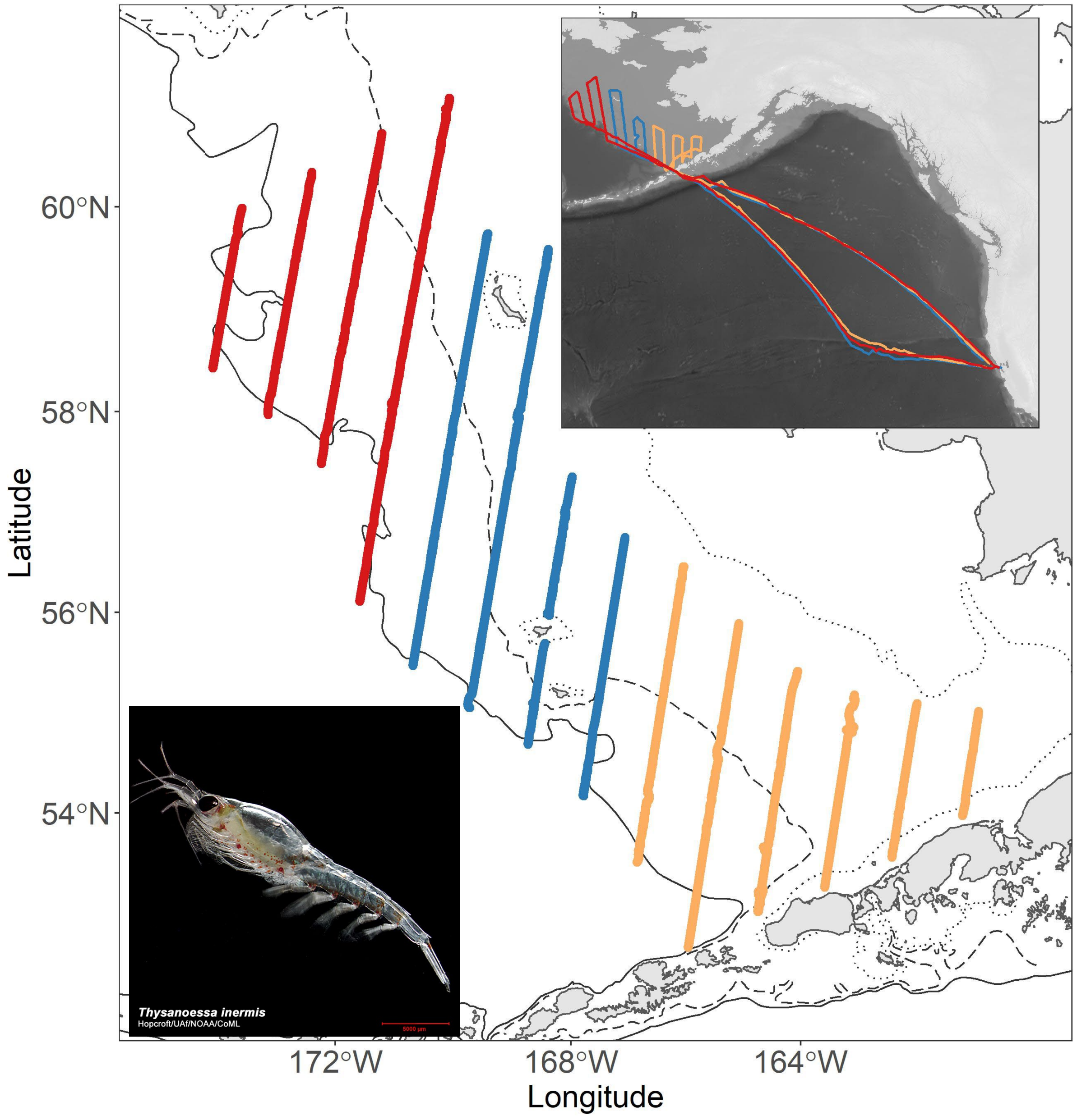
EBS krill index update: 2020 Saildrone acoustic survey

Mike Levine, Alex De Robertis

EBS krill index update: 2020 Saildrone acoustic survey

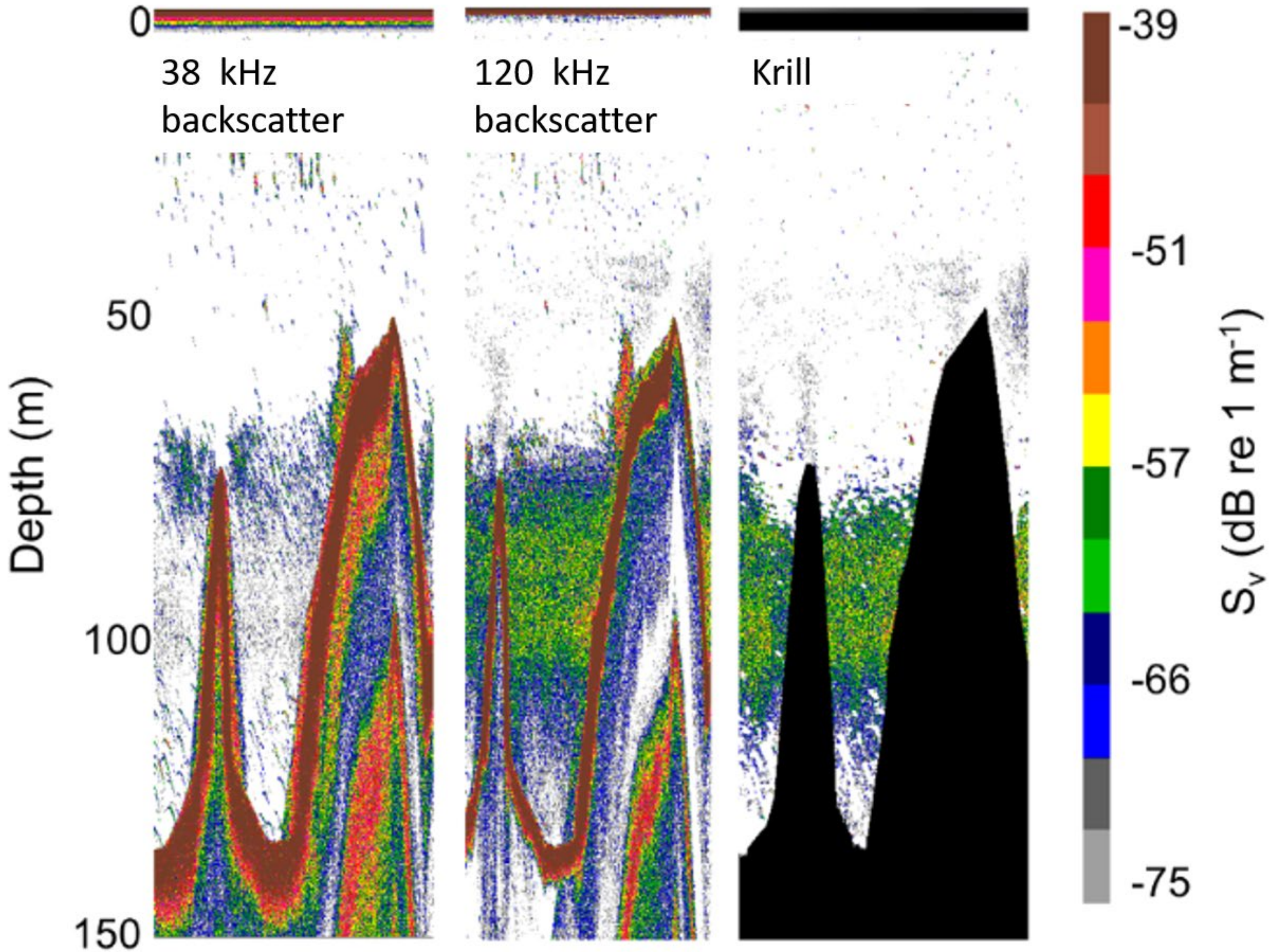
- Euphausiids (‘krill’) are a key food for many species of importance in the EBS, including walleye pollock.
- MACE provides an estimate of krill abundance in the EBS going back to 2004
- Estimate used in Bering Sea ESR, Bering Sea Report Card, ESP
- In 2020, the midwater pollock survey was conducted by 3 Saildrones equipped due to Covid-19 pandemic; these estimates are included in pollock stock assessment time series
- Can we estimate krill abundance as well?



