Defence, Safety & Security
Oude Waalsdorperweg 63
2597 AK Den Haag
P.O. Box 96864
2509 JG The Hague
The Netherlands

www.tno.nl

T +31 88 866 10 00

TNO report

TNO 2022 R12205

Analysis of the effects on harbour porpoises from the underwater sound during the construction of the Borssele and Gemini offshore wind farms

Date	November 2022
Author(s)	C.A.F. de Jong, F.P.A. Lam, A.M. von Benda-Beckmann, T.S. Oud (TNO); S.C.V. Geelhoed, T. Vallina, T. Wilkes (Wageningen Marine Research); J.A. Brinkkemper, R.C. Snoek (WaterProof BV)
Classification report	TNO Publiek
Number of pages	115 (incl. appendices)
Number of appendices	7
Sponsor	RWS
Project name	zaak 31163293 "Frequentieweging/meting analyse"
Project number	060.42899

Samenvatting

Effecten van het onderwatergeluid bij het heien van funderingen voor offshore windparken (platforms en windturbines) op zeezoogdieren worden momenteel ingeschat op basis van overschrijding van drempelwaarden voor de blootstelling van de dieren aan onderwatergeluid. Dit gebeurt via het Kader Ecologie en Cumulatie (KEC). Deze studie in opdracht van het Wozep programma van Rijkswaterstaat geeft meer inzicht in de respons van bruinvissen op heigeluid en in de relatie tussen heigeluid en de gedragsrespons.

Op basis van de gegevens van akoestische monitoring tijdens de aanleg van de Borssele windparken in 2019 en 2020 en tijdens de aanleg van de Gemini windparken in 2015 is onderzocht welke geluidsmaat (gewogen of niet gewogen voor de verschillende soort-specifieke frequentie-afhankelijke gehoorgevoeligheid) eventuele gedragsveranderingen van de bruinvis het beste verklaart.

Statistische analyse van de detectie van bruinvis echolocatie-geluiden (*porpoise positive minutes* gemeten door CPOD-apparatuur) als functie van de afstand tot de heipaal laat zien dat bruinvissen minder vaak gedetecteerd worden tijdens het heien tot op afstanden tot tenminste 7 km bij Borssele, waar het heigeluid binnen een geluidsnorm moest blijven, en tot tenminste 15 km bij Gemini, waar nog geen geluidsnorm van toepassing was. Deze afstanden zijn aanzienlijk kleiner dan uitgerekend (50% verstoringskans op circa 30 km afstand) op basis van de huidige KEC methodiek.

Uit de analyse van de detectie van bruinvis-echolocatiegeluiden als functie van ongewogen en gewogen single strike sound exposure level van de heiklappen (SEL_{SS}) volgen verschillende drempelwaarden waarboven bruinvissen minder vaak gedetecteerd worden. Vanwege de toegepaste mitigatiemaatregelen en maskering door omgevingsgeluid was het niet mogelijk om de SEL_{SS} waarden bij Borssele over dezelfde frequentiebandbreedte te bepalen als bij Gemini. Daardoor zijn de dosis-effect relaties niet direct vergelijkbaar. De logische aanname dat een voor de gehoorgevoeligheid van dieren gewogen maat een betere voorspelling geeft voor de gedragsrespons dan een ongewogen maat, wordt niet bevestigd door de analyse van de meetgegevens van de Borssele en Gemini projecten.

De studie laat ook zien dat er problemen zijn met de praktische implementatie van frequentiegewogen SEL_{SS}, door de onzekerheid bij zowel het modelleren als het meten van de hoogfrequente componenten van heigeluid, waarbij omgevingsgeluid hoogfrequent heigeluid maskeert.

De metingen bij Borssele, waar het heigeluid is gemitigeerd en waar de drukke scheepvaart resulteert in een hoog niveau van achtergrondgeluid, laten zien dat het niet altijd duidelijk is of heigeluid de belangrijkste bron van verstoring is. De analyse van de detectie van bruinvis echolocatie-geluiden als functie van het ongewogen en gewogen geluidniveau (SPL) van het onderwatergeluid bij Borssele en Gemini laat een duidelijke afname van detecties zien bij toenemende SPL waarden. Dat suggereert dat SPL wellicht een volledigere maat is voor het voorspellen van een gedragsrespons dan SEL_{SS}. Maar voor het voorspellen van het totale SPL ten gevolge van alle bronnen is meer informatie en zijn meer modellen nodig dan beschikbaar zijn, waardoor deze maat nog niet direct toepasbaar is voor effectstudies.

Summary

Effects of underwater noise on marine mammals during the piling of foundations for offshore wind farms (platforms and wind turbines) are currently estimated on the basis of exceeding threshold values for the exposure of the animals to underwater noise. This is done through the Ecology and Cumulation Framework (KEC). This study, commissioned by the Wozep program of Rijkswaterstaat, provides more insight into the response of harbour porpoises to pile-driving noise and into the relationship between pile-driving noise and the behavioural response.

Based on the data from acoustic monitoring during the construction of the Borssele wind farms in 2019 and 2020 and during the construction of the Gemini wind farms in 2015, it was investigated which acoustic metric (weighted or not weighted for the various species-specific frequency-dependent hearing sensitivity) best explains behavioural changes in the harbour porpoise.

Statistical analysis of the detection of porpoise echolocation sounds (porpoise positive minutes measured by CPOD equipment) as a function of the distance from the pile, shows that porpoises are detected less often during pile driving at distances of at least 7 km near Borssele, where the pile-driving noise had to remain within a noise limit, and up to at least 15 km at Gemini, where no noise limit was yet applicable. These distances are considerably smaller than calculated (50% probability of disturbance at a distance of approximately 30 km) based on the current KEC methodology.

The analysis of the detection of harbour porpoise echolocation sounds as a function of unweighted and weighted single strike sound exposure level of the piling hits (SELSS) shows various threshold values above which harbour porpoises are detected less often. Due to the applied mitigation measures and masking by ambient noise, it was not possible to determine the SELSS values at Borssele over the same frequency bandwidth as at Gemini. Therefore, the dose-effect relationships are not comparable. The logical assumption that a measure weighted for the hearing sensitivity of animals gives a better prediction for the behavioural response than an unweighted measure is not immediately confirmed by the analysis of the measurement data from the Borssele and Gemini projects.

The study also shows that there are problems with the practical implementation of frequency-weighted SEL_{SS}, due to the uncertainty in both modelling and measurement of the high-frequency components of pile-driving noise, where ambient noise masks high-frequency pile-driving noise.

The measurements at Borssele, where pile-driving noise has been mitigated and where busy shipping results in a high level of background noise, show that it is not always clear whether pile-driving noise is the main source of disturbance. The analysis of the detection of harbor porpoise echolocation sounds as a function of the unweighted and weighted sound level (SPL) of the underwater sound at Borssele and Gemini shows a clear shows a clear decrease in detections with increasing SPL values. That suggests that SPL may be a more complete measure for predicting behavioral response than SEL_{SS}. However, predicting the total SPL as a result of all sources requires more information and models than are available, which means that this measure is not yet directly applicable for effect studies.

Contents

Appendices

Acronyms

AdBm	noise mitigation system, by AdBm Technologies, Austin, TX
ADD	acoustic deterrent device
AIC	Akaike information criterion
BIC	Bayesian information criterion
CPOD	porpoise click detector, by Chelonia Limited, UK
DBBC	double big bubble curtain
EIA	environmental impact assessment
GAM	generalized additive model
GAMM	generalized additive mixed model
HSD	Hydro-Sound-Damper, by OffNoise-HSD-Systems GmbH
ISO	International Organization for Standardization
KEC	Kader Ecologie en Cumulatie
PAM	passive acoustic monitoring
PPM	porpoise positive minutes per hour (CPOD)
RBINS	Royal Belgian Institute of Natural Sciences
SEL_{SS}	Single strike sound exposure level
$SEL_{SS,VHF}$	Single strike sound exposure level, weighted for porpoise hearing (VHF), see (Southall, et al., 2019)
SPL_{1s}	sound pressure level, averaged over 1 s
$SPL_{1s,VHF}$	sound pressure level, averaged over 1 s, weighted for porpoise hearing (VHF), see (Southall, et al., 2019)
VHF	‘very high frequency’, referring to the hearing group of VHF cetaceans to which the harbour porpoise belongs (Southall, et al., 2019)
Wozep	wind op zee ecologisch programma

1 Introduction

In 2016, the Dutch Ministry of Economic Affairs and Climate Policy commissioned Rijkswaterstaat to set up an integrated research program to reduce the knowledge gaps regarding the effects of offshore wind farms on the North Sea ecosystem. This *Wind op Zee Ecologisch Programma* (Wozep) runs from 2016 to 2023 and the results of the studies carried out are used in the Framework for Assessing Ecological and Cumulative Effects (*Kader Ecologie en Cumulatie*; KEC). The KEC framework is used to determine the cumulative effects of current and planned wind farms on protected species, to get a view on the possible long-term effects of future upscaling of offshore wind energy.

The Wozep project aims to:

- Reduce uncertainties of assumptions and knowledge gaps in the KEC, environmental impact reports (EIA) and appropriate assessments;
- Reduce uncertainties of assumptions and knowledge gaps regarding long-term effects due to scaling up of wind energy at sea;
- Gain insight in the effectiveness of mitigation measures to reduce adverse effects.

The effects of the underwater sound due to the impact piling of the foundations for wind turbines and platforms on marine mammals are currently assessed on the basis of threshold values above which the exposure of these animals to underwater sound imposes a risk of significant disturbance. In the KEC, see (Heinis F. , de Jong, von Benda-Beckmann, & Binnerts, 2019), the threshold value for disturbance of harbour porpoises by piling sound is set at an unweighted, broadband single strike sound exposure level (SEL_{SS}) of 140 dB re 1 μPa^2s. In the 2021 update of KEC (KEC 4.0), this has been replaced by using a dose-effect relationship, that describes a probability of disturbance as function of SELss exposure.
In the discussion of the knowledge gaps in (Heinis, de Jong, & Rijkswaterstaat Underwater Noise Working Group, 2015) it was recognized that the effect of the signal waveform and frequency content on the dose-effect relationship needs to be investigated further. Linking threshold values to the species-specific hearing threshold in the way proposed by (Tougaard, Wright, & Madsen, 2015) may have an effect on the estimate of the number of disturbed animals.

The monitoring programme during the construction of the Borssele and Gemini offshore wind farms, see (Brinkkemper, et al., 2021) and (Geelhoed, Friedrich, Joost, Machiels, & Ströber, 2018), provided an opportunity to investigate which acoustic metric (unweighted or weighted for the hearing sensitivity) provides the best prediction of behavioural response of marine mammals to piling sounds.

This report describes the behavioural response analysis for harbour porpoises (*Phocoena phocoena*) and the potential consequences of the results for the KEC. Effects on seals are reported in (Brasseur, Aarts, & Schop, 2022).

2 Motivation

The KEC describes a staged procedure to determine the cumulative effects of impulsive underwater sound on the harbour porpoise population, which is schematically illustrated in Figure 1.

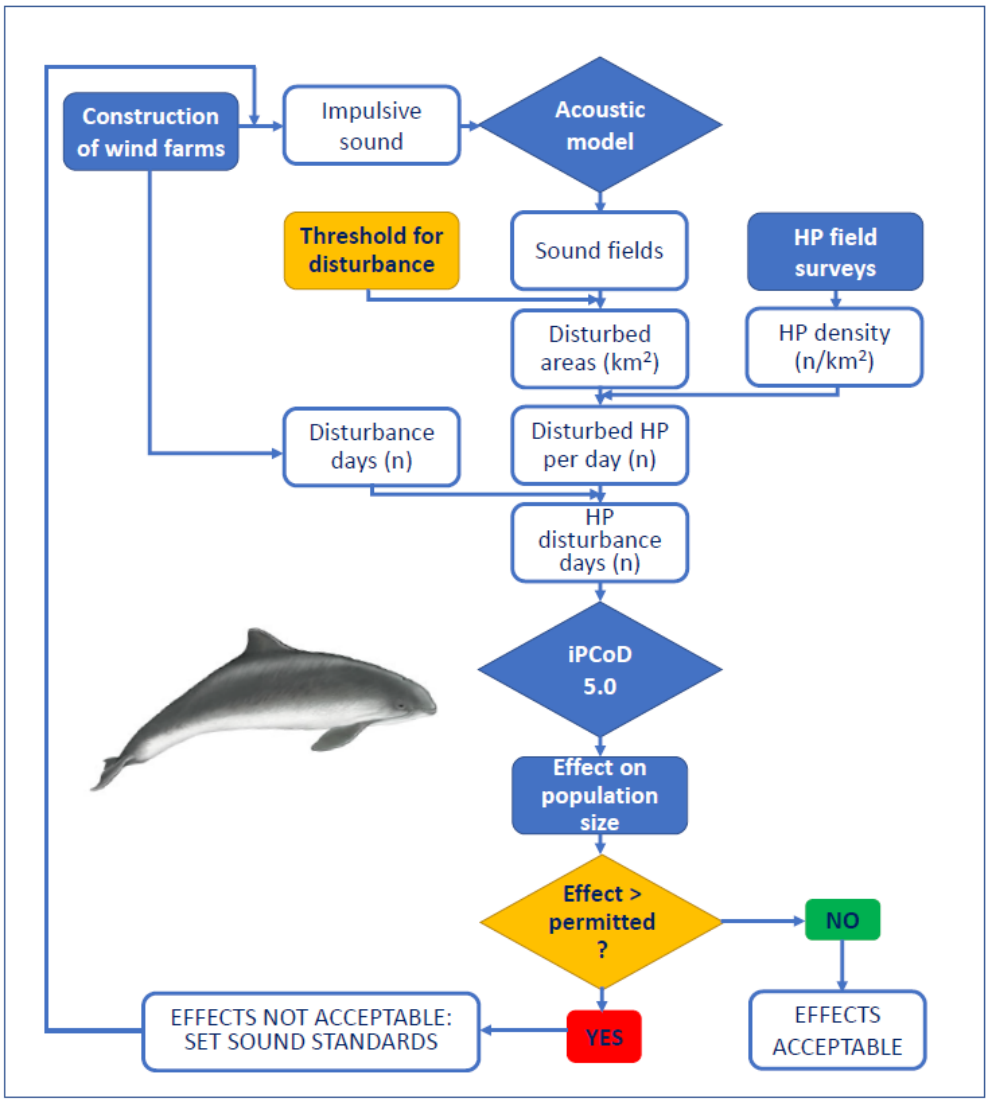

Figure 1 Schematic representation of the stages in the staged procedure for determining and assessing the cumulative effects of impulsive underwater sound on harbour porpoises during the construction of wind farms, see (Heinis F. , de Jong, von Benda-Beckmann, & Binnerts, 2019).

As indicated by the yellow rectangle in Figure 1, the size of the area around the piling location in which porpoises may be disturbed by piling sound is determined by comparison of the calculated sound field with a 'threshold value for disturbance'. The KEC 3.0 approach (Heinis F. , de Jong, von Benda-Beckmann, & Binnerts, 2019) assumed that porpoises are likely to show significant behavioural disturbance when they are exposed to the sound piling strikes when the maximum broadband unweighted single strike sound exposure level (SEL_{SS}) exceeds a threshold value of 140 dB re 1 μPa^2s. This threshold value was tentatively based on observed

reductions in harbour porpoise presence around the piling location during the construction of the Borkum West II wind farm (Diederichs, et al., 2014) and on observations of jumping out of the water during exposure studies in the pool of SEAMARCO (Kastelein, van Heerden, Gransier, & Hoek, 2013).

An extensive study of the effects of pile-driving on harbour porpoises looking at the first seven wind farms in German waters (Brandt, et al., 2018) concluded that '*Declines were found at sound exposure levels exceeding 143 dB re 1 µPa²s (the sound exposure level exceeded during 5% of the piling time, SEL_{05}) and up to 17 km from piling*'. This suggested that the current KEC threshold value represents a 'worst-case' assumption.

That became clearer in the assessment of the impact of the unmitigated piling noise during the construction of the Gemini wind farms in Dutch waters.

a) Analysis of the data from passive acoustic monitoring (PAM) of porpoise presence during the construction of the Gemini wind farms (Geelhoed, Friedrich, Joost, Machiels, & Ströber, Gemini T-c: aerial surveys and passive acoustic monitoring of harbour porpoises 2015, 2018) led to the conclusion that '*the avoidance distance of harbour porpoises lies in the range of 10-20 km, which is supported by the aerial surveys that suggest changes in distribution in a radius from < 15 km up to 25 km around a pile driving location during piling. The avoidance distance might be restricted by the length of the piling event, that lasted on average too short to allow harbour porpoise to swim further away during this period.'*
b) The monitoring data from the Gemini wind farm construction were applied by (Nabe-Nielsen, et al., 2018) to calibrate their DEPONS model in which individual animals respond by being chased away from the sound source in response to observed piling sound. In their model, the strength of the response decreases linearly with decreasing distance from the pile. Their analysis led to the conclusion that porpoises responded to the piling noise at Gemini up to 8.9 km from the construction sites.
c) Predictions and measurements of the broadband unweighted single strike sound exposure level for some Gemini piles suggest that the threshold value SEL_{SS} = 140 dB re 1 µPa²s was likely exceeded up to a distance of 40 to 50 km from the pile (de Jong, et al., 2019).

The apparent difference between the observations and the predicted distance at which porpoises are disturbed by piling sound is likely caused by simplifying worst-case assumptions that were necessarily made for KEC due to knowledge gaps.

As a first step, it was recognized that the use of single value thresholds for noise exposure for behavioural responses leads to uncertainties in predicting effects, as argued by (Tyack & Thomas, 2019) and (Southall, et al., 2021). Taking into account the inherent variability in the response of different individuals by means of dose-response relationships, can possibly reduce the risk that effects are underestimated in the assessment. Deriving appropriate dose-response relationships requires availability of measurement data, which is currently very limited. However, Graham et al. (2019) derived a dose-response relationship for the disturbance of harbour porpoises by piling sound from measurements made during the construction of the Beatrice wind farm in the UK. The relationship is expressed in their paper as a function of an 'audiogram-weighted' SEL_{SS} but a dose-effect relationship based on

unweighted SEL_{SS} was also presented at the INPAS Symposium in Amsterdam (June 2018). In the KEC 4.0, it was decided to adopt, as the worst-case scenario, the dose-effect relationship derived by Graham et al. (2019) for the response of harbour porpoises to the turbine foundation that was piled first. Possible habituation, leading to a reduced probability of disturbance after successive piling days was disregarded, as a precautionary measure.

It was also recognized during the KEC development that the effect of the signal form and frequency content on the dose-effect relationship needs to be investigated further. Linking threshold values to the species-specific hearing threshold in the way proposed by (Tougaard, Wright, & Madsen, 2015) may have an effect on the estimate of the number of disturbed animals. In 2017, TNO carried out an initial study for Wozep reviewing the possible application of frequency weighting in the assessment of the impact of underwater sound on harbour porpoise, see (de Jong & von Benda-Beckmann, 2018). At that time, it was not yet possible to draw firm conclusions about the effects of (impulsive) underwater sound on porpoises and seals and the appropriate metrics to quantify these. Studies were proposed in (de Jong & von Benda-Beckmann, 2018) to fill some of the apparent knowledge gaps.

One proposed study of *porpoise behavioural response to sound pulses with different frequency content* was carried out by SEAMARCO. The results were published in Aquatic Mammals (Kastelein, de Jong, Tougaard, Helder-Hoek, & Defillet, 2022). This study concluded that frequency weighting of the sound exposure level (SEL) will improve prediction of behavioural responses. However, it remained unclear whether the weighting for predicting auditory effects is also the best weighting to predict behavioural effects.

Other proposals from (de Jong & von Benda-Beckmann, 2018) concerned further analysis of available data from field measurements. The passive acoustic monitoring during the construction of the Borssele wind farms provided an opportunity for such investigations. It was found that the circumstances at Borssele were not ideal for studying the responses of porpoises to piling sound. The piling sound levels were relatively low, because the permit for construction required piling sound mitigation, and the background sound levels in the area are quite high, due to heavy shipping. Therefor the project was extended with an analysis of the already available monitoring data from the construction of the Gemini Offshore Wind Park. These analyses are described in this report.

2.1 Acoustic metrics

KEC 4.0 (Heinis, de Jong, & von Benda-Beckmann, 2022) copied the assumption from KEC 1.0 (Heinis, de Jong, & Rijkswaterstaat Underwater Noise Working Group, 2015) that the broadband single strike sound exposure level (SEL_{SS}) is an appropriate metric for the prediction of behavioural disturbance.

Single strike sound exposure level (abbreviation SEL_{SS}, symbol L_E) is an alternative name for *time-integrated squared sound pressure* (ISO 18405, 2017). For an acoustic signal with sound pressure $p(t)$, SEL_{SS} is defined by:

$$L_E = 10\log_{10}\left\{\frac{\int_{t=0}^{t=T} p^2(t)\mathrm{d}t}{p_0^2 t_0}\right\}\ \mathrm{dB\ re\ 1\ \mu Pa^2 s} \quad (1)$$

with reference exposure $p_0^2 t_0 = 1\ \mu\mathrm{Pa}^2\mathrm{s}$. Figure 2 provides an example for illustration. The time duration T over which the integration is taken must include the full length of the single strike sound, see (ISO 18406, 2017). The ISO standard also specifies that $\mathrm{SEL_{SS}}$ is calculated as a broadband value (single number for a stated bandwidth) and in decidecade[1] frequency band levels covering at least the frequency range from 20 Hz to 20 kHz.

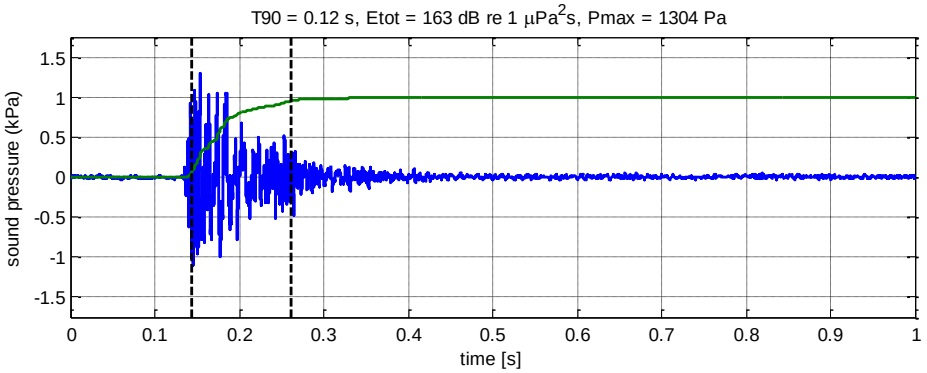

Figure 2 Example of the acoustic pressure received at a hydrophone for a single piling strike (blue line). The green line shows the cumulative time-integrated squared sound pressure (sound exposure) as a function of integration time, scaled to the total sound exposure in the signal. The thick black dashed lines indicate the start and end times of a time window (T_{90}) that contains 90% of the exposure energy.

