UiT The Arctic University of Norway

Faculty of Biosciences, Fisheries and Economics, Department of Arctic Marine Biology

Assessing the deterrence effect of target-specific acoustic startle technology on killer whales and humpback whales during interactions with Norwegian purse seine herring fishery

Elida Langstein

Master’s thesis in Biology BIO-3950, May 2023

Assessing the deterrence effect of target-specific acoustic startle technology on killer whales and humpback whales during interactions with Norwegian purse seine herring fishery

Elida Langstein

Master of Science in Biology – Marine Ecology and Resource Biology

May 2023

Supervisors:

Audun Rikardsen: UiT The Arctic University of Norway

Deanna Leonard: Institute of Marine Research

Thomas Götz: University of St Andrews

Martin Biuw: Institute of Marine Research

Emma Vogel: UiT The Arctic University of Norway

Cover photo: Killer whale (*Orcinus orca)* feeding at herring (*Clupea harengus)* which has aggregated around a purse seine fishing net. Photo: Audun Rikardsen

Content

1 Abstract 1
2 Introduction 2
3 Method 7
3.1 Study area 7
3.2 The fishing operations 8
3.3 Study design and protocol 9
3.4 Acoustic system 11
3.5 Audio analysis 12
3.6 Data analysis 12
4 Results 15
4.1 Killer whales and zones 15
4.2 Humpback whales and zones 21
4.3 Killer whales and distances 24
5 Discussion 27
5.1 Deterrence effect 28
5.2 Recovery and attenuation 30
5.3 Alternative mitigation methods 32
5.4 Study setup: strengths and limitations 33
5.5 Conclusion 35
5.6 Future research and implementation 36
6 Works cited 36
Appendix 1 – Killer whales and zones 43
A - Model selection 43
B - Model specifics 45
C – Model diagnostics 48
Appendix 2 – Humpback whales and zones 50
A – Model selection 50
B – Model specifics 53
C – Model diagnostics 55
Appendix 3 – Killer whales and distances 58
A – Model selection 58
B Model specifics 61
C Model diagnostics 62

List of figures and tables

Figure 1: Inset map of the Kvænangen fjord system. The top left map shows where in Norway the study was conducted marked with the red square, while the bottom right map is zoomed in to the study area. The experiment sites are marked as red dots. Coordinates given in UTM (zone 35), scale 1:200 000. 8

Figure 2: Illustration of the experimental set up in the field. Vestbris is given as an example of fishing vessel. The playback vessel was placed 90 degrees from the bow of the fishing vessel. The size of zones A, B, C and D were determined based on the size of the fishing vessel. The diagonal arrow represents the mean distance from the playback vessel to the furthest edge of the observation area for the zones. Sector size is the length of each zone. ……..11

Table 1: Overview table of the experiment trials that were conducted and used in the analysis, divided into 3 different datasets: zones and killer whales, distances and killer whales and zones and humpback whales.......... 15

Table 2: Model summary for the zone data on killer whales, from the generalized linear mixed models A) m13a (given as m13) and B) m14a (given as m14). Both models examine changes in the number of surfacings as a function of treatment (pre, playback, post) and zones (A,B,C,D). To the left, model M13 is with interaction between zones and treatment, while m14 to the right is without this interaction term and look at a more global effect. Significant values are highlighted in bold.…. 17

Figure 3: Boxplot of observed number of killer whale surfacings in the different zones A, B, C and D during the different treatment periods before, during and after playback. The X-axis show the different zones A, B, C and D, while the y-axis represents the number of surfacings. The different colors represent the different treatment periods before(red), during(green) and post playback(blue). Boxes show the medians and the interquartile ranges (IQR), where the lowest whisker is the smallest observation over the 25% quartile- 1.5x IQR , and the upper whisker is the largest observation below the 75% quartile + 1.5 x IQR. The dots represent the outliers, however 3 outliers (counts over 30 in before playback in zone A and in before and after playback in zone B) do not show in this figure as the y-axis was trimmed for aesthetic purposes. Significance values are derived from the model results, where ***<0.001 (p-value). 18

Table 3: The summary of pairwise comparisons (killer whales and zones) of the fixed effects “treatment” in the generalized linear mixed model m13a. Significant values are highlighted in bold. 19

Figure 4: Expected reduction of surfacings in percent across zones (killer whales) from before playback to during playback based on the model. The Y-axis show the percents, while the X-axis represents the different zones and all zones combined. The dots represent the expected value, while the error bars show the confidence intervals (95%CIs) with the respective borders. The colors represent the different zones A(yellow), B(blue), C(green), D(purple) and all zones pooled(red) 20

Figure 5: The expected surfacing counts (killer whales) based on model m13a (see Table. 1A). The x-axis shows the different treatment periods before, during and post playback, in which all are 5 minutes (total 15 minutes). The y-axis gives the mean expected number of surfacings. The colors represent the different zones A(red), B(blue), C(green) and D(purple). The dots represent the estimate, while the error bars represent the 95% confidence intervals. 21

Table 4: Summary of the generalized linear mixed models m13b and m14b for humpback whales and pooled zones. Random effects are the different trials. ∆AIC for m13 and m14= 3.017. 22

Figure 6: Number of humpback whale surfacings observed divided treatment (pre, playback, post) and on two zones: AC and BD. Red indicates the pre-playback period, while green represents the playback period and blue the post-playback period. 23

Figure 7: Expected surfacing counts based on the models m13b and m14b (humpback whales and pooled zones), between the different treatment periods (pre, playback, post) and pooled zones A(closest) and B(furthest). The experiment lasted for 15 minutes where each treatment type was 5 minutes long. The Y-axis show the expected mean surface count, while the X-axis show the different treatment periods. Red bars show zone A and blue bars show zone B. The dots are the expected values while the bars represent the 95% confidence interval.. 24

Figure 8: Observed surfacings of killer whales divided by different treatment periods (pre, playback, post), and two distance bins illustrated in a boxplot. The X-axis represents the two distance bins 0-50 m and 50-100 m. The X-axis show the expected mean of surfacing observations. Each bar represents different treatment periods, where red is before playback, green is during, and blue is post playback. 25

Table 5: Pairwise comparisons retrieved from model m1 (killer whales and distances). Surfacing counts between the treatment periods within each distance bin is compared. Significant values highlighted in bold...................... 26

Figure 9. Expected surfacings (killer whales) based on the GLMM m1 (table 3), between the treatment periods(pre, playback, post), and time bins 0-50 m and 50-100 m from playback vessel. The experiment lasted for 15 minutes where each treatment type was 5 minutes long. The Y-axis show the expected mean surface count, while the X-axis show the different treatment periods. Red bars show distance bin 0-50 m and blue bars show distance bin 50-100 m. The dots are the expected values while the bars represent the 95% confidence interval………27

Figure 10: Figure 10: Time series illustration of all observations on killer whales and zones data, divided into 1 minute time intervals (15 minutes total). Surfacing counts are shown on the y-axis while each minute is represented on the x-axis. Playback starts at 5 minues and end at 10 minutes (marked with the dotted line). This figure is presented for illustrative purposes. A command was used when coding this figure in R (geom_jitter from the package ggplot), which avoided overlapping of datapoints. No analysis is performed on the figure. The dashed line indicates the start and end of the playback session. Each zone is given one color, and the grey areas represent the 95% Cis……………………………………………………………………………………………...31

Acknowledgements

Working on this project has been a great learning process for me, and an adventure that I feel very lucky to have been a part of. I feel grateful to have had the chance to work with and be guided by so many talented and dedicated people. Special thanks to my supervisors who have supported me with knowledge and guidance along the way: Audun Rikardsen, Thomas Götz, Deanna Leonard, Martin Biuw and Emma Vogel. My deepest gratitude in particular to Thomas for your expertise and help with the statistical modeling, and Deanna for your excellent guidance in academic writing along the way.

Thanks to Vincent Janik, for your guidance both during the fieldwork and after.

A big thanks to Maria Tenningen, Lise Sivle and all the other people who contributed to the project.

Thank you to all the fishers who let us perform our experiments around your fishing vessels during fishing operations.

Thank you also to the fishery directory (Fiskeridirektoratet) for being very helpful and assisting us in the field.

This project was funded by FHF – Norwegian Seafood Research Fund and the acoustic equipment was provided by Genuswave LTD.

1 Abstract

Norwegian fishers report an increase in negative interactions with marine mammals: An issue that is also on the rise, globally. Efforts to mitigate these interactions using currently available methods have yielded little success for larger cetaceans. This study examines the efficacy of a novel target-specific acoustic startle technology (TAST) in deterring killer whales (*Orcinus orca*) and humpback whales (*Megaptera novaeangliae*) during interactions with purse seine fisheries for Norwegian spring-spawning herring (*Clupea harengus)* in Northern Norway. This was done by conducting controlled exposure playback experiments in proximity of fishing vessels and feeding whales and recording the presence of the whales before, during and after the acoustic playbacks. The results suggest that TAST may be effective in reducing the likelihood of negative killer whale interactions with the fishery, but on the effect was much less clear in terms of humpback whale behavior. Overall, the expected number of surfacings of killer whales was reduced by 85% during exposure to TAST, with no evidence of habituation. Additionally, the effect of TAST appears to attenuate quickly over distance with a strong recovery (return of whales) during the post-exposure observation phase, which is positive for preventing harm due to long-term habitat avoidance. This thesis provides strong indications that TAST can be used as a safe and effective tool for mitigating whale-fisheries conflicts in the marine environment.

2 Introduction

The increasing intensity and frequency of negative operational interactions between marine mammals and commercial fisheries has become an issue of global concern (Johnson, 2005; Read et al., 2006; Moore & Van der Hoop, 2012; Robbins et al., 2015). Over the past three decades, human consumption from fisheries and aquaculture has nearly doubled (FAO, 2022), while the protections in place after industrial whaling have led to the recovery of up to 50% of depleted cetacean populations (Lotze et al., 2011; Wedekin et al., 2017; Zerbini et al., 2019). As a result, these populations are now competing with commercial fisheries for the same prey species, leading to a rise in negative interactions and posing significant challenges for fishery managers. Preventing these negative interactions has proven to be difficult, as the few tools developed to-date have been largely ineffective or have been shown to decline in effectiveness over time (Tixier, Gasco, et al., 2015, Hamilton & Baker, 2019; Lucas & Berggren, 2022). In some cases, the tools have even been found to be harmful to marine mammals (Götz & Janik, 2013). Consequently, there is an urgent need for cost-effective methods to reduce bycatch events such as entanglements and prevent costly losses to harvesters due to fishery losses or damaged gear. This study investigates the effectiveness of new acoustic-deterrent technology to mitigate negative interactions, preventing harm to marine mammals and fisheries alike.

An operational interaction between a fishery and a marine mammal is defined as direct contact in time and space, often when both are in pursuit of the same prey species (Beverton, R.1985 in Read et al., 2006; Hamer et al., 2012). Operational interactions may be neutral or have negative or positive outcomes to either or both parties (Mul et al., 2020). In some cases, cetaceans may benefit from fisheries by depredating on fish already captured or by targeting fish that has aggregated around a fishing net (Gilman et al., 2006; Tixier, Authier et al. 2015,). Feeding aggregations of cetaceans can also benefit fishers by leading them to commercially important target species (Escalle et al., 2015). Negative consequences for cetaceans are typically entanglement in fishing gear (Johnson, 2005; Cassoff et al., 2011; Moore & Van der Hoop, 2012; Van der Hoop et al., 2017) and bycatch (Alverson, 1994; Lewison et al., 2004; Read et al., 2006; Hamer et al., 2012; Reeves et al., 2013; Gray & Kennelly, 2018), while for the fisheries, the negative consequences can be loss of catch, loss or damage of fishing gear and increased operation time (Lien & Aldrich, 1982; Tixier et al., 2019).

Entanglement events and bycatch is a concern for population-level viability for cetaceans (Volgenau et al., 2011; Moore & Van Der Hoop, 2012) and a serious animal welfare issue (Robbins, 2012). According to one estimate, 650,000 marine mammals end up as bycatch in fisheries each year (Read et al., 2006). The global scale of illegal fishing activities, coupled with underreporting and the likelihood that many bycatch incidents go undiscovered, suggests that these figures are likely an underestimate (Read et al., 2006; Moore & Van Der Hoop, 2012; Moore, 2014). In the western North Atlantic population of humpback whales (*Megaptera novaeangliae*), entanglement in fishing gear may represent up to 5.4% of natural annual mortality (Volgenau et al., 2011) and it is estimated that 78% of the population may have experienced a previous entanglement event (Robbins, 2012). A study in the same area found entanglement to be the leading cause of human-caused mortalities for eight different large whale species (Van der Hoop et al., 2013). In Mexico, fishery bycatch of the vaquita (*Phocoena sinus*) has resulted in the species becoming the most endangered marine mammal in the world (D'Agrosa et al., 2000).

Entanglement events involve almost all existing fishing gear (Gilman et al., 2006), such as fixed gear like pots and gill nets (Johnson, 2005), longline gear (Gilman et al., 2006), trawl nets (Bonizzoni et al., 2022) and purse seines (Read et al., 2006). The impacts can be acute or chronic, ranging from direct drowning in fishing nets to more gradual impacts from wounds and stress that increase morbidity over time (Cassoff et al., 2011; Moore & Van Der Hoop, 2012; Moore et al., 2013; van der Hoop et al., 2017). Entangled cetaceans can often carry fishing gear for months or years, suffering from infections, exhaustion, and emaciation before dying (Moore & Van Der Hoop, 2012; Moore et al., 2013; Van der Hoop et al., 2017). In addition to concerns for animal welfare and population viability, there may be financial losses in the fishing industry due to negative interactions between marine mammals and fisheries, resulting from loss of catch, gear damage or loss off gear, and downtime for repairs (Lien & Aldrich, 1982). Bycatch and entanglement events may also cause unwanted attention and lead to a negative reputation for fisheries.

Attempts to prevent bycatch, entanglement and depredation have been made through various mitigation methods including implementing physical barriers, modification of fishery practices, culling, relocation, or deterrence devices (Shivik et al., 2003; Quick et al., 2004; Hamer et al., 2008; Götz & Janik, 2013). Reviews of such mitigation strategies are quite mixed in terms of their effectiveness, as the outcomes are highly dependent on the type of fishery operation, target species and location (Hamer et al., 2012; Dawson et al., 2013;

Hamilton & Baker, 2019; Lucas & Berggren, 2022). Physical barriers between prey and predator have high costs and are difficult to implement (Quick et al., 2004). Elimination and relocation of predators have shown to be ineffective and may have adverse effects on populations (Yodzis, 2001; Thompson et al., 2007) and ecosystem functioning (Estes & Palmisano, 1974; Terborgh et al., 2001; Casini et al., 2008), while also causing ethical concerns. Acoustic deterrence methods are also problematic as they have been shown to lead to hearing damage (Götz & Janik, 2013), habituation (Dawson et al., 2013; Götz & Janik, 2013; Tixier, Gasco et al., 2015), habitat exclusion (by driving animals away from their natural habitats) (Johnston, 2002; Morton & Symonds, 2002; Kastelein et al., 2005; Thompson et al., 2007; Götz & Janik, 2013), and masking of sounds used for communication, orientation, or prey detection (Götz & Janik, 2013).

Despite having its limitations, acoustic technology has been the most widely tested of all mitigation efforts (Lucas & Berggren, 2022), and is considered the most benign means of addressing the problem (Götz & Janik, 2013). Sounds produced by acoustic deterrent devices (ADD) are assumed to work by either driving the animal away by exceeding the auditory pain threshold, or through aversion from acoustic stimuli itself without causing pain (Götz & Janik, 2015). The effectiveness of ADDs varies greatly between studies, with some finding that ADDs such as pingers can be effective for preventing bycatch of some small odontocetes (Dawson et al., 2013; Lucas & Berggren, 2022), while other studies find them insufficient for preventing depredation from pinnipeds, large baleen whales and odontocetes (Hamer et al., 2012; Götz & Janik, 2013; Hamilton & Baker, 2019). In some cases, marine mammals have been attracted to ADDs designed to deter other species. For example, pinnipeds have been shown to be attracted to pingers deployed on gillnets to prevent bycatch of small odontocetes, increasing the risk of seals being entangled (Dawson et al., 2013). Studies conducted on the West Coast of Scotland demonstrated that ADDs are a significant and chronic source of underwater noise (Findlay et al., 2018), and may potentially lead to hearing damage on target and non-target species (Götz & Janik, 2013; Findlay et al., 2018) as well as have a broader negative impact on the marine ecosystem (Findlay et al., 2018).

The purse-seine fishery in Northern Norway, targeting Norwegian spring spawning (NSS) herring, has experienced an increase in interactions with killer whales (*Orcinus orca*) and humpback whales since 2011 (Jourdain & Vongraven, 2017; Bjørge, 2022; Lindbæk, 2022). Beginning in the winter of 2010/11, the overwintering areas for the herring began to shift to more in-shore areas in specific northern Norwegian fjords during winter (Rikardsen, 2019; Salthaug & Stenevik, 2020). As a result, large masses of herring have entered these fjords, attracting large numbers of purse-seine fishing vessels as well as killer whales and humpback whales (Jourdain & Vongraven, 2017; Rikardsen, 2019; Bjørge et al., 2022). From 2012-2016, the main area for these herring aggregations were in the fjords around Kvaløya close to Tromsø, but from the winter of 2017/2018, the overwintering hotspots shifted northward to the Kvænangen fjord system and western Finnmark, where the main coastal herring fishery takes place today (Rikardsen, 2019; Fiskeridirektoratet, 2022; Martinussen, 2022; Norsk Sildesalgslag, 2022). Fishers have reported an increase in the level of interaction with whales since 2017 (Bjørge et al., 2022), with some fishers describing an explosive increase in the concentration of killer whales, and to some extent also humpback whales, resulting in more frequent problems with whales caught in the fishing nets (Lindbæk, 2022). Killer whales and humpback whales are not only drawn to the overwintering herring in the fjords but have also learned to associate fishing vessels with an abundance of prey. This behavior suggests that whales have learned that active fishing operations provide an easy opportunity to feed on herring that aggregate outside the fishing net (Rikardsen, 2019; Mul et al., 2020). Killer whales have been shown to change their foraging behavior when in the vicinity of herring fishing vessels (Van Opzeeland et al., 2005) and may follow fishing vessels for hundreds of kilometers (Towers et al., 2019). This suggests a “dinner-bell effect” in response to vessel-specific acoustic signals or acoustic sounds from ongoing fishing operations such as winching and pumping (Mul et al., 2020).

The interactions between fishing operations and marine mammals in the NSS fishery have led to several incidents with negative outcomes for both whales and the fishing industry, with several instances of killer whales or humpbacks becoming entangled in purse seines, resulting in injuries and deaths in both species (Rikardsen, 2019; Fiskeridirektoratet, 2020; Bjørge & Sivle, 2021). These events have also led to lost catches and damage to fishing gear for the fishing industry (Bjørge & Sivle, 2021). Additionally, the issue has caused unwanted negative attention and perceptions of the fishery (Bjørge, 2022). Following from these occurrences, a new law was passed by the Directory of Fisheries in Norway in 2016 (Fiskeridirektoratet, 2020), which prioritizes disentanglement of whales over to the harvesting of fish. This new law has serious cost implications for the fishery as it means that the fishers may have to release all of their catch if they have a whale entanglement (Bjørge et al., 2022).

ADDs have been suggested as a possible mitigation approach in the NSS fishery. However, there is a general lack of knowledge on the effect of such deterrence devices on killer whales and humpback whales, and ADDs such as pingers have earlier shown little effect on both species (Todd et al., 1992; Tixier, Gasco, et al., 2015; Basran et al., 2020). In 2021, The Norwegian Seafood Research Fund (Fiskeri- og havbruksnæringens forskningsfinansiering, FHF) provided funding to the Norwegian Institute of Marine Research and UiT The Arctic University of Norway to investigate the potential of Acoustic Deterrent Devices (ADDs) as a deterrent (Bjørge & Sivle, 2021). This resulted in a proposal to investigate the target-specific startle technology (TAST) developed by the Sea Mammal Research Unit (SMRU) at the University of St. Andrews.

Götz & Janik (2011; 2015; 2016) developed TAST by utilizing the acoustic startle response, with species-specific sounds that trigger autonomous reflexes associated with a flight response. The mammalian startle reflex is a fast motor response that is elicited if short acoustic stimuli have a sudden onset (short rise time) and exceeds the auditory threshold by 60-90 dB of the targeted animal (Koch & Schnitzler, 1997). Sounds with long rise times are less likely to trigger a startle response (Schakner et al., 2017).

The startle reflex has been studied intensively in behavioral psychology and neuroscience for many decades as a reflex-arc model for simple learning mechanisms (Koch & Schnitzler, 1997). Repeated eliciting of the acoustic startle reflex has been shown to first lead to a behavioral response, and then further lead to a learned response of either habituation or sensitization (Götz & Janik, 2011, 2013; 2015). In fact, triggering the startle reflex with TAST has already been demonstrated to induce fear conditioning and a flight response in pinnipeds and echolocating odontocetes without leading to habituation (Götz & Janik, 2011; Götz et al., 2020). Habituation is defined as a decrease in behavioral response (Groves & Thompson, 1970), while sensitization is an increase in the behavioral response (Plappert et.al, 1999 in Schakner & Blumstein, 2013). The structure of a sound in an acoustic deterrent device can be constructed based on the psychological characteristics that contribute to aversion in different species. The startle response can be elicited in target species without affecting other species by adapting the startle signals to each target species and in a pattern that reduces the exposure to a level below what is expected to result in hearing damage (Götz & Janik, 2015, 2016; Götz et al., 2020). Also, the magnitude of the startle response and the subsequent sensitization can be manipulated by adjusting the sounds produced by TAST (Hiley et al.,2021).

The first attempt to harness the startle reflex in the context of developing TAST was conducted with grey seals in 2011 (Götz & Janik), and further developed as a practical tool and tested in the field on harbor seals preying on farmed fish in 2015 (Götz & Janik). The startle reflex resulted in a flight response and interrupted foraging behavior and led to sensitization and avoidance behavior in most of the tested grey seals. It was unclear whether the startle reflex could be triggered in a similar way in echolocating odontocetes, as they possess the ability to regulate their auditory sensitivity (Nachtigall & Supin, 2015 in Götz et al., 2020). However, further studies conducted in 2020 by Götz and Janik demonstrated that the startle reflex could be triggered and result in an avoidance behavior in two species of echolocating odontocetes; bottlenose dolphins *(Tursiops truncates)* and false killer whales *(Pseudorca crassidens).* Shortly thereafter, a study demonstrated a strong avoidance response to TAST in harbor porpoises (Hiley et al., 2021).