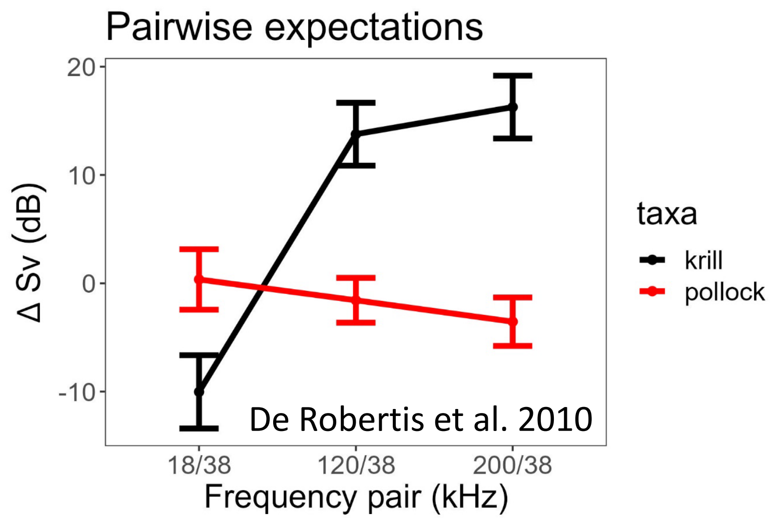
De Robertis et al. 2021

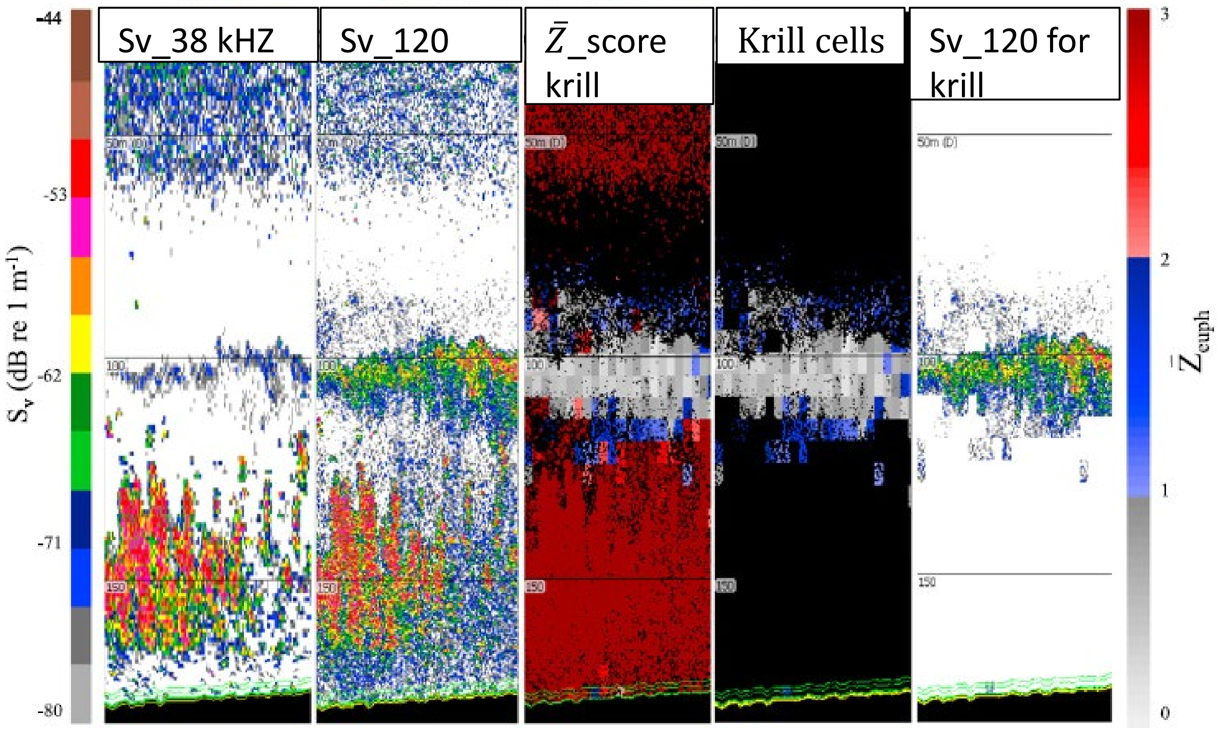
Multifrequency krill identification

- Krill have a strong frequency response
- They form continuous layers



Multifrequency krill identification (Z-score)





Ressler et al. 2012

Compute pairwise frequency differences for each acoustic cell (i.e. $Sv_{120\ kHz} - Sv_{38\ kHz}$ etc.)

