2019-09-03

Use of HF radar for replciating wave-current combined wave conditions for testing of wave energy converters

Wang, D

http://hdl.handle.net/10026.1/16004

https://proceedings.ewtec.org/?

Use of HF radar for replicating wave-current combined wave conditions for testing of wave energy converters

Daming Wang, Daniel Conley, Martyn R. Hann, Keri M. Collins and Deborah Greaves

Abstract—**Wave tank testing is a useful tool to assess the performance of Wave Energy Converters (WEC) at different technology readiness levels (TRL). At early TRL the use of systematically varying wave conditions is acceptable, however at later stages there is a need to use testing conditions representative of potential prototype deployment sites. Environmental data at these deployment sties can be collected by various instruments, such as buoys, ADCP, radars etc. The most commonly used method to re-create the measured environmental testing conditions in a wave tank is to represent a sea state using the measured significant wave height and peak period and a parametric wave spectrum such as the JONSWAP or Pierson-Moskowitz spectrum. Although a useful tool, these parametric spectra represent a simplification which omits much of the site specific spectral information, such as the directional information of the sea state, which has the potential to significantly impact WEC performance. In most wave tanks it is possible to reproduce directly a measured spectrum. However this raises questions about which measured cases to reproduce. The use of HF radar data provides the opportunity to select sea states for wave tank testing that are more representative of a potential deployment site. It is necessary to develop a methodology to systematically select a set of representative experimental test cases from the wave datasets obtained from HF radar. Previous research has demonstrated the use of K-means clustering to obtain representative wave cases from measured wave spectrum. Here the expansion of this method is demonstrated. HF radar data obtained at Wave Hub, a wave energy test site in Cornwall, UK, and buoy data obtained close to Long Island, USA are used in this study.**

Keywords— **HF radar, K-means clustering method, wave spectrum, WEC**

I. Introduction

WAVE energy converters (WEC) are devices which are able to capture energy within waves and transform it into other kind of energy such as electricity [1]. The first WEC in history was invented in 1799 [2]. After centuries' development, there are mainly five types of WECs which are oscillating wave surge converter (OWSC), oscillating water columns (OWC), overtopping devices, point absorbers and attenuators. However, even by 2019, the commercialization of WECs is not as successful as other types of marine renewable energy such as offshore wind energy or tidal energy converters because of the high cost. Although not exploited yet, the potential of WECs is large , it is expected the cost for WEC devices will decrease in the near future to be competitive with conventional power plants [3]. It has been reported that in the UK alone, the wave energy devices have a potential to provide more than 50TWh/yr [4].

Due to the high financial and technical risks during the development of WECs, Technology Readiness Levels (TRL) [5], an approach established by the US Space Agency (NASA), are widely used to describe the state of the WEC programme. According to the TRL concept, the development of WECs from concept validation to final commercialization is divided into 5 stages, the first two stages rely on small-scale tank model testing with scale from 1:100 to 1:10 [6]. During tank testing, the concept of WECs will be validated and the performance at small scale can be verified.

Owing to the importance of tank testing for the development of WECs, it is necessary to determine representative sea states based on limited resources. Traditionally the sea states can be represented by

ID number 1615, conference track WRC. This work has been conducted as part of a PhD funded by the University of Plymouth.

Daming Wang is with School of Engineering, University of Plymouth, Drake Circus, Plymouth, Devon, U.K. PL4 8AA (e-mail: daming.wang@plymouth.ac.uk).

Daniel Conley is with School of Biological and Marine Sciences, University of Plymouth, Drake Circus, Plymouth, Devon, U.K. PL4 8AA (e-mail: daniel.conley@plymouth.ac.uk).

Martyn R. Hann is with School of Engineering, University of Plymouth, Drake Circus, Plymouth, Devon, U.K. PL4 8AA (e-mail: martyn.hann@plymouth.ac.uk).

Keri M. Collins is with School of Engineering, University of Plymouth, Drake Circus, Plymouth, Devon, U.K. PL4 8AA (e-mail: keri.collins@plymouth.ac.uk).

Deborah Greaves is with School of Engineering, University of Plymouth, Drake Circus, Plymouth, Devon, U.K. PL4 8AA (e-mail: deborah.greaves@plymouth.ac.uk).

parametric wave spectra such as JONSWAP or Pierson-Moskowitz, by giving the significant wave height and the peak spectrum [7]. However, the directional information of the sea states and the details of the shape of the wave spectrum are omitted. To solve the problem, we use the site-specific directional wave spectrum directly. There exist several instruments that can measure site-specific directional wave spectrum such as buoy, Acoustic Doppler Current Profiler (ADCP) and HF radar. Obtaining the directional wave spectrum from those instruments is the first step due to the fact that each instrument can provide thousands of hourly or half-hourly directional wave spectra, it is then important to identify a certain number of representative directional wave spectra of the sea states to be used for tank testing of WECs.

In order to find out the representative directional wave spectrum, Hamilton [8] used the K-means clustering method on 2456 non-directional wave spectra measured at Port Hedland, Australia in 1992 to get the representative sea states. This method was later extended into more forms by Draycott [9] to identify 20 and 40 representative sea states from 64673 buoy-measured half-hourly directional wave spectra provided by European Marine Energy Centre (EMEC). The K-means clustering method performed well in selecting representative sea states.

