Supplementary Information for*:*

Model-informed classification of broadband acoustic backscatter from zooplankton in an *in situ* mesocosm

Muriel Dunn[1,2]*, Chelsey McGowan-Yallop[3]*, Geir Pedersen[4], Stig Falk-Petersen[5], Malin Daase[6], Kim Last[3], Tom J. Langbehn[7], Sophie Fielding[8], Andrew S. Brierley[9], Finlo Cottier[3,6], Sünnje L. Basedow[6], Lionel Camus[1] & Maxime Geoffroy[2,6].

[1]Akvaplan-niva AS, Fram Centre, Postbox 6606, Stakkevollan, 9296 Tromsø, Norway

[2]Centre for Fisheries Ecosystems Research, Fisheries and Marine Institute of Memorial University of Newfoundland, St. John's A1C 5R3, Canada

[3]Scottish Association for Marine Science, Oban, Argyll PA37 1QA, UK

[4]Institute for Marine Research, 5005 Bergen, Norway

[5]Independent scientist, 9012 Tromsø

[6]Department of Arctic and Marine Biology, UiT The Arctic University of Norway, 9036 Tromsø, Norway

[7]Department of Biological Sciences, University of Bergen, 5020 Bergen, Norway

[8]British Antarctic Survey, Natural Environment Research Council, High Cross, Cambridge CB30ET, United Kingdom

[9]Pelagic Ecology Research Group, School of Biology, Scottish Oceans Institute, Gatty Marine Laboratory, University of St Andrews, St Andrews, KY16 8LB, United Kingdom