↓

Compute *Z*-score (number of standard deviations from the mean expectation) for each class + frequency pair

↓

Compute $\bar{Z}$; identify cell as consistent with a given taxa if $\bar{Z} \leq 2$

↓

For krill cells: integrate in 0.5 nmi (length) * 20 m (depth) survey intervals

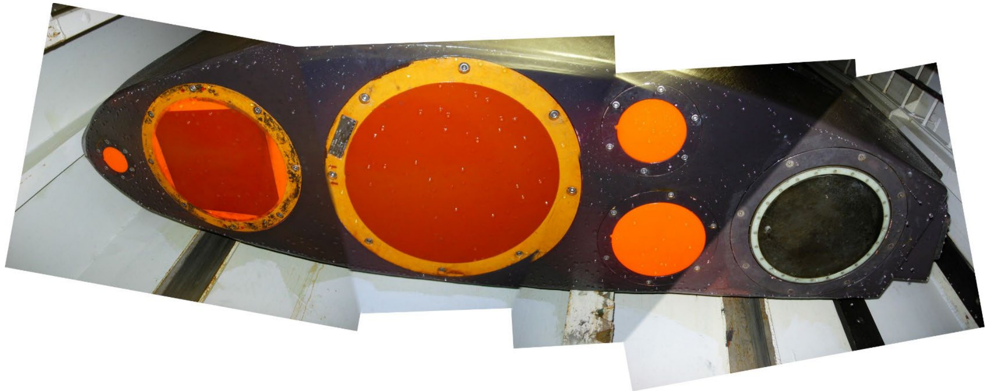
↓

Reject marginal krill cells (cases where interval average is marginally consistent with krill)

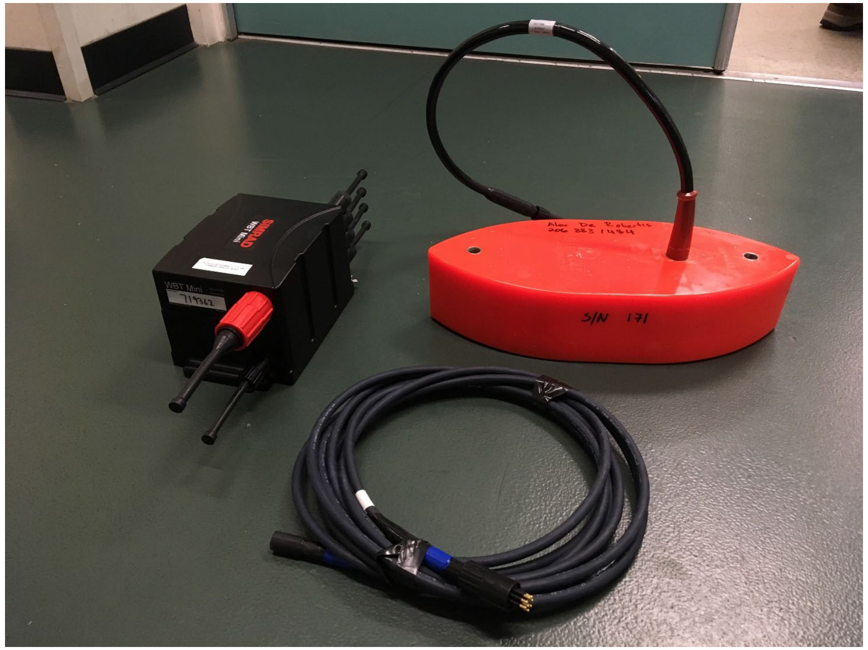
Can we do this with less?

- Krill abundance estimate is based on a 4-frequency identification method
  - (18 kHz/38 kHz/120 kHz/200 kHz)

- Can we produce a comparable estimate with less acoustic information (38 kHz/200 kHz)?

