7.A(1)rev1
153rd SSC

Relative efficacy of paired tori-lines and night-setting with blue-dyed bait as seabird bycatch mitigation measures deployed in the Hawaii-based shallow-set longline fishery

August 2024

Western Pacific Regional Fishery Management Council
1164 Bishop Street, Suite 1400, Honolulu, Hawaii

A Report of the Western Pacific Regional Fishery Management Council
1164 Bishop Street, Suite 1400, Honolulu, Hawaii 96813 USA

Prepared by Milani Chaloupka [1,2], Eric Gilman [3], Asuka Ishizaki [4], Christi Smith [5], Eric Kingma [6], Yonat Swimmer [7], Lynn Rassel [8]

[1] Ecological Modelling Services Pty Ltd, Australia
[2] Marine Spatial Ecology Lab, School of the Environment, University of Queensland, Australia
[3] Fisheries Research Group, Honolulu, USA
[4] Western Pacific Regional Fishery Management Council, Honolulu, USA
[5] AIS Inc, USA
[6] Hawaii Longline Association, Honolulu, USA
[7] NOAA Fisheries, Pacific Islands Fisheries Science Center, Honolulu, USA
[8] NOAA Fisheries, Pacific Islands Regional Office, Honolulu, USA

This report may be cited as follows:

Chaloupka, M., Gilman, E., Ishizaki, A., Smith, C., Kingma, E., Swimmer, Y., Rassel, L. 2024. Relative efficacy of paired tori-lines and night-setting with blue-dyed bait as seabird bycatch mitigation measures deployed in the Hawaii-based shallow-set longline fishery. Western Pacific Regional Fishery Management Council. Honolulu, Hawaii.

Published in the United States by the Western Pacific Regional Fishery Management Council

ISBN: 978-1-950193-46-2

Cover Photo: Paired tori-line and an albatross observed through the electronic monitoring video camera during the experimental trip.

Contents

FIGURES

ACKNOWLEDGEMENTS

The project team thanks the F/V Lady Luck owner Calvin Huynh, captain Kevin Huynh and crew who participated in the experiment. We are extremely grateful to Calvin Huynh for authorizing the release of government confidential observer data records and to Lesley Hawn of the National Marine Fisheries Service Pacific Islands Regional Office for confirming that this report meets their confidentiality requirements. This project was funded by the NOAA Cooperative Research Program through a NOAA & Council Award NA20NMF4410013 — and we are grateful to John Wang (NMFS PIFSC) for helping towards securing that funding. We thank and acknowledge John Peschon (NMFS PIRO) who helped design and implemented an observer data query, Adam Briles (fisheries observer) who collected data during the experiment and Matthew Carnes (United Fishing Agency, AIS Inc) who installed the EM system and built and installed the paired tori-lines. We also thank AIS Inc staff for conducting EM reviewing and Integrated Monitoring for supplying and preparing the EM system.

EXECUTIVE SUMMARY

Night-setting is a well-known seabird bycatch mitigation measure (Jiménez et al 2020) and has been effective in reducing albatross interactions in the Hawaii-based shallow-set longline fishery. Tori-lines are also known to be an effective seabird bycatch measure for many pelagic fisheries including for the Hawaii-based deep-set tuna longline fishery (Gilman et al 2021, Chaloupka et al 2021), but have not been evaluated in the Hawaii-based shallow-set fishery since trials were conducted in the late 1990s and early 2000s that identified safety and operational concerns.

The Hawaii based shallow-set fishery participants have expressed interest in exploring alternative combinations of seabird mitigation techniques that could allow for greater operational flexibility in the start of gear setting time to include the sunset hours, and tori-lines were identified as a potential alternative to night-setting. Vessels are required to start setting at least one hour after local sunset under existing regulations applicable to the Hawaii-based shallow-set longline fishery set. So, the Western Pacific Regional Fishery Management Council, in collaboration with the Hawaii Longline Association (HLA), National Marine Fisheries Service (NMFS) Pacific Islands Regional Office and NMFS Pacific Islands Fishery Science Center, conducted a pilot study to evaluate alternative methods of discouraging seabird interactions while providing operational flexibility during setting in the shallow-set longline fishery.

The study evaluated seabird bycatch risk in fishing sets that used paired tori-lines starting prior to local sunset (1-2 hours earlier than the existing night-setting requirement) that were then compared to sets deployed with both night-setting and blue-dyed bait. The existing regulatory requirement for strategically discharging offal during setting was withheld throughout the study to reduce potential confounding factors. The study was conducted under a NMFS Experimental Fishing Permit issued to the HLA. The study was designed as a limited scale field study to provide a preliminary evaluation of bycatch risk during sunset hours while using an alternative mitigation measure compared to the existing night-setting operation.

The field experiment comprised 44 sets deployed during 2 trips between February and April 2024 that was undertaken by one Hawaii-based shallow-set longline vessel targeting swordfish. We used 39 of those 44 sets since 5 sets were not consistent with the prescribed experimental protocols. The sets were deployed without variability in gear designs and gear deployment methods other than the deployment of paired tori-lines, the time-of-day of setting, and the use of blue-dyed fish bait. All the seabird interactions during this experiment involved black-footed (*Phoebastria nigripes*) and Laysan albatrosses (*Phoebastria immutabilis*). The Laysan albatross accounted for 83% of the captures, 83% of contacts and 73% of attempts. We combined both species records into a generic \`albatross\` category.

We used a Bayesian modelling workflow to statistically estimate: (1) the number of \`albatrosses\` captured during gear setting, as recorded by observers during the gear haulback and (2) albatross interactions during setting (attempts to contact a baited hook, actual contacts with a baited hook) recorded using onboard video-based electronic monitoring systems. There were 23 captured albatrosses, 66 albatross gear contacts and 215 attempts to contact the gear. Most noteworthy, ca 96% of the 23 albatross captures occurred on sets deployed with paired tori-lines

and only 1 albatross was caught on sets deployed with night-setting (in combination with blue-dyed fish bait).

While the two albatross species in this fishery were primarily at-risk during twilight hours, some night-time captures occurred. Of the 22 albatrosses caught on tori-line sets, 21 were captured before nautical dusk or within one hour after sunset. Only 2 albatrosses were captured at night—one during a tori-line set and one during a night-set with blue-dyed bait. This highlights the higher albatross capture risk associated with daytime and twilight fishing periods, even when using tori-lines.

We found that night-setting (with blue-dyed fish bait) was an overwhelmingly more effective seabird bycatch mitigation measure than paired tori-lines (with sets beginning during the daytime but concluding during the night). Specifically, we found that:

- Albatrosses were ca 38 times (95% HDI[1]: 5 - 945) more likely to be **captured** when paired tori-lines were deployed in partial daytime sets compared to the night-sets with blue-dyed fish bait

While sample sizes were too small to derive robust estimates of relative efficacies at reducing albatross attempts and contacts, our findings suggest that:

- Albatrosses were > 1000 times (95% HDI: 1 - ∞) more likely to **contact** a baited hook when paired tori-lines were deployed in partial daytime sets compared to the night-sets (with blue-dyed fish bait)
- Albatrosses were > 3000 times (95% HDI: 2 - ∞) more likely to **attempt** to contact a baited hook when paired tori-lines were deployed in partial daytime sets compared to the night-sets (with blue-dyed fish bait)

[1] HDI = highest posterior density interval (the shortest uncertainty interval)

1. INTRODUCTION

Incidental capture in pelagic longline fisheries is a major threat to the conservation of pelagic seabirds and reduces fishing efficiency (Phillips et al 2016, Bakker et al 2018, Dias et al 2019). Several gear technologies have been developed to mitigate seabird bycatch in pelagic longline fisheries (ISSF 2024, ACAP 2023).

Regional fisheries management organizations (RFMOs) allow their member fisheries to select from a range of options, including night-setting and deployment of tori-lines, to mitigate seabird bycatch in designated areas (IATTC 2011, ICCAT 2011, WCPFC 2018, SIOFA 2022, IOTC 2023). All options included within a category might be assumed to have equivalent efficacies at mitigating seabird bycatch risk. Single and paired tori-lines and night-setting can significantly reduce seabird bycatch (Jimenez et al 2020, Gilman et al 2021, ACAP 2023). A tori-line, initially developed by Japanese tuna longline fishers (Brothers et al 1999), is a line with streamers that is towed from the stern of the vessel as crew set baited hooks that can deter seabirds from scavenging bait from hooks.

Following successful trials in the Hawaii-based deep-set tuna longline fishery (Gilman et al 2021, Chaloupka et al 2021), the Western Pacific Regional Fishery Management Council recommended replacing blue-dyed fish bait and strategic offal discharge with tori-lines for stern-setting deep-set longline vessels. The new regulations became effective April 2024 (50 CFR 665.815). In the process of developing the regulatory amendment for the Hawaii-based deep-set fishery under the Pelagic Fishery Ecosystem Plan, the Council considered whether blue-dyed bait may also be removed from the suite of measures required for the Hawaii-based shallow-set longline fishery.

The Hawaii-based shallow-set longline fishery has been required to use seabird mitigation measures since 2004. The primary seabird bycatch mitigation measure required in the fishery is night-setting in combination with blue-dyed fish bait and strategic offal discharges. Alternatively, shallow-set vessels may side-set in combination with using weighted branchlines and bird curtains. Prior to the seabird bycatch mitigation regulations fishers had adjusted their setting time according to the moon phase to take advantage of the diel vertical migration of swordfish to increase fishing efficiency. Under the night-setting requirement, fishers are required to start setting at least one after local sunset and finish setting no later than local sunrise. This reduces the time available for fishers to deploy their fishing gear and impacts fishing efficiency. Fishers may compensate for the lost gear setting time by deploying gear at faster vessel speeds, leading to greater fuel consumption.

At its 185th meeting in March 2021, the Council recommended further research to develop appropriate seabird mitigation measures for the Hawaii-based shallow-set fishery. The Council noted that removing both blue-dyed bait and offal discharge would result in some vessels not being in compliance with international seabird requirements of the Western and Central Pacific Fisheries Commission and the Inter-American Tropical Tuna Commission and concluded that additional research to identify alternative mitigation measures was warranted due to the differences in operational characteristics between the Hawaii-based shallow-set and deep-set longline fisheries. In promoting further research, the Council recommended identifying a combination of mitigation measures that would continue to ensure seabird deterrence during dusk compared to the existing night-setting suite of measures.

An experimental field trial using a commercial pelagic longline vessel, an onboard observer and a video-based electronic monitoring system was undertaken to evaluate whether fishing operations starting prior to local sunset (ca 1 to 2 hours earlier than the existing night-setting requirement) using paired tori-lines provides an equivalent efficacy at mitigating seabird bycatch as night-setting with blue-dyed fish bait in the Hawaii-based shallow-set longline fishery. The existing requirement for longline fishing vessels to employ strategic offal discharges during the set was withheld during the study to reduce potentially confounding factors. It was a limited trial intended as an interim step to determine whether alternative seabird mitigation measures to the current night-setting requirement might be warranted such as the use of tori-lines in the Hawaii-based shallow-set longline fishery.

