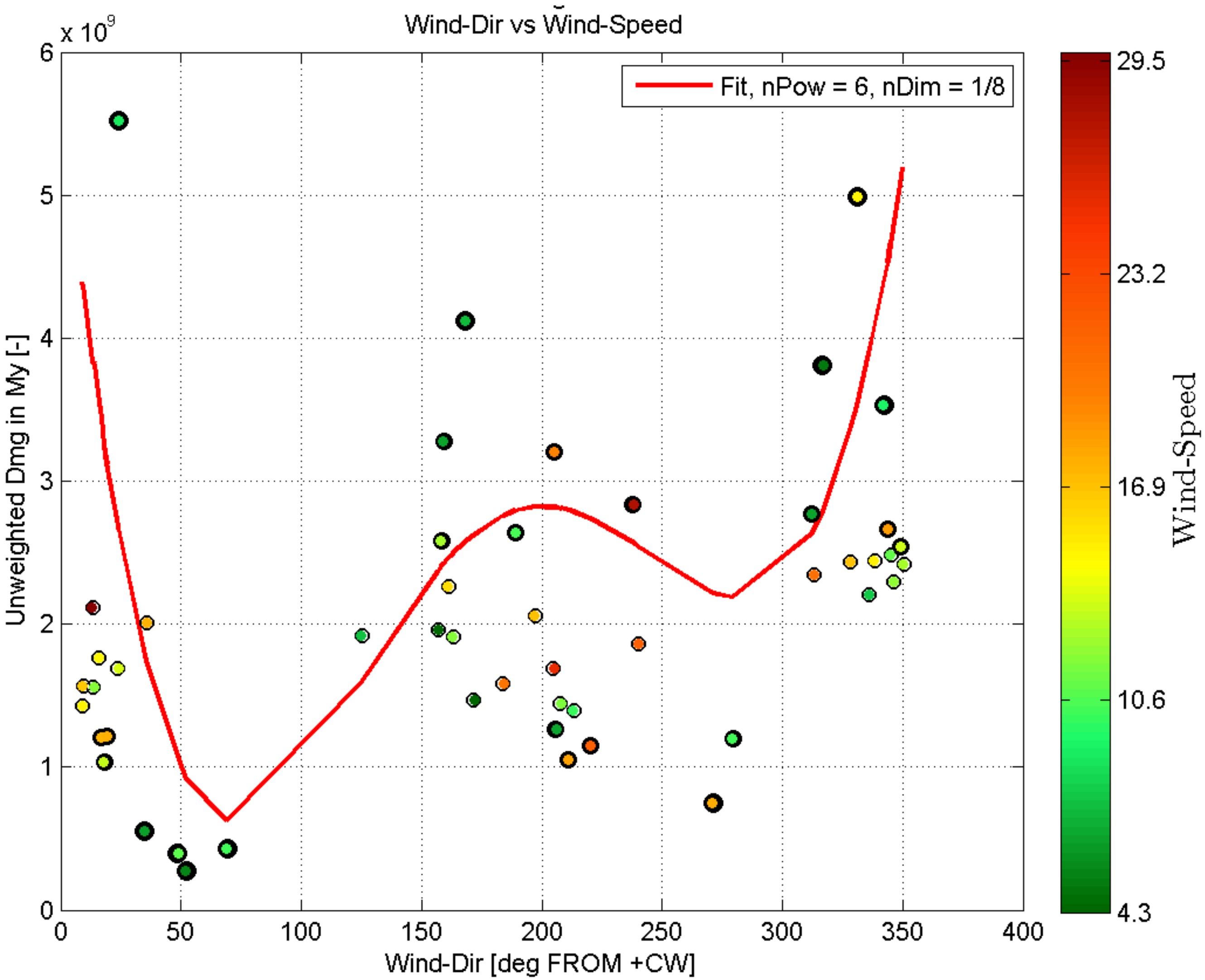
PRINCIPLE POWER

Results F/A Tower-Base: Damage Dependence on Wind (50 bins)



