**R1 Comment, Line 144:** what about the instances where multiple ships were broadcasting (cumulative exposure)? Did you have a covariate for multiple ships to see if there was any interaction/effect by including whether there were multiple sources?

I think we had a covariate for “ships” that was 1 if 1 ship was present, 2 if 2 or more were present? But that also included other sources (submarines, recovery boats). I guess we could say that this is something we’re looking into as a follow-up study? Liz, thoughts on this?

Yes, we couldn’t report the exact number of ships or the number actively using sonar at any given time, so for both Phase A and Phase B there was a “ships” category that was 0, 1 or 2, with two simply meaning more than one. I do think you included the ships variable and it was significant during Phase B, correct? So that would be a proxy for what the reviewer is asking.

**R1 Comment, Line 190:** can we assume that the detection range (6.5 km) will be the same at PMRF as it is at AUTEC? As you mentioned earlier, the two ranges have different bathymetries and depths. Has it been estimated what the detection range is at PMRF?

Hasn’t been estimated at PMRF, yes, different bathymetries and depths . . . but same species and bottom-mounted HP setup. Could probably say that the results are unlikely to be sensitive to this parameter since it only affects edge hydrophones. If 6.5km is an underestimate, then we would expect detection rates at the edges to be even lower.

Yes, we use the 6.5 km (actually I think we use 6 km) estimated detection radius, although we haven’t measured at PMRF yet. Although the bathymetry is different, in the southern portion of the range (where the bathymetry is most complex) the phones are spaced closer than 6 km so we have high confidence we aren’t missing groups on the range. In this area we usually detect groups on more than one phone. And in the northern part of the range where the phones are spaced a little further apart, the bathymetry is deep and generally less complex, so propagation becomes largely spherical and so we still feel confident we aren’t missing groups, although they are more likely to be detected on a single phone rather than multiple phones.

**R1 Comment, Line 193:** what do you mean “different combinations of hydrophones were used”? Are you saying that the range operators only had certain hydrophones on/recording for one SCC and then a different hydrophone set for another SCC? Please clarify. Does this also mean you generated different probability models pre-SCC for each SCC as opposed to combining them all (so one M1 per SCC)? – I see the answer to my question in section 2.2.2. Maybe introduce on line 193 a little more clearly (ie: “different combinations of hydrophones were used due to...”)?

**Response:** We have modified the text to explain that “Due to recording capacity and occasional instrument failure, not all hydrophones were recorded during all SCCs. Because different combinations of hydrophones were used during different SCCs, we created separate tessellations for each SCC.”

That is basically it. We don’t record ALL the phones at PMRF, the subset of 62 is based on the digital input cards that convert the incoming analog signal from the hydrophones to a digital signal that we record. The two card we used each had 32 slots, so we recorded 62 hydrophones plus two time-code signals. Over the years, different hydrophones would go out (or the entire string in 2017!) so we would switch which phones were recorded. Generally we tried to keep as consistent as possible. But since you created the tessellations relative to each SCC, that was accounted for. So perhaps this could be rephrased as “Due to recording capacity, a subset of PMRF hydrophones were recorded each year. While this subset was held as consistently as possible, because of occasional hydrophone failures a small number of the hydrophones that were recorded changed across years. These slight differences in hydrophone combinations during different SCCs led to the creation of separate tessellations for each SCC.” Something like that – to indicate that the majority of the phones were consistent across SCCs but just a small number changed.

**R1 Comment,** Line 197: please include how long it had been since the last exercise ie “no other naval activity was known to occur for at least 1 week prior” as we saw with McCarthy et al. 2011 and Stanistreet et al. 2022 that it takes time for baseline foraging to resume (sometimes as much as 1 week)

Liz can you confirm how long since last exercise?

Oh boy. I can try and dig up old exercise schedules, but to get that information we have to download the schedules while we are at PMRF and I don’t think this was done regularly in 2013 and 2014. So I don’t think we have any way of confirming that there wasn’t other MFAS (or other kinds of exercises) prior to the SCCs. Maybe we just have to state that, that we can only confirm there wasn’t other activity during our “Before” recordings (because those were unclassified and therefore there wasn’t other activity) and we don’t know otherwise.

R2 Comment: The use of the term “MFAS” (mid frequency active sonar) should be reconsidered. At least it should be defined but my preference would be to replace it with the frequency band meant. Mid frequency means different things between different navies and if you ask a fishery acoustician it means something very different. Please define or replace MFAS.

Liz, can you provide the frequency band that we’re talking about, or a citation that defines it?

For the US Navy MFAS is defined as signals in the 1-10 kHz bandwidth. You can cite one of the EIS documents, e.g. US Navy (2019) Atlantic Fleet Training and Testing (AFTT) Environmental Impact Statement (EIS)/Overseas Environmental Impact Statement (OEIS) Phase III.

R2 Comment: I propose to replace the term “general navy activity” with “navy activity without 1-10 kHz sonar” (or whatever frequency range MFAS means here) to increase readability. Or at least it should be defined the first time it is mentioned.

Will do once the above is sorted.

We should actually specify that Phase A is “Navy activity without surface ship hull-mounted sonar”, since there can be sources with frequencies in the MFAS range during Phase A, just not that one.

R2 Comment: The dB metrics used for received and source level are not clearly defined. I assume it is rms levels. Add proper definition of even better reference the ISO standard (ISO 18405:2017).

I can’t access this. Liz can you confirm dB definitions?

Yes everything I reported was in rms following the ISO standard. I guess the first instance of dB should say “dB re 1 µPa root-mean-square (rms)”.  
  
R2 Comment: In cases were estimated max received level at the hydrophones were lower than 140dB my understanding is that the phones have enough dynamic range to record it. Did you do any comparison between the measured and estimated levels for those cases?

Not as part of this study, but Liz I assume you’ve looked into this to check that prop modelling is reasonable?

I personally haven’t actually done that, but assume in the history of the lab it has been done. Do you want me to find out? That is partially not relevant because we are modeling at 1000 km (or just above the seafloor if it’s shallower than that at a given phone) and most of the phones are much deeper than that, so the level measured on the phone would not match what was modeled. But I can ask Steve if they ever modeled to a phone and then compared the measured levels.