* These authors contributed equally: Muriel Dunn and Chelsey McGowan-Yallop

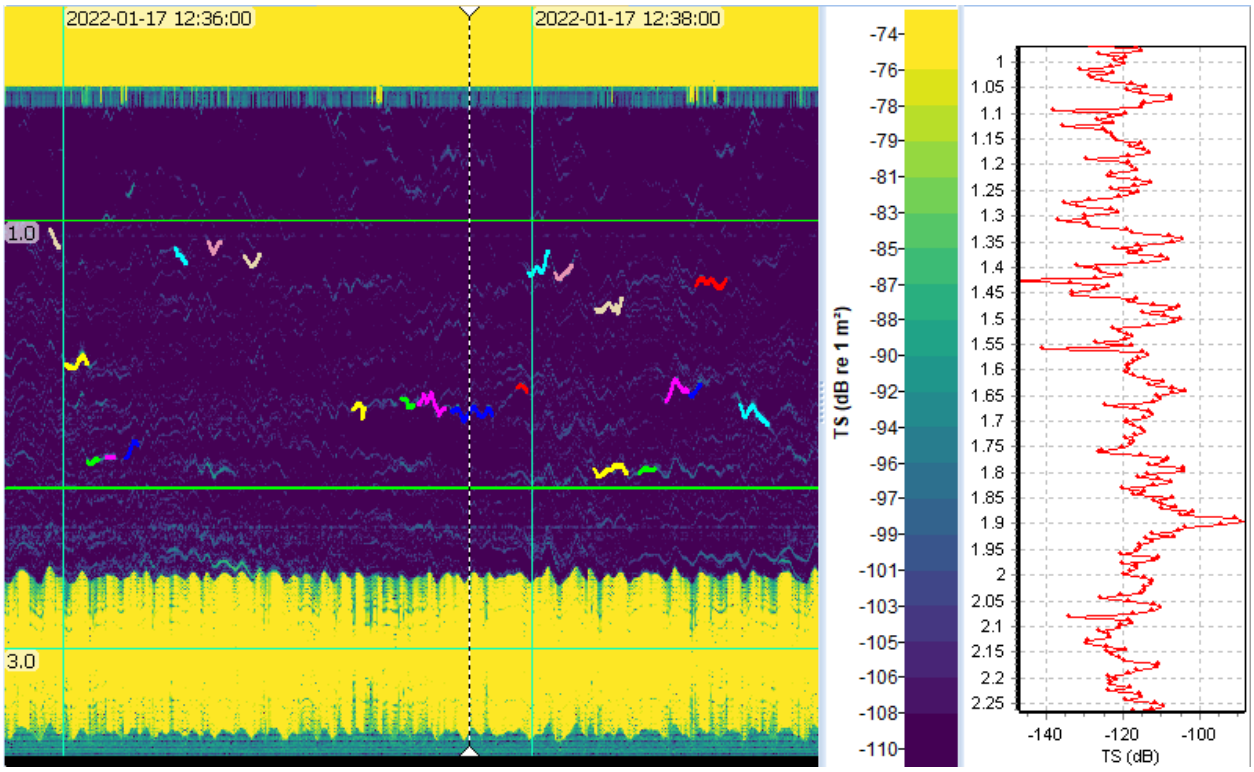

Figure S1. Left: Pulse compressed target strength echogram excerpt of 3.5 minutes from the total 3 hours of data collected during the AZKABAN mesocosm experiment. The bottom of the net is at 3 m. The tracks were identified from single target detections and are highlighted as bright coloured tracks superimposed on the echogram. The analysis domain for the experiment was bounded by the green horizontal lines at 1 m and 2.25 m to minimize interference from near-field and net bottom. Image was exported directly from Echoview using the "Export to Image" option. *Right:* Graphed ping at the dotted vertical line in the left panel. The graphed ping is bounded by the analysis domain limits (1 m and 2.25 m). Note the stronger signal at 1.9 meters shows the dark blue track that coincides with the dotted verticak line.

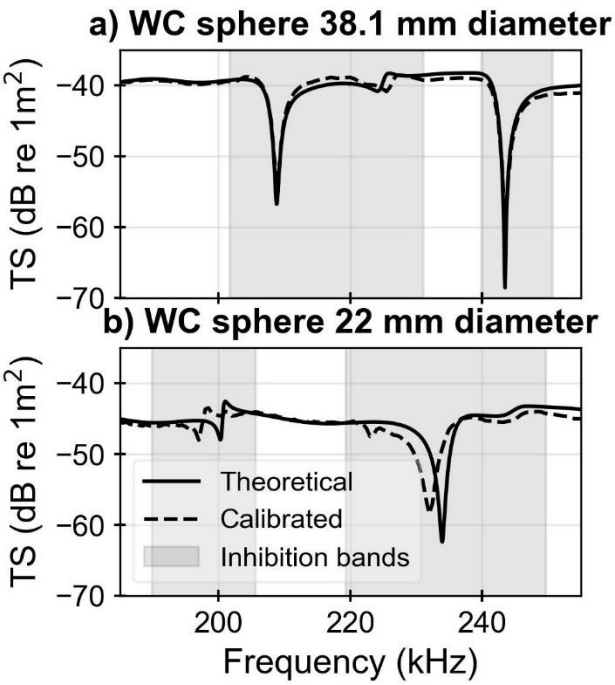

Figure S2. Theoretical target strength spectra (solid line) and calibrated target strength spectra (dashed line) for the a) 38.1 mm diameter tungsten carbide (WC) sphere and b) 22 mm diameter tungsten carbide (WC) sphere. The inhibition bands for each sphere are indicated by the grey area and were not used for to calculate the final calibration parameters.

The calibration results from the 38.1 mm sphere were preferred over the 22 mm when both spheres had available calibrations values because of the lower root mean square values (~0.2), a measure of error, and a higher signal-to-noise ratio. We observed a lateral shift in the nulls between 220 - 240 kHz and between 195- 205 kHz with the 22 mm sphere which had a spark-eroded suspension point (Supplementary Information Figure S1b). Large inhibition bands were used to reduce the effect of the spark-eroded suspension point (Renfree *et al.*, 2020).