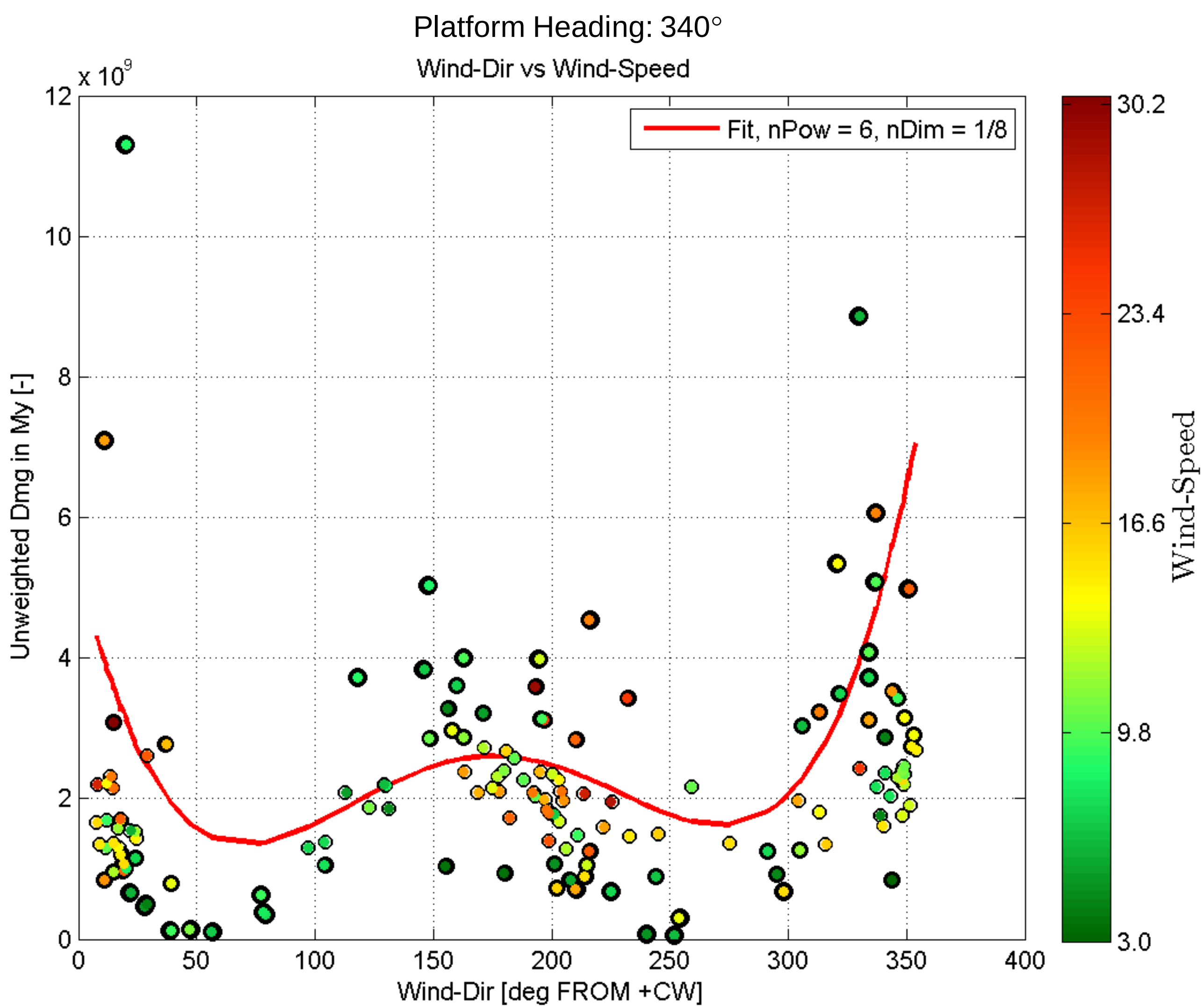
PRINCIPLE
POWER

1-D Regression (Wind-Direction, 50 bins)



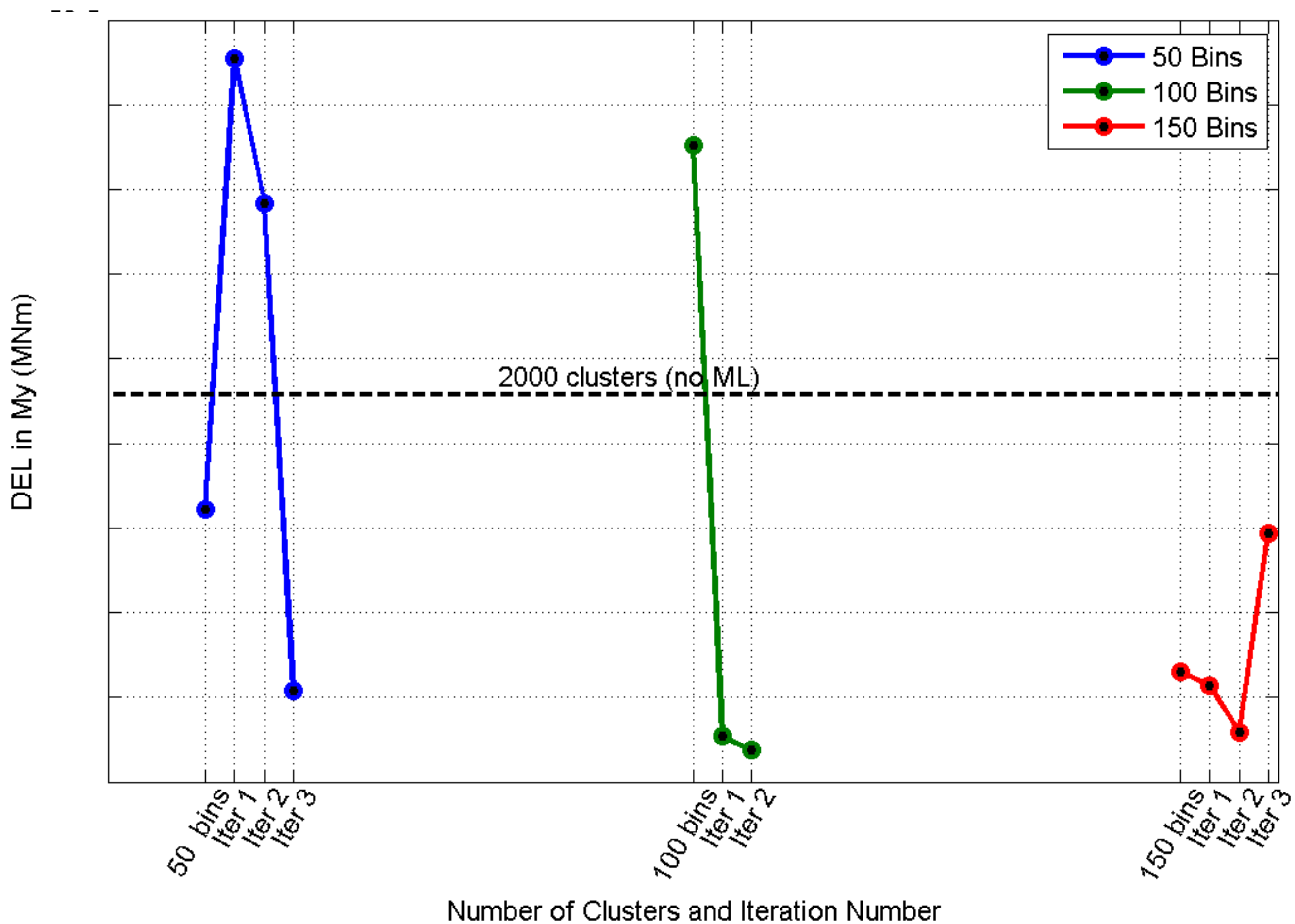
PRINCIPLE POWER

1-D Regression (Wind-Direction, 150 bins)