The main objective of this study was to investigate if TAST can be used to deter two different species: killer whales and humpback whales, during purse seine fishing operations for NSS herring. This was done by conducting controlled exposure playback experiments in proximity of fishing vessels and feeding whales and recording the presence of the whales before, during and after the acoustic playbacks.

Specifically, we aimed to answer the following questions:

1. Are killer whales and humpback whales deterred from the area during acoustic playback?
2. Is there evidence of habituation during the experiment?
3. Do whales return to the study area after playback has ended?
4. Is there a difference in deterrence effect over different distances?

3 Method

3.1 Study area

The experiments were conducted east and northeast of Skjervøy (70°02´N, 20°58´E) on the coast of Northern Norway (Figure 1) during November and December 2022. This area is a part of the Kvænangen fjord system, and the study area covered approximately 240 km² of the fjord. Fishing in the fjords occurs from approximately November until January (Rikardsen, 2019; Bjørge, 2022), as it is linked to the timing of overwintering for the NSS herring (Jourdain & Vongraven, 2017; Bjørge et al., 2022; Fiskeridirektoratet, 2022; Martinussen, 2022; Norsk Sildesalgslag, 2022). The relatively confined area of the fjord combined with prevailing offshore winds and the narrow window of fishing activity creates an ideal opportunity to use a small boat to perform field experiments during real-time fishing operations with feeding whales around. Although certain wind directions can create quite rough conditions even inside the fjord, there are typically sheltered areas within the fjord allowing the experiments to be carried out daily.

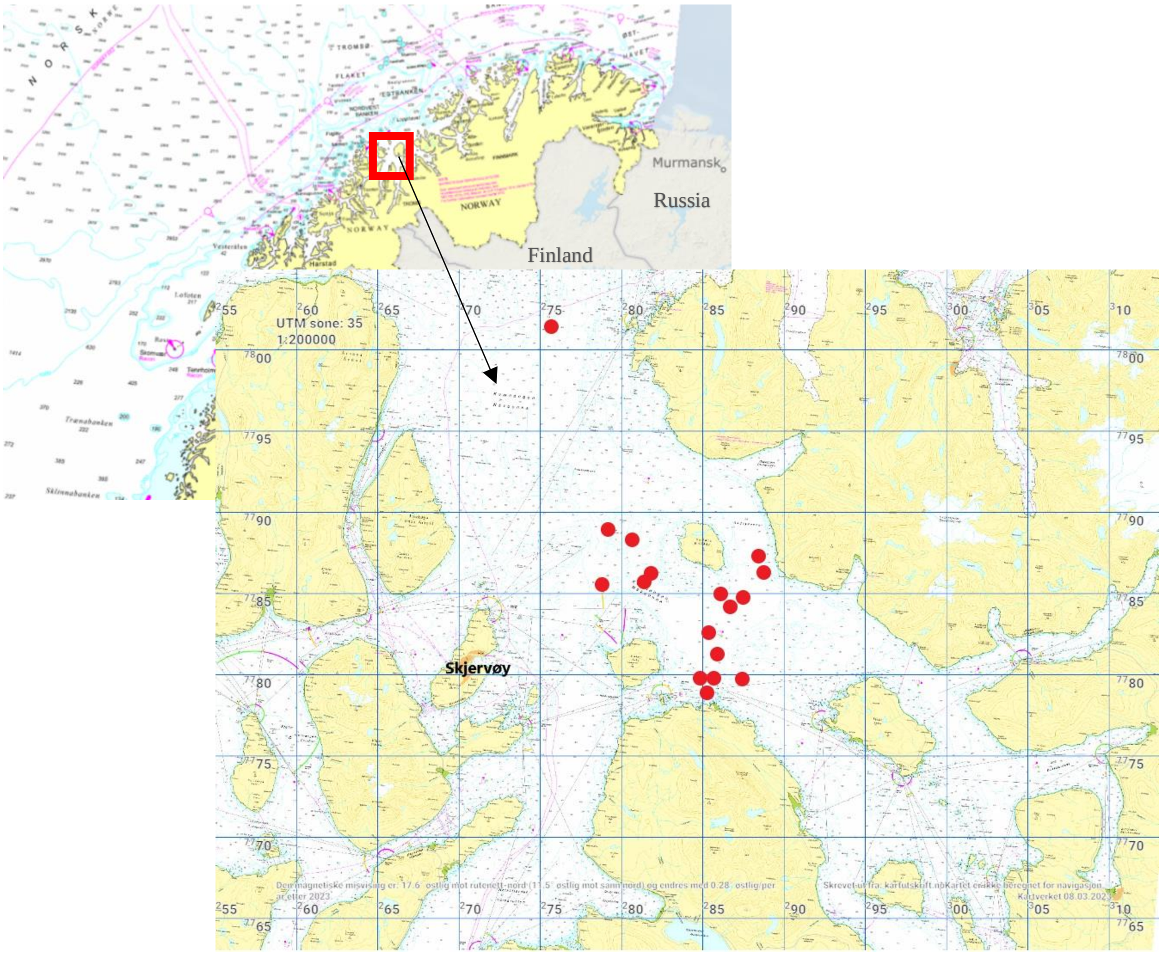


Figure 1: Inset map of the Kvænangen fjord system. The top left map shows where in Norway the study was conducted marked with the red square, while the bottom right map is zoomed in to the study area. The experiment sites are marked as red dots. Coordinates given in UTM (zone 35), scale 1:200 000.

3.2 The fishing operations

For small and medium fishing vessels, it would normally take 1-2 hours from the net setting until the net was hauled and the herring was pumped onboard the vessel. For larger vessels, and if a separate pumping vessel conducted the pumping, the process could be longer. The fishing operations occured during dark hours, as this is the time when herring become accessible by moving upward in the water column (Huse & Korneliussen, 2000). Due to variability in the time and place of fishing operations, an opportunistic and flexible approach was required to be ready and on-site when opportunities arose. Surveillance and information systems like Barentswatch and MarineTraffic were used to anticipate the location of fishing vessels. Once a vessel sets a purse seine for herring, they operate in darkness and use indicator lights to signal whether they are fishing, the size of the vessel, and whether the vessel is at speed (Lovdata, 1975). We approached fishing vessels after making contact over the VHF radio. The experiments relied on visual observations which required some visible light to detect whales at the surface, so we stood by until the net was retraced (pursed up alongside) and the fishing vessel had turned their working lights on. At this point, the process of pumping the herring from the net and into the fishing vessel would begin. This stage lasted between 30-40 minutes, which provided enough time for us to perform the experiments under consistent fishing activity while there were whales around. Sometimes, the vessel would stand by with the net hauled up alongside the ship and wait for a nearby pumping vessel to retrieve the herring. After the pumping process was finished, the whales would typically move away, often to search out another vessel still actively fishing.

3.3 Study design and protocol

The experiments were conducted from a small boat, identified herein as the "playback vessel" (Arronet 30, 9 meters in length), which was positioned 90 degrees from the bow of the fishing vessel at 30-60 meters from the fishing net (Figure 2). While the aim was to get as close as possible, the position of the playback vessel relative to the fishing vessel varied due to current, wind and frequent re-orientations of the fishing vessels.

The experiments were performed by a 4-person team consisting of a driver, a TAST operator, and two observers. The experiment lasted for a total of 15 minutes, wherein whale surfacings were recorded for 5 minutes prior to the acoustic playback, 5 minutes during the playback, and 5 minutes after the playback, a method which follows a standard controlled exposure experiment methodology (Tyack et al., 2003). A surfacing was strictly defined as a whale breaking the surface with an exhalation. For all experiments, the ID of the fishing vessel, vessel size, the absence or presence of a pumping vessel, observer ID, vessel distance, light conditions (visibility resulting from ambient light and/or artificial light from the fishing vessel), and sea state were recorded. All observations were recorded on a smartphone, either with voice recording or video. If there were no whales around during the pre-playback period, we would either wait around until whales arrived or abort the experiment and move on to another fishing vessel.

Spatial zones observations: One observer was responsible for recording surfacings in spatial zones between the playback vessel and the fishing vessel. The zones were defined as quadrants A, B, C and D based on the length of the fishing vessel (Figure 2). The middle of the fishing vessel, typically where the net sits, demarcated the division of the zones. For each surfacing, species and zone were recorded.

Distance observations: The second observer recorded the distances of each surfacing from the playback vessel. The distances to the surfacing animals were recorded within the 180-degree sector area between the sound source and the fishing vessel. All observations were made by naked eye and distances were estimated and calibrated using a range finder. The maximum distance measurable varied with light conditions but ranged up to 120 meters away. For each surfacing, distance and species were recorded.

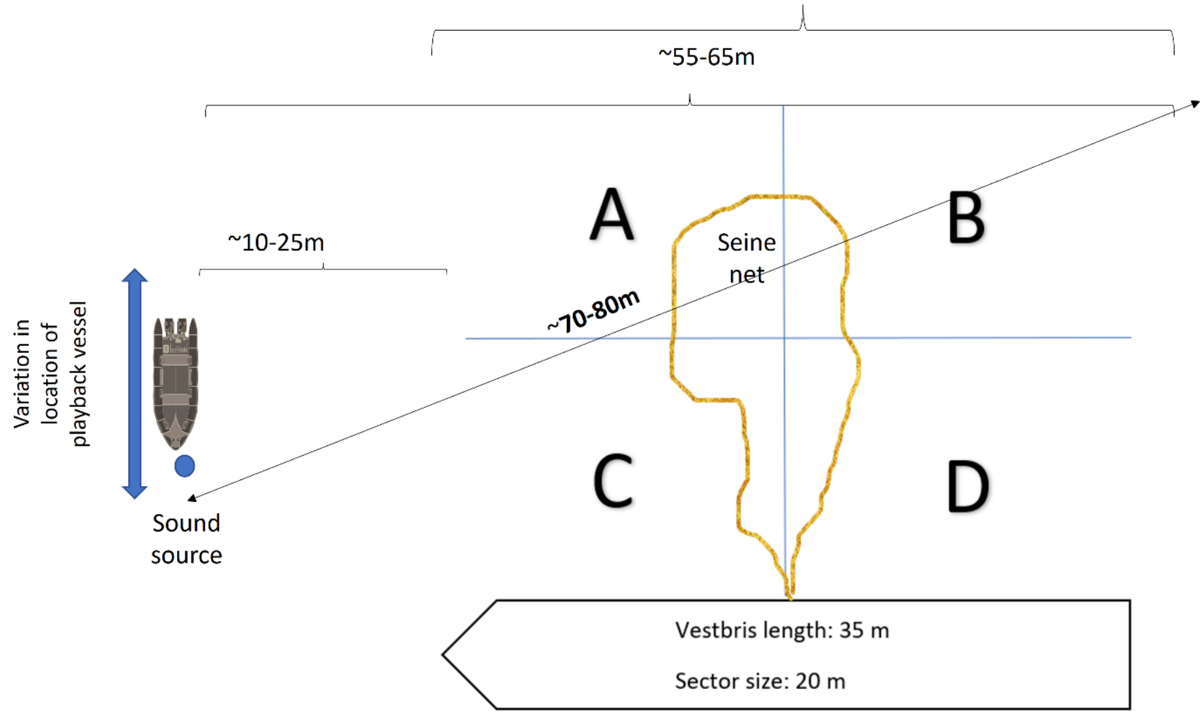


Figure 2: Experimental set up in the field. Vestbris is given as an example of fishing vessel. The playback vessel was placed 90 degrees from the bow of the fishing vessel. The size of zones A, B, C and D were determined based on the size of the fishing vessel. The diagonal arrow represents the mean distance from the playback vessel to the furthest edge of the observation area for the zones. Sector size is the length of each zone.

3.4 Acoustic system

The TAST (target-specific startle technology) system was provided by Genuswave LTD. The unit consisted of a TAST device from Genuswave LTD and a Lubell 9161T transducer. The TAST playbacks were produced by the speaker lowered 20 meters into the water from the playback vessel. The sounds consisted of 0.2s long sound signals with an overall duty cycle of ~1.2%, so ~0.4% for each signal type. Only short signals were played to increase the chance of triggering a startle response which has led to a learned avoidance behavior in other marine mammals (Götz & Janik, 2015; Götz et al., 2020). Signals within a sequence were emitted at randomized intervals. The sound emitted from the unit was a mix of three different signals:

1. Low-frequency (LF), centered at 1kHz, ~2 octaves, similar to Götz & Janik (2015) with am amplitude of 180 dB re 1 µPa
2. High-frequency (HF), ~5 to 20kHz as tested for odontocetes in Hiley et al. (2021) with an amplitude of 183-184 dB re 1 µPa

3. Broadband (BB), from ~0.7kHz to 20kHz, with modest peaks at the resonances at 1kHz and 10kHz with an amplitude of 183-184 dB re 1 μPa

Based on published audiograms for killer whales where the median threshold for 20 kHz has shown to be 53 dB re 1 μPa (Branstetter et al., 2017), we considered it likely that the signals used in this study could trigger a startle response. Little information is available on the hearing of humpback whales, but one study suggested that auditory threshold at 1 kHz may be between 60-80 dB re 1 μPa (Erbe, 2002). Based on the sound signals they produce, it is likely that killer whales have more sensitive hearing at high frequencies than humpback whales (Southall et al., 2007). We were interested in investigating whether the range of different frequencies (LF, HF and BB) could trigger a reaction in both species during the experiments.

3.5 Audio analysis

To extract the timings of surfacings and the observation-period start and end times, the audio recordings were analyzed using the audio software Audacity. Whenever the observers voice indicated a surfacing or other important time event, a mouse click allowed the event to be labeled on to the audio file with the recorded information on species (killer whale or humpback whale) and distance or zone. Each observation was time-stamped with seconds (0-900 seconds) over the duration of each 15-minute experiment. The datapoints were then manually exported into Excel datasheets, one for the zone-counts and one for the distance-counts. The datasheets were completed with additional information for each experiment including distance measurements to the bow and net of the fishing vessel; light conditions; presence of pumping vessel; the ID of the fishing vessel; and the ID of the observer. When available, the video recordings were used to double-check the accuracy of the audio-recordings.

3.6 Data analysis

The surfacings were transcribed as count data within three different treatment periods (pre, playback and post). Generalized linear models were chosen as the model type, as these models are well suited for ecological non-normal data with random effects (Bolker et al., 2009). Two different datasets were analyzed in this study; one was based on the data acquired from surfacing counts within the different spatial zones, and the other was based on surfacing counts where distances were recorded. The analysis on zones was done separately for the two species killer whales and humpback whales, and the data were coded into a count of surfacings per zone per treatment period. For the distance analysis, only the data on killer whales was selected for analysis due to limited data on humpback whales. Distances were categorized into distance bins and the aim was to investigate if there were any differences in surfacing counts between the treatment periods and two different distance ranges from the playback device. For simplicity, we binned the data into two distance intervals from 0 to 50 meters and from 50 to 100 meters.

Evaluation of the candidate models was performed in a 3-step selection process (appendix 1A,2A and 3A) (Smith et al., 2009). This selection process was done separately for the zone analysis and the distance analysis, and also separately for the two different species of whales. In each step, a model selection table was made based on the Akaike Information Criterion (AIC) and Akaike weights (AIC weight). Likelihood ratio tests (Zeileis & Hothorn, 2002) were applied to the top two models, to ascertain if there was sufficient support for the top model being substantially better than the second best model. In cases where there was no significant difference between the two top models, we chose the simplest model.

Step 1, assessing distribution family type and the zero-inflation argument:

Different model combinations of family distribution type (negative binomial 1, negative binomial 2 or Poisson distribution) and the presence of the zero-inflation argument was assessed. We tested both Poisson and negative binomial distribution models as these distributions are considered appropriate for count data (Hoef & Boveng, 2007). Treatment periods and zone/distance were treated as fixed effects and the surface count as the response variable. Zones/distance nested within each trial ID, vessel size, light conditions and presence of pumping vessel were included as random effect in all the candidate models in this step.

Step 2, assessing random effects:

Plots for covariates were first assessed to trim down the number of candidate models for model selection in this step. Based on these plots, two different random effect specifications were chosen as consistent in the candidate models in this step: Trial ID or zones/distances nested within trial ID. Combinations of these two random effect specifications and the remaining random effects (vessel size, light conditions, and the presence of a pumping vessel) were then assessed based on AIC.

Step 3, assessing relationship between fixed effects:

Fixed effects and interaction terms between fixed effects zones or distance were assessed based on AIC. The response variable (surfacing counts) and one fixed effect (treatment period) were constant, and different combinations of the relationship between treatment and zones were compared (treatment*zones, treatment+zones or just treatment).

Once the best supported candidate model had been identified, predictions from this model were used to compute estimated marginal means of the response variable, with the packages emmeans (Lenth, 2022) and ggeffects (Lüdecke, 2018). The model fit was assessed with interpretations of a Q-Q plot of the residuals; plot of residuals vs. the predicted values; plot of scaled residuals VS predicted values for treatment and zones/distance; an autocorrelation function (ACF) which measured the correlation between observations; and finally, a zero-inflation test to check for any zero-inflation problem. These model diagnostics were performed with the package DHARMa (Hartig, 2022) which is specifically developed for generalized mixed linear models. Results from the summaries of the top models were used to calculate and manually design graphs to present the estimated reduction effect surfacings (in percentage) during the playback period.

4 Results

Table 1 provides a summary over all the experiments (N=18) in which data were used in the 3 different analyses (1: zones and killer whales, 2: zones and humpback whales, 3: distances and killer whales) presented in the results. There were insufficient data for a distance analysis for humpbacks. The number of experiments, the variation in the number and sizes of fishing vessels, the light conditions (good or poor), and the time of day (dusk or dawn) differed between the experiments (Table 1).

Table 1: Overview of the experiment trials that were conducted and used in the analysis, divided into 3 different datasets: zones and killer whales, distances and killer whales and zones and humpback whales.

	Zones Killer whale	Distances Killer whale	Zones Humpback whale
Number of trials	**15**	**14**	**9**
Number of fishing vessels	10	7	7
Good light conditions	7	5	5
Poor light conditions	8	9	4
Pumping vessel present	2	3	2
Dusk	8	8	4
Dawn	7	6	5
Number of observers	1	2	1

4.1 Killer whales and zones

For the first step in the model selection for killer whales and zones, a minor auto-correlation problem in the residuals was detected post diagnostics. Assessing only the models where zones nested within trial were included as a random effect seemed to fix this problem.

By doing so, we acknowledged that the zones were not absolute, and varied in size dependent on the trial setup (e.g., differences in artificial light from the vessel, our position relative to the net). The top model (including the zero inflation argument and with a negative binomial distribution type 2) had substantially more support than the second model (no zero inflation, distribution family: negative binomial) (likelihood-ratio test: df=0, $\chi 2$=0.98, p<0.0001).

In the second step, where the random effects structure was selected, the top model (surfacings~treatment*zones+1|trial/zones) had more support over the second model (surfacings~treatment*zones+1|trial/zones+1|vessel size). While the likelihood ratio test did not indicate a significant difference between these two models (likelihood-ratio test, df=1, $\chi 2$=0.95, p =0.33), we chose the top model as it was also the simplest of the two.

In the last step when testing whether an interaction between the fixed effects treatment and zones improved the fit of the model, the model without this interaction (m14a: surfacing count~treatment+zones+1|trial/zones) had the best fit. The difference in AIC between this model and the second best (m13a: surfacings~treatment*zones+1|trial/zones) was 2.24. While the likelihood ratio test was not significant (likelihood-ratio test, df=6, $\chi 2$=11.80, p =0.07), m14a was the simplest of the two (Appendix 1B).

Model summaries (Table 2) shows results of the marginal R^2 and conditional R^2 which compared the goodness of fit for the model without (marginal) and with (conditional) random effects present (Nakagawa et al., 2017). The conditional R^2 was higher than the marginal R^2 for both m13 and m14, showing that the random variables improved fit of the variance with 0.197 (m13) and 0.192 (m14) (Table 2). The population variance $\sigma 2$ was 0.80 for m13 and 1.03 for m14. The random slope variance (τ_{00}) (Pinheiro & Bates, 2000) was lowest for zones nested within trials (0.17 in m13 and 0.13 in m14), and a bit higher for trial ID (0.31 in m13 and 0.34 in m14). The second model investigated the relationship between the different zones (m13a) and the top model looked at a more global effect (m14a) (Table 2). Based on this, both of the two top models were chosen to be included in the results.

Table 2: Model summary for the zone data on killer whales, from the generalized linear mixed models A) m13a (given as m13) and B) m14a (given as m14). Both models examine changes in the number of surfacings as a function of treatment (pre, playback, post) and zones (A,B,C,D). To the left, model M13 is with interaction between zones and treatment, while m14 to the right is without this interaction term and look at a more global effect. Significant values are highlighted in bold.

A

	GLMM with interaction term(m13)	
Predictors	*Estimate*	*p*
Count Model		
(Intercept)	5.56	**<0.001**
treat [during]	0.08	**<0.001**
treat [post]	0.22	**<0.001**
zones [B]	2.30	**0.007**
zones [C]	0.88	0.701
zones [D]	1.33	0.376
treat [during] × zones [B]	2.93	**0.042**
treat [post] × zones [B]	3.57	**0.004**
treat [during] × zones [C]	1.20	0.773
treat [post] × zones [C]	2.27	0.092
treat [during] × zones [D]	3.04	**0.045**
treat [post] × zones [D]	2.75	**0.028**
(Intercept)	23.20	
Zero-Inflated Model		
(Intercept)	0.09	**<0.001**
Random Effects		
σ^2	0.80	
$\tau_{00\ zones:trial}$	0.17	
$\tau_{00\ trial}$	0.31	
N_{zones}	4	
N_{trial}	15	
Marginal R^2 / Conditional R^2	0.475 / 0.672	
AIC	863.294	

B

	GLMM with global effect(m14)	
Predictors	*Estimate*	*p*
Count Model		
(Intercept)	4.04	**<0.001**
treat [during]	0.15	**<0.001**
treat [post]	0.46	**<0.001**
zones [B]	4.15	**<0.001**
zones [C]	1.12	0.671
zones [D]	2.19	**0.003**
(Intercept)	7.83	
Zero-Inflated Model		
(Intercept)	0.05	**0.003**
Random Effects		
σ^2	1.03	
$\tau_{00\ zones:trial}$	0.13	
$\tau_{00\ trial}$	0.34	
N_{zones}	4	
N_{trial}	15	
Marginal R^2 / Conditional R^2	0.390 / 0.582	
AIC	863.090	

While the number of sightings in all zones were relatively high during the 5-minute period pre-playback, there was a dramatic and significant drop in surfacings following the initiation of playbacks (Figure 3). This was significant in all zones ($p< 0.0001$). Once playbacks ended, there was a significant increase in surfacings in all zones except for zone A ($p=0.09$) (Figure 3). There was a visible decrease in surfacings from pre- to post playback in all zones, but only zone A (closest to playback vessel and furthest from fishing vessel) showed a significant difference ($p< 0.001$).