The study was conducted under a National Marine Fisheries Service (NMFS) Experimental Fishing Permit (EFP) issued to the Hawaii Longline Association on 24 March 2022 and reissued on 20 September 2023 (NMFS 2023). To allow for experimental sets with partial daytime sets while using paired tori-lines, the EFP exempted the participating vessel from the existing requirement for blue-dyed fish bait and strategic offal discharge while setting — as well as the requirement to set at least one hour after local sunset. The findings from this study should therefore provide sufficient evidence to determine whether paired tori-lines are a viable seabird bycatch mitigation measure in the Hawaii-based shallow-set longline fishery.

2. METHODS

2.1. Experimental Setting

The experiment comprised 44 longline sets deployed during 2 trips undertaken by the Hawaii-based pelagic longline fishing vessel Lady Luck between February and April 2024. Five (5) of the 44 sets made during the 2 experimental fishing trips were discarded because those sets did not follow the experimental protocols — 2 sets that were deployed with tori-line were excluded because the tori-line broke and was used for only part of the set, and 3 sets were excluded that deployed a tori-line *with* night-setting (the set began after the local time of nautical dusk as well as later than 1 hour after the time of local sunset).

Here the longline set is the sampling unit and is treated as a random effect in all subsequent models used to estimate any treatment-specific effects on albatross bycatch rates. Henceforth, the 39-set experiment comprised 2 treatments: (1) paired tori-lines and (2) night-setting with blue-dyed thawed fish bait. Tori-lines were deployed for 17 of the 39 sets with sets beginning prior to the local time of nautical dusk and night-setting (with blue-dyed bait) for 22 sets made between nautical dusk and nautical dawn — nautical dawn and dusk were calculated for each set using the set start geolocation and the date using the `lubridate` (Grolemund & Wickham 2011) and `suncalc` (Thieurmel & Elmarhraoui 2022) R packages. The start-time for the 17 paired tori-line sets was ca 1-2 hours prior to the set time of `*1 hour after local sunset*` and < 2 hours prior to nautical dusk. Other gear designs and fishing methods were consistent for all sets, including 75g weighted swivels, ca 7m leaders (distance between weight and hook), use of mackerel bait, use of green chemical lightsticks on some branchlines, the deployment of 5 hooks between floats, and no deployment of offal or spent bait during setting.

This study used a modified version of a light-weight short-streamer design tested in the Hawaii-based deep-set longline fishery (Gilman et al 2021). Compared to the design used in the deep-set longline fishery, the aerial extent was set at ca 65m to account for the slower gear sink rate and associated greater distance astern needed. The drag section length was also adjusted in response to the longer aerial extent. All other components of the design were consistent with the design for the deep-set longline fishery. A summary of the tori-line design used is given in the Supplementary Material.

2.2. Data

The study used data sourced from (1) onboard observer records of seabirds captured on the gear and (2) review of seabird interactions with the longline gear recorded using an onboard video-based electronic monitor system (EM: *see* Gilman et al 2020) following pre-determined criteria for classifying an interaction as an attempt or a contact (Gilman et al 2021). Only seabird captures that were likely to have occurred during the set, and not during the gear haulback or soak, were included in the study analysis. Data collected through the Pacific Islands Regional Observer Program that would otherwise be confidential under the Magnuson-Stevens Fishery Conservation and Management Act are included in this report with authorization from the vessel owner to release the data for the EFP trips for the purpose of the study analysis.

2.2.1. Captures on the Gear

There were 19 Laysan and 4 black-footed albatross captures recorded by onboard observers during the gear haulback — so Laysan albatross captures accounted for ca 83% of all seabird captures. All 23 albatrosses captured during setting were retrieved during gear haulback. For each seabird capture recorded by the onboard observer during the gear haulback that was determined to have been captured during the set, we estimated whether the capture occurred prior to or after the time of nautical dusk by using information from the observer data on the haulback direction, the number of the float nearest to the branchline on which the seabird was captured, the total number of floats deployed in that set, and the duration of the setting operation.

All sets were retrieved in the opposite direction as deployed. The value of the ratio of the number of floats from the start of a set on which a capture occurred to the total number of deployed floats was used to estimate the proportion of the duration of the set when the capture event occurred. This latter value of the amount of time into the set when the seabird captured occurred was then added to the time of day of the start of the set to estimate when the seabird captured occurred, which was compared to the local time of nautical dusk and to onset time of `*1 hour after local sunset*`.

2.2.2. Interactions with the Gear/Baited Hook

The EM-derived data comprised the number of albatross attempt and contact interactions, which were for either the black-footed albatross or the Laysan albatross exposed to the Hawaii-based shallow-set longline fishery. The Laysan albatross accounted for ca 83% of attempts to contact a baited hook and ca 73% of the actual contacts with a baited hook. Therefore, the 2 albatross species records were combined into a generic albatross species category as there were too few data to estimate species-specific interaction rates with baited hooks. Nonetheless, any inference regarding attempted or actual contacts with a baited hook metric is based mainly on data for the Laysan albatross rather than the black-footed albatross. Seabird scan counts during setting were also derived from the EM data.

2.2.3. Response Metrics

The focus in this study was on the following albatross interactions or response metrics (see Figure 1): (1) number of albatross-attempts to contact a baited hook recorded for each set, (2) number of albatross-contacts with a baited hook recorded for each set and (3) number of albatrosses recorded as captured during setting for each set. There were too few attempts or contacts > 2 or 3 per set to model meaningfully (Figure 1a-b), so these response metrics were more appropriately restructured as a binary or Bernoulli response (0,1) variable with the attempt rate being recoded as either 0 for no attempts and 1 for one or more attempts per set. The same procedure was applied to the albatross contacts data. The focus in this report is on the statistical modelling of the observer-sourced captures rather than the EM-sourced interaction data since there were zero interactions (attempts, contacts) recorded for the night-time-deployed sets — perhaps because EM failed to identify such seabird interactions at night.

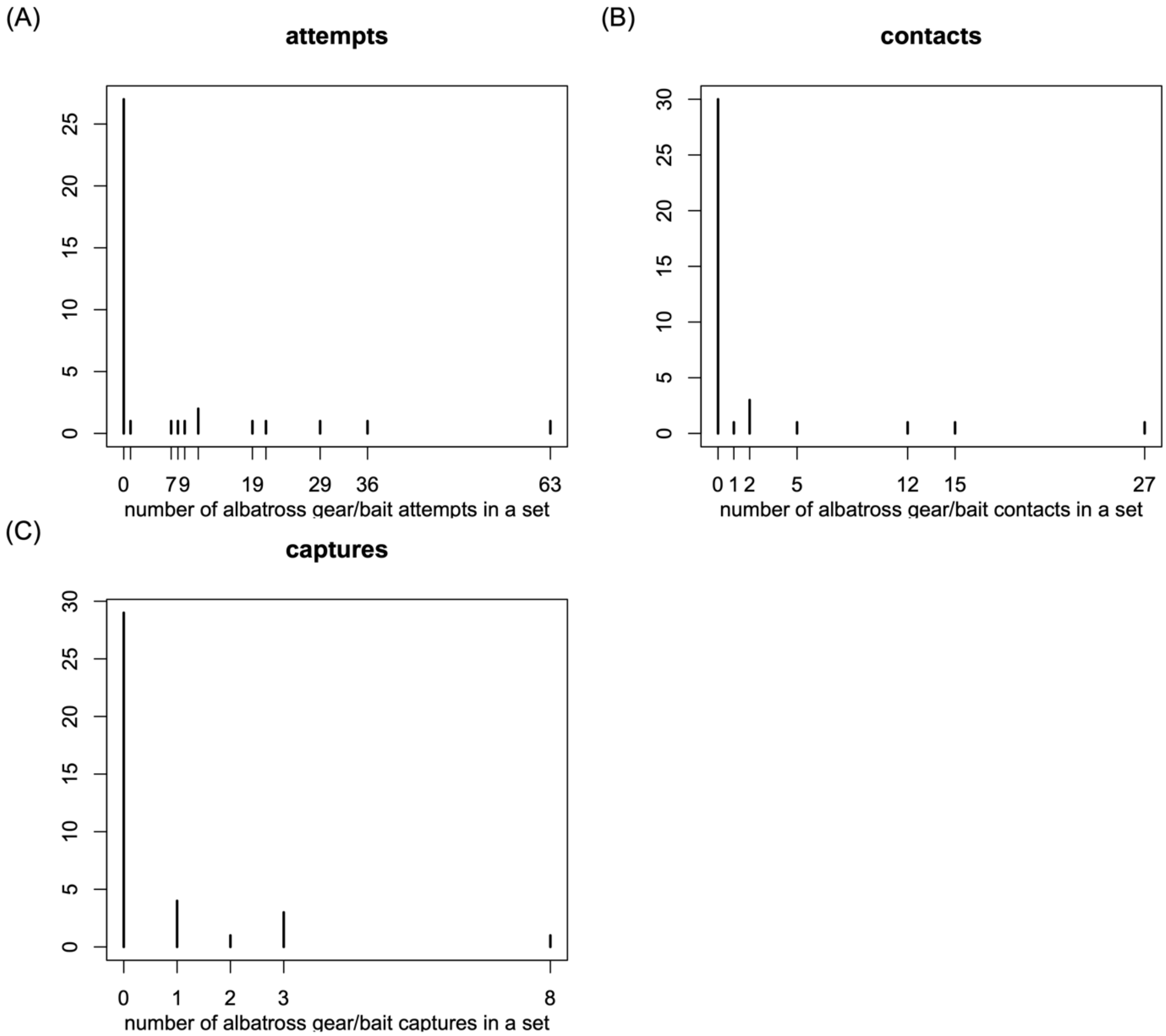

Figure 1 RESPONSE METRICS. Summary of albatross interactions with the gear/bait for each of the 39 sets.. (**Panel A**) shows frequency of albatross attempts for each set. (**Panel B**) shows frequency of albatross contacts for each set. (**Panel C**) shows frequency of albatross captures (drownings) for each set.

2.3. Statistical Modelling Approaches

2.3.1. Data Exploration and Marginal Response Summaries

We summarized the treatment-specific proportion of sets with ≥ 1 albatross capture using the posterior mean capture rate and highest posterior density interval (HDI) by using the `binom` R package (Dorai-Raj 2014) to sample from a binomial likelihood with a Bayes-Laplace prior (Tuyl et al 2008) that accounts for zero recorded captures for some sets — see Gilman et al (2024) for a recent example using this approach to estimate odontocete depredation rates in the Hawaii-based tuna longline fisheries. This procedure was also applied to summarize the proportion of sets with at least 1 contact and the proportion of sets with at least one attempt. We also estimated both the odds ratio and risk ratio (relative risk) association summaries for the binary structured capture data using the `epitools` R package (Aragon 2020) — this was not possible for the contact or attempt metrics since there were zero interactions recorded for the 22 sets deployed as night-setting with blue-dyed fish bait.

The following sections now proceed with more specific statistical modelling of all recorded albatross captures (not just restructured into a binary response) and both the binary structured interaction response metrics (attempt: 0 v ≥ 1 attempt, contact: 0 v ≥ 1 contact).