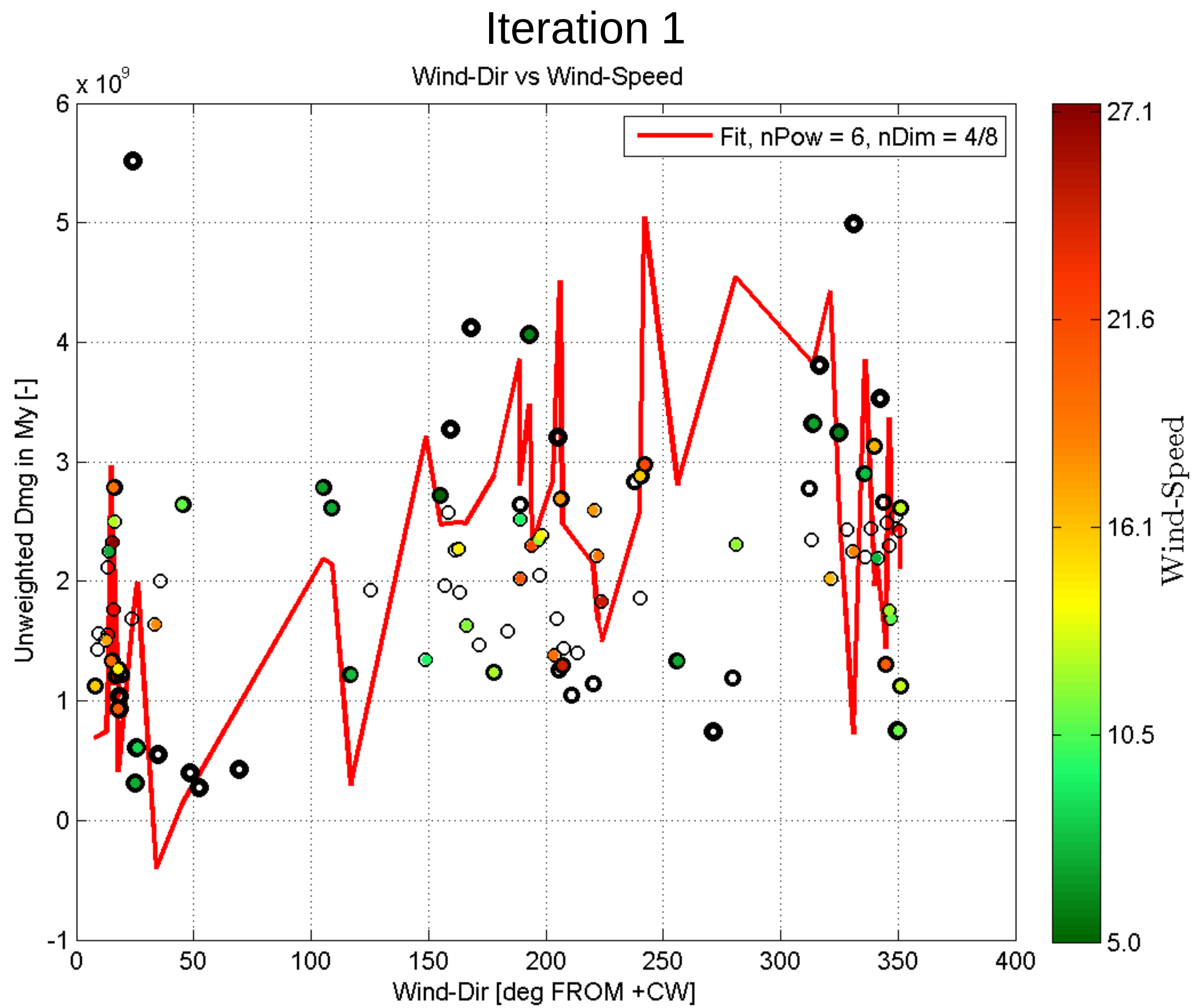
PRINCIPLE POWER

1-D Regression (Wind-Direction) DEL Results



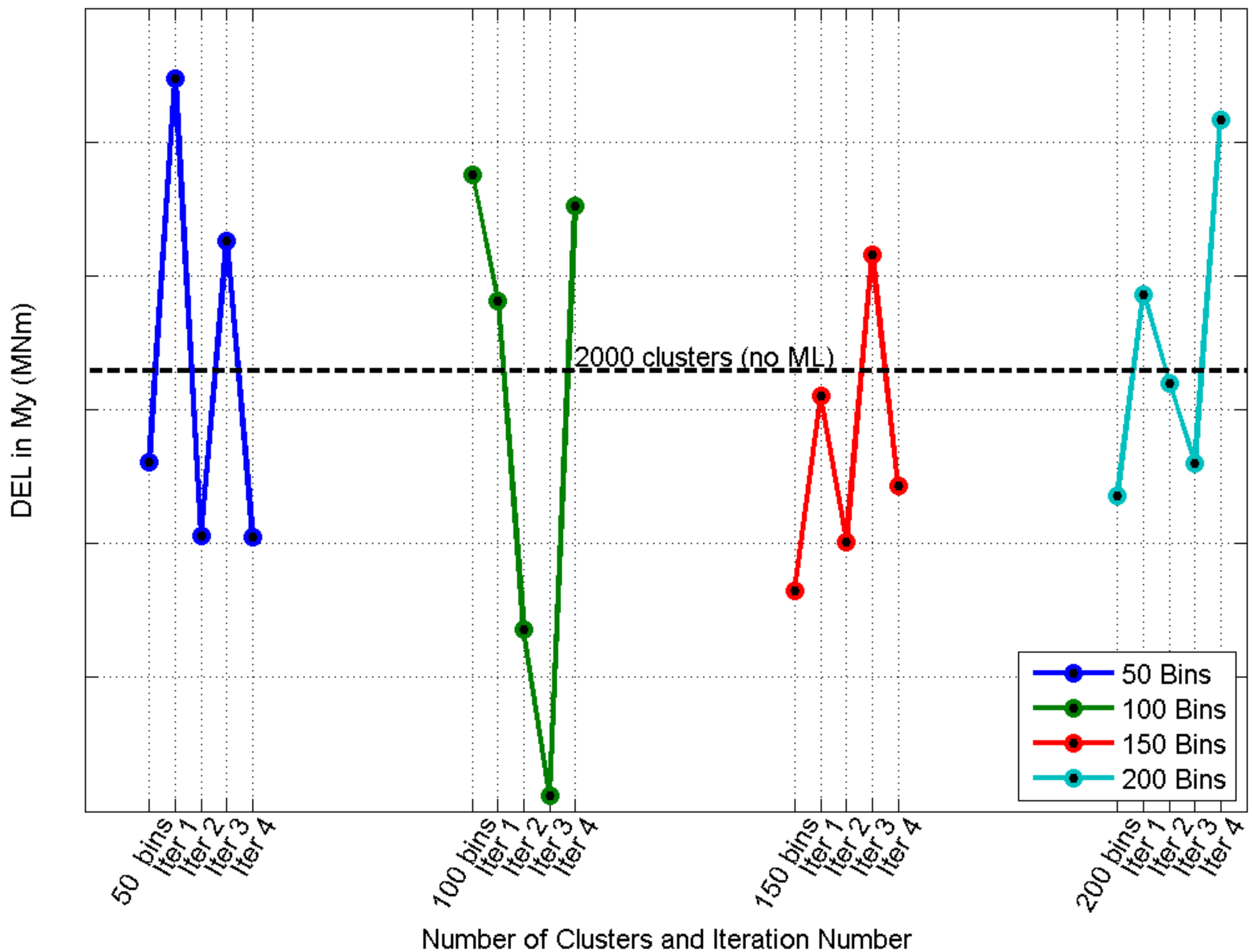
Conservative?

