1 **QUANTIFYING THE RESPONSE OF**

2 **BLAINVILLE’S BEAKED WHALES TO US NAVAL**

3 **SONAR EXERCISES IN HAWAII**

4

## 5 Eiren K. Jacobson1, E. Elizabeth Henderson2, David L. Miller1, Cornelia S.

6 **Oedekoven1, David J. Moretti3, Len Thomas1**

7 1 *Centre for Research into Ecological and Environmental Modelling, School of Mathematics and*

8 *Statistics, University of St Andrews, St Andrews, Scotland*

9 2 *Naval Information Warfare Center Pacific, San Diego, CA, USA*

10 3 *Naval Undersea Warfare Center, Newport, RI, USA*

## 11 Correspondence:

12 Eiren Jacobson

13 University of St Andrews

14 The Observatory

15 Buchanan Gardens

16 St Andrews

17 Fife

18 KY16 9LZ

19 Scotland

20 Email: [eiren.jacobson@st-andrews.ac.uk](mailto:eiren.jacobson@st-andrews.ac.uk)

21

## 22 Draft 16 January 2022

23 **Abstract**

24 Behavioral responses of beaked whales (family Ziphiidae) to naval use of mid-frequency

25 active sonar (MFAS) have been quantified for some species and regions. We describe the

26 effects of MFAS on the probability of detecting diving groups of Blainville’s beaked whales

27 on the US Navy Pacific Missile Range Facility (PMRF) in Hawaii and compare our results

28 to previously published results for the same species at the Atlantic Undersea Test and

29 Evaluation Center (AUTEC) in the Bahamas. We use passive acoustic data collected at

30 bottom-mounted hydrophones before and during six naval training exercises at PMRF along

31 with modelled sonar received levels to describe the effect of training and MFAS on foraging

32 groups of Blainville’s beaked whales. We use a multi-stage generalized additive modelling

33 approach to control for the underlying spatial distribution of vocalizations under baseline

34 conditions. At an MFAS received level of 150 dB re 1 *µ*Pa the probability of detecting groups

35 of Blainville’s beaked whales decreases by 77% (95% CI 67%-84%) compared to periods when

36 general training activity was ongoing and by 87% (95% CI 81%-91%) compared to baseline

37 conditions. Our results indicate a more pronounced response to naval training and MFAS

38 than has been previously reported. [196/200]

39 **KEYWORDS**

40 Blainville’s beaked whales, *Mesopolodon densirostris*, mid-frequency active sonar, passive

41 acoustic data, behavioral response, generalized additive model

42 **1 Introduction**

43 Beaked whales (family Ziphiidae) are a group of deep-diving cetaceans that rely on sound to

44 forage, navigate, and communicate (Aguilar de Soto et al., 2012; Johnson et al., 2004; Macleod

45 and D’Amico, 2006) and are sensitive to anthropogenic noise (Southall et al., 2016). Multiple

46 mass strandings of beaked whales have been associated with high-intensity anthropogenic

47 sound sources including naval sonar (Bernaldo de Quirós et al., 2019; D’Amico et al., 2009).

48 These acute events have motivated research into whether and how beaked whales respond

49 to different types and intensities of anthropogenic noise (e.g., Aguilar de Soto et al., 2006;

50 Cholewiak et al., 2017; Tyack et al., 2011). Anthropogenic sound can disrupt the foraging

51 dive cycles of beaked whales (Falcone et al., 2017), potentially leading to cumulative sublethal

52 impacts resulting from reduced foraging opportunities (New et al., 2013; Pirotta et al., 2018),

53 or to symptoms similar to decompression sickness that can lead to injury or death (Hooker

54 et al., 2009, 2012).

55 Echolocation clicks produced by diving groups of Blainville’s beaked whales indicate foraging

56 activity and can be recorded by hydrophones (Johnson et al., 2006). Research on Blainville’s

57 beaked whales *(Mesoplodon densirostris)* using data from bottom-mounted hydrophones on a

58 U.S. Navy range in the Bahamas has shown decreases in time spent foraging and movement

59 away from naval sonar sources (Joyce et al., 2019; Tyack et al., 2011). Naval sonar can

60 be broadcast from various platforms, including vessels, helicopters, buoys, submarines, and

61 torpedoes (U.S. Department of the Navy, 2018). Most research has focused on the impacts of

62 mid-frequency active sonar (MFAS) broadcast from naval vessels. Separately, researchers

63 have shown that, in the absence of MFAS, beaked whales may alter their behavior in response

64 to vessel noise (Aguilar de Soto et al., 2006; Pirotta et al., 2012).

65 The U.S. Navy is interested in quantifying the effects of sonar on beaked whales for the

66 purpose of risk assessments and permitting associated with training activities (e.g., U.S.

67 Department of the Navy, 2017). There are different experimental and analytical ways of

68 quantifying responses to sonar (see Harris et al., 2018 for a review). Here, we focus on

69 analyses of observational data from cabled hydrophone arrays collected concurrently with

70 naval training exercises. Examples of these from previous studies include McCarthy et al.

71 (2011) who used data from the cabled hydrophone array at the U.S. Navy’s Atlantic Undersea

72 Test and Evaluation Center (AUTEC) in the Bahamas collected before, during, and after

73 naval training exercises involving MFAS. The authors used separate generalized additive

74 models (GAMs) for each period, and modelled the acoustic detection of groups of Blainville’s

75 beaked whales (group vocal periods; GVPs) as a function of location on the range and time.

76 They found that the number of GVPs was lower during the exercises than before or after.

77 Building on this work, Moretti et al. (2014) used a GAM to examine the presence or absence

78 of GVP starts within 30-min periods (i.e., whether or not a GVP started within each 30-min

79 period) on the AUTEC range as a smooth function of MFAS received level. They compared

80 the expected probability of detecting animals when no sonar was present to the expected

81 probability of detecting animals across sonar received levels to estimate the probability of

82 disturbance. They found that the probability of detecting groups of Blainville’s beaked whales

83 was reduced by 50% at 150 dB re 1 *µ*Pa, which they interpreted as a 50% probability of

84 disturbance.

85 Our primary objective was to replicate the effort of Moretti et al. (2014) with the same

86 species on a different U.S. Navy training range in a different oceanic environment. We used

87 a spatially-referenced dataset of Blainville’s beaked whale foraging dives recorded at the

88 PMRF off the island of Kauai, Hawaii (Fig. 1). Passive acoustic detections of the presence or

89 absence of GVP starts within 30-min periods were collected via a cabled hydrophone array

90 at PMRF before and during training exercises involving MFAS broadcast from navy ships.

91 Unlike AUTEC, which is situated in a deep isolated basin surrounded by steep slopes, the

92 Pacific Missile Range Facility (PMRF) in Hawaii is located on the side of an ancient volcano,

93 with a steep slope to the deep ocean floor. Previous work in this region has shown that

94 Blainville’s beaked whales are present year-round at this site, prefer sloped habitats, and that

95 acoustic detections decrease during multi-day training events involving MFAS (Henderson

96 et al., 2016; Manzano-Roth et al., 2016). As we expected the density of Blainville’s beaked

97 whales at PMRF to be low and spatially variable, our methods needed to explicitly account

98 for differences in underlying beaked whale presence across the range. An additional objective

99 was to isolate the effect of general training activity from the effect of MFAS, so that beaked

100 whale response to MFAS could be quantified relative to pre-training baseline periods and to

101 periods when general training activities were present on the range.

102 **2 Methods**

# 103 2.1 Data Collection and Processing

## 104 2.1.1 Acoustic detection of beaked whales

105 The Pacific Missile Range Facility (PMRF) is an instrumented U.S. Navy range extending

106 70 km NW of the island of Kauai, Hawaii and encompassing 2,800 km2. The range includes

107 a cabled hydrophone array (Fig. [1)](#_bookmark0) with hydrophones at depths ranging from approximately

108 650 m to 4,700 m. We used data collected before and during six Submarine Command

109 Courses (SCCs) at PMRF. SCCs are training exercises that involve several different naval

110 platforms, occur biannually in February and August, and typically last 6-7 days. MFAS is

111 broadcast from naval vessels during part of the training exercises. Acoustic recordings were

112 made for a minimum of two days before each SCC as well as during the exercise. During data

113 collection, hydrophones sampled at a rate of 96 kHz. Up to 62 hydrophones were recorded

114 simultaneously by the Naval Information Warfare Center (NIWC).

115 A beaked whale echolocation detector from the Navy Acoustic Range WHale AnaLysis

116 (NARWHAL) algorithm suite (Martin et al., 2020) was run on the recordings. This detector

117 first compared signal-to-noise ratio (SNR) thresholds within the expected frequency range

118 of beaked whale clicks (16-44 kHz) versus the bandwidth outside the click in a running

119 16,384-pt fast Fourier transform (FFT) spectrogram. The detected clicks were then passed

120 to a 64-pt FFT stage that measured power, bandwidth, slope, and duration characteristics

121 to classify the clicks to species. This process was followed by an automated routine in

122 MATLAB (*MATLAB*, 2017) to group detections of individual beaked whale echolocation

123 clicks into GVPs (Henderson et al., 2016). If a group of whales was detected by more than

124 one hydrophone, the GVP was assigned to the hydrophone that recorded the most clicks.

125 The data were then aggregated to indicate the presence or absence of the start of a GVP for

126 each hydrophone within each half-hour period. We used half-hour periods to approximate

127 the typical vocal period of Blainville’s beaked whales during deep foraging dives (Tyack et

128 al., 2006).

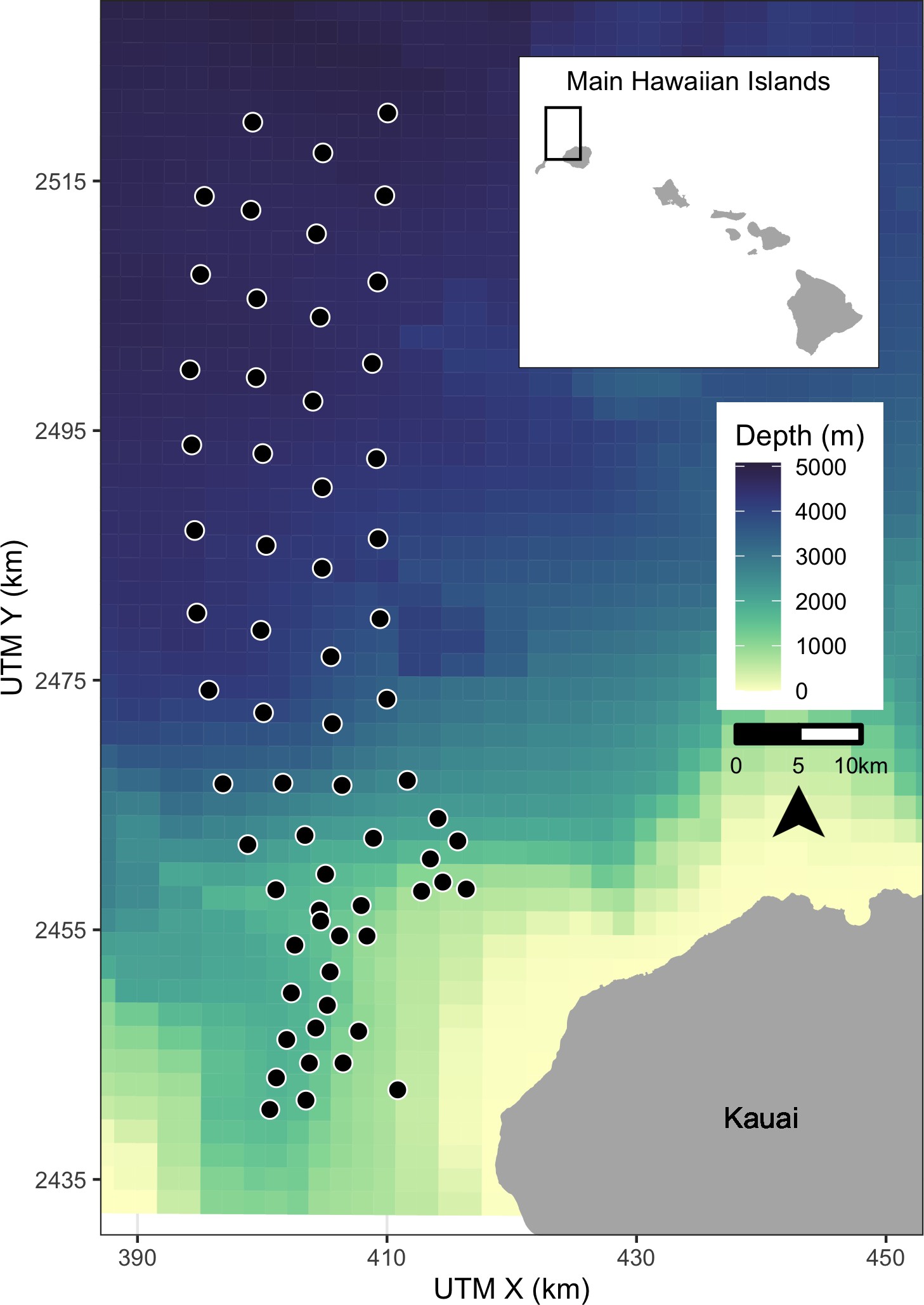


Figure 1: Map of hydrophones (black points) at the Pacific Missile Range Facility near the island of Kauai, Hawaii. For security reasons, the approximate rather than exact locations are shown here. Color scale indicates bathymetry. Inset map shows range location (black rectangle) relative to the main Hawaiian Islands.