2.3.2. Modelling the Albatross Captures

The statistical modelling approach used here was based on a Bayesian inference workflow (Gabry et al 2019) based on either a generalized linear mixed model (GLMM)[2] or a generalised additive mixed regression model structure (GAMM: Fahrmeir & Lang 2001, Kammann & Wand 2003) with an appropriate response-specific likelihood such as Poisson or negative binomial for the count data (Congdon 2003) — see Gilman et al (2021) and Chaloupka et al (2021) for recent application of this approach for the assessment of seabird bycatch risk in the Hawaii-based pelagic deep-set longline fishery. The Bayesian approach to statistical modelling provides a powerful way to account for uncertainty in the data, the model parameters and the model structure using probability theory (van de Schoot et al 2021).

The Bayesian modelling workflow used here comprised: (1) assessment of potential prior and model likelihood sensitivity for the priors used for (2) a robust statistical model accounting for experimental design constraints and potential predictors of interaction rates followed by (3) graphical posterior predictive checks of the adequacy of the statistical model(s) fitted to the set-specific albatross capture data. Prior and likelihood sensitivity diagnostics were implemented using the `priorsense` R package (Kallioinen et al 2024) while the `bayesplot` R package was used for implementing the posterior predictive check tests. Residual diagnostics of the best-fit GLMM model were also assessed using the `DHARMa` (Hartig 2022) and `DHARMa.helpers` (Rodríguez-Sánchez 2024) packages for R.

Here the response metrics are count data (set-specific number of albatross seabirds caught on baited hooks) and so appropriately modelled using a regression model structured with either Poisson or negative binomial likelihood (Congdon 2003) — we also tested for zero-inflated negative binomial to address potential excess zero captures. Specifically, GLMMs and GAMMs with Poisson (or other) likelihood were fit to the albatross captures data while accounting for potentially informative predictors using the Stan computation engine (Carpenter et al 2017) via the `brms` interface (Bürkner 2017). All models here were implemented using weakly informative regularizing priors (Lemoine 2019).

The predictors included treatment (tori-line, night-setting with blue-dyed bait), number of hooks per set as a fishing effort metric, specific trip, wind force category (< Beaufort 4 or ≥ Beaufort 4), cloudiness (overcast or not), mean density of seabirds observed attending the vessel during setting and moonlight intensity. Cubic smoothing splines (Wood 2006) were used to

[2] We tested for nonlinear effects (number of hooks per set, seabird density attending the vessel) using a GAMM or simply accounting for proportional effort as an offset and so using a GLMM with no other nonlinear effects explored.

account for possible nonlinear functional form of the covariates such as seabird density near the vessel. Any potential spatial effect of the individual set geolocations was accounted for explicitly by including the specific trip as a factor (see Figure 2). The random effect structures (intercepts-only) included in the models were 39 individual sets used in this experiment to account for any heterogeneity in the capture rates not accounted for by the other predictors.

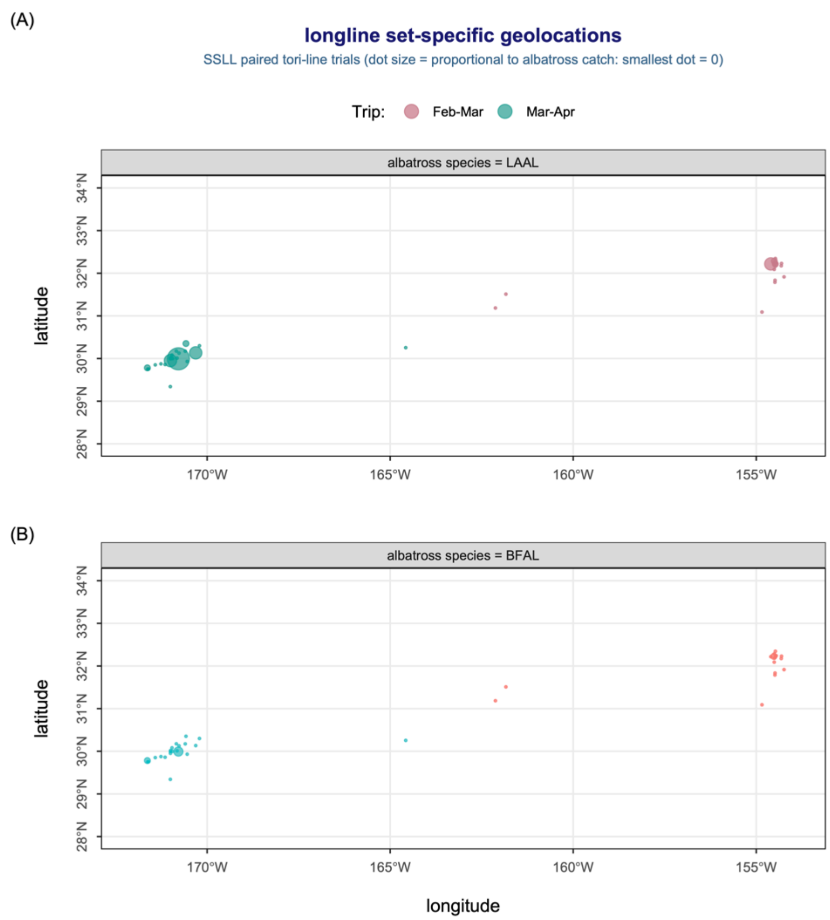

Figure 2 SET GEOLOCATIONS. Set-specific geolocations for each fishing trip for the 39 sets and showing the relative proportion of (A) Laysan and (B) black-footed albatrosses caught on the longline gear at each location.

Model selection was based on leave-one-out cross-validation metrics to estimate any comparative difference in expected predictive accuracy between the various models fitted such as (1) which model likelihood was most appropriate for this data set such as Poisson, negative binomial or zero-inflated negative binomial for the count or capture data (Vehtari et al 2017) or (2) alternative functional forms of the fishing effort (number of hooks) as an offset (proportional) or as nonlinear and hence nonproportional (Davies & Jonsen 2011). The weight of evidence in favour of one model over any other candidate models was also assessed using Bayesian stacking, which is the Bayesian analogue of model averaging (Yao et al 2018).

The posterior samples for all models were sourced from 4 chains and 9,000 iterations after a warmup of 1000 iterations per chain. Therefore, the posterior for each estimate comprised 5,000 samples or draws that were used to derive the uncertainty intervals (HDIs or highest posterior density intervals: Kruschke & Liddell 2018) using the `tidybayes` package for R (Kay 2023). This summary display showing both the full posterior distribution of the parameter coupled with the summary metrics (median, 95% HDI) helps to support a more precise form of communicating parameter uncertainty (van der Bles et al 2019).

A probability statement about the *existence* of a particular effect and its direction, such as tori-line effects, can be determined with those 5,000 draws using the probability of direction metric proposed recently by Makowski et al (2019) — also known as the maximum probability of an effect and derived here using the `BayestestR` package for R (Makowski et al 2019).

Model convergence was assessed using parameter-specific diagnostics such as multiple chain rank plots, bulk and tail effective sample size metrics and a rank-based *Rhat* statistic (Vehtari et al 2021) — all diagnostics reflected convergence of all models used here. Further evaluation of the best-fit-model was assessed using graphical posterior predictive checks (Gabry et al 2019) including using a hanging rootogram as one of the checks (Kleiber & Zeileis 2016). All inference was made using the best-fit model.

Finally, the estimated effects summaries based on the best-fit conditional regression models were then adjusted for variable sample size of the treatments using the predicted marginal means approach (Searle et al 1980, Lenth 2016) and implemented using the `emmeans` package for R (Lenth 2020). The `ggplot2` (Wickham 2016), `ggdist` (Kay 2024) and `colorspace` (Zeileis et al 2020) packages for R were used for all summary graphics while the `patchwork` package for R (Pedersen 2020) was used for all multi-panel arrangements.

2.3.3. Modelling the Albatross Attempt and Contact Rates

Here the response metrics are binary data (for example: 0 = no attempts, 1 = ≥ 1 attempt to contact a baited hook during the set) and so are appropriately modelled using a regression model with Bernoulli likelihood — which is a special case of a binomial likelihood but now with a single trial (Congdon 2003). Specifically, GLMMs with Bernoulli likelihood were fit to the albatross interaction data (ATTEMPT = ≥ 1 albatross attempt to attack the gear, CONTACT = ≥ 1 albatross contact with a baited hook) while accounting for potentially informative predictors using the Stan computation engine (Carpenter et al 2017) via the `brms` interface (Bürkner 2017).

The predictors included just the treatment (tori-line, night-setting) since there were zero recorded attempts or contacts on any night-setting with blue bait-deployed sets — so meaningless to include any potential predictors in the model other than the treatment-specific effect. The random effect structure (intercepts-only) was included in the model and again comprised the 39 individual sets to account for any heterogeneity in the interactions rates not accounted for by the treatment effect.

All models here were implemented using weakly informative regularizing priors (Lemoine 2019) and compared using leave-one-out cross-validation metrics (Vehtari et al 2017) with the R2D2M2 prior proposed recently for multilevel regression models applied to sparse binary response data (Aguilar & Bürkner 2023). The weight of evidence in favour of one model over any other candidate models was also assessed using Bayesian stacking, which is the Bayesian analogue of model averaging (Yao et al 2018). We also derived the binary response classification rate or accuracy metric for the best-fit model using the confusionMatrix() function in the `caret` package for R (Kuhn 2008) as an appropriate measure of model adequacy.

3. Results

3.1. Marginal Response Summaries

The datasets for all 3 metrics were sparse and this does result in valid bycatch rate estimates, but wide uncertainty intervals as summarized in Table 1, where all 3 metrics are collapsed into simple (0,1) binary form. The expected capture, contact and attempt rates were all significantly higher on sets deployed with tori-lines than for sets deployed with night-setting (and blue-dyed fish bait). For instance, the expected capture rate (≥ 1 albatross) for sets deployed with paired tori-lines (partial daytime sets) was ca 47% (95% HDI: 26% - 69%) in this study. Whereas the expected albatross capture rate for night-sets (with blue-dyed bait) was ca 8% (95% HDI: 1% - 19%). And so on for the binomial rate estimates for the `contact` and `attempt` metrics (Table 1).

Table 1. Posterior mean capture and interaction rates and 95% highest posterior density interval (HDI) estimated by sampling from a binomial likelihood with a Bayes-Laplace prior to account for the zeros. N = number of sets, `yes` = number of those sets with that treatment for that specific metric (0, ≥ 1 albatross)

metric	treatment	yes	N	posterior mean	95% HDI lower	95% HDI upper
capture	tori-line	8	17	0.47	0.26	0.69
	night-set	1	22	0.08	0.01	0.19
contact	tori-line	8	17	0.47	0.26	0.69
	night-set	0	22	0.04	0.00	0.12
attempt	tori-line.	11	17	0.63	0.42	0.84
	night-set	0	22	0.04	0.00	0.12

The odds ratio[3] (Greenland 2004, Naimi & Whitcomb 2020) for the capture metric = 17.2 [(95% CI: 1.9-861.8); p=0.005 (two-sided Fisher exact)] — so, again, sets deploying paired tori-lines (partial daytime sets) were significantly more likely to catch albatrosses than night-setting (with blue-dyed bait). The risk ratio (or relative risk: Naimi & Whitcomb 2020) = 10.4 [(95% CI: 1.4-74.99); p=0.005], which reflects approximately a 900% increase in risk of an albatross caught on a set deployed with the paired tori-lines (partial daytime sets) than on a set with night-setting (blue-dyed fish bait). Based on these simple summaries of association-type statistical

[3] Using the `yes` and N data in Table 1 to structure the 2 x 2 contingency table. And similarly for calculating the risk ratio.

measures, night-setting was far more effective at mitigating albatross captures compared to longline sets deployed with the paired tori-lines.