In 2019, (Southall, et al., 2019) proposed updated auditory weighting functions for assessing the effects of sound exposure on marine hearing (permanent and temporary hearing threshold shifts). One of the aims of the present study was to investigate whether the proposed auditory weighting functions are appropriate for quantifying marine mammal behavioural response to sound exposure as well. Because it is unlikely that animals are disturbed by sound outside their hearing range, it is reasonable to take that into account in the assessment of dose-response relationships, by applying some form of auditory frequency weighting.

For harbour porpoises we use the auditory weighting function for the “very high-frequency cetaceans” (VHF) marine mammal hearing group, proposed in (Southall, et al., 2019). The frequency weighting is given by the function

$$W(f) = C + 10\log_{10}\left\{\frac{(f/f_1)^{2a}}{[1+(f/f_1)^2]^a[(1+(f/f_2)^2]^b}\right\}\ \mathrm{dB} \quad (2)$$

Here f is the frequency in kHz and the parameter values for VHF-weighting are $C = 1.36\ \mathrm{dB}$, $a = 1.8$, $b = 2$, $f_1 = 12\ \mathrm{kHz}$ and $f_2 = 140\ \mathrm{kHz}$.
This is a generic function based on measured harbour porpoise audiograms, as illustrated in Figure 3. Note that the shape of the audiograms and the inverted weighting function is similar, but differences between the individual curves can be observed of the order of 10 dB. Uncertainty in selection of the appropriate weighting function leads to uncertainty in the calculation of weighted levels.

[1] Also known as ‘one-third octave (base-10)’, see (ISO 18405, 2017).

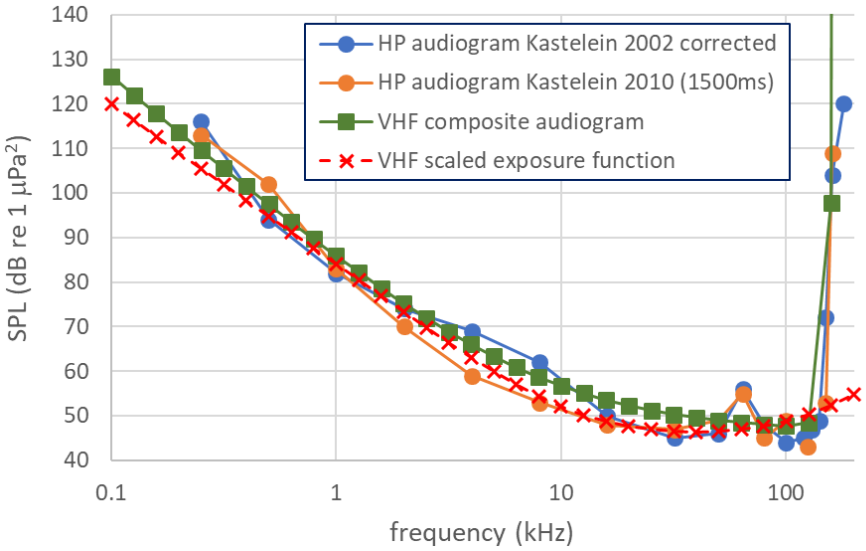

Figure 3 Measured harbour porpoise audiograms (hearing threshold versus frequency), from (Kastelein, Hoek, de Jong, & Wensveen, 2010), compared with the composite audiogram proposed by (Southall, et al., 2019) and the scaled inverse of the frequency weighting (exposure) function: $46.4\ \mathrm{dB} - W(f)$.

From a practical point of view, using the same weighted metrics for assessing physiological, auditory and behavioural effects has great benefits.

The acoustic data in this study are quantified in terms of the metrics summarized in Table 1. The SEL_{SS} metrics are determined per piling strike. The SPL_{1s} metrics have been selected by the Jomopans[2] project for characterizing the measured ambient sound field. In addition to the piling strike sounds, this includes the underwater sound produced by other sources, such as the operation of vessels associated with the wind farm construction, passing ships, and surface waves. For the correlation analysis with the CPOD porpoise detections, which are quantified in terms of 'porpoise positive minutes per hour', the SEL_{SS} and SPL_{1s} metrics are summarized in the 50th percentile (median), 95th percentile, 100th percentile (maximum) and the power average of all levels during each hour of pile driving.

Table 1 Overview of metrics considered in this study.

Metric	Description	symbol	unit
SEL_{SS}	Unweighted broadband single strike sound exposure level	L_E	dB re 1 µPa²s
$SEL_{SS,VHF}$	Broadband single strike sound exposure level, frequency weighted for very high frequency cetaceans (VHF)	$L_{E,\mathrm{VHF}}$	dB re 1 µPa²s
Distance	Distance to the pile location	R	km
SPL_{1s}	Unweighted broadband sound pressure level, averaged over duration $T = 1\ s$	$L_{p,1s}$	dB re 1 µPa²
$SPL_{1s,VHF}$	Maximum value over duration $T = 1\ s$ of the time-weighted ('fast', i.e. exponentially weighted with a 0.125 s time constant) broadband sound pressure level, frequency weighted for very high frequency cetaceans (VHF)	$L_{p,1s,\mathrm{VHF}}$	dB re 1 µPa²

[2] https://northsearegion.eu/jomopans/

2.2 Dose-response relationships

A ‘dose-response relationship’ quantifies the probability of a specified response to an acoustic exposure as a function of the acoustic dosage.

The KEC 4.0 approach includes a (tentative) model that quantifies the probability that the behaviour of harbour porpoises is ‘significantly disturbed’ as a function of the unweighted broadband single strike sound exposure level (abbreviation: SEL_{SS}; symbol L_E) of piling sound to which they are exposed. This probability is largely unknown, but a first estimation of a dose-response relationship was proposed based on the work of (Graham, et al., 2019). They have derived a dose-response relationship from changes in detections of porpoise clicks (expressed as ‘detection positive hours’) in the area and period around piling activities during the construction of the Beatrice Offshore wind farm in the UK. In the (Graham, et al., 2019) paper this relationship is expressed in terms of an ‘audiogram-weighted’ SEL_{SS}, but at the INPAS Symposium in Amsterdam (June 2018) they have also presented dose-response relationship based on unweighted broadband SEL_{SS}.
A ‘worst-case’ relationship, based on the curve derived by (Graham, et al., 2019) for the disturbance by the first piling event of an installation sequence, has been proposed for implementation in KEC 4.0 (Heinis, de Jong, & von Benda-Beckmann, 2022). Possible habituation, leading to a reduced probability of disturbance when there are successive piling days, has been disregarded as a precautionary measure.

The proposed dose-response relationship is described with a logistic function

$$\mathrm{P_{resp}}(L_E) = \frac{1}{1+\mathrm{e}^{-k(L_E - L_{E,50\%})}} \tag{3}$$

with parameters $k = 0.1482/\mathrm{dB}$ and $L_{E,50\%} = 144.4\ \mathrm{dB}$, see Figure 4.

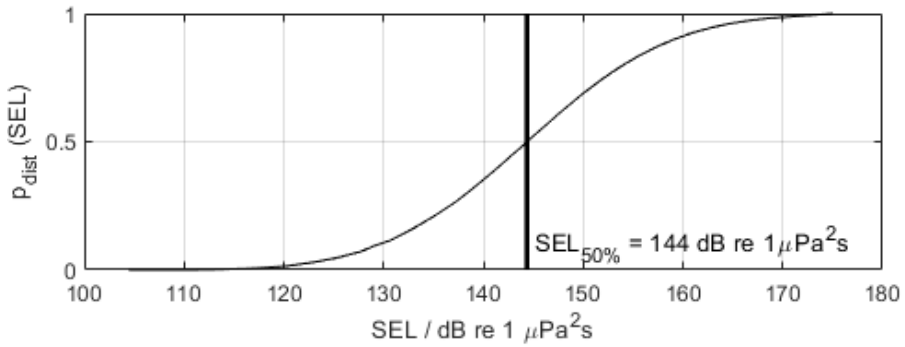

Figure 4 The dose-response relationship for the disturbance of porpoises by piling sound, based on an analysis of measurements during the construction of the Beatrice Offshore wind farm in the UK (Graham, et al., 2019). This represents the observed response to the piling for the first turbine.

In the KEC approach (Heinis, de Jong, & von Benda-Beckmann, 2022) it is assumed that this relationship quantifies the probability of the occurrence of a significant behavioural response in harbour porpoises as function of the unweighted broadband single strike sound exposure level (SEL_{SS}) of piling sound to which they are exposed. Behaviour with a score of ‘5’ or higher on the behavioural response severity scale of (Southall, et al., 2007) is considered ‘significant’. These are behaviours such as changes in swimming behaviour and breathing, avoiding a

particular area and changes in calling or clicking behaviour (for the purposes of communication or foraging). The number of days during which individual porpoises are subject to such a response is assumed to affect the harbour porpoise population development through factors such as survival and reproductive success (the *vital rates*).

Note that (Graham, et al., 2019) defined 'response' as a proportional decrease in harbour porpoise click detections by more than 50% in the period (12 or 24 hours) after cessation of piling. Therefore, the dose-response relationship proposed in KEC 4.0 must be considered as a temporary solution for the lack of more appropriate data. It was considered as an appropriate first proxy. The SELss at which the probability of disturbance equals 50% is of the same order of magnitude as the disturbance threshold (SELss = 140 dB re 1 μPa^2s) assumed for KEC 3.0 (Heinis F. , de Jong, von Benda-Beckmann, & Binnerts, 2019). Observations of (Brandt, et al., 2018) during construction of the first seven offshore wind farms in German waters showed an onset of declined porpoise detections at sound levels exceeding SELss = 143 dB re 1 μPa^2s. This confirms that the KEC 4.0 dose-response relationship, which predicts a 50% probability of disturbance at SELss = 144 dB re 1 μPa^2s, provides a worst-case estimation of porpoise disturbance.

One of the aims of the present study is to investigate the applicability of this dose-response relationship with the observations made during the monitoring for the Borssele and Gemini wind farms.

2.3 Effect of frequency weighting on SELss-metrics

Incorporating frequency weighting in the exposure assessment will lead to different predictions of the (effective) disturbance distance if the decay of weighted SELSS,VHF with range differs significantly from the decay of unweighted SELSS with distance.

A first comparison was made in (de Jong & von Benda-Beckmann, 2018) for the piling sound measured during the construction of the Gemini wind farm, see Figure 5.
Table 2 lists the corresponding broadband values, which are dominated by sound at frequencies below 1 kHz for SELSS and at frequencies above 1 kHz for SELSS,VHF. Figure 6 shows that the broadband weighted SELSS,VHF decays significantly steeper with range than the unweighted SELSS. The difference increases from 6 dB at 7 km to about 15 dB at 32 and 66 km distance. Note, however, that in this example a significant part of this decrease is related with the reduced frequency bandwidth at larger distances, where higher frequencies are masked by background noise.

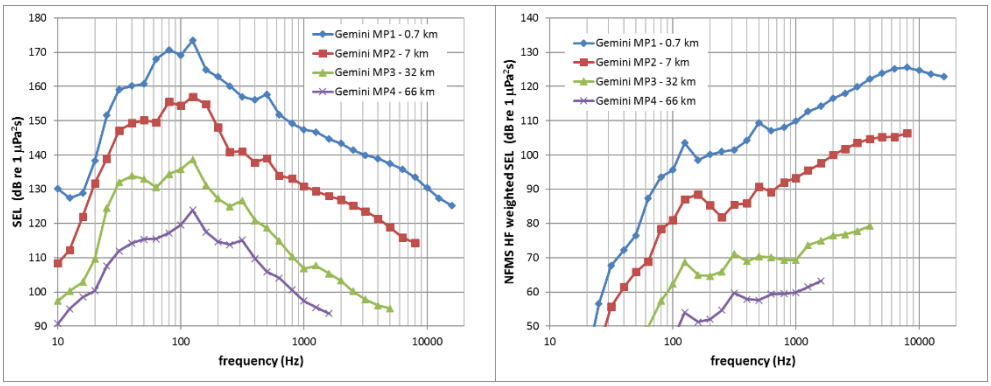

Figure 5 Decidecade spectra of mean single strike sound exposure levels measured at 4 distances from pile U8 for the Gemini wind farm, from (Binnerts, et al., 2016). Left graph: unweighted SEL_{SS}; right graph: frequency-weighted $SEL_{SS,VHF}$. The upper part of the frequency range is omitted for the distant locations where the piling sound was below background noise. Source: (de Jong & von Benda-Beckmann, 2018).

Table 2 Unweighted and weighted broadband values of single strike sound exposure level (with varying frequency bandwidth, see Figure 5) as measured during piling for the Gemini wind farm (pile U8).

	unit	**MP1**	**MP2**	**MP3**	**MP4**
Distance	km	0.7	7	32	66
Unweighted SEL_{SS}	dB re 1 µPa²s	178	163	144	128
Porpoise-weighted $SEL_{SS,VHF}$	dB re 1 µPa²s	133	112	84	67

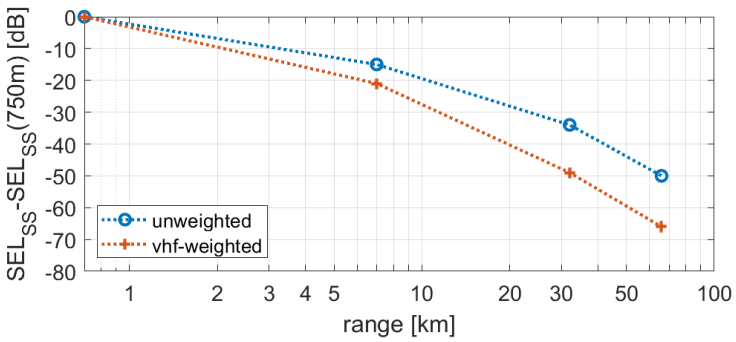

Figure 6 Decay of unweighted and weighted broadband SEL_{SS} (with varying frequency bandwidth, see Figure 5) with range as measured during piling for the Gemini wind farm (pile U8), relative to the values measured at 0.7 km.

3 Borssele monitoring

3.1 Wind farm construction

The Borssele wind farms are located offshore of the southwestern coast of the Netherlands, see Figure 7. Monopile turbine foundations were installed between October 2019 and June 2020. Ørsted installed 94 turbines in Borssele lots ('kavels') I and II, the Blauwwind consortium installed 77 turbines in Borssele lots III and IV. The site decision ('kavelbesluit') specifies limits for a maximum SEL_{SS} at 750 m distance from the pile, dependent on the number of turbines per lot and the season. In order to meet this requirement, noise mitigation measures were applied. Borssele I and II were constructed with a combination of the Hydro-Sound-Damper[3] (HSD) system and a double big bubble curtain[4] (DBBC). Borssele III and IV were constructed with the AdBm[5] noise mitigation system and a DBBC.

In the same period there were piling activities for construction of the Seamade and NorthWester2 offshore wind farms in Belgian waters, just across the border, see Figure 7. These concern 60 turbine piles for Seamade and 24 piles for NorthWester2, all constructed with DBBC noise mitigation.

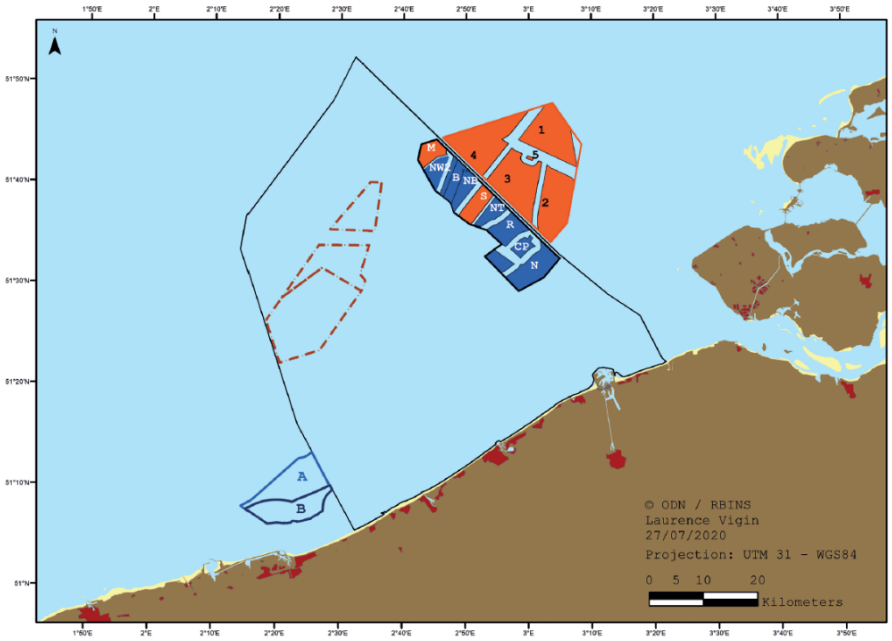

Figure 7 Overview of the locations of the Borssele (1 to 4) offshore wind farm sites next to the Belgian wind farms, of which Seamade (S) en NorthWester2 (NW2) were constructed in the same period as the Borssele wind farms. Figure provided by RBINS.

[3] https://www.offnoise-solutions.com/the-hydro-sound-damper-system-hsd-system/

[4] https://www.bsh.de/EN/TOPICS/Offshore/Environmental_assessments/Underwater_sound/_Module/Karussell/_documents/Artikel_Gr_Blasenschleier.html

[5] https://adbmtech.com/

3.2 Piling data

For this study, Ørsted and Van Oord (constructor for the Blauwwind consortium) kindly provided the following information from the piling of all Borssele turbine foundations:

- Pile location and geometry (length, diameter, wall thickness).
- Hammer logs (time, hammer energy and pile penetration per piling strike).
- Information about applied deterrent devices (ADDs) and noise mitigation systems, including start and stop times of DBBCs.

RBINS kindly provided information from the Belgian wind farm construction projects:

- Pile location and geometry (length, diameter, wall thickness).
- Information about applied ADDs and noise mitigation systems, including start and stop times of DBBCs.
- Hammer logs (time, hammer energy and pile penetration per piling strike) for the SeaMade wind farm. These hammer logs could not be made available for the NorthWester2 offshore wind farm.

Figure 8 shows an overview of the time windows between the first and the last piling strike for all turbine foundations in the four wind farms in the period between August 2019 and June 2020. It is clear from this figure that piling occurs in day and night. The construction of the Belgian wind farms was already three months underway when the first pile at Borssele was installed.

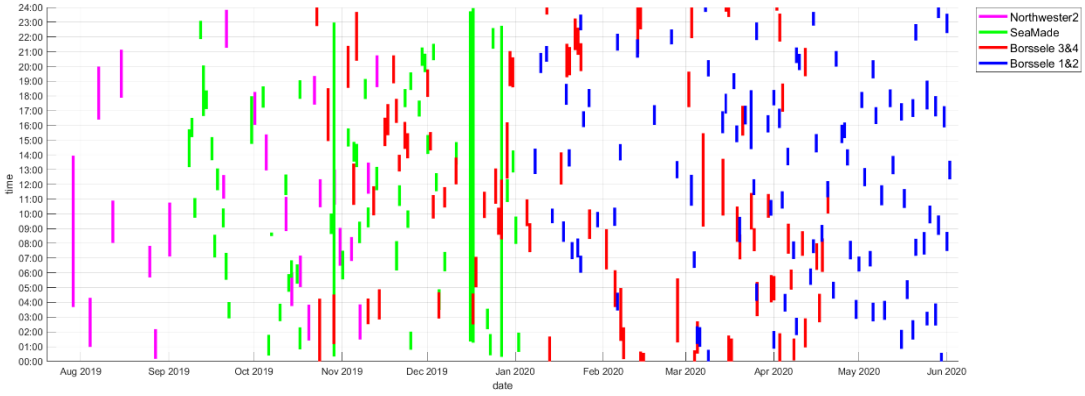

Figure 8 Calendar of the piling time windows for the four offshore wind farms in the Borssele area. The markers along the x-axis indicate the start of each month. The colours refer to the four wind farms (see legend).

3.3 Passive acoustic monitoring

WaterProof BV and Wageningen Marine Research carried out underwater sound measurements during the installation of the Borssele wind farms (Brinkkemper, et al., 2021), see Figure 9. 16 porpoise detectors (Chelonia CPODs) and seven underwater sound recorders (Ocean Instrument SoundTraps, recording sound over the frequency range 20 Hz – 20 kHz) were deployed from October 2019 up to September 2020. Figure 10 provides an overview of the data availability from the acoustic recorders.