4-D Regression (Wind Speed+Direction, Wind-Sea Tp+Direction)



PRINCIPLE POWER

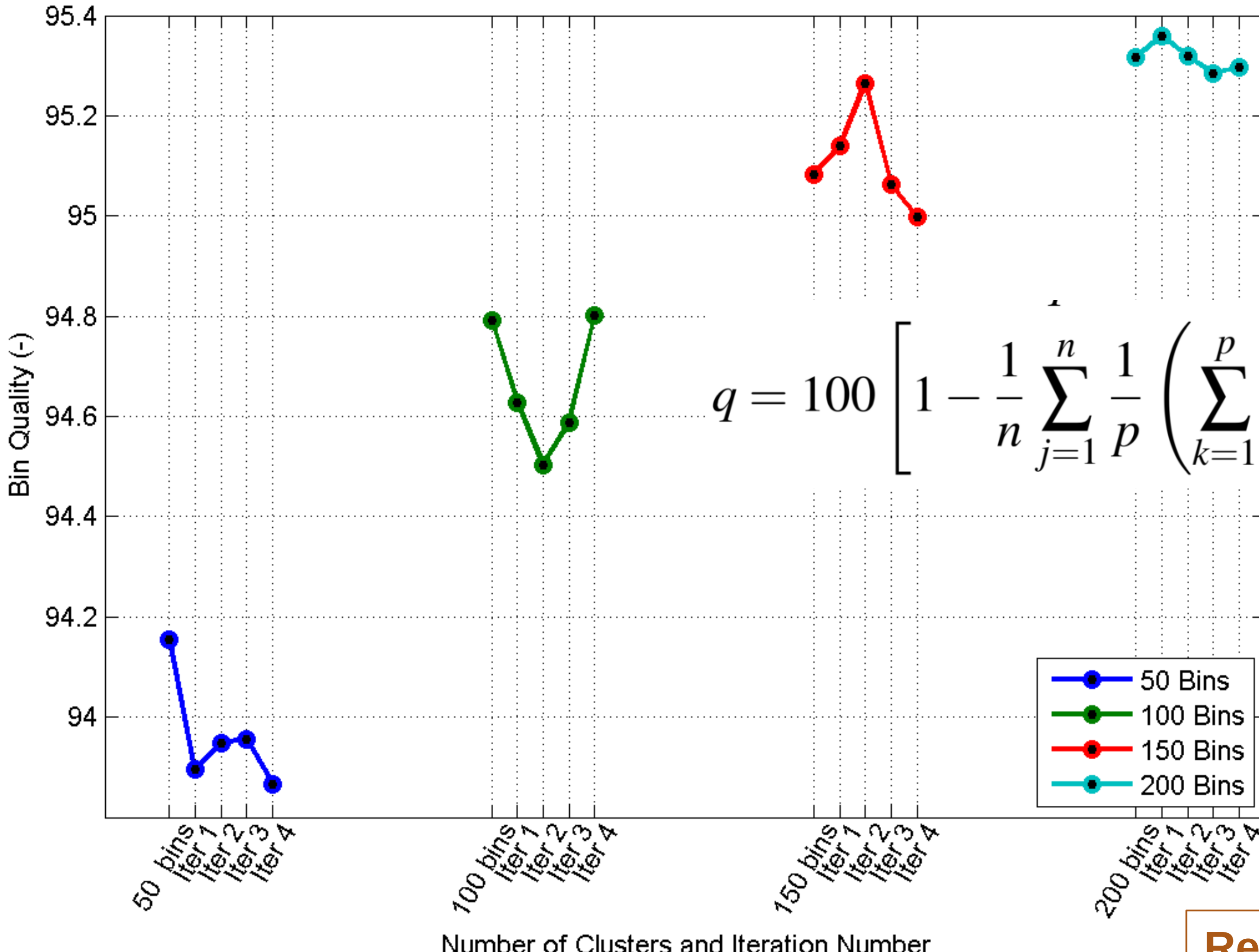
Results, 4-D Regression
(Wind Direction, Wind Speed, Wave Direction, Wave Tp)



Conservative?

PRINCIPLE POWER

"Quality" Measure as a Proxy for Representativeness



$$q = 100\left[1 - \frac{1}{n}\sum_{j=1}^{n}\frac{1}{p}\left(\sum_{k=1}^{p}|\hat{X}_k - X_{j,k}|\right)\right]$$

Representative?

PRINCIPLE POWER

Wrap-Up

- **A machine learning-based algorithm is proposed to try and find the most *representative* and *conservative* set of environmental conditions to estimate fatigue damage on a floating offshore wind turbine.**
- **While a 1-D linear regression (based on wind-direction) is easily identified, it does not lead to conservative damage estimations.**
- **A 4-D linear regression (based upon wind and wind-seas) leads to a more wildly behaving fit, but finds better conservativeness.**
- **The values of *representativeness* and *conservativeness* may be opposed to each other.**
- **In the future, we hope to improve algorithm to find conservativeness with smaller number of conditions**
  - More regularization?
  - Learning rate?