## 129 2.1.2 Modelling received levels of hull-mounted mid-frequency active sonar

130 For security reasons, classified data regarding activity that occurred on the range during each

131 SCC was passed from PMRF to one author with clearance (E.E.H.). These data indicated

132 the locations of the ships during the training periods and the start and stop times of each

133 individual training event. However, no information was provided on the start and stop

134 of sonar use; hence, periods of active sonar were determined from the range hydrophone

135 recordings by running a sonar detector from the NARWHAL algorithm suite tuned to MFAS.

136 The hydrophone recordings cannot reliably be used to determine received level when the

137 received level exceeds 140 dB re 1 *µ*Pa due to voltage constraints at the analog to digital

138 recorder interface. Additionally, the hydrophones are mostly 4-5 km deep, whereas Blainville’s

139 beaked whales begin clicking when they have reached depths of approximately 200-500 m and

140 spend most of their foraging dive at depths of 1-1.5 km (Johnson et al., 2004, 2006; Madsen

141 et al., 2013). Therefore, we used an acoustic modeling approach to estimate the maximum

142 received level of hull-mounted MFAS during each half-hour period around the location of

143 each hydrophone at a depth of 1,000 m.

144 First, the locations of all surface ships were noted at the start of each half-hour period and

145 the closest ship to each hydrophone was determined. MFAS propagation was modelled using

146 the parabolic equation propagation model in the program Peregrine (OASIS, Heaney and

147 Campbell, 2016). Acoustic transmission loss was estimated using a 200 Hz band around the

148 center frequency of the sonar (3.5 kHz). A nominal source level of 235 dB re 1 *µ*Pa @ 1 m was

149 assumed (U.S. Department of the Navy, 2018). The transmission loss was estimated along

150 the radial from the ship to the hydrophone from a distance of 1 km before the hydrophone to

151 1 km past the hydrophone in 200 m increments and converted to received levels based on

152 the source level of the sonar. The maximum modeled received level along that radial was

153 determined for each hydrophone and half-hour period. However, if the distance between the

154 ship and the hydrophone was less than the depth of the water column, the parabolic equation

155 would overestimate transmission loss at that angle. In these cases, a simple sonar equation

156 was used to estimate transmission loss instead (Urick, 1983). For hydrophones shallower

157 than 1,000 m the received level was estimated at a point 20 m above the sea floor with a +/-

158 10 m buffer, while for hydrophones deeper than 1,000 m the received level was estimated at a

159 depth of 1,000 m with a +/- 10 m buffer. This process resulted in an estimate of received

160 level for each hydrophone and half-hour period. Uncertainty in the modeled received levels

161 was not considered.

162 **2.2 Spatial Modelling**

163 **Summary**

164 We first used tessellation to determine the area effectively monitored by each hydrophone

165 (section 2.2.1). Then, we used a three-stage GAM approach to control for the underlying

166 spatial distribution of Blainville’s beaked whales when modelling the effects of training

167 activities and of MFAS. For the first model (M1), we used pre-activity data to create a spatial

168 model of the probability of GVPs across the range prior to the onset of naval activity (2.2.2).

169 We used the predicted values from this first model as an offset in a second model (M2) created

170 using data from when naval activity was present on the range, but MFAS was not (2.2.3).

171 We then used the predicted values from this second model as an offset in a third model (M3)

172 created using data when naval activity and MFAS were present on the range (2.2.4). Finally,

173 we used posterior simulation to calculate confidence intervals and quantified the change in

174 the probability of detecting GVPs when naval activity was present and across received levels

175 of MFAS (2.2.5). Analyses were undertaken in R (R Core Team, 2018). Code and data are

176 available at [https://github.com/eirenjacobson/MdMFASResponsePMRF.](https://github.com/eirenjacobson/MdMFASResponsePMRF)

## 177 2.2.1 Determining hydrophone effort

178 For security reasons, randomly jittered locations and depths of hydrophones at PMRF

179 were used. We projected the coordinates of each hydrophone into Universal Transverse

180 Mercator Zone 4. Because the beaked whale detection algorithm assigned GVPs to the

181 hydrophone that recorded the most echolocation clicks, and because the spatial separation of

182 the hydrophones was not uniform, effort was not the same for all hydrophones. This meant

183 that some hydrophones may have detected more GVPs because they were further away from

184 other hydrophones, not because they were located in higher-density areas. To account for

185 this, we used a Voronoi tessellation implemented in the R package deldir (Turner, 2019) to

186 define a tile for each hydrophone that contained all points on the range that were closest to

187 that hydrophone. We assumed that beaked whale groups occur within the tessellation tile

188 of the hydrophone to which the GVP is assigned, and that the area of each tessellation tile

189 influences the GVP detection rate at that hydrophone. For hydrophones on the outside of

190 the range, i.e., not surrounded by other hydrophones, we used a cutoff radius of 6.5 km to

191 bound the tessellation tiles. This distance was based on the estimated maximum detection

192 distance of individual Blainville’s beaked whale clicks at a U.S. Naval range in the Bahamas

193 (Marques et al., 2009). Different combinations of hydrophones were used during different

194 SCCs, so separate tessellations were created for each SCC.

## 195 2.2.2 M1: Modelling the pre-activity probability of dive detection

196 In the first model, we used data collected prior to SCCs, when no naval ships were present on

197 the range and no other naval activity was known to occur, to model the spatial distribution of

198 GVP detections across the range. Because of the way that GVPs were assigned to hydrophones

199 (see Section 2.1.1) the data were not continuous in space. To account for this, we used a

200 Markov random field (MRF) implemented in the R package mgcv (Wood, 2017) to model

201 the spatial distribution of GVP detections. Markov random fields (Rue and Held, 2005)

202 model correlation in space between discrete spatial units (henceforth, “tiles”). The correlation

203 between two tiles is dictated by distance, as measured by the number of other tiles one needs

204 to pass through to travel between two tiles (“hops”); correlation is strongest between a tile

205 and its direct neighbors (those tiles it shares a border with) and decreases with additional

206 hops. This was appropriate for our data as we did not know where in each tile a given GVP

207 occurred, but we assumed that it did occur in that tile.

208 We modelled the probability of a GVP at tile *i* during SCC *s* (*µ*M1*,i,s*) as a Bernoulli random

209 variable. The linear predictor (on the logit scale) was given as:

logit (*µ*M1*,i,s*) = *β*M1*,*0 + *f* (MRF*i,s*) + *f* (Depth*i*) + log*e Ai,s* (1)

210 where *β*M1*,*0 is an intercept, *f* (MRF*i,s*) denotes the Markov random field used to smooth space

211 in the *s*th SCC, *f* (Depth*i*) is a smooth of depth at the location of each hydrophone (using a

212 thin plate spline; Wood (2003)) and log*e Ai,s* is an offset for the area (in km2) of each tile for

213 each SCC, *Ai,s*. The offset term accounts for changes in probabilities of GVP detection due

214 to the different areas monitored by each hydrophone. Because the hydrophone tessellation

215 changed between SCCs (as there were different sets of hydrophones recorded during each SSC),

216 separate MRFs were used for each SCC, but a single smoothing parameter was estimated

217 across all MRFs. This allowed for different spatial smooths for each SCC, but constrained

218 the smooths to have the same amount of wiggliness. The smooth of depth was shared across

219 SCCs. We used this model to predict the baseline probability of a GVP detection at each

220 hydrophone.

## 221 2.2.3 M2: Modelling the effect of Naval activity

222 For the second model, we used data collected prior to the onset of hull-mounted MFAS used

223 during SCCs, when other naval training activities occurred at PMRF. Various vessels were

224 present on the range during these periods, and other noise sources, including range tracking

225 pingers, torpedoes, and submarines, may have been present. We used data collected when

226 training activity was present on the range, but hull-mounted MFAS was not used, to model

227 the effect of general naval activity on beaked whale GVPs.

228 We used the predicted baseline probability of a GVP detection at each hydrophone from M1

229 as an offset to control for the underlying spatial distribution of GVPs. The model for the

230 data when naval activity was present was intercept-only, with an offset derived from M1. This

231 meant that the spatial distribution of GVPs was not allowed to change, but that we expected

232 a uniform relative change in GVPs when naval activity was present. We again modelled

233 the probability of GVP presence at tile *i* (*µ*M2*,i*) as a Bernoulli random variable, with the

234 following linear predictor:

logit (*µ*M2*,i,s*) = *β*M2*,*0 + log*e ξ*M1*,i,s,* (2)

235 where *β*M2*,*0 is an intercept and *ξ*M1*,i,s* is the prediction (on the logit scale) for tile *i* during

236 SCC *s* using model M1, included as an offset term.

## 237 2.2.4 M3: Modelling the effect of hull-mounted MFAS

238 For the third model, we used data collected when hull-mounted MFAS was present on the

239 range to model the effect of sonar on beaked whales. We excluded data collected during

240 breaks in training activities when sonar was not being used. The probability of a GVP

241 when sonar was present was modeled as a function of the maximum received level (modeled

242 at each hydrophone for each half-hour period; see section 2.2.1). We assumed that as the

243 maximum received level increased, the probability of dives decreased and modeled this using

244 a monotonically decreasing smooth implemented in the R package scam (Pya and Wood,

245 2015). To ensure that the model predictions were the same at a maximum received level

246 of 0 dB and when only naval activity was present, we did not include an intercept. The

247 probability of GVP presence at tile *i* (*µ*M3*,i*) was modelled as a Bernoulli random variable

248 where the linear predictor was:

logit (*µ*M3*,i,s*) = *f* (MaxRL*i,s*) + log*e ξ*M2*,i,s,* (3)

249 where *f* (MaxRL*i,s*) was modeled as a monotonic decreasing smooth, *ξ*M2*,i,s* denotes the prediction

250 (on the logit scale) for tile *i* during SCC *s* when naval training activities were present on the

251 range using model M2.

252 **2.2.5 Uncertainty propagation**

253 We used posterior simulation (sometimes referred to as a parametric bootstrap, Wood et al.,

254 2017) to propagate uncertainty through M1, M2, and M3. This consisted of sampling from

255 the posterior distribution of the parameters for each model in turn, calculating predictions

256 using these parameters and then refitting the subsequent model with updated offsets. We

257 generated 5,000 sets of posterior samples. Following this procedure through from M1 to M2

258 to M3 incorporated uncertainty from each model in the final predictions of the probability of

259 detecting a GVP given different combinations of covariates.

260 The prediction grid contained all possible combinations of covariates within the realized

261 covariate space; i.e., each hydrophone for each SCC with associated location, hydrophone

262 depth, and area of the tessellation tile, presence/absence of naval activity, and, if naval

263 activity was present, then either sonar absence or sonar received level. Based on the resulting

264 final posterior distribution of results (for model M3) we used 2.5%, 50%, and 97.5% quantiles

265 to obtain median predictions and credible intervals (CIs). Details of the procedure are given

266 in Appendix S1.

## 267 2.2.6 Quantifying the change in probability of GVPs

268 Finally, we calculated the expected change in the probability of detecting a GVP at each

269 hydrophone *P* (GVP) relative to either the probability of detecting a GVP when no general

270 naval training activity was present and no MFAS was present (∆M3:M1), or relative to probability

271 of detecting a GVP when general naval training activity was present but no MFAS was

272 present (∆M3:M2).

273 Using the *Nb* posterior samples, we calculated the expected *P* (GVP) under each set of

274

covariates as

*P* (GVP) = logit−1(*µ*M)*,* (1)

for each M = M1, M2, and M3. Then, we calculated the change in *P* (GVP) for each set of covariates between M3 and M1 (∆M3:M1) and between M3 and M2 (∆M3:M2) for each realization of the posterior simulation.

∆*M* 3:*M* 1

∆*M* 3:*M* 2

= *P* (GVP)M3 − *P* (GVP)M1

*P* (GVP)M1

= *P* (GVP)M3 − *P* (GVP)M2

*P* (GVP)M2

(2)

(3)

275 For each received level we calcualted the 2.5th, 50th, and 97.5th quantiles of ∆M3:M1 and

276 ∆M3:M2 to create 95% CIs of change in *P* (GVP) across possible received levels.

277 **3 Results**

# 278 3.1 Data Collection and Processing

279 Data were collected before and during six SCCs: two each in 2013, 2014, and 2017 (Table [1).](#_bookmark1)

280 The number of hydrophones for which recordings were available for each SCC varied from 49

281 to 61. A total of 190,561 30-min observations were made.