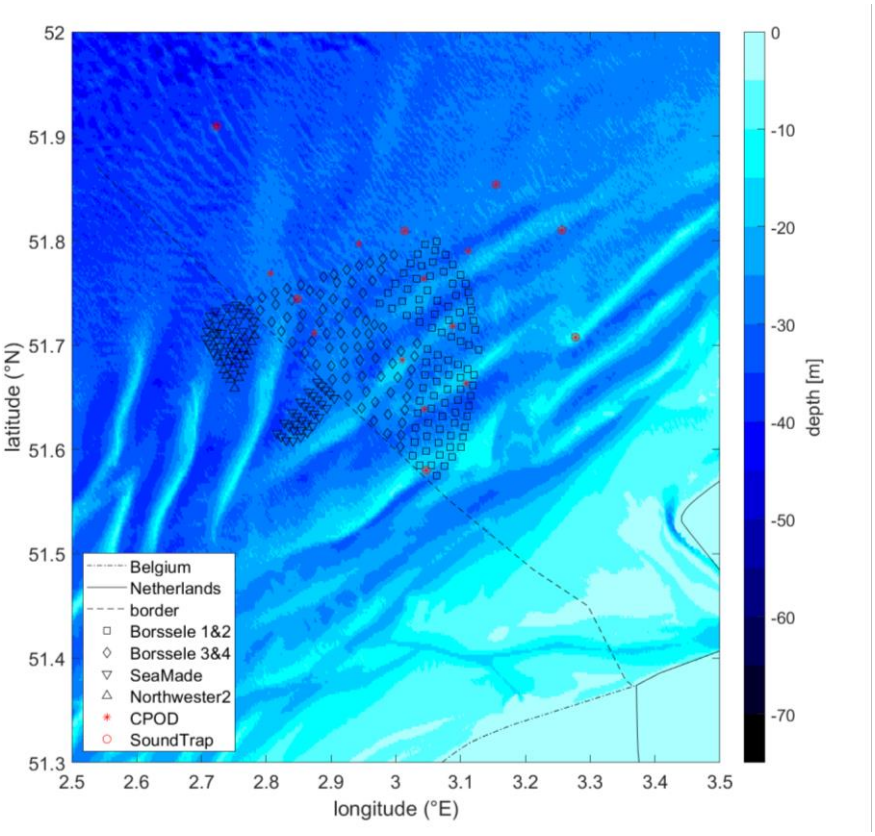

Figure 9 Overview of the bathymetry and of the locations of the piles for the four wind farms and of the sensors. The underwater sound was monitored with 7 acoustic recorders (SoundTrap ST300HF, Ocean Instrument, N.Z.). Harbour porpoise presence was monitored with 16 continuous porpoise detectors (CPOD version 1, Chelonia Ltd., U.K.), see (Brinkkemper, et al., 2021).

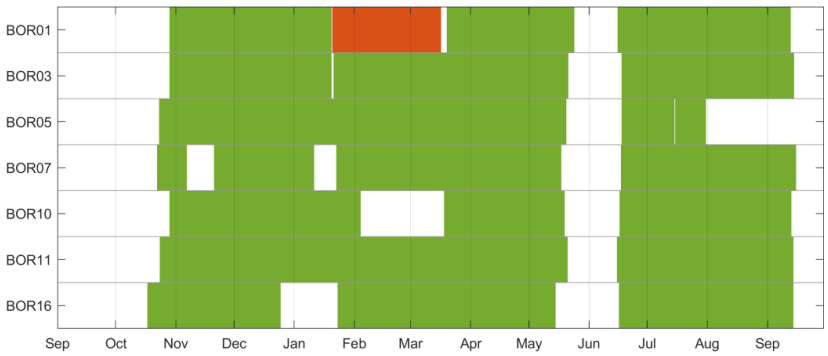

Figure 10 Overview of the time periods over which data are available from the 7 SoundTrap acoustic recorder stations, with (white) no data, (red) corrupt data, and (green) good quality data. Figure from (Brinkkemper, et al., 2021). Data gaps are due to technical issues and to Covid-19 restrictions that prohibited timely servicing of the recorders at the end of May 2020.

All CPODs recorded acoustic activity almost continuous, with on average 328.5 deployment days (320-335 days), which was mainly restricted by the initial deployment and recovery date of the CPOD, see (Brinkkemper, et al., 2021).

Figure 11 gives an overview of the distances of the piles from the sound recording stations. This shows that the distance of stations 1, 3, 10 and 11 to the piles is always more than 10 km.

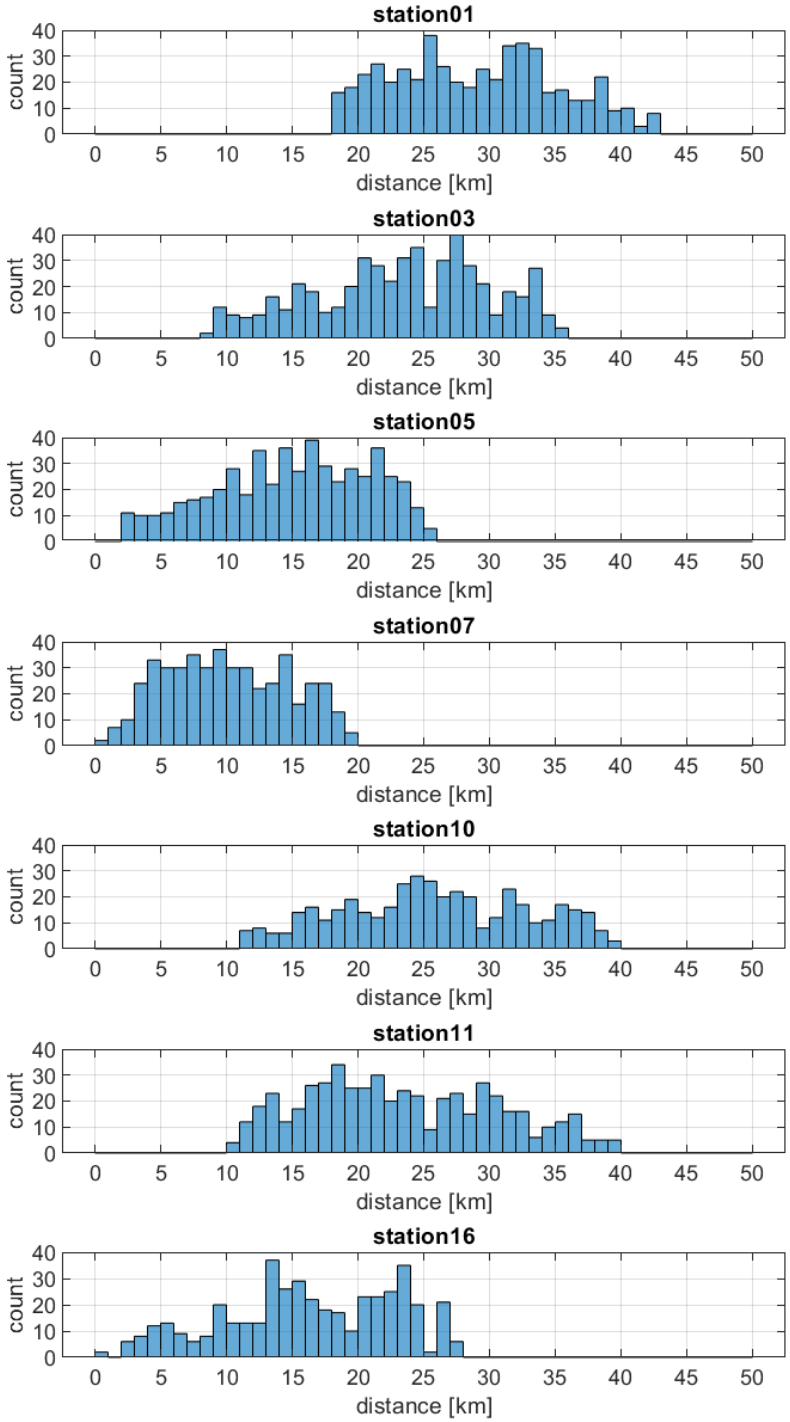

Figure 11 Histograms of the distances between the pile locations and stations with sound recorders.

3.4 Acoustic data processing

WaterProof carried out the processing of the acoustic measurements, including calibration of the recorder signals, detection of the piling strikes and calculation of the agreed metric (Table 1). The processing, including an analysis of benchmark test of the processing routines between WaterProof and TNO is reported in (Brinkkemper J. , 2021).

The processing resulted in values of SEL_{SS} and $SEL_{SS,VHF}$ for all detected piling strikes and of SPL_{1s} and $SPL_{1s,VHF}$, for all seven recorder stations and for the complete recording period.

For the behavioural response analysis, these data were aggregated into hourly statistic metrics: 50% (median) and 95% and 100% percentiles of the decibel levels and the power averaged level, i.e. the level in decibel of the arithmetic average of the underlying power quantity (mean-square sound pressure or sound exposure).

3.5 Analysis of ambient sound (SPL_{1s}) during piling at Borssele

The seven SoundTrap recorders in the Borssele area have continuously recorded underwater sound, independent of the presence of piling activities. WaterProof has converted the acoustic recordings into time series of the unweighted ($L_{p,1s}$) and the maximum value of the time- and frequency-weighted SPL ($L_{p,1s,PCW}$ and $L_{p,1s,VHF}$), calculated following the approach of (Tougaard & Beedholm, 2019) for the VHF and PCW mammal hearing groups, applying an exponential temporal weighting with an integration time of 125 ms, see (Brinkkemper J. , 2021).

The processed data have been analysed by TNO, to evaluate the relevance of piling sound in relation to other sounds in the area, such as these from ships, and of flow noise due to tidal currents, see (de Jong, 2021).

Figure 12 gives an example of the underwater sound ($L_{p,1s}$) measured in the windfarm area on one typical day (5 November 2019). This shows that:

- At measurement station 07 the one-third octave band spectra of $L_{p,1s}$ show the piling sound from Borssele pile F01, at 3.1 km from the recorder, predominantly in the frequency bands between 50 and 500 Hz.
- The piling sound from Seamade pile MC5, at 18.4 km from the recorder, is visible as well, in the same frequency bands, at lower $L_{p,1s}$ levels.
- Very early on the same day, between 1 and 3 o'clock, there is another peak in the underwater sound, but according to the piling logs of the wind farms there was no piling at that time. The sound is from a ship that is likely using a dynamic positioning system to stay at its location.
- From 4 o'clock until 19 o'clock, there are horizontal lines in the spectrogram which are associated with mechanical processes on a ship or platform.
- In the lowest frequency bands (<50 Hz) the tidal flow along the hydrophones causes a periodic increase and decrease of flow noise (period 6 hours).
- The piling noise is visible, but not very prominent, in the time series of the unweighted and PCW-weighted $L_{p,1s}$.
- The unweighted broadband SPL clearly measures the periodic tidal flow noise pattern. Nevertheless, the piling of the nearby turbine pile in the Borssele wind farm is visible, because it coincides with slack tide. The low frequency tidal flow noise is reduced in the PCW- and VHF-weighted metrics.

- The piling noise is nearly invisible in the VHF-weighted $L_{p,1s}$ and completely invisible in the hourly statistical metrics of VHF-weighted $L_{p,1s}$ (at this recorder on this day).

Analysis of the sound recorded by all stations over the full recording period (de Jong, 2021), suggests that nearby piling is clearly present in the underwater sound, but it is clearly not the dominant or only source of underwater sound in the area. The sound recordings at the sensors outside the wind farm area (see Figure 9) appear to be dominated by passing ships rather than by the sound from distant piling activities.

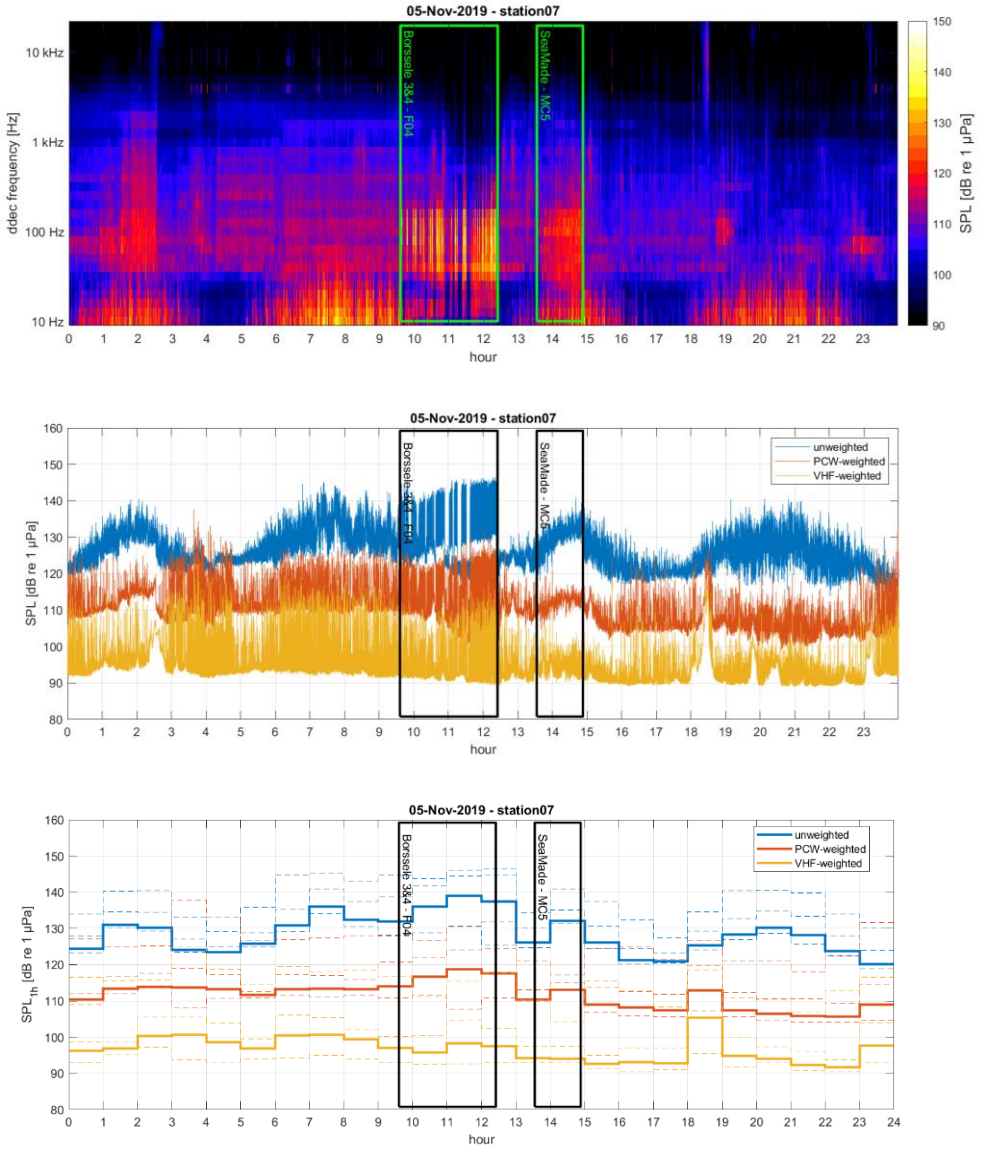

Figure 12 Underwater sound recorded at station 07 on 5 November 2019.
Upper: decidecade spectrogram of $L_{p,1s}$. Middle: unweighted and weighted broadband $L_{p,1s}$. Lower: unweighted and weighted broadband values of $L_{p,1h}$ (solid lines).
The dashed lines give the three hourly percentiles (50%, 95% and 100%) for the $L_{p,1s}$ metrics. PCW-weighted refers to weighting for seal (phocid pinniped in water) hearing.

Figure 13 shows a comparison of the hourly unweighted and weighted broadband SPL recorded during the piling events as function of the distance from the recorder to the pile.

- The piling events cause a significant increase of the unweighted $L_{p,1\mathrm{h}}$ at distances up to about 10 km from the pile. At larger distances the difference with the recorded $L_{p,1\mathrm{h}}$ outside the piling hours is negligible.
- For the VHF-weighted $L_{p,1\mathrm{h}}$ the difference with the recorded $L_{p,1\mathrm{h}}$ outside the piling hours is also very small at distances shorter than 10 km.

Due to the effectiveness of the applied noise mitigation measures, in combination with the noisy environment (lots of ship traffic), the piling sound is only clearly observed at relatively short distances from the piles, generally at less than about 10 km. The piling noise is predominantly observed in the frequency bands between 50 and 500 Hz. Because harbour porpoise hearing is not very sensitive at these low frequencies, the distances at which piling noise is observed in the porpoise-weighted SPL are much smaller than 10 km. The SPL due to piling is generally not much larger than the SPL caused by passing ships.

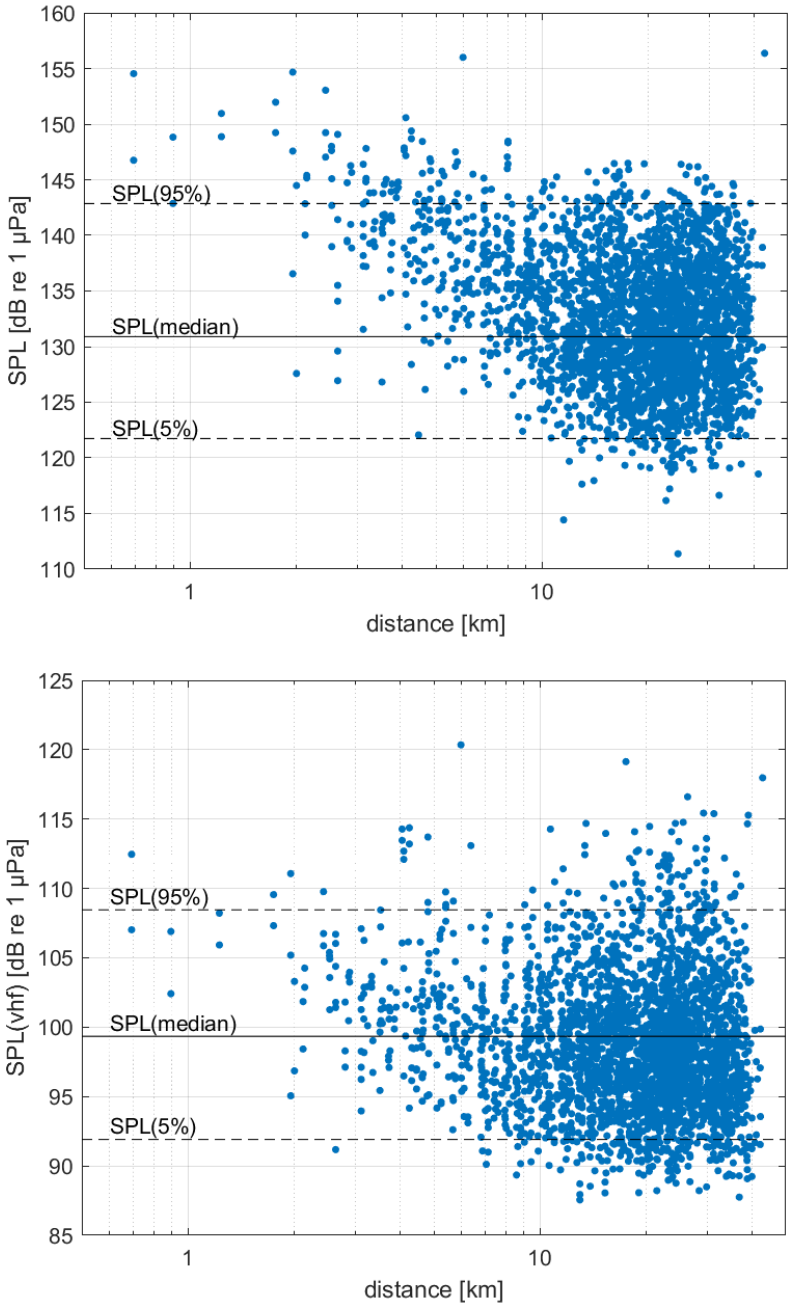

Figure 13 Broadband $L_{p,1h}$, at the seven sound recorders in the hours when piling took place, as a function of distance to the pile. The horizontal lines show the 5th, 50th (median) and 95th percentiles of the $L_{p,1h}$ in the hours when no piling occurred. Upper graph: unweighted $L_{p,1h}$; lower graph: VHF-weighted $L_{p,1h,VHF}$.

3.6 Sound exposure modelling

TNO's Aquarius 4 piling sound model (de Jong, et al., 2019) was used to calculate the sound exposure at the 16 CPOD locations from all individual piling strikes on each pile for which the piling log was available. The results were calibrated to the sound exposure level spectra measured by the seven sound recorders, to correct for the effectiveness of the applied sound mitigation that could not be directly included in the modelling. The modelling is described in Appendix A and in (Oud & de Jong, 2021).

Figure 14 shows the resulting decay of the modelled weighted and unweighted SEL_{SS} with increasing distance from the pile. The marker colours illustrate the modelled effect of the different noise mitigation systems on the SEL_{SS}. Variation of SEL_{SS} values at the same distance are related with variations in the bathymetry along the trajectories between pile and recorder locations. Note that this modelling is limited to the Borssele and Seamade piles, because the piling logs for NorthWester2 were not available.

The modelled decay with distance in Figure 14 is similar for the unweighted SEL_{SS} and the porpoise-weighted $SEL_{SS,VHF}$. Because the noise mitigation and the reduced frequency bandwidth of the modelled spectra have removed the high frequency piling sound (above 500 Hz), the decay with increasing range is not affected by higher frequencies. These higher frequencies were responsible for the faster decay of the weighted $SEL_{SS,VHF}$ that was observed in the unmitigated piling for the Gemini wind farm, as shown in Figure 6. Consequently, SEL_{SS}(50 Hz – 500 Hz) and $SEL_{SS,VHF}$(50 Hz – 500 Hz) are strongly correlated over the range of distances at which piling sound is observed, which explains why there is no clear difference in the suitability of both metrics for prediction of behavioural response of marine mammals to piling sounds with noise mitigation, see Chapter 5.

To determine whether this conclusion holds for unmitigated piling sound as well, this study was extended with an analysis of the monitoring data from the construction of the Gemini wind farms, where piling sound was not mitigated, see Chapter 4. The data volume could be further expanded via international collaboration with researchers in for example Germany, UK and Denmark.