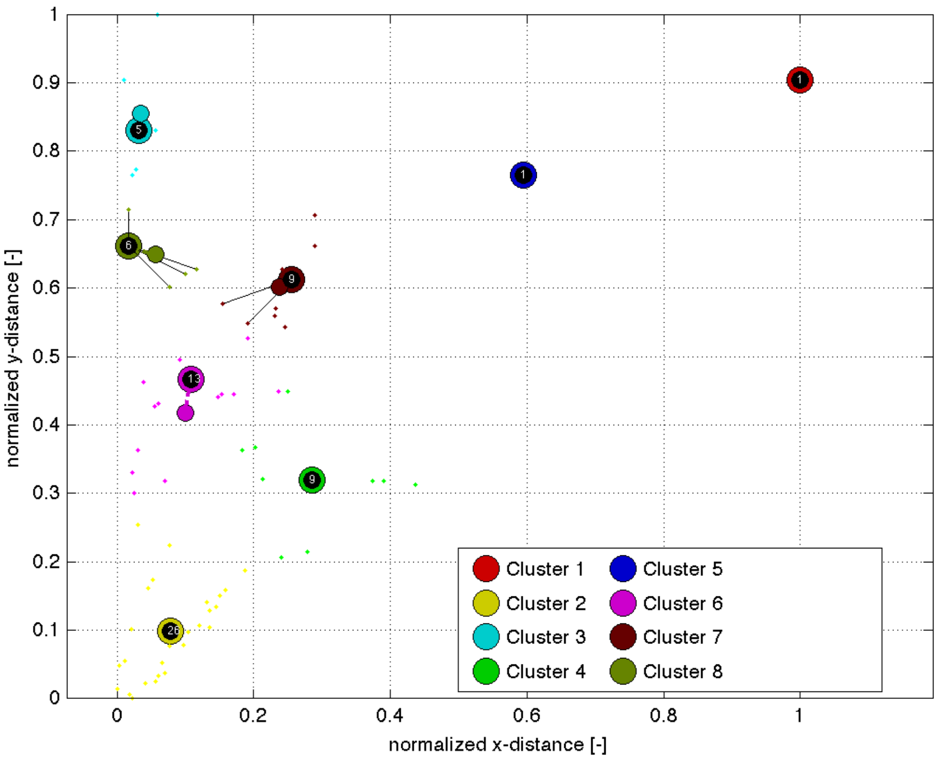
PRINCIPLE
POWER

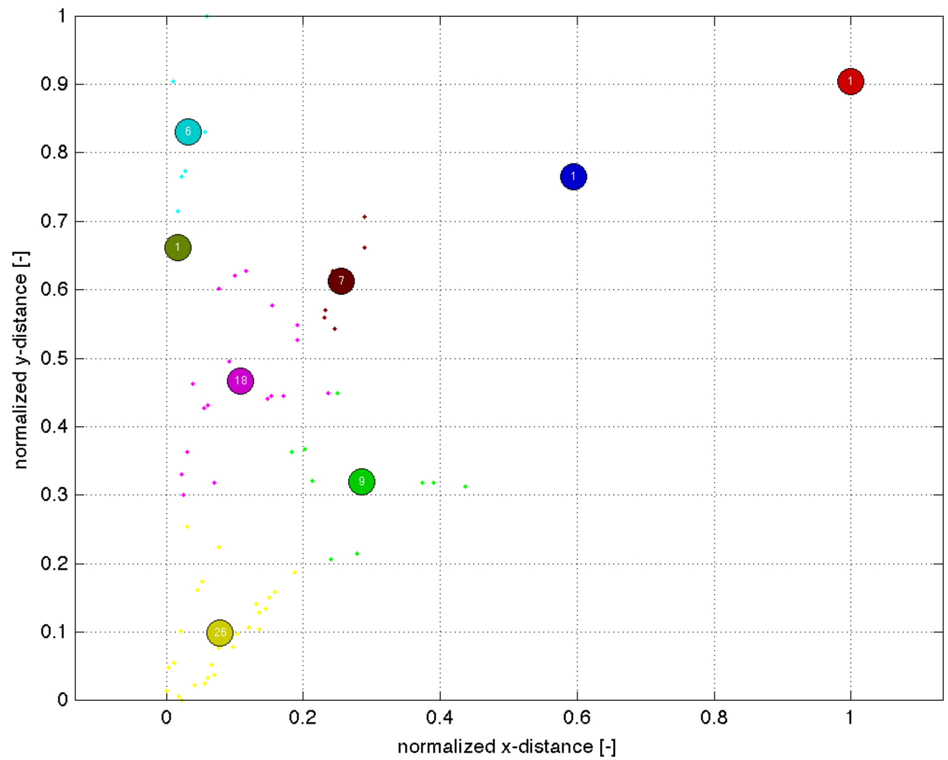
PRINCIPLE
POWER

Questions?

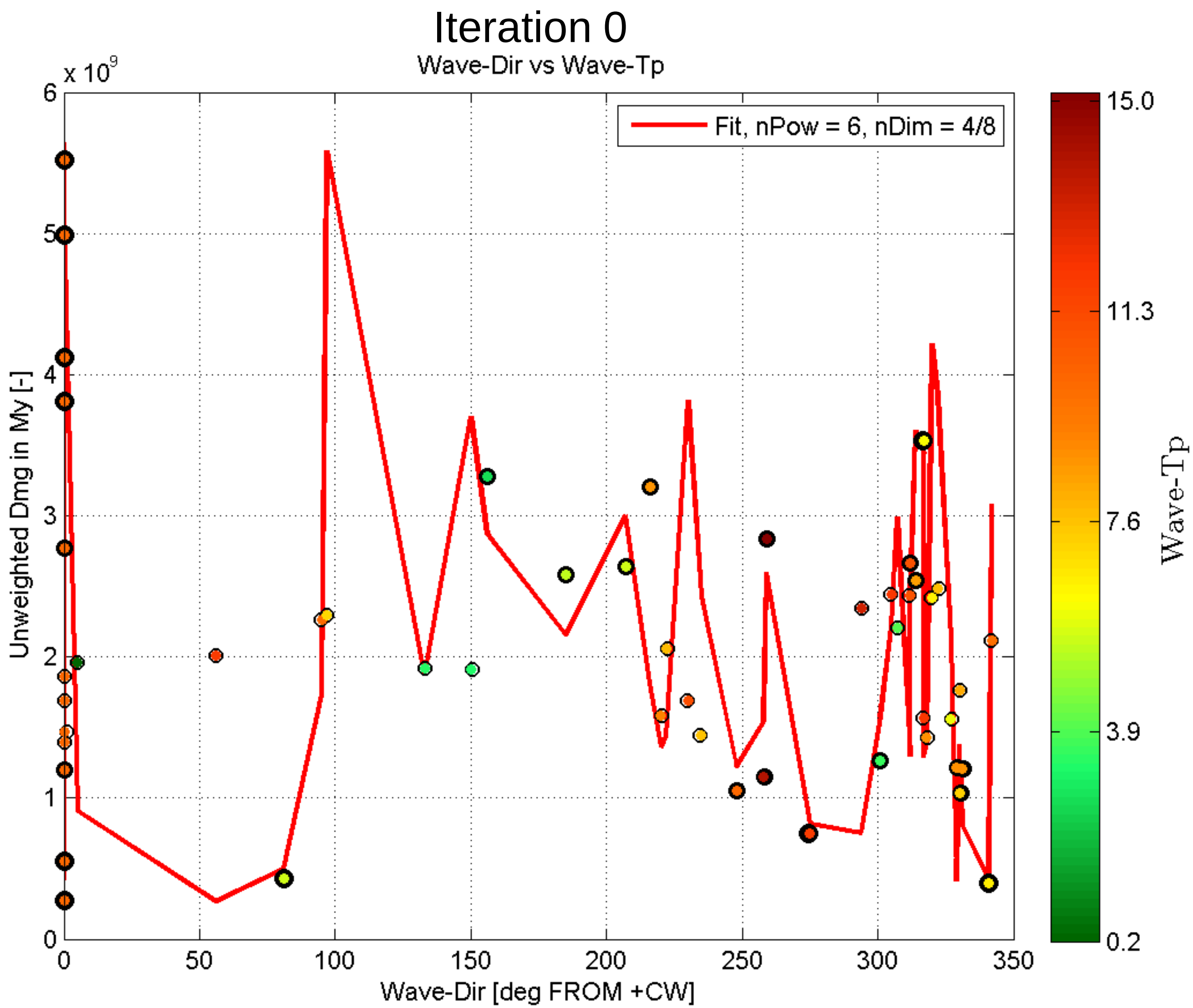
Sam Kanner
skanner@principlepowerinc.com