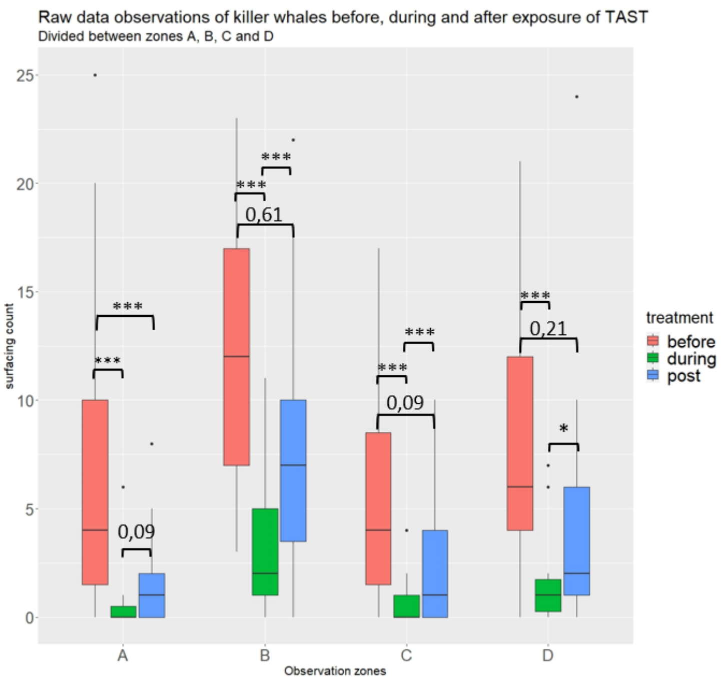


*Figure 3: Boxplot of observed number of killer whale surfacings in the different zones A, B, C and D during the different treatment periods before, during and after playback. The X-axis show the different zones A, B, C and D, while the y-axis represents the number of surfacings. The different colors represent the different treatment periods before(red), during(green) and post playback(blue). Boxes show the medians and the interquartile ranges (IQR), where the lowest whisker is the smallest observation over the 25% quartile- 1.5x IQR , and the upper whisker is the largest observation below the 75% quartile + 1.5 x IQR. The dots represent the outliers, however 3 outliers (counts over 30 in before playback in zone A and in before and after playback in zone B) do not show in this figure as the y-axis was trimmed for aesthetic purposes. Significance values are derived from the model results, where ***<0.001 (p-value).*

A pairwise comparison of the different treatment periods and zones was provided with model m13a (Table 3) and show a significant reduction during playback across all zones ($p<0.0005$). The recovery of surfacings was significant in all zones (Zone B, $p<0.0005$, zone C, $p<0.005$, zone D, $p<0.05$) except for zone A ($p=0.09$) (Table 3).

Table 3: The summary of pairwise comparisons of the fixed effects "treatment" in the generalized linear mixed model m13a. Significant values are highlighted in bold.

contrast	zones	ratio	SE	df	lower.CL	upper.CL	null	t.ratio	p.value
playback/pre	A	0.08	0.03	164	0.03	0.22	1	-5.78	**0.000**
post/pre	A	0.22	0.07	164	0.10	0.49	1	-4.46	**0.000**
post/playback	A	2.81	1.38	164	0.88	8.95	1	2.11	0.090
playback/pre	B	0.22	0.07	164	0.11	0.45	1	-5.16	**0.000**
post/pre	B	0.77	0.21	164	0.40	1.47	1	-0.95	0.608
post/playback	B	3.43	1.02	164	1.70	6.94	1	4.15	**0.000**
playback/pre	C	0.09	0.04	164	0.03	0.26	1	-5.46	**0.000**
post/pre	C	0.49	0.17	164	0.22	1.10	1	-2.08	0.097
post/playback	C	5.33	2.48	164	1.77	16.00	1	3.60	**0.001**
playback/pre	D	0.23	0.08	164	0.11	0.51	1	-4.35	**0.000**
post/pre	D	0.59	0.18	164	0.28	1.24	1	-1.68	0.218
post/playback	D	2.55	0.95	164	1.06	6.14	1	2.51	**0.035**

P value adjustment: tukey method for comparing all estimates.

The reduction effect during the playback period is illustrated in Figure 4, and show that there was a 85% reduction across all zones when playback was initiated. For zones A and C (closest zones to the playback vessel), a reduction of 92 and 90% was estimated, while expected reductions were slightly lower with 78% and 76% for zones B and D (furthest from the playback vessel) (Figure 4). There was also a significant difference in surfacings between zone A and B (p=0.007) during the pre-playback period (Figure 4).

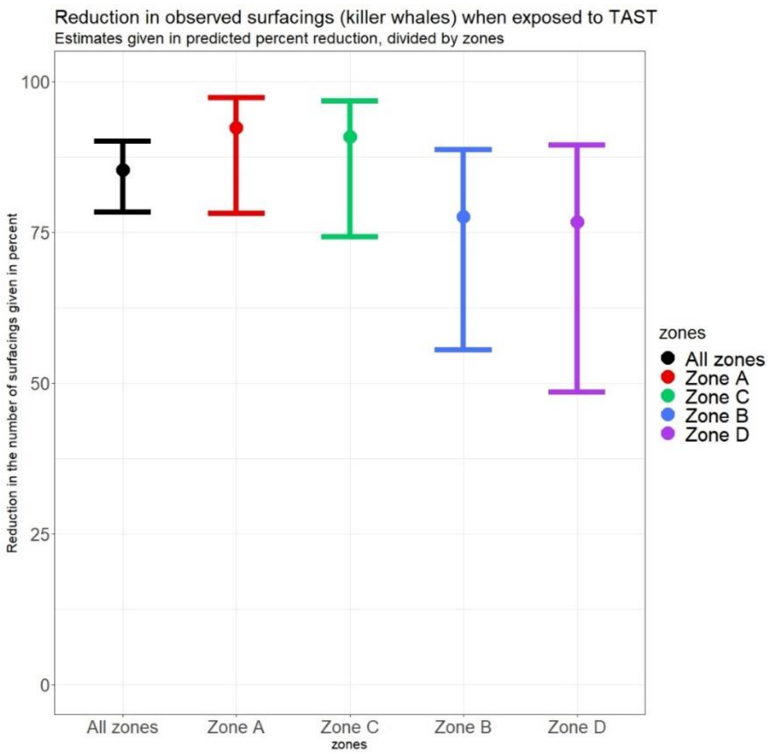


Figure 4: Expected reduction of surfacings in percent across zones from before playback to during playback based on the model. The Y-axis show the percents, while the X-axis represents the different zones and all zones combined. The dots represent the expected value, while the error bars show the confidence intervals (95%CIs) with the respective borders. The colors represent the different zones A(yellow), B(blue), C(green), D(purple) and all zones pooled(red)

During the playback period there was a significant difference between zone A and zone B ($p=0.042$) and D ($p=0.045$) (Figure 5). In the post-playback period, there was a significant difference between zone A and B ($p=0.004$) and A and D ($p=0.028$) (Table 1 and Figure 5). Zones A and C were fairly similar in the pre-payback and playback treatment periods ($p>0.7$) and had a non-significant difference in the post-playback period ($p=0.092$) (Table 1). Across all treatments, zones A and C significantly differ from zone B and D (Figure 4). Figure 5 provides an illustration of the expected surfacings in the different treatment periods, divided on the four different spatial zones based on model m13a, and it shows clearly and highly visibly how the counts of surfacings were greatly reduced during the playback, before they again increased when playbacks were over. Also, the difference between zone A+C (closest to playback vessel) and B+D (furthest) was clearly visible (Figure 5).

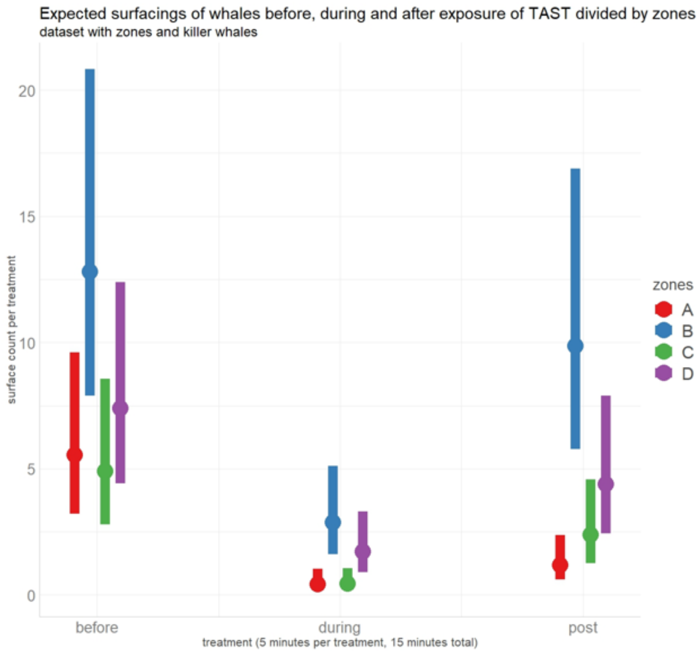


Figure 5: The expected surfacing counts (killer whales) based on model m13a (see Table. 2A). The x-axis shows the different treatment periods before, during and post playback, in which all are 5 minutes (total 15 minutes). The y-axis gives the mean expected number of surfacings. The colors represent the different zones A(red), B(blue), C(green) and D(purple). The dots represent the estimate, while the error bars represent the 95% confidence intervals.

4.2 Humpback whales and zones

There were, in general, fewer humpback whales present in the vicinity of fishing vessels compared to killer whales, and only 9 trials had sufficient observations to be used in this analysis (Table 1). Based on the results of the models for killer whales and zones, the zones A+C and B+D were most similar to each other (Figure 8). The 4 original zones A, B, C and D were for these reasons pooled into two zones: AC (closest to the playback vessel) and BD (on the opposite side of the seine from the playback vessel), and the models for humpbacks were fitted for these pooled zones.

There was strong support for the top model in the first step of the model selection (no zero inflation, distribution family: negative binomial type 2, likelihood-ratio test, df=0, $\chi 2=2.90$, $p = >0.0001$). In the second step, when selecting random effects, a likelihood-ratio test provided no strong support that either of the top two models were better (df=1, $\chi 2=0$, p = 1.00), however, the AIC weight for the top model (surfacings~treatment*zones+1|trial) was 0.534, compared to 0.126 for the second model (surfacings~treatment*zones+1|trial +1|light). We chose the top model, as it was the simplest model of the two. By including only trial as a random effect in the top model, we acknowledge that the light conditions varied for each trial.

In the third step, the relationship between the fixed effects treatment and zones were investigated, resulting in a strong support for the top model m14b (surfacings~treatment+zones+1|trial) over the second model m13b (surfacings~treatment*zones+1|trial) (likelihood-ratio test, df= 1, $\chi 2=5.60$, $p <0.05$), and the difference in AIC was 3.07 where m14b had the best AIC score (appendix 2A). Based on this, the top model m14 was chosen to be presented in the results. In m14b, the conditional R2 was higher than the marginal R2 (0.412 higher), showing that the random effects strongly improved the overall fit by reducing the residual variance (Table 4). The population variance $\sigma 2$ was 0.73, and the random slope variance ($\tau 00$) for the random effect trial ID was 0.64.

Table 4. Summary of the generalized linear mixed models m13b and m14b for humpback whales and pooled zones. Random effects are the different trials. ∆AIC for m13 and m14= 3.017.

	GLMM with interaction term(m13)		**GLMM for global effect(m14)**	
Predictors	*Estimate*	*p*	*Estimate*	*p*
(Intercept)	1.54	0.309	1.65	0.186
treat [during]	0.60	0.310	0.64	0.113
treat [post]	0.72	0.481	0.52	**0.033**
zones [B]	2.04	0.056	1.83	**0.016**
treat [during] × zones [B]	1.08	0.896		
treat [post] × zones [B]	0.58	0.386		
Random Effects				
σ^2	0.73		0.73	
τ_{00}	0.64_{trial}		0.64_{trial}	
ICC	0.46		0.46	
N	9_{trial}		9_{trial}	
Marginal R^2 / Conditional R^2	0.109 / 0.523		0.111 / 0.523	
AIC	212.466		209.449	

The observed results from the experiments are illustrated in Figure 6, and show no striking pattern except that there was a visible difference between the two different binned zones (AC and BD) (Figure 6).

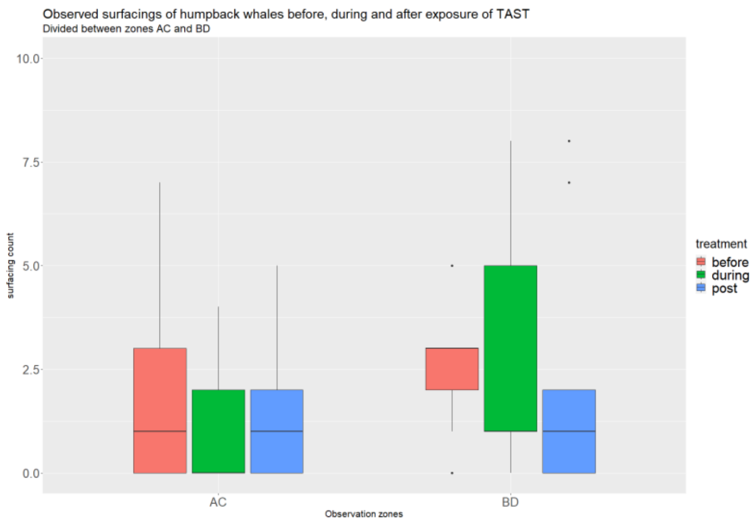


Figure 6: Number of humpback whale surfacings observed divided on two zones: AC and BD. Red indicates the pre-playback period, while green represents the playback period and blue the post-playback period.

Results from the model m14 however showed a significant difference in surfacings between zones AC and BD (p=0.016), and also a significant difference between the periods pre- and post-playback (p=0.033) (Table 4). Otherwise, the results show a small effect of the TAST on humpback whales during the playback period, and a further small reduction after playbacks were ended. By using results given in summaries of model m13b and m14b (Table 4), a figure representing the changes in surfacings between the treatment periods (pre, playback and post) given in percentage was illustrated in figure 7. A reduction in surfacings of 37% during playback, and a further reduction to 48% in the post-playback (compared to pre-playback) was calculated (Figure 7).

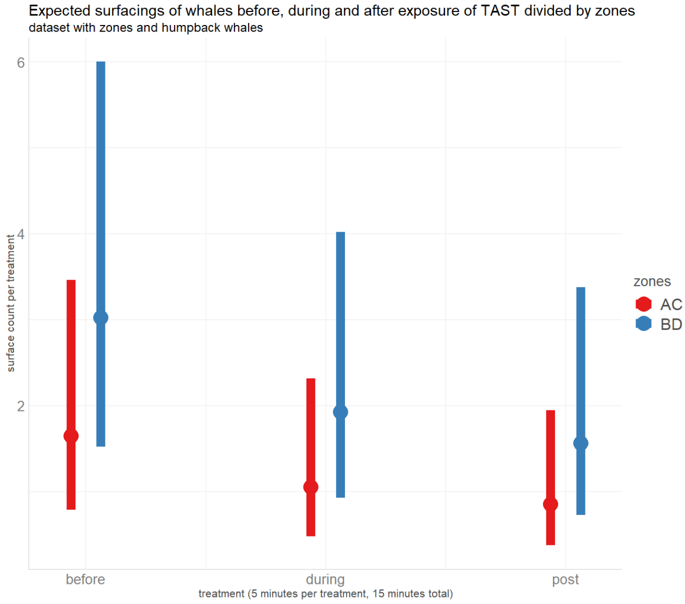


Figure 7: Expected surfacing counts based on the GLMM, between the different treatment periods before, during and post, and pooled zones A(closest) and B(furthest). The experiment lasted for 15 minutes where each treatment type was 5 minutes long. The Y-axis show the expected mean surface count, while the X-axis show the different treatment periods. Red bars show zone A and blue bars show zone B. The dots are the expected values while the bars represent the 95% confidence interval.

4.3 Killer whales and distances

For the first step in the model selection for killer whales and distances, a likelihood ratio test comparing the top model (no zero inflation argument, distribution family: negative binomial type 1) and the second best model (with zero inflation argument, distribution family: negative binomial type 1) from a model selection table (appendix 3A) showed no difference in the fit between the two models (p=1.00). However, when including the third and fourth best model in a likelihood ratio test, there was a strong improvement in the top 2 models from the third model (p= 0.05) and the fourth model (p>0.0005). The difference in AIC was 2.98 between the top model and the second best, resulting in the top model being selected in this step.

In the second step where random effects were selected, a likelihood ratio test between the two top models from a model selection table (Appendix 3A) did not show any difference (df=1, $\chi2$=0, p =1.0). However, the top model (distance nested within trial as random effect) had a better score on AIC weight (0.367) and was the simplest model compared to the second best model (distance nested within trial and pumping vessel as random effect) (AIC weight=0.098). Distance bins nested within trial as a random effect account for variation between the two distance bins in all trials.

In the last step, the relationship between the fixed effects treatment and distance was assessed. The top model (based on AIC) m1 (surfacings~treatment*distance+1|trial/distance) was compared with m2 (surfacings~treatment+distance+1|trial/distance) in a likelihood ratio test, resulting in moderate support for model m1 (likelihood-ratio test, df=2, $\chi 2$=5.35, p = 0.07). The difference in AIC was 0.3, and the AIC weight was relatively similar (m1: 0.538 and m2: 0.462). Both models were included in the results, as one model looked at interactions between each distance bin and treatment period (m1), and the other model looked at a more global effect (m2). The conditional R2 was higher than the marginal R2 for both m1 (0.302 higher) and m2 (0.270 higher), showing an improvement in the overall fit by reducing residual variance when random effects were included (Appendix 3B). The population variance was 0.40 for m1 and 0.45 for m2.

A significant difference was detected between all three treatment periods at distances 0-50 meters (Figure 8). This was not the case for the distance bin 50-100 meters, where the differences between treatment periods did not appear to be significant.

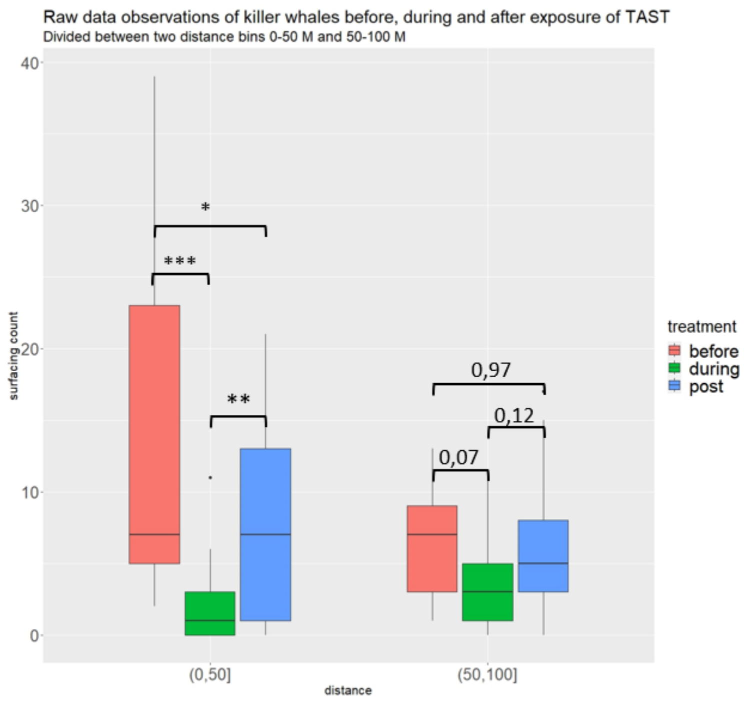


Figure 8: Observed surfacings of killer whales divided by different treatment periods (pre, playback, post), and two distance bins illustrated in a boxplot. The X-axis represents the two distance bins 0-50 m and 50-100 m. The X-axis show the expected mean of surfacing observations. Each bar represents different treatment periods, where red is before playback, green is during, and blue is post playback

There was a significant difference in surfacings between the two distance bins during playback (p=0.030) (Appendix 3B). A pairwise comparison illustrated in Table 5 was provided from model m1 and showed a great reduction in surfacings at both distances between the pre-playback and playback periods, with significant evidence at 0-50 m (p<0.001) (Table 5) and moderate evidence at 50-100 m (p=0.06) (Table 5).

Table 5. Pairwise comparisons retrieved from model m1 (killer whales and distances). Surfacing counts between the treatment periods within each distance bin is compared. Significant values highlighted in bold

contrast	dist	ratio	SE	df	asymp.LCL	asymp.UCL	null	z.ratio	p.value
playback/pre	(0,50]	0.1931285	0.05964886	Inf	0.09364305	0.3983063	1	-5.3241659	**0.000**
post/pre	(0,50]	0.5431155	0.12068896	Inf	0.32263261	0.9142735	1	-2.7470265	**0.017**
post/playback	(0,50]	2.8121976	0.93210179	Inf	1.29323604	6.1152453	1	3.1195278	**0.005**
playback/pre	(50,100]	0.4982096	0.15455283	Inf	0.24079993	1.0307845	1	-2.2459619	0.064
post/pre	(50,100]	0.9400023	0.25011740	Inf	0.50384389	1.7537263	1	-0.2325338	0.971
post/playback	(50,100]	1.8867606	0.59299177	Inf	0.90326212	3.9411212	1	2.0199799	0.107

P value adjustment: tukey method for comparing all estimates.

The global model (model m2 without the interaction term between distance and treatment) showed an overall reduction of 69% from pre-playback to playback (Appendix 3B). The interaction model (m1) investigated the interaction between the distance bins and the treatment periods, and showed that for the nearer distance bin (0-50m), an even larger reduction in surfacings (of 81%) was expected, while for the further distances (50-100m) the expected reduction was 51% (Figure 9).