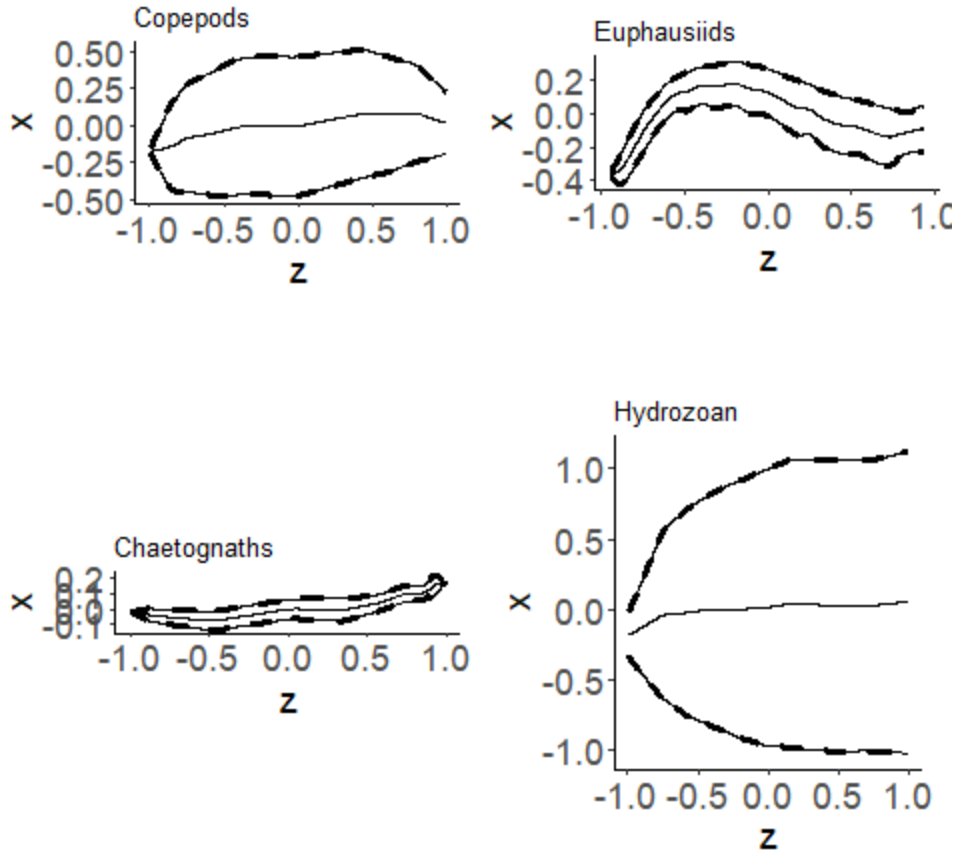

Figure S3: Shape files for each species scattering model ensembles. The shapes are selected as a representative and the best picture of an individual from each species.