Step 7: Re-cluster observations based on new locations



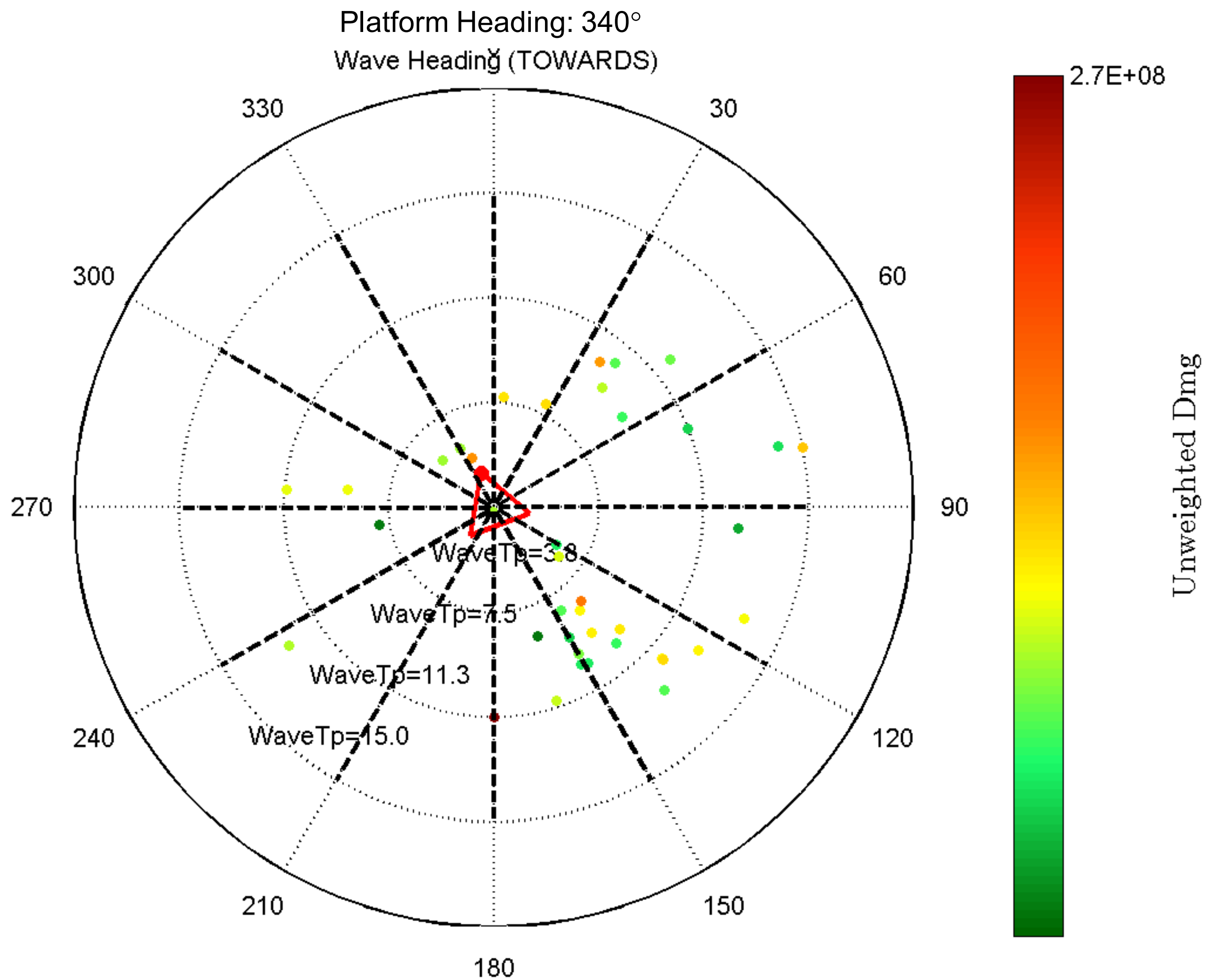


4-D Regression, Wind-Sea Dependence



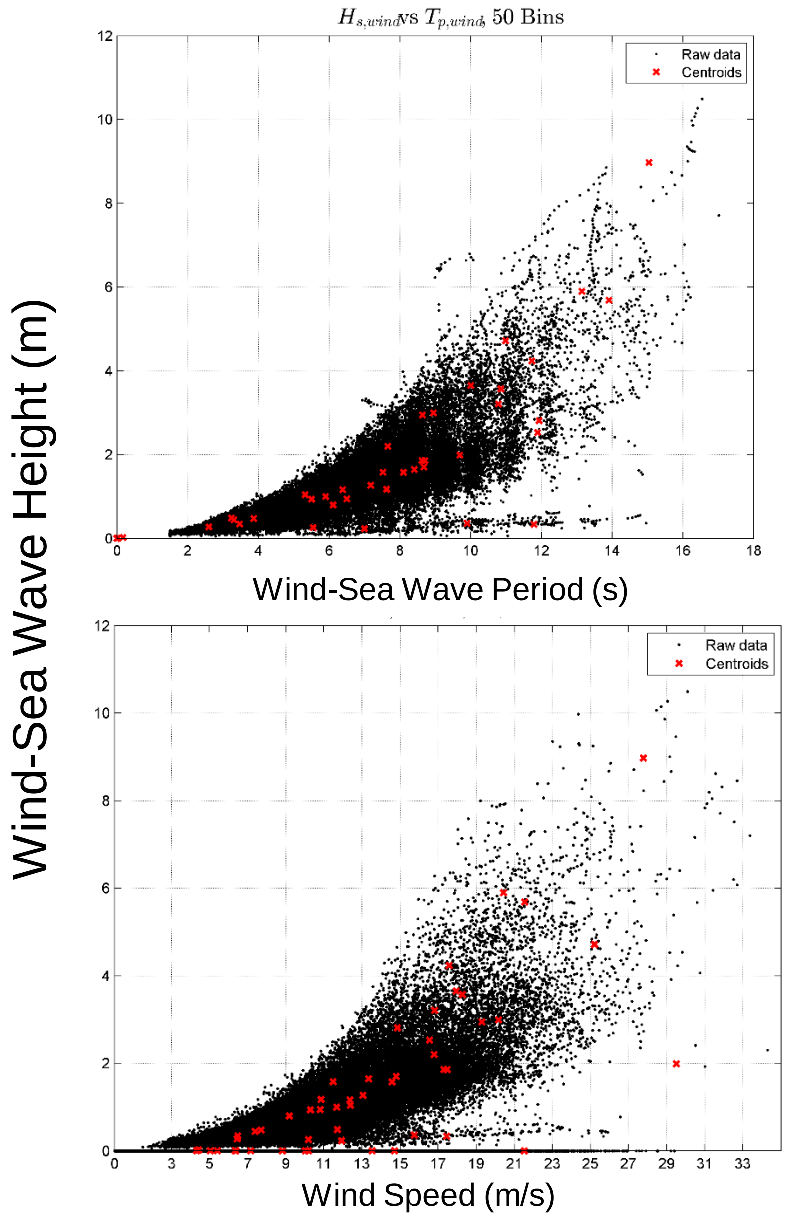
PRINCIPLE
POWER

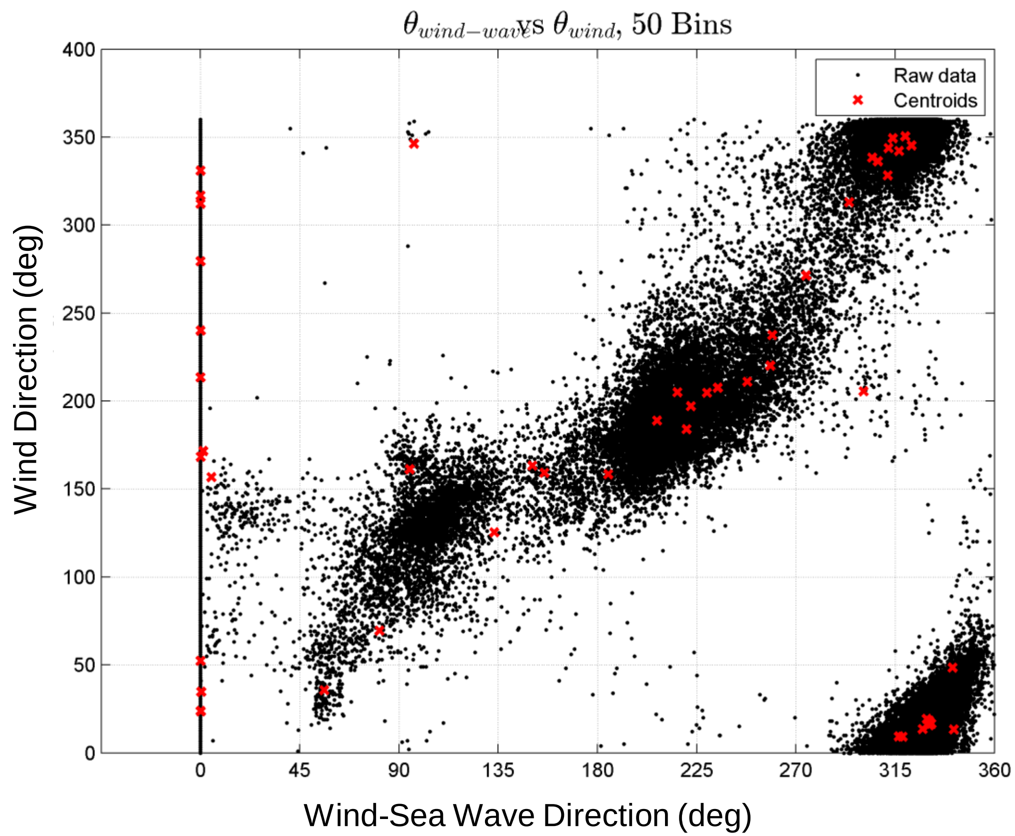
Results F/A: Damage Dependence on Wind-Sea (50 bins)



PRINCIPLE POWER

Metocean Data: Wind-Sea Waves





PRINCIPLE
POWER

Step 5 (cont.) Coordinate descent determines relevant parameters

1. **Again, trying to find β such that:**

$$\min_{(\beta_0,\beta)\in\mathbb{R}^{p+1}} R_\lambda(\beta_0,\beta) = \min_{(\beta_0,\beta)\in\mathbb{R}^{p+1}} \left[\frac{1}{2N}\sum_{i=1}^{N}(y_i-\beta_0-x_i^\top\beta)^2+\lambda P_\alpha(\beta)\right]$$