A. *K-means clustering method*

The K-means clustering method is used to cluster a total of *N* instances into *K* groups, ensuring that the most similar instances are put in the same group. In order to find the preferred clustering results, the sum of squared error (SSE) of all instances is minimised. The equation for SSE is given in equation (1) [10]:

$$\begin{aligned} \mathrm{SSE} &= \sum_{k=1}^{K} \sum_{\forall x_i \in C_k} ||x_i - \mu_k||^2 \\ &= \sum_{k=1}^{K} \sum_{\forall x_i \in C_k} d(x_i, \mu_k)^2 \end{aligned} \tag{1}$$

In which x_i is the data instance, C_k is the set of instances in cluster k, μ_k is the vector mean of cluster k. d is the Euclidean distance between two p-dimensional instances, $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})$ and $x_j = (x_{j1}, x_{j2}, \dots, x_{jp})$.

$$\begin{aligned} d(x_i, x_j) = (&|x_{i1} - x_{j1}|^2 \\ &+ |x_{i2} - x_{j2}|^2 + \ldots + |x_{ip} \\ &- x_{jp}|^2)^{1/2} \end{aligned} \tag{2}$$

μ_k is defined as:

$$\mu_k = \frac{1}{N_k} \sum_{\forall x_i \in C_k} x_i \tag{3}$$

The K-means clustering method provides the easiest way to minimize SSE, the flow chart is shown in Fig.1.

When SSE does not decrease by relocation of the cluster centres, it indicates the current partition is optimal. The iteration will then stop. The K-means method is widely used for multi-dimensional data point clustering, thus can be extended and applied to multidimensional instances such as non-directional and directional spectra.

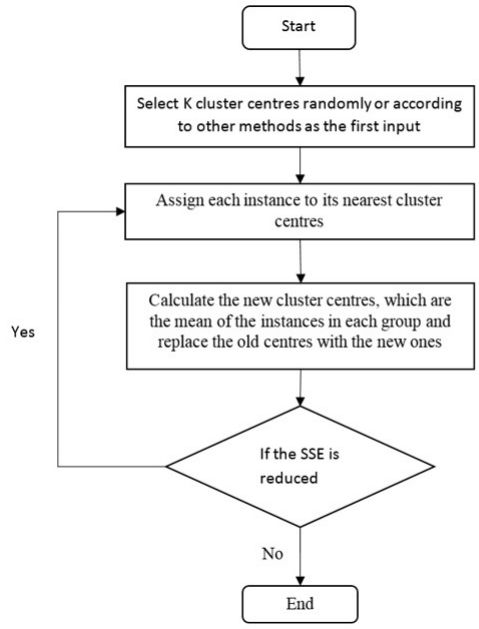

Fig. 1. Flow chart for K-means clustering method.

The main disadvantage of K-means clustering method is that the clustering results will be affected by the initial K cluster centres for the first iteration which is usually selected randomly. To get rid of this uncertainty, the common way is to repeat the whole clustering procedure iteratively until the SSE has reduced below the threshold. Alternatively, we can use a K-means++ algorithm that can reduce the uncertainty for initial centres generation, according to Arthur and Vassilvitskii [11], the K-means++ algorithm uses an heuristic to find the initial centres for K-means clustering which improves the running time and the quality of the final solution. However, the K-means++ method still cannot find out the best result in one run, it still needs repetition. In addition, the K-means method will not tell us how large the number of groups is, this value is defined subjectively.

B. *High frequency radar system*

High frequency radar is a shore-based remote sensing system which was commonly used to measure the surface current information. The working principle of HF radar is based on the continuous transmission of vertically polarized electromagnetic waves, which are scattered by the water wave surface. The scattered waves that travel back and are received by the radar are named back scattered waves [12]. By analysing the back scattered wave

spectra, two nearly symmetric dominant peaks can be seen from the spectra [13]. The surface current information can be found by analysing the frequency shift of the two peaks, thus constitutes the foundation for surface current measurement [14]. By using two HF radars covering the same ocean area together and processing the back-scattered spectrum with certain inversion algorithms, the directional wave spectra can be obtained ([15], [16]). An example is the two-phased-array Wellen Radars (WERA) system located in the southwest coast of the UK which overlooks a marine renewable testing field (Wave Hub). The directional spectra used in this paper are all measured by this radar system.

TABLE I
METHODS USED FOR SEA STATES REGROUPING

Method	Description
A	*Hs,Te binning method*
B	*Hs,Te,θ_m binning method*
C	*S(f) clustering method*
D	*S(f,θ) clustering method*
E	*Hs,Te clustering method*
F	*Hs,Te,v, θ_m ,P,Sp,σ_θ clustering method*
G	*E+C two-step method*
H	*F+D two-step method*

Hs is the significant wave height, *Te* is the energy period, θ_m is the mean wave direction, *S(f)* is the non-directional wave spectrum, *S(f,θ)* is the directional wave spectrum, *v* is the spectral bandwidth, *P* is the wave power, *Sp* is the wave steepness, σ_θ is the directional spreading.

II. METHODS TO GET REPRESENTATIVE SEA STATES

Data measured by HF radar from Wave Hub from April 2nd 2012 to December 4th 2012 was used as an example data set. Each hourly directional spectrum is in the units of $m^2/(Hz \cdot rad)$, there are 30 angular directions ranging from 0 rad to 29π/15 rad with an interval of π/15 rad and 92 frequencies ranging from 0.03Hz to 0.28Hz. To make sure the regrouping method is reliable and independent of the location and measuring instruments, we introduce a control group, which is buoy-measured directional wave spectra. The buoy data was collected from station 44025, which was deployed close to Long Island, South of Islip, New York. It is owned and maintained by National Data Buoy Centre [17]. From 2008 to 2017, there are 74896 hourly directional wave spectra available. In the year 2017 alone, there are 8402 data instances. The unit of buoy measured directional wave spectra is in $m^2/(Hz \cdot degree)$, there are also 30 angular directions ranging from 0° to 348° with an interval of 12°. There are 47 frequency values ranging from 0.02Hz to 0.49Hz.