282 The exact timing of activities during these exercises varied (Fig. [2).](#_bookmark2) For most SCCs, pre-

283 activity data were available immediately preceding the onset of Naval training activity;

284 however, in February 2013 the only available pre-activity data were collected almost a month

285 prior to the onset of Naval training activity. In some SCCs, weekends or other breaks in

286 training resulted in a break in training activity on the range during the days preceding MFAS

287 use. MFAS was used for 3-4 days during each training event.

288 Across all SCCs, hydrophones, and conditions, a total of 2,458 GVPs were identified. The

Table 1: Number of hydrophones (HPs) used and number of observations made (no. 30-min periods) during each Submarine Commander Course (SCC) before the exercise began, when naval activity was present, and when naval activity and mid-frequency active (MFA) sonar were present.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| SCC | HPs | Pre-Exercise | Phase A | Phase B |
| Feb13 | 61 | 113 | 183 | 134 |
| Aug13 | 61 | 204 | 113 | 99 |
| Feb14 | 60 | 514 | 102 | 138 |
| Aug14 | 61 | 262 | 115 | 133 |
| Feb17 | 59 | 450 | 96 | 109 |
| Aug17 | 49 | 270 | 106 | 113 |

289 average probability of detecting a GVP during each half-hour period was therefore 1.3%. The

290 spatial distribution of GVPs differed during the pre-activity phases of SCCs (Fig. [S2.1;](#_bookmark0) top

291 panel).

292 Modelled maximum received levels ranged from 38 to 186 dB re 1 *µ*Pa, with a median value

293 when MFAS was present of 147 dB re 1 *µ*Pa. The intensity and spatial distribution of MFAS

294 received levels varied across the range and across SCCs (Fig. [S2.2).](#_bookmark2)

295 Based on the observed data, the probability of detecting a GVP changed by -38% when

296 general naval training activity was present compared to when naval activity was absent, by

297 -61% when naval activity and MFAS were present compared to when only naval activity was

298 present, and by -76% when naval activity and MFAS were present compared to when neither

299 naval activity nor sonar were present (Fig. [S2.3).](#_bookmark3) The highest modelled received level at

300 which a GVP was observed was 164 dB re 1 *µ*Pa.

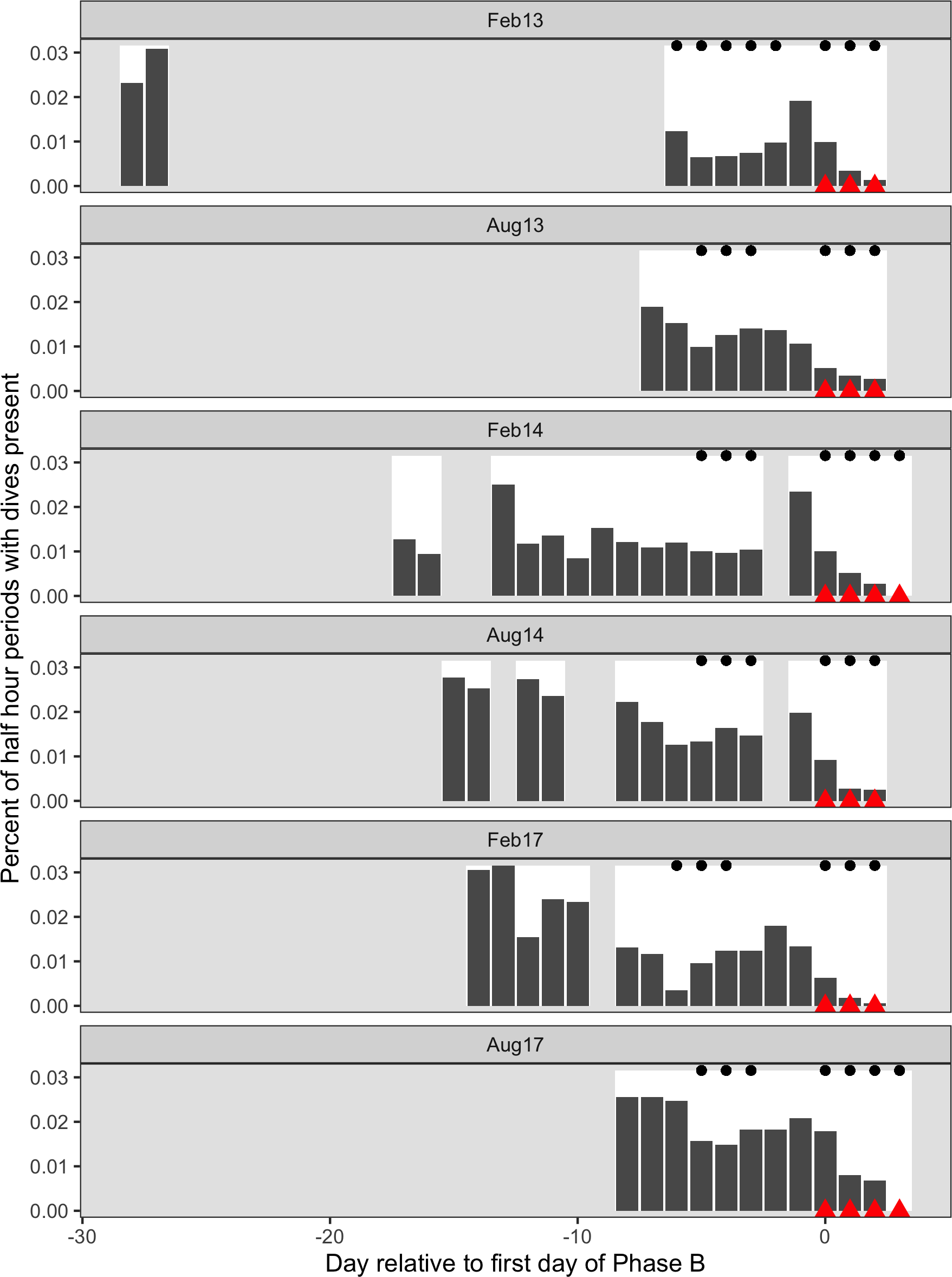


Figure 2: Time series of six recorded Naval training activities at the Pacific Missile Range Facility. The time series are aligned relative to the first day that mid-frequency active sonar (MFAS; red triangles) was used in each exercise (horizontal axis). Days with white background indicate days for which recordings and data were available. Dark gray bars indicate the proportion of 30-min periods on each day, across all hydrophones, when group vocal periods (GVPs) were detected (vertical axis). Black dots indicate days when naval training activity was present on the range.

301 **3.2 Spatial Modelling**

302 We created separate tessellations for each SCC (Fig. [S2.4).](#_bookmark4) In August 2017, data were

303 available from fewer hydrophones, and so in some cases the tessellated tiles, with bounding

304 radius of 6,500 m, did not completely cover the range. Hydrophone depths varied from

305 approximately 650 to 4720 m.

306 M1 fitted a spatial model of *P* (GVP) to data collected prior to the onset of naval training

307 activity. This model used a MRF smooth to account for the spatial structure of the range

308 and a spline on depth, with an offset for the log of the area effectively monitored by each

309 hydrophone. Both the MRF and spline on depth were significant at the *α* = 0.05 level (*p*-value

310 < 2E-16), indicating that GVPs varied space. The model explained 13.5% of deviance in the

311 dataset, and visual inspection of observed versus predicted values indicated a good fit to the

312 data (Fig. [S2.5).](#_bookmark5) The model M1 predicted highest *P* (GVP) at hydrophone depths between

313 1,500 and 2,000 m (Fig. S[2.6).](#_bookmark5)

314 M2 used the predicted values from M1 as an offset and fitted a model of to data when naval

315 activity was ongoing, as indicated by the presence of naval ships on the range. This model

316 was intercept-only, and *P* (GVP) when naval activity was ongoing was significantly different

317 from the baseline period at the *α* = 0.05 level (*p*-value < 2E-16). The expected *P* (GVP)

318 decreased by a median of 44% (95% CI 38% - 49%) when naval activity was present compared

319 to when it was absent.

320 M3 used the predicted values from M2 as an offset and fitted a model to data when naval

321 activity and MFAS were present. This model used a monotonically decreasing spline on

322 modelled MFAS received level (Fig. S[2.7)](#_bookmark5) and did not include an intercept term. The smooth

323 on MFAS received level was significant at the *α* = 0.05 level (*p*-value = 2E-10) and the

324 model explained 20% of deviance in the data.

325 We did not make inference on sonar received levels below 100 dB re. 1 *µ*Pa because Blainville’s

326 beaked whales are unlikely to perceive MFAS below received levels of approximately 80

327 dB re. 1 *µ*Pa (Pacini et al., 2011) and because very little data (9 hrs, or 1% of the data

328 collected when MFAS was present) were collected at received levels below 100 dB re. 1 *µ*Pa.

329 Similarly, we did not make inference on sonar received levels above 165 dB re. 1 *µ*Pa because

330 no GVPs were observed above this received level and therefore M3 predicted *P* (GVP) = 0

331 (95% CI 0-1).

332 For MFAS received levels between 100 and 165 dB re. 1 *µ*Pa, change in *P* (GVP) was calculated

333 relative to the pre-activity baseline period (∆M3:M1; Fig. 4 left panel) and to the period when

334 naval activity was present on the range (∆M3:M2; Fig. 4 right panel) using the posterior

335 samples. For illustration purposes, ∆M3:M1 and ∆M3:M2 calculated using five individual posterior

336 samples are shown in Fig. [S2.8.](#_bookmark6) At a received level of 150 dB, the posterior median of ∆M3:M1

337 was -87% (95% CI -91% - -81%) and the posterior median of ∆M3:M2 was -77% (95% CI -84%

338 - -67%). Relative to when only naval training is present, ∆M3:M2 predicts a 50% reduction in

339 *P* (GVP) at a MFAS received level of 132 dB re 1 *µ*Pa.

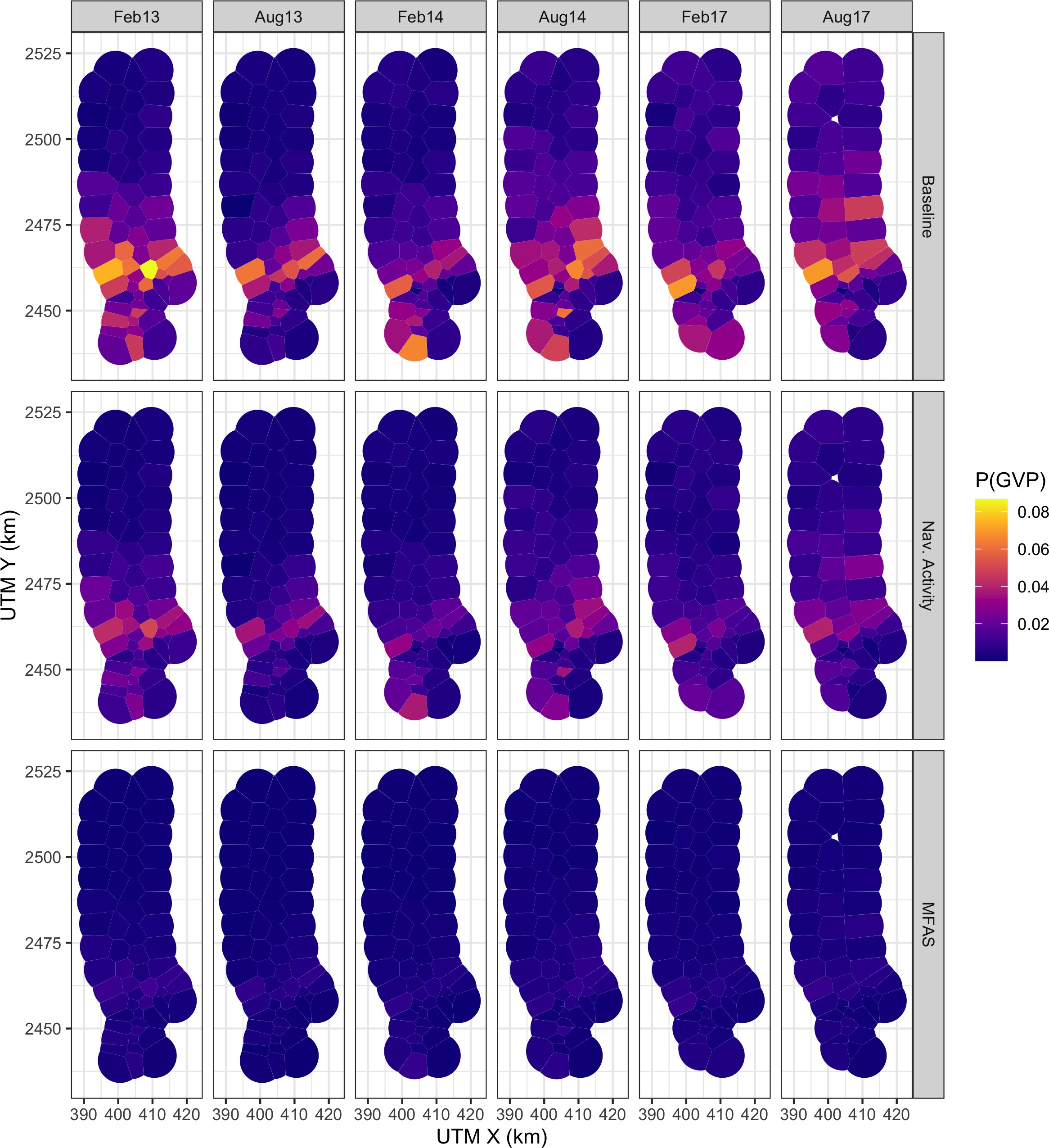


Figure 3: Map of expected probability of detecting a GVP (color scale) at each hydrophone during each SCC (columns) prior to the onset of naval training activity, during naval training activity when no MFAS was present, and during naval training activity when MFAS was present at a received level of 150 dB re 1 *µ*Pa rms (rows).

