Supplementary Material: Euphausiid classification algorithm

Matching acoustic records to zooplankton samples

Multi-frequency acoustic data along with zooplankton samples were collected from the CCGS W.E. Ricker and CCGS John P. Tully between 2009 and 2015. Both vessels were equipped with EK60 split-beam echosounders and included 38 kHz and 120 kHz transducers (ES38B and ES120-7 models) that were calibrated on an annual basis following recommended procedures (Demer *et al.*, 2015). Acoustic data were collected following the same protocols as described in the Methods section. Zooplankton were collected using several sampling instruments as part of oceanographic missions that were simultaneously recording acoustic data. Net sampling included vertical paired Bongo hauls (2 x 0.25 m^2 opening and 230-micron nylon mesh) as well as oblique and horizontal hauls using an Hydrobios midi MultiNet (0.25 m^2 square opening) and a BIONESS (1 m^2 opening) with 333-micron nylon mesh. Surveys occurred off the west coast of Vancouver Island, British Columbia, and in the Strait of Georgia, including adjacent inlets.

At each daytime zooplankton sampling station, acoustic data were scrutinized to detect the presence of a distinct zooplankton scattering layer between depths of 50 m to 250 m on the 120 kHz echogram, using a -70 dB threshold on the S_v (20logR) data. Data with no apparent discrete layer, as well as those including fish schools, individual fish targets, or overlapping diffuse scattering layers were discarded. For each net sample associated with the selected acoustic region, zooplankton specimens were identified to species or lowest identifiable taxon, counted, and placed into a predefined size class with associated mean length and weight. For vertical hauls, stations were only used for acoustic classification when more than 50% of the total biomass of the catch was comprised of euphausiids larger than 5 mm. The 50% cut-off was chosen to ensure the retention of net samples which passed through multiple layers of smaller plankton (given the fine mesh) and to account for net avoidance of euphausiids. For targeted horizontal net hauls that were towed through a distinct zooplankton scattering layer, only samples that were comprised of 75% euphausiids larger than 5 mm were retained for further analysis. After applying these selection criteria, 38 acoustic regions were selected and associated with euphausiids (2 from BIONESS samples, 15 from MultiNet samples, and 21 from vertical Bongo net samples).

Difference in mean volume backscattering strength

Acoustic data were pre-processed as described in the Methods section. Mean volume backscattering strength (MVBS) values were calculated for each 5 ping by 5 sample cells. Matching cells at 38 kHz and 120 kHz with MVBS values below -70 dB at either frequency were removed. The difference in mean volume backscattering strength (ΔMVBS) between the 120 kHz and 38 kHz data was calculated in the remaining matched cells: $\Delta MVBS_{120-38} = MVBS_{120} - MVBS_{38}$. There was no significant difference in mean ΔMVBS between euphausiid layers based on reference samples dominated by *Euphausia pacifica* and reference samples dominated by *Thysanoessa spinifera*, as there was significant overlap in animal size and high variability of ΔMVBS values between samples. Overall mean $\Delta MVBS_{120-38}$ was 13.1 ± 3.1 dB (Figure S1). There was no significant difference between scattering layers sampled with the BIONESS, MultiNet, or Bongo nets.

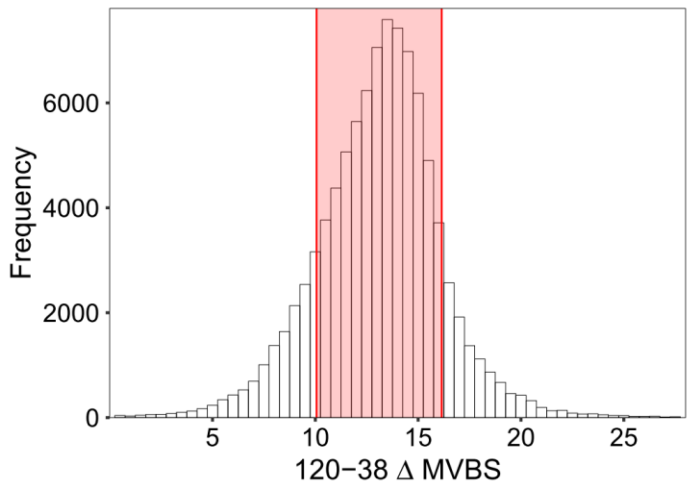

Figure S1: Distribution of all $\Delta MVBS_{120\text{-}38}$ values for scattering layers associated with euphausiid samples. The red region represents one standard deviation around the mean.

Classification of euphausiids

The range of $\Delta MVBS_{120\text{-}38}$ that represent euphausiids, defined as one standard deviation around the mean, was used to create a classification mask to exclude any acoustic data falling outside of that range. A conservative approach to use one standard deviation around the mean was used to avoid misclassification, given the high variability of $\Delta MVBS_{120\text{-}38}$ values and the likelihood of having other organisms (e.g. crustacean nekton) other than *E. pacifica* and *T. spinifera* overlap in their frequency response. The workflow of the classification algorithm is illustrated in Figure S2, and can be applied to cleaned, but unsupervised, acoustic data to obtain volumetric (S_v) and integrated area backscatter (e.g. NASC) associated with euphausiids in the Northeast Pacific.