C. *Eight regrouping methods*

In previous work by Draycott [9], eight methods were created and tested on the EMEC buoy measured directional wave spectra. In this paper, the same methods are used on both the HF radar and long island buoy data in order to check if there is similar conclusions. All of the eight method are shown in TABLE I.

What is worth mentioning is that method C to method H are all based on K-means clustering methods, while method A and B are binning methods. For method A, each directional spectrum is plotted in *Hs-Te* two-dimensional space. By defining the size of equally sized *Hs-Te* bins, the total data set can be divided into a number of bins, the centroid of the data in each bin is used as the representative sea states. Method B is similar to method A but introduces the third dimension θ_m, the mean wave direction. So the bins are three-dimensional bins.

Method C is non-directional spectrum K-means clustering method, similar to multi-dimensional point clustering method, the difference between two non-directional spectra $S_i(f)$ and $S_j(f)$, $f = (f_1, f_2, \dots, f_p)$ is given in equation (4):

$$d\big(S_i(f), S_j(f)\big) = (|S_i(f_1) - S_j(f_1)|^2 + \big|S_i(f_2) - S_j(f_2)\big|^2 + \dots + |S_i(f_p) - S_j(f_p)|)^{1/2} \tag{4}$$

For method D, with the same manner, the difference between two directional spectra $S_i(f,\theta), S_j(f,\theta)$, $f = (f_1, f_2, \dots, f_p)$, $\theta = (\theta_1, \theta_2, \dots, \theta_q)$ can be calculated by the following equation (5):

$$d\left(S_i(f,\theta), S_j(f,\theta)\right) = \frac{1}{pq}\left[\sum_{m=1}^{p}\sum_{n=1}^{q}(|S_i(f_m,\theta_n) - S_j(f_m,\theta_n)|^2)\right]^{1/2} \tag{5}$$

For method E and F, normalized parameters are used for the K-means clustering, which are significant wave height *Hs*, energy period *Te*, spectral bandwidth *v*, wave mean direction θ_m , power *P*, peak steepness S_P , directional spreading parameter σ_θ. They are all normalized by their mean of the data set to eliminate the influence of different units. Method G and H are both two-step methods, the first step is to create *K*/2 sub-clusters by regrouping method E and F and the second step is to use C and D to split each sub-cluster into two groups to create in total *K* groups.

D. *Two metrics to check the quality of regrouping*

In order to check the quality of the regrouping results, two metrics, metric one [18] and metric two [19] are used. Metric one focuses on the difference between each instance's certain parameter with its cluster mean. After *K*

clusters are created, $= 1, \dots K$, each group k has a certain number of members $M(k)$, $m = 1, \dots M(k)$ represents each group's member. For a certain variable v for analysis $(v = H_s, T_e, \dots S(f), S(f,\theta))$, $d = 1, \dots D(v)$ represents how many discrete values different variable v has, because metric one represents distance from the cluster mean, a lower value I better. Metric one is defined in equation (6).

$$Met(v) = \sum_{k=1}^{K} \frac{1}{K} \sum_{m=1}^{M(k)} \frac{1}{M(k)} \sum_{d=1}^{D(v)} \frac{|v_{k,m,d} - \mu_{k,d}(v)|}{\sum_{d=1}^{D(v)} \mu_{k,d}(v)} \quad (6)$$

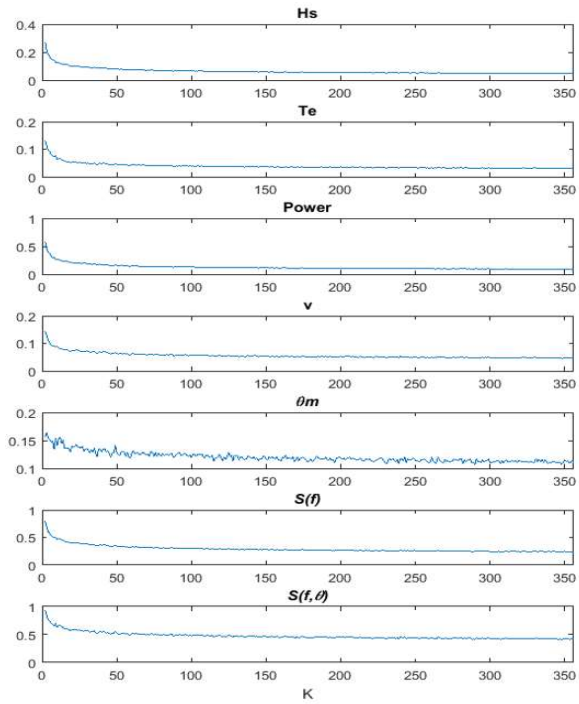

Fig. 2. Metric one for Long Island buoy measured data with K for seven parameters.

Metric two considers not only the difference between each instance's certain parameter and their cluster mean, but also the cluster means for all different clusters. It consists of two parameters S_W and S_B . S_W indicates the difference between each parameter and their cluster mean and S_B indicates the difference of certain mean parameter between different clusters. For a good regrouping result, it has to satisfy that the within cluster difference is low and in-between cluster different is large. It means a large $S_W^{-1}S_B$ value. The purpose is to test WECs in a limited number of sea states with the largest difference between them, in order to make sure the WECs have been tested in as many different sea states as possible.