While outside the objective of the study, an assessment of the marginal summary response of the expected albatross capture rate for a modified treatment of the portion of each set using paired tori-lines that occurred during the night [4] was compared with the expected capture rate for the treatment of night-setting with blue-dyed bait. Here, the posterior mean capture rate in the tori-line portion of sets at night was 0.11 (95% HDI: 0.003 - 0.24) compared to 0.08 (95% HDI; 0.01 - 0.19) for night-setting with blue-dyed bait, with no significant difference in the relative risk ratio = 1.01 (95% CI: 0.87 - 1.18). Again, it is important to recognize that this is weak inference as only one albatross was caught on a single night-based paired tori-line set and only one albatross was caught on a single set deployed with night-setting (and blue-dyed fish bait).

There were insufficient data for the attempt and contact measures to undertake this summary approach using the odds ratio and risk ratio — but it was possible using the binomial estimator with a Bayes-Laplace prior that accounts for the zero counts. So, the next few sections below address these issues within a fully Bayesian regression modelling framework that accounts also for potentially informative predictors of the bycatch or interaction rates.

3.2. Modelling the Albatross Capture Rates

The best-fit Bayesian GLMM with zero-inflated negative binomial likelihood (ZINB) identified by LOOcv and Bayesian stacking metrics fitted the albatross captures data (Figure 1c) adequately as shown for example by (1) the prior sensitivity diagnostics shown in Figure S1, (2) the residual diagnostics test shown in Figure S2 and, more importantly, (3) the posterior predictive check tests summarised in Figure S3.

This ZINB model had higher predictive accuracy than a GLMM with either Poisson or negative binomial likelihood and accounted for > 81% of the weight of evidence compared to the other 2 models. Moreover, this GLMM was a better fit to the data that using a GAMM with a nonproportional or nonlinear functional form for fishing effort (number of hooks per set) — the GLMM accounted for fishing effort as an offset, which implies that catch is proportional to effort.

All model diagnostics such as multiple chain rank plots, and the effective posterior sample size (ESS) metrics coupled with the rank-based diagnostic statistic Rhat < 1.01 (Vehtari et al 2021) reflected convergence of this model. All inference was then based on this model, conditional on the fixed effects or predictors (treatment, trip, cloudiness, wind force category) and the fishing-set-specific random effects included in the model that is shown in Figures 3 and S4.

[4] All 17 sets using tori-lines were deployed starting during daytime and ended during the night – so this amended treatment considers only the portion of the tori-line sets that were made during the night.

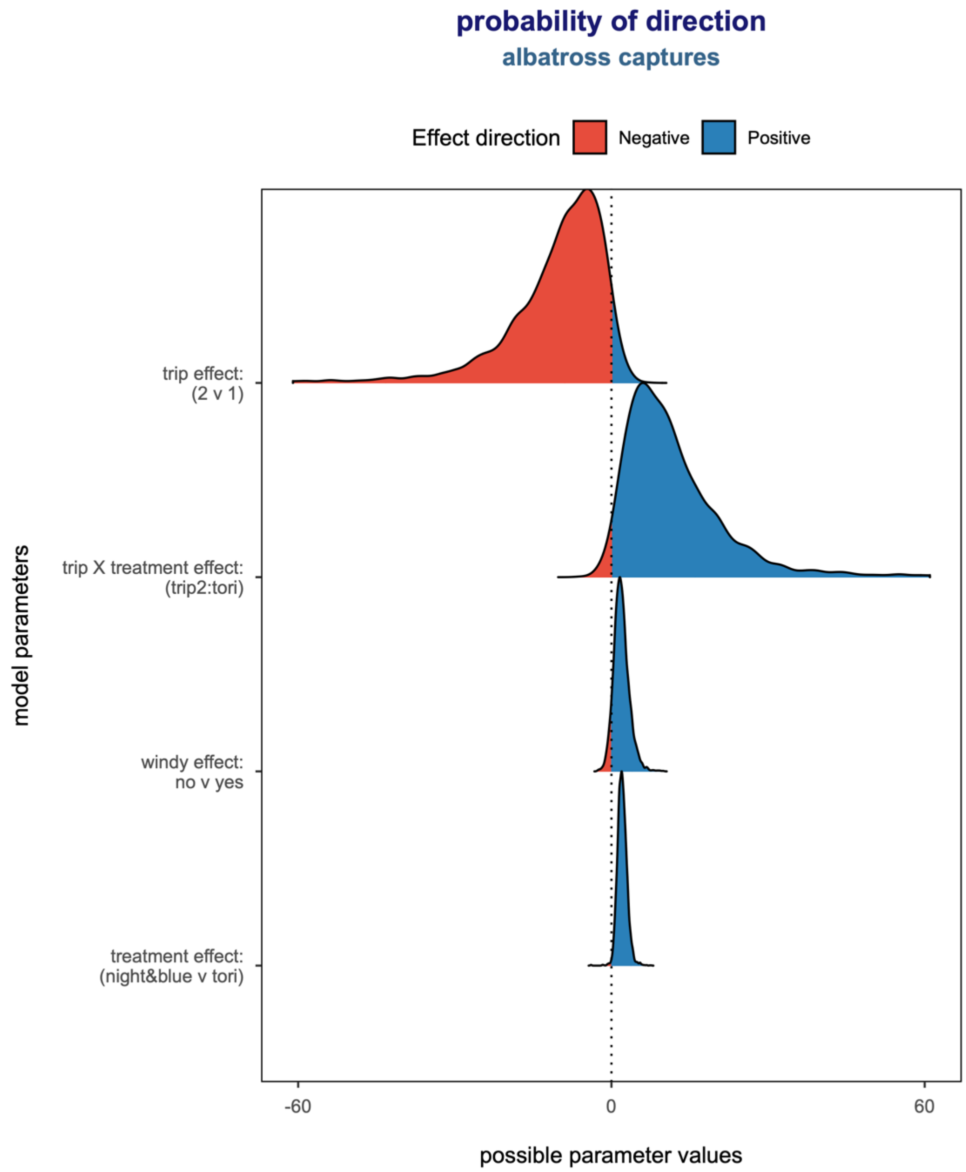

Figure 3 GLMM PARAMETER SUMMARY. Probability of direction plot for selected parameters estimated from the best-fit GLMM with ZINB likelihood for the capture rate of albatrosses on the shallow-set longline gear. Polygons show the density summary of the posterior draws and coloured given the estimated direction (positive or negative) of the effect or parameter. The proportion of the polygon that does not include zero is a statement about the probability of the proposed direction of the effect.

There was a significant difference in albatross capture rates between the 2 trips (see Figure 2 for trip specific geolocations) with capture rates being lower during the first trip (Figure 3: negative direction and we can be > 95% sure of that effect). More importantly, Figure 3 shows that the treatment effect for tori-lines (partial daytime sets) had a > 0.94 probability of being positive (and so higher catch) compared to the treatment effect for night-setting (with blue-dyed fish bait) and this is also reflected in the higher capture rates for the interaction term between

Trip & Treatment where captures rates were much higher for tori-lines during the second trip (see Figure S4d) — and we can be ca 98% sure of that effect (Figure 3). And we can also be >99% sure that higher albatross captures occurred during sets exposed to windy conditions (≥ Beaufort 4).

The predicted GLMM-adjusted marginal treatment effect for the albatross capture rate averaged over the 2 trips and wind condition is summarised in Figure 4 for sets deployed with paired tori-lines or night-setting (with blue-dyed fish bait). It was estimated that the expected median capture rate for albatrosses on tori-line sets was ca 1.35 (95% HDI: 0.36-3.29) albatrosses per set compared to ca 0.04 (95% HDI: 0-0.18) albatrosses per set for the night-setting sets. The posterior marginal median for the ratio (TORI/NIGHT) of those 2 treatment specific effects is ca 37.6 (95% HDI: 4.7-945) suggesting that tori-lines were ca 38 times more likely to catch an albatross in this study than on sets deployed with night-setting and blue-dyed fish bait (Figure 4).

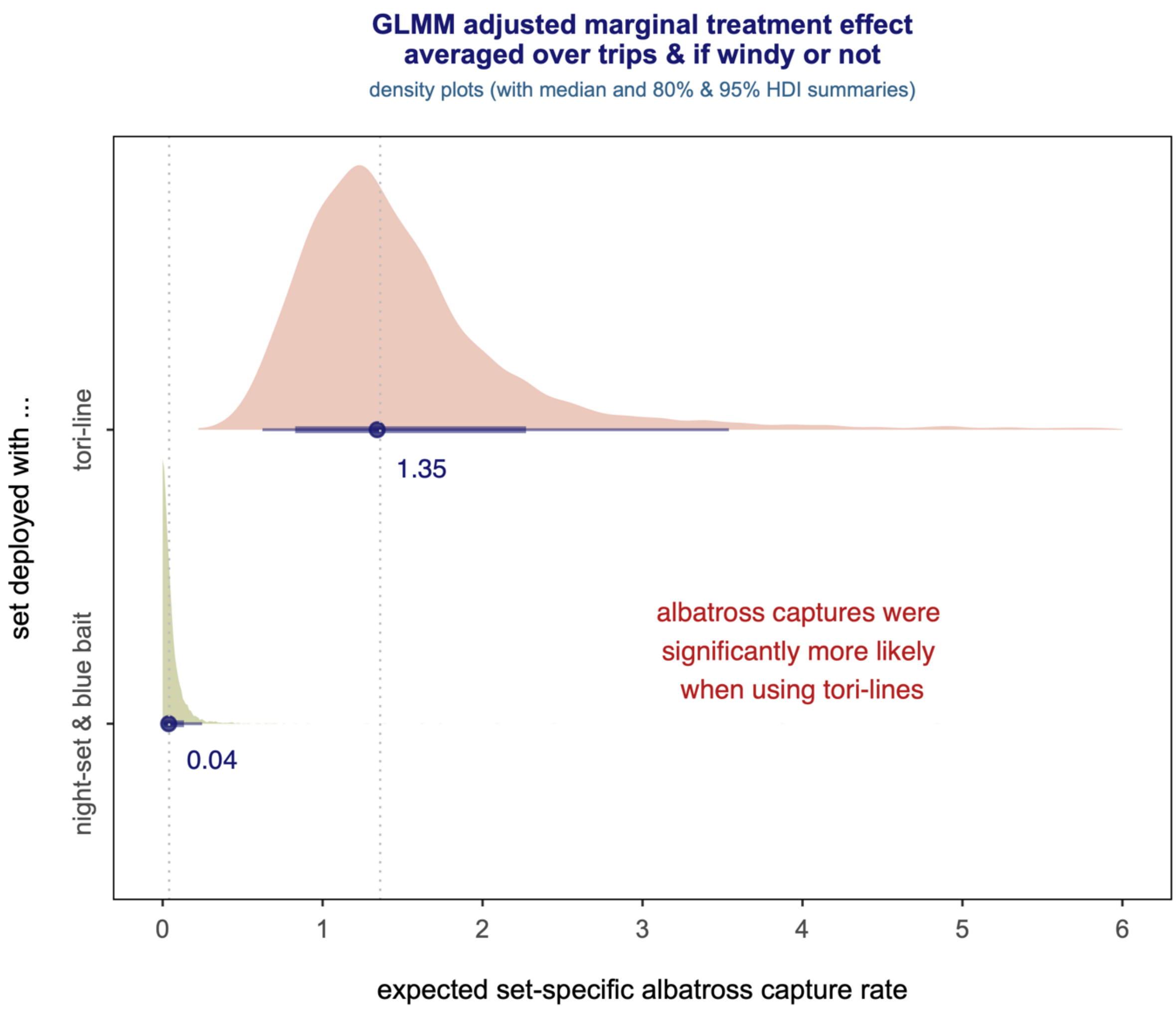

Figure 4 CAPTURES. Summary of the estimated marginal treatment-specific effect derived from the best-fit GLMM with ZINB likelihood for the capture rate of albatrosses on the shallow-

set longline gear. Coloured polygon shows the density distribution summary, solid dot (+ numeric label) = estimated marginal median of the density polygon, thick horizontal line below each polygon shows the 80% highest posterior density interval for the density polygon while the thin horizontal line is the 95% HDI.