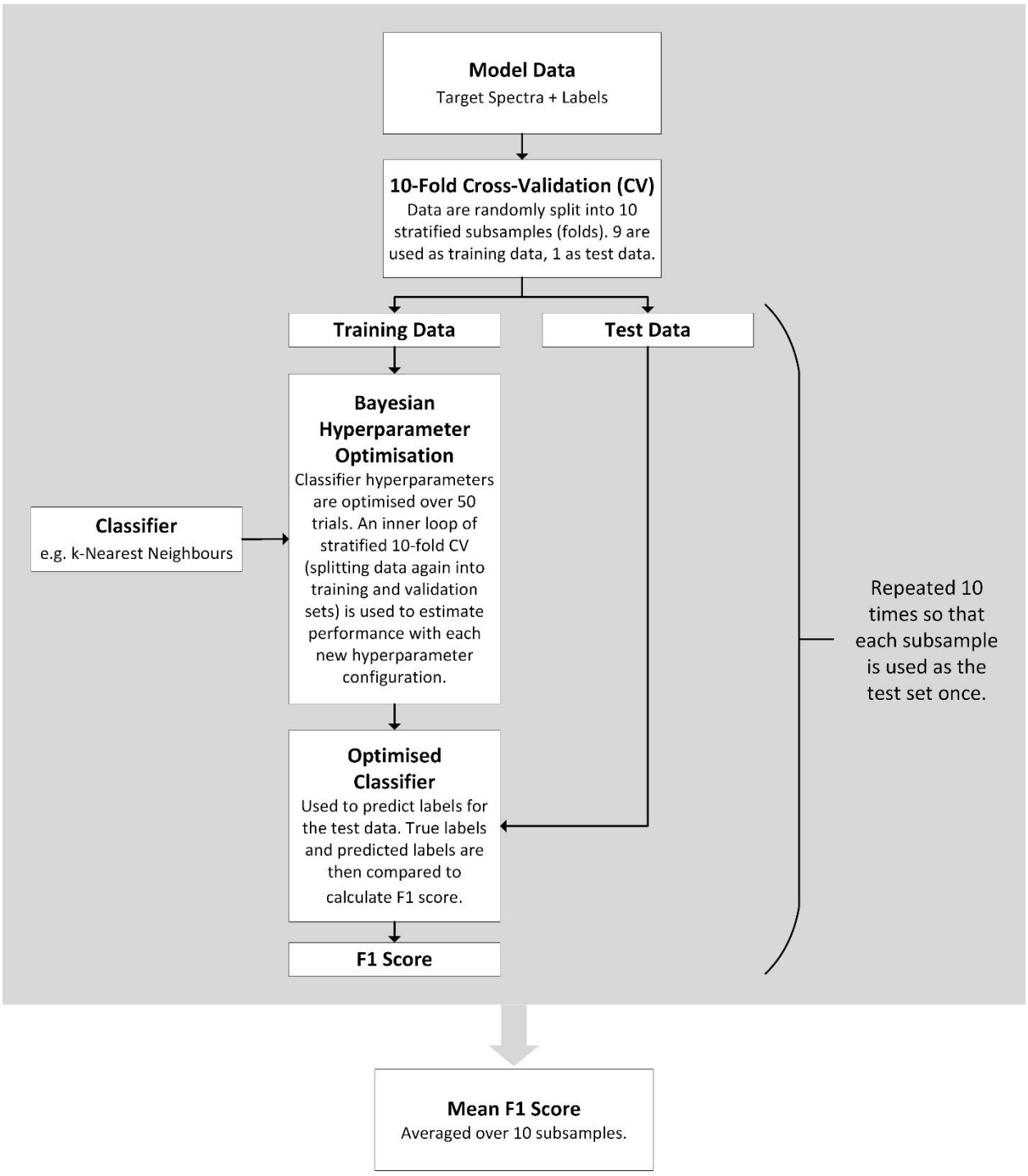

Figure S4: Nested cross-validation procedure used to estimate performance for a classifier trained on the entire labelled modelled dataset. The 'Model Data' block includes the target spectra results calculated by the scattering models for each taxonomic group, the labels are the associated taxonomic group parameterised for the scattering model result. F1 score is used as the evaluation metric.

Table S1: Single Target Detection - wideband algorithm parameters for target analysis (target detection) and for target removal in noise level analysis (noise level).

Parameter	**Target detection**	**Noise level**
Compensated TS Threshold (dB re $1m^2$)	-120.0	-120.0
Pulse Length Determination Level (dB re 1W)	6.0	5.0
Min. Normalised Pulse Length	0.5	0.05
Max. Normalised Pulse Length	1.5	1.5
Min. Target Separation (m)	0.5	0.001
Max. Beam Compensation (dB re $1m^2$)	39.0	18.0
Max SD of Minor-Axis Angle (deg)	0.6	0.5
Max SD of Major-Axis Angle (deg)	0.6	0.5

We used the pulse length determination level parameter to threshold the signal strength of the broadband targets relative to the background noise. The pulse length determination level is the value below the peak value at which pulse length is measured (Soule *et al.*, 1996). Here, the pulse length determination level parameter provided some noise filtering by ensuring that accepted peaks have a prominence of at least 6 dB re 1W (Table 2). If the noise level were within 6 dB re 1W of the peak value, the normalised pulse length should exceed the thresholds typically used in single target detection, and the peak would be rejected. A maximum two-way beam compensation of 6 dB re $1m^2$ was used to restrict all targets to the 3 dB re $1m^2$ beam angle. A minimum target separation of 0.4 m was used to reduce the likelihood of incorporating multiple targets into each target spectra. We further filtered the detected targets to be within the 4 degrees off-axis angle and 6 dB re $1m^2$ beam compensation.

Table S2: Target tracking algorithm parameters.

	Major axis (degree)	Minor axis (degree)	Range (m)	TS (dB re $1m^2$)	Ping gap
Track detection					
Alpha	0.7	0.7	0.5		
Beta	0.5	0.5	0.3		
Target gates					
Exclusion distance (m)	0.15	0.15	0.1		

Missed ping expansion	0	0	0		
Weights	30	30	60	60	0
Track acceptance					
Minimum number of single targets in a track			4		
Minimum number of pings in track			4 pings		
Maximum gap between single targets			2 pings		

Table S3. Confusion matrix averaged over all folds in the outer stratified 10-fold cross-validation for kNN (mean ± standard deviation) for normalized modelled target spectra.

	Predicted species			
Species	Chaetognaths	Copepods	Euphausiids	Hydrozoans
Chaetognaths	56.6 ± 5.0	8.9 ± 1.4	13.1 ± 2.5	22.3 ± 2.9
Copepods	2.7 ± 1.2	93.3 ± 3.1	3.8 ± 2.4	0.2 ± 0.4
Euphausiids	11.8 ± 3.6	12.5 ± 3.7	69.2 ± 4.1	6.5 ± 1.7
Hydrozoans	23.7 ± 5.2	1.7 ± 1.3	11.0 ± 3.3	63.6 ± 5.2

Table S4. Confusion matrix averaged over all folds in the outer stratified 10-fold cross-validation for LightGBM (mean ± standard deviation) for normalized modelled target spectra.

	Predicted species			
Species	Chaetognaths	Copepods	Euphausiids	Hydrozoans
Chaetognaths	53.2 ± 4.9	6.8 ± 1.5	13.4 ± 3.7	26.6 ± 4.3
Copepods	2.6 ± 1.4	90.8 ± 2.7	5.2 ± 2.0	1.4 ± 1.1
Euphausiids	9.1 ± 2.3	8.9 ± 2.6	72.2 ± 3.8	9.8 ± 2.2
Hydrozoans	18.9 ± 5.6	1.6 ± 1.1	10.4 ± 2.2	69.1 ± 5.3

Table S5. Confusion matrix averaged over all folds in the outer stratified 10-fold cross-validation for SVM (mean ± standard deviation) for normalized modelled target spectra.