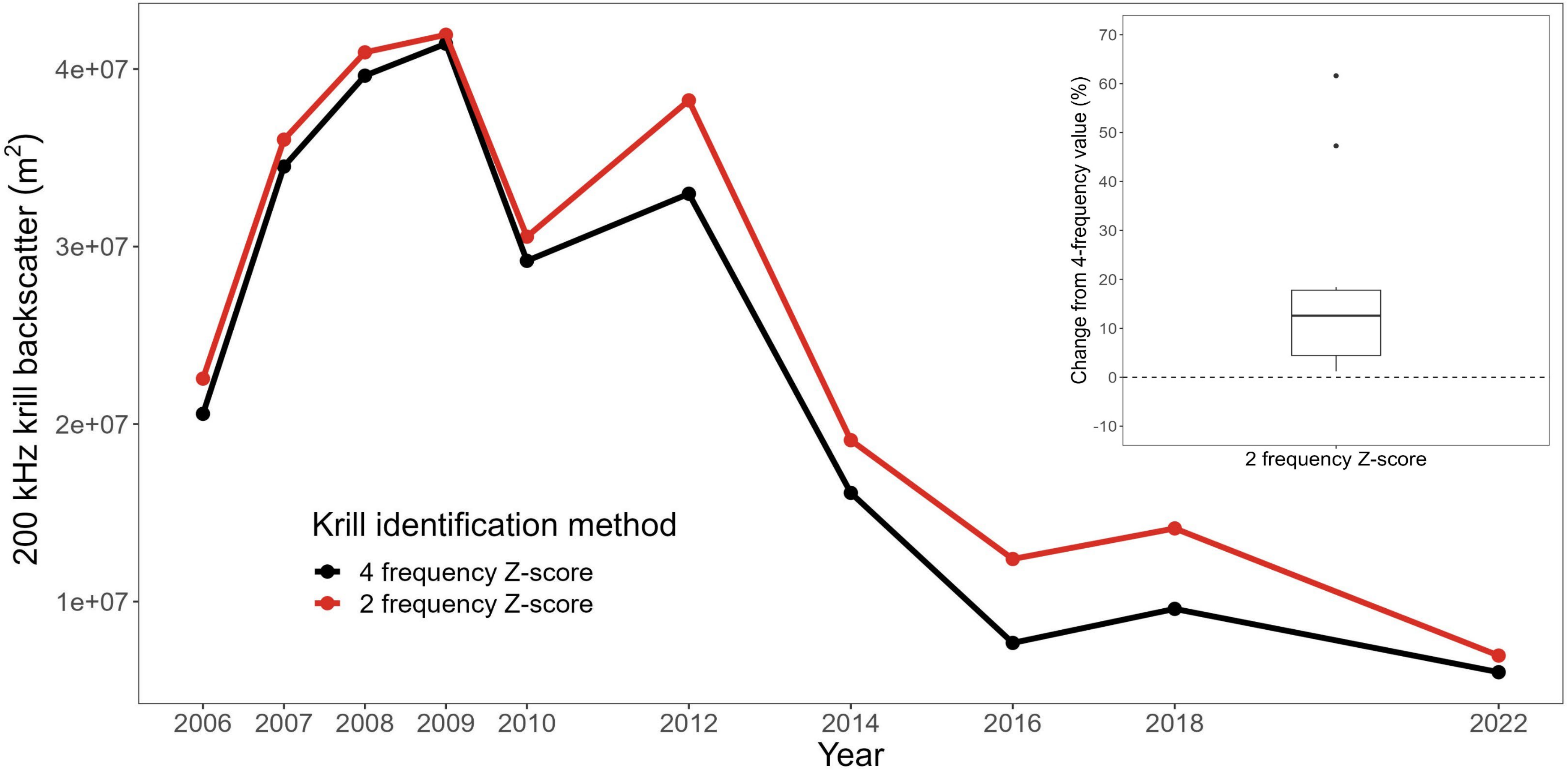


VS



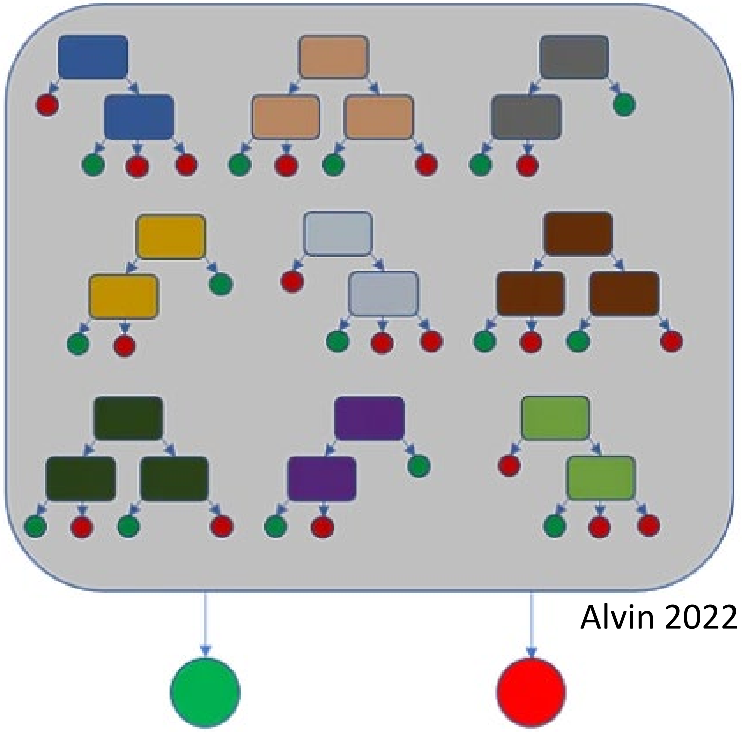
Z-score identification struggles with fewer frequencies

- Can we replicate the 4-frequency approach with only 2 frequencies?
- Test dataset: 10 EBS surveys from 2006-2022



- Estimates are biased high
  - Fewer pairwise frequency comparisons (1 vs 6)
  - Potentially missing important acoustic contrasts

Random forest classification



Alvin 2022

- Random forests: a collection of many classification decision trees (i.e. yes/no)
- Each tree is trained on a different sample of the data, and selects *n* predictors randomly the total available
- Classification is based on 'wisdom of the crowd'- the majority vote among the trees
- Models are simple to train, require minimal tuning, and generally show low bias and variance
- Model accuracy can be assessed via cross-validation

Random forest classification

Training data:
Krill identifications from 10 EBS surveys; classified using 4-frequency Z score method

Predictors:
Sv_200 kHz
ΔSv 200 kHz – 38 kHz
Seafloor depth at cell location
Cell distance off seafloor

Sv_38 kHz
Latitude
Longitude
Time of day
Cell Proportion of water column depth

Prediction:
For each acoustic cell:
krill/ not krill

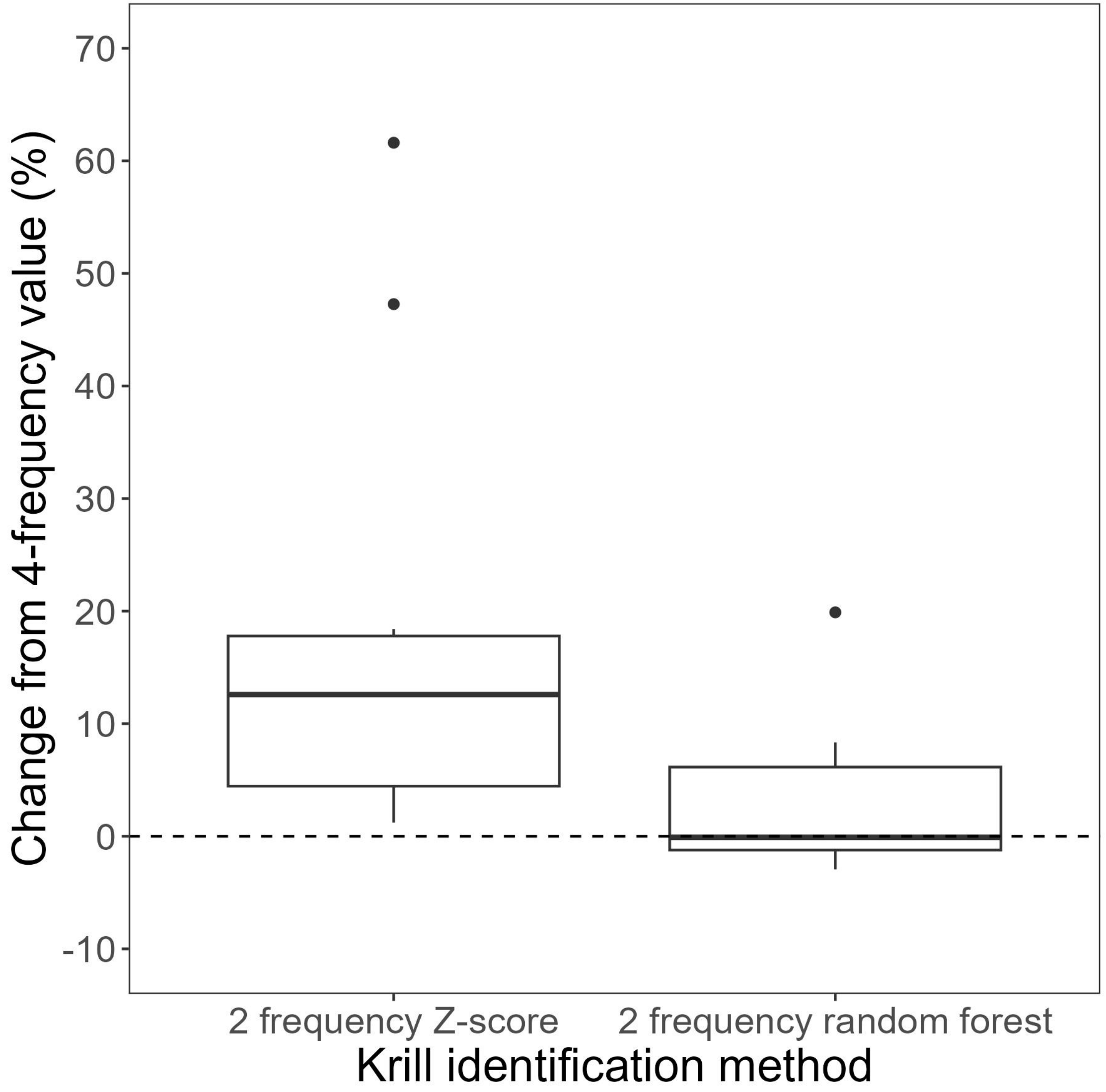
- Model constructed with R packages *caret* and *ranger*
- 100,000 observations in training set
- Cross validation by year (withhold one year from training, and then use this for testing)
  - How well does model generalize across years out of sample?
- Model results:
  - Accuracy (how many predictions did model get right?): 94.2%
  - Kappa (proportion of predictions beyond what would be expected by chance): 82%