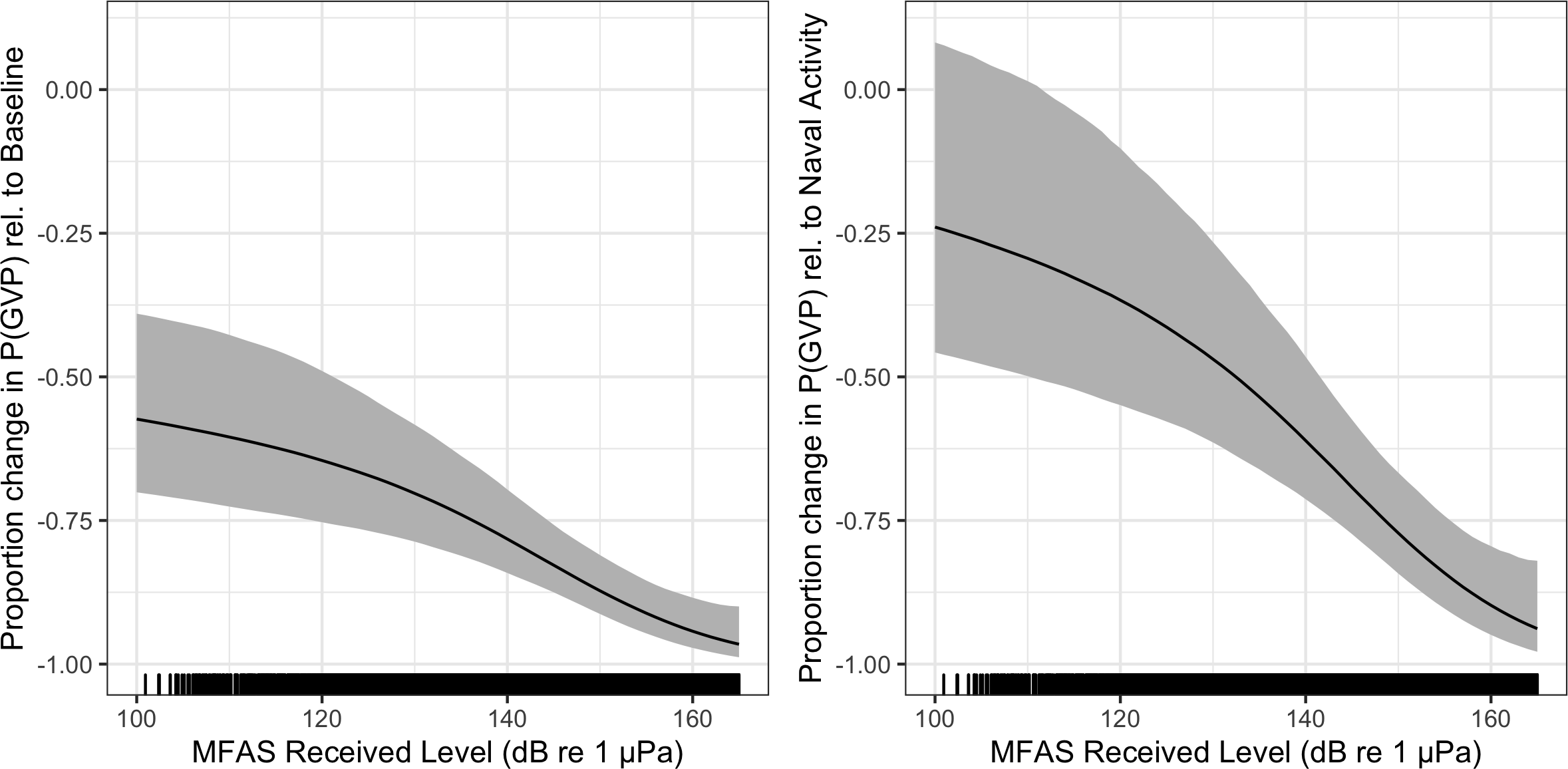


Figure 4: Median (black line) and 95% CI (gray shading) expected change in the probability of detecting a group vocal period (vertical axis) with increasing MFAS received level (horizontal axis) relative to when naval training activity but no MFAS was present on the range (left panel) and to when neither naval training activity nor MFAS were present on the range.

340 **4 Discussion**

341 We used a series of three linked models to quantify the response of Blainville’s beaked whales

342 to naval training exercises involving MFAS: the first model was fitted to pre-exercise baseline

343 data, the second was fitted to data collected when naval training exercises were ongoing but

344 no MFAS was present, and the third model was fitted to data collected during naval training

345 exercises that used MFAS. We found that the probability of acoustic detections of Blainville’s

346 beaked whales decreased when both naval training exercises and naval training exercises

347 using MFAS were present (Fig. [4).](#_bookmark4)

348 The methods presented here are spatially explicit and account for the spatial confounding

349 of animal distribution and naval training activity. The data used in this study are from an

350 undesigned experiment, where the spatial intensity of the treatments (naval activity and

351 MFAS) were not applied randomly with respect to either the study area or Blainville’s beaked

352 whale foraging activity. We did not want the spatial distribution of training exercises and

353 MFAS to influence our understanding of the baseline spatial distribution of Blainville’s beaked

354 whales. Due to the spatial confounding of animal distribution and naval training activity at

355 PMRF, fitting a single model to all of the data leads to greater uncertainty in estimating the

356 impact of sonar, since changes in distribution due to MFAS could be explained as variability

357 in spatial distribution by the MRF (Appendix S3). Our three-stage modelling approach

358 addresses this issue while propagating uncertainty between the models. To our knowledge,

359 this is a novel application of GAMs.

360 The analytical approach outlined in this article could be applied to other species, regions, and

361 types of disturbance where experimental design is not possible. The use of Markov random

362 fields for the spatial term is useful for cases where exact location data are not available,

363 avoiding the inappropriate use of continuous-space smoothers. Shape-constrained smoothing

364 (in our case, monotonically decreasing smooth) is also well-suited to the kind of data we

365 modelled here, ensuring that values can only stay constant or decrease over time (or any

366 other covariate). Finally, the use of a multi-stage posterior sampling scheme for quantifying

367 uncertainty extends to other situations where multiple models are fitted and the results of

368 one part feed into another. Simulation-based approaches such as these bypass the need to

369 derive (often complex) mathematical expressions for variance (or shortcut them by assuming

370 independence).

371 The expected change in the probability of a GVP when MFAS was present included CIs

372 that reflect several sources of uncertainty (Fig. 4). The small number of GVPs when

373 MFAS was present—and therefore sparse coverage of data points across the range of received

374 levels—makes it difficult to estimate the effect of MFAS received level precisely. GVPs were

375 detected in only 1.7% of half-hour periods in the baseline dataset, in 1% of periods (n = 448)

376 when naval activity was present, and 0.2% (n = 50) when MFAS was present. Additional

377 data—particularly at relatively low and relatively high source levels, where uncertainty is

378 greatest—may reduce uncertainty in the expected probability of GVPs across different source

379 levels. It is also possible that contextual factors that we did not include in this analysis,

380 such as distance to sound source (DeRuiter et al., 2013; Falcone et al., 2017), may provide

381 additional explanatory power and reduce uncertainty. Finally, the observed uncertainty may

382 reflect true individual variation in response due to variables like age and sex (see Harris et

383 al., 2018, sec. 2.2 for a review of relevant publications).

384 The model M3, which modelled the effect of received level on *P* (GVP), was constrained to be

385 monotonically decreasing with no intercept term, so that the predicted *P* (GVP) would be

386 the same or lesser when MFAS was present compared to when only naval training activity

387 was present. However, it is possible that *P* (GVP) could be higher at relatively low MFAS

388 received levels than when only naval training is present, since animals may move away from

389 high-intensity areas, resulting in increased densities in lower-intensity areas. In our dataset,

390 some hydrophones had lesser observed *P* (GVP) at low levels of MFAS and some had greater

391 (Fig. [S2.3).](#_bookmark3) Due to small sample size at low intensities, we cannot determine whether

392 observed increases in *P* (GVP) when MFAS was present at relatively low intensities was due

393 to sampling error or to avoidance of high-intensity areas. The version of the model fitted

394 as a single GAM (Appendix S3) predicted the change in *P* (GVP) to be > 0, i.e., increased

395 relative to when only naval training activity was present, at MFAS received levels below 103

396 dB re 1 *µ*Pa (Fig. [S3.1).](#_bookmark0)

397 We excluded data collected between training activity within an SCC (13% of the available

398 data) as we did not consider it to be true baseline data since naval activity and/or MFAS had

399 recently (within hours or days) been present. It would be interesting to explore the complete

400 dataset, including these interim periods, to investigate the timescales on which beaked whales

401 respond to naval activity (e.g., Jones-Todd et al., 2021; Joyce et al., 2019). We might expect

402 that time since training activity or MFAS could lead to recovery of *P* (GVP) towards baseline

403 levels, perhaps modulated by the cumulative exposure to training and MFAS.

404 In a regulatory context, a dose-response function as presented in Fig. [4](#_bookmark4) is often interpreted as

405 representing the proportion of a population that responds (vertical axis) to a given received

406 level (horizontal axis) (Tyack and Thomas, 2019). However, the metric used in this study—the

407 change in the probability of detecting a GVP within a 30-min period—may not directly

408 correspond to the proportion of the population that is affected. It may instead reflect a

409 change in the proportion of time that all individuals in the population spent foraging in the

410 study area. These two interpretations have different implications for understanding sub-lethal

411 impacts of MFAS. In the former interpretation, given exposure to a certain received level,

412 some of the population is affected and some of the population is not. In the latter, the entire

413 exposed population is affected. With our data, we cannot distinguish between these possible

414 scenarios.

415 In comparison to the risk function developed by Moretti et al. (2014) for Blainville’s beaked

416 whales at AUTEC, our risk function for PMRF predicts a more intense response to naval

417 sonar. This may be because Moretti et al. were not able to account explicitly for the effects

418 of naval training activities that did not include MFAS. Their baseline period consisted of

419 19 hours of data before the onset of MFAS; as at PMRF, it is likely that training activities

420 during this period included sound sources other than MFAS. Therefore, their risk function is

421 likely more analogous to our expected change in the probability of a detection when MFAS is

422 present relative to when naval training activity was present (Fig. [4).](#_bookmark4) Future research will

423 investigate the specific causes of changes in the probability of detecting GVPs before the

424 onset of MFAS. The reduction in detection of foraging dives could be a response to general

425 naval training activity on the range, or to specific sound sources that have not previously

426 been studied. Alternatively, it is possible that Blainville’s beaked whales are semi-resident

427 on the range and have become habituated to SCC activity; they may move off the range in

428 anticipation of MFAS. Resident animals that are frequently exposed to training activity and

429 transient animals that only encounter MFAS occasionally are likely to respond differently

430 to sonar. It is not known how resident the Blainville’s beaked whales are at PMRF, and

431 offshore animals may be detected on the northern hydrophones.

432 Blainville’s beaked whales occur in multiple ocean basins and have been studied on U.S.

433 Navy training ranges in both the Atlantic (AUTEC) and the Pacific (PMRF) Oceans. The

434 AUTEC range is located in a deep basin bounded to the south, east, and west by shallow

435 waters and with maximum depths of 2,000 m. In contrast, the PMRF occurs across a steep

436 slope and into deep water, over 5,000 m in depth. Although the environments at PMRF

437 and AUTEC are different, the foraging dive behavior of Blainville’s beaked whales is similar:

438 dives occur in waters over steep slopes with gradients ranging from 3-23%, although dives

439 occur in deeper waters (2,000-3,000 m, Henderson et al., 2016) at PMRF that at AUTEC

440 (Hazen et al., 2011; 500-1,300 m, MacLeod and Zuur, 2005). Resident Blainville’s beaked

441 whales off the island of Hawaii also occur in slightly shallower waters than at PMRF, from

442 980-1,410 m (Baird, 2011; Baird et al., 2008). It seems likely the location of the mesopelagic

443 scattering layer (indicating the presence of prey) along the slope that drives the location of

444 Blainville’s beaked whales rather than the bathymetric depth; this is supported by the fact

445 that dive depths are similar across areas, occurring on average down to 1,050-1,150 m for

446 46-60 min (Baird et al., 2008; Joyce et al., 2017; Schorr et al., 2009). Documented responses

447 to MFAS activity are comparable at both ranges, with individuals and groups moving to the

448 periphery of the range or off the range and returning 2-4 days after the cessation of the sonar

449 (Joyce et al., 2019; Manzano-Roth et al., 2016; McCarthy et al., 2011).