	Predicted species			
Species	Chaetognaths	Copepods	Euphausiids	Hydrozoans
Chaetognaths	36.8 ± 3.9	20.5 ± 3.1	19.0 ± 4.9	23.7 ± 2.5
Copepods	0.9 ± 1.2	79.7 ± 3.0	16.5 ± 2.5	2.9 ± 1.4
Euphausiids	4.0 ± 2.1	18.4 ± 4.9	70.2 ± 4.1	7.4 ± 2.6
Hydrozoans	24.7 ± 4.7	5.8 ± 2.9	14.7 ± 2.6	54.8 ± 5.6

Table S6. Classifier F1 scores estimated through nested cross-validation for kNN (mean ± standard deviation) for 5 different continuous frequency bandwidths.

	70 kHz (45 - 90 kHz)	**120 kHz (90 - 170 kHz)**	**200 kHz (185 - 255 kHz)**	**333 kHz (283 - 383 kHz)**	**Full (45-383 kHz)**
Class-weighted	0.78 ± 0.02	0.86 ± 0.01	0.70 ± 0.02	0.64 ± 0.02	0.92 ± 0.02
Chaetognaths	0.60 ± 0.03	0.76 ± 0.03	0.58 ± 0.04	0.59 ± 0.03	0.88 ± 0.02
Copepods	0.93 ± 0.01	0.90 ± 0.02	0.87 ± 0.02	0.82 ± 0.03	0.95 ± 0.01
Euphausiids	0.76 ± 0.03	0.85 ± 0.02	0.70 ± 0.03	0.57 ± 0.03	0.89 ± 0.03
Hydrozoa	0.81 ± 0.01	0.92 ± 0.02	0.66 ± 0.04	0.57 ± 0.04	0.96 ± 0.01

Table S7. Classifier F1 scores estimated through nested cross-validation for kNN (mean ± standard deviation) for normalized scattering models with copepod material properties from Antarctica.

	Predicted species			
Species	Chaetognaths	Copepods	Euphausiids	Hydrozoans
Chaetognaths	57.0 ± 6.7	10.0 ± 2.6	13.8 ± 4.4	19.1 ± 4.8
Copepods	3.7 ± 2.5	91.3 ± 2.8	4.7 ± 1.7	0.3 ± 0.6
Euphausiids	10.7 ± 3.1	13.3 ± 5.3	68.9 ± 6.2	7.1 ± 2.0
Hydrozoans	24.5 ± 3.6	2.8 ± 1.5	11.1 ± 3.1	61.6 ± 4.2

Code S1. Optimised kNN model as determined by Bayesian hyperparameter optimisation using HyperOpt-Sklearn 1.0.3.

```
KNeighborsClassifier(algorithm='kd_tree', leaf_size=36, metric='l2', n_jobs=-1,
                     n_neighbors=1, p=1.6822439326937353, weights='distance')
```

Code S2. Optimised LightGBM model as determined by Bayesian hyperparameter optimisation using HyperOpt-Sklearn 1.0.3.

```
LGBMClassifier(boosting_type='goss', colsample_bytree=0.5097647467361791,
                learning_rate=0.012631457372918417, max_delta_step=0,
                max_depth=7, min_child_weight=1, n_estimators=3400,
                num_leaves=46, objective='binary',
                reg_alpha=0.0002621872624005705, reg_lambda=1.8302350814675785,
                scale_pos_weight=1, seed=0, subsample=0.810277827141478)
```

Code S3. Optimised SVM model as determined by Bayesian hyperparameter optimisation using HyperOpt-Sklearn 1.0.3.

```
SVC(C=3.003839970561586, coef0=0.4612291525680026,
    decision_function_shape='ovo', degree=1, random_state=3, shrinking=False,
    tol=5.715337840164192e-05)
```

Code S4. Optimised kNN model as determined by Bayesian hyperparameter optimisation using HyperOpt-Sklearn 1.0.3. trained with copepod parametrised with Antarctic waters material properties

```
KNeighborsClassifier(leaf_size=22, n_jobs=-1, n_neighbors=4,
                     p=2.6840910661386785, weights='distance')
```

References.

Renfree, J.S., Andersen, L.N., Macaulay, G., Sessions, T.S., Demer, D.A. 2020. Effects of sphere suspension on echosounder calibrations. *ICES Journal of Marine Science*, 77: 2945-2953.