Random forest classification

Removing isolated krill ID's:

- Analogous to removing marginal krill in *Z*-score method
- We assumed that krill should be found in spatially extensive layers (not lone cells)
- We required any krill identified by the random forest model to be touching at least two other krill
- This removed 10.3 % of krill cells (+/- 7.4 %), comprising 6.8 % of krill backscatter (+/- 6.8 %)

- As compared to Z-score identification, random forest ID shows lower bias and higher precision



2020 Saildrone estimate

1) Identify krill in survey intervals

Sv_38 kHz

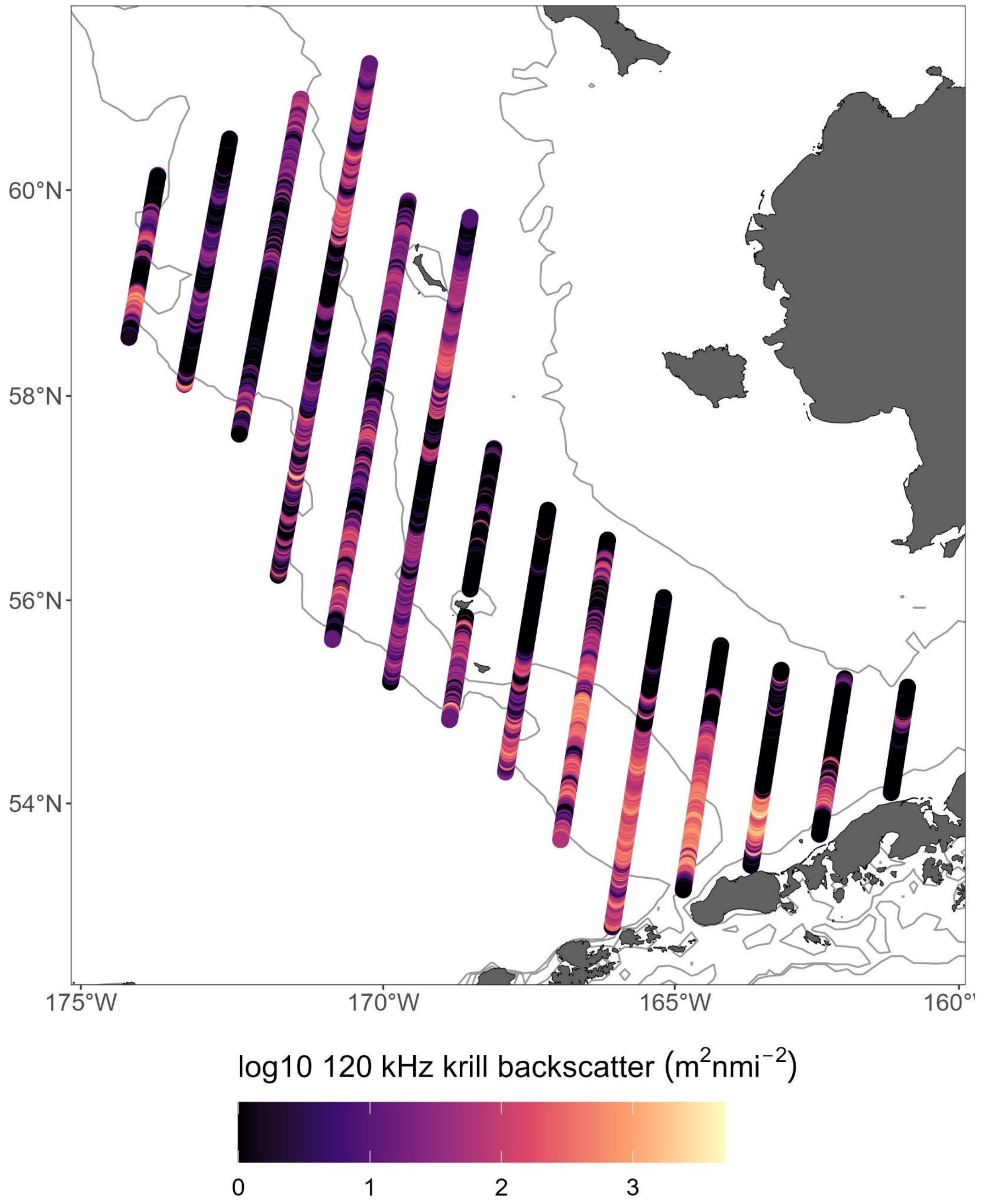
Sv_200 kHz

Random forest classification

Remove isolated krill IDs

Integrate 200 kHz krill backscatter in acoustic-trawl survey intervals

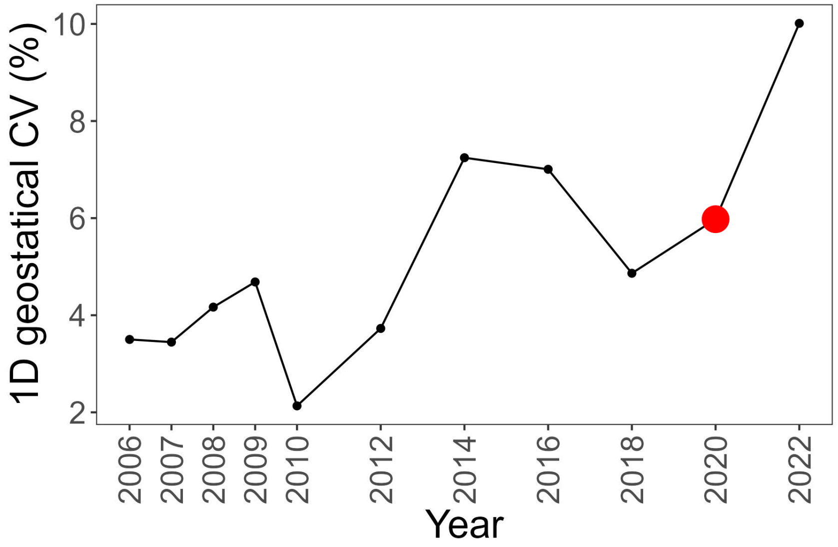
Convert 200 kHz backscatter to 120 kHz backscatter (based on empirical relationships in De Robertis et al 2012)