450 The similarities in Blainville’s beaked whale behavioral responses to navy training activity

451 across different ranges and environments at similar received levels may indicate the intrinsic

452 nature of the response. The findings presented here and in Moretti et al. (2014) may be

453 applicable to other species and regions, though species-specific dive behaviors and regional

454 differences in oceanography likely modulate the impact of MFAS. For example, existing

455 findings already demonstrate that Cuvier’s respond in a similar manner by reducing their

456 foraging dives and moving away from sonar sources (DeRuiter et al., 2013; Falcone et al.,

457 2017). Conducting a similar analysis of Cuvier’s beaked whale responses at the Southern

458 California Anti-Submarine Warfare Range (SOAR) would further support our understanding

459 of how different populations and species respond to naval sonar.

460 **Acknowledgements**

461 This study was funded by the US Navy Living Marine Resources Program (Contract

462 No. N39430-17-P-1983).

463 **Authors’ Contributions**

464 Conceptualization: E.E.H., D.J.M, L.T.

465 Data curation: E.K.J., E.E.H.

466 Formal analysis: E.K.J., E.E.H., C.S.O.

467 Funding acquisition: E.E.H., D.J.M., L.T., E.K.J.

468 Investigation: E.E.H.

469 Methodology: E.K.J., E.E.H., D.L.M., C.S.O., L.T.

470 Software: E.K.J., D.L.M., C.S.O.

471 Supervision: L.T.

472 Visualization: E.K.J., C.S.O.

473 Writing – original draft: E.K.J., E.E.H., D.L.M.

474 Writing – review & editing: E.K.J, E.E.H., D.L.M., C.S.O., D.J.M., L.T.

475 **ORCID**

476 Eiren K. Jacobson: <https://orcid.org/0000-0003-0147-8367>

477 E. Elizabeth Henderson: <https://orcid.org/0000-0002-3212-1080>

478 David L. Miller: <https://orcid.org/0000-0002-9640-6755>

479 Cornelia S. Oedekoven: <https://orcid.org/0000-0002-5610-7814>

480 Len Thomas: <https://orcid.org/0000-0002-7436-067X>

481 **References**

482 Aguilar de Soto, N., Johnson, M., Madsen, P. T., Tyack, P. L., Bocconcelli, A., &

483 Fabrizio Borsani, J. (2006). Does intense ship noise disrupt foraging in deep-diving

484 Cuvier’s beaked whales (*Ziphius cavirostris*)? *Marine Mammal Science*, *22* (3), 690–699.

485 <https://doi.org/10.1111/j.1748-7692.2006.00044.x>

486 Aguilar de Soto, N., Madsen, P. T., Tyack, P., Arranz, P., Marrero, J., Fais, A., Revelli, E.,

487 & Johnson, M. (2012). No shallow talk: Cryptic strategy in the vocal communication

488 of Blainville’s beaked whales. *Marine Mammal Science*, *28* (2), E75–E92. [https:](https://doi.org/10.1111/j.1748-7692.2011.00495.x)

489 [//doi.org/10.1111/j.1748-7692.2011.00495.x](https://doi.org/10.1111/j.1748-7692.2011.00495.x)

490 Baird, R. W. (2011). Short note: Open-ocean movements of a satellite-tagged Blainville’s

491 beaked whale (*Mesoplodon densirostris*): Evidence for an offshore population in Hawai‘i?

492 *Aquatic Mammals*, *37* (4), 506–511. <https://doi.org/10.1578/AM.37.4.2011.506>

493 Baird, R. W., Webster, D. L., Schorr, G. S., McSweeney, D. J., & Barlow, J. (2008). Diel

494 variation in beaked whale diving behavior. *Marine Mammal Science*, *24* (3), 630–642.

495 <https://doi.org/10.1111/j.1748-7692.2008.00211.x>

496 Bernaldo de Quirós, Y., Fernandez, A., Baird, R. W., Brownell, R. L., Aguilar de Soto, N.,

497 Allen, D., Arbelo, M., Arregui, M., Costidis, A., Fahlman, A., Frantzis, A., Gulland,

498 F. M. D., Iñíguez, M., Johnson, M., Komnenou, A., Koopman, H., Pabst, D. A., Roe,

499 W. D., Sierra, E., . . . Schorr, G. (2019). Advances in research on the impacts of

500 anti-submarine sonar on beaked whales. *Proceedings of the Royal Society B: Biological*

501 *Sciences*, *286* (1895), 20182533. <https://doi.org/10.1098/rspb.2018.2533>

502 Cholewiak, D., DeAngelis, A. I., Palka, D., Corkeron, P. J., & Van Parijs, S. M. (2017).

503 Beaked whales demonstrate a marked acoustic response to the use of shipboard

504 echosounders. *Royal Society Open Science*, *4* (12), 170940. [https://doi.org/10.1098/](https://doi.org/10.1098/rsos.170940)

505 [rsos.170940](https://doi.org/10.1098/rsos.170940)

506 D’Amico, A., Gisiner, R., Ketten, D., Hammock, J., Johnson, C., Tyack, P., & Mead, J.

507 (2009). Beaked whale strandings and naval exercises. *Aquatic Mammals*, *35*, 452–472.

508 <https://doi.org/10.1578/AM.35.4.2009.452>

509 DeRuiter, S. L., Southall, B. L., Calambokidis, J., Zimmer, W. M. X., Sadykova, D.,

510 Falcone, E. A., Friedlaender, A. S., Joseph, J. E., Moretti, D., Schorr, G. S., Thomas,

511 L., & Tyack, P. L. (2013). First direct measurements of behavioural responses by

512 Cuvier’s beaked whales to mid-frequency active sonar. *Biology Letters*, *9* (4), 20130223.

513 <https://doi.org/10.1098/rsbl.2013.0223>

514 Falcone, E. A., Schorr, G. S., Watwood, S. L., DeRuiter, S. L., Zerbini, A. N., Andrews,

515 R. D., Morrissey, R. P., & Moretti, D. J. (2017). Diving behaviour of Cuvier’s beaked

516 whales exposed to two types of military sonar. *Royal Society Open Science*, *4* (8),

517 170629. <https://doi.org/10.1098/rsos.170629>

518 Harris, C. M., Thomas, L., Falcone, E. A., Hildebrand, J., Houser, D., Kvadsheim, P. H.,

519 Lam, F.-P. A., Miller, P. J. O., Moretti, D. J., Read, A. J., Slabbekoorn, H., Southall,

520 B. L., Tyack, P. L., Wartzok, D., & Janik, V. M. (2018). Marine mammals and sonar:

521 Dose-response studies, the risk-disturbance hypothesis and the role of exposure context.

522 *Journal of Applied Ecology*, *55* (1), 396–404. <https://doi.org/10.1111/1365-2664.12955>

523 Hazen, E. L., Nowacek, D. P., St. Laurent, L., Halpin, P. N., & Moretti, D. J. (2011). The

524 relationship among oceanography, prey fields, and beaked whale foraging habitat in

525 the Tongue of the Ocean. *PLoS ONE*, *6* (4), e19269. [https://doi.org/10.1371/journal.](https://doi.org/10.1371/journal.pone.0019269)

526 [pone.0019269](https://doi.org/10.1371/journal.pone.0019269)

527 Heaney, K. D., & Campbell, R. L. (2016). Three-dimensional parabolic equation modeling

528 of mesoscale eddy deflection. *The Journal of the Acoustical Society of America*, *139* (2),

529 918–926. <https://doi.org/10.1121/1.4942112>

530 Henderson, E. E., Martin, S. W., Manzano-Roth, R., & Matsuyama, B. M. (2016). Oc-

531 currence and habitat use of foraging Blainville’s beaked whales (*Mesoplodon den-*

532 *sirostris*) on a U.S. Navy range in Hawaii. *Aquatic Mammals*, *42* (4), 549–562.

533 <https://doi.org/10.1578/AM.42.4.2016.549>

534 Hooker, S. K., Baird, R. W., & Fahlman, A. (2009). Could beaked whales get the bends?

535 Effect of diving behaviour and physiology on modelled gas exchange for three species:

536 Ziphius cavirostris, Mesoplodon densirostris and Hyperoodon ampullatus. *Respiratory*

537 *Physiology & Neurobiology*, *167* (3), 235–246. [https://doi.org/10.1016/j.resp.2009.04.](https://doi.org/10.1016/j.resp.2009.04.023)

538 [023](https://doi.org/10.1016/j.resp.2009.04.023)

539 Hooker, S. K., Fahlman, A., Moore, M. J., Soto, N. A. de, Quirós, Y. B. de, Brubakk, A.

540 O., Costa, D. P., Costidis, A. M., Dennison, S., Falke, K. J., Fernandez, A., Ferrigno,

541 M., Fitz-Clarke, J. R., Garner, M. M., Houser, D. S., Jepson, P. D., Ketten, D. R.,

542 Kvadsheim, P. H., Madsen, P. T., . . . Tyack, P. L. (2012). Deadly diving? Physiological

543 and behavioural management of decompression stress in diving mammals. *Proceedings.*

544 *Biological Sciences*, *279* (1731), 1041–1050. <https://doi.org/10.1098/rspb.2011.2088>

545 Johnson, M., Madsen, P. T., Zimmer, W. M. X., Aguilar de Soto, N., & Tyack, P. L.

546 (2004). Beaked whales echolocate on prey. *Proceedings of the Royal Society of London.*

547 *Series B: Biological Sciences*, *271*, S383–S386. <https://doi.org/10.1098/rsbl.2004.0208>

548 Johnson, M., Madsen, P. T., Zimmer, W. M. X., Soto, N. A. de, & Tyack, P. L. (2006).

549 Foraging Blainville’s beaked whales (*Mesoplodon densirostris*) produce distinct click

550 types matched to different phases of echolocation. *Journal of Experimental Biology*,

551 *209* (24), 5038–5050. <https://doi.org/10.1242/jeb.02596>

552 Jones-Todd, C. M., Pirotta, E., Durban, J. W., Claridge, D. E., Baird, R. W., Falcone, E.

553 A., Schorr, G. S., Watwood, S., & Thomas, L. (2021). Discrete-space continuous-time

554 models of marine mammal exposure to Navy sonar. *Ecological Applications*, *0* (0),

555 e02475. <https://doi.org/10.1002/eap.2475>

556 Joyce, T. W., Durban, J. W., Claridge, D. E., Dunn, C. A., Fearnbach, H., Parsons, K. M.,

557 Andrews, R. D., & Ballance, L. T. (2017). Physiological, morphological, and ecological

558 tradeoffs influence vertical habitat use of deep-diving toothed-whales in the Bahamas.

559 *PLOS ONE*, *12* (10), e0185113. <https://doi.org/10.1371/journal.pone.0185113>

560 Joyce, T. W., Durban, J. W., Claridge, D. E., Dunn, C. A., Hickmott, L. S., Fearnbach, H.,

561 Dolan, K., & Moretti, D. (2019). Behavioral responses of satellite tracked Blainville’s

562 beaked whales (*Mesoplodon densirostris*) to mid-frequency active sonar. *Marine*

563 *Mammal Science*, 1–18. <https://doi.org/10.1111/mms.12624>

564 Macleod, C. D., & D’Amico, A. (2006). A review of beaked whale behaviour and ecology

565 in relation to assessing and mitigating impacts of anthropogenic noise. *Journal of*

566 *Cetacean Research and Management*, *7* (3), 211–221.

567 MacLeod, C. D., & Zuur, A. F. (2005). Habitat utilization by Blainville’s beaked whales

568 off Great Abaco, northern Bahamas, in relation to seabed topography. *Marine Biology*,

569 *147* (1), 1–11. <https://doi.org/10.1007/s00227-004-1546-9>

570 Madsen, P. T., Aguilar de Soto, N., Arranz, P., & Johnson, M. (2013). Echolocation

571 in Blainville’s beaked whales (*Mesoplodon densirostris*). *Journal of Comparative*

572 *Physiology A*, *199* (6), 451–469. <https://doi.org/10.1007/s00359-013-0824-8>

573 Manzano-Roth, R., Henderson, E. E., Martin, S. W., Martin, C., & Matsuyama, B.

574 (2016). Impacts of U.S. Navy training events on Blainville’s beaked whale (*Mesoplodon*

575 *densirostris*) foraging dives in Hawaiian waters. *Aquatic Mammals*, *42* (4), 507–518.