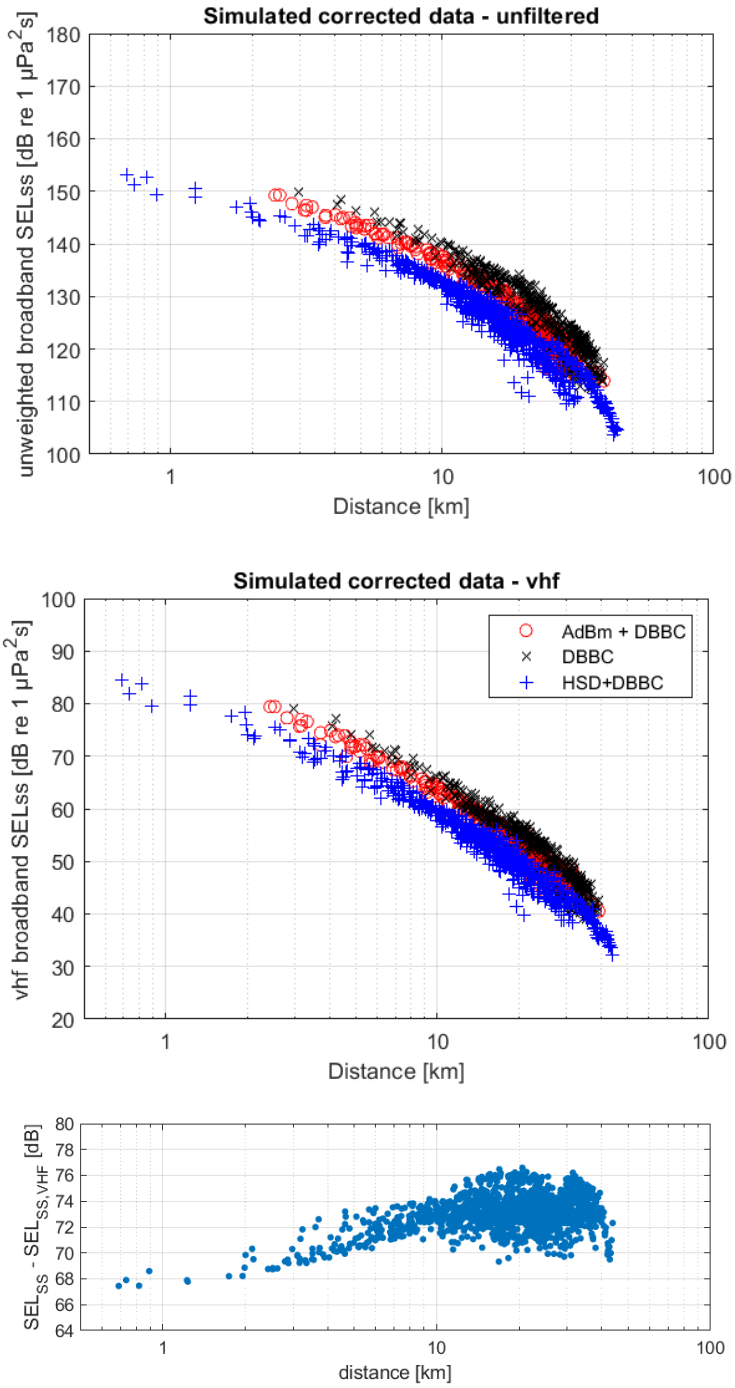

Figure 14 Modelled broadband (50 Hz – 500 Hz) SELss as a function of the pile-recorder distance (one marker for each pile recorder combination), using the model calibration. Upper: unweighted SEL_{SS}; Middle: porpoise-weighted $SEL_{SS,VHF}$. The three mitigation configurations are indicated by three different markers and colours. The lower figure shows the difference between the unweighted and weighted SELss as a function of distance, illustrating that the weighted SELss drops faster with increasing distance.

4 Gemini monitoring

4.1 Wind farm construction

The Gemini Offshore Wind Park (consisting of two parts: Buitengaats and ZeeEnergie) is located off the northern coast of the Netherlands, see Figure 7. 150 monopile turbine foundation piles and 8 platform piles were installed between 1 July 2015 and 17 October 2015. At that time, the permits for offshore wind farm constructions did not specify noise limits, and the piling took place without noise mitigation.

Figure 15 Overview of the locations of the Gemini offshore wind farm sites. Figure from: https://www.noordzeeloket.nl/functies-gebruik/windenergie/doorvaart-medegebruik/ten-noorden-waddeneilanden-inclusief-gemini/.

4.2 Piling data

The following data from the Gemini monitoring research program[6] were made available for this study:

- A detailed overview of pile locations (UTM zone 31N, ETRS89) and hammer logs (time, hammer energy and pile penetration per piling strike) for 142 out of the 150 piles in the Gemini Offshore Wind Park.
- Data from the monitoring of underwater sound in the vicinity of Gemini offshore wind park in the (T-0) period prior to construction (Lucke, 2015).
- Data from the T-0 passive acoustic monitoring of harbour porpoises (Geelhoed, et al., Gemini T-0: passive acoustic monitoring and aerial surveys of harbour porpoises, 2015).
- Data from the monitoring of underwater sound in the (T-c) period during the Gemini construction of three selected piles (Remmers & Bellmann, 2016).

[6] See https://www.geminiwindpark.nl/ecological-reports.html

- Data from the T-c passive acoustic monitoring of harbour porpoises (Geelhoed, Friedrich, Joost, Machiels, & Ströber, 2018).

Figure 8 shows an overview of the time windows between the first and the last piling strike for all turbine foundations in the Gemini Offshore Wind Farms. Simultaneous pile driving for piles in the Buitengaats and ZeeEnergie parts of Gemini occurred on nine occasions, generally with short overlap in time (ca 30 min).

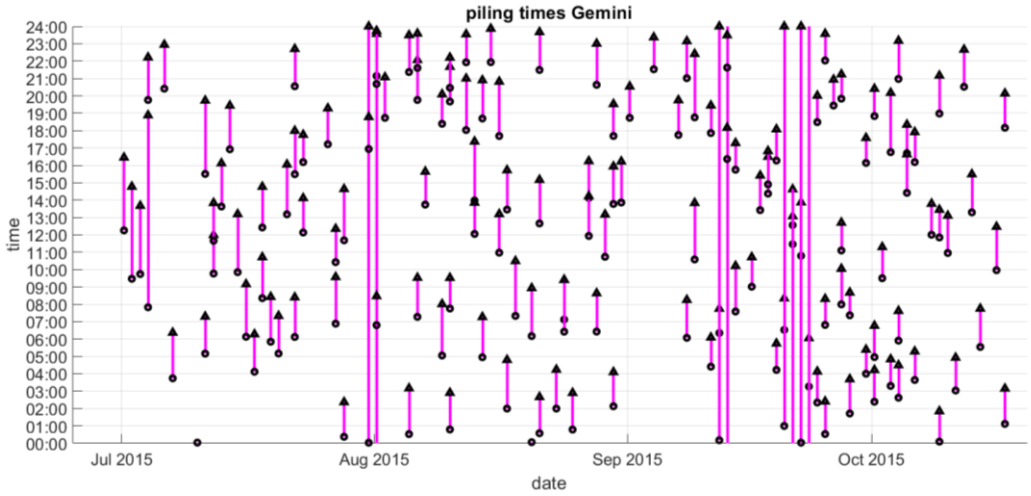

Figure 16 Calendar of the piling time windows for the Gemini Offshore Wind Park construction.

4.3 Passive acoustic monitoring

Wageningen Marine Research carried out aerial surveys and passive acoustic monitoring of harbour porpoises prior to and during the construction of the Gemini Offshore Wind Park, see (Geelhoed, et al., Gemini T-0: passive acoustic monitoring and aerial surveys of harbour porpoises, 2015) and (Geelhoed, Friedrich, Joost, Machiels, & Ströber, 2018). 15 CPODs were deployed for the passive acoustic monitoring of harbour porpoises, as shown in Figure 17.

In the T-0 monitoring, between September 2011 and July 2014, 6,881 days of CPOD recordings were obtained during two distinct one-year sampling periods. The T-c monitoring covered the period from June 2015 until January 2016.

Porpoise clicks that are detected by the CPOD are applied as an indicator of harbour porpoise presence. Detections are quantified in terms of ‘porpoise positive minutes’ (PPM), aggregated per hour (PPM/h).

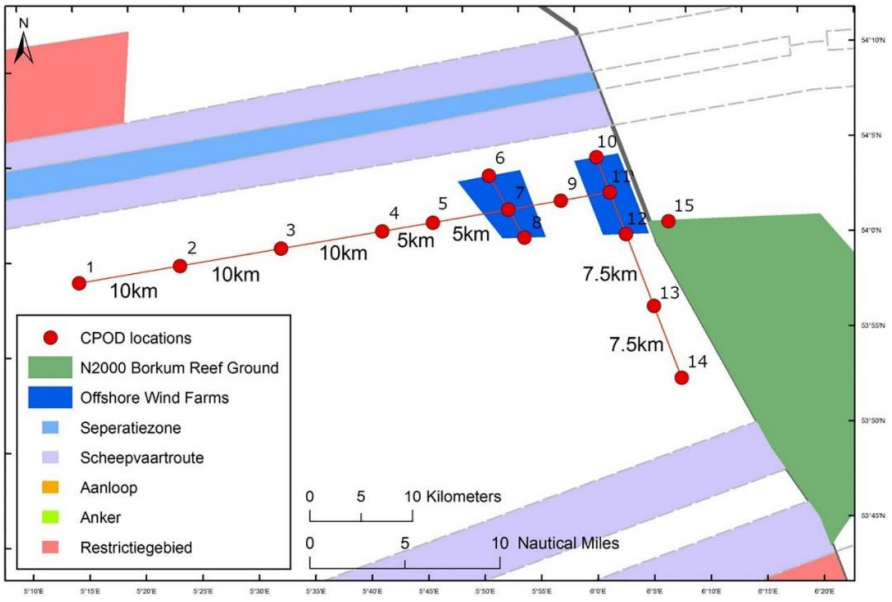

Figure 17 Location of the CPODs for the passive acoustic monitoring of porpoises in the area of the Gemini Offshore Wind Park, from (Geelhoed, Friedrich, Joost, Machiels, & Ströber, Gemini T-c: aerial surveys and passive acoustic monitoring of harbour porpoises 2015, 2018). This shows the configuration for the T-c monitoring. The configuration for the T-0 monitoring was similar, but with slightly different locations and a different numbering.

WMR deployed AMAR acoustic recorders at two locations close to CPOD locations 1 and 9 (Figure 17) in two periods prior to the construction: July until September 2013 and March until July 2014.

Underwater sound monitoring during the construction period was limited to measurements of sound during the piling for turbine foundations U8 and Z2 and platform (for offshore high voltage station OHVS1) pile B3, see (Remmers & Bellmann, 2016). These data have been used for the development and validation of TNO's Aquarius 4 piling sound model (de Jong, et al., 2019).

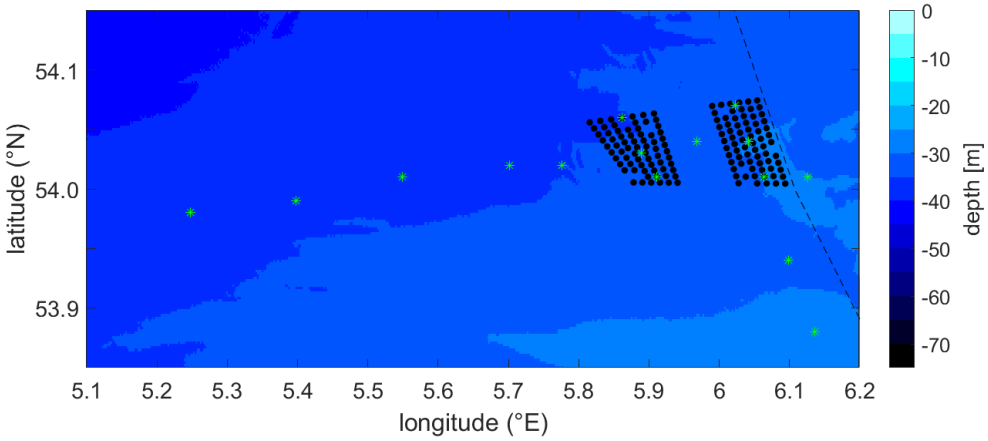

Figure 18 Overview of the bathymetry and of the locations of the piles (black dots) of the Gemini Offshore Wind Farm for which piling data were available and of the CPODs (green ⋆ markers).

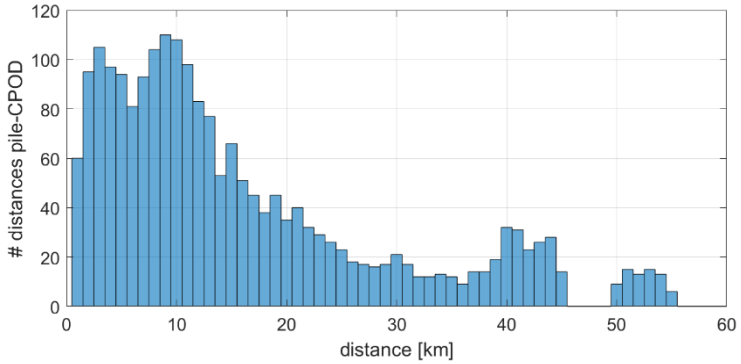

Figure 19 Histogram of the distances between the pile locations and CPOD stations.

4.4 Sound exposure modelling

TNO's Aquarius 4 piling sound model (de Jong, et al., 2019) was used to calculate the sound exposure at the 15 CPOD locations from all individual piling strikes on each pile for which the piling log was available. Unlike the modelling for Borssele (§3.6), no calibration was needed, because no mitigation measures were applied and the Gemini sound measurement data were already used for the development and validation of TNO's Aquarius 4 piling sound model (de Jong, et al., 2019).

Figure 20 shows the resulting decay of the modelled weighted and unweighted SEL_{SS} with increasing distance from the pile. The modelled decay with distance for the Gemini piling sound is clearly different for the unweighted SEL_{SS} and the porpoise-weighted $SEL_{SS,VHF}$. Figure 21 shows the difference between the unweighted and weighted SELss as a function of distance, illustrating that the weighted SELss drops faster with increasing distance than observed for the mitigated piling sound in the Borssele study (Figure 14). This suggested that these data might be more useful to study the suitability of both metrics for prediction of behavioural response of marine mammals to piling sounds with noise mitigation, see Chapter 5.

Note, however, that a part of the decrease of the weighted $SEL_{SS,VHF}$ with increasing distance in Figure 20 is related with underpredictions of the high-frequency content by the current Aquarius code, see (de Jong, et al., 2019). The reason for this underprediction were too complex to solve in the scope of this project, see Appendix B.

The Aquarius SEL-calculations do not include the effect of masking by background sound. Analysis of the background sound (§3.5) demonstrates that the VHF-weighted SEL_{SS} is much more susceptible to masking than the unweighted SEL_{SS}. Consequently, the lowest predicted values of $SEL_{SS,VHF}$ are likely not detected by porpoises. This is taken into account in the behavioural response modelling (§5.3) by setting a lower threshold for SEL_{SS}, based on an estimation of the masking by background sound over a representative pulse duration of about 100 ms. The lower threshold for $SEL_{SS,VHF}$ is set at 80 dB re 1 μPa^2s and the lower threshold for unweighted SEL_{SS} is set at 100 dB re 1 μPa^2s.

Note that the lowest calculated unweighted SEL_{SS} value for piling strikes is about 130 dB, hence well above the assumed threshold, while the lowest calculated weighted $SEL_{SS,VHF}$ value (~60 dB) is well below the assumed threshold, which accounts for masking effects.

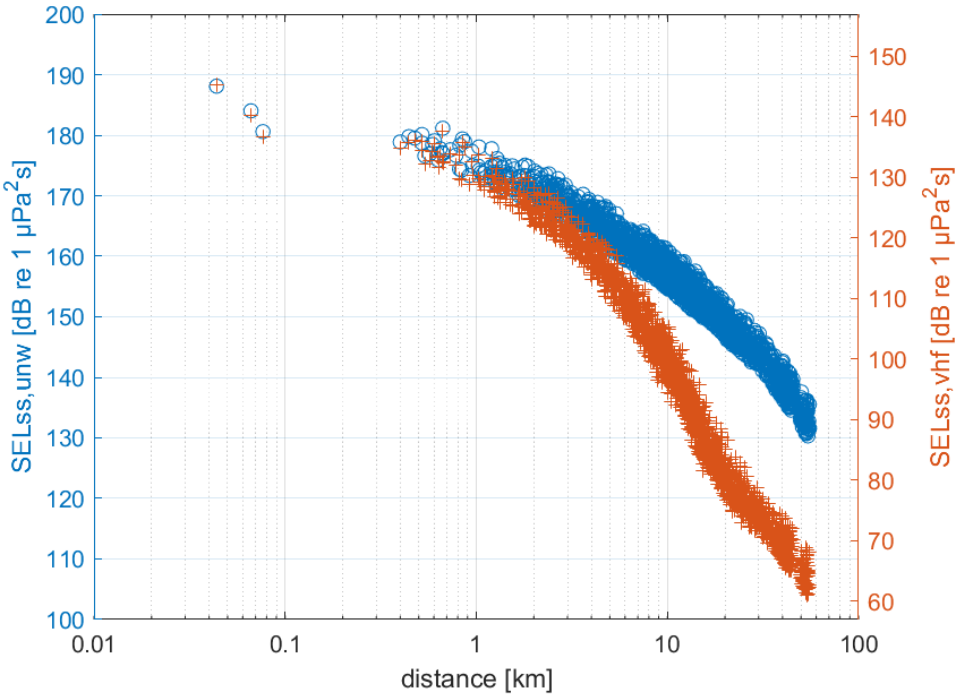

Figure 20 Modelled broadband (10 Hz – 20 kHz) SELss as a function of the pile-recorder distance (one marker for each pile recorder combination), using the model calibration. This shows the unweighted SEL_{SS} (blue) and weighted $SEL_{SS,VHF}$ (red) results of the Aquarius calculations for the Gemini piling. The y-axes are shifted so that the dB-value of unweighted and weighted SELss at the shortest distance overlap.

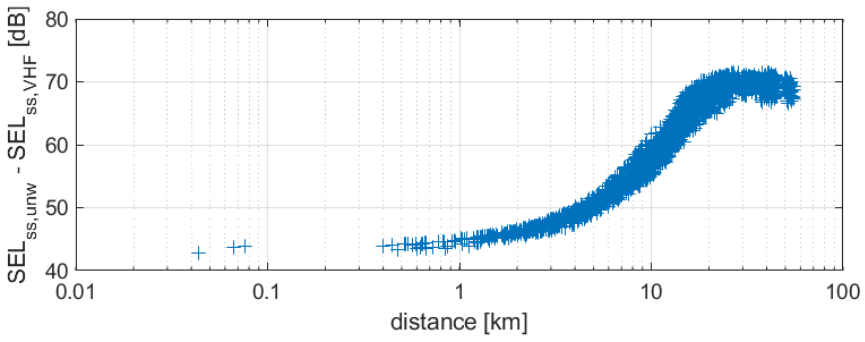

Figure 21 Difference between unweighted SEL_{SS} and weighted $SEL_{SS,VHF}$.

5 Porpoise behavioural response to piling sound

5.1 Analysis of passive acoustic monitoring

Detection of harbour porpoise clicks and buzzes by “Continuous Porpoise Detector” devices (CPOD version 1, Chelonia Ltd., U.K.) is used to quantify harbour porpoise acoustic activity around the CPOD locations. Typically, the maximum range at which CPODs can detect porpoise clicks is about 300 m, depending on CPOD settings as well as on background noise (Clausen, Tougaard, Carstensen, Delefosse, & Teilman, 2019). Details of the used CPODs and processing are described in (Geelhoed, Friedrich, Joost, Machiels, & Ströber, 2018) and (Brinkkemper, et al., 2021). CPOD data can also be used to quantify foraging behaviour, so-called feeding buzzes (Berges, Geelhoed, Scheidat, & Tougaard, 2019).

On the severity scale for ranking observed behavioural responses of free ranging marine mammals proposed by (Southall, et al., 2007), a reduction of acoustic activity (‘cessation or modification of vocal behaviour’) was scored as ‘4’ (on a scale from 0 to 9) when it is ‘moderate (duration ≈ duration of source operation)’ and ‘5’ when it is ‘prolonged (duration > duration of source operation). The updated severity scale in (Southall, et al., 2021) scores ‘brief/minor changes in vocal rates’ as ‘4’ and ‘sustained changes in vocal rates’ as ‘6’. This remains a qualitative interpretation. A reduction in acoustic activity can indicate that less porpoises pass the detector, or that porpoises stay in the area, but change their behaviour, for example cease foraging. In KEC we have chosen a conservative approach by qualifying a reduction of harbour porpoise acoustic activity due to exposure to piling sound as a ‘significant behavioural change’ for which the possible effect on porpoise vital rates has been quantified in the expert elicitation for the Interim PCOD model (Booth, Heinis, & Harwood, 2018).

Harbour porpoise acoustic activity is quantified by the number of minutes per hour in which at least one porpoise click train was detected (‘porpoise positive minutes per hour’; PPM/h). In the Dutch North Sea, CPODs typically measure a few porpoise positive hours per day, so that the PPM/h dataset contains a lot of zeros, so-called zero inflation. Harbour porpoise click detections vary with day in the year, time of the day and location, and possibly also with changing environmental factors such as tide, temperature, wind and wave height. In addition, porpoise detections can be reduced by disturbing events, such as (but not limited to) exposure to pile driving sound.

Due to the strong variation, with many zeros, in the PPM/h recordings at the CPOD locations, it is not trivial to determine the effect of exposure to piling sound on the detections. This requires statistical modelling, taking into account all relevant confounding variables. In practice, this is limited to the confounding variables that can be quantified, such as day in the year, time of the day, location, tide, temperature, wind and wave height. Piling sound exposure is quantified in terms of unweighted (SEL_{SS}) or weighted sound level ($SEL_{SS,VHF}$), or distance to the piling location.

Modelling of the probability of detecting porpoise clicks (PPM/h) involves the assumption that unknown variables do not significantly affect the model.

The modelling ignores effects of interactions between porpoises and their prey and predators, and of (sound from) other anthropogenic disturbances than piling, such as for example passing ships and construction activities. It is questionable if the influence of construction activities can reliably be ignored, because disturbance by (the underwater sound from) human activities in the area before piling provides a likely explanation why porpoise detections decrease prior to the first piling sound, as, for example, observed in (Brandt, et al., 2018). This effect is studied in the Chapter 6.

5.2 Borssele analysis

Details of the statistical modelling of porpoise detections based on the data from the monitoring during the construction of the Borssele wind farms are described in Appendix D.

The detection of porpoise clicks was recorded in the period from 15 October 2019 to 31 October 2020 using 16 CPODs. Figure 22 illustrates the temporal and spatial distribution of porpoise click detections.

More than 90% of the PPM/h-values were equal to zero. Such a large number of zeros complicates fitting a model to the measured values of PPM/h. Therefore, the presence (PPM/h>0) or absence (PPM/h=0) of porpoise positive minutes was modelled instead, using a Bernoulli generalized additive mixed model (GAMM) with a complementary log-log link function.

To study the effect of piling sound on porpoise presence, the main variable in the models was either distance to the piling location, or the VHF-weighted $SEL_{SS,VHF}$(50-500Hz), or the unweighted SEL_{SS}(50-500Hz).

The hourly mean of the unweighted and weighted SELss-values from the piling sounds received at the CPOD locations was obtained from sound exposure modelling, as described in §3.6. When no piling event took place, sound exposure levels were set to 0 dB.