2. **The minimum of the residual:**

$$\frac{\partial R_\lambda}{\partial\beta_j}\Big|_{\beta=\tilde\beta} = -\frac{1}{N}\sum_{i=1}^{N}x_{ij}(y_i-\tilde\beta_o-x_i^\top\tilde\beta)+\lambda(1-\alpha)\beta_j$$

3. **Update the guess of β:**

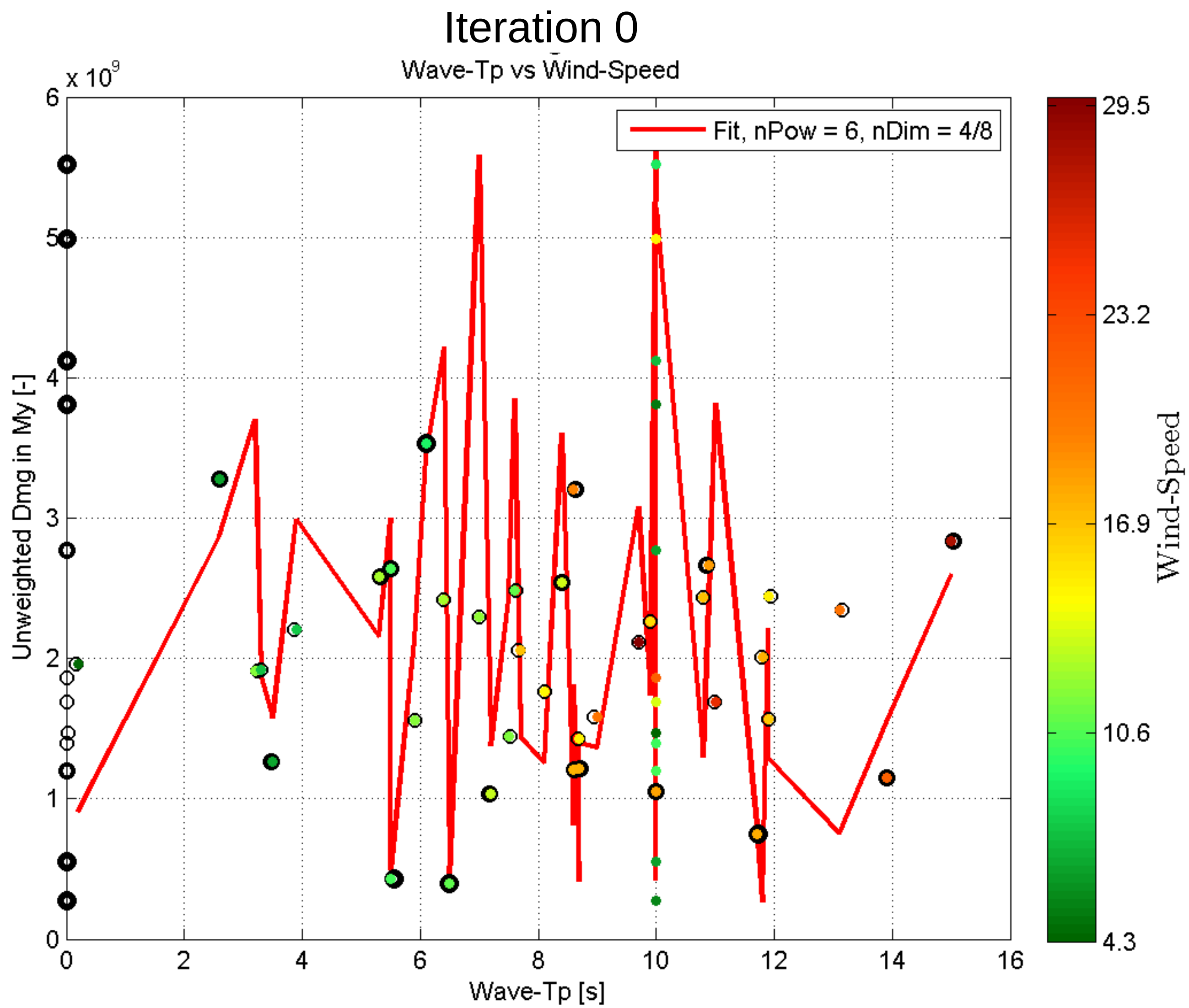
$$\tilde\beta_j \leftarrow \frac{S\left(\frac{1}{N}\sum_{i=1}^{N}x_{ij}(y_i-\tilde y_i^{(j)}),\ \lambda\alpha\right)}{1+\lambda(1-\alpha)} \qquad \tilde y_i^{(j)} = \tilde\beta_0+\sum_{\ell\neq j}x_{i\ell}\tilde\beta_\ell$$