The average marginal effect difference in terms of number of albatrosses per set is shown in Figure 5, which reflects the finding shown in Figure 4. Overall, it is quite evident for this experiment that albatross capture rates were substantially higher on the paired tori-line sets than on those sets deployed as night-setting (with blue-dyed fish bait).

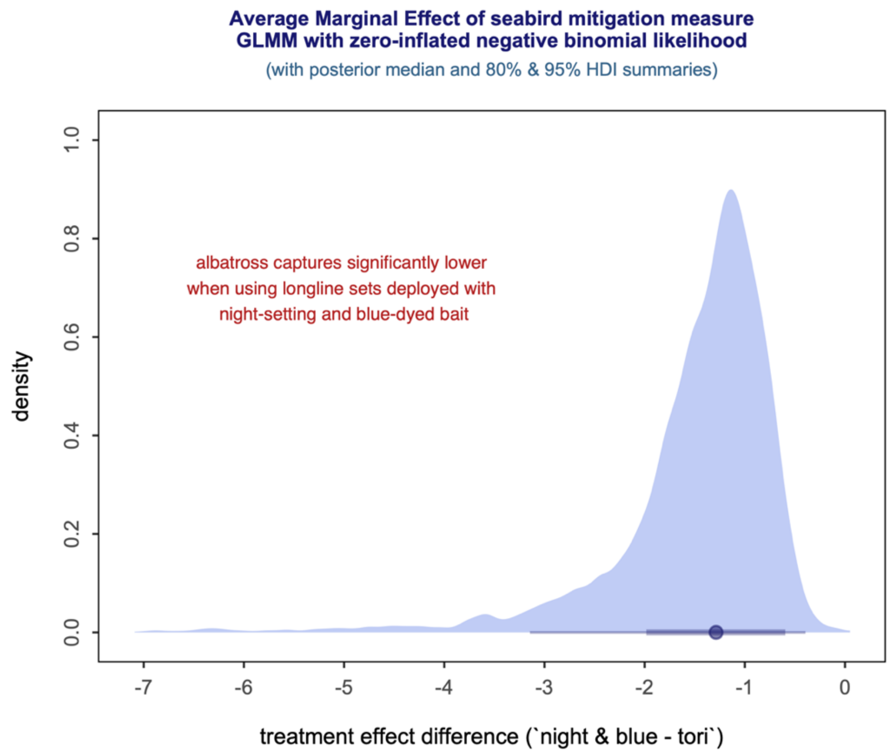

Figure 5 AVERAGE MARGINAL EFFECT (CAPTURES). Summary of the estimated average marginal effect of the difference in albatross capture rates between the 2 treatments derived from the best-fit model (see estimated treatment-specific effects shown in Figure 4). Coloured polygon shows the density distribution summary, solid dot = estimated average treatment effect difference of the density polygon, thick horizontal line below each polygon shows the 80% highest posterior density interval for the density polygon while the thin horizontal line is the 95% HDI.

3.3. Modelling the Albatross contact and Attempt Rates

GLMMs with Bernoulli likelihood were fit to the albatross interaction data (ATTEMPT = ≥ 1 albatross attempt to attack the gear, CONTACT = ≥ 1 albatross contact with a baited hook) with a single predictor (Treatment) and sets as the random effect. There were too few records to include more potential predictors. Bernoulli GLMMs fitted the binary data well for both measures (attempt, contact) as evidenced by prior sensitivity diagnostics, residual diagnostics and posterior predictive checks.

Overall, the binary response classification rate or prediction accuracy metric for the `attempt` GLMM was 84% (95% CI: 69% - 94%) and 77% (95% CI: 61% - 89%) for the `contact` GLMM. We can > 99% sure that paired tori-line sets had higher albatross interaction rates than night-setting for both the `attempt` and the `contact' metrics.

We also used two different priors on the fixed parameters in the model (weakly informative, R2D2M2) but LOOcv and Bayesian stacking metrics showed that predictive accuracy was higher for both attempt and contact data using weakly informative priors rather than a R2D2M2 prior — nonetheless, we show summary treatment-specific estimates based on both types of prior. The predicted GLMM-estimated marginal treatment-specific effect for both the albatross attempt and contact probabilities are summarized in Figure 6.

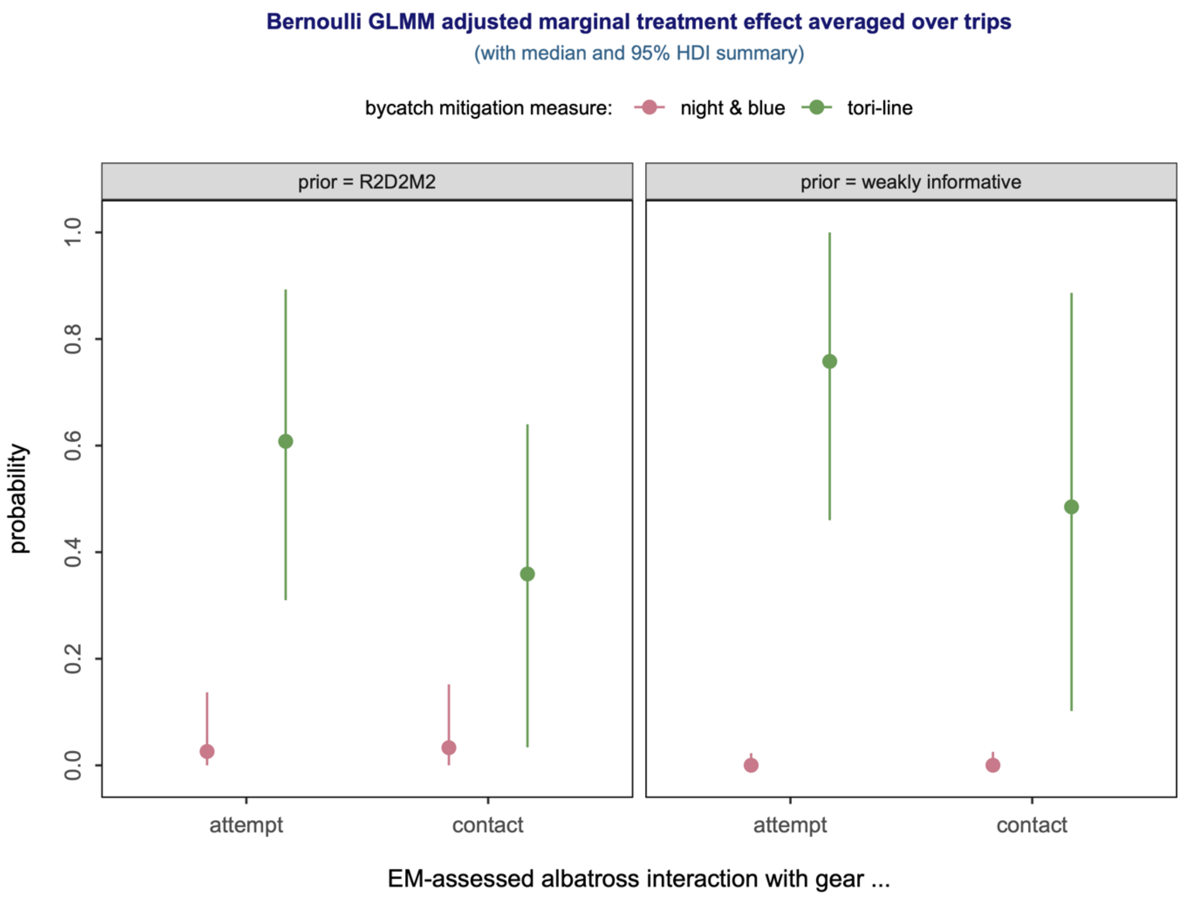

Figure 6 ATTEMPT & CONTACT RATES. (**Left Panel**) showing the treatment-specific marginal effect for the two metrics: (1) at least 1 `attempt` and (2) at least 1 `contact` with the gear/bait derived from a GLMM with Bernoulli likelihood using a R2D2M2 prior on the treatment parameter for the interaction rate of albatrosses with the shallow-set longline gear. (**Right Panel**) shows the same as in the (left panel) but with the Bernoulli GLMM fitted instead with a

weakly informative prior, which was the best-fit model. solid dot = estimated median and vertical bar = 95% credible interval.

The best-fit model-based estimates for both metrics using weakly informative priors are shown in the right-side panel in Figure 6 — the posterior marginal median probability of at least 1 albatross *attempt* at a baited hook was: ca 0.76 (95% HDI: 0.46-1.00) while the probability of at least 1 albatross *contact* at a baited hook was: ca 0.49 (95% HDI: 0.10-0.89). Overall, it is quite evident for this experiment that albatross interaction probabilities were substantially higher on the paired tori-line sets than on those sets deployed with night-setting.

3.4. Time-of-day of Seabird Captures

Of the 22 albatross captures that occurred during tori-line (partial-day) sets, 21 were estimated to have been captured prior to both nautical dusk and 1 hour after sunset. Thus, of the 23 total albatross captures that occurred during setting, 21 were captured during the daytime (prior to nautical dusk and prior to 1 hour after sunset) during sets with deployed paired tori-lines and 2 were captured at night (after nautical dusk and after 1 hour after sunset), one in a tori-line set and one in a night-set (with blue-dyed fish bait).