E. The influence of K on the regrouping quality

In previous research from Draycott [9], the regrouping quality (metric two) increases when K increased. This effect can be also observed for buoy data measured in Long Island. Taking 74896 directional wave spectra from 2008 to 2017 for analysis, and for simplicity check method C only, the non-directional wave spectra clustering. K is increased by 1 each time during K-means clustering starting from K=2 to K=356 and results are shown in Fig.2 and Fig.3. It can be seen the regrouping quality increases clearly as K increases, which means it is better to use a large K to identify the representative sea states. However, due to the time constraints of the tank testing, it is not possible to increase K without limit.

TABLE II

NUMBER OF BINS FOR HF RADAR DATA

METHOD	NO. OF H_s BINS	NO. OF T_e BINS	NO. OF θ_m BINS	NO. OF NON-EMPTY BIN
A (K=10)	3	4	-	10
B(K=10)	2	3	2	9
A(K=20)	6	4	-	19
B(K=20)	4	3	3	21

TABLE III

NUMBER OF BINS FOR LONG ISLAND BUOY DATA

METHOD	NO. OF H_s BINS	NO. OF T_e BINS	NO. OF θ_mBINS	NO. OF NON-EMPTY BIN
A (K=10)	3	4	-	10
B(K=10)	2	3	2	11
A(K=20)	6	4	-	18
B(K=20)	4	3	3	27

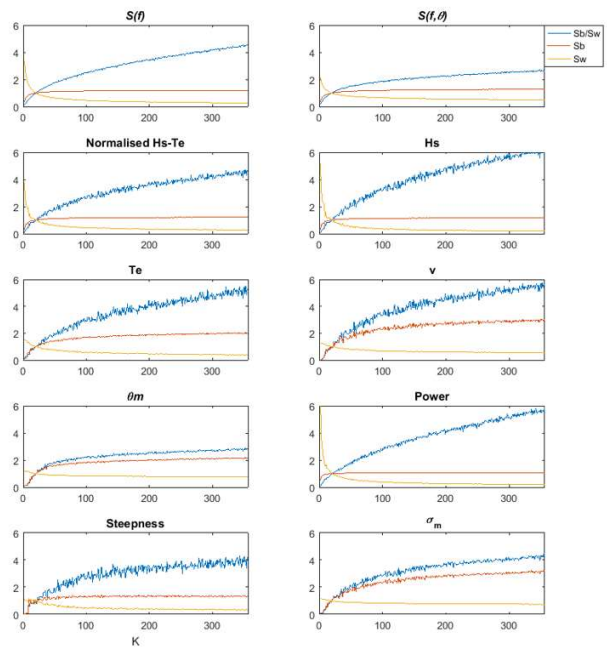

Fig. 3. Metric two for Long Island buoy measured data with K for seven parameters. All the values were normalized by the same parameter when K=20 for ease of visualisation.

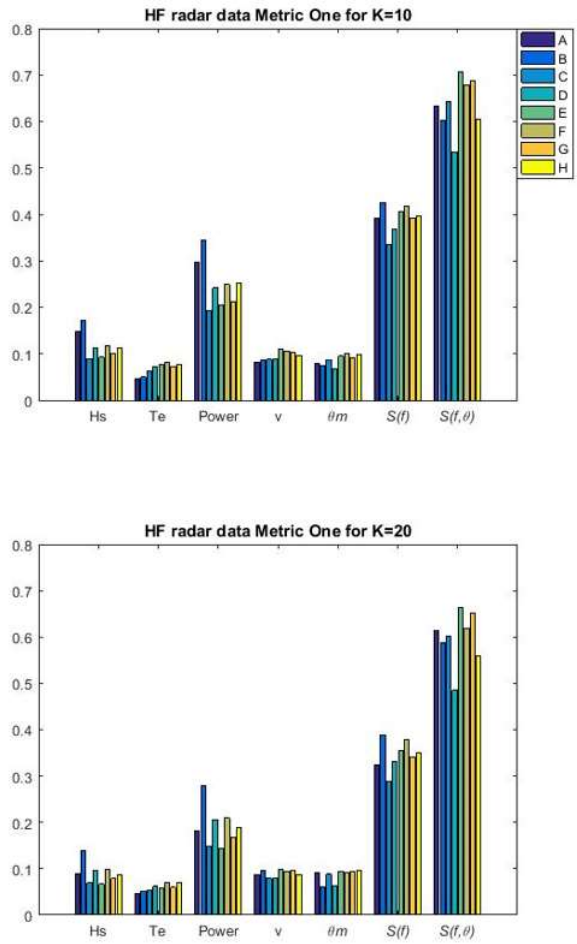

Fig. 4. Metric one for HF radar data for methods A to H.

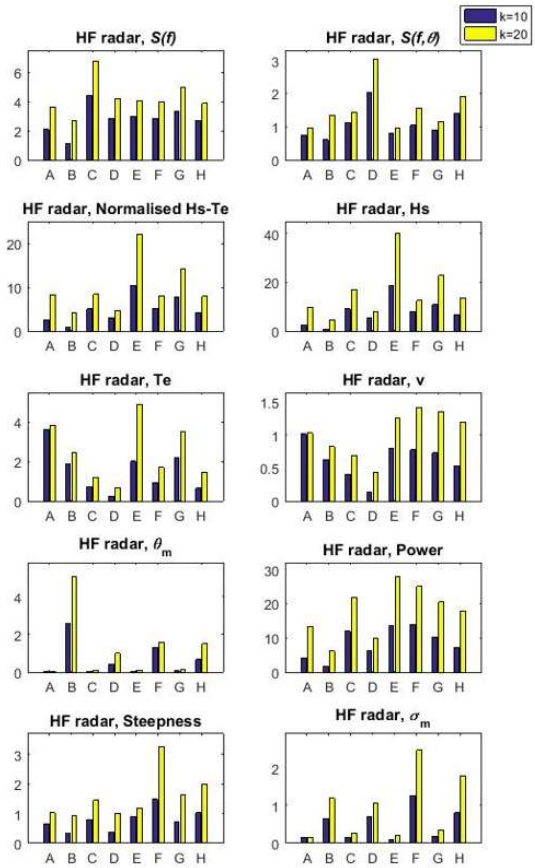

Fig. 5. Metric two for HF radar data for methods A to H.