$$S(z,\gamma)\quad \operatorname{sign}(z)(|z|-\gamma)_+ = \begin{cases} z-\gamma & \text{if } z>0 \text{ and } \gamma<|z| \\ z+\gamma & \text{if } z<0 \text{ and } \gamma<|z| \\ 0 & \text{if } \gamma\geq|z|. \end{cases}$$

Friedman, J., R. Tibshirani, and T. Hastie. Regularization paths for generalized linear models via coordinate descent. Journal of Statistical Software, Vol 33, No. 1, 2010. http://www.jstatsoft.org/v33/i0

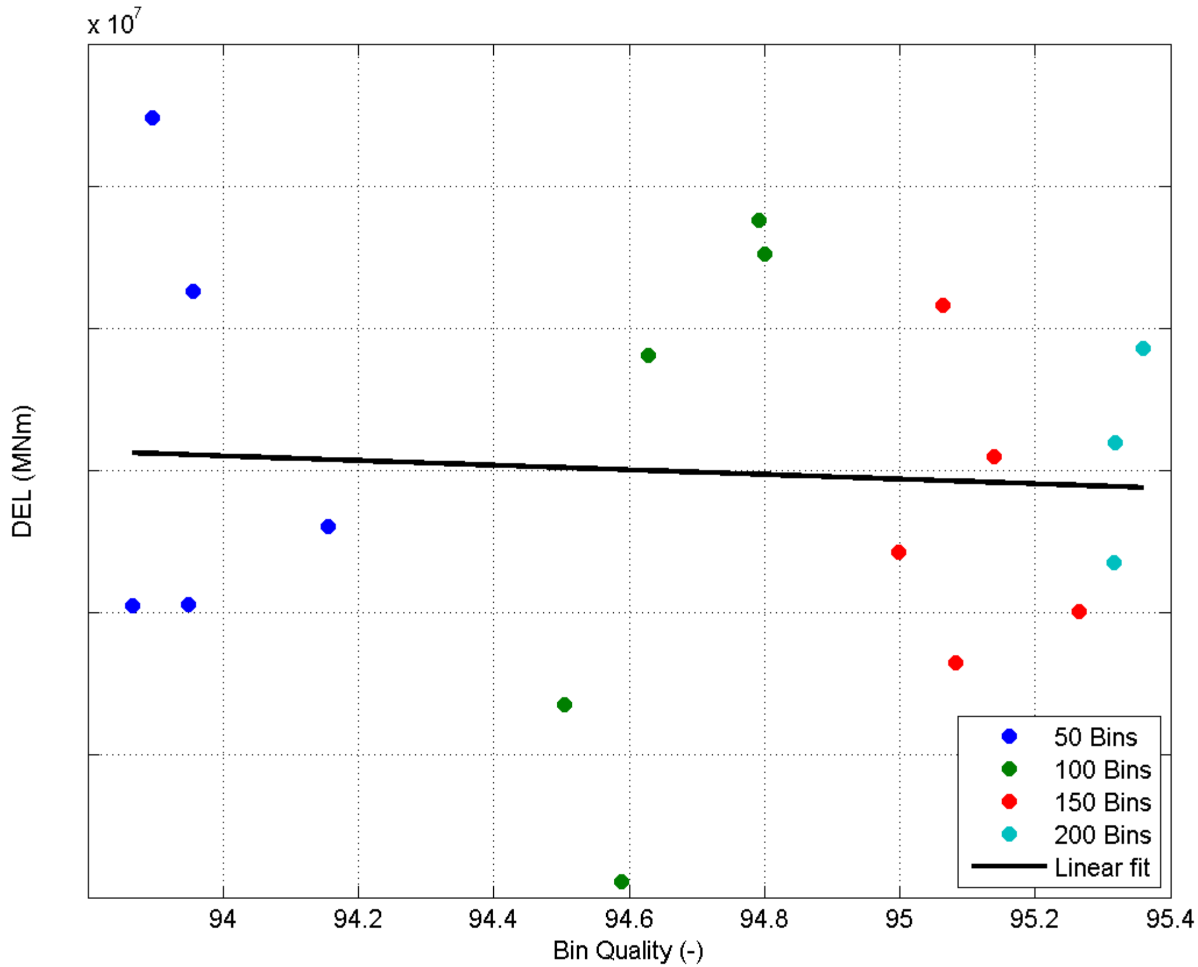
PRINCIPLE POWER

4-D Regression, Wind-Sea Dependence



PRINCIPLE POWER

Competing Interests: Representativeness vs Conservativeness



PRINCIPLE POWER