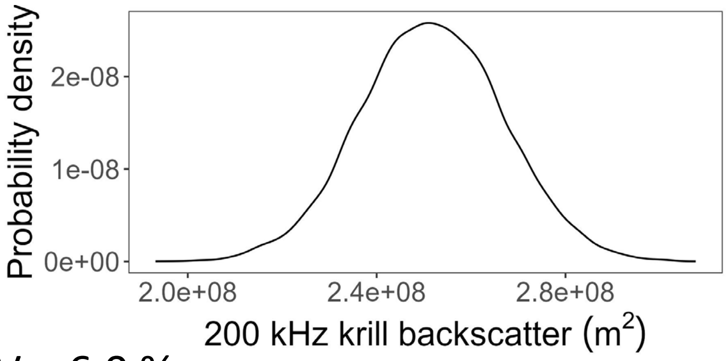


2020 Saildrone estimate

2) Estimate uncertainty

1-d geostatistical uncertainty



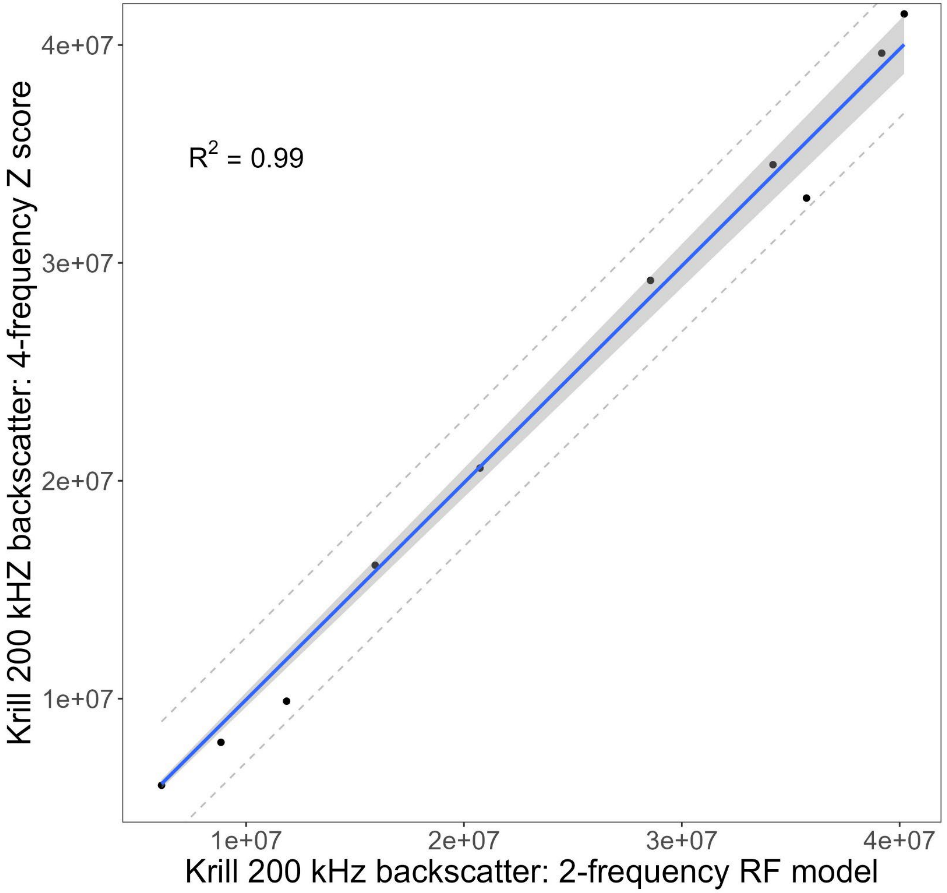


CV = 6.0 %

Approach based on De Robertis et al. 2021

+

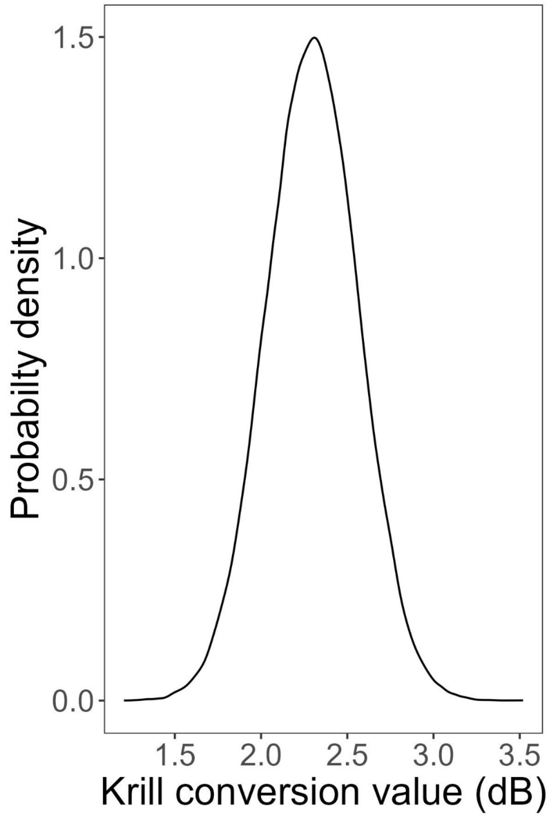
Random forest uncertainty
(via Monte-Carlo simulation)