F. Quality analysis for different regrouping methods

Here, all of the eight regrouping methods were tested using both HF radar data and Long Island buoy data. HF radar data comprises all of 3161 directional wave spectra. To save calculation time, the year 2017 is used for Long Island buoy data analysis, which comprises 8402 directional wave spectra. *K*=10 and *K*=20 are selected for comparison.

For method A and B, it is necessary to define the size of bins in order to create *K* non-empty bins. However, it is not an easy task, because the binning method might create empty bins. After consideration, the bins are created in TABLE II and TABLE III. It can be seen that the targeted K and the K achieved are sometimes different. For clustering method C to H, the K-means++ method is used for initial centres generation and the number of replicates is 100.

After running eight regrouping methods for HF radar data and Long Island buoy data, the regrouping results were then analysed by using metric one and metric two. Although the data set were measured by different instruments in different locations, the results showed great similarities. To save space, only the results for HF radar data are shown. The results are shown in Fig.4 and Fig.5.

From Fig.4, it can be seen that for the same parameter, K=20 has always a better quality than K=10, this is already verified in Fig. 2 and Fig.3. It can be seen that the metric one for one-dimensional parameters (*Hs, Te, Power, v,* θ_m) is always lower than that of non-directional spectra *S(f)*, the parameter for non-directional spectra is always lower than that of the directional spectra *S(f,* θ *)*. It can be explained like this, according to the definition of metric one, the difference between each instance with its cluster mean is defined as the summation of the absolute value of the difference between the two. But during the integration from *S(f,* θ *)* to *S(f)*, the differences between different directions were cancelled out. As a result, a part of the absolute differences between two instances disappeared during the integration and the metric one result for *S(f)* is always lower than that of *S(f,* θ*)*. Similarly, the metric one result for one-dimensional parameters is always lower than that of non-directional spectra *S(f).*

It can be seen that method C and method D resemble the best overall performance compared with other methods, it is because methods C and D are related with non-directional spectra and directional spectra respectively, which contain more detailed information compared with other methods that only consider several one-dimensional parameters.

It can also be seen that if a regrouping method uses a certain parameter during regrouping, then the quality for that parameter for this method is always good.

For metric two in Fig.5, It can be seen that the results of metric two are very different from metric one. For metric one, the values of different parameters ranges from 0 to 1, but the values for metric two ($S_w^{-1}S_B$) vary a lot among the

10 different parameters, so it is meaningless to plot metric two together in the same figure as it was done for metric one. Similar conclusion from metric one can be found in metric two; the quality for K=20 is better than that of K=10 and the quality of regrouping results is also closely related with the parameters participated in regrouping. For example method C provides the best quality for parameter *S(f)* and method D provides the best quality for parameter $S(f,\theta)$, and method E (clustering by normalized *Hs-Te*) provides the best quality for the parameter of normalised *Hs-Te*. However, there is no overall best method for regrouping as was seen in metric one.

The main difference between the metrics is that for one-dimensional parameters, metric two was lower (poorer) for methods C and D. The main reason is that the inter-cluster differences S_B of methods C and D are very low compared with other methods. From metric one it can be seen that methods C and D have the lowest intra-cluster differences which means a low value of S_W for one-dimensional parameters. But the value of S_B for them are also very low, which led to a low value of $S_W^{-1}S_B$. Methods C and D are still preferred because according to the definition of K-means clustering method, their only purpose is to find the most similar data instances and put them inside the same group. During the iterations of K-means clustering methods, it won't consider the differences between different clusters. As a result, it is unfair to 'blame' K-means method for not maximising the difference between different clusters. As a result, metric two is used more as an exploration tool rather than for numerical optimisation[9].So method C and method D show the overall best performance on representativeness compared with other regrouping methods.

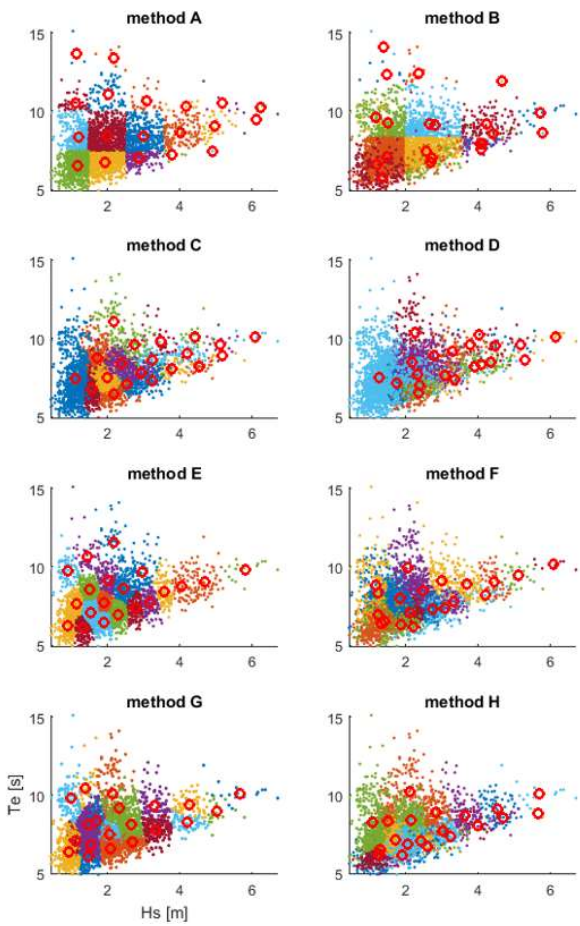

Fig. 6. Each group's wave spectra and the representative sea states (red circle) in *Hs-Te* space for HF radar data when *K*=20.