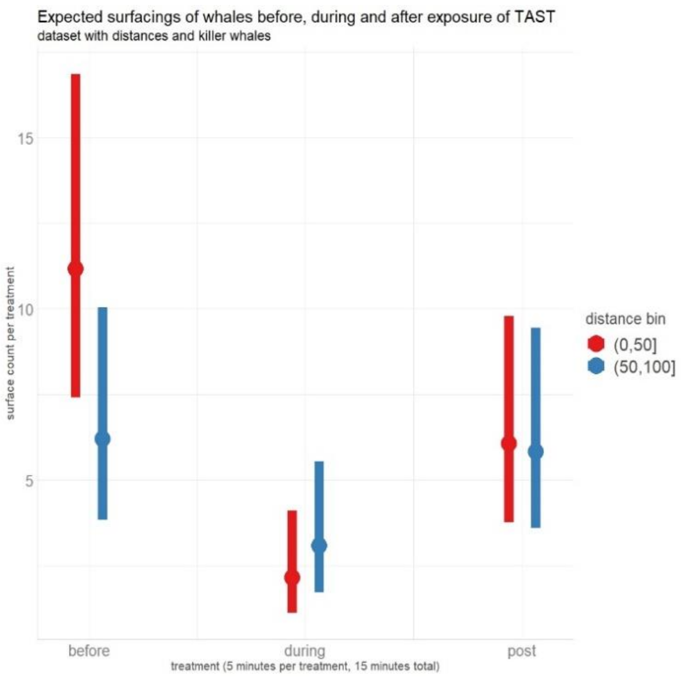


Figure 9. Expected surfacings (killer whales) based on the GLMM m1 (table 3), between the treatment periods(pre, playback, post), and time bins 0-50 m and 50-100 m from playback vessel. The experiment lasted for 15 minutes where each treatment type was 5 minutes long. The Y-axis show the expected mean surface count, while the X-axis show the different treatment periods. Red bars show distance bin 0-50 m and blue bars show distance bin 50-100 m. The dots are the expected values while the bars represent the 95% confidence interval

5 Discussion

The purpose of this study was to assess the effect of the new acoustic deterrence technology, TAST, as a tool to deter killer whales and humpback whales and reduce the risk of negative operational interactions with the NSS herring fishery in Norway. We found a clear and significant reduction in the number of killer whales in the zones and within 50m from the sound source, with an up to 90% reduction in killer whale surfacings during the playbacks. The response appeared to be much less strong at distances greater than 50 meters, and this is probably a result of sound attenuation having reduced the signal strength sufficiently for whales to respond less at these distances. There was a relatively quick recovery of whales returning to the experimental site after playback was ended. While the effect was strongly significant for killer whales, the results for the humpback whales looked quite different, with no significant effect from TAST detected. This study is the first study to show a significant and consistent deterrent effect from an acoustic device on killer whales without evidence of habituation.

5.1 Deterrence effect

The successful deterrence effect observed in killer whales can be mainly explained by the nature of a startle reflex. Our findings with killer whales build on this growing base of knowledge (Götz & Janik, 2011, 2015, 2016, Götz et al., 2020, Hiley et al., 2021) and further strengthen the theory that the startle reflex can effectively be harnessed to elicit a flight response and avoidance behavior in marine mammals.

The limited response found with humpback whales was not entirely unexpected, as the species-specific startle response in humpbacks has not been successfully detected in previous attempts (Götz and Janik personal comment). The fact that we did detect a weak response could be an indication of a startle reflex and small flight response. But, due to the limited sample size of humpbacks, it was difficult to detect a consistent trend from overall variability, and this could have led to the insignificant results. However, there are other possible explanations. The playback setup used for our experiments was ideal for killer whales and less ideal for humpbacks, which likely have lower hearing sensitivity than killer whales based on the sounds they use (Erbe, 2022). The setup was likely not optimal for producing the low-frequency sounds at a sufficient amplitude.

While the startle itself is autonomous, the magnitude of the startle response has been shown to be modulated by emotional state, for example when it induces conditioned fear in animals (Götz & Janik, 2015). This means that the startle magnitude is an important factor for how strong the avoidance behavior is. It is also possible that the limited behavioral response observed with humpbacks was not associated with a response to the sound at all, but rather with a response to the behavior of the killer whales. Previous studies have shown that humpback whales will follow killer whales around to find food (Jourdain & Vongraven, 2017); thus, the humpbacks may simply follow the killer whales when they move away.

The dual process theory of habituation states that there are two observed behaviors after repeated exposure to a stimulus: habituation or sensitization (Groves & Thompson, 1970). Short sound pulses, with a short rise time at the startle threshold within hearing range of the animal have been shown to trigger a startle, which when triggered several times will likely lead to sensitization and avoidance behavior. However, in some cetacean species including killer whales, repeated acoustic stimuli has resulted in the animal moderating their hearing sensitivity (Nachtigall & Supin, 2013, 2014). This physiological response was detected in a study on the effect of TAST on sea lions, when short pulses were combined with longer pulses (Schakner et al., 2017). The short pulses led to sensitization, but when combined with a pure tone or longer sound pulse, it led to conditioned reduction in hearing sensitivity followed by habituation. In our study, the combination of short pulses, low duty cycle, and rapid onset was likely reducing the likelihood of such reductions in hearing sensitivity, as well as preventing or at least reducing the risk of habituation. The combination of short signals (0.2 seconds); mixed frequencies (low-frequency, high-frequency, and broadband); randomized intervals; and low duty cycles (0.4% of each signal type) likely induced sensitization and avoidance (Götz & Janik, 2011, 2015, 2016) and contributed to the behavioral response observed in the killer whales. The use of broadband signals may have improved the likelihood of eliciting a startle response as they have been shown to be more potent than pure tones and triggering responses at much lower amplitude and causing higher startle magnitudes (Götz et al., 2020).

A previous study testing the deterrence effect of an ADD on depredating killer whales within 50 meters from the sound source during longline fishery, was conducted in a similar time frame to our study (two and a half week) (Tixier, Gasco et al., 2015). The killer whales in this study were exposed to three different sound pulses (not based on the startle response) ranging from 50 ms to 1 s, given at both fixed and random intervals. Sounds were played in sequences for 20 and 15 minutes, with 15-minute breaks in between playbacks. The reduction effect was strong during the 3 first experiments (less than 10 playback sequences), but then after this the animals habituated to the sound and no effect was detected (Tixier, Gasco, et al., 2015). The rapid drop in effective deterrence was thought to be the result of inter-matriline social transfer of the behavior that can make habituation happen extremely rapid in a population (Ford et al.,1998 in Tixier, Gasco, et al., 2015). The ecological benefit of feeding is a strong motivation for sustaining hearing disturbances and was thought to have contributed to the observed habituation (Tixier, Gasco, et al., 2015). This study also differed in that the overall duty cycle was almost ~10 times higher than ours (Tixier, Gasco, et al., 2015), and shorter duty cycles are less likely to lead to habituation (Götz & Janik, 2015). Sounds with long rise times are less likely to trigger a startle response (Schakner et al., 2017), which again may have caused a quicker habituation process. The longer sounds may have led to learned behavior of reduced hearing sensitivity (Nachtigall & Supin, 2013) which would reduce the efficiency of the deterrent, similar to the study on sea lions (Schakner et al., 2017). For fish predators a successful ADD will have to cause a “perceived risk” sufficient to override the benefits of continuing the feeding (Schakner et.al 2013).

5.2 Recovery and attenuation

There was a significant and strong recovery of killer whales returning to the study area after playback ended, with a mean recovery of 50% of the whales within 5 minutes after playback was ended. The recovery response was strongest for zones B and D, which were the zones furthest away from the playback device, indicating that the whales returned earlier to the more distant areas at the end of the playbacks, but appeared to avoid the area closest to sound source for a bit longer. Although the recovery for killer whales was clear and significant, it was not a full recovery, as the surfacing rates remained lower than in the pre-playback period. It is worth noting that the rate of surfacings appeared to increase 1-2 minutes after playbacks ended (Figure 10), suggesting that a full recovery may have occurred if the post-playback observation period had been longer. The level of recovery suggests that it is unlikely that the animals moved away by chance or due to other factors such as the presence of the playback vessel.

Although it was not possible to show whether the same individual killer whales were deterred and then returned to the study area, the return of whales, generally, suggests that the effect is likely short-term and therefore less likely to result in habitat exclusion. This can be positive for future implementation of TAST in the NSS purse seine fishery, as we only want to deter the whales away from the closest vicinity of the net, during the most critical phase of the fishery, and not exclude them completely from the area.

The deterrence effect with killer whales was limited beyond 50 meters and in the two more distant zones. This indicated that that the amplitude at distances beyond 50 meters attenuated to the extent that the deterrence effect was substantially reduced. It could also be that the herring and the seine net was masking and attenuating the acoustic signal. It was difficult to say whether the sound level is reduced to the point that the whales do not respond at all, as we were not able to observe them at further distances due to the lack of light. It was almost impossible to observe beyond 100 meters, and most likely the reduced visibility made it more difficult (and therefore the distance observations were probably more inaccurate) to observe the distances from 50-100 meters during the experiments.

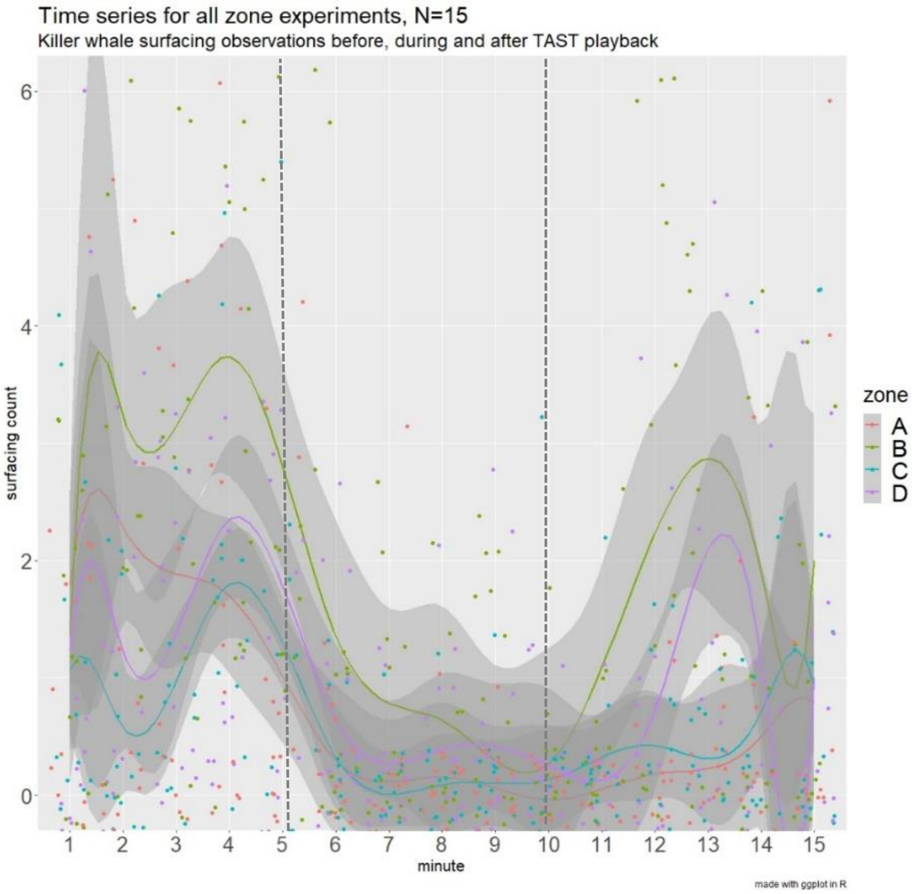


Figure 10: Time series illustration of all observations on killer whales and zones data, divided into 1 minute time intervals (15 minutes total). Surfacing counts are shown on the y-axis while each minute is represented on the x-axis. Playback starts at 5 minues and end at 10 minutes (marked with the dotted line). This figure is presented for illustrative purposes. A command was used when coding this figure in R (geom_jitter from the package ggplot), which avoided overlapping of datapoints. No analysis is performed on the figure. The dashed line indicates the start and end of the playback session. Each zone is given one color, and the grey areas represent the 95% Cis.

The rapid recovery we observed is contrary to many other studies which have shown that ADDs can lead to a full habitat exclusion of both target and non-target species (Johnston, 2002; Kastelein et al., 2005; Thompson et al., 2007; Götz & Janik, 2013). One study found full habitat exclusion by killer whales in response to an ADD (Morton & Symonds, 2002), however this ADD was not targeting the startle response. The ADD used for this study was originally designed to deter pinnipeds. Odontocete hearing is generally 15 to 30 dB more sensitive than pinniped hearing in the frequency band 4-40 kHz (Götz & Janik, 2013), and the perceived loudness may have been high enough for killer whales to exclude them from their habitat in this study (Morton & Symonds, 2002). Another study using TAST to deter porpoises showed that the animals would move away up to 3 km, with a mean distance of 1.78 km, with no re-sightings of porpoises within 1000 meters; however, habitat exclusion of the targeted species was the aim in this particular study (Hiley et al., 2021). Our results in this study may be explained by the sound pulses that were adjusted to the sensation level of 60-90 dB, above the acoustic threshold for killer whales. By staying within this sensation level, it is likely that we triggered the startle reflex, but not a strong enough reaction to trigger a full habitat exclusion. This is in fact the whole idea of the TAST system; that the sounds can be specifically tailored to the species of interest in each case, and the subsequent response magnitude of the species can be moderated by adjusting the sound signals. Our findings give further evidence that this technology does indeed work on killer whales.

5.3 Alternative mitigation methods

In the last years there has been an increase in interest of finding potential solutions that can mitigate and reduce negative operational interactions between fisheries and marine mammals, resulting in several literature reviews on existing peer-reviewed studies on the topic; However, none of these reviews have deemed ADDs to be efficient mitigation method for negative operational interactions (Northridge, 1991; Hamer et al., 2012; Hamilton & Baker, 2019; Lucas & Berggren, 2022;). There are a few other methods that have shown to be effective in mitigating interactions between purse seine fishery and killer whales (Tixier, Garcia et al., 2015 Tixier et al., 2019) but they necessitate changes in fishery practices, watchkeeping and avoidance (Tixier, Garcia et al., 2015). In the narrow and confined fjord systems of Norway, where the densities of whales and fishing vessels are high, these alternatives would be difficult to implement. Killer whales have shown to be attracted to fishing vessels from up to 20 kilometers away (Mul et al., 2020), which makes avoidance measures impossible. The modified fishery practices described by Tixier and Garcia et al. (2015) may be feasible for larger fishing vessels operating in the more open-ocean areas.

5.4 Study setup: strengths and limitations

Most experiments in this study were conducted during calm conditions, but a few times there were currents and wind that made it challenging to stay in position. As the season for the NSS herring fishery starts in November, we assumed that there would still be plenty of fishing vessels around when we started the field work in mid-November. However, it turned out that many of the fishing vessels had already filled their quota, and after the first week there were not many vessels left. The dark season in arctic regions of Northern Norway has a very short period of daylight, which was also a challenge as most experiments were conducted during the dark hours. The darkness made it somewhat challenging to accurately estimate the distances to each surfacing with a range finder due to low light. Video could also improve the accuracy of the distance estimations, and it was not possible to measure every surfacing with a range finder in real time. This led to many distance measurements being only approximate. We did one test-experiment onboard a larger fishing vessel, and experienced that the visibility range was even shorter when standing on board the fishing vessel compared to the small boat. Infrared cameras or binoculars would be useful tools in the future to observe surfacings in the dark

Ideally, we would conduct these experiments from onboard a fishing vessel, as this would allow more time to conduct the experiments and allow for longer exposure and recovery periods to further test for habituation. We were able to perform one 30-minute-long trial onboard a large fishing vessel as a pilot study, with good results. This was conducted by lowering the speaker 20 meters into the water column in the same way as we did from the small boat. In the future, testing from a fishing vessel could be conducted with this method. For our experiments in this study, we had to wait until the working lights on the fishing vessels to be turned to have enough visibility to detect the whales. However, the fishing vessels will only turn on the lights when the seine net is pursed up alongside the fishing vessel. When the net is already closed, the risk of entrapping a whale in the net is obviously reduced.

The observer effect in the distance measurement observations was larger than expected. This shows that the distance measurement data strongly depended on the observers' skills and ability to use a consistent method throughout all experiments. The surfacing count data with the zones, which only had one observer, had a lot more consistent data and was in the end more reliable than the distance measurements. If a future experiment will have more than one observer, training in a consistent method should be practiced beforehand.

The study was likely affected by the high degree of variation in conditions among the experimental trials. Each trial differed due to variation in fishing vessels, weather and light conditions, the number of whales present, the presence or absence of a pumping vessel, and the observer effect. Thorough record keeping ensured that these variables were detected, and included in the analysis and therefore can be accounted for in the models to the greatest extent possible.

Due to the study design as a controlled exposure field experiment, the control-trials were implemented in each experiment by observing for 5 minutes before playback. In a more ideal scenario, full control experiments would be conducted, where the same procedure was followed, with the speaker lowered into the water, but without sounds emitted. This would reveal any potential response to the playback vessel or speaker, or other unknown effects. On one occasion, due to a technical glitch in the playback device, we performed an accidental control experiment where we lowered the speaker into the water without playing any sound. There seemed to be a slight dip in the number of surfacings between minute 6 and 12, similar to the pattern in the actual experiments. More experiments and analysis would be required to investigate this, but it could be that the killer whales learned to recognize when the speaker was lowered into the water, as this was also one of the final experiments we attempted. full control experiments maybe be useful to investigate whether killer whale may sensitize to the presence of a speaker.

A strength to this study is that the setting of the NSS fishery provided a unique opportunity to perform a real-time controlled exposure experiments in the field with large numbers of whales and fishing vessels interacting with each other in the relatively confined area of a fjord. The close interaction with fishing vessels also created opportunities to observe, experience and document operational interactions. The fieldwork also created good opportunities to communicate and create connections with fishers. We were able to collaborate and work around their fishing operations and learn from them about their experiences with interactions with whales.

5.5 Conclusion

This study is the first of its kind to show a significant response to acoustic deterrence device by killer whales without leading to habitat exclusion. Despite the limited number of trials, the complexity of the data, with many random effects, the modelled effect for killer whales was strong and significant. These findings leave little room for speculation. Killer whales show a strong avoidance response to TAST and that response appears to be short-term, limiting the potential negative effects of deterrence.

The response for humpback whales was much less clear, but these results were also affected by a small sample size and most likely the sound source levels being insufficiently high. An effective method to deter humpback whales with TAST is still a work in progress, and further investigations regarding the hearing properties of humpback whales are needed. The project of which this study is a part, also aimed to find the startle response in humpback whales by tagging whales with a device that records video, audio, and movement of the whale and conducting focal follows where candidate sounds were tested. The results from these studies may be helpful in the future when conducting further testing on humpback whales with TAST.

We believe that TAST has great potential for practical application in the fjords of Northern Norway to reduce the risk of negative operational interactions with whales. Problems with negative operational interactions with cetaceans are also frequent in the fishing industry for capelin and mackerel, and also during the fishery at the herring spawning ground off Helgeland and Møre (Bjørge et al., 2022), and TAST can also be suited as a potential deterrence method here. It may also be implemented in aquaculture, where both seals and whales stay close to the fish farms to depredate, and occasionally become entangled in mooring devices, or destroy equipment and release fish (Bjørge et al., 2022). TAST as a seal deterrent device on fish farms has already been implemented in Scotland with promising results (Götz & Janik, 2016). Although the problems associated with marine mammal bycatch are recognized today, it is still an unresolved issue in many parts of the world (FAAO, 2021). The rate of negative operational interactions is likely to continue increasing worldwide, and a way forward is needed to mitigate the negative consequences of this. It is possible that TAST can be a future mitigation and can be implemented in other types of fisheries or industry in the future with proper adjustments and testing.

5.6 Future research and implementation

The critical phase in herring purse seine fishery, in which deterrence would be most useful in future applications, starts when the net is being set and lasts through the early phase of the hauling of the net while the seine is still open. Further trials will be necessary to test for an effect during this phase. Given that the effect of TAST attenuated quickly after 50 meters, it will be necessary to further explore the range of the effect and even consider multiple speakers. Investigating sound source pressure levels in order to elicit a startle response in the targeted species at some desired distance (e.g., 200 meters) could be interesting. Further testing directly from fishing vessels will be essential, as will consultation with fishers to inform future development. For TAST to be implemented as a tool in the fishery and used by fishers, it must be efficient, rugged, and easily deployed with minimal impact to fishing operations. There have also been concerns regarding ADD's and their potential effect on non-target species in different taxa such as fish. Previous studies with TAST have showed that the signals did not significantly exceed the hearing threshold of fish, which again suggest that adverse effects are unlikely (Götz & Janik, 2015), but further investigation on the potential effect on herring should be explored. This study is just the first step in a long process if TAST is to be implemented as a mitigation strategy in the NSS fishery, but so far, the results are promising.

6 Works cited

Alverson, D. L. (1994). *A global assessment of fisheries bycatch and discards* (FAO Fisheries Department, Issue. https://www.fao.org/3/T4890E/T4890E00.htm

Barton, K. (2023). *MuMIn: Multi-Model Inference*. In https://CRAN.R-project.org/package=MuMIn

Basran, C. J., Woelfing, B., Neumann, C., & Rasmussen, M. H. (2020). Behavioural Responses of Humpback Whales (*Megaptera novaeangliae*) to Two Acoustic Deterrent Devices in a Northern Feeding Ground off Iceland. *Aquatic mammals*, *46*(6), 584-602. https://doi.org/10.1578/AM.46.6.2020.584

Bjørge, A. (2022). Enkelte hvaler ser ut til å «spesialisere seg» på å beite rundt fiskebåter. *Fiskeribladet*. https://www.fiskeribladet.no/debatt/enkelte-hvaler-ser-ut-til-a-spesialisere-seg-pa-a-beite-rundt-fiskebater/2-1-1252475

Bjørge, A., Moan, A., Ryeng, K. A., & Wiig, J. R. (2022). Low anthropogenic mortality of humpback (*Megaptera novaeangliae*) and killer (*Orcinus orca*) whales in Norwegian purse seine fisheries despite frequent entrapments. *Marine Mammal Science*, *n/a*(n/a). https://doi.org/https://doi.org/10.1111/mms.12985

Bjørge, A., & Sivle, L. D. (2021). *Eksisterende teknologi for å holde hval og sel borte fra fiskeredskap og oppdrettsanlegg og sammenfatning av erfaringer i Norge og internasjonalt*. Havforskningsinstituttet.