4. CONCLUSION

So, did tori-lines work as an effective seabird bycatch mitigation measure in this trial compared to the use of night-setting with blue-dyed fish bait? No, they did not. Specifically:

- albatrosses were ca 37.6 times (95% HDI: 4.7 - 945) more likely to be **captured** when paired tori-lines were deployed compared to the night-sets (with blue-dyed fish bait)

- albatrosses were > 1000 times (95% HDI: 1 - 52e9) more likely to **contact** a baited hook when paired tori-lines were deployed compared to the night-sets (with blue-dyed fish bait)

- albatrosses were > 3000 times (95% HDI: 1.9 - 23e9) more likely to **attempt** to contact a baited hook when paired tori-lines were deployed compared to the night-sets (with blue-dyed fish bait)

While deployment of tori-lines was shown to reduce seabird bycatch risk for the Hawaii-based deep-set longline fishery (Gilman et al 2021, Chaloupka et al 2021), this study demonstrated unequivocally that it is not as effective a seabird bycatch mitigation method as night-setting with blue-dyed fish bait in the Hawaii-based shallow-set longline fishery. The results may be in part attributed to the tori-line design used in this study (*see* Supplementary Materials). The study utilized a light-weight short-streamer design that was found to be effective in the Hawaii-based deep-set longline fishery, with the length adjusted to account for the greater aerial coverage needed due to the slower gear sink rate in the shallow-set longline fishery. However, the Hawaii-based shallow-set fishery operates at higher latitudes than the deep-set tuna longline fishery and is generally subject to higher wind conditions, which we found can increase seabird bycatch risk (Figure S4c) — perhaps because seabirds are more agile and able to access baited hooks during stronger wind conditions, or because tori-lines are less likely to be maintained adequately over the set-deployment footprint where baited hooks are vulnerable to scavenging seabirds.

Results also suggest that the main albatross species exposed to this fishery are primarily daytime foragers (Fernandez & Anderson 2000, Weimerskirch & Guionnet 2002, Kappes et al 2015) and might be predominantly more at risk of capture during crepuscular daylight hours, and that night-setting reduces but does not eliminate capture risk as these exposed albatross species also forage at night such as during periods of high lunar illumination (Conners et al 2015).

The findings illustrate the problem with a smörgåsbord or menu approach to regional seabird bycatch management. Some seabird management measures adopted by regional fisheries management organizations allow parties to select amongst a menu of options of seabird bycatch mitigation methods, in some cases with two lists of options (IATTC 2011, ICCAT 2011, WCPFC 2018, SIOFA 2022, IOTC 2023). Our findings suggest that not all options within a category have equivalent efficacies in all fisheries for paired tori-lines and night-setting with blue-dyed fish bait — at least we found that to be the case when using light-weight short-streamer paired tori-lines in the Hawaii-based shallow-set longline fishery.

5. REFERENCES

ACAP (2023) ACAP Review of Mitigation Measures and Best Practice Advice for Reducing the Impact of Pelagic Longline Fisheries on Seabirds. Agreement on the Conservation of Albatrosses and Petrels, Hobart, Australia

Aguilar J, Bürkner P (2023) Intuitive joint priors for Bayesian linear multilevel models: the R2D2M2 prior. Electronic Journal of Statistics 17: 1711–1767

Aragon T (2020) epitools: Epidemiology tools. R package v0.5-10.1. http://CRAN.R-project.org/ package=epitools

Bakker V et al (2018) The albatross of assessing and managing risk for long-lived pelagic seabirds. Biol. Conserv. 217, 83–95

Bergh M, Pikitch E, Skalski J, Wallace J (1990) The statistical design of comparative fishing experiments. Fisheries Research 9: 143-163

Brothers N, Cooper J, Lokkeborg S (1999) The Incidental Catch of Seabirds by Longline Fisheries: Worldwide Review and Technical Guidelines for Mitigation. FAO Fisheries Circular No. 937. Food and Agriculture Organization of the United Nations, Rome

Bürkner P (2017) brms: An R Package for Bayesian multilevel models using Stan. Journal of Statistical Software 81: 1-28

Carpenter B, Gelman A, Hoffman M, Lee D, Goodrich B, Betancourt M, Brubaker M, Guo J, Li P, Riddell A (2017) Stan: a probabilistic programming language. Journal of Statistical Software 76(1): 1-32

Chaloupka M, Gilman E, Carnes M, Ishizaki A, Brady C, Swimmer Y, Wang J, Ellgen S, Kingma E (2021) Could tori lines replace blue-dyed bait to reduce seabird bycatch risk in the Hawaii deep-set longline fishery? Western Pacific Regional Fishery Management Council. Honolulu, Hawaii. pp 36

Congdon P (2003) Applied Bayesian Modelling. Wiley and Sons Ltd, UK

Conners M, Hazen E, Costa D, Shaffer S (2015) Shadowed by scale: subtle behavioral niche partitioning in two sympatric, tropical breeding albatross species. Movement Ecology 3: 28

Davies T, Jonsen I (2011) Identifying nonproportionality of fishery-independent survey data to estimate population trends and assess recovery potential for cusk (*Brosme brosme*). Canadian Journal of Fisheries and Aquatic Sciences 68: 413-425

Dias M, Martin R, Pearmain E, Burfield I, Small C, Phillips R, Yates O, Lascelles B, Borboroglu P, Crozall J (2019) Threats to seabirds: A global assessment. Biological Conservation. 237: 525–537

Fahrmeir L, Lang S (2001) Bayesian inference for generalised additive mixed models based on Markov random field priors. Applied Statistics 50: 201-220.

Fernández P, Anderson D (2000). Nocturnal and diurnal foraging activity of Hawaiian albatrosses detected with a new immersion monitor. Condor 102: 577–584

Gabry J, Mahr T (2024) bayesplot: Plotting for Bayesian Models. R package version 1.11.1. https://mc-stan.org/bayesplot

Gabry J, Simpson D, Vehtari A, Betancourt M, Gelman A (2019) Visualization in Bayesian workflow. Journal of the Royal Society Series A 182: 1-14

Gilman E, Chaloupka M, Ishizaki A, Carnes M, Naholowaa H, Brady C, Ellgen S, Kingma E (2021) Tori lines mitigate seabird bycatch in a pelagic longline fishery. Reviews in Fish Biology and Fisheries 31: 653–666

Gilman E, Chaloupka M, Pacini A, Kingma E (2024) Exploring odontocete depredation rates in a pelagic longline fishery. PLoS ONE 19(3): e0301072

Gilman E, Chaloupka M, Wiedoff B, Willson J (2014) Mitigating seabird bycatch during hauling by pelagic longline vessels. PLoS ONE 9(1): e84499

Gilman E, Evans T, Pollard I, Chaloupka M (2023) Adjusting time-of-day and depth of fishing provides an economically viable solution to seabird bycatch in an albacore tuna longline fishery. Scientific Reports 13: 2621

Gilman E, de Ramón Castejón V, Loganimoce E, Chaloupka M (2020) Capability of a pilot fisheries electronic monitoring system to meet scientific and compliance monitoring objectives. Marine Policy 113: 103792

Gilman E, Passfield K, Nakamura K (2014) Performance of regional fisheries management organizations: ecosystem-based governance of bycatch and discards. Fish and Fisheries 15: 327-351

Greenland S (2004) Model-based estimation of relative risks and other epidemiologic measures in studies of common outcomes and in case-control studies. American Journal of Epidemiology 160: 301–305
Grolemund G, Wickham H (2011) Dates and times made easy with lubridate. Journal of Statistical Software 40(3): 1-25
Hartig F (2022) DHARMa: Residual diagnostics for hierarchical (multi-level/mixed) regression models. R package version 0.4.6. https://CRAN.R-project.org/package=DHARMa
IATTC (2011) Resolution to Mitigate the Impact on Seabirds of Fishing for Species Covered by the IATTC. Resolution C-11-02. Inter-American Tropical Tuna Commission, La Jolla, USA
ICCAT (2011) Supplemental Recommendation by ICCAT on Reducing Incidental By-catch of Seabirds in ICCAT Longline Fisheries. Recommendation 11-09. International Commission for the Conservation of Atlantic Tunas, Madrid
IOTC (2023) On Reducing the Incidental Bycatch of Seabirds in Longline Fisheries. Resolution 23/07. Indian Ocean Tuna Commission, Mahe, Seychelles
ISSF (2024) Inputs to Multispecies Bycatch Management Strategy Evaluation. By Gilman E, Murua H, Chaloupka M. ISSF Technical Report 2024-04. International Seafood Sustainability Foundation, Pittsburgh, PA, USA
Jiménez S, Domingo A, Winker H, Parker D, Gianucad D, Neves T, Coelho R, Kerwath S (2020) Towards mitigation of seabird bycatch: large-scale effectiveness of night setting and tori lines across multiple pelagic longline fleets. Biological Conservation 247: 108642
Kallioinen N, Paananen T, Bürkner P, Vehtari A (2024) Detecting and diagnosing prior and likelihood sensitivity with power-scaling. Statistics and Computing 34: 57
Kammann E, Wand M (2003) Geoadditive models. Applied Statistics 52: 1-18
Kappes M, Shaffer S, Tremblay Y, Foley D, Palacios D, Robinson P, Bograd S, Costa D (2010). Hawaiian albatrosses track interannual variability of marine habitats in the North Pacific. Progress in Oceanography 86: 246–260
Kay M (2023) tidybayes: Tidy data and geoms for Bayesian models. R package version 3.0.6. http://mjskay.github.io/tidybayes. https://doi.org/10.5281/zenodo.1308151
Kay M (2024) ggdist: Visualizations of distributions and uncertainty in the Grammar of Graphics. IEEE Transactions on Visualization and Computer Graphics 30: 414-424
Kleiber C, Zeileis A (2016) Visualizing count data regressions using rootograms. The American Statistician 70: 296–303
Kuhn M (2008) Building predictive models in R using the caret package. Journal of Statistical Software 28(5): 1–26
Kruschke J, Liddell T (2018) The Bayesian New Statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. Psychonomic Bulletin & Review 25: 178–206
Lemoine N (2019) Moving beyond noninformative priors: why and how to choose weakly informative priors in Bayesian analyses. Oikos 128: 912-928
Lenth R (2016) Least-squares means: the R package lsmeans. Journal of Statistical Software 69:1–33
Lenth R (2020) emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.5.2-1. https://CRAN.R-project.org/package=emmeans
Makowski D, Ben-Shachar M, Lüdecke D (2019) bayestestR: Describing effects and their uncertainty, existence and significance within the Bayesian framework. Journal of Open Source Software 4(40): 1541. https://doi.org/10.21105/joss.01541
Melvin E, Guy T, Read L (2013) Reducing seabird bycatch in the South African joint venture tuna fishery using bird-scaring lines, branch line weighting and nighttime setting of hooks. Fisheries Research 147:72–82.
Melvin E, Dietrich K, Suryan R, Fitzgerald S (2019) Lessons from seabird conservation in Alaskan longline fisheries. Conservation Biology 33: 842-852
Naimi A, Whitcomb B (2020) Estimating risk ratios and risk differences using regression, American Journal of Epidemiology 189: 508–510
NMFS (2023) Experimental Fishing Permit. Letter from Sarah Malloy, Acting Regional Administrator, Pacific Islands Regional Office, National Marine Fisheries Service, to Eric Kingma, Executive Director,