G. *Representative sea states in Hs-Te space*

Here, all the directional spectra used for regrouping in *Hs-Te* space are plotted together with their representative cases, see Fig.6. It can be seen that after regrouping, each method creates K (or close to K for binning methods) groups; each group has a different number of data instances inside with the same colour. The distribution of the representative cases in *Hs-Te* space is not uniform. Taking method C for example, all the data instances and their representative cases in *S(f)* space are shown in Fig. 7. The largest group consists of 740 data instances inside, with the lowest average energy (the lowest average *Hs* and *Te*). The same phenomenon can be observed for Long Island buoy data, where there are 5042 data instances in only one group with the lowest average energy. Having so many data inside a single group causes non-uniform distribution of the representative sea states in *Hs-Te* space. Another issue is that all of the representative sea states for K-means clustering methods from C to H have relatively low Te values compared with binning method A and B.

H. *Methods to obtain representative sea states with uniform Hs-Te distribution*

In previous work from Draycott [20], a two-step method was created, which is a combination of method A and method D. The first step was to create *K*/2 groups by method A, the second step was to create sub-groups by method D from step one. However, there is still subjectivity introduced by method A. For method A, the size of each bin is created subjectively and it is hard to guarantee there will be K non-empty bins for the next step. In this work, we create two new methods I and J.

The new methods I and J are also two-step methods, the first step is to use method C or method D, depending on the types of WECs. For WECs which are un-related with wave direction (point absorber, attenuator i.e.), method I is preferred because the performance of WECs is independent of the wave direction. In this case method C

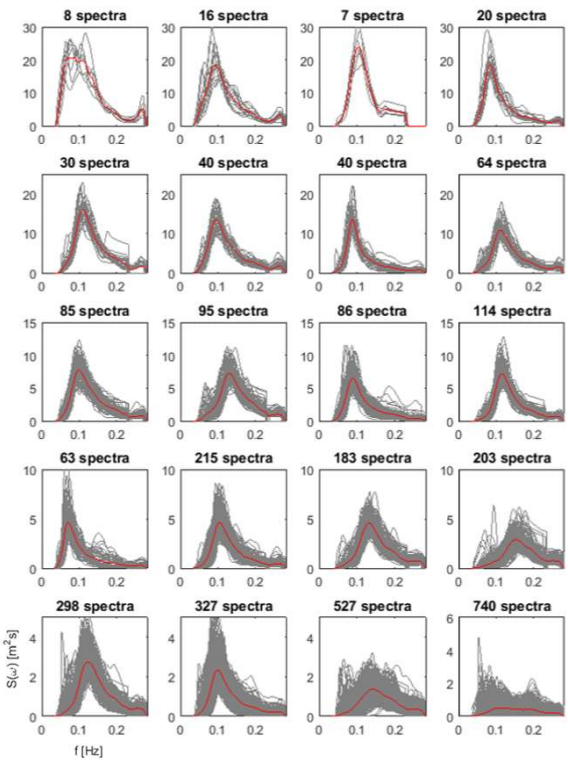

Fig.7. Each group's wave spectra and their representative sea states in *S(f)* space, title represents the number of each group's members for method C.

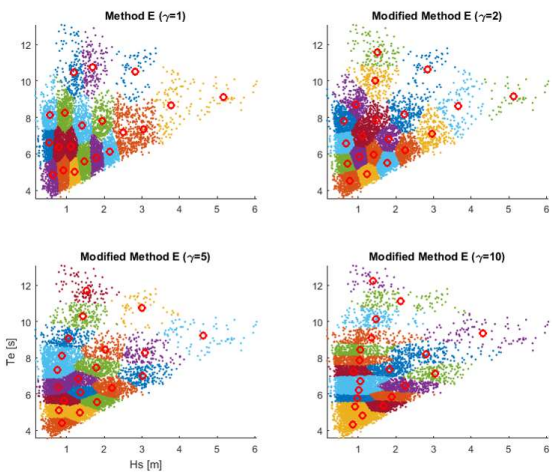

Fig. 8. The influence of γ on the regrouping results for weighted method E for Long Island buoy data.

is used for the first step. For WECs which are sensitive to wave direction (OWSC, overtopping devices i.e.), method J is preferred because the performance of WECs will be influenced by the incoming wave direction, the influence of wave direction cannot be neglected. In this case method D is used for the first step. The first step is to create *K*/2 sub-groups. Using these sub-groups as starting point, step two is to create more sub-groups from those obtained in step one by a new method, which is named weighted method E.