576 <https://doi.org/10.1578/AM.42.4.2016.507>

577 Marques, T. A., Thomas, L., Ward, J., DiMarzio, N., & Tyack, P. L. (2009). Estimating

578 cetacean population density using fixed passive acoustic sensors: An example with

579 Blainville’s beaked whales. *The Journal of the Acoustical Society of America*, *125* (4),

580 1982–1994. <https://doi.org/10.1121/1.3089590>

581 Martin, C. R., Henderson, E. E., Martin, S. W., Helble, T. A., Manzano-Roth, R. A.,

582 Matsuyama, B. M., & Alongi, G. A. (2020). *FY18 annual report on Pacific missile*

583 *range facility marine mammal monitoring*. Retrieved from Naval Information Warfare

584 Center Pacific San Diego United States website: [https://apps.dtic.mil/sti/citations/](https://apps.dtic.mil/sti/citations/AD1091141)

585 [AD1091141](https://apps.dtic.mil/sti/citations/AD1091141)

586 *MATLAB*. (2017). Natick, Massachusetts: The MathWorks Inc.

587 McCarthy, E., Moretti, D., Thomas, L., DiMarzio, N., Morrissey, R., Jarvis, S., Ward,

588 J., Izzi, A., & Dilley, A. (2011). Changes in spatial and temporal distribution and

589 vocal behavior of Blainville’s beaked whales (*Mesoplodon densirostris*) during multiship

590 exercises with mid-frequency sonar. *Marine Mammal Science*, *27* (3), E206–E226.

591 <https://doi.org/10.1111/j.1748-7692.2010.00457.x>

592 Moretti, D., Thomas, L., Marques, T., Harwood, J., Dilley, A., Neales, B., Shaffer, J.,

593 McCarthy, E., New, L., Jarvis, S., & Morrissey, R. (2014). A risk function for behavioral

594 disruption of Blainville’s beaked whales (*Mesoplodon densirostris*) from Mid-Frequency

595 Active Sonar. *PLoS ONE*, *9* (1), e85064. <https://doi.org/10.1371/journal.pone.0085064>

596 New, L. F., Moretti, D. J., Hooker, S. K., Costa, D. P., & Simmons, S. E. (2013). Using

597 energetic models to investigate the survival and reproduction of beaked whales (family

598 Ziphiidae). *PLoS ONE*, *8* (7), e68725. <https://doi.org/10.1371/journal.pone.0068725>

599 Pacini, A. F., Nachtigall, P. E., Quintos, C. T., Schofield, T. D., Look, D. A., Levine, G. A., 600 & Turner, J. P. (2011). Audiogram of a stranded Blainville’s beaked whale (*Mesoplodon* 601 *densirostris*) measured using auditory evoked potentials. *Journal of Experimental* 602 *Biology*, *214* (14), 2409–2415. <https://doi.org/10.1242/jeb.054338>

603 Pirotta, E., Booth, C. G., Costa, D. P., Fleishman, E., Kraus, S. D., Lusseau, D.,

604 Moretti, D., New, L. F., Schick, R. S., Schwarz, L. K., Simmons, S. E., Thomas, L.,

605 Tyack, P. L., Weise, M. J., Wells, R. S., & Harwood, J. (2018). Understanding the

606 population consequences of disturbance. *Ecology and Evolution*, *8* (19), 9934–9946.

607 <https://doi.org/10.1002/ece3.4458>

608 Pirotta, E., Milor, R., Quick, N., Moretti, D., Di Marzio, N., Tyack, P., Boyd, I., & 609 Hastie, G. (2012). Vessel noise affects beaked whale behavior: Results of a dedicated 610 acoustic response study. *PLoS ONE*, *7* (8), e42535. [https://doi.org/10.1371/journal.](https://doi.org/10.1371/journal.pone.0042535) 611 [pone.0042535](https://doi.org/10.1371/journal.pone.0042535)

612 Pya, N., & Wood, S. N. (2015). Shape constrained additive models. *Statistics and*

613 *Computing*, *25* (3), 543–559. <https://doi.org/10.1007/s11222-013-9448-7>

614 R Core Team. (2018). *R: A Language and Environment for Statistical Computing*.

615 Retrieved from <https://www.R-project.org/>

616 Rue, H., & Held, L. (2005). *Gaussian Markov Random fields: Theory and Applications*.

617 London: Chapman & Hall.

618 Schorr, G. S., Baird, R. W., Hanson, M. B., Webster, D. L., McSweeney, D. J., & Andrews, 619 R. D. (2009). Movements of satellite-tagged Blainville’s beaked whales off the island of 620 Hawai‘i. *Endangered Species Research*, *10*, 203–213. <https://doi.org/10.3354/esr00229>

621 Southall, B., Nowacek, D., Miller, P., & Tyack, P. (2016). Experimental field studies to 622 measure behavioral responses of cetaceans to sonar. *Endangered Species Research*, *31*, 623 293–315. <https://doi.org/10.3354/esr00764>

624 Turner, R. (2019). *Deldir: Delaunay triangulation and Dirichlet (Voronoi) tessellation*.

625 Retrieved from [https://CRAN.R-project.org/package=deldir](https://CRAN.R-project.org/package%3Ddeldir)

626 Tyack, P. L., Johnson, M., Soto, N. A., Sturlese, A., & Madsen, P. T. (2006). Extreme

627 diving of beaked whales. *Journal of Experimental Biology*, *209* (21), 4238–4253. [https:](https://doi.org/10.1242/jeb.02505)

628 [//doi.org/10.1242/jeb.02505](https://doi.org/10.1242/jeb.02505)

629 Tyack, P. L., & Thomas, L. (2019). Using dose–response functions to improve calculations 630 of the impact of anthropogenic noise. *Aquatic Conservation: Marine and Freshwater* 631 *Ecosystems*, *29* (S1), 242–253. <https://doi.org/10.1002/aqc.3149>

632 Tyack, P. L., Zimmer, W. M. X., Moretti, D., Southall, B. L., Claridge, D. E., Durban, J. 633 W., Clark, C. W., D’Amico, A., DiMarzio, N., Jarvis, S., McCarthy, E., Morrissey, R., 634 Ward, J., & Boyd, I. L. (2011). Beaked whales respond to simulated and actual navy 635 sonar. *PLoS ONE*, *6* (3), e17009. <https://doi.org/10.1371/journal.pone.0017009>

636 U.S. Department of the Navy. (2017). *Criteria and thresholds for U.S. Navy acoustic and* 637 *explosive effects analysis (phase III)*. Retrieved from [https://www.goaeis.com/portals/](https://www.goaeis.com/portals/goaeis/files/eis/draft_seis_2020/supporting_technical/Criteria_and_Thresholds_for_U.S._Navy_Acoustic_and_Explosive_Effects_Analysis_June2017.pdf) 638 [goaeis/files/eis/draft\_seis\_2020/supporting\_technical/Criteria\_and\_Thresholds\_](https://www.goaeis.com/portals/goaeis/files/eis/draft_seis_2020/supporting_technical/Criteria_and_Thresholds_for_U.S._Navy_Acoustic_and_Explosive_Effects_Analysis_June2017.pdf)

639 [for\_U.S.\_Navy\_Acoustic\_and\_Explosive\_Effects\_Analysis\_June2017.pdf](https://www.goaeis.com/portals/goaeis/files/eis/draft_seis_2020/supporting_technical/Criteria_and_Thresholds_for_U.S._Navy_Acoustic_and_Explosive_Effects_Analysis_June2017.pdf)

640 U.S. Department of the Navy. (2018). *Final environmental impact statement/overseas* 641 *environmental impact ttatement Hawaii-Southern California training and testing*. Re- 642 trieved from [https://www.hstteis.com/portals/hstteis/files/hstteis\_p3/feis/section/](https://www.hstteis.com/portals/hstteis/files/hstteis_p3/feis/section/HSTT_FEIS_3.07_Marine_Mammals_October_2018.pdf) 643 [HSTT\_FEIS\_3.07\_Marine\_Mammals\_October\_2018.pdf](https://www.hstteis.com/portals/hstteis/files/hstteis_p3/feis/section/HSTT_FEIS_3.07_Marine_Mammals_October_2018.pdf)

644 Urick, R. J. (1983). *Principles of Underwater Sound* (Third Edition, Reprint 2013). New

645 York: McGraw-Hill, Inc.

646 Wood, S. N. (2003). Thin plate regression splines. *Journal of the Royal Statistical Society:*

647 *Series B (Statistical Methodology)*, *65* (1), 95–114. [https://doi.org/10.1111/1467-9868.](https://doi.org/10.1111/1467-9868.00374)

648 [00374](https://doi.org/10.1111/1467-9868.00374)

649 Wood, S. N. (2017). *Generalized Additive Models: An Introduction with R* (2nd ed.).

650 Chapman; Hall/CRC.

651 Wood, S. N., Li, Z., Shaddick, G., & Augustin, N. H. (2017). Generalized additive 652 models for gigadata: Modeling the U.K. Black smoke network daily data. *Journal of* 653 *the American Statistical Association*, *112* (519), 1199–1210. [https://doi.org/10.1080/](https://doi.org/10.1080/01621459.2016.1195744)

654 [01621459.2016.1195744](https://doi.org/10.1080/01621459.2016.1195744)

655 **S1: Uncertainty estimation details**

We used posterior simulation to propagate uncertainty through M1, M2, and M3. Each model was fitted via restricted maximum likelihood (REML), so the resulting estimates were empirical Bayes estimates. In this case we generated 5,000 samples from the (approximately

multivariate normal) posterior of the model parameters. We generated a sample of the

model parameters, ***β***∗ ∼ MVN(***β***ˆ*,* **V*β***ˆ), where ***β***ˆ is the estimate of the model coefficients

and **V*β***ˆ is the posterior covariance matrix. Here the ***β*** for each model included the

coefficients for the smooth terms in the model and fixed effects (e.g., intercept) if present. We then used the matrix that maps the model parameters to the predictions on the linear predictor scale (**X***p*; Wood et al. 2017; section 7.2.6), along with the inverse link function, to generate predictions for each posterior sample. Denoting the vector of predictions ***µ***∗, we calculate as follows:

***µ***∗ = *g*−1(***η***∗) = *g*−1(**X***p****β***∗ + ***ξ***)*,*

656 where *g* was the link function, ***η***∗ was the linear predictor and ***ξ*** was any offset used by 657 this prediction. Variance estimates can be obtained by taking the empirical variance of 658 the resulting predictions (Wood et al. 2017; section 7.2.6). The prediction grid contained 659 all possible combinations of covariates within the realized covariate space; i.e., each 660 hydrophone for each SCC with associated location, hydrophone depth, and area of the 661 tessellation tile, presence/absence of naval activity, and, if naval activity was present, then 662 either sonar absence or sonar received level between 35 and 190 dB in intervals of 5 dB. 663 This procedure was repeated for each model, with refitting to updated offsets from the 664 previous model.

665 An algorithm for calculating the variance from our multi-stage approach is as follows. First

666 define *Nb* as the number of samples to take (*Nb*=5,000 here), let **X***p,*M*j* for *j* = 1*,* 2*,* 3 be

667 the matrix that maps coefficients to the predictions for model M*j*. For *Nb* times:

Draw a sample from the posterior of M1:

|  |  |
| --- | --- |
| 668 | 1. |
| 669 | 2. |
| 670 | 3. |
| 671 | 4. |
| 672 | 5. |
| 673 |  |
| 674 | 6. |
| 675 | 7. |

***β***˜M1 ∼ MVN(***β***ˆM1*,* **V**M1).

Calculate a new offset for M2, ***ξ***˜M1 = **X***p,*M1***β***˜M1 + log*e* **A**. Refit M2 with ***ξ***˜M1 as the offset, to obtain M2t.

˜ ˆ

Draw a sample from the posterior of M2t:

***β***M2 ∼ MVN(***β***M2 *,* **V**M2 )

Calculate a new offset for M3,

***ξ***˜M2 = **X***p,*M2***β***˜t

+ ***ξ***˜M1 (predictions for the sonar data

locations for M2t, when no sonar was present). Refit M3 with offset ***ξ***˜M2 to obtain M3t.

M2

Predict ***µ***M1 , ***µ***M2 , and ***µ***M3 over prediction grid and store them.

676 We then calculated summary statistics (means and variances) of the *Nb* values of ***µ***M1 , ***µ***M2 , 677 and ***µ***M3 we generated. The empirical variance of the *Nb* values of ***µ***M3 gave the uncertainty, 678 incorporating components from all three models. We took appropriate pointwise quantiles 679 (e.g., 2.5th and 97.5th for a 95% interval) to form confidence bands for the functional 680 relationships between sonar received level and estimated probability of detecting GVPs.