Hawaii Longline Association, September 20, 2023. Pacific Islands Regional Office, National Marine Fisheries Service, Honolulu.
Pedersen T (2020) patchwork: The composer of plots. R package version 1.1.0. https://CRAN.R-project.org/package=patchwork
Phillips R, Gales R, Baker G, Double M, Favero M, Quintana F, Tasker M, Weimerskirch H, Uhart M, Wolfaardt A (2016) The conservation status and priorities for albatrosses and large petrels. Biological Conservation 201: 169–183
Rodríguez-Sánchez F (2024) DHARMa.helpers: Helper functions to check models not (yet) directly supported by DHARMa R package version 0.0.1. https://github.com/Pakillo/DHARMa. helpers
Searle S, Speed F, Milliken G (1980) Population marginal means in the linear model: An alternative to least squares means. The American Statistician 34: 216-221
SIOFA (2022) Conservation and Management Measure on Mitigation of Seabirds Bycatch in Demersal and Pelagic Longlines and other Demersal Fishing Gears Fisheries (Mitigation of Seabirds Bycatch). CMM 2022/13. Southern Indian Ocean Fisheries Agreement, Saint-Denis, La Reunion
Thieurmel B, Elmarhraoui A (2022) suncalc: Compute sun position, sunlight phases, moon position and lunar phase. R package version 0.5.1. https://CRAN.R-project.org/package=suncalc
Tuyl F, Gerlach R, Mengersen K. Comparison of Bayes-Laplace, Jeffreys, and other priors: The case of zero events. Am Stat. 2008; 62: 40–44.
van de Schoot R, Depaoli S, King R, Kramer B, Märtens K, Tadesse M, Vannucci M, Gelman A, Veen D, Willemsen J, Yau C (2021) Bayesian statistics and modelling. Nature Reviews Methods Primers 1(1): 1-26. https://doi.org/10.1038/s43586-021-00017-2
van der Bles A, van der Linden S, Freeman A, Mitchell J, Galvao A, Zaval L, Spiegelhalter D (2019) Communicating uncertainty about facts, numbers and science. Royal Society Open Science 6: 181870
Vehtari A, Gelman A, Gabry J (2017) Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Statistics and Computing 27: 1413–1432
Vehtari A, Gelman A, Simpson D, Carpenter B, Bürkner P (2021) Rank-normalization, folding, and localization: an improved Rhat for assessing convergence of MCMC (with Discussion). Bayesian Analysis 16: 667–718
WCPFC (2018) Conservation and Management Measure to Mitigate the Impact of Fishing for Highly Migratory Fish Stocks on Seabirds. CMM 2018-03. Western and Central Pacific Fisheries Commission, Kolonia, Federated States of Micronesia
Weimerskirch H, Guionnet T (2002) Comparative activity pattern during foraging of four albatross species. Ibis 144: 0–50
Wickham H (2016) ggplot2: Elegant Graphics for Data Analysis. 2nd Edition. Springer-Verlag, New York, USA
Wood S (2006) Generalized Additive Models: an introduction with R. Chapman and Hall/CRC, Boca Raton.
Yao Y, Vehtari A, Simpson D, Gelman A (2018) Using stacking to average Bayesian predictive distributions (with Discussion). Bayesian Analysis 13: 917–1003
Zeileis A, Fisher J, Hornik K, Ihaka R, McWhite C, Murrell P, Stauffer R, Wilke C (2020) colorspace: A toolbox for manipulating and assessing colors and palettes. Journal of Statistical Software 96: 1-4

SUPPLEMENTARY MATERIAL

SM1. Tori-line design

Two 65 m-long, red, 3 mm diameter, 12-strand, single-braid Dyneema (AmSteel AS-78 Dyneema, thermoplastic polyethylene) ropes will be used for the aerial sections of the tori-lines. Two 50 cm-long streamers will be attached every 1 m along each of the aerial sections, with the first streamer on each aerial section attached at 2.5 m from the point of attachment to the tori poles on the vessel deck, and the last streamers attached at 62.5 m, for a total of 45 streamers on each aerial section. The aerial sections were attached to the rail on top of the bait shed. The height from the water surface to tori-line attachment point on the bait shed roof was 5 meters. Streamers on each aerial section were 50 cm-long, 5 cm-wide, 6 mm thick, clear polyethylene sheeting dyed fluorescent orange. One 100 cm-long strip was spliced through the Dyneema rope on each aerial section to create the two 50 cm-long streamers. The drag sections of the tori-line was 100 m-long, and made of 6 mm diameter, 'Blue Steel' 12-strand polyolefin fiber rope. A weak link installed on both tori-line poles, using monofilament nylon (polyamide), was in between the aerial sections of the tori-line and the tori pole. Safety lines were incorporated to retain the tori-line if the weak link breaks. When a vessel sets at 11.1 km/hr (6 knots), this would produce ca. 12.5 kg of drag, resulting in a100 m-long aerial section, which should be sufficient to maintain the aerial section up off the ocean surface. Each of the aerial and drag sections was connected by splicing them onto a 6 mm stainless steel swivel, was then covered first with heat shrink wrap and then with electrical tape to prevent entanglement with the fishing gear.

SM2. EM System

The two-camera electronic monitoring system and onboard server recorded video. In addition to video, the onboard system GPS recorded the vessel position every thirty seconds, which informs the vessel direction and speed. The wireless, satellite enabled video transfer EM system from Integrated Monitoring allowed for system health checks during fishing, and seamless data transfer from the vessel to the video storage cloud. Review was completed using Integrated Monitoring's web portal review software, *Monitor* by AIS Inc reviewers to collected the required data. A custom data entry template was developed to capture data fields. See Gilman et al (2021) for the EM analyst data recording protocols. The only modification was that seabird scan counts were made by the EM analyst within an area 100 m astern of the vessel. The analyst conducted scan counts at the start of the set when the first hook is deployed and every hour thereafter until the last hook was retrieved.

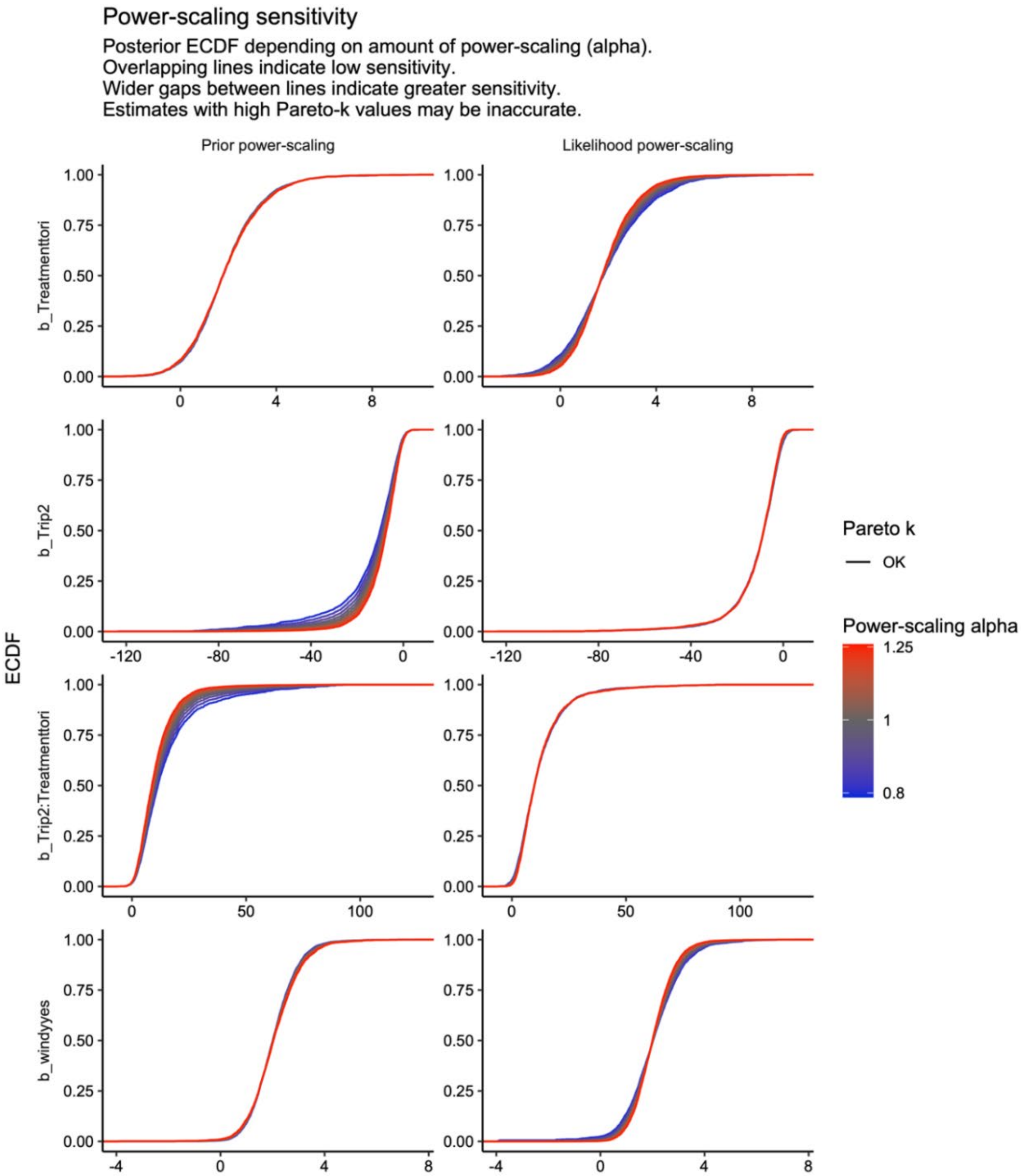

Figure S1 MODEL SENSITIVITY DIAGNOSTICS. Shows the sensitivity of the posterior derived from the best-fit albatross capture rate model to power-scaled perturbations of the priors (**left side panels**) and the likelihood (**right side panels**) for each model parameter (treatment, trip, treatment:trip interaction, wind category) — cumulative Jensen-Shannon distance divergence measure was used for assessment. Model inference is based on a posterior derived from the model priors and model likelihood (zero-inflated negative binomial) so that assessing the posterior sensitivity to perturbations to the parameter-specific priors and likelihood is an important part of a Bayesian modelling workflow. Importantly, there was no prior-data conflict or likelihood non-informativity for the treatment effect derived from this ZINB GLMM (see specifically the top row of panels). There was some prior-data conflict apparent for the trip and trip:treatment parameters.