For weighted method E, we introduce a weighting parameter γ. Then the difference of two data instances can be calculated by equation (7):

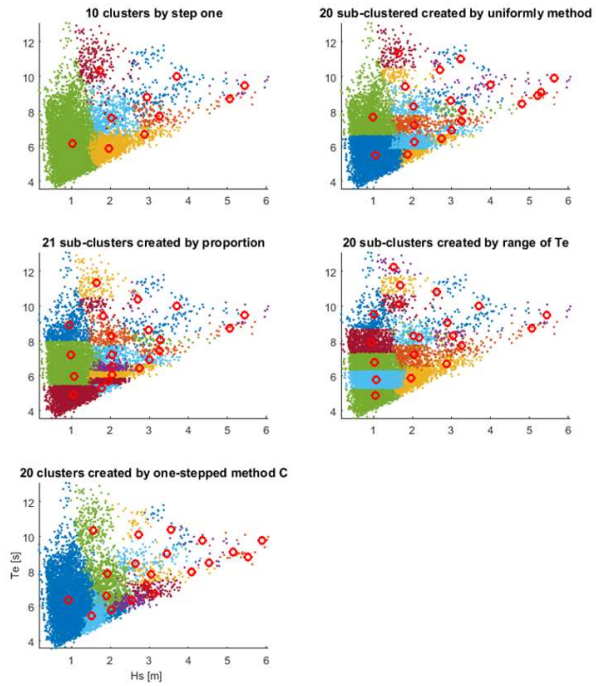

Fig. 9. Different methods to divide sub-groups for Long Island buoy data.

$$d(x_i, x_j) = (|x_{i1} - x_{j1}|^2 + \gamma \cdot |x_{i2} - x_{j2}|^2)^{1/2} \qquad (7)$$

In which x_1 and x_2 represent normalized *Hs* and *Te* respectively, γ is the weighting parameter ($\gamma \geq 1$). When $\gamma = 1$, the method is actually method E. The purpose of using the weighting parameter is to amplify the importance of *Te* during K-means clustering. In method E, *Hs* and *Te* were treated equally. However, the unit for *Hs* (m) and the unit for *Te* (s) are not the same. During the normalization procedure, the relative difference between different units can be eliminated, but there still exist the fact that different units are treated with equal importance. So it is necessary to re-balance the importance of *Hs* and *Te* for the K-means clustering.

Taking Long Island buoy data (*K*=20) for example, we use four weighting parameters for calculation. The final regrouping results and the representative sea states are heavily influenced by the weighting parameter, see Fig. 8. It can be seen that the larger γ is, the higher is the representative *Te*. When γ increases, the influence of *Hs* decreases. And it can be noticed that the maximum *Hs* decreases when increasing γ. It is necessary to find a proper γ to guarantee the balance between *Hs* and *Te*. Here we use γ=10.

After solving the low representative *Te* by the weighted method E, the next step is to solve the non-uniformity in *Hs*-*Te* space. The idea is simple, after getting *K*/2 sub-groups from step one, each sub-group will not be divided

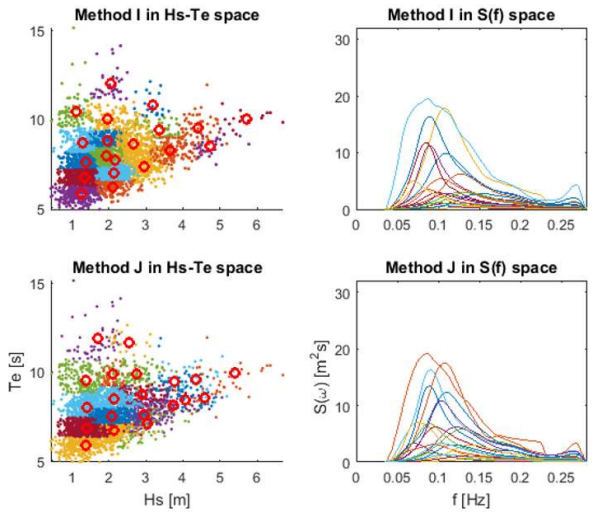

Fig. 10. Representative cases for HF radar data for method I and J in *Hs-Te* space and *S(f)* space.

evenly. Simply speaking, the more data instances contained in the sub-group, the more sub-groups it should be divided into. There are several ways to divide them, we can either simply divide each sub-group into two, or divide the sub-group according to the proportion of the data instances in it related with the total data set (150) or using a new method, which is to decide the number of groups for each sub-group based on their *Te* range. An example is shown in Fig.9. For Long Island buoy data, *K*=20, after the first step (method C), 10 sub-groups are created, then by using different methods are divided into 20 groups. For comparison, a one-step method C result is also plotted. It can be seen that dividing sub-groups by the range of *Te* provides the best uniformity in *Hs-Te* space.

Apply method I and method J to HF radar data to find out the representative sea states (*K*=20). In order to check the quality for method I and J, here we use metric two. The reason for that is method I and J all considered not only the representativeness within the same group but also the differences between the representative cases in different groups (a uniform distribution of the representative cases in *Hs-Te* space). As a result, it is not proper to judge the quality for regrouping with metric one. Only the result for Long Island buoy data is shown because the results measured by both instruments in different locations are similar. See Fig. 11. It can be seen that the quality for *Te* increased significantly for method I and J. This is expected, a good range of *Te*. They also showed very good quality for *S(f)* and *S(f,θ)*; this is a feature inherited from method C and D from the first step. The quality for bandwidth *v* also increased significantly. It can be seen that method I and J preserved good quality for many parameters and showed very good overall performance compared with other eight methods. It is also noticed that the quality for *Hs* is not very good. It is because the quality for *Hs* was negatively affected when using the weighted method E for the second step; when the quality of *Te* increased, the