681 **S2: Supplementary Tables and Figures**

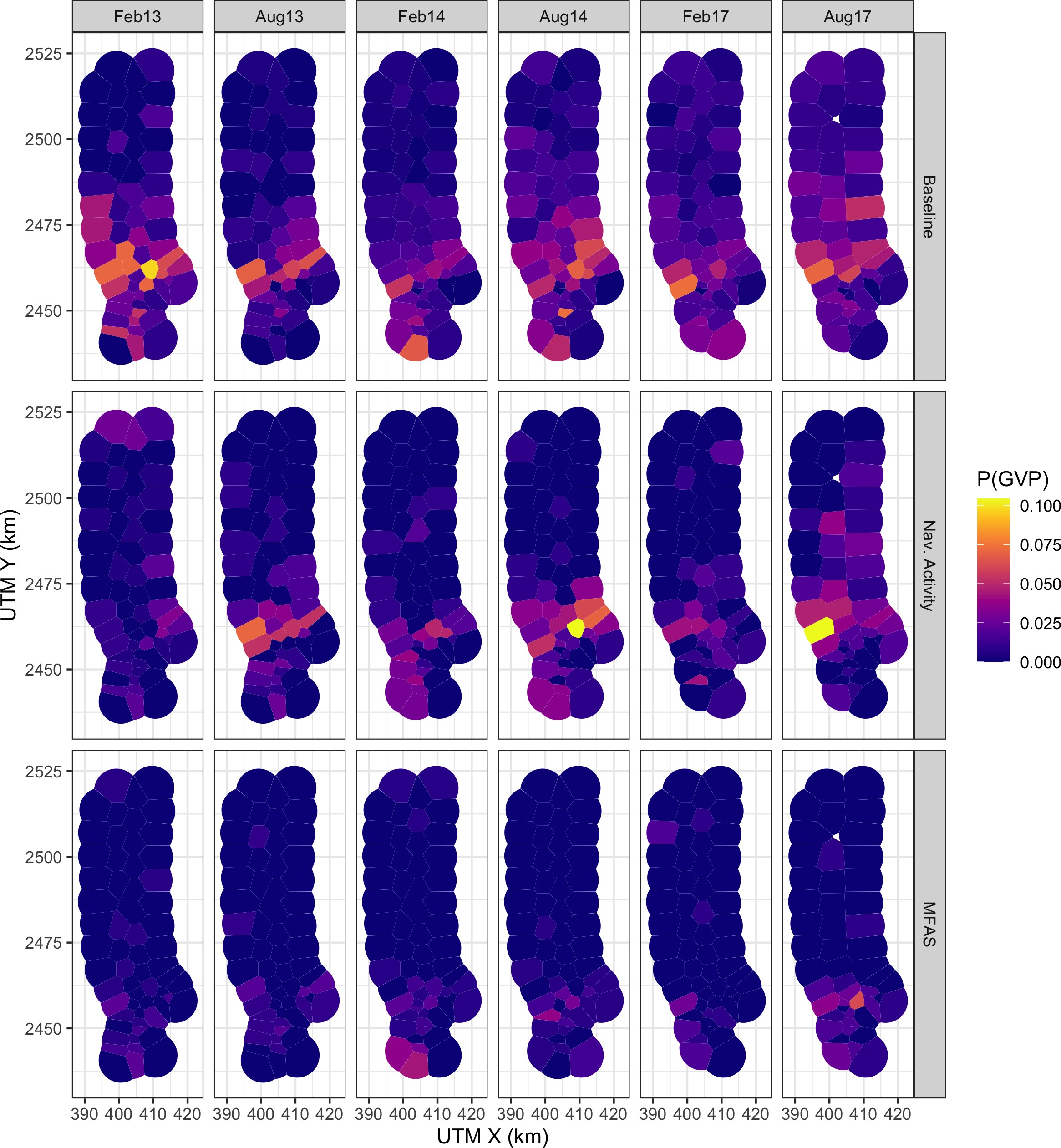


Figure S2.1: Map of observed probability of detecting a GVP at each hydrophone (color scale) during the baseline period, when naval activity was present, and when MFAS was present (rows) for each SCC (columns). Note that values of the probability of detecting a GVP are not corrected for effort (size of the 3h8ydrophone tile).

2525

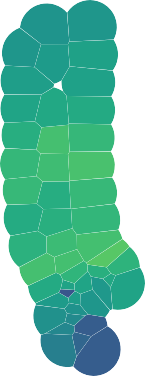
Feb13

Aug13

Feb14

Aug14

Feb17



Aug17

2500

UTM Y (km)

2475

MedRL

160



150

140

130

2450

390 400 410 420 390 400 410 420 390 400 410 420 390 400 410 420 390 400 410 420 390 400 410 420

UTM X (km)

Figure S2.2: Median received level (dB re. 1 *µ*Pa) when MFAS was present (color scale) for all hydrophones and SCCs.

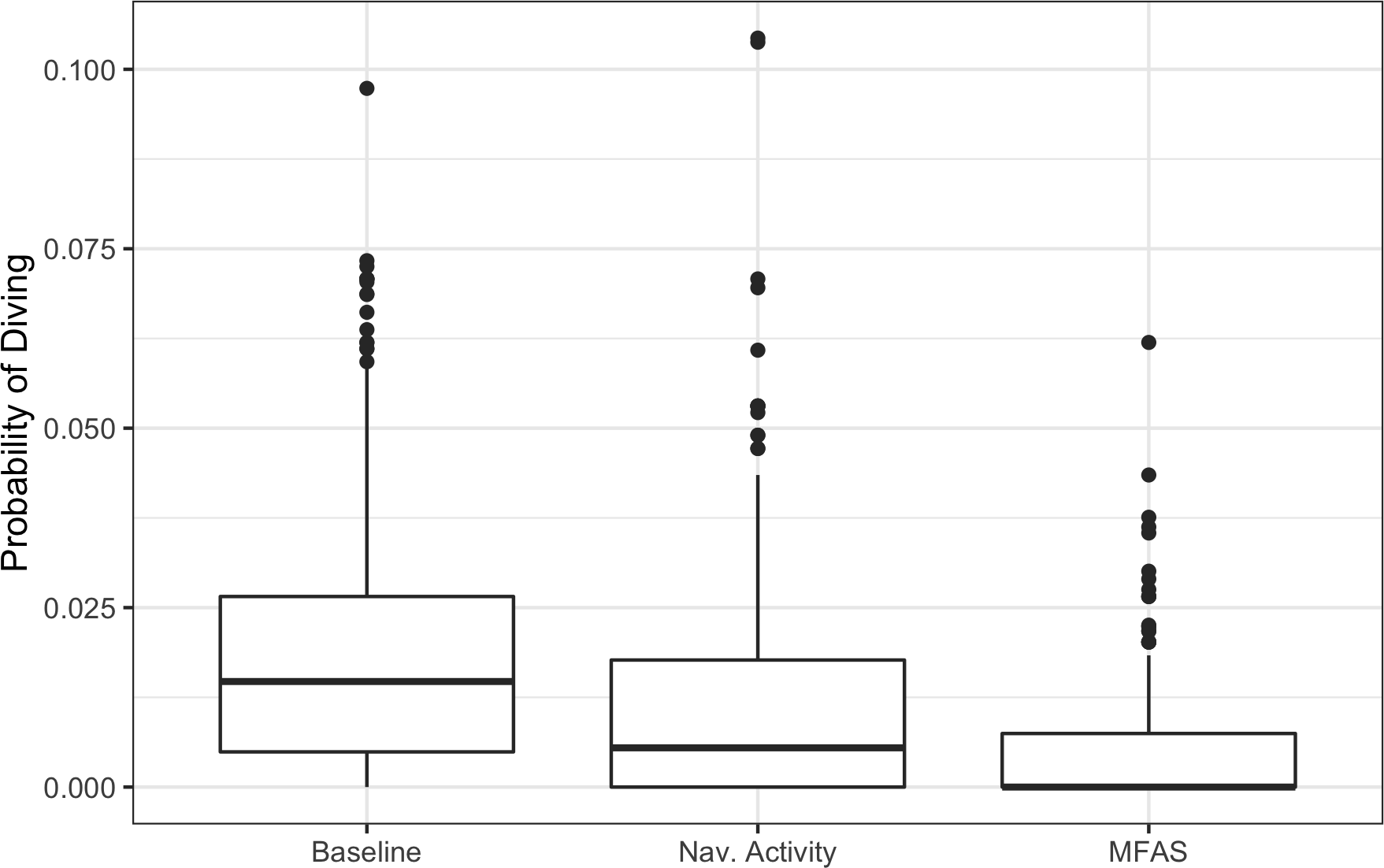


Figure S2.3: Boxplot of observed probability of a GVP for all hydrophones and SCCs (vertical axis) during baseline period, when naval activity was present, and when MFAS was present (horizontal axis). Each data point represents one hydrophone during one SCC and one phase of the training exercise.

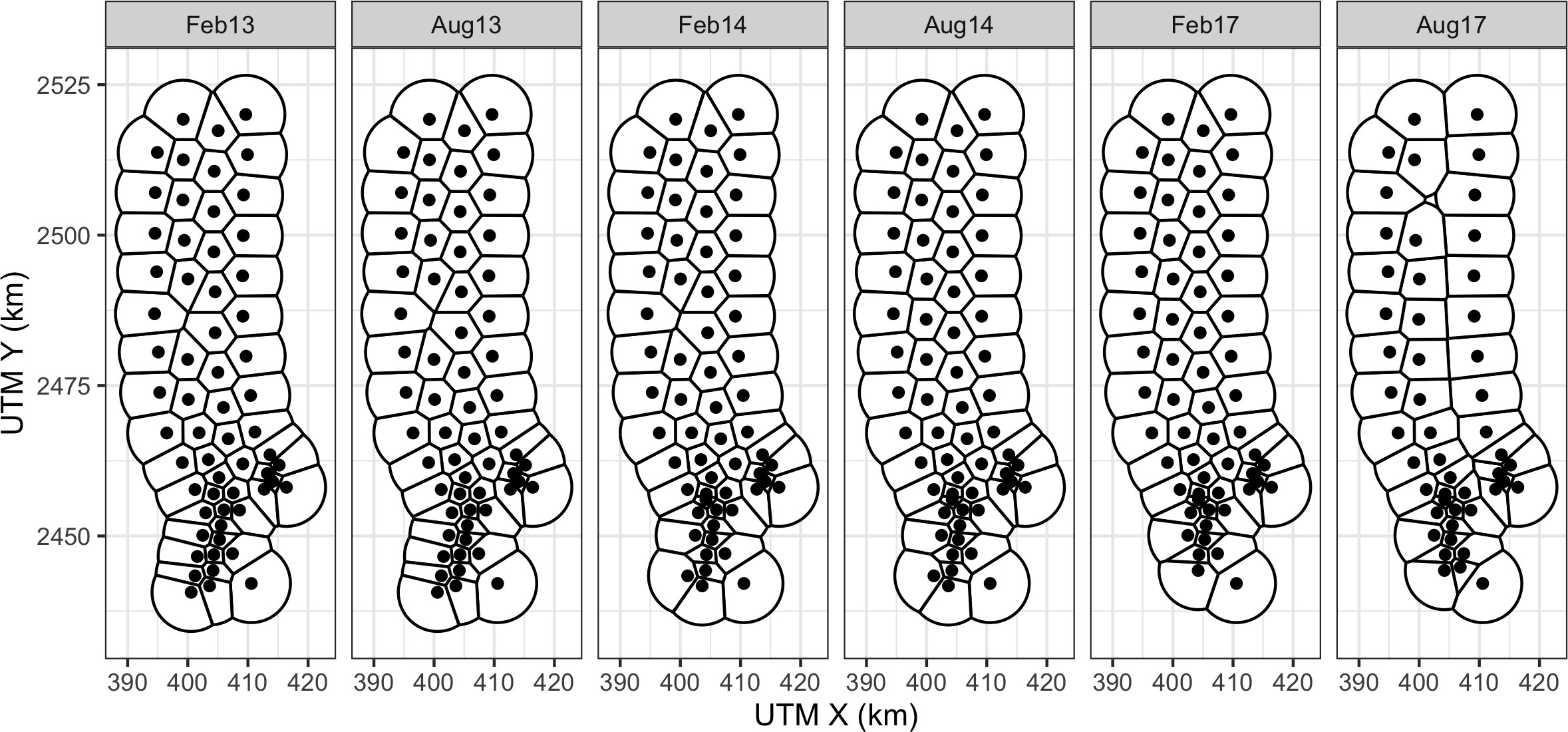


Figure S2.4: PMRF range tessellations for each of six recorded SCCs. Black lines indicate boundaries of hydrophone tiles. Black dots indicate approximate hydrophone locations.