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J.-S. S. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends Ecol Evol*, *24*(3), 127-135. https://doi.org/10.1016/j.tree.2008.10.008

Bonizzoni, S., Hamilton, S., Reeves, R. R., Genov, T., & Bearzi, G. (2022). Odontocete cetaceans foraging behind trawlers, worldwide. *Reviews in Fish Biology and Fisheries*, *32*(3), 827-877. https://doi.org/10.1007/s11160-022-09712-z

Branstetter, B. K., St. Leger, J., Acton, D., Stewart, J., Houser, D., Finneran, J. J., & Jenkins, K. (2017). Killer whale (*Orcinus orca*) behavioral audiograms. *The Journal of the Acoustical Society of America*, *141*(4), 2387-2398. https://doi.org/10.1121/1.4979116

Casini, M., Lövgren, J., Hjelm, J., Cardinale, M., Molinero, J.-C., & Kornilovs, G. (2008). Multi-level trophic cascades in a heavily exploited open marine ecosystem. *PROC R SOC B*, *275*(1644), 1793-1801. https://doi.org/10.1098/rspb.2007.1752

Cassoff, R., Moore, K., McLellan, W., Barco, S., Rotstein, D., & Moore, M. (2011). Lethal entanglement in baleen whales. *Diseases of Aquatic Organisms*, *96*(3), 175-185. https://doi.org/10.3354/dao02385

D'Agrosa, C., Lennert-Cody, C. E., & Vidal, O. (2000). Vaquita Bycatch in Mexico's Artisanal Gillnet Fisheries: Driving a Small Population to Extinction. *Conservation Biology*, *14*(4), 1110-1119. https://doi.org/10.1046/j.1523-1739.2000.98191.x

Dawson, S. M., Northridge, S., Waples, D., & Read, A. J. (2013). To ping or not to ping: the use of active acoustic devices in mitigating interactions between small cetaceans and gillnet fisheries. *Endangered species research*, *19*(3), 201-221. https://doi.org/10.3354/esr00464

Erbe, C. (2002). *Hearing Abilities of Baleen Whales*. Contractor report (W7707-01-0828) for Defence Research & Development Canada.

Escalle, L., Capietto, A., Chavance, P., Dubroca, L., Delgado De Molina, A., Murua, H., Gaertner, D., Romanov, E., Spitz, J., Kiszka, J. J., Floch, L., Damiano, A., & Merigot, B. (2015). Cetaceans and tuna purse seine fisheries in the Atlantic and Indian Oceans: interactions but few mortalities. *Marine Ecology Progress Series*, *522*, 255-268. https://www.int-res.com/abstracts/meps/v522/p255-268/

Estes, J. A., & Palmisano, J. F. (1974). Sea Otters: Their Role in Structuring Nearshore Communities. *Science*, *185*(4156), 1058-1060. https://doi.org/10.1126/science.185.4156.1058

FAAO, 2021. Fishing operations-guidelines to prevent and reduce bycatch of marine mammals in capture fisheries. *FAO Technical Guidelines for Responsible Fisheries 2021, Vol 1.* https://doi.org/10.4060/cb2887en

Findlay, C. R., Ripple, H. D., Coomber, F., Froud, K., Harries, O., Van Geel, N. C. F., Calderan, S. V., Benjamins, S., Risch, D., & Wilson, B. (2018). Mapping widespread and increasing underwater noise pollution from acoustic deterrent devices. *Mar Pollut Bull*, *135*, 1042-1050. https://doi.org/10.1016/j.marpolbul.2018.08.042

Fiskeridirektoratet, 2020. Handlingsplan for hendelser med hval. https://www.fiskeridir.no/Yrkesfiske/Tema/Sjoepattedyr/Hval/_/attachment/download/6e38294d-0fef-4164-ab6c-8d0cecd1324b:0910f58a9be3e9535425754264b9bfac5d09d8fd/handlingsplan-for-hendelser%20med-hval-revidert-04.03.2020.pdf

Fiskeridirektoratet, 2022. J-243-2022: Forskrift om midlertidig åpning for fiske etter norsk vårgytende sild med not innenfor fjordlinjene i Kvænangen og Reisafjorden i Troms

og Finnmark fylke, (2022). https://www.fiskeridirektoratet.no/Yrkesfiske/Regelverk-og-reguleringer/J-meldinger/Gjeldende-J-meldinger/j-243-2022
Gilman, E., Brothers, N., McPherson, G., & Dalzell, P. (2006). A review of cetacean interactions with longline gear. *The journal of cetacean research and management*, *8*(2), 215-223.
Gray, C. A., & Kennelly, S. J. (2018). Bycatches of endangered, threatened and protected species in marine fisheries. *Reviews in Fish Biology and Fisheries*, *28*(3), 521-541. https://doi.org/10.1007/s11160-018-9520-7
Groves, P. M., & Thompson, R. F. (1970). Habituation: A dual-process theory. *Psychol Rev*, *77*(5), 419-450. https://doi.org/10.1037/h0029810
Götz, T., & Janik, V. M. (2011). Repeated elicitation of the acoustic startle reflex leads to sensitisation in subsequent avoidance behaviour and induces fear conditioning. *BMC Neurosci*, *12*(1), 30-30. https://doi.org/10.1186/1471-2202-12-30
Götz, T., & Janik, V. M. (2013). Acoustic deterrent devices to prevent pinniped depredation: efficiency, conservation concerns and possible solutions. *Marine ecology. Progress series (Halstenbek)*, *492*, 285-302. https://doi.org/10.3354/meps10482
Götz, T., & Janik, V. M. (2015). Target-specific acoustic predator deterrence in the marine environment. *Anim Conserv*, *18*(1), 102-111. https://doi.org/10.1111/acv.12141
Götz, T., & Janik, V. M. (2016). Non-lethal management of carnivore predation: long-term tests with a startle reflex-based deterrence system on a fish farm. *Anim Conserv*, *19*(3), 212-221. https://doi.org/10.1111/acv.12248
Götz, T., Pacini, A. F., Nachtigall, P. E., & Janik, V. M. (2020). The startle reflex in echolocating odontocetes: basic physiology and practical implications. *J Exp Biol*, *223*(5). https://doi.org/10.1242/jeb.208470
Hamer, D. J., Childerhouse, S. J., & Gales, N. J. (2012). Odontocete bycatch and depredation in longline fisheries: A review of available literature and of potential solutions. *Marine Mammal Science*, *28*(4), E345-E374. https://doi.org/https://doi.org/10.1111/j.1748-7692.2011.00544.x
Hamer, D. J., Ward, T. M., & McGarvey, R. (2008). Measurement, management and mitigation of operational interactions between the South Australian Sardine Fishery and short-beaked common dolphins (*Delphinus delphis*). *Biological Conservation*, *141*(11), 2865-2878. https://doi.org/10.1016/j.biocon.2008.08.024
Hamilton, S., & Baker, G. B. (2019). Technical mitigation to reduce marine mammal bycatch and entanglement in commercial fishing gear: lessons learnt and future directions. *Reviews in Fish Biology and Fisheries*, *29*(2), 223-247. https://doi.org/10.1007/s11160-019-09550-6
Hartig, F. (2022). *DHARMa: Residual Diagnostics for Hierarchical (Multi-Level/Mixed) Regression Models*. In https://CRAN.R-project.org/package=DHARMa
Hiley, H. M., Janik, V. M., & Götz, T. (2021). Behavioural reactions of harbour porpoises *Phocoena phocoena* to startle-eliciting stimuli: movement responses and practical applications. *Marine Ecology Progress Series*, *672*, 223-241. https://doi.org/10.3354/meps13757
Hoef, J. M. V., & Boveng, P. L. (2007). Quasi-Poisson vs. Negative Binomial Regression: How Should We Model Overdispersed Count Data? *Ecology*, *88*(11), 2766-2772. https://doi.org/10.1890/07-0043.1
Huse, I., & Korneliussen, R. (2000). Diel variation in acoustic density measurements of overwintering herring (*Clupea harengus*). *ICES Journal of Marine Science*, *57*(4), 903-910. https://doi.org/10.1006/jmsc.2000.0577

Johnson, A. (2005). Fishing gear involved in entanglement of Right and Humpback whales. *Marine Mammal Science*, *21*(4), 10. https://bycatch.org/sites/default/files/Johnson_etal_2005_0.pdf

Johnston, D. W. (2002). The effect of acoustic harassment devices on harbour porpoises (*Phocoena phocoena*) in the Bay of Fundy, Canada. *Biological Conservation*, *108*(1), 113-118. https://doi.org/10.1016/S0006-3207(02)00099-X

Jourdain, E., & Vongraven, D. (2017). Humpback whale (Megaptera novaeangliae) and killer whale (*Orcinus orca*) feeding aggregations for foraging on herring (*Clupea harengus*) in Northern Norway. *Mammalian biology : Zeitschrift für Säugetierkunde*, *86*(1), 27-32. https://doi.org/10.1016/j.mambio.2017.03.006

Kastelein, R. A., Verboom, W. C., Muijsers, M., Jennings, N. V., & Van der Heul, S. (2005). The influence of acoustic emissions for underwater data transmission on the behaviour of harbour porpoises (*Phocoena phocoena*) in a floating pen. *Mar Environ Res*, *59*(4), 287-307. https://doi.org/10.1016/j.marenvres.2004.05.005

Koch, M., & Schnitzler, H.U. (1997). The acoustic startle response in rats—circuits mediating evocation, inhibition and potentiation. *Behavioural Brain Research*, *89*(1), 35-49. https://doi.org/https://doi.org/10.1016/S0166-4328(97)02296-1

Lenth, R. V. (2022). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. In https://CRAN.R-project.org/package=emmeans

Lewison, R. L., Crowder, L. B., Read, A. J., & Freeman, S. A. (2004). Understanding impacts of fisheries bycatch on marine megafauna. *Trends in Ecology & Evolution*, *19*(11), 598-604. https://doi.org/https://doi.org/10.1016/j.tree.2004.09.004

Lien, J., & Aldrich, D. (1982). *Damage to inshore fishing gear in Newfoundland and Labrador by whales and sharks during 1981*. https://dx.doi.org/10.5281/zenodo.5914697

Lindbæk, E. (2022). Sildefiskere hevder bestanden av spekkhoggere har eksplodert. Havforsker er ikke enig. *Fiskeribladet*. https://www.fiskeribladet.no/fiskeri/sildefiskere-hevder-bestanden-av-spekkhoggere-har-eksplodert-havforsker-er-ikke-enig-/2-1-1170970

Sjøveisreglene, (1975). Forskrift om forebygging av sammenstøt på sjøen. https://lovdata.no/dokument/SF/forskrift/1975-12-01-5/KAPITTEL

Lucas, S., & Berggren, P. (2022). A systematic review of sensory deterrents for bycatch mitigation of marine megafauna. *Reviews in Fish Biology and Fisheries*. https://doi.org/10.1007/s11160-022-09736-5

Lüdecke, D. (2018). *ggeffects: Tidy Data Frames of Marginal Effects from Regression Models*. In Journal of Open Source Software, vol 3(26).

Martinussen, T. M. (2022). Rekord på Kvænangen - aldri før er det låssatt så mye sild på en uke. *Fiskeribladet*. https://www.fiskeribladet.no/fiskeri/rekord-pa-kvanangen-aldri-for-er-det-lassatt-sa-mye-sild-pa-en-uke/2-1-1358222

Moore, M. J., & Van der Hoop, J. M. (2012). The Painful Side of Trap and Fixed Net Fisheries: Chronic Entanglement of Large Whales. *Journal of Marine Biology, 2012*, 1-4. https://doi.org/10.1155/2012/230653

Moore, M. J., Van der Hoop, J. M., Barco, S. G., Costidis, A. M., Gulland, F. M., Jepson, P. D., Moore, K. T., Raverty, S., McLellan, W. A. (2013). Criteria and case definitions for serious injury and death of pinnipeds and cetaceans caused by antrhopogenic trauma. *Diseases of Aquatic Organisms, 103(*3*)*, 229-264. https://dx.doi.org/10.3354/dao02566

Moore, M. J. (2014). How we all kill whales. *ICES Journal of Marine Science, 71*(4), 760-763. https://doi.org/10.1093/icesjms/fsu008

Morton, A. B., & Symonds, H. K. (2002). Displacement of Orcinus orca (L.) by high amplitude sound in British Columbia, Canada. *ICES Journal of Marine Science*, *59*(1), 71-80. https://doi.org/10.1006/jmsc.2001.1136
Mul, E., Blanchet, M., McClintock, B., Grecian, W., Biuw, M., & Rikardsen, A. (2020). Killer whales are attracted to herring fishing vessels. *Marine Ecology Progress Series*, *652*, 1-13. https://doi.org/10.3354/meps13481
Nachtigall, P. E., & Supin, A. Y. (2013). A false killer whale reduces its hearing sensitivity when a loud sound is preceded by a warning. *J Exp Biol*, *216*(Pt 16), 3062-3070. https://doi.org/10.1242/jeb.085068
Nakagawa, S., Johnson, P. C. D., & Schielzeth, H. (2017). The coefficient of determination R(2) and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *J R Soc Interface*, *14*(134). https://doi.org/10.1098/rsif.2017.0213
Norsk Sildesalgslag. (2022). *Årsrapport 2021*. 32 pages. https://www.sildelaget.no/media/172663376/aarsmelding-2021-final-spread.pdf
Northridge, S. (1991). An updated world review of interactions between marine mammals and fisheries. *FAO Fisheries Technical Paper*, *251*, 58.
Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-Effect Models in S and S-PLUS*. In *Springer, esp. pp 100,461*
Quick, N. J., Middlemas, S. J., & Armstrong, J. D. (2004). A survey of antipredator controls at marine salmon farms in Scotland. *Aquaculture*, *230*(1), 169-180. https://doi.org/10.1016/S0044-8486(03)00428-9
Read, A. J., Drinker, P., & Northridge, S. (2006). Bycatch of Marine Mammals in U.S. and Global Fisheries. *Conservation Biology*, *20*(1), 163-169. https://doi.org/https://doi.org/10.1111/j.1523-1739.2006.00338.x
Reeves, R. R., McClellan, K., & Werner, T. B. (2013). Marine mammal bycatch in gillnet and other entangling net fisheries, 1990 to 2011. *Endangered species research*, *20*(1), 71-97. https://doi.org/10.3354/esr00481
Rikardsen, A.H. (2019). *Winter Whales* (Vol. 1). To-foto, Harstad, Norway. 150 pages. ISBN13-9788293178163
Robbins, J. (2012). *Scar-Based Inference Into Gulf of Maine Humpback Whale Entanglement: 2010* Provincetown, Massachusetts]. Provincetown.
Robbins, J., Knowlton, A. R., & Landry, S. (2015). Apparent survival of North Atlantic right whales after entanglement in fishing gear. *Biological Conservation*, *191*, 421-427. https://doi.org/https://doi.org/10.1016/j.biocon.2015.07.023
Salthaug, A., & Stenevik, E. K. (2020). Kommer silda inn i fjordene i Nord-Norge i vinter? https://www.hi.no/hi/nyheter/2020/november/kommer-silda-inn-i-fjordene-i-nord-norge-i-vinter
Schakner, Z. A., & Blumstein, D. T. (2013). Behavioral biology of marine mammal deterrents: A review and prospectus. *Biological Conservation*, *167*, 380-389. https://doi.org/10.1016/j.biocon.2013.08.024
Schakner, Z. A., Götz, T., Janik, V. M., & Blumstein, D. T. (2017). Can fear conditioning repel California sea lions from fishing activities? *Animal conservation*, *20*(5), 425-432. https://doi.org/10.1111/acv.12329
Shivik, J. A., Treves, A., & Callahan, P. (2003). Nonlethal techniques for managing predation: Primary and secondary repellents : Human-Carnivore Conflict: Local Solutions with Global Applications. *Conservation Biology*, *17*(6), 1531-1537.
Smith, G. M., Ieno, E. N., Saveliev, A. A., Walker, N., & Zuur, A. F. (2009). *Mixed Effects Models and Extensions in Ecology with R* (1. Aufl. ed.). New York, NY: Springer-Verlag.

Southall, B. L., Bowles, A. E., Ellison, W. T., Finneran, J. J., Gentry, R. L., Greene, C. R., Kastak, D., Ketten, D. R., Miller, J. H., Nachtigall, P. E., Richardson, W. J., Thomas, J. A., & Tyack, P. L. (2007). Marine Mammal Noise-Exposure Criteria: Initial Scientific Recommendations. *Aquatic mammals*, *33*(4), 411-509.

Terborgh, J., Lopez, L., Nuñez, P., Rao, M., Shahabuddin, G., Orihuela, G., Riveros, M., Ascanio, R., Adler, G. H., Lambert, T. D., & Balbas, L. (2001). Ecological Meltdown in Predator-Free Forest Fragments. *Science*, *294*(5548), 1923-1926. https://doi.org/10.1126/science.1064397

Thompson, P. M., Mackey, B., Barton, T. R., Duck, C., & Butler, J. R. A. (2007). Assessing the potential impact of salmon fisheries management on the conservation status of harbour seals (*Phoca vitulina*) in north-east Scotland. *Animal conservation*, *10*(1), 48-56. https://doi.org/10.1111/j.1469-1795.2006.00066.x

Tixier, P., Authier, M., Gasco, N., & Guinet, C. (2015). Influence of artificial food provisioning from fisheries on killer whale reproductive output. *Animal conservation*, *18*(2), 207-218. https://doi.org/https://doi.org/10.1111/acv.12161

Tixier, P., Burch, P., Richard, G., Olsson, K., Welsford, D., Lea, M.-A., Hindell, M. A., Guinet, C., Janc, A., Gasco, N., Duhamel, G., Villanueva, M. C., Suberg, L., Arangio, R., Söffker, M., & Arnould, J. P. Y. (2019). Commercial fishing patterns influence odontocete whale-longline interactions in the Southern Ocean. *Scientific Reports*, *9*(1), 1904. https://doi.org/10.1038/s41598-018-36389-x

Tixier, P., Gasco, N., Duhamel, G., & Guinet, C. (2015). Habituation to an acoustic harassment device (AHD) by killer whales depredating demersal longlines. *ICES Journal of Marine Science*, *72*(5), 1673-1681. https://doi.org/10.1093/icesjms/fsu166

Tixier, P., Vacquie Garcia, J., Gasco, N., Duhamel, G., & Guinet, C. (2015). Mitigating killer whale depredation on demersal longline fisheries by changing fishing practices. *ICES Journal of Marine Science*, *72*(5), 1610-1620. https://doi.org/10.1093/icesjms/fsu137

Todd, S., Lien, J., & Verhulst, A. (1992). Orientation of humpback whales (*Megaptera novaeangliae*) and minke whales (*Balaenoptera acutorostrata*) to acoustic alarm devices designed to reduce entrapments in fishing gear. In J. A. Thomas, R. A. Kastelien, & A. Y. Supin (Eds.), *Marine Mammal Sensory Systems* (pp. pp. 727-739). Plenum Press.