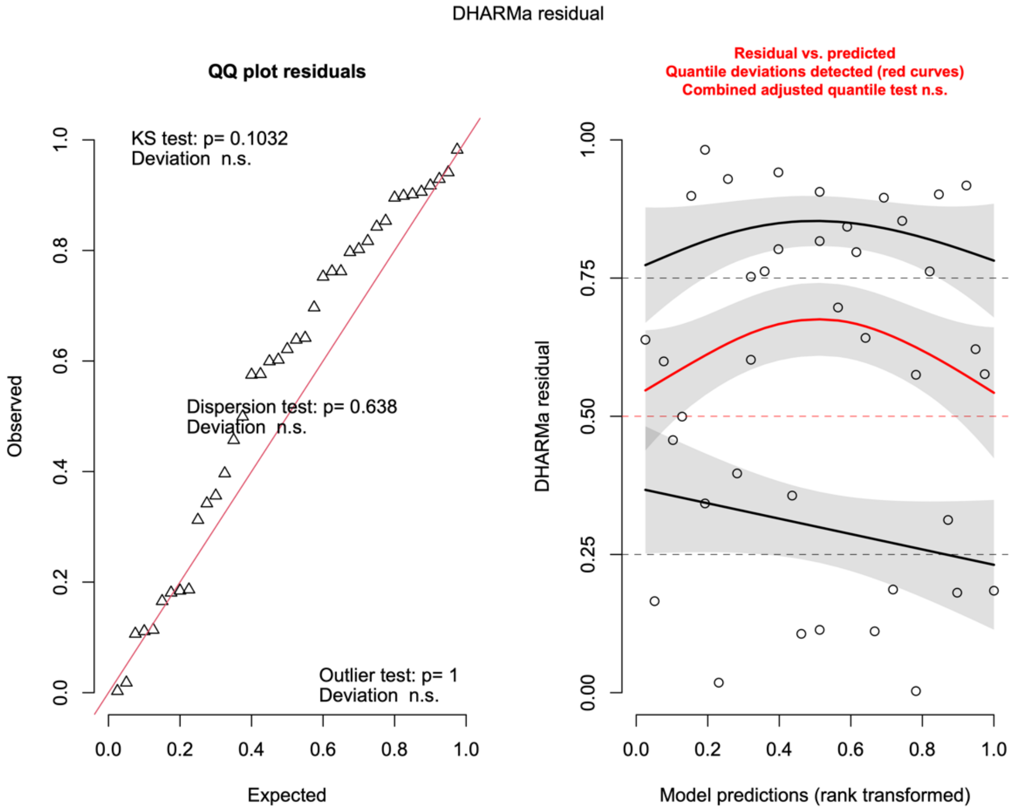

Figure S2 RESIDUAL DIAGNOSTICS. Simulation-based assessment of posterior predictive draws sourced from the best-fit albatross capture rate model (GLMM with ZINB likelihood). (**Panel A**) shows the QQ-plot to detect overall deviations from the expected distribution of the standardised and scaled residuals with all non-significant tests for expected residual distribution, dispersion and outlier test functions. (**Panel B**) shows residuals plotted against predicted values with 3 superimposed quantile regression curves shown for the 0.25, 0.50 and 0.75 quantiles, with any significant deviation for those curves shown in red — the median residuals (the 0.5 quantile curve) do show a deviation from the expected residual pattern but no such deviation apparent for either the lower or upper quantile curves. Given the sparse data set here, we consider this model is stable and adequately supported by these residual diagnostic checks — moreover, the posterior predictive checks in Figure S3 show the model more than adequate accounted for this sparse data set.

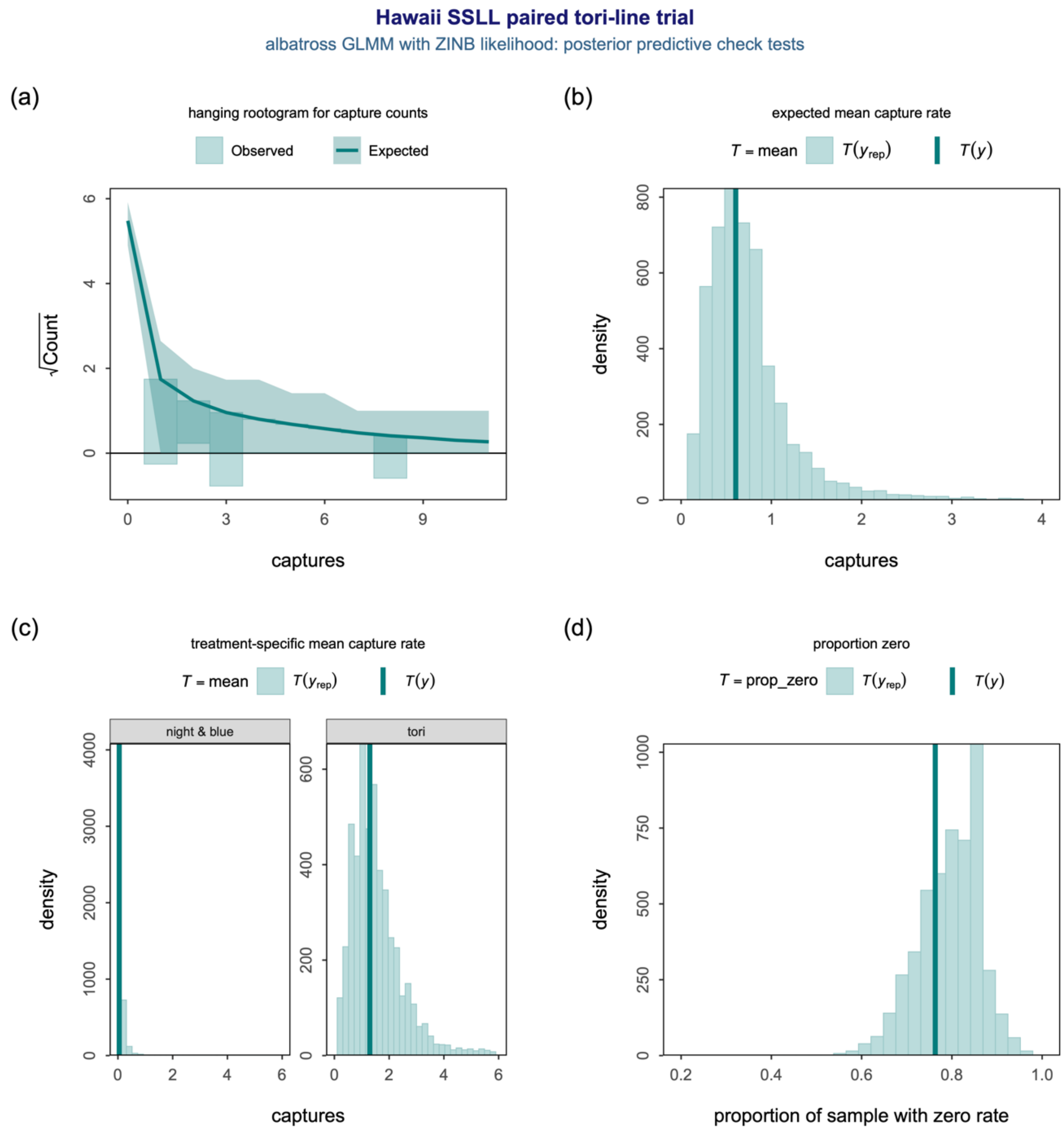

Figure S3 MODEL PREDICTIVE ACURACY. Posterior predictive check tests for 1000 draws from the best-fit albatross capture rate GLMM with ZINB likelihood. (Panel A) hanging rootogram showing the model recovers the distributional form adequately for such a sparse albatross capture data set with small deviations from the observed capture frequencies. (Panel B) shows check for the mean observed rate (solid vertical line) and the histogram of the expected rates (T(y)). (Panel C) shows the same check as for (B) now conditional on whether tori-lines were deployed or not. (Panel D) shows the observed (solid line) and expected proportion of zeroes in the data.

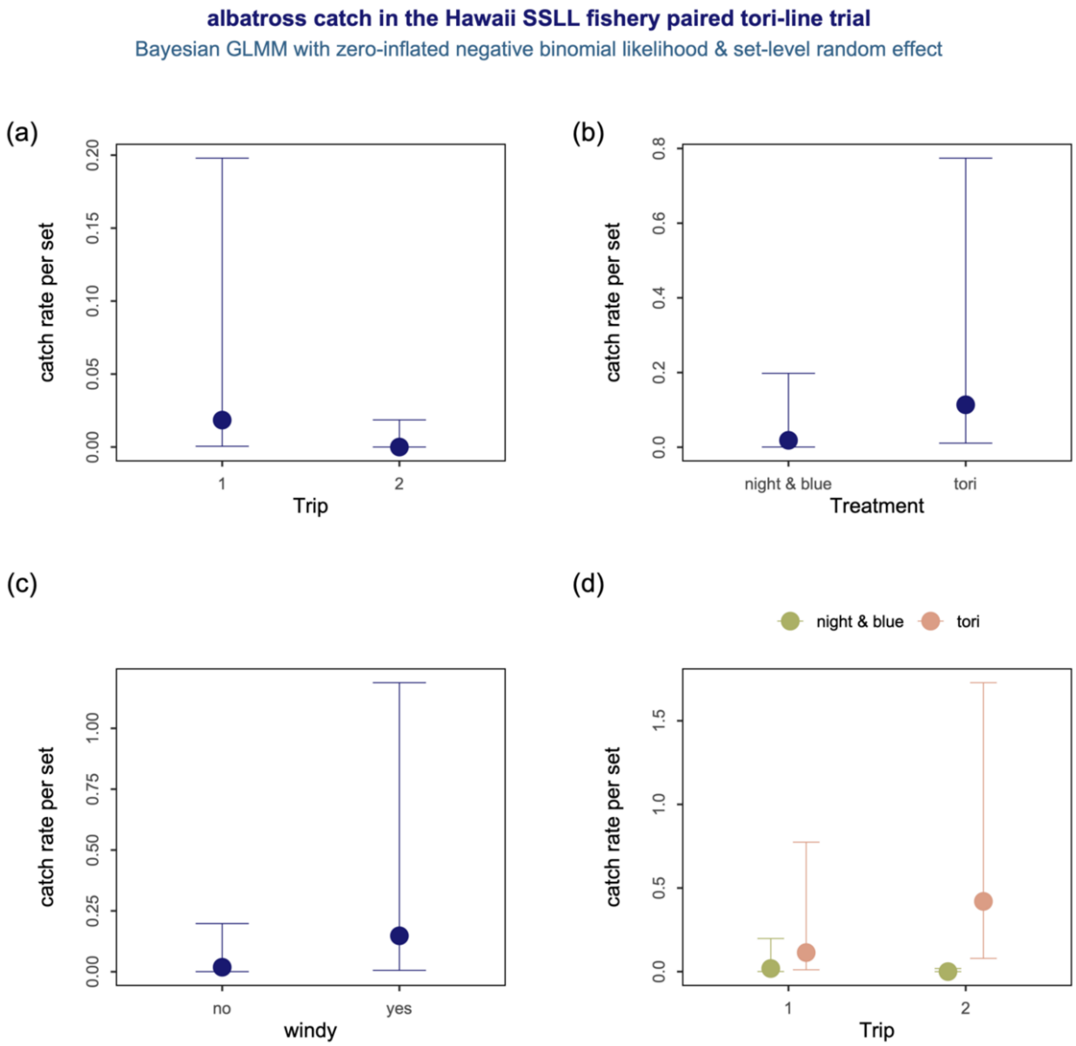

Figure S4 ZINB GLMM SUMMARY (CAPTURES). Summary of the estimated conditional albatross set-specific capture rates given fishing trip, treatment (bycatch mitigation measure), windiness category (no = < Beaufort 4 or yes = ≥ Beaufort 4), and the interaction term between trip and treatment: (**a**) shows the estimated conditional trip-specific effect, (**b**) shows the conditional treatment-specific effect, (**c**) shows the conditional effect of wind category on expected capture rates and (**d**) shows the conditional trip-dependent treatment-specific effects. solid dot = estimated median and vertical bar = 95% credible interval.