CV = 13.0 %

+

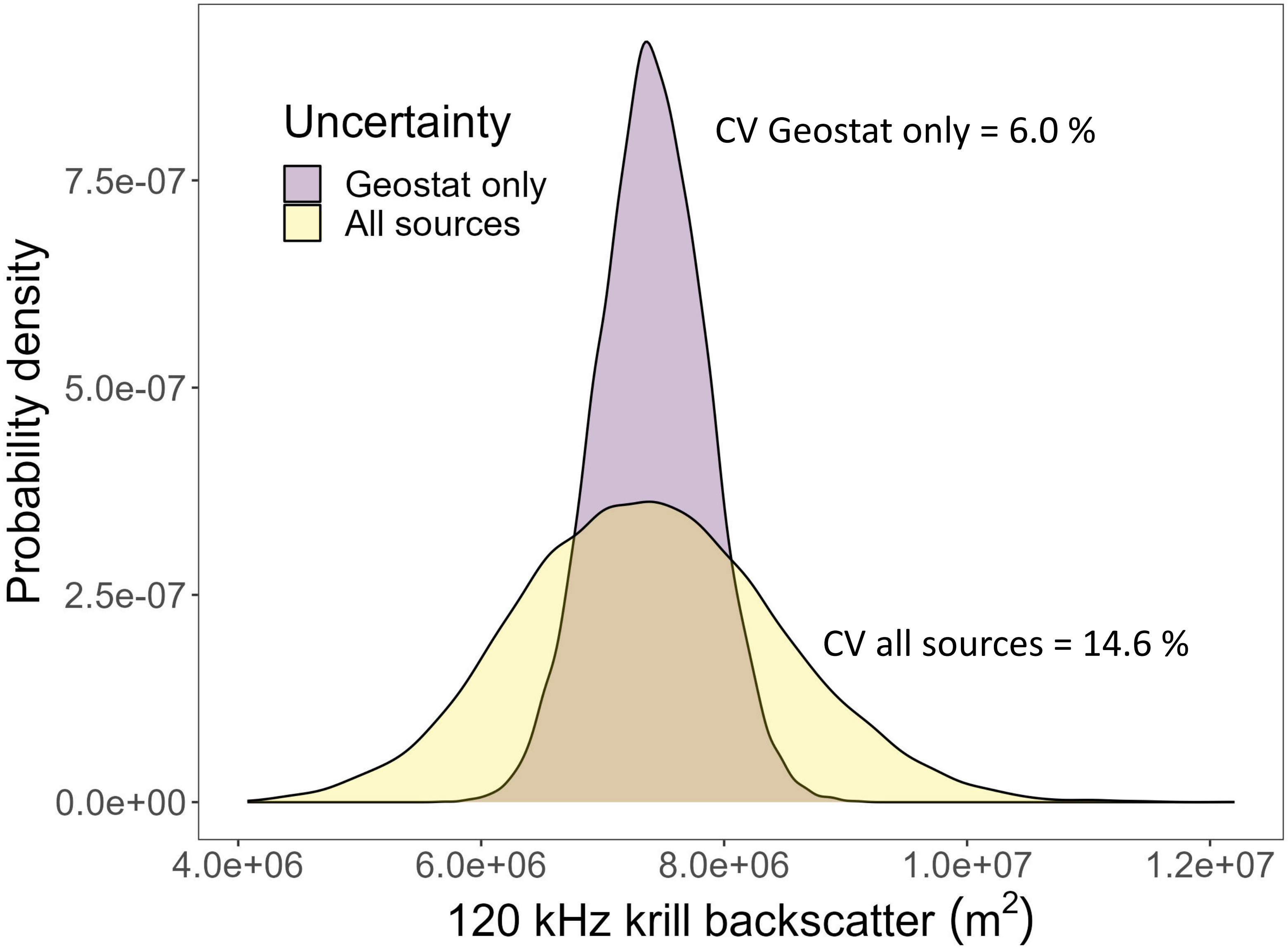
200 kHz to 120 kHz conversion uncertainty
(via Monte-Carlo simulation)



CV = 14.6 %

2020 Saildrone estimate

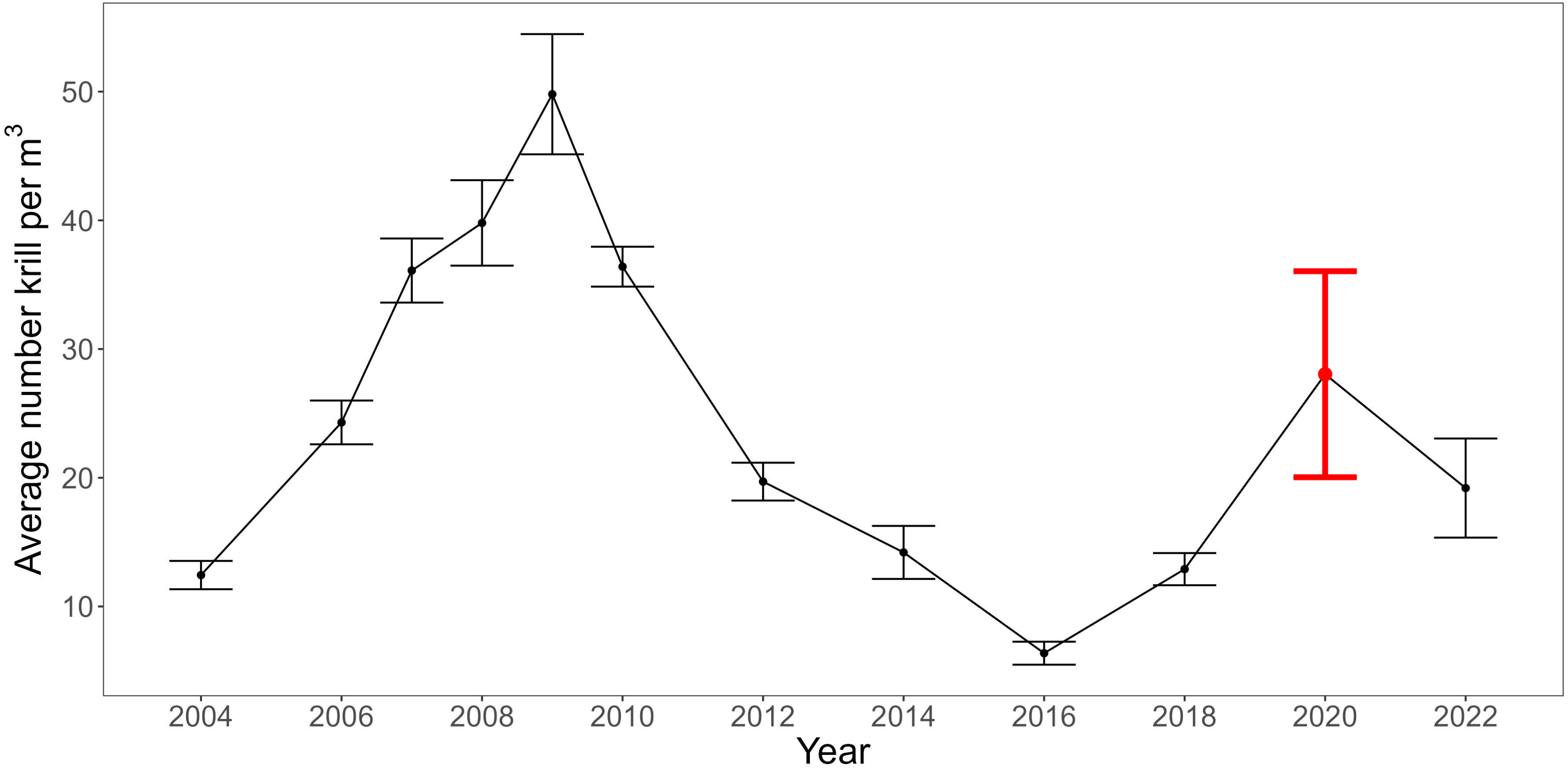
2) Estimate uncertainty



2020 Saildrone estimate

3) Convert backscatter to biomass

Net catches from 2004-2016 used to estimate average length and species composition → $\boxed{\dfrac{backscatter\ per\ m\char`^3}{backscatter\ per\ krill}}$ → krill / m^3