Towers, J. R., Tixier, P., Ross, K. A., Bennett, J., Arnould, J. P. Y., Pitman, R. L., & Durban, J. W. (2019). Movements and dive behaviour of a toothfish-depredating killer and sperm whale. *ICES Journal of Marine Science*, *76*(1), 298-311. https://doi.org/10.1093/icesjms/fsy118

Tyack, P., Gordon, J., & Thompson, D. (2003). Controlled Exposure Experiments to Determine the Effects of Noise on Marine Mammals. *Marine Technology Society journal*, *37*(4), 41-53. https://doi.org/10.4031/002533203787537087

Van der Hoop, J., Corkeron, P., & Moore, M. (2017). Entanglement is a costly life-history stage in large whales. *Ecol Evol*, *7*(1), 92-106. https://doi.org/10.1002/ece3.2615

Van der Hoop, J., Moore, M. J., Barco, S. G., Cole, T. V. N., Daoust, P.-Y., Henry, A. G., Mcalpine, D. F., McLellan, W. A., Wimmer, T., & Solow, A. R. (2013). Assessment of Management to Mitigate Anthropogenic Effects on Large Whales. *Conservation Biology*, *27*(1), 121-133. https://doi.org/https://doi.org/10.1111/j.1523-1739.2012.01934.x

Van Opzeeland, I. C., Corkeron, P. J., Teo, L., Tiu, S., & Van Parijs, S. M. (2005). Acoustic Behaviour of Norwegian Killer Whales, Orcinus orca, During Carousel and Seiner Foraging on Spring-Spawning Herring. *Aquatic mammals*, *31*(1), 110. https://doi.org/10.1578/AM.31.1.2005.110

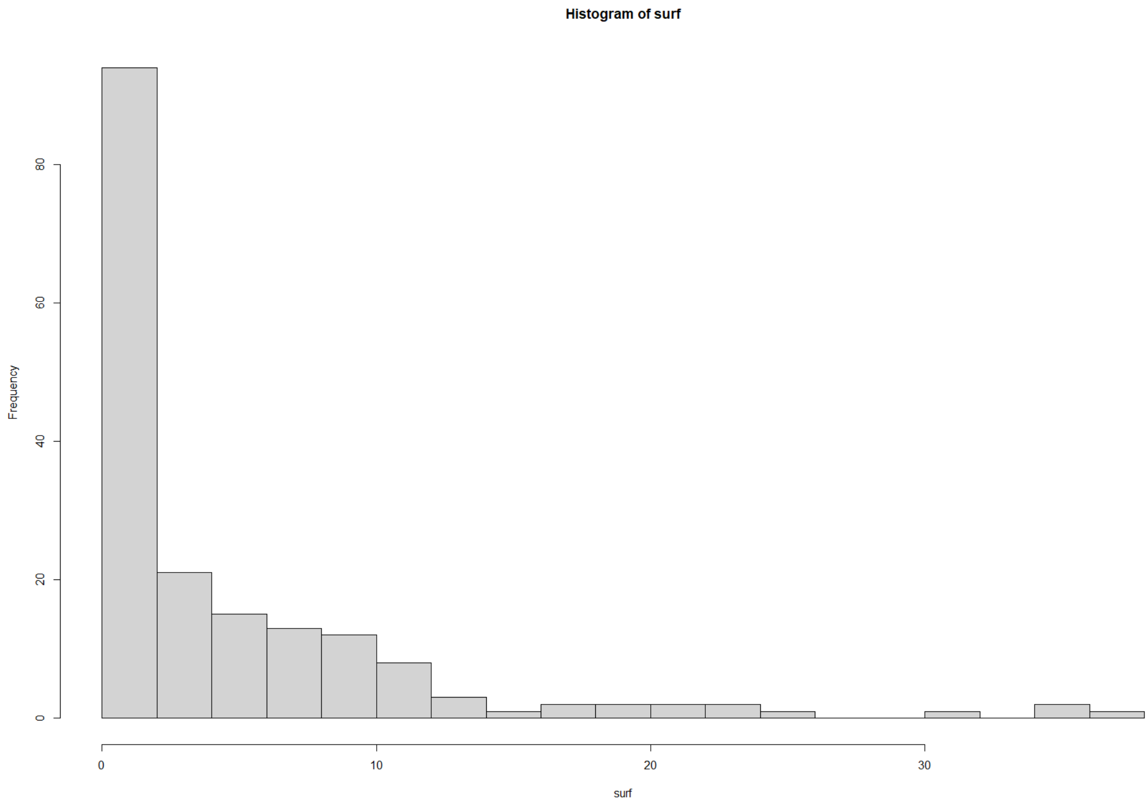
Volgenau, L., Kraus, S., & Lien, J. (2011). Erratum: The impact of entanglements on two substocks of the western North Atlantic humpback whale, Megaptera novaeangliae. *Canadian Journal of Zoology*, *73*, 1689-1698. https://doi.org/10.1139/z95-201

Yodzis, P. (2001). Must top predators be culled for the sake of fisheries? *Trends Ecol Evol*, *16*(2), 78-84. https://doi.org/10.1016/S0169-5347(00)02062-0

Zeileis, A., & Hothorn, T. (2002). *Diagnostic Checking in Regression Relationship*. In R News. https://CRAN.R-project.org/doc/Rnews/

Appendix 1 – Killer whales and zones

A - Model selection



Histogram of data distribution (counts of observed surfacings per treatment period per zone)

Step 1: Selecting for error distribution family (Poisson, negative binomial 1 or 2) and and the zero-inflation argument. Table created with MuMIn (Barton, 2023)

```
Model selection table
   cnd((Int)) zi      family                random df   logLik   AICc  delta weight
m3   1.955             n1(lg)    c(t)+c(v)+c(l)+c(p) 17 -417.315  872.4   0.00  0.220
m4   1.978      -3.165 n1(lg)    c(t)+c(v)+c(l)+c(p) 18 -416.214  872.7   0.27  0.192
m6f  1.774      -2.407 n2(lg)  c(t/z)+c(v)+c(l)+c(p) 19 -415.170  873.1   0.68  0.156
m5   1.813             n2(lg)    c(t)+c(v)+c(l)+c(p) 17 -417.853  873.5   1.08  0.128
m4d  1.916      -2.808 n1(lg)  c(t/z)+c(v)+c(l)+c(p) 19 -415.661  874.1   1.66  0.096
m6   1.874      -2.769 n2(lg)    c(t)+c(v)+c(l)+c(p) 18 -416.995  874.2   1.83  0.088
m3c  1.955             n1(lg)  c(t/z)+c(v)+c(l)+c(p) 18 -417.315  874.9   2.47  0.064
m5e  1.753             n2(lg)  c(t/z)+c(v)+c(l)+c(p) 18 -417.448  875.1   2.74  0.056
m2b  1.709      -1.946 ps(lg)  c(t/z)+c(v)+c(l)+c(p) 18 -438.426  917.1  44.69  0.000
m1a  1.605             ps(lg)  c(t/z)+c(v)+c(l)+c(p) 17 -482.086 1002.0 129.54  0.000
m2   2.093      -1.705 ps(lg)    c(t)+c(v)+c(l)+c(p) 17 -486.820 1011.4 139.01  0.000
m1   1.984             ps(lg)    c(t)+c(v)+c(l)+c(p) 16 -526.799 1088.9 216.53  0.000
Abbreviations:
 zi: zero inflation argument
 family: n1(lg) = 'nbinom1(log)', n2(lg) = 'nbinom2(log)', ps(lg) = 'poisson(log)'
Models ranked by AICc(x)
Random terms:
 c(t)   : cond(1 | trial)
 c(v)   : cond(1 | vessel)
 c(l)   : cond(1 | light)
 c(p)   : cond(1 | pump)
 c(t/z): cond(1 | trial/zones)

 Note: After addressing a minor auto-correlation problem in the residuals post
 diagnostics, only the models where zones nested within treatment were assessed
 (model m1,m2,m3,m4,m5 and m6 were therefore not included). This random effect
 specification
 acknowledges that zones were not absolute.
```

```
Likelihood ratio test

Model 1: surf ~ treat * zones + (1 | trial/zones) + (1 | vessel) + (1 |
    light) + (1 | pump)
Model 2: surf ~ treat * zones + (1 | trial/zones) + (1 | vessel) + (1 |
    light) + (1 | pump)
  #Df  LogLik Df  Chisq Pr(>Chisq)
1  19 -415.17
2  19 -415.66  0 0.9824  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Step 2: Select random effects

```
Model selection table 2

    cnd((Int)) zi((Int)) random effect          df   logLik  AICc delta weight
m14      1.716    -2.425 c(t/z)                 16 -415.647 866.6  0.00  0.414
m11      1.774    -2.407 c(t/z)+c(v)            17 -415.170 868.1  1.49  0.197
m12      1.716    -2.425 c(t/z)+c(l)            17 -415.647 869.1  2.44  0.122
m13      1.716    -2.425 c(t/z)+c(p)            17 -415.647 869.1  2.44  0.122
m8       1.774    -2.407 c(t/z)+c(v)+c(l)       18 -415.170 870.6  3.96  0.057
m9       1.716    -2.425 c(t/z)+c(l)+c(p)       18 -415.647 871.5  4.91  0.036
m10      1.716    -2.425 c(t/z)+c(p)+c(l)       18 -415.647 871.5  4.91  0.036
m7       1.774    -2.407 c(t/z)+c(v)+c(l)+c(p) 19 -415.170 873.1  6.46  0.016
Models ranked by AICc(x)
Random terms:
 c(t/z): cond(1 | trial/zones)
 c(v)  : cond(1 | vessel)
 c(l)  : cond(1 | light)
 c(p)  : cond(1 | pump)
```

Best choice: m14 with zones nested within trial as random effects.

```
Likelihood ratio test

Model 1: surf ~ treat * zones + (1 | trial/zones)
Model 2: surf ~ treat * zones + (1 | trial/zones) + (1 | vessel)
  #Df  LogLik Df Chisq Pr(>Chisq)
1  16 -415.65
2  17 -415.17  1 0.954     0.3287
>
```

Step 3: Select fixed effects

```
Model selection table
m13: surf~treat*zones+1|trial/zones
m14: surf~treat+zones+1|trial/zones
m15: surf~treat+1|trial/zones

    cnd((Int)) zi((Int)) df   logLik  AICc delta weight
m14      1.397    -2.911 10 -421.545 864.4  0.00  0.754
m13      1.716    -2.425 16 -415.647 866.6  2.24  0.246
m15      1.994    -2.473  7 -437.676 890.0 25.61  0.000
Models ranked by AICc(x)
(all models):
Zero inflation argument
Family= negative binomial 2

m14 and m13 used in results
```

```
Likelihood ratio test

Model 1: surf ~ treat + zones + (1 | trial/zones)
Model 2: surf ~ treat * zones + (1 | trial/zones)
  #Df  LogLik Df  Chisq Pr(>Chisq)
1  10 -421.54
2  16 -415.65  6 11.796    0.06669 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

B - Model specifics

Expected surfacing estimates based on model m13(interaction model):

```
                                        2.5 %    97.5 %   Estimate
cond.(Intercept)                   3.21854341 9.6124551 5.56220317
cond.treatduring                   0.03211866 0.1831063 0.07668852
cond.treatpost                     0.10995009 0.4232529 0.21572365
cond.zonesB                        1.24997772 4.2497868 2.30480777
cond.zonesC                        0.46149153 1.6825455 0.88118130
cond.zonesD                        0.70711249 2.5053101 1.33099064
cond.treatduring:zonesB            1.04128831 8.2179023 2.92527019
cond.treatpost:zonesB              1.51448442 8.4152504 3.56998118
cond.treatduring:zonesC            0.35432206 4.0406824 1.19653789
cond.treatpost:zonesC              0.87578347 5.8597398 2.26536162
cond.treatduring:zonesD            1.02747030 8.9880227 3.03890217
cond.treatpost:zonesD              1.11628379 6.7800531 2.75108403
zi.(Intercept)                     0.03228748 0.2426470 0.08851248
cond.Std.Dev.(Intercept)|zones:trial 1.26146621 2.1217702 1.51892943
cond.Std.Dev.(Intercept)|trial     1.39781823 2.5178967 1.74388813
```

Expected surfacing estimates based on model m14(global effect model):

```
                                       2.5 %    97.5 %   Estimate
cond.(Intercept)                  2.43580329 6.7148787 4.04427045
cond.treatduring                  0.09901087 0.2168437 0.14652604
cond.treatpost                    0.31661173 0.6751853 0.46235441
cond.zonesB                       2.48960964 6.9166694 4.14967551
cond.zonesC                       0.66373787 1.8901634 1.12007725
cond.zonesD                       1.29572228 3.6852726 2.18519790
zi.(Intercept)                    0.00814666 0.3636731 0.05443089
cond.Std.Dev.(Intercept)|zones:trial 1.17549233 2.2957131 1.44276839
cond.Std.Dev.(Intercept)|trial    1.42171149 2.6040566 1.78658677
```

A

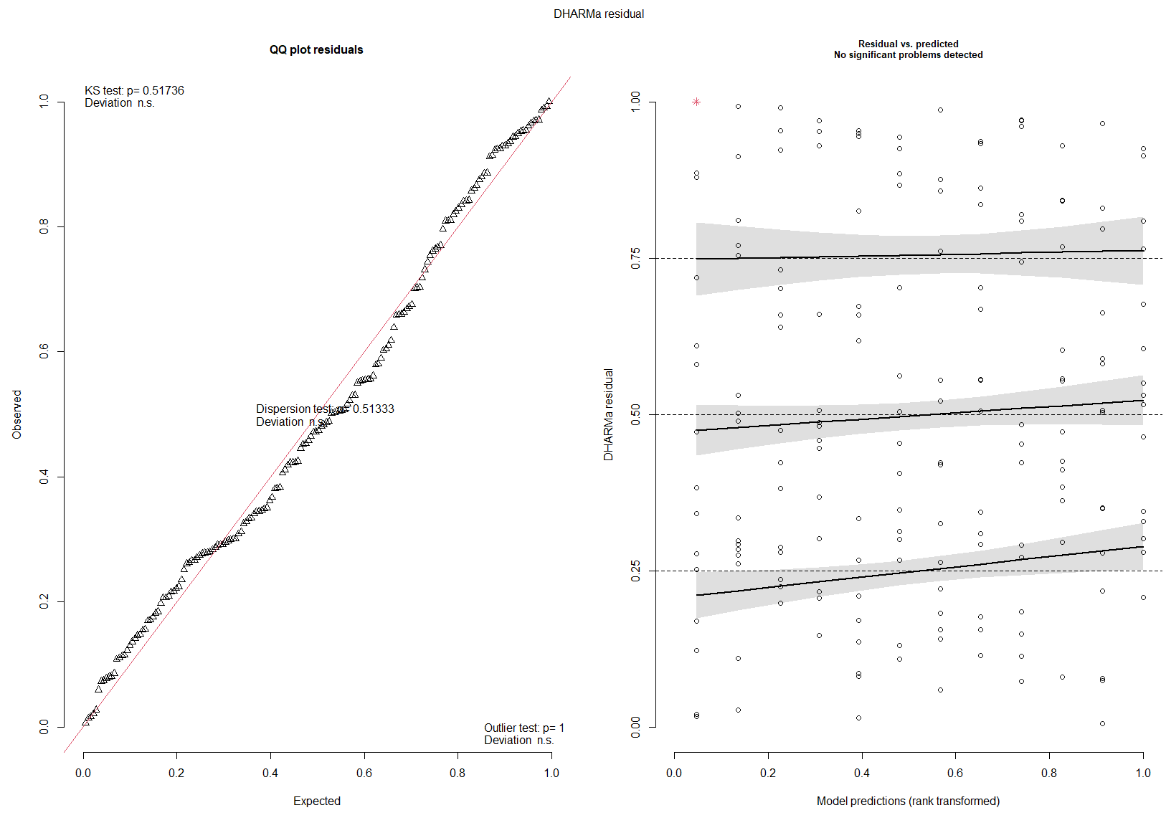
	GLMM with interaction term(m13)	
Predictors	*Estimate*	*p*
Count Model		
(Intercept)	5.56	**<0.001**
treat [during]	0.08	**<0.001**
treat [post]	0.22	**<0.001**
zones [B]	2.30	**0.007**
zones [C]	0.88	0.701
zones [D]	1.33	0.376
treat [during] × zones [B]	2.93	**0.042**
treat [post] × zones [B]	3.57	**0.004**
treat [during] × zones [C]	1.20	0.773
treat [post] × zones [C]	2.27	0.092
treat [during] × zones [D]	3.04	**0.045**
treat [post] × zones [D]	2.75	**0.028**
(Intercept)	23.20	
Zero-Inflated Model		
(Intercept)	0.09	**<0.001**
Random Effects		
σ^2	0.80	
$\tau_{00\ zones:trial}$	0.17	
$\tau_{00\ trial}$	0.31	
N_{zones}	4	
N_{trial}	15	
Marginal R^2 / Conditional R^2	0.475 / 0.672	
AIC	863.294	

B

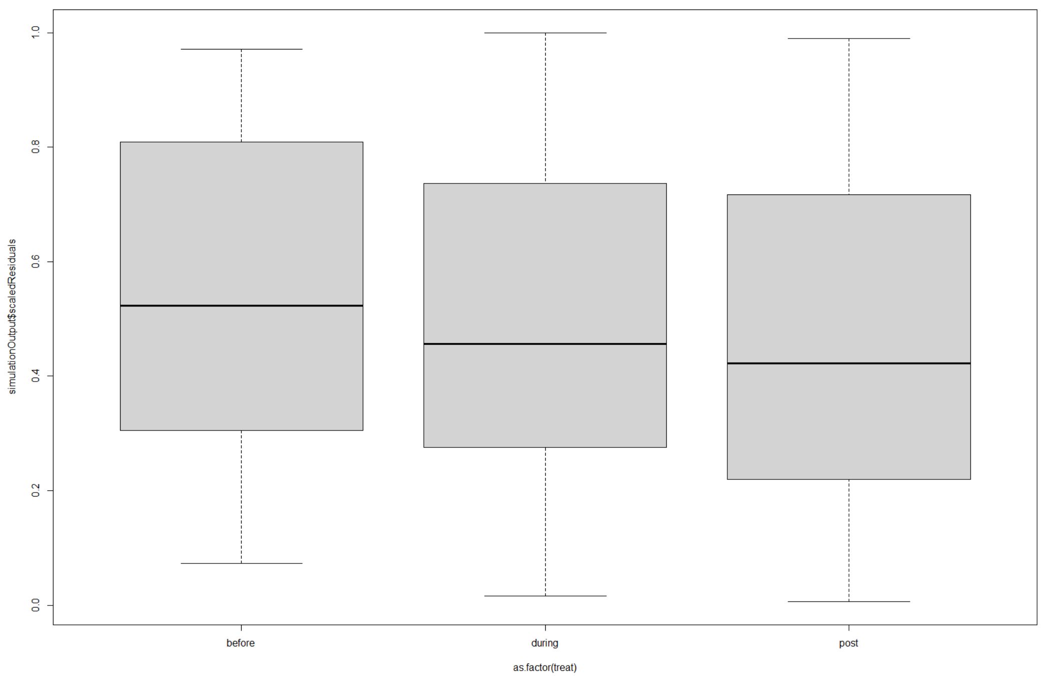
	GLMM with global effect(m14)	
Predictors	*Estimate*	*p*
Count Model		
(Intercept)	4.04	**<0.001**
treat [during]	0.15	**<0.001**
treat [post]	0.46	**<0.001**
zones [B]	4.15	**<0.001**
zones [C]	1.12	0.671
zones [D]	2.19	**0.003**
(Intercept)	7.83	
Zero-Inflated Model		
(Intercept)	0.05	**0.003**
Random Effects		
σ^2	1.03	
$\tau_{00\ zones:trial}$	0.13	
$\tau_{00\ trial}$	0.34	
N_{zones}	4	
N_{trial}	15	
Marginal R^2 / Conditional R^2	0.390 / 0.582	
AIC	863.090	

Summary of the generalized linear mixed models A) m13a (given as m13) and B) m14a (given as m14). Both models examine changes in the number of surfacings as a function of treatment (before, during, after) and zones. To the left, model M13 is with interaction between zones and treatment, while m14 to the right is without interaction and look at the global effect. The predictors are the treatment types before, during and post, and the different zones A, B, C and D. Significant values are highlighted in bold.

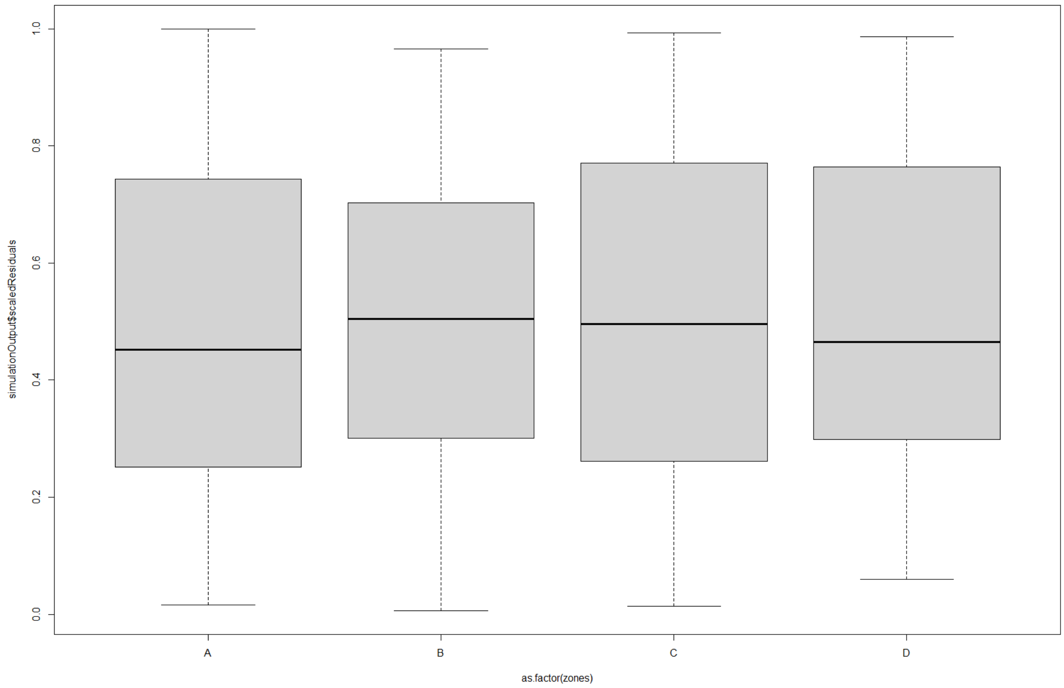
C – Model diagnostics



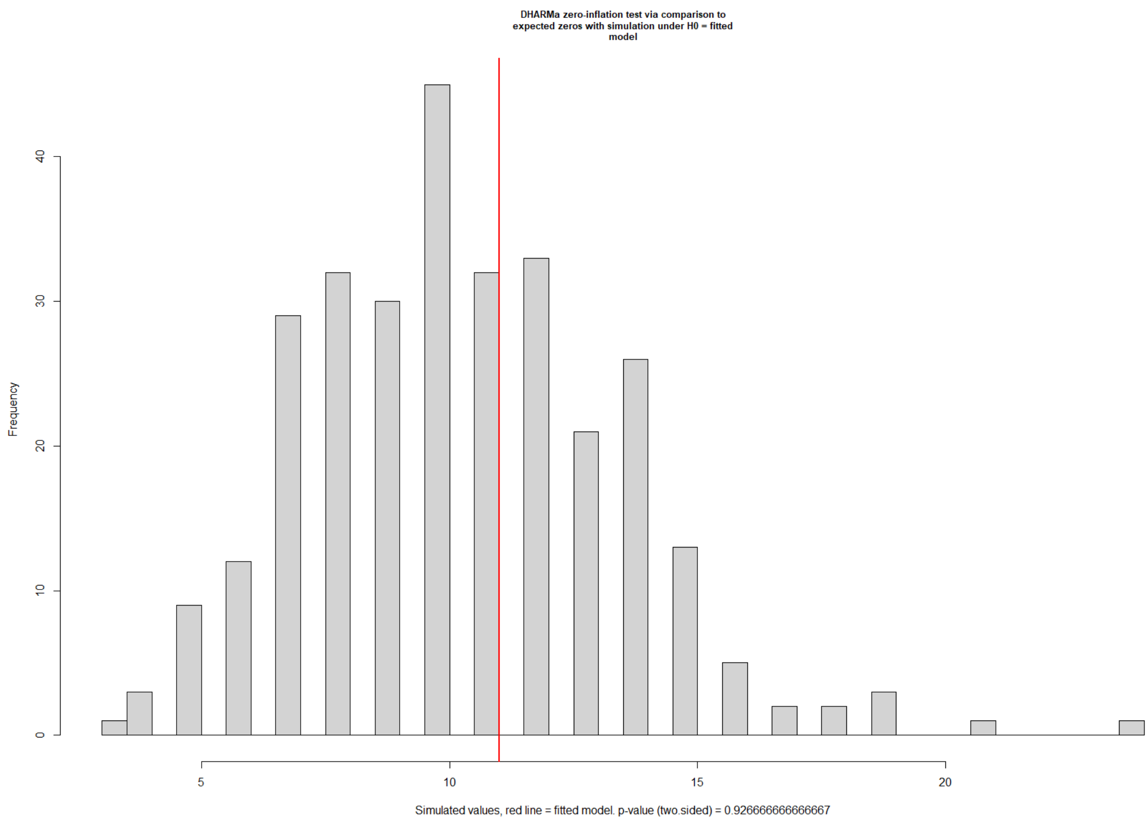
QQplot of residuals. Looks good.