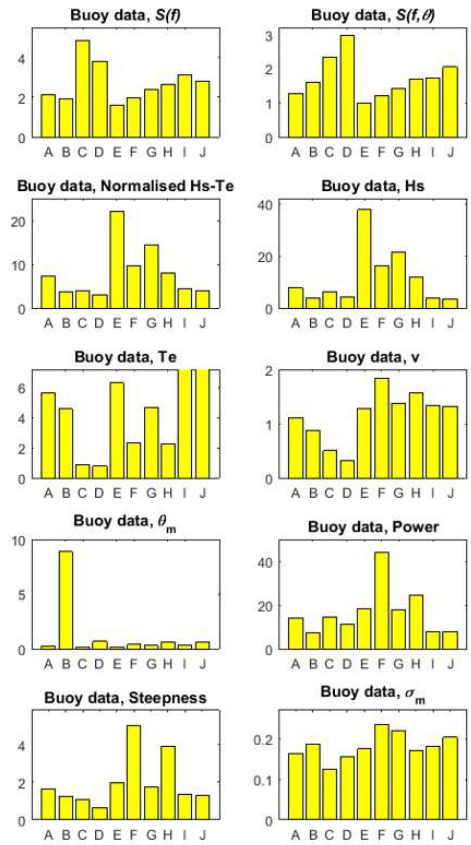

Fig. 11. Metric two for Long Island buoy data for method A to J for *K*=20

quality of *Hs* decreased naturally. It can be seen from Fig.10 that the representative sea states are uniformly distributed in *Hs-Te* space and the maximum value of representative *Te* is large and the non-directional spectra in *S(f)* space are distinct and have large peak values. As a result, method I and J showed a good capacity on finding representative sea states and having a good distribution for the representative cases in *Hs-Te* space.

III. CONCLUSIONS

In order to select representative sea states for WEC tank testing, data measured from two instruments (HF radar and buoy) were examined. In the beginning, eight regrouping methods, including two binning methods and six K-means clustering methods, were evaluated. The K-means method showed good applicability in sea states regrouping which were obtained in different sea states by different instruments. The quality of the results, as described by two metrics, showed that methods C and D provide the best overall performance when considering only the within cluster differences. However, we want to find a good uniformity for the representative cases *in Hs-Te* space and relatively large representative *Te* values for tank testing, methods C and D alone are not enough. Then methods I and J were created, they are both two-step methods that combined method C and D with a weighted

method E. The new methods provide very good results are suitable for finding the representative sea states for WEC tank testing.

ACKNOWLEDGEMENT

The authors would like to acknowledge the MaRINET2 project for supporting this research.

REFERENCES

[1] K. Koca, A. Kortenhaus, and H. Oumeraci, *Recent Advances in the Development of Wave Energy Converters*. 2013.

[2] D. Ross, "Power from sea waves," ed. United Kingdom,Oxford University Press, 1995.

[3] R. M. F. Robin Pelc, "Renewable energy from the ocean," *Marine Policy,* vol. 26, pp. 471–479, 2002.

[4] "House of Commons，Science and technology-seventh report: wave and tidal energy," LondonApril 30, 2001.

[5] M. Héder, "From NASA to EU: the evolution of the TRL scale in Public Sector Innovation."

[6] D. M. Ingram, *Protocols for the equitable assessment of marine energy converters*. 2011.

[7] K. Hasselmann *et al.*, "Measurements of wind-wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP)," *Ergänzungsheft 8-12,* 1973.

[8] L. J. Hamilton, "Characterising spectral sea wave conditions with statistical clustering of actual spectra," *Applied Ocean Research,* vol. 32, no. 3, pp. 332-342, 2010/07/01/ 2010.

[9] S. Draycott, "On the Re-creation of Site-Specific Directional Wave Conditions," 2017.

[10] J. Han and M. Kamber, "Data Mining: Concepts and Techniques, chapter Mining association rules in large databases," ed: Morgan Kaufmann Publisher, 2001.

[11] D. Arthur and S. Vassilvitskii, "k-means++: The advantages of careful seeding," in *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, 2007, pp. 1027-1035: Society for Industrial and Applied Mathematics.

[12] S. F. Barstow, J.-R. Bidlot, and S. Caires, "Measuring and analysing the directional spectrum of ocean waves," ed: COST Office, 2005.

[13] D. D. Crombie, "Doppler spectrum of sea echo at 13.56 Mc./s," *Nature,* vol. 175, no. 4459, p. 681, 1955.

[14] H.-H. Essen, K.-W. Gurgel, and T. Schlick, "Measurement of ocean wave height and direction by means of HF radar: an empirical approach," *Deutsche Hydrografische Zeitschrift,* vol. 51, no. 4, pp. 369-383, 1999.

[15] L. R. Wyatt, "Limits to the Inversion of HF Radar Backscatter for Ocean Wave Measurement," *Journal of Atmospheric and Oceanic Technology,* vol. 17, no. 12, pp. 1651-1666, 2000.

[16] J. J. Green and L. R. Wyatt, "Row-Action Inversion of the Barrick–Weber Equations," *Journal of Atmospheric and Oceanic Technology,* vol. 23, no. 3, pp. 501-510, 2006.

[17] "https://www.ndbc.noaa.gov/," *National Data Buoy Center website*.

[18] S. Draycott, T. Davey, D. Ingram, J. Lawrence, A. Day, and L. Johanning, "Applying site specific resource assessment: Methodologies for replicating real seas in the FLOWAVE facility," presented at the International Conference on Ocean Energy (ICOE), 2014-11-06, 2014.

[19] O. Maimon and L. Rokach, *Data Mining and Knowledge Discovery Handbook*. Springer-Verlag, 2005.

[20] S. Draycott, T. Davey, D. M. Ingram, J. Lawrence, A. Day, and L. Johanning, "Applying Site-Specific Resource Assessment: Emulation of Representative EMEC seas in the FloWave Facility," 2015/7/27/,