Additionally, the model included month of the year, hour of the day, water temperature, tidal flow magnitude, wind speed and CPOD location as covariates (see Appendix D for details). There was no collinearity between covariates.

The fitted model predicts the probability of presence (PPM/h>0) of porpoise positive minutes as function of the considered variables. Independent of the choice of the main variable, the model shows consistent trends for the covariates.
The probability of PPM/h>0:

- increased from January to November,
- was lower during the day and higher during the night,
- was higher at lower water temperatures,
- decreased somewhat with increasing tidal flow magnitude,
- decreased with increasing wind speed,
- was a bit higher to the south than to the north of the Borssele area.

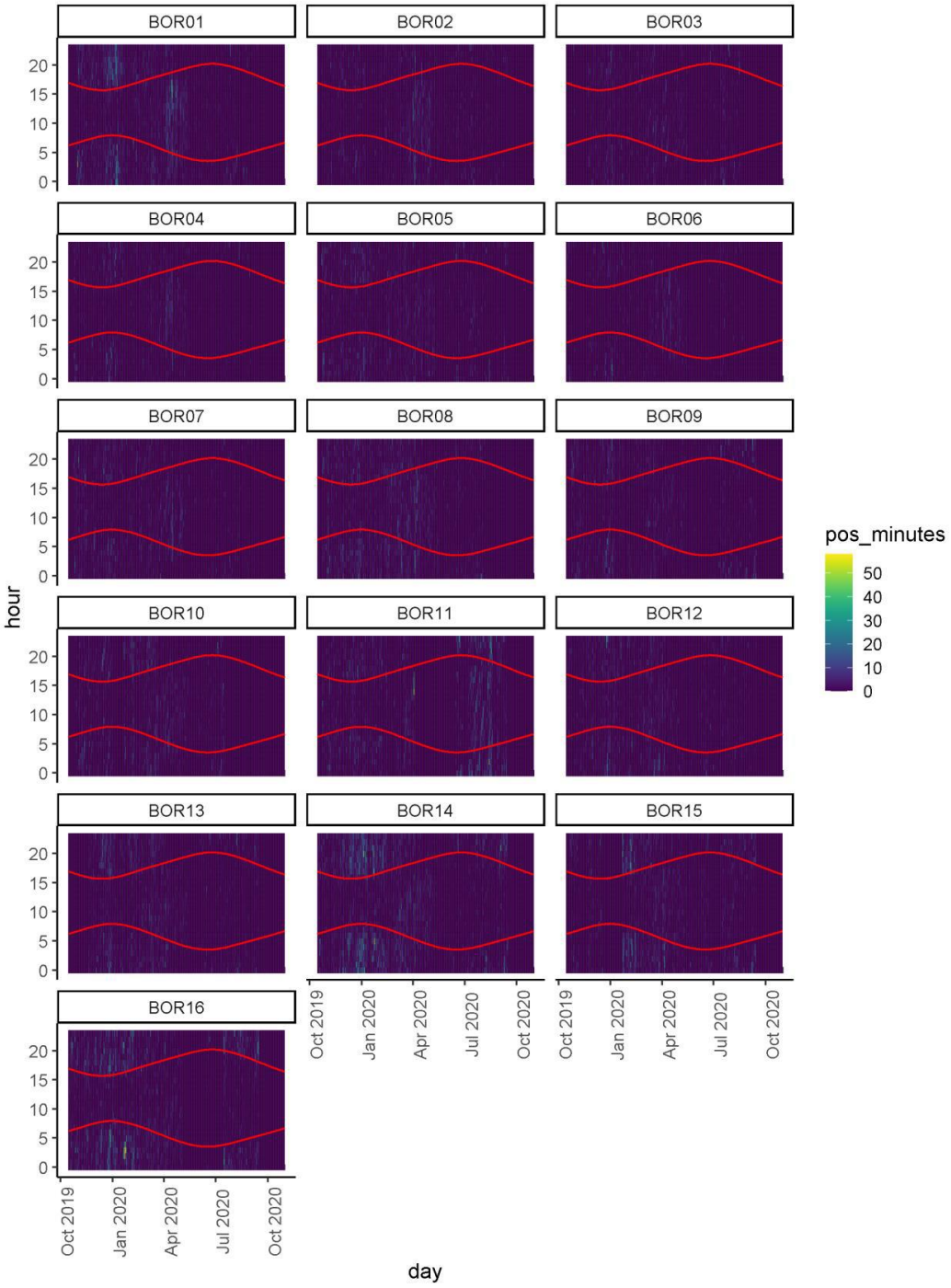

Figure 22 Harbour porpoise acoustic activity per CPOD location, expressed as porpoise positive minutes per hour. The red lines indicate sunset and sunrise. Figure from (Brinkkemper J. , 2021).

Figure 23 shows model plots for the probability of PPM/h>0 as function of the three main variables, for mean values of the covariates.

The models for SEL_{SS} and $SEL_{SS,VHF}$ were fitted to the hourly power averaged level of the values per piling strike.

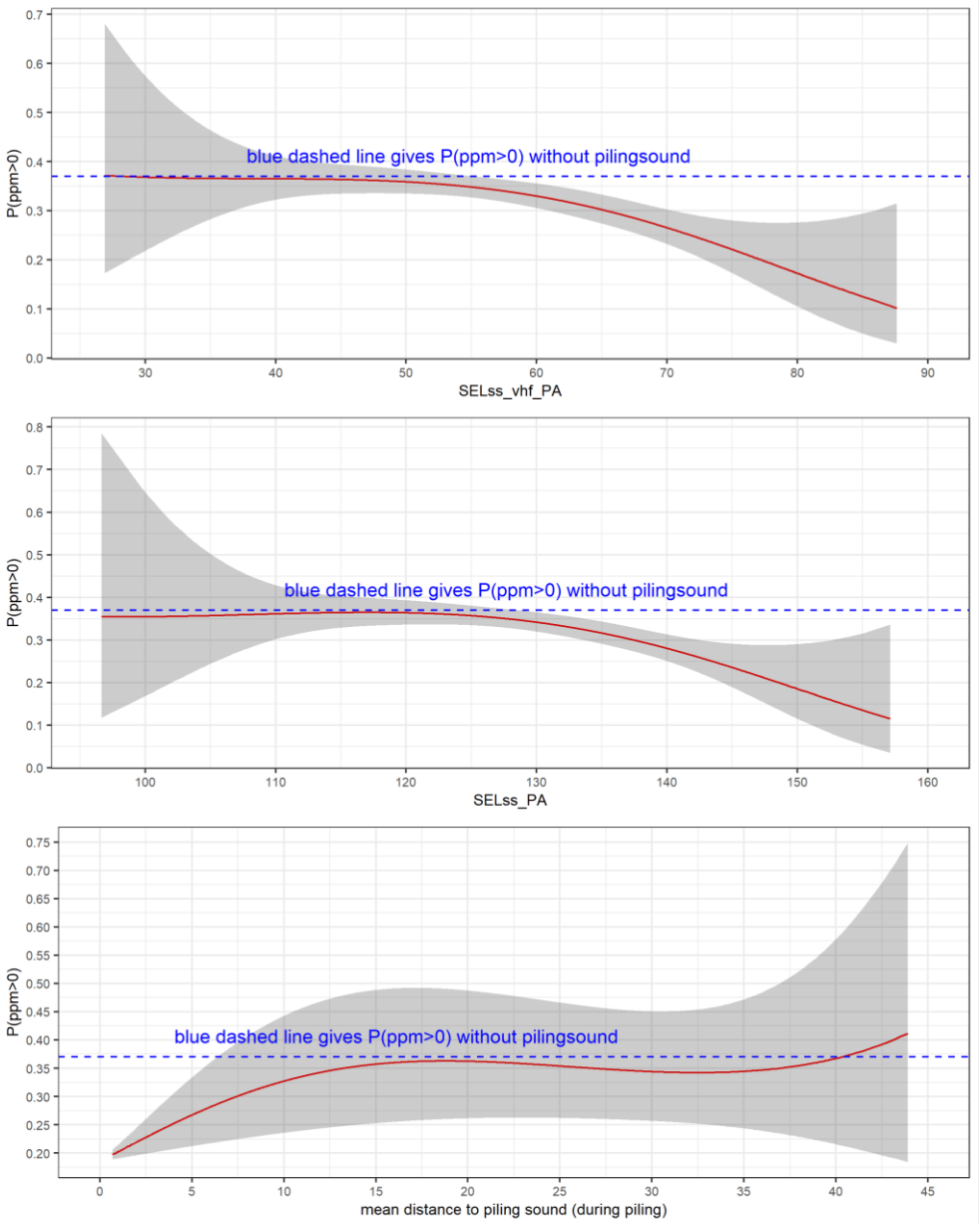

Figure 23 Comparison of the main effects of the three Bernoulli models for the Borssele data. Upper: VHF-weighted $SEL_{SS,VHF}$(50-500Hz); Middle: unweighted SEL_{SS}(50-500Hz); Lower: distance to pile. The blue dashed lines give the average probability for detection porpoise clicks (PPM/h>0) in the area based on a model for the time periods in which no piling took place. The shaded area indicates the 95% confidence interval of the model predictions. Note that the vertical scales of the plots differ.

The three models had a very similar fit and model diagnostics, which indicates that this study does not lead to a direct answer to the question which of the main variables predicts porpoise behavioural response to piling sound best.

The curves in Figure 23 show that the data suggest a 95% probability that porpoise presence is reduced, compared with the reference without piling sound when:

- weighted broadband $SEL_{SS,VHF}$(50-500Hz) > 55 dB re 1 µPa^{2}s;
- unweighted broadband SEL_{SS}(50-500Hz) >127 dB re 1 µPa^{2}s;
- distance to the pile ≤ 7 km.

5.3 Gemini analysis

To assess the influence of pile driving on harbour porpoise presence, a similar approach to the analysis of the Borssele data was taken. Details and additional information on porpoise detections based on the data from the monitoring during the construction of the Gemini wind farms are given in appendix E.

Data

The detection of porpoise clicks was recorded in the period from 23 June 2015 to 17 February 2016 using 15 CPODs. Figure 24 illustrates the temporal and spatial distribution of porpoise click detections. Harbour porpoises were detected at all stations throughout the study period.

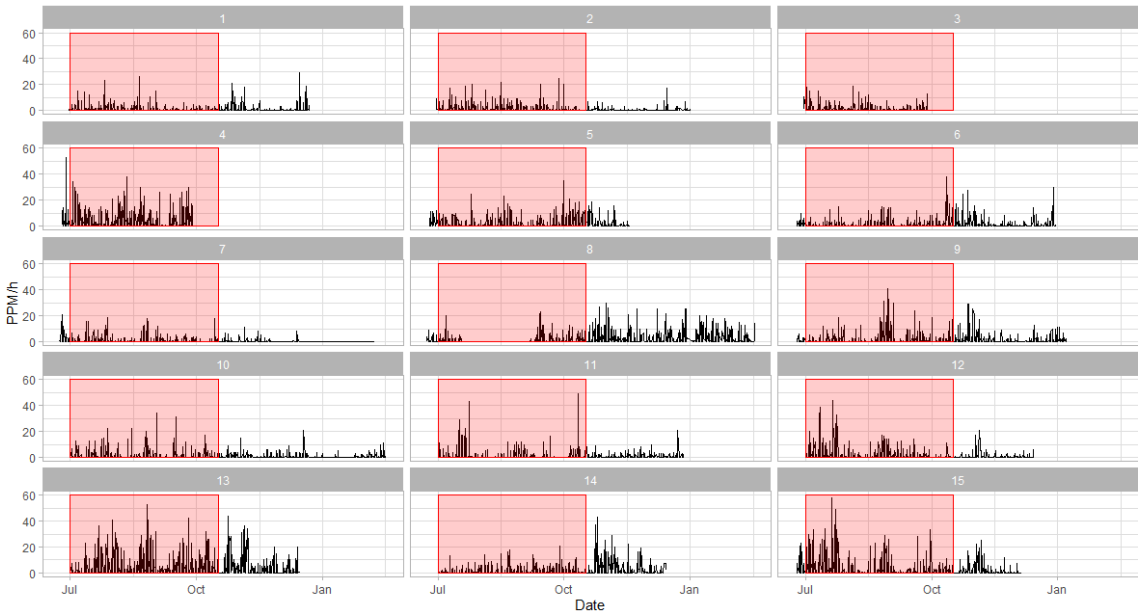

Figure 24 Porpoise acoustic activity in Porpoise positive minutes per hour (PPM/h) for the 15 CPOD locations during the study period 2015-06-23 to 2016-02-17. The shaded red time period indicates the piling period from start to end.

Extra control ('zero') data was added from earlier recordings in the same area, to increase the observation period outside the piling period from 9 to 96% of the total number of cumulative observational hours (n=168,160). These data were collected by WMR in the years 2010 to 2015, see (Geelhoed, et al., 2015) and (Geelhoed, Friedrich, Joost, Machiels, & Ströber, 2018). This addition of data was important to have a balanced data set, and create a thorough understanding of harbour porpoise presence and the variables that influence them in absence from any piling activity. A summary of this dataset is given in appendix D.2.

Methodology

The influence of piling on harbour porpoises was studied by looking at porpoise-click positive minutes in an hour. Both general porpoise presence (there are more than 0 porpoise positive minutes) and porpoise activity (the number of porpoise positive minutes in an hour) were considered as response variables. To stay consistent with previous analyses, results presented here are binomial analyses on general porpoise presence in an hour (yes/no, or 1/0). Results of the analysis on porpoise activity are presented in appendix D.5.

The detected porpoise clicks were aggregated per CPOD per hour and transformed into a binary response variable (the presence (PPM/h>0) or absence (PPM/h=0) of porpoise positive minutes). Porpoise presence per hour (yes/no) was modelled statistically by a generalized additive mixed model (GAMM) with a binomial error structure and a complementary logit link function and the restricted maximum likelihood method.

The main variable in the models was either the unweighted SEL_{SS}(10Hz-20kHz), the VHF-weighted $SEL_{SS,VHF}$(10Hz-20kHz), or the distance to the piling event. As the main predictor variables of our interest were strongly correlated, three separate models were created. The hourly mean of the unweighted and weighted SELss-values from the piling sounds received at the CPOD locations were obtained from sound exposure modelling, as described in §4.4. When no piling event took place, sound exposure levels were set to an approximate value of minimum observed background noise level, set at 100 dB for unweighted SEL_{SS}, and 80 dB for weighted $SEL_{SS,VHF}$.

For the three models, the same control variables were included. Tidal height was included, as harbour porpoises have demonstrated varying levels of activity depending on the tides in previous studies (IJsseldijk, Camphuysen, Nauw, & Aarts, 2015). As a proxy for annual patterns of porpoise presence in the area, both day of the year and water temperature were considered. Including both variables lead to collinearity issues. Consequently, two models that included either of the variables were compared. The model that included water temperature had the best fit and was selected for further analysis. Additionally, the model included the hour of the day to account for diurnal activity. Wind speed was included in the model to correct for the deficiency of CPODs to detect clicks for higher wind speeds. Lastly, the CPOD was included as a factorial predictor effect, as there might be a general spatial preference for some locations over others.

Most variables were fitted with a cubic spline basic function. Only the hour in the day was fitted with a cyclic spline, as the 24^{th} hour of the day must represent an almost similar value as the first hour of the day. The variables were modelled with a maximum of 4 basic functions (k) for interpretability. The model formulations for Gemini and Borssele differ only slightly. Details on the model structure and how this model is different from the statistical model in the Borssele analysis are shown in Appendix E.1.

Results

The models significantly diverged from the control models and had appropriate diagnostic plots, meaning that model results could be interpreted. As in the Borssele analysis, model fit and diagnostics were very similar for the three models, so that there is no clear result which predictor variable (SEL_{SS}, $SEL_{SS,VHF}$ or distance) is most appropriate for the analysis of harbour porpoise presence. Therefore, no conclusive answer to the question which of the main sound variables predicts porpoise behavioural response to piling sound best.

Results: General observations

When piling took place, mean overall porpoise presence (p PPM/h>0) and porpoise activity (PPM) decreased. When considering the effect for the CPOD locations separately, there was a spatial effect of the harbour porpoises moving away from the piling sites. For the CPODs in the construction zone, mean porpoise presence

and activity always decreased, but for the locations just outside the construction zone, mean porpoise presence and activity increased when pile driving events took place (appendix D.3). This could be explained by animals moving away from the construction zone, increasing relative abundance at these CPOD locations.

Results: Predictor variables
Figure 25 shows model plots for the probability of porpoise presence (p PPM/h>0) as a function of the three main variables, with control variables at the mean level. The reference line in the SEL_{SS} and $SEL_{SS,VHF}$ plots was established by making the corresponding models predict the harbour porpoise levels at background noise levels, where there is the assumption of no disturbance (100 and 80 dB, respectively). The reference line is the distance plot is the mean chance of harbour porpoise presence when no piling is taking place anywhere.

The curves in Figure 25 suggest a 95% probability that that porpoise presence is reduced relative to the observed maximum probability (solid line) when:

- weighted broadband $SEL_{SS,VHF}$(10Hz-20kHz) $\geq$ 100 dB re 1 µPa^{2}s;
- unweighted broadband SEL_{SS}(10Hz-20kHz) $\geq$ 146 dB re 1 µPa^{2}s;
- distance to the pile $\leq$ 15 km (relative to (p PPM>0 without piling).

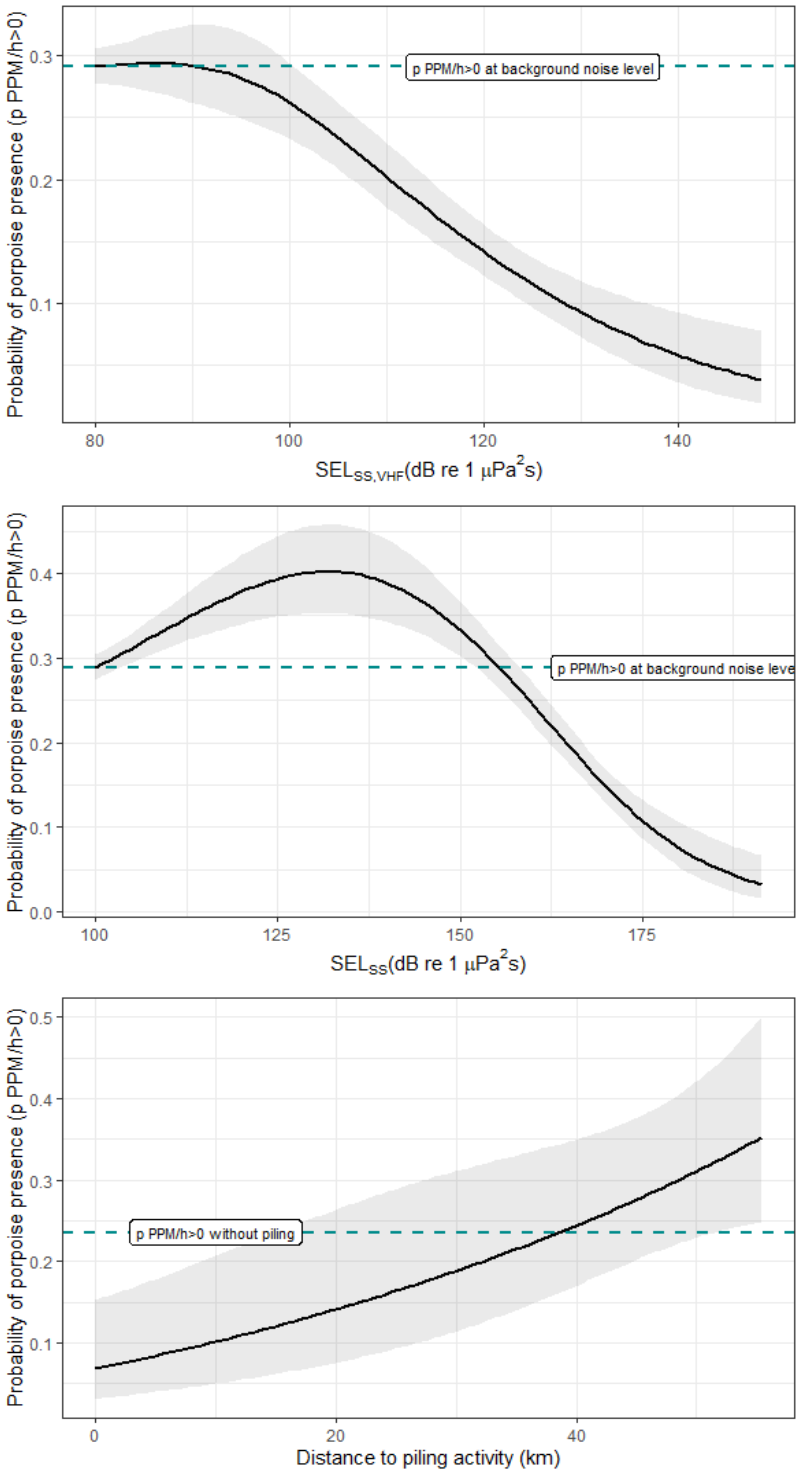

Figure 25 Comparison of the main effects of the three binomial models for the Gemini data. Upper: VHF-weighted $SEL_{SS,VHF}$(10Hz-20kHz); Middle: unweighted SEL_{SS}(10Hz-20kHz); Lower: distance to pile. The shaded area indicates the 95% confidence interval of the model predictions. The dashed lines for the unweighted SEL_{SS} and VHF-weighted $SEL_{SS,VHF}$ model are predicted probabilities of porpoise presence at the lowest exposure levels (where no disturbance is assumed) according to corresponding models. The reference line for the distance model is the mean probability of porpoise presence when no piling took place. Note that the vertical scales of the plots differ. Note also that the curve for unweighted SELss between 100 dB and 130 dB is uncertain, due to a lack of intermediate data points (see text below).

Results: Control variables
The fitted model predicts the probability of porpoise presence as function of all considered variables. Independent of the choice of the main variable, the model shows very similar effects of the control variables. Control variable plots can be found in appendix E.4.

The probability of PPM/h>0:
- slowly increased with temperature, with an optimum of around 9 degrees, and then stabilized. Putting this on a temporal scale, it would say that the porpoises avoid coming to the areas at the coldest months;
- decreased somewhat with increasing tidal flow magnitude;
- decreased with increasing wind speed;
- was lower during the day and higher during the night;
- varied between CPOD locations without clear spatial trend.