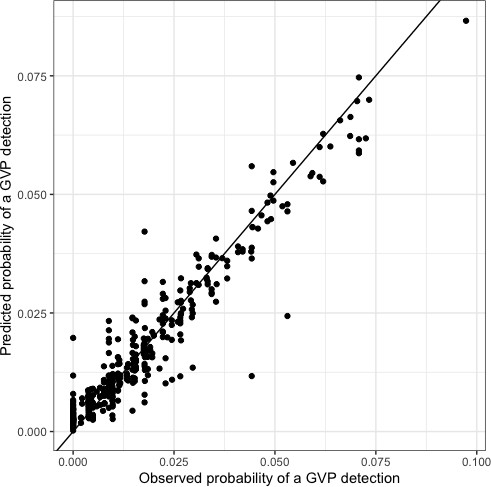
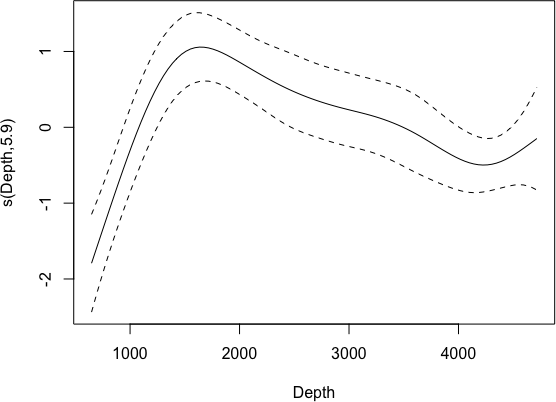


Figure S2.5: Observed (horizontal axis) versus M1 predicted (vertical axis) probability of detecting a GVP at each hydrophone during the baseline period.

682



683

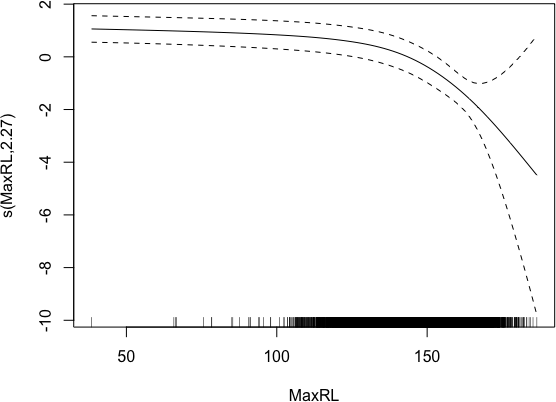
684

685

686

687

Figure S2.6: Spline for the relationship between P(GVP) and depth from M1 on the logit-link scale. Solid line: best fit; dashed lines: 95% CIs.



688

689

690 Figure S2.7: Spline for the relationship between P(GVP) and maximum received level from 691 M3 on the logit-link scale. Solid line: best fit; dashed lines: 95% CIs.

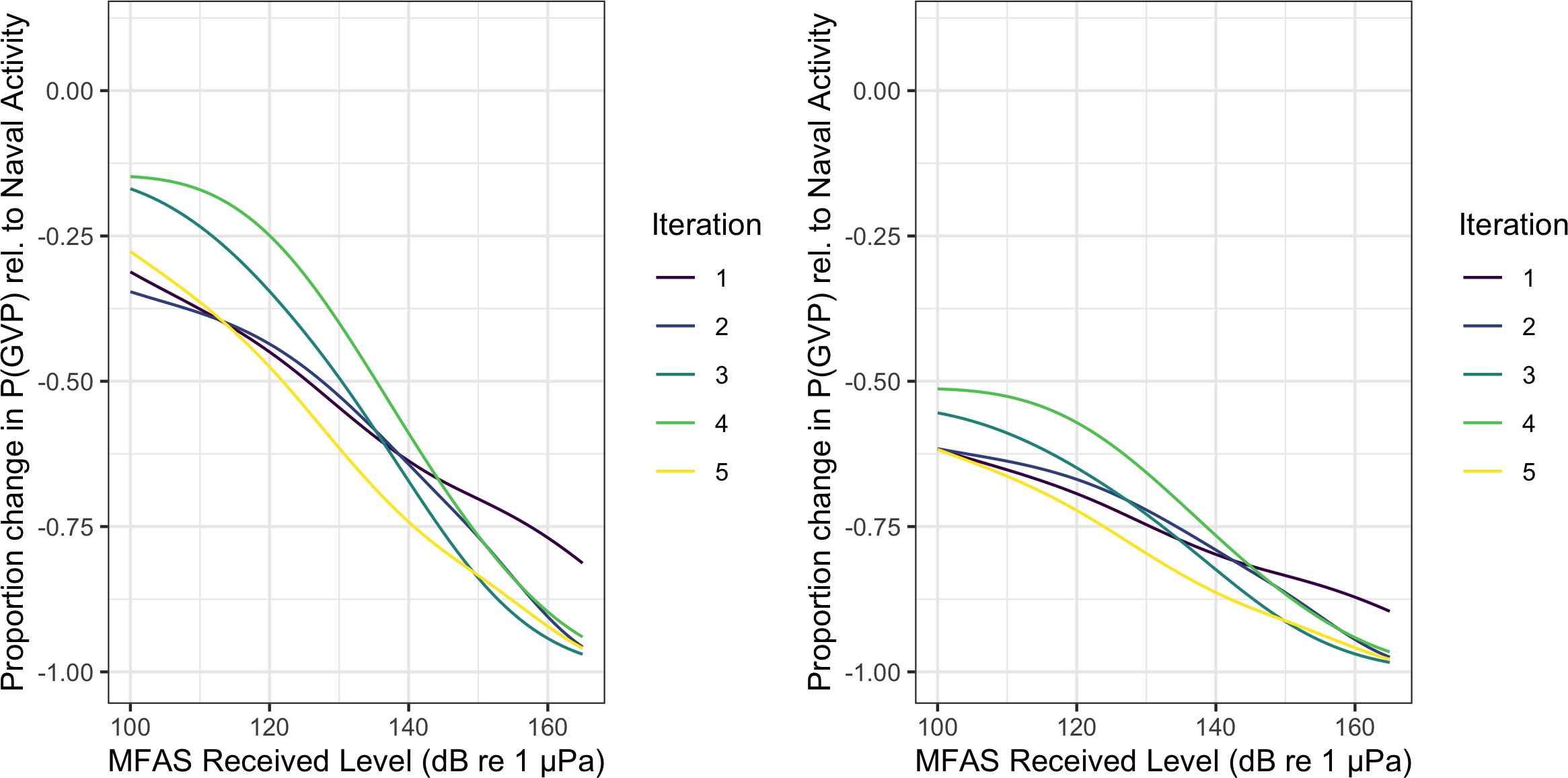


Figure S2.8: Example of five iterations (colored lines) of the 5,000 posterior samples of the expected change in the probability of detecting a group vocal period (vertical axis) with increasing MFAS received level (horizontal axis) relative to when naval training activity but no MFAS was present on the range (left panel) and to when neither naval training activity nor MFAS were present on the range.

692 **S3: Single GAM**

693 A single GAM could be used to quantify the effect of naval sonar on Blainville’s beaked

694 whales. Here, we present such a model and compare the results to the results obtained using

695 the multi-stage model presented in the main text of the manuscript.

We modelled the probability of a GVP at tile *i* in SCC *s* at time *t* as a Bernoulli trial:

GVP*i,s,t* ∼ Bin(1*, µi,s,t*). The linear predictor on the logit scale was given as:

logit (*µi,s,t*) = *β*0 + *β*1NavTrain*t* + *f* (MRF*i,s*) + *f* (Depth*i*) + *f* (MaxRL*i, t*)Sonar*t* + log*e Ai*)

696 where *β*0 is an intercept, *β*1NavTrain*t* is the effect of naval training times an indicator

697 variable for whether naval training was present or absent at time *t*, *f* (MRF*i,s*) denotes the

698 Markov random field used to smooth space, *f* (Depth*i*) is a smooth of depth (using a thin 699 plate spline; Wood et al. 2003), *f* (MaxRL*i,t*)Sonar*t* is a smooth of sonar received level (using 700 a thin plate spline) times an indicator variable for whether sonar was present or absent at 701 time *t*, and log*e Ai* is an offset for the area (in km2) of each tile, *Ai*.

702 We fit the model to the same data used in M1, M2, and M3 (see Methods section of main

703 manuscript for details) using mgcv (Wood, 2017).

704 This single GAM (Fig. [S3.1)](#_bookmark0) predicts a 41% (95% CI 34%-46%) decrease in *P* (GVP) when 705 naval training is present compared to the baseline period, whereas the multi-stage GAM (Fig. 706 4) predicts a decrease of 44% (95% CI 38%-49%). The single GAM predicts that at a MFAS 707 received level of 150 dB re 1 *µ*Pa, *P* (GVP) decreases by 87% (95% CI 71%-95%) relative to 708 when only naval training is present, whereas the multi-stage model predicts the same

709 decrease of 87% with a narrower credible interval (95% CI 81%-92%). Relative to when only 710 naval training is present, the single GAM predicts a 50% reduction in *P* (GVP) at a MFAS 711 received level of 120 dB re 1 *µ*Pa, whereas the multi-stage model predicts a 50% reduction at 712 a MFAS received level of 132 dB re 1 *µ*Pa.

713 The major difference between this single GAM and the multi-stage model presented in the 714 main text of the manuscript is that here, the spatial smooth is constructed using data from 715 the baseline, naval training, and MFAS periods of each SCC. Therefore, the spatial

716 distribution of MFAS may influence the predicted distribution of Blainville’s beaked whales.

717 Using a single GAM leads to similar point estimates of the impact of sonar with greater

718 uncertainty than the multi-stage model.

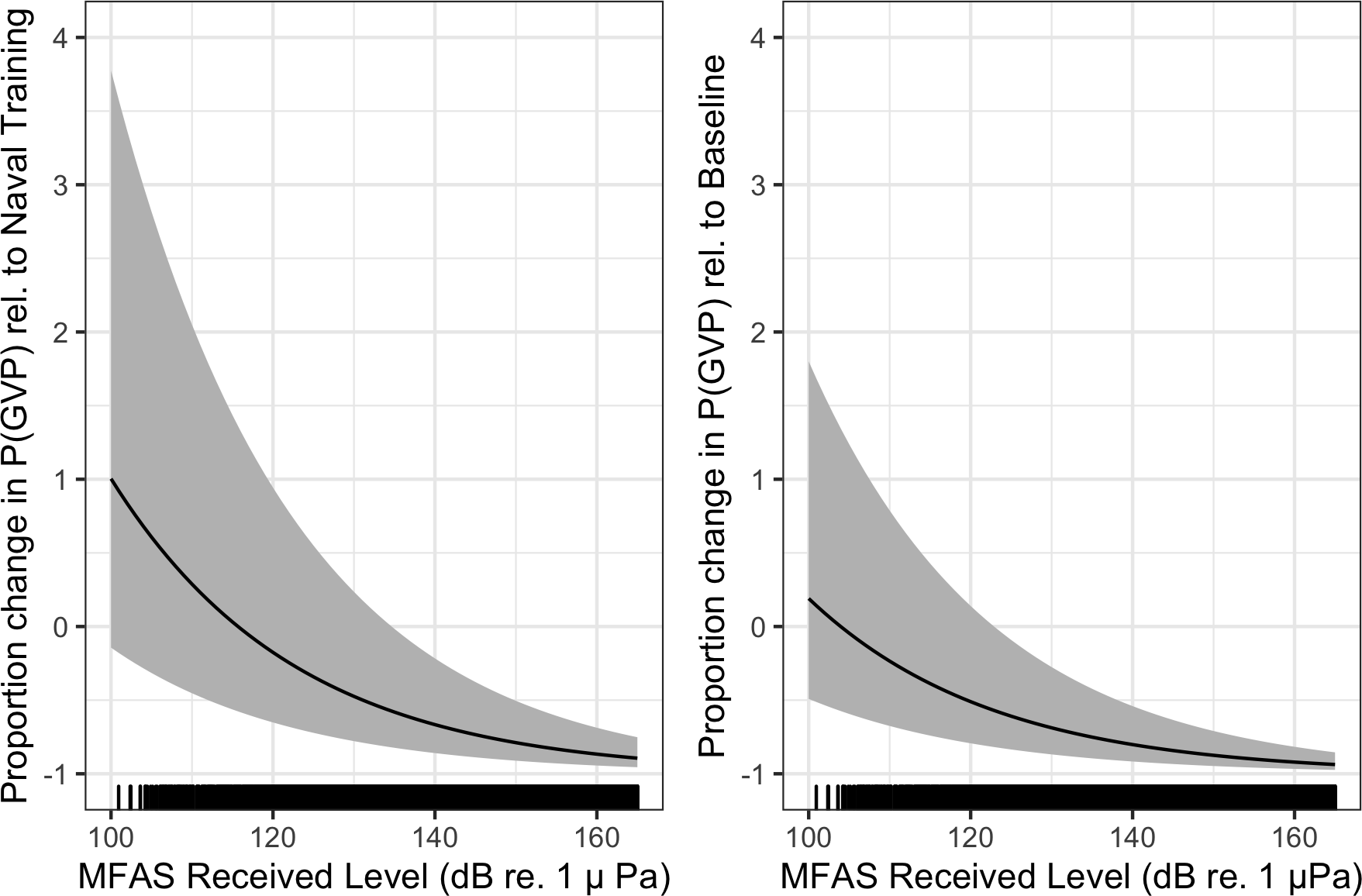


Figure S3.1: Results from a single GAM: Median (black line) and 95% CIs (gray shading) expected change in the probability of detecting a group vocal period (vertical axis) with increasing MFAS received level (horizontal axis) relative to when naval training activity but no MFAS was present on the range (left panel) and to when neither naval training activity nor MFAS were present on the range.