2020 :
28.0 ± 8.0 krill/m^3 (95% CI)

Conclusions

- Random forest model could mimic the proven 4-frequency approach with less information
- This approach was less biased than simply applying the existing method with less information
- It provides a mechanism to deal with noisy or missing acoustic data, at the cost of increased uncertainty
- This method may be useful in other krill estimates where we have fewer frequencies than usual (summer 2021 GOA, for example)
- For 2020- krill abundance was average in timeseries, with higher uncertainty

Questions?

Acknowledgements: Patrick Ressler provided many details about the 4-frequency approach that were crucial to this work.

Citations

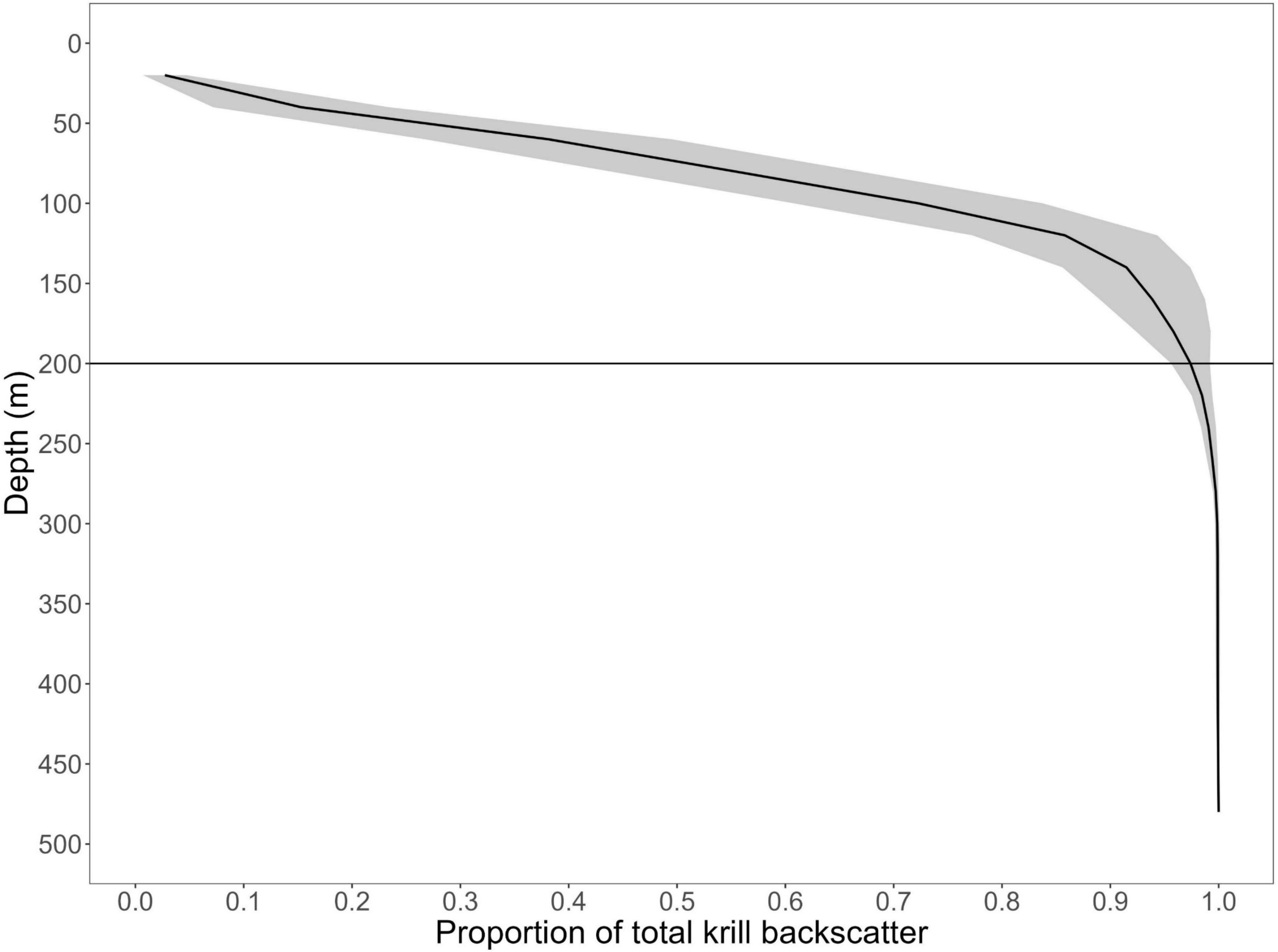
De Robertis A. , D. R. McKelvey and P. H. Ressler. Development and application of an empirical multifrequency method for backscatter classification. Canadian Journal of Fisheries and Aquatic Sciences 2010 Vol. 67 Issue 9 P. 1459-1474.

De Robertis A., M. Levine, N. Lauffenburger, T. Honkalehto, J. Ianelli, C. C. Monnahan, R. Towler, D. Jones, S. Stienessen, D. McKelvey. 2021. Uncrewed surface vehicle (USV) survey of walleye pollock, *Gadus chalcogrammus*, in response to the cancellation of ship-based surveys. ICES Journal of Marine Science 2021 Issue 0 P. 1-12.

Ressler, P.H., A. De Robertis, J. D. Warren, J. N. Smith and S. Kotwicki. 2012. Developing an acoustic survey of euphausiids to understand trophic interactions in the Bering Sea ecosystem. Deep-Sea Research Part II 2012 Vol. 65-70 P. 184-195.

Saildrone images: Richard Jenkins

Logunova, Inna. Random Forest Classifier: Basic Principles and Applications. serokell.io, 06/23/2022, https://serokell.io/blog/random-forest-classification



Depth (m)
0
50
100
150
200
250
300
350
400
450
500
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Proportion of total krill backscatter