5.4 Discussion

For both Borssele (Figure 23) and Gemini (Figure 25) there is a clear effect of all three proposed predictor values (VHF-weighted SELss, unweighted SELss and distance) on the probability of porpoise presence.

For both Gemini and Borssele, there is an effect of VHF-weighted SELss on the probability of porpoise presence. Although absolute values of $SEL_{SS,VHF}$ at Gemini and Borssele cannot be compared due to different calculation methods, there is a resemblance in the effects' shape.

The model for unweighted SELss at Gemini suggests that the probability of encountering porpoises is highest at exposure to SELss around 132 dB.
The reduction of the probability towards lower exposure levels could be a modelling artefact, because there are no data for SELss values between 100 dB (background level in case of no piling) and 130 dB (the lowest calculated SELss of piling strikes, see Figure 20). It could also be that disturbed porpoises distribute to the edge of sound tolerance levels, although the unweighted SELss-model for Borssele and the other models for weighted SELss and distance do not seem to confirm this.
In future studies, we could possibly account for fleeing behaviour when creating a dose-response curve.

For both Gemini and Borssele, there is an effect of distance on the probability of porpoise presence, with a resemblance in the effects' shape. The results for distance should however be interpreted with care, as the 95% confidence interval is larger than for the SELss models. It is noteworthy, however, that the distances for disturbance seem larger for Gemini than Borssele, demonstrating the effectiveness of sound mitigation measures in Borssele.

In general, the probability of porpoise presence (p PPM/h > 0) :
- Is higher during summer and fall, when the water temperature is higher.
- Decreases somewhat with increasing tidal height (Gemini) or tidal flow (Gemini).
- Decreases with increasing wind speed, following our expectations, and most likely reflecting inability of CPODs to detect clicks at high wind speed levels.
- Was higher during the day and lower during the night in the Gemini area, opposite to what was observed in the Borssele analysis.

5.5 Dose-response relationship

For application in the framework for assessing ecological and cumulative effects of offshore wind farm construction (KEC), the analysis results must be translated into dose-response relationships that quantify the probability of disturbance as a function of (unweighted or weighted) sound exposure level.

It can be assumed that the relative difference between the probability of presence $P_{(\mathrm{PPM/h>0})}$ of porpoise positive minutes in a baseline situation (without piling) and in a situation with piling sound is generally independent of the covariate variables. In the KEC modelling it is assumed that the behavioural response of porpoises to piling sound does not depend on day of the year, hour of the day or location. Tidal flows and wind may affect the background sound in which porpoises are exposed to piling sound. This may have an effect on their response in situations where the piling sound is not much louder than the background. This effect is tentatively ignored, because it would require a significant modification of the KEC approach to include these effects, and it is in line with precautionary approach.

Hence, a dose-response relationship is estimated based on the predicted effect curves for mean values of the covariates, shown in Figure 23. As 'response' we quantify a *probability of disturbance* (P_{dist}) based on the relative difference between the probability of presence of porpoise positive minutes in a baseline situation (without piling) and in a situation with piling sound:

$$P_{\mathrm{dist}}(x) \approx 1 - \frac{P^{\mathrm{piling}}_{\mathrm{PPM/h>0}}(x)}{P^{\mathrm{baseline}}_{\mathrm{PPM/h>0}}}$$

The variable x can refer to unweighted SEL_{SS} or VHF-weighted $SEL_{SS,VHF}$.

For unweighted SEL_{SS} this dose-response relationship can be calculated from the Bernoulli model curves for Borssele (in Figure 23) and Gemini (Figure 25).

The unweighted SEL_{SS} is dominated by sound below 500 Hz, therefore the SEL_{SS}(10Hz-20kHz) and SEL_{SS}(50-500Hz) values of the two studies are comparable. The resulting dose-response curves are shown in Figure 26. They are compared with the dose-response curve for porpoise behavioural disturbance that is proposed in KEC 4.0.

For Gemini, we have tentatively selected as baseline the maximum value of the mean probability that PPM>0 s baseline, rather than the value observed 'without piling' dominated by data from years prior to the piling. This also accounts for the lack of data at SELss values between the arbitrary 'background' value (100 dB) and the lowest values for received piling sounds (130 dB).

This comparison confirms that the KEC 4.0 curve provides a conservative estimate of the dose-response relationship, as required by the limited availability of data when it was proposed. As explained in (Heinis, de Jong, & von Benda-Beckmann, 2022), it was decided then to adopt, as the worst-case scenario, the dose-effect relationship derived by Graham et al. (2019) for the response of harbour porpoises to the turbine foundation that was piled first. Possible habituation, leading to a reduced probability of disturbance when there are successive piling days, was disregarded as a precautionary measure. In the models for Borssele and Gemini the time from the start of construction was not included as covariate. Because the

piling for the nearby Belgian wind farms started well before the monitoring at Borssele, it is unlikely that adding this covariate would have led to reliable results.

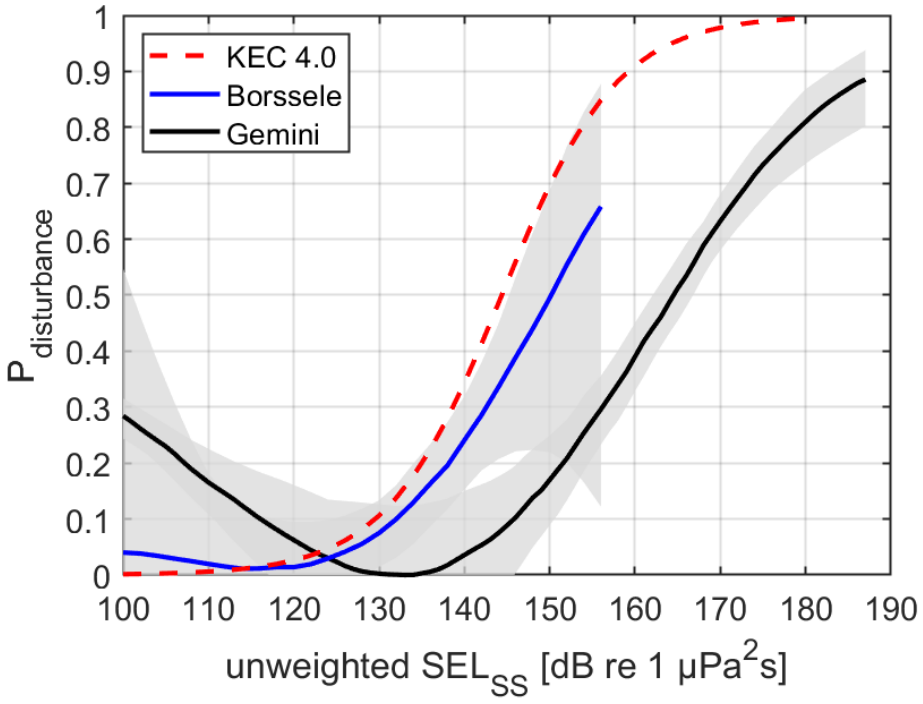

Figure 26 Dose-response relationship, quantifying the probability of disturbing porpoises as function of the unweighted broadband SEL_{SS} of piling sound to which they are exposed. The red dashed line shows the dose-response relationship proposed in the current KEC 4.0. The blue line is calculated from the model fitted to the Borssele monitoring data and the black line from the model fitted to the Gemini monitoring data. The shaded areas show the 95% confidence intervals of the models. Note that the curve for Gemini between 100 dB and 130 dB is uncertain, due to a lack of intermediate data points (see Figure 25).

There are multiple possible explanations why the modelled probability of disturbance by piling sound at Gemini appears to be substantially lower than at Borssele. The available data do not directly allow to determine the most likely explanation.

- The porpoises crossing the Gemini area during the piling period may be less susceptible to disturbance than the porpoises in the Borssele area.
- The Gemini piling was without noise mitigation, the Borssele piling with noise mitigation. Hence:
  - The same SEL_{SS} value corresponds with much larger distance from the pile at Gemini than at Borssele. Porpoises may be less susceptible to disturbance by piling sound at large distance, or it may be more difficult to detect the effects of piling sound at larger distance given the low number of CPODs at these distances.
  - Porpoises fleeing over the larger distances around the Gemini piling may be still detected by CPODs while fleeing, which may lead to the false interpretation that they are not disturbed and hence to underestimation of the disturbance distance.
  - The same SEL_{SS} value corresponds with a different sound spectrum at Gemini than at Borssele. Due to the applied mitigation measures at Borssele, the higher frequencies in the spectrum are more reduced than the lower frequencies. Hence, at the same SEL_{SS} value, the sound contains more high

frequencies at Gemini. Because porpoises are more sensitive to high-frequency sound, one would expect a stronger response to piling sound at the same SEL_{SS} value at Gemini than at Borssele. This is not confirmed by the dose-response curves shown in Figure 26. These suggest a stronger response at Borssele, which is not understood.

The dose-response relationships for weighted $SEL_{SS,VHF}$ calculated from the Bernoulli model curves for Borssele (in Figure 23) and Gemini (Figure 25) cannot be directly compared, due to the difference in frequency bandwidth of the modelled weighted $SEL_{SS,VHF}$.

Making the $SEL_{SS,VHF}$ values comparable requires an estimation of the acoustic energy in the frequency bandwidth between 500 Hz and 20 kHz that is missing in the Borssele data.

A first indication of that missing energy can be made on the basis of a second model from (Oud & de Jong, 2021) that included energy up to 1250 Hz instead of 500 Hz. Instead of repeating the Bernoulli modelling with this second data set, which was prohibited by time, a correction was derived, see Figure 27, that was applied to scale the $SEL_{SS,VHF}$ axis of the dose-response curve.

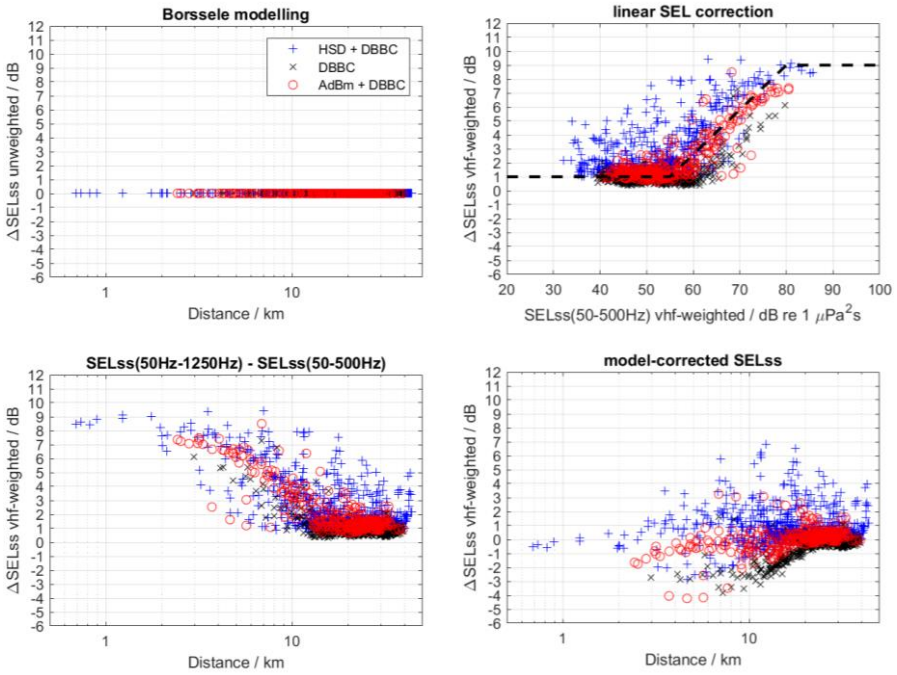

Figure 27 Difference between SEL_{SS}(50-1250Hz) and SEL_{SS}(50-500Hz) in the results from Borssele piling sound modelling (left figures; upper: unweighted and lower: weighted). Upper right figure: proposed linear correction (black dashed line) for the observed difference as a function of $SEL_{SS,VHF}$(50-500Hz). Lower right figure: error associated with the modelled correction.

The $SEL_{SS,VHF}$(50-1250Hz) dose-response curve from the Borssele model is compared with the $SEL_{SS,VHF}$(10Hz-20kHz) from the Gemini model in Figure 28. Note that the corrected $SEL_{SS,VHF}$ values for Borssele are likely still too low for a reliable comparison, due to missing energy values in higher frequency bands. Reducing the bandwidth of the Gemini spectra (Figure B.4) to 50-1250Hz lowers the $SEL_{SS,VHF}$ by about 15-20 dB, at the shorter distances (MP1 & MP2).

This cannot fully explain the observed 30 dB difference between the curves for Borssele and Gemini.

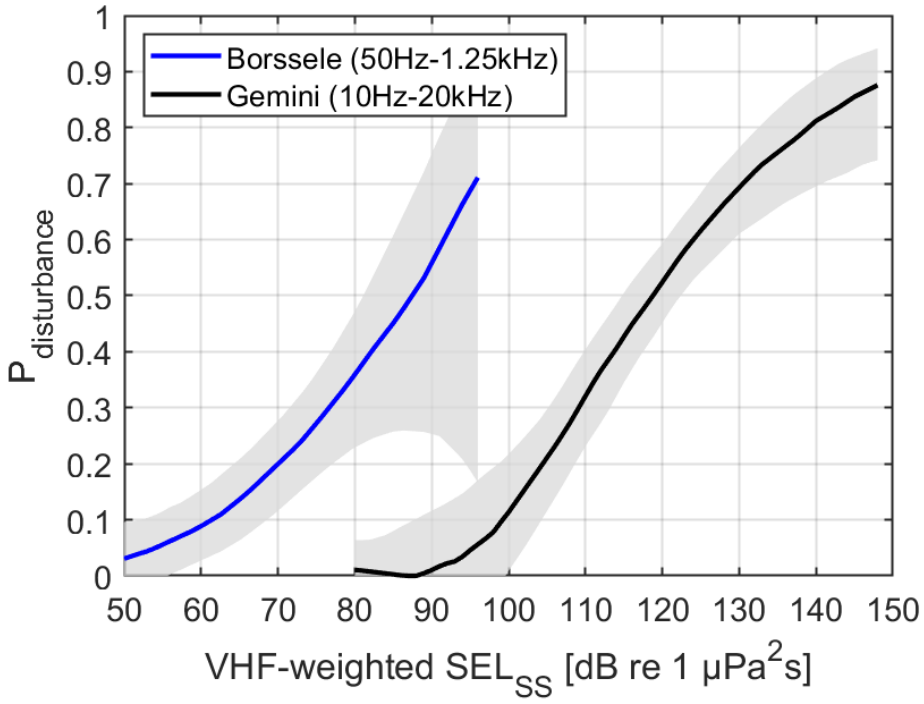

Figure 28 Dose-response relationship, quantifying the probability of disturbing porpoises as function of the weighted broadband $SEL_{SS,VHF}$ of piling sound to which they are exposed. The blue line is calculated from the model fitted to the Borssele monitoring data, corrected to $SEL_{SS,VHF}$(50-1250Hz), and the black line from the model fitted to the Gemini monitoring data, for $SEL_{SS,VHF}$(10Hz-20kHz). The shaded areas show the 95% confidence intervals of the models.

6 Porpoise behavioural response to ambient sound

In KEC and in the analyses presented in Chapter 5 it is assumed that piling sound is disturbing porpoises, leading to reduced detections of porpoise clicks. Disturbance by other factors than piling sound is ignored in these analyses. That assumption cannot explain observations that "porpoise detections decline several hours before the start of piling" (Brandt, et al., 2018) and that "higher vessel activity increased the probability of observing a response, which could indicate either a response of porpoises to vessels, a masking of porpoise detections on the CPOD by vessel noise, or both" (Graham, et al., 2019). It is not always clear if piling sound is the main source of disturbance.

Because other activities than piling have not been registered in detail during the monitoring periods, it is not possible to include these in the statistical modelling. However, at some locations the received underwater sound was monitored continuously together with the detection of porpoise clicks. These simultaneous recordings have been used to study the statistical relationship between received sound level (from all present sources, natural and anthropogenic) and porpoise click detections.

The total ambient sound at the CPOD locations is quantified in terms of the hourly unweighted and VHF-weighted values of broadband SPL: $L_{p,1\mathrm{h}}$ and $L_{p,1\mathrm{h,VHF}}$ as defined in Table 1. Porpoise presence is monitored in terms of CPOD porpoise positive minutes per hour (PPPM/h).

6.1 Borssele analysis

Ambient sound was recorded near 7 of the 16 CPOD locations in the Borssele area (see Chapter 3). This allows for a statistical modelling of the relationship between porpoise detections and received SPL at these locations.

Like in the SEL analysis (§5.2), a Bernoulli model with complementary log-log link function was built to model the presence (PPM>0) or absence (PPM=0) of the porpoise positive minutes. Details of the statistical modelling are described in Appendix F.

The main variable in the models was either the unweighted SPL(10Hz-20kHz), or the VHF-weighted $\mathrm{SPL_{VHF}}$(10Hz-20kHz), and the model included month of the year, hour of the day, wind speed, water temperature and tidal flow magnitude as covariates (low-rank thin-plate smoothers) and CPOD location as a random intercept.

In addition to the models of the relation between porpoise detections and ambient sound ("analysis C" in Appendix F), additional models were developed to investigate the effect of piling sound on the observed relationship. In "analysis B", two separate models were built, one for the time periods in which piling took place, and one for the periods without piling, and in "analysis A" the mean distance to the pile was added as a variable for the time periods in which piling took place.

The results of the three analyses (A, B and C) are very similar. Figure 29 shows the results for analysis C. The VHF-weighted SPL shows that the higher the $\mathrm{SPL_{VHF}}$,

the lower the probability of a porpoise positive minute. This relationship is quite linear, and the narrow confidence interval indicates that it is significant. The models for unweighted SPL show an approximately linear negative relation between PPM and unweighted SPL between a SPL of 130 dB to 140 dB re 1 µPa^2, but outside of that range the relationship appears to become flat, with a large uncertainty.

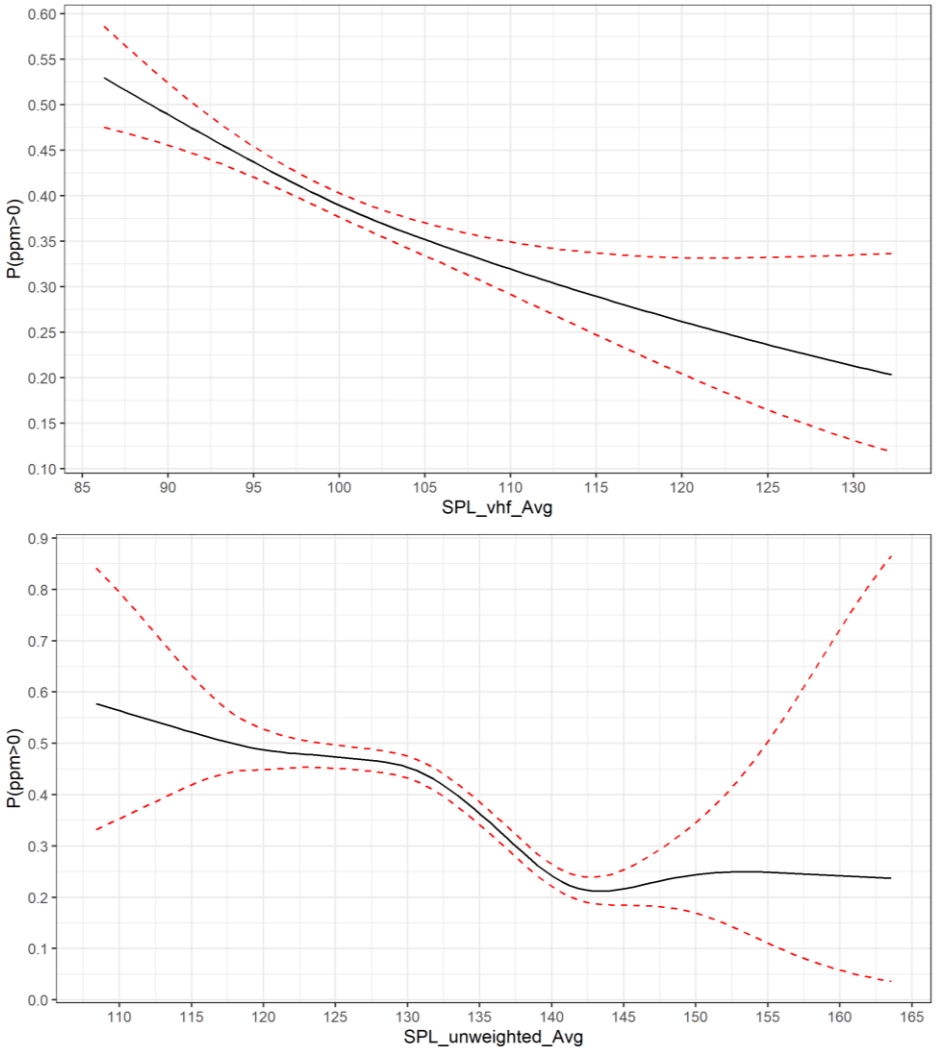

Figure 29 Predictor effect plot of the main effect of the VHF weighted SPL and unweighted SPL, from the Bernoulli models of analysis C. The plots show the predicted probability of occurrence of a porpoise positive minute (PPM>0) on the y-axis, while ignoring all random effects/smoothers (which average out to 0). The red dashed lines indicate the 95% confidence interval.

The uncertainty in the predictor-effect plot for unweighted broadband SPL may be partially explained because the unweighted SPL is regularly dominated by flow noise due to tidal currents, see §3.5.

Figure 30 shows that the unweighted and VHF-weighted SPL are much less cleanly related to each other than the unweighted and weighted SELss of piling strikes (Figure 21).