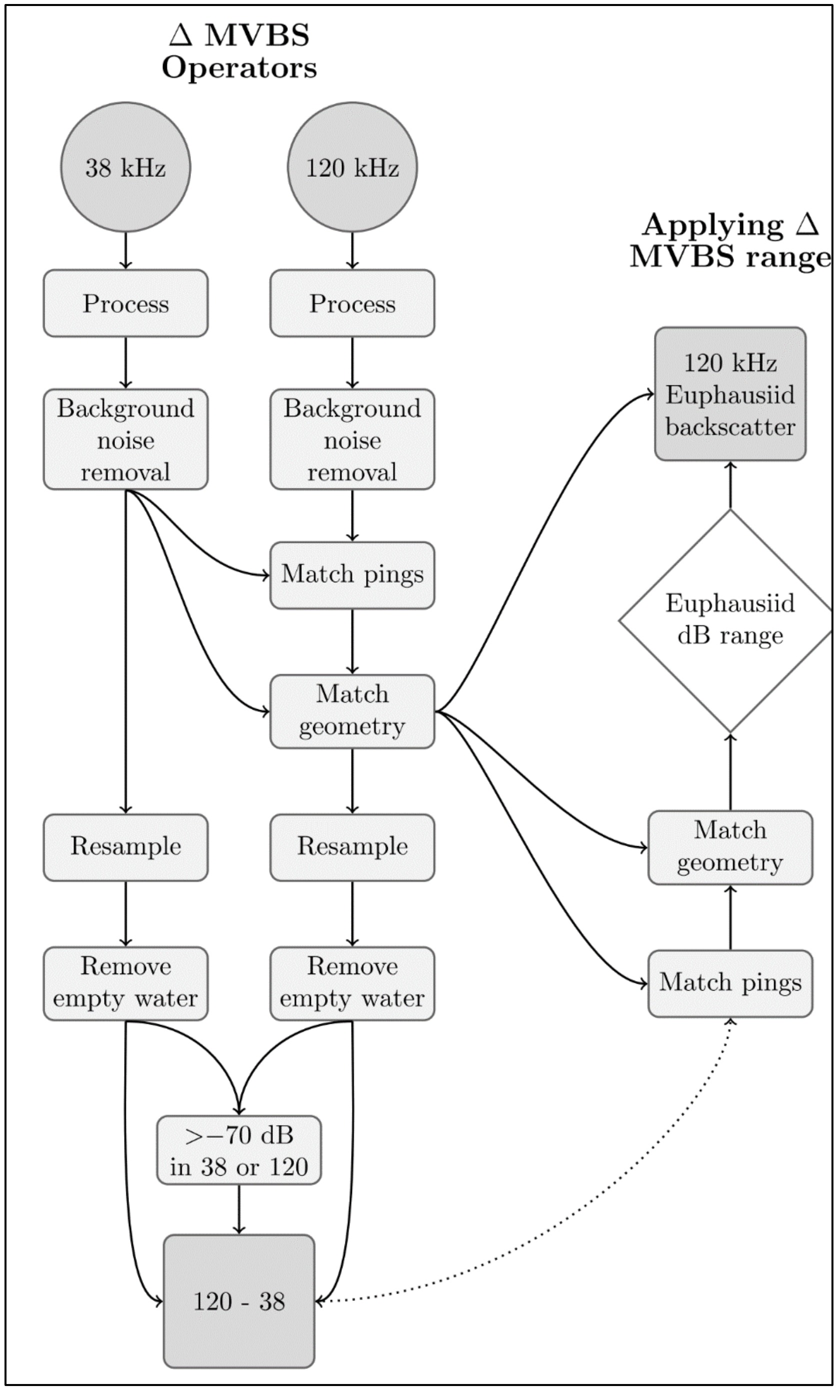

Figure S2. Workflow of the operators used to classify euphausiids from acoustic data collected at 38 kHz and 120 kHz (circles). The rectangles represent the operators applied to the data. "Process" is the cleaned data after removal of noise artifacts, surface and bottom exclusions, as well as user's identified bad data. Resample is the average of 5 pings by 5 samples. The diamond represents a mask that is applied to the data to exclude any values that are outside the defined range of $\Delta MVBS_{120-38}$ for euphausiids, as applied to the original resolution of the acoustic data.

Literature Cited

Demer, D. A., Berger, L., Bernasconi, M., Bethke, E., Boswell, K., Chu, D., Domokos, R., *et al.* 2015. Calibration of acoustic instruments. ICES Cooperative Research Report, ICES Cooperative Research Report, 326.