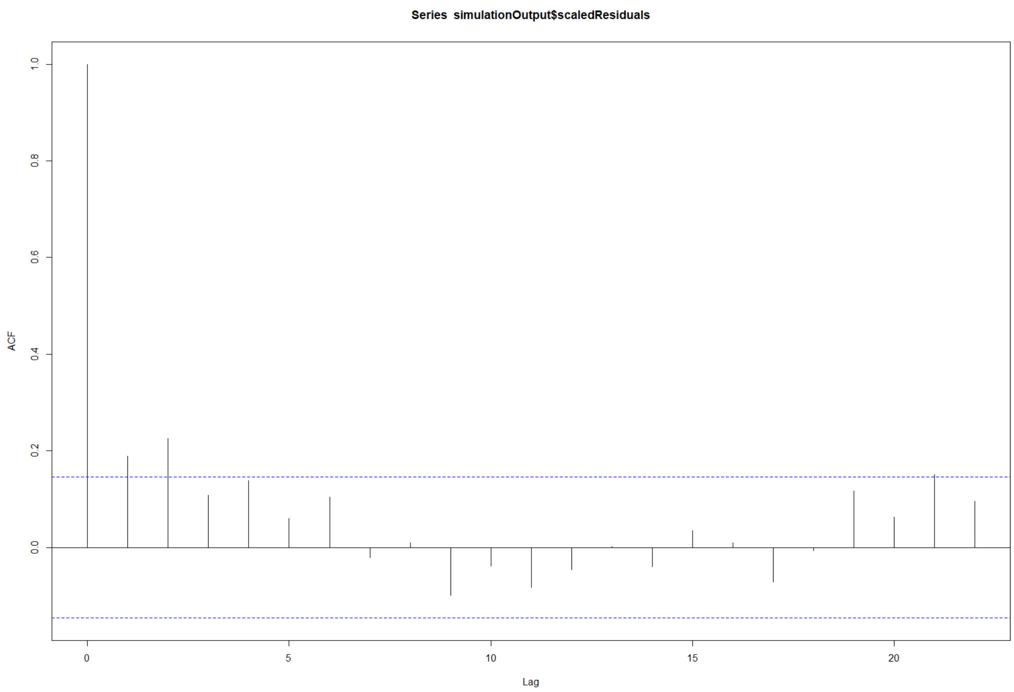
Residuals VS predicted values (treatment). Looks good.



Residuals vs predicted values zones. Looks good.



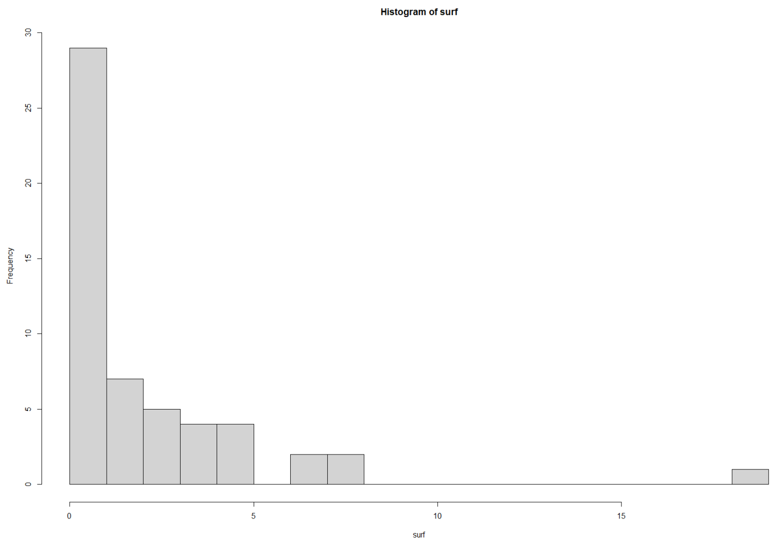
No zero-inflation problem detected.



ACF output of the scaled residuals.

Appendix 2 – Humpback whales and zones

A – Model selection



Histogram of data distribution (counts of observed surfacings per treatment period per zone)

Step 1: Error distribution and zero inflation argument with two possible random effect specifications based on trial design

```
Model selection table
cnd((Int)) zi   family  random              df   logLik  AICc delta

m3   0.4301          n1(lg)   c(t)+c(v)+c(l)+c(p) 11  -98.233 224.8  0.00
m5   0.3815          n2(lg)   c(t)+c(v)+c(l)+c(p) 11  -99.681 227.6  2.90
m3c  0.4301          n1(lg) c(t/z)+c(v)+c(l)+c(p) 12  -98.233 228.1  3.32
m5e  0.3792          n2(lg) c(t/z)+c(v)+c(l)+c(p) 12  -99.679 231.0  6.22
m6   0.3815  -19.870 n2(lg)   c(t)+c(v)+c(l)+c(p) 12  -99.681 231.0  6.22
m4d  0.4301  -21.060 n1(lg) c(t/z)+c(v)+c(l)+c(p) 13  -98.233 231.6  6.81
m1   0.3230          ps(lg)   c(t)+c(v)+c(l)+c(p) 10 -103.552 232.2  7.47
m1a  0.3243          ps(lg) c(t/z)+c(v)+c(l)+c(p) 11 -103.343 235.0 10.22
m2   0.4377  -2.621  ps(lg)   c(t)+c(v)+c(l)+c(p) 11 -103.354 235.0 10.24
m4   0.4301  -20.700 n1(lg)   c(t)+c(v)+c(l)+c(p) 12
m2b  0.3243  -15.680 ps(lg) c(t/z)+c(v)+c(l)+c(p) 12
m6f  0.3815  -18.670 n2(lg) c(t/z)+c(v)+c(l)+c(p) 13

Abbreviations:
zi: zero inflation argument
 family: n1(lg) = ‘nbinom1(log)’, n2(lg) = ‘nbinom2(log)’, ps(lg) = ‘poi
sson(log)’
Models ranked by AICc(x)
Random terms:
 c(t)   : cond(1 | trial)
 c(v)   : cond(1 | vessel)
 c(l)   : cond(1 | light)
 c(p)   : cond(1 | pump)
 c(t/z): cond(1 | trial/zones)
```

```
Likelihood ratio test

Model 1: surf ~ treat * zones + (1 | trial) + (1 | vessel) + (1 | light) +
    (1 | pump)
Model 2: surf ~ treat * zones + (1 | trial) + (1 | vessel) + (1 | light) +
    (1 | pump)
  #Df  LogLik Df  Chisq Pr(>Chisq)
1  11 -98.233
2  11 -99.681  0 2.8953  < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Step 2: Random effects selection

```
Model selection table
    cnd((Int)) random                df  logLik  AICc delta weight
m14     0.4301   c(t)                 8 -98.233 215.7  0.00  0.534
m12     0.4301   c(t)+c(l)            9 -98.233 218.6  2.89  0.126
m10     0.4301   c(t)+c(v)            9 -98.233 218.6  2.89  0.126
m11     0.4301   c(t)+c(p)            9 -98.233 218.6  2.89  0.126
m13     0.4301   c(t)+c(l)+c(p)      10 -98.233 221.6  5.92  0.028
m8      0.4301   c(t)+c(v)+c(p)      10 -98.233 221.6  5.92  0.028
m9      0.4301   c(t)+c(v)+c(l)      10 -98.233 221.6  5.92  0.028
m7      0.4301   c(t)+c(v)+c(l)+c(p) 11 -98.233 224.8  9.09  0.006
Models ranked by AICc(x)
Random terms:
 c(t): cond(1 | trial)
 c(l): cond(1 | light)
 c(v): cond(1 | vessel)
 c(p): cond(1 | pump)
```

```
Likelihood ratio test

Model 1: surf ~ treat * zones + (1 | trial)
Model 2: surf ~ treat * zones + (1 | trial) + (1 | light)
Model 3: surf ~ treat * zones + (1 | trial) + (1 | vessel)
Model 4: surf ~ treat * zones + (1 | trial) + (1 | pump)
  #Df  LogLik Df Chisq Pr(>Chisq)
1   8 -98.233
2   9 -98.233  1     0     0.9999
3   9 -98.233  0     0     <2e-16 ***
4   9 -98.233  0     0     <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Step 3: Fixed effects

```
 Model selection table

m13: surf~treat*zones+1|trial
m14: surf~treat+zones+1|trial
m15: surf~treat+1|trial


    cnd((Int)) df   logLik  AICc delta weight
 m14     0.5002  6  -98.724 211.2  0.00  0.755
 m15     0.8713  5 -101.526 214.3  3.07  0.163
 m13     0.4301  8  -98.233 215.7  4.43  0.082
 Models ranked by AICc(x)
  (all models):
   No zero inflation argument
   Family: negative binomial 1


 m14 used in results
```

```
Likelihood ratio test

Model 1: surf ~ treat + zones + (1 | trial)
Model 2: surf ~ treat + (1 | trial)
  #Df   LogLik Df  Chisq Pr(>Chisq)
1   6  -98.724
2   5 -101.526 -1 5.6022    0.01794 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

B – Model specifics

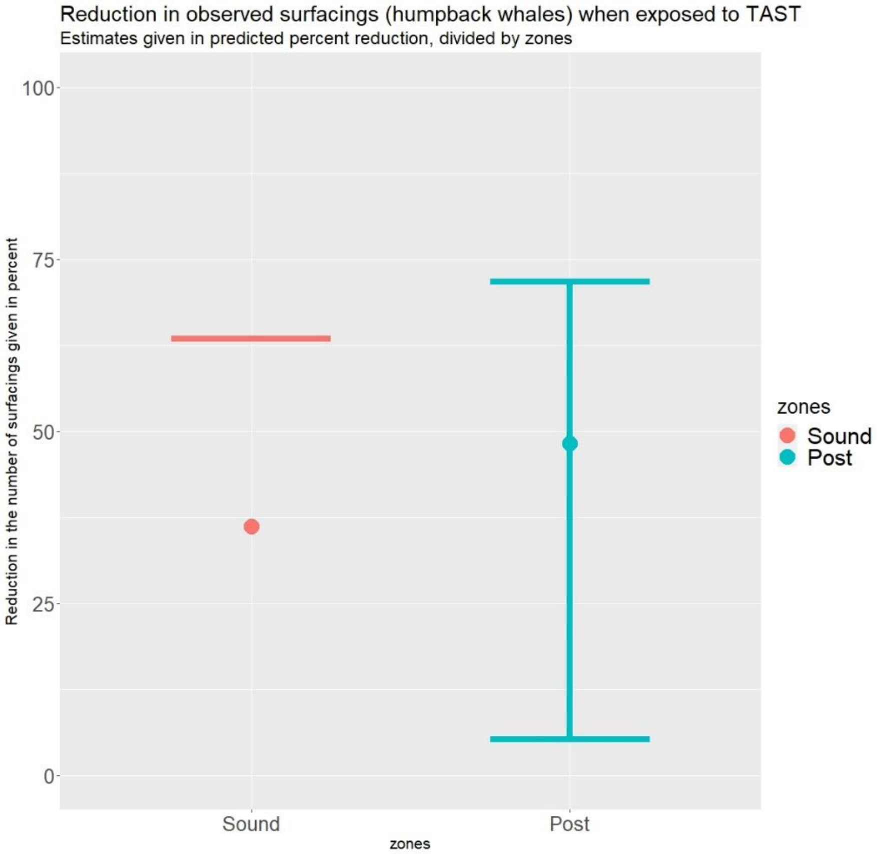
	GLMM for global effect(m14)	
Predictors	*Estimate*	*p*
(Intercept)	1.65	0.186
treat [during]	0.64	0.113
treat [post]	0.52	**0.033**
zones [BD]	1.83	**0.016**
Random Effects		
σ^2	0.73	
$\tau_{00\ trial}$	0.64	
ICC	0.46	
N_{trial}	9	
Marginal R^2 / Conditional R^2	0.111 / 0.523	

Summary of model m14b

Table of model contrasts in model m14b:

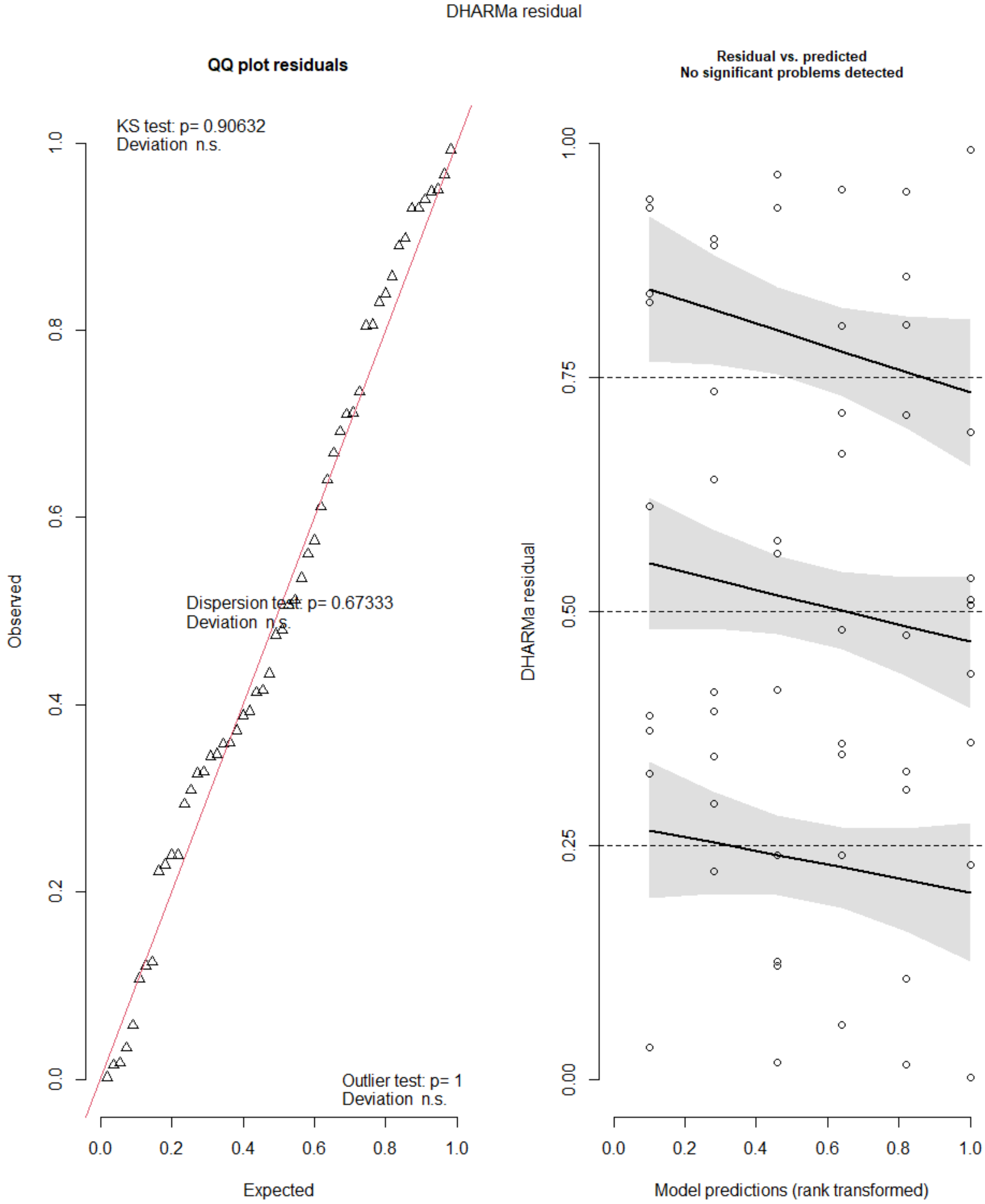
contrast	zones	ratio	SE	df	asymp.LCL	asymp.UCL	null	z.ratio	p.value
during / before	AC	0.6381178	0.1809106	Inf	0.3283461	1.240137	1	-1.584557	0.252
post / before	AC	0.5176299	0.1597709	Inf	0.2511015	1.067062	1	-2.133408	0.083
post / during	AC	0.8111824	0.2696390	Inf	0.3722037	1.767895	1	-0.629545	0.804
during / before	BD	0.6381178	0.1809106	Inf	0.3283461	1.240137	1	-1.584557	0.252
post / before	BD	0.5176299	0.1597709	Inf	0.2511015	1.067062	1	-2.133408	0.083
post / during	BD	0.8111824	0.2696390	Inf	0.3722037	1.767895	1	-0.629545	0.804

P value adjustment: tukey method for comparing all estimates.

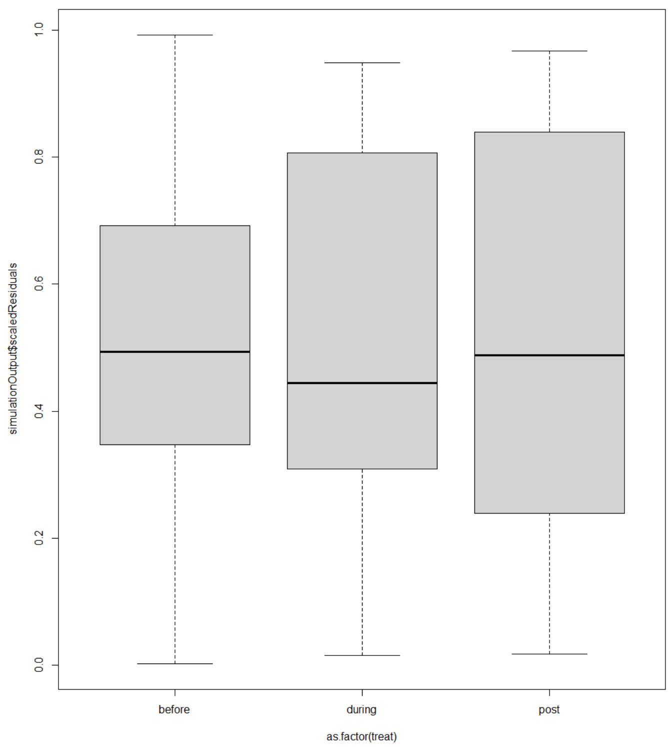


Expected reduction in surfacings given in percentage, compared to pre-playback. Red dot represents the expected reduction from pre-PB to playback, while the blue dot represents the expected reduction from pre-PB to post-PB.

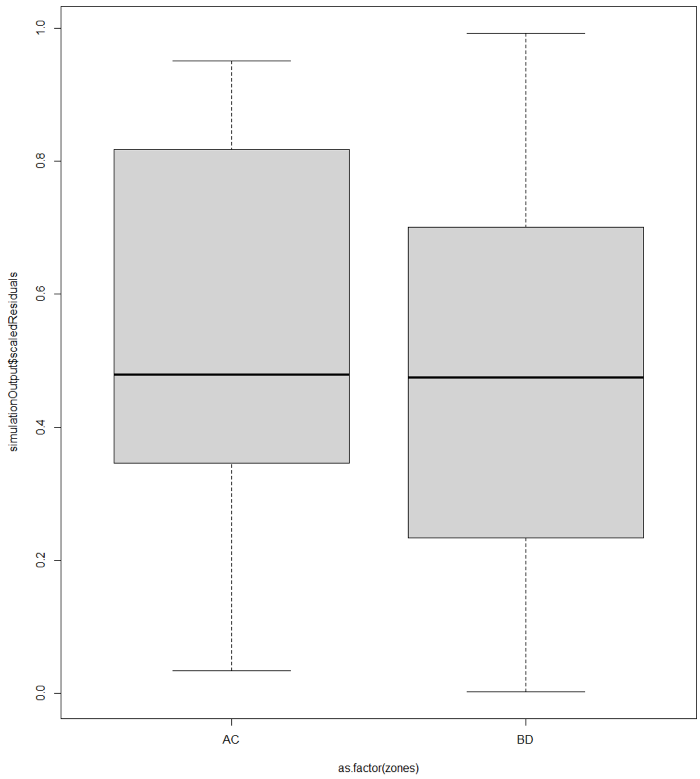
C – Model diagnostics



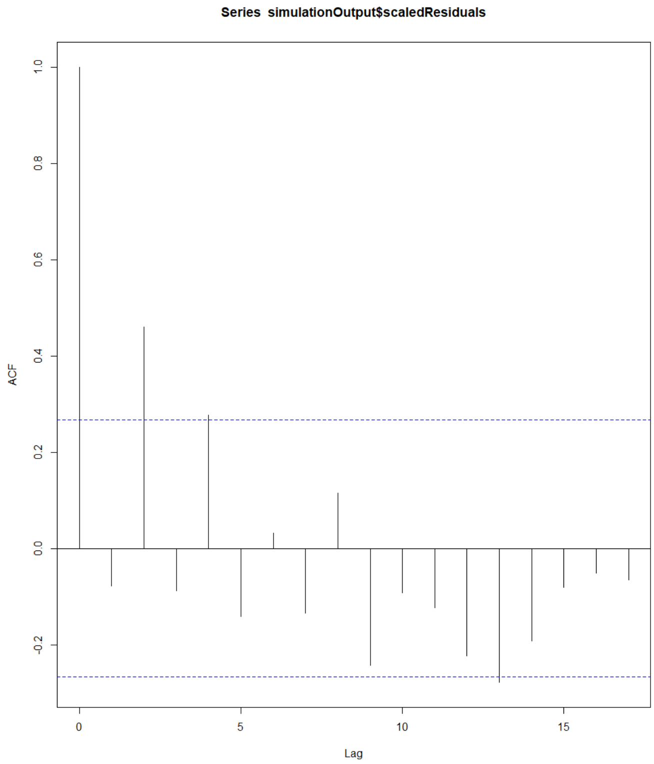
QQplot of scaled residuals.



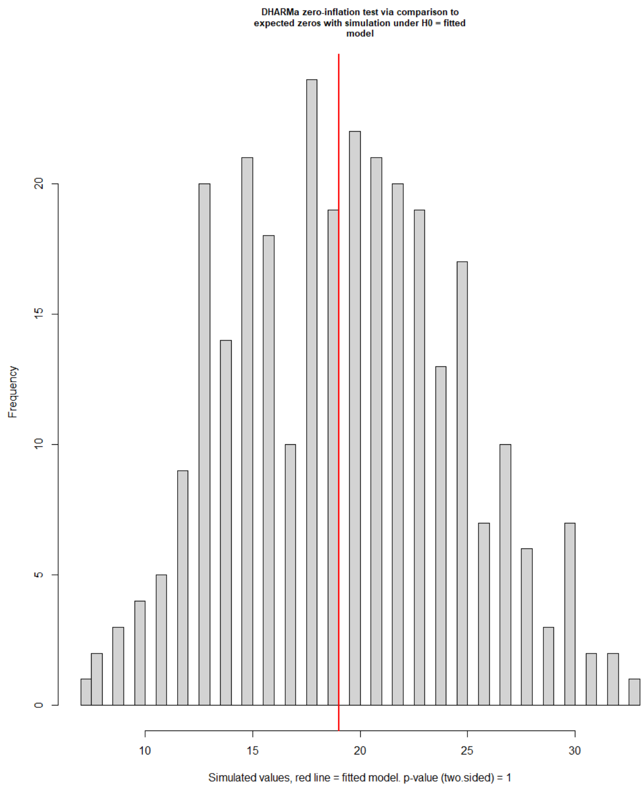
Plot of scaled residuals VS predicted values (treatment).



Plot of scaled residuals vs predicted values (zones)



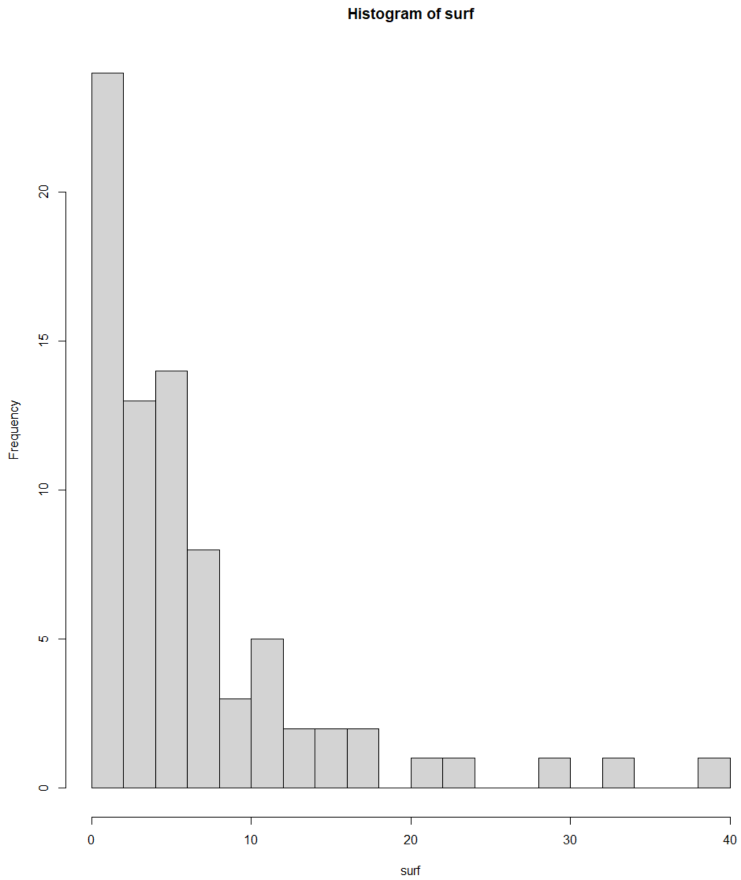
ACF of the scaled residuals.



Test for zero inflation argument- no zero-inflation problem detected.

Appendix 3 – Killer whales and distances

A – Model selection



Histogram of data distribution (counts of observed surfacings per treatment period per distance bin)

Step 1: Error distribution and zero-inflation argument