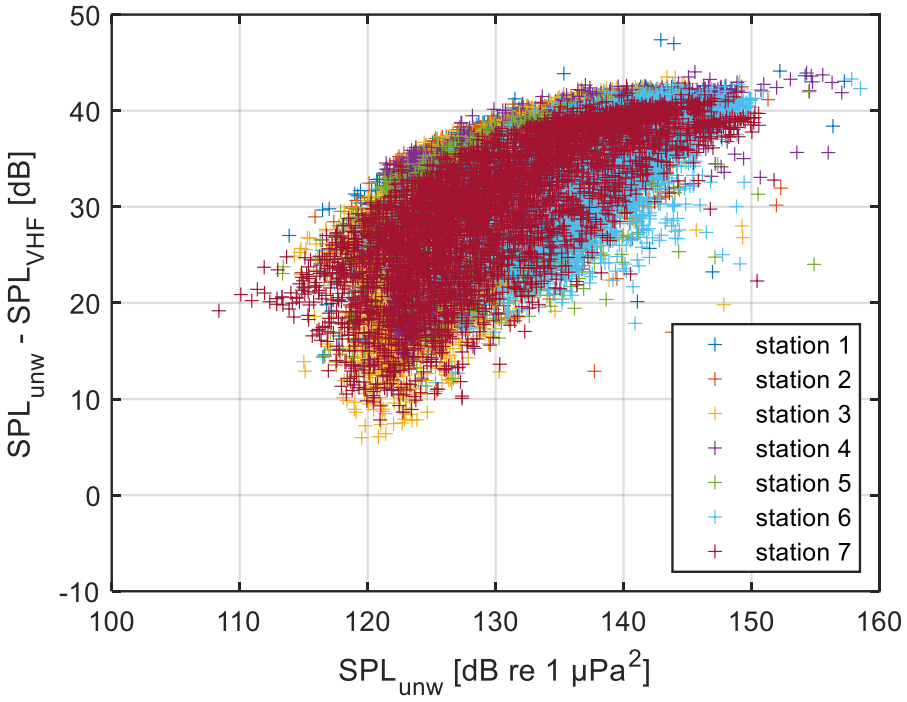

Figure 30 Difference between unweighted and VHF-weighted hourly SPL measured at the 7 stations in the Borssele area.

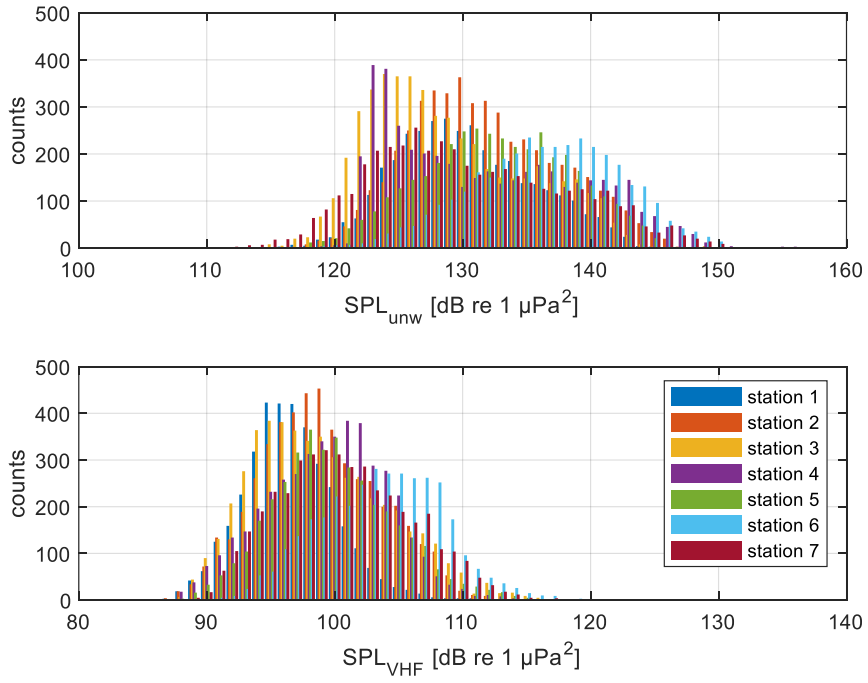

Figure 31 Histograms of the hourly mean SPL, unweighted (upper graph) and vhf-weighted (lower graph), as measured at the seven stations in the Borssele area.

Interpretation of the modelled relationship between porpoise detections and unweighted or VHF-weighted SPL needs further investigation. It cannot yet be excluded that background noise has affected porpoise click detections (Clausen, Tougaard, Carstensen, Delefosse, & Teilman, 2019), so that the reduced detections at higher SPL are not necessarily an indication of reduced porpoise presence. The available data do not allow studying this effect, because the SPL recordings do not extend to the frequency range of the clicks (>100 kHz).

The piling covariates (whether it be the mean distance to piling sound effect in analysis A or the SPL × piling active interaction in analysis B) were not found to be relevant.

As to be expected, the effects of the other covariates (wind speed, tidal flow magnitude, and the temporal smoothers) are very similar to what was observed in the SEL analysis (§5.2).

6.2 Gemini analysis

In Gemini, the ambient sound was recorded by underwater sound recorders (AMAR) near CPOD-location 1 and 8 (see Chapter 3). These data allowed for statistical modelling of the relationship between porpoise detections and SPL at these locations. The SPL analysis setup was consistent with Gemini's SEL analysis (§5.3). Porpoise presence per hour (yes/no) was modelled statistically by a generalized additive mixed model (GAMM) with a binomial error structure and a complementary logit link function, and the restricted maximum likelihood method.

The model predictor variable was either the unweighted SPL ($SPL_{unweighted}$, 10Hz-20kHz) or the VHF-weighted SPL (SPL_{VHF}, 10Hz-20kHz). The same control variables as in the SEL analysis (tidal height, water temperature, hour of the day, wind speed, and CPOD) were included. Most variables were fitted with a cubic spline basic function, but the hour in the day was fitted with a cyclic spline. The variables were modelled with a maximum of 4 basic functions (k) for interpretability.

Results

The models significantly diverged from the control models and had appropriate diagnostic plots, meaning that model results could be interpreted. The model fits for $SPL_{unweighted}$ and SPL_{VHF} were very similar. Figure 32 shows model plots for the probability of porpoise presence (p PPM/h>0) as a function of $SPL_{unweighted}$ and SPL_{VHF}, with control variables at the mean level. The shape is formed as a potential slight increase in the first part of the curve, at the lowest SPL values, (but not definite, see confidence intervals), after which a logistic-like decrease follows. The probability of porpoise presence drops from 106 dB re 1 μPa^2s for $SPL_{unweighted}$ and from 77 dB re 1 μPa^2s for SPL_{VHF}. The shape of the relationship is very similar to that of the SEL analysis in Gemini. The consistency through these analyses testifies to the robustness of the fitted model shape and the potential biological validity of the observed relation.

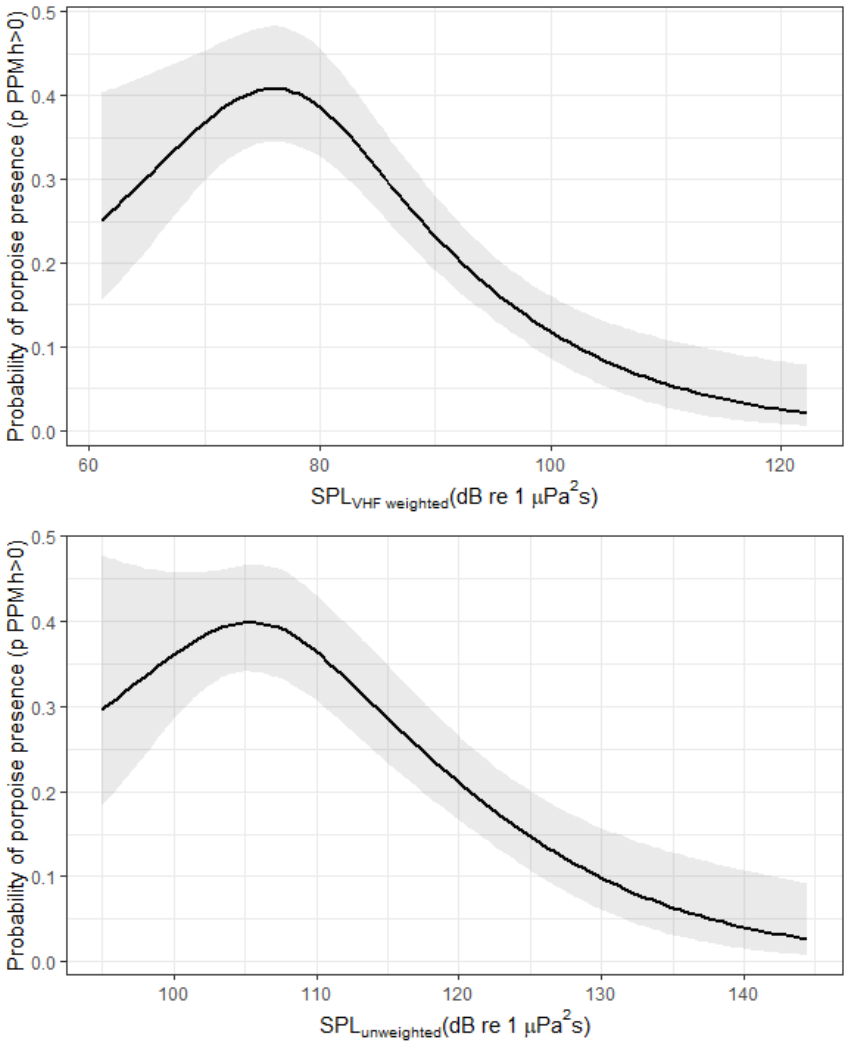

Figure 32 Predictor effect plot of the main effect of the unweighted SPL and VHF weighted SPL and, from binomial models of analysis D. The plots show the predicted probability of occurrence of a porpoise positive minute (p PPM>0) on the y-axis, while all other effects are at mean level. The grey shaded area represents the 95% confidence interval.

In contrast with the observations at Borssele, these plots show a clear maximum in the probability of porpoise presence, which allows for an estimation of the threshold value above which there is 95% probability that porpoise presence is reduced:

- weighted broadband SPL_{VHF}(10Hz-20kHz) > 83 dB re 1 µPa;
- unweighted broadband SPL(10Hz-20kHz) >112 dB re 1 µPa.

These threshold values are at the lower end or below the range of SPL values measured at Borssele (Figure 26). Hence, the reduction of porpoise presence with increasing SPL beyond these threshold values appears to be consistent in the observations at Borssele and Gemini. The reduction towards lower SPL values does not seem to be significant.

As for the Borssele analysis, interpretation of the modelled relationship between porpoise detections and unweighted or VHF-weighted SPL needs further investigation.

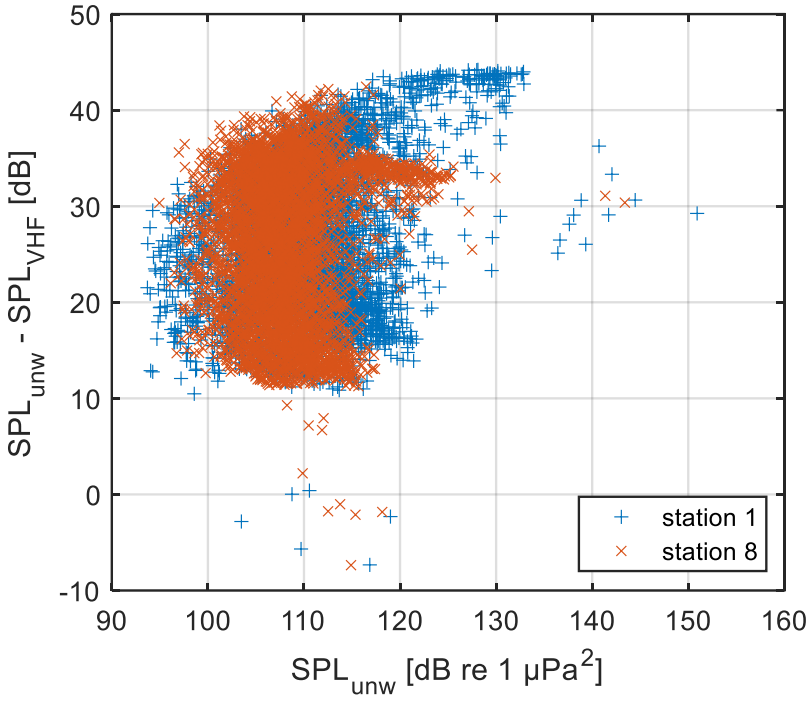

Figure 33 Difference between unweighted and VHF-weighted hourly SPL measured at the 2 stations in the Gemini area.

Figure 31 shows histograms of the recorded hourly SPL values at both stations at Gemini in 2013 and 2014. This gives an impression of the percentage of the time in which a given threshold level is exceeded. For example, Figure 32 shows that the probability of occurrence of porpoise positive minutes is at less than half of its maximum at unweighted SPL values greater than 121 dB. The histogram shows that this condition occurs about 2% of the time at station 1 and 6% of the time at station 8. For the weighted SPL_{VHF}, this threshold is at 92 dB, and it is exceeded about 17% of the time at station 1 and 20% of the time at station 8.

These percentages of time during which porpoise are disturbed locally due to exposure to underwater sound in a situation without piling, are of the same order of magnitude as the percentage of the time per day that they can be exposed to piling sound (about 2-3 hours per 24 hours, corresponding with 8-13% of the time).

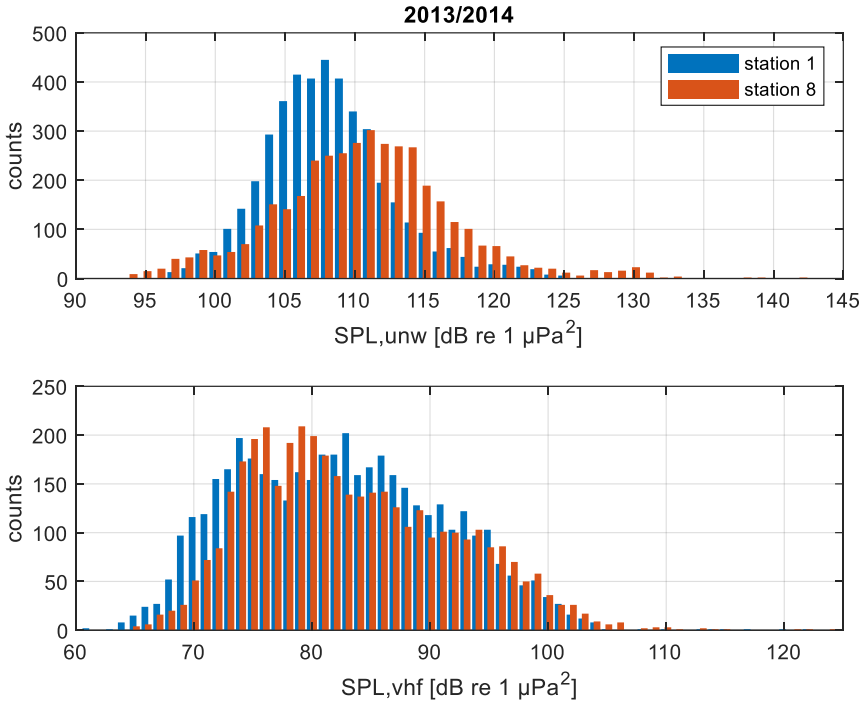

Figure 31 Histograms of the hourly mean SPL, unweighted (upper graph) and vhf-weighted (lower graph), as measured at the two stations (1 and 8) in 2013 and 2014. The levels are significantly lower than these observed in the Borssele area (Figure 31).

7 Summary and conclusions

7.1 Porpoise response to piling sound in the Borssele area

The passive acoustic monitoring during the construction of the Borssele offshore wind farms provided the opportunity to further investigate the effects of piling sound on harbour porpoises. The construction of the Borssele wind farms coincided with construction activities in Belgian wind farms, just across the border. These activities could (partly) be included in the analysis thanks to the cooperation from the Royal Belgian Institute of Natural Sciences (RBINS).

In general, the conditions at Borssele were not particularly favourable for an undisturbed observation of effects. The underwater sound in the area is dominated by busy nearby ship traffic, while the piling activities were carried out with additional noise mitigation, driven by underwater noise limits in the permits. Moreover, piling activities in the area started well before the monitoring started, so there are no reference data for the period prior to the wind farm construction (T-0).

From the acoustic data analysis, it could be concluded that piling sound is clearly detected at distances up to about 10 km from the piles. Thanks to the effectiveness of the applied noise mitigation measures, the piling sound at larger distances often disappears in the high background sound levels from nearby ship traffic. Due to masking by background sound, the piling sound could only be quantified reliably in the frequency range between 50 and 500 Hz.

Statistical analysis of the harbour porpoise acoustic detections (PPM: porpoise positive minutes recorded by the CPODs) as a function of the distance to the piling indicates that porpoises avoid piling up to a distance of about 15 km. However, the statistical significance of the observed reduction in PPM is already lost at 7 km.

The probability of porpoise detections starts to decrease significantly when the unweighted broadband SEL_{SS}(50 Hz – 500 Hz) of the piling sound exceeds 130 dB re 1 μPa^2s and when the porpoise-weighted broadband (50 Hz – 500 Hz) piling $SEL_{SS,VHF}$(50 Hz – 500 Hz) exceeds 55 dB re 1 μPa^2s.

The three Bernoulli models, modelling the presence (PPM>0) or absence (PPM=0) of the porpoise detections as a function of either distance, unweighted SEL_{SS} or porpoise-weighted $SEL_{SS,VHF}$, showed only small differences in the Bayesian and Aikake information criteria (BIC and AIC), indicating that these three metrics predicted the porpoise detections equally good. The similar fit and model diagnostics of the three Bernoulli models underpin this conclusion. However, the distance based Bernoulli model has a larger confidence interval than the SEL-based models. Therefore, it can be concluded that the two SEL-based Bernoulli models performed slightly better than the distance-based Bernoulli model.

The Borssele study did not provide a conclusive answer to the question whether frequency-weighted $SEL_{SS,VHF}$ is a better predictor for porpoise disturbance than the currently used unweighted SEL_{SS}. Distance, SEL_{SS} and $SEL_{SS,VHF}$ are strongly correlated at the distances where the piling sound can be detected, due to the applied sound mitigation measures and the high background noise levels in the area. Differences in the trend of SEL_{SS}(distance) and $SEL_{SS,VHF}$(distance) are

smaller than the variability in the modelled sound exposure levels at the CPOD locations for the various piles.

7.2 Porpoise response to piling sound in the Gemini area

After analysing the data from the monitoring during the construction of the Borssele wind farms, we recommended to repeat the analysis for offshore piling projects with more predominant piling sound. The data from the monitoring during the construction of the Gemini wind farm in 2015, without sound mitigation and in an area with less ship traffic in the vicinity, were available and suited for such an analysis. Predictions of the distance at which porpoises were expected to avoid piling sound during the construction of the Gemini wind farm, based on unweighted SEL_{SS}, appeared to be much larger than measured.

From the three different models in the SEL acoustic data analysis, it could be concluded that piling sound negatively influenced the probability of porpoise presence in the Gemini area. When piling took place, harbour porpoises appeared to move away from the CPOD locations closest to the piling locations to locations further away. The analysis showed no clear distance after which the probability of porpoise presence stabilizes. However, the distances at which porpoise presence was reduced are clearly larger than in Borssele, which is in line with the expected effect of sound mitigation efforts in Borssele.

For unweighted SEL values, the probability of porpoise presence decreased above 132 dB re 1 μPa^2s. For weighted SEL values, the likelihood of encountering porpoises decreases above 90 dB re 1 μPa^2s, corresponding to the predicted porpoise presence probability level at background noise levels.

Although the shapes of harbour porpoise presence in relation to the SEL variables generally follow expectations, there were some difficulties explaining the 'bump' in the porpoise presence in the unweighted SEL analysis. The bump in these types of relations is not uncommon (e.g., (Sarnocińska, et al., 2020)). It was considered that this might be due to the effect that porpoises that avoid high noise levels concentrated around locations where the noise levels are acceptable. However, this effect is not clearly observed in the weighted SEL or Borssele analysis. The shape of the relation introduces difficulty in choosing a biologically appropriate reference level. As a reference one might consider the probability of porpoise presence at the lowest sound exposure level (complete left of the relation in the plot), but one could also consider using the highest probability of porpoise presence in the curve after which the harbour porpoise presence starts to decrease. The latter was considered most appropriate here, but it is open to debate and further study. One solution to this problem could possibly be the fitting of a non-additive generalized linear model to the data, forcing the relationship to take on a sigmoid shape where no 'bump' can occur.

The three models fitting either distance, unweighted SEL_{SS}, or weighted $SEL_{SS,VHF}$ showed only minor differences in the model diagnostic plots, and information criteria BIC and AIC, indicating that these three variables are comparable in predicting porpoise presence. However, the distance-based binomial model had larger confidence intervals than the SEL-based models. Therefore, the two SEL-based binomial models performed slightly better than the distance-based binomial model. Like the Borssele analysis, the Gemini analysis did not provide a conclusive answer

to whether frequency-weighted $SEL_{SS,VHF}$ is a better predictor for porpoise disturbance than the currently used unweighted SEL_{SS}.

It could be that differences observed between the porpoise responses to the Borssele and Gemini piling have to do with the different environments. The different 'levels' at which changes in behaviour occur could therefore be a spatial effect. If we want more general conclusions for the impact of pile driving for wind farm construction on harbour porpoise presence, it would be valuable to do a combined analysis of the Gemini and Borssele data.

7.3 Porpoise response to ambient sound in the Borssele area

It is likely that porpoises respond to other underwater sound from human activities as well as to piling sounds. Previous studies, such as (Brandt, et al., 2018), show that porpoises avoid offshore wind farm construction sites before the start of piling.

Statistical analysis of the harbour porpoise detections (PPM/h) against the measured broadband (10 Hz – 20 kHz) sound pressure level (SPL) of continuous ambient sound at seven locations in the Borssele area shows that detections decrease with increasing SPL.

Analysis of the ambient sound recordings showed that the unweighted broadband SPL is regularly dominated by flow noise due to tidal currents, see (de Jong, 2021). Swimming porpoises do not experience flow noise in the same manner as the sound recorders. That might explain why the relationship between PPM/h and unweighted SPL is less certain. The fitted model shows a linear decrease of PPM/h and with increasing unweighted SPL in the range between 130 dB and 140 dB re 1 µPa, but outside of that range the relationship becomes flat, with a large uncertainty.

The fitted model for the porpoise-weighted SPL_{VHF} shows a reasonably clear linearly decrease of PPM/h with increasing SPL_{VHF} over the full range of measured values of SPL_{VHF}. Hence, a ‘baseline’ is lacking, therefore it is not possible to derive an absolute dose-response function for porpoise presence as a function of SPL_{VHF}.

Nevertheless, this analysis suggests that the response of the porpoises in the Borssele area is not limited to piling sound exposure. This suggests that SELss or distance based dose-response functions for exposure to piling sound are only meaningful where piling sound clearly dominates the sound exposure.