```
Model selection table
    cnd    zi      family   random                        df   AIC   delta
m3c  2.414         n1(lg) c(t/d)+c(v)+c(l)+c(p)+c(o)  13  463.5   0.00
m4d  2.414 -20.930 n1(lg) c(t/d)+c(v)+c(l)+c(p)+c(o)  14  466.5   2.98
m5   2.50          n2(lg)   c(t)+c(v)+c(l)+c(p)+c(o)  12  466.6   3.08
m3   2.483         n1(lg)   c(t)+c(v)+c(l)+c(p)+c(o)  12  466.9   3.32
m5e  2.414         n2(lg) c(t/d)+c(v)+c(l)+c(p)+c(o)  13  466.9   3.40
m6   2.515 -3.656  n2(lg)   c(t)+c(v)+c(l)+c(p)+c(o)  13  469.4   5.84
m6f  2.452 -3.100  n2(lg) c(t/d)+c(v)+c(l)+c(p)+c(o)  14  469.7   6.14
m2b  2.372 -2.338  ps(lg) c(t/d)+c(v)+c(l)+c(p)+c(o)  13  495.8   32.31
m1a  2.319         ps(lg) c(t/d)+c(v)+c(l)+c(p)+c(o)  12  502.3   38.73
m2   2.531 -2.198  ps(lg)   c(t)+c(v)+c(l)+c(p)+c(o)  12  543.6   80.10
m1   2.497         ps(lg)   c(t)+c(v)+c(l)+c(p)+c(o)  11  560.3   96.74
m4   2.483 -20.480 n1(lg)   c(t)+c(v)+c(l)+c(p)+c(o)  13
Abbreviations:
zi: zero inflation argument
 family: n1(lg) = 'nbinom1(log)', n2(lg) = 'nbinom2(log)', ps(lg) = 'poisson(log)'
Models ranked by AICc(x)
Random terms:
 c(t/d): cond(1 | trial/dist)
 c(v)  : cond(1 | vessel)
 c(l)  : cond(1 | light)
 c(p)  : cond(1 | pump)
 c(o)  : cond(1 | obs)
 c(t)  : cond(1 | trial)
```

```
Likelihood ratio test

Model 1: surf ~ treat * dist + (1 | trial/dist) + (1 | vessel) + (1 |
    light) + (1 | pump) + (1 | obs)
Model 2: surf ~ treat * dist + (1 | trial/dist) + (1 | vessel) + (1 |
    light) + (1 | pump) + (1 | obs)
  #Df  LogLik Df Chisq Pr(>Chisq)
1  13 -215.93
2  14 -215.93  1     0     0.9999
```

```
No difference (in lrt) between model 1 and 2, but models 3 and 4 show reduc
ed fit:
3  12 -218.91  -2   5.9716  0.0505 .
4  12 -219.03   0   0.2373  <2e-16 ***
```

Step 2: Random effects selection

```
Model selection table
    cnd((Int)) random                              df   logLik  AICc delta weight
m23       2.41  c(t/d)                              9 -215.925 452.5  0.00  0.367
m13       2.414 c(t/d)+c(p)                        10 -215.925 455.1  2.64  0.098
m11       2.414 c(t/d)+c(v)                        10 -215.925 455.1  2.64  0.098
m17       2.414 c(t/d)+c(o)                        10 -215.925 455.1  2.64  0.098
m12       2.414 c(t/d)+c(l)                        10 -215.925 455.1  2.64  0.098
m14       2.414 c(t/d)+c(p)+c(v)                   11 -215.925 457.9  5.35  0.025
m18       2.414 c(t/d)+c(p)+c(o)                   11 -215.925 457.9  5.35  0.025
m8        2.414 c(t/d)+c(v)+c(l)                   11 -215.925 457.9  5.35  0.025
m16       2.414 c(t/d)+c(v)+c(l)                   11 -215.925 457.9  5.35  0.025
m10       2.414 c(t/d)+c(p)+c(l)                   11 -215.925 457.9  5.35  0.025
m15       2.414 c(t/d)+c(p)+c(l)                   11 -215.925 457.9  5.35  0.025
m9        2.414 c(t/d)+c(l)+c(p)                   11 -215.925 457.9  5.35  0.025
m19       2.414 c(t/d)+c(o)+c(l)                   11 -215.925 457.9  5.35  0.025
m22       2.414 c(t/d)+c(v)+c(o)                   11 -215.925 457.9  5.35  0.025
m21       2.414 c(t/d)+c(v)+c(p)+c(o)              12 -215.925 460.7  8.15  0.006
m20       2.414 c(t/d)+c(v)+c(l)+c(o)              12 -215.925 460.7  8.15  0.006
m7        2.414 c(t/d)+c(v)+c(l)+c(p)+c(o)         13 -215.925 463.5 11.04  0.001
Models ranked by AICc(x)
Random terms:
 c(t/d): cond(1 | trial/dist)
 c(p)  : cond(1 | pump)
 c(v)  : cond(1 | vessel)
 c(o)  : cond(1 | obs)
 c(l)  : cond(1 | light)


Best selection was m23 with distance bin nested within each trial as rand
om effect.
```

```
Likelihood ratio test

Model 1: surf ~ treat * dist + (1 | trial/dist)
Model 2: surf ~ treat * dist + (1 | trial/dist) + (1 | pump)
  #Df  LogLik Df Chisq Pr(>Chisq)
1   9 -215.93
2  10 -215.93  1     0     0.9999
```

Step 3: Fixed effects selection

```
Model selection table
m1: surf~treat*dist+1|trial/dist
m2: surf~treat+dist+1|trial/dist

  cnd((Int)) df   logLik  AICc delta weight
m1      2.414  9 -215.925 452.5   0.0  0.538
m2      2.295  7 -218.600 452.8   0.3  0.462
Models ranked by AICc(x)
(all models):
No zero inflation argument
Family= negative binomial 1
```

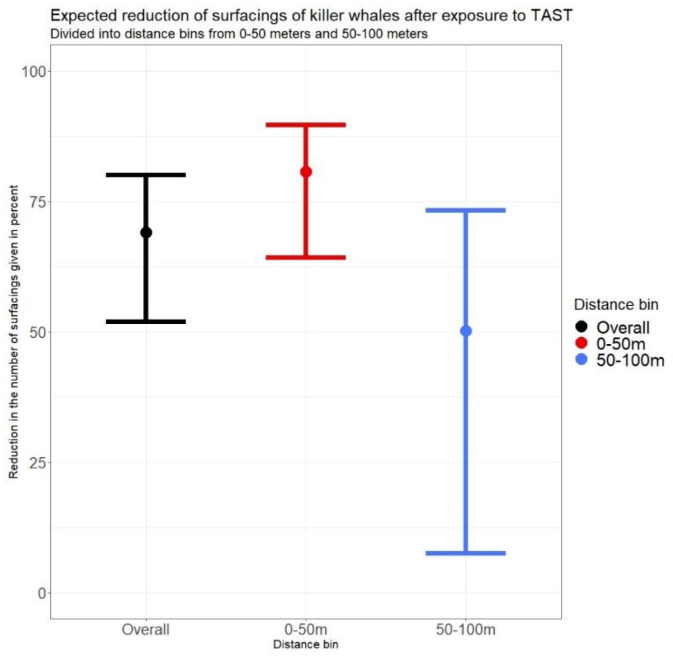
```
Likelihood ratio test

Model 1: surf ~ treat * dist + (1 | trial/dist)
Model 2: surf ~ treat + dist + (1 | trial/dist)
  #Df  LogLik Df Chisq Pr(>Chisq)
1   9 -215.93
2   7 -218.60 -2 5.349    0.06894 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

B Model specifics

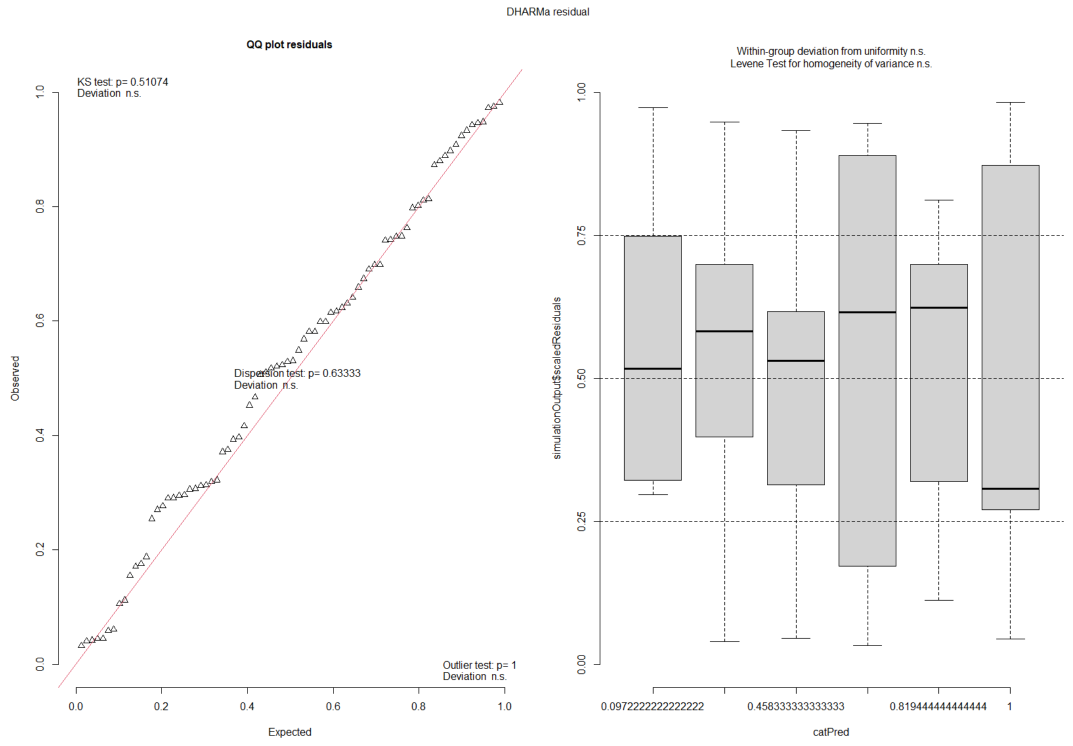
	m1 with interaction term		m2 for global effect	
Predictors	*Estimate*	*p*	*Estimate*	*p*
(Intercept)	11.18	**<0.001**	9.92	**<0.001**
treat [during]	0.19	**<0.001**	0.30	**<0.001**
treat [post]	0.54	**0.006**	0.68	**0.032**
dist [51-100]	0.56	0.051	0.79	0.341
treat [during] × dist [51-100]	2.58	**0.030**		
treat [post] × dist [51-100]	1.73	0.113		
Random Effects				
σ^2	0.40		0.45	
τ_{00}	0.24 $_{dist:trial}$		0.23 $_{dist:trial}$	
	0.06 $_{trial}$		0.04 $_{trial}$	
N	2 $_{dist}$		2 $_{dist}$	
	13 $_{trial}$		13 $_{trial}$	
Marginal R^2 / Conditional R^2	0.289 / 0.591		0.278 / 0.548	
AIC	449.851		451.200	

Summary of model m1 and m2.

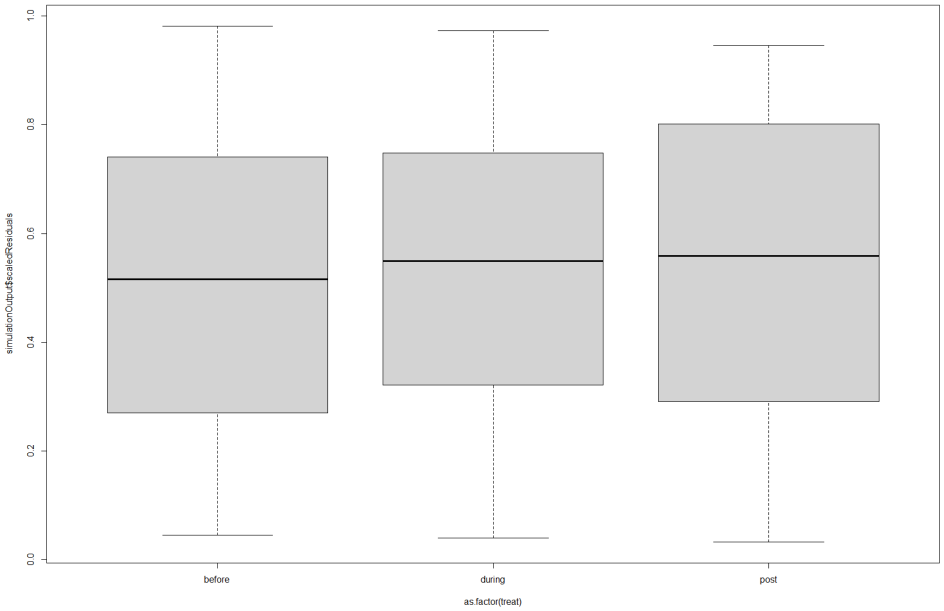


Expected reduction in surfacings from the pre-PB period to the playback period given in percentage reduction.

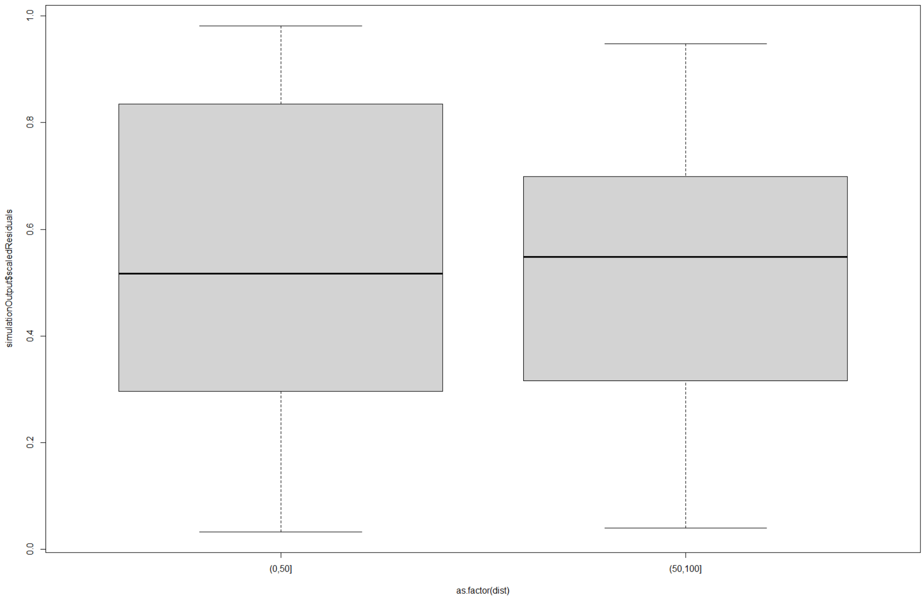
C Model diagnostics



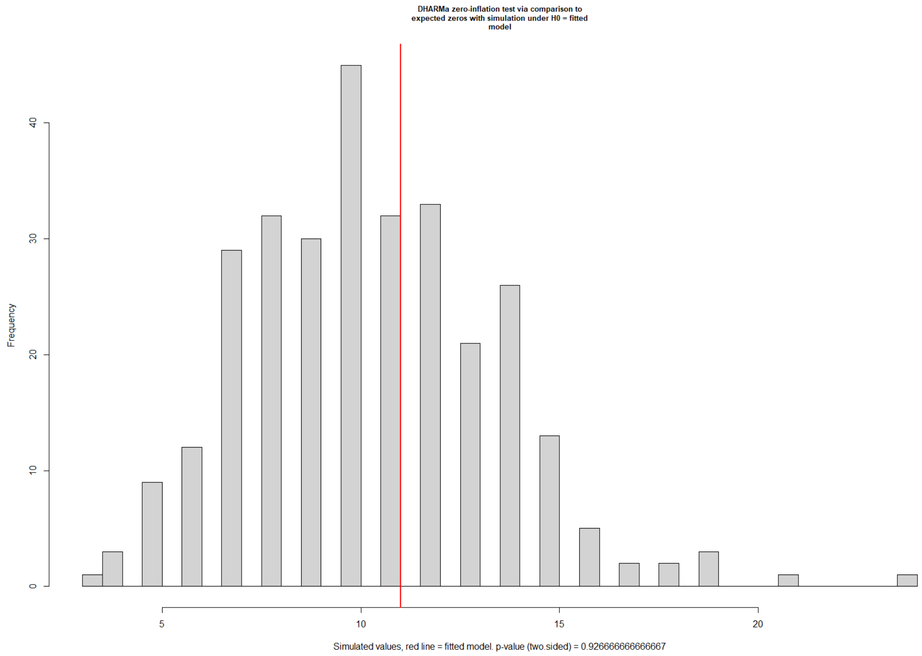
QQ-plot. Looks good.



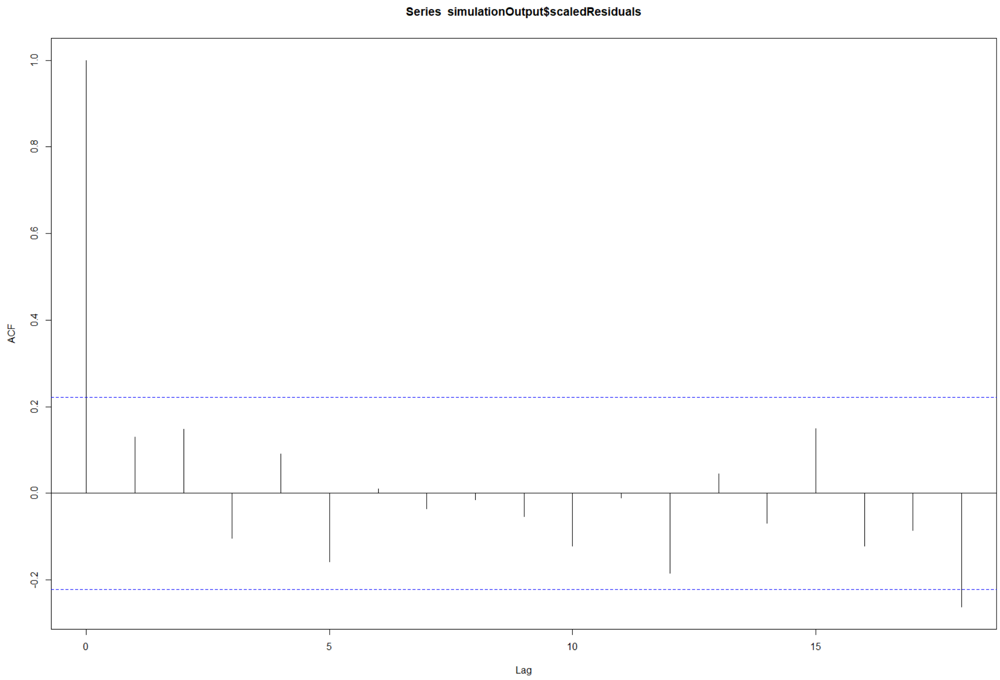
The residuals plotted against predicted values.



Residuals plotted against distance bins. Some imbalance here.



Zero-inflation test. No zero-inflation problem.



ACF test of residuals.