7.4 Porpoise response to ambient sound in the Gemini area

Statistical analysis of the harbour porpoise detections (pPPM/h) against the measured broadband (10 Hz – 20 kHz) sound pressure level (SPL) of continuous ambient sound at two locations in the Gemini area shows that detections decrease with increasing SPL, which is consistent with the results in the Borssele area.

The probability of porpoise presence drops from 112 dB re 1 μPa^2s for $SPL_{unweighted}$ and from 83 dB re 1 µPa for SPL_{vhf}. These threshold values are at the lower end or below the range of SPL values measured at (Figure 26). Hence, the decrease of porpoise presence with increasing SPL beyond these threshold values appears to be consistent in the observations at Borssele and Gemini.

7.5 Conclusions about frequency weighting and dose-response functions

Data from the monitoring during the construction of the Borssele and Gemini offshore wind farms was analysed to investigate if an acoustic metric weighted for the hearing sensitivity provides a better prediction of behavioural response of harbour porpoises to piling sounds than an unweighted metric.

The current version of KEC (Heinis, de Jong, & von Benda-Beckmann, 2022) uses a dose-response function for the behavioral disturbance of porpoises by piling sound that is based on the unweighted broadband single strike SEL_{SS}.
This dose-response functions was derived from previously published results of a study in Scotland (Graham, et al., 2019).

Analysis of the data from Borssele and Gemini monitoring leads to the conclusion that the currently used dose-response relationship predicts a larger response than observed at these two locations, leading to a precautionary prediction of the impact of the piling sound on porpoises.

Due to the applied sound mitigation measures during the piling at Borssele, the measured SEL_{SS} levels are lower than at Gemini, resulting in shorter disturbance ranges.

Though there are valid reasons to assume that behavioural response is better predicted by sound exposure metrics that incorporate a weighting for the frequency-dependent sensitivity of animal hearing than by unweighted metrics, the analysis of the data from Borssele and Gemini does not confirm this.

This study highlights problems associated with the practical implementation of hearing-weighted metrics. This is complicated by uncertainties in numerical modelling of piling sound (see Appendix B) as well as by uncertainties in the measurement of piling sound at higher frequencies where it is subject to masking by other sounds. Consequently, the data analysis has not led to a consistent dose-response relationship for weighted sound exposure metrics. More work is needed. As a first step, a comprehensive statistical analysis of the combination of the data from Borssele and Gemini might result in more stable and consistent dose-response relationships, though the situation in the two areas is quite different, with mitigated piling sound in a high background sound at Borssele and unmitigated piling sound in the much quieter environment at Gemini.

Another lesson learned from the observations at Borssele and Gemini is that it is not always clear if piling sound is the main source of disturbance, particularly when piling sound is reduced by bubble curtains or other mitigation measures.
A consistent decrease of porpoise presence with increasing SPL is observed at Borssele and Gemini. The (unweighted or weighted) SPL-metric has the advantage that it accounts for all sound, including piling as well as ships and construction activities. Hence, it likely provides a completer metric of the impact of sound exposure on porpoise behaviour. However, prediction of the SPL of all sound requires more information and models than currently available, consequently it is not straight away suitable for impact assessment.

8 Acknowledgement

This study if performed as part of the Wozep ecological programme, a central and long-term research programme initiated by Rijkswaterstaat on behalf of the Ministry of Economic Affairs and Climate. Niels Kinneging (RWS) was the project officer for this project (zaak 31163293). This report describes the results of a complex effort that could only be performed by combining the expertise from various disciplines (WaterProof, WMR, and TNO), and thanks to data contributions from RWS, Van Oord, Ørsted, RBINS and Gemini Wind Park.

9 References

Ainslie, M. (2010). *Principles of Sonar Performance Modeling.* Springer-Praxis.

Allaire, J., Xie, Y., McPherson, J., Luraschi, J., Ushey, K., Atkins, A. & Iannone, R. (2019). *Rmarkdown: Dynamic documents for R.* Retrieved from https://github.com/rstudio/rmarkdown.

Bakka, H., Rue, H., Fuglstad, G.-A., Riebler, A., Bolin, D., Illian, J. & Lindgren, F. (2018). Spatial modelling with R-INLA: A review. *WIRES (Invited Extended Review)*. Retrieved from http://arxiv.org/abs/1802.06350

Berges, B., Geelhoed, S., Scheidat, M. & Tougaard, J. (2019). *Quantifying harbour porpoise foraging behaviour in CPOD data: identification, automatic detection and potential application.* Wageningen Marine Research report C039/19.

Binnerts, B., de Jong, C., Ainslie, M., Nijhof, M., Müller, R. & Jansen, H. (2016). *Validation of the Aquarius models for prediction of marine pile driving sound.* report TNO 2016 R11338.

Booth, C., Heinis, F. & Harwood, J. (2018). *Updating the Interim PCoD Model: Workshop Report - New transfer functions for the effects of disturbance on vital rates in marine mammal species.* Report Code SMRUC-BEI-2018-011, submitted to the Department for Business, Energy and Industrial Strategy (BEIS).

Brandt, M., Dragon, A.-C., Diederichs, A., Bellmann, M., Wahl, V., Piper, W., . . . Nehls, G. (2018). Disturbance of harbour porpoises during construction of the first seven offshore wind farms in Germany. *Mar. Ecol. Prog. Ser., 596*, 213-232.

Brasseur, S., Aarts, G. & Schop, J. (2022). *Measurement of effects of pile driving in the Borssele wind farm zone on the seals in the Dutch Delta area- version II.* Texel: Wageningen Marine Research, DRAFT REPORT.

Brinkkemper, J. (2021). *Processing underwater pile-driving noise measured during construction OWF Borssele: Methodology and benchmark.* WaterProof report WP2020_1246_R1r2.

Brinkkemper, J., Geelhoed, S., Berges, B., Noort, C., Nieuwendijk, D. & Verdaat, J. (2021). *Underwater sound measurements – During the installation of the Borssele OWF.* WaterProof report WP2019_1173_R9r1.

Brooks, M., Kristensen, K., van Benthem, K., Magnusson, A., Berg, C., Nielsen, A. & Bolker, B. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal, 9*(2), 378-400. Retrieved from https://journal.r-project.org/archive/2017/RJ-2017-066/index.html

Clausen, K., Tougaard, J., Carstensen, J., Delefosse, M. & Teilman, J. (2019). Noise affects porpoise click detections – the magnitude of the effect depends on logger type and detection filter settings. *Bioacoustics, 28*(5), 443-358.

de Jong, C. (2021). *Borssele underwater noise SPL analysis (zaak 31163293 "Metingen onderwatergeluid Windparken Borssele", WP2).* memorandum TNO 2021 M11756.

de Jong, C. & von Benda-Beckmann, A. (2018). *Wozep underwater sound: frequency sensitivity of porpoises and seals.* report TNO 2017 R11238.

de Jong, C., Binnerts, B., Prior, M., Colin, M., Ainslie, M., Mulder, I. & Hartstra, I. (2019). *Wozep – WP2: update of the Aquarius models for marine pile driving sound predictions.* report TNO 2018 R11671.

de Jong, C., Binnerts, B., Robinson, S. & Wang, L. (2021). *Guidelines for modelling ocean ambient noise. Report of the EU INTERREG Joint Monitoring Programme for Ambient Noise North Sea (Jomopans).* The Hague: TNO.

Diederichs, A., Pehlke, H., Nehls, G., Bellmann, M., Gerke, P., Oldeland, J., . . . Rose, A. (2014). *Entwicklung und Erprobung des Großen Blasenschleiers zur Minderung der Hydroschallemissionen bei Offshore-Rammarbeiten (BMU Förderkennzeichen 0325309A/B/C).* Husum: BioConsult SH.

Dunn, P. & Smyth, G. (1996). Randomized Quantile Residuals. *Journal of Computational and Graphical Statistics, 5*(3), 236-244.

Garnier, S. (2018). *Viridis: Default color maps from 'matplotlib'.* Retrieved from https://CRAN.R-project.org/package=viridis.

Geelhoed, S., Friedrich, E., Joost, M., Jühre, H., Kirkwood, R., van Leeuwen, P., . . . Verdaat, H. (2015). *Gemini T-0: passive acoustic monitoring and aerial surveys of harbour porpoises.* IMARES Report number C144/15.

Geelhoed, S., Friedrich, E., Joost, M., Machiels, M. A. & Ströber, N. (2018). *Gemini T-c: aerial surveys and passive acoustic monitoring of harbour porpoises 2015.* Wageningen University & Research report C020/17.

Graham, I. M., Merchant, N. M., Farcas, A., Barton, T. R., Cheney, B., Bono, S. & Thompson, P. M. (2019). Harbour porpoise responses to pile-driving diminish over time. *R. Soc. open sci. 6: 190335*.

Heinis, F., de Jong, C., & Rijkswaterstaat Underwater Noise Working Group. (2015). *Cumulative effects of impulse underwater noise on marine mammals.* report TNO 2015 R10335-A.

Heinis, F., de Jong, C. & von Benda-Beckmann, A. (2022). *Framework for Assessing Ecological and Cumulative Effects 2021 (KEC 4.0) – marine mammals.* TNO 2021 R12503.

Heinis, F., de Jong, C. & von Benda-Beckmann, A. (2022). *Framework for Assessing Ecological and Cumulative Effects 2021 (KEC 4.0) – marine mammals.* The Hague: TNO 2021 R12503-UK.

Heinis, F., de Jong, C., von Benda-Beckmann, S. & Binnerts, B. (2019). *Kader Ecologie en Cumulatie – 2018, Cumulatieve effecten van aanleg van windparken op zee op bruinvissen.* HWE & TNO. Retrieved from https://www.noordzeeloket.nl/publish/pages/157579/kec_update_2018_effecten_impulsief_geluid_op_bruinvissen_20190124def.pdf

IJsseldijk, L., Camphuysen, K., Nauw, J. & Aarts, G. (2015). Going with the flow: Tidal influence on the occurrence of the harbour porpoise (Phocoena phocoena) in the Marsdiep area, The Netherlands. *Journal of Sea Research, 103*, 129-137.

ISO 18405. (2017). *Underwater acoustics - Terminology.* Geneva, Switzerland. Retrieved from https://www.iso.org/obp/ui/#iso:std:iso:18405:ed-1:v1:en

ISO 18406. (2017). *Underwater acoustics — Measurement of radiated underwater sound from percussive pile driving.* Geneva: ISO.

Kastelein, R., de Jong, C., Tougaard, J., Helder-Hoek, L. & Defillet, L. (2022). Behavioral responses of a harbor porpoise (Phocoena phocoena) depend on the frequency content of pile-driving sounds. *Aquatic Mammals, 48*(2), 97-109.

Kastelein, R., Hoek, L., de Jong, C. & Wensveen, P. (2010). The effect of signal duration on the underwater detection thresholds of a harbor porpoise (Phocoena phocoena) for single frequency-modulated tonal signals between 0.25 and 160 kHz. *J. Acoust. Soc. Am., 128*(5), 3211–3222.

Kastelein, R., van Heerden, D., Gransier, R. & Hoek, L. (2013). Behavioral responses of a harbor porpoise (Phocoena phocoena) to playbacks of broadband pile driving sounds. *Mar. Environ. Res, 92*, 206-214.

Lindgren, F. & Rue, H. (2015). Bayesian spatial modelling with R-INLA. *Journal of Statistical Software, 63*(19), 1-25. Retrieved from http://www.jstatsoft.org/v63/i19/

Lucke, K. (2015). *Measurement of underwater sound at the GEMINI Windpark site. (Unpublished Report).* Perth, Australia: Centre for Marine Science and Technology, Curtin University,.

Nabe-Nielsen, J., van Beest, F. M., Grimm, V., Sibly, R. M., Teilmann, J. & Thompson, P. M. (2018). Predicting the impacts of anthropogenic disturbances on marine populations. *Conservation Letters. e12563*.

Oud, T. & de Jong, C. (2021). *Borssele piling underwater noise modelling (zaak 31163293 "Metingen onderwatergeluid Windparken Borssele", WP2).* memorandum TNO 2021 M11758.

R Core Team. (2020). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rees, T. (2003). "C-squares," a new spatial indexing system and its applicability to the description of oceanographic datasets. *Oceanography, 16*. doi:10.5670/oceanog.2003.52

Remmers, P. & Bellmann, M. (2016). *Ecological monitoring of underwater noise during piling at Offshore Wind Farm Gemini.* ITAP report 2571-15.

RStudio Team. (2020). *RStudio: Integrated development environment for R.* Boston, MA: RStudio, PBC. Retrieved from http://www.rstudio.com/

Rue, H., Martino, S. & Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations (with discussion). *Journal of the Royal Statistical Society B, 71*, 319-392.

Sakamoto, Y., Ishiguro, M. & Kitagawa, G. (1986). *Akaike information criterion statistics.* Dordrecht, the Netherlands: D. Reidel.

Sarnocińska, J., Teilmann, J., Balle, J., van Beest, F., Delefosse, M. & Tougaard, J. (2020). Harbor porpoise (Phocoena phocoena) reaction to a 3D seismic airgun survey in teh North Sea. *Frontiers in Marine Science*(6), 824.

Schwarz, G. (1978). Estimating the dimension of a mode. *The Annals of Statistics*, 461-464.

Southall, B. L., Finneran, J. J., Reichmuth, C., Nachtigall, P., Ketten, D., Bowles, A. E., . . . Tyack, P. L. (2019). Marine Mammal Noise Exposure Criteria:Updated Scientific Recommendations for Residual Hearing Effects. *Aquatic Mammals*, 125-232.

Southall, B., Bowles, A., Ellison, W., Finneran, J., Gentry, R., Greene Jr, C., . . . Tyack, P. (2007). Marine Mammal Noise Exposure Criteria: Initial Scientific Recommendations. *Aquatic Mammals, 33*(4), 411-521.

Southall, B., Nowacek, D., Bowles, A., Senigaglia, V., Bejder, L. & Tyack, P. (2021). Marine mammal noise exposure: Assessing the severity of marine mammals behavioural responses to human noise. *Aquatic Mammals, 47*(5).

Tierney, L. (2018). *Codetools: Code analysis tools for R.* Retrieved from https://CRAN.R-project.org/package=codetools.

Tougaard, J. & Beedholm, K. (2019). Practical implementation of auditory time and frequency weighting in marine bioacoustics. *Applied Acoustics, 145*, 137-143.

Tougaard, J., Wright, A. & Madsen, P. (2015). Cetacean noise criteria revisited in the light of proposed exposure limits for harbour porpoises. *Marine Pollution Bulletin, 90*, 196–208.

Tyack, P. & Thomas, L. (2019). Using dose-response functions to improve calculations of the impact of anthropogenic noise. *Aquatic Conserv: Mar. Freshw. Ecosyst., 29*(S1), 242-253.

Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis.* Retrieved from https://ggplot2.tidyverse.org.

Wood, S. (2003). Thin-plate regression splines. *Journal of the Royal Statistical Society (B), 65*(1), 95-114.

Wood, S. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B), 73*(1), 3-36.

Wood, S. (2017). *Generalized additive models: An introduction with R* (2nd ed.). Chapman-Hall/CRC.

Xie, Y. (2016). *Bookdown: Authoring books and technical documents with R markdown.* Retrieved from https://bookdown.org/yihui/bookdown.

A Sound exposure modelling

A.1 Borssele acoustic modelling

To be able to investigate dose-response relationships for all 16 CPOD locations, the acoustic exposure at the 9 locations without SoundTrap acoustic recorder (see Figure 9) needs to be determined. For this purpose, TNO has modelled the piling sound at all 16 locations. This modelling could only be done for the Borssele and Seamade piles, because the piling logs for NorthWester2 were not available.

The details of the acoustic modelling, using the Aquarius 4 piling noise model (de Jong, et al., 2019), are described in (Oud & de Jong, 2021).

The SEL_{SS} values measured at the 7 SoundTrap stations have been used to calibrate the model. This calibration was needed to account for the noise mitigation measures (AdBm, HSD and DBBC), applied to meet the noise limits set in the permits for the wind farms, that cannot yet be modelled explicitly in Aquarius 4. This calibration also somewhat reduces uncertainties in the modelling associated with the not well specified acoustical properties of the local sediment and of the hammer impact force.

The following procedure was followed for the model calibration:

1 For each pile location x_i, sound exposure level $L_{E,\mathrm{SS}}^{Aq4}(x_j, x_i, E_{\mathrm{ref}}, f_n)$ is calculated at the locations x_j of the 7 sound recorders, at decidecade band centre frequencies. The Aquarius 4 model is used, for a reference hammer strike energy $E_{\mathrm{ref}} = 2000$ kJ and ignoring the noise mitigation system.

2 The modelled $L_{E,\mathrm{SS}}^{Aq4}(x_j, x_i, E_{\mathrm{ref}}, f_n)$ is scaled to the actual strike energy E_m for each recorded hammer strike:

$$L_{E,\mathrm{SS}}^{Aq4}(x_j, x_i, E_m, f_n) = L_{E,\mathrm{SS}}^{Aq4}(x_j, x_i, E_{\mathrm{ref}}, f_n) + 10\log_{10}(E_m/E_{\mathrm{ref}}) \text{ dB}$$

3 The modelled spectra $L_{E,\mathrm{SS}}(x_j, x_i, E_m, f_n)$ are compared with measured spectra $L_{E,\mathrm{SS}}^{\mathrm{meas}}(x_j, x_i, E_m, f_n)$ and the difference is quantified for each strike:

$$\Delta L(x_j, x_i, E_m, f_n) = L_{E,SS}^{Aq4}(x_j, x_i, E_m, f_n) - L_{E,SS}^{meas}(x_j, x_i, E_m, f_n)$$

4 An average calibration spectrum $\overline{\Delta L(x_j, x_i, f_n)}$ is calculated by averaging $\Delta L(x_j, x_i, E_m, f_n)$ over the series of pile strikes (m) for each pile - recorder combination.

5 In the ideal case, the $\overline{\Delta L(x_j, x_i, f_n)}$ spectra would not vary over the SoundTrap locations x_j, resulting in a single $\overline{\Delta L(x_i, f_n)}$ correction for the SELs for each individual pile at location x_i. However, initial comparisons showed large variability over x_j, see (Oud & de Jong, 2021). The uncertainty increases significantly with increasing distance between pile and SoundTrap recorder, possibly due to uncertainty in the propagation calculations but clearly also due to decreasing signal-to-noise ratios (SNR). Therefore, we decided to limit the calibration to pile-recorder distances smaller than 7.5 km and piling strikes at hammer energy 1200 kJ and above, to reduce SNR problems. This led to a

data significant reduction, so that initial attempts to determine a $\overline{\Delta L}$ estimation for individual piles were discarded in favour of a more generic $\overline{\Delta L}$ estimation. Because different noise mitigation measures were used, we decided to determine a $\overline{\Delta L_{\mathrm{mit}}(f_n)}$ correction for the three different applied noise mitigation measure combination (DBBC, AdBm + DBBC and HSD + DBBC), averaged over the individual $\overline{\Delta L(x_j, x_i, f_n)}$ values for all pile – SoundTrap combinations with a distance smaller than 7.5 km for each mitigation. The resulting calibration spectra are shown in Figure A.1.

6 The resulting three $\overline{\Delta L_{\mathrm{mit}}(f_n)}$ estimations are then subtracted from the calculated sound exposure level spectrum $L_{E,\mathrm{SS}}(x_j, x_i, E_m, f_n)$ at the locations of the 16 CPODs, for each individual hammer strike on each pile:

$$L_{E,SS}(x_j, x_i, E_m, f_n) = L_{E,SS}^{Aq4}(x_j, x_i, E_m, f_n) - \overline{\Delta L_{mit}(f_n)}$$

7 Finally, for each sensor location the 50% percentile, 95% percentile, 100% percentile and power average values of $\mathrm{SEL_{SS}}$ and $\mathrm{SEL_{SS,VHF}}$ are calculated for each hour in which pile driving took place. These were used for the statistical analyses reported in Chapter 5.

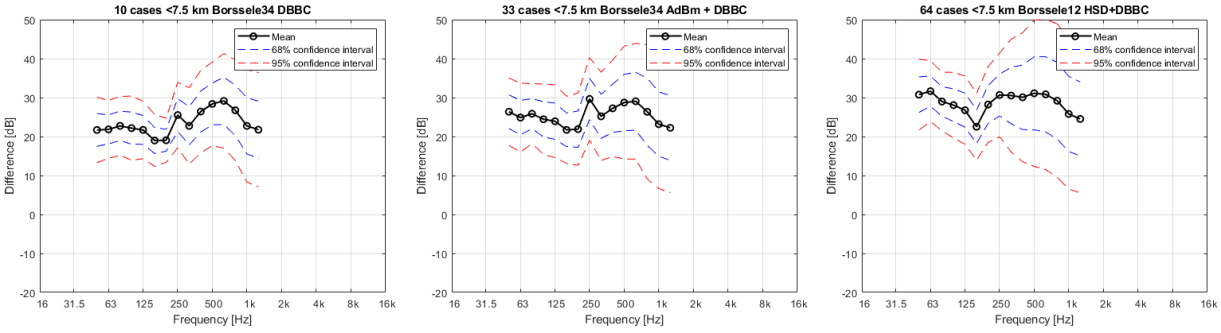

Figure A.1 Model calibration spectra $\overline{\Delta L_{\mathrm{mit}}(f_n)}$ for the three applied noise mitigation measure combinations: left: DBBC, middle: AdBm + DBBC and right: HSD + DBBC, with the associated confidence intervals. See (Oud & de Jong, 2021) for a detailed description.

The model calibration spectra (Figure A.1) are limited to the 50 Hz to 1250 Hz decidecade bands. Outside of these bands, piling sounds could not be reliably detected in the recordings, because they were masked by background noise, see also §3.5. All calibration curves show similar characteristics, including a minimum model-data difference between 125 and 250 Hz and slightly larger differences above 200 Hz.

Because of the increasing width of the confidence intervals with increasing frequency, only the frequency content below 500 Hz was included in the calculation of the weighted and unweighted broadband $\mathrm{SEL_{SS}}$(50 Hz – 500 Hz) metrics.

At frequencies below 125 Hz, adding a mitigation measure on top of DBBC results in a higher model-data difference, indicating a more effective noise reduction. The AdBm system adds about 3 to 4 dB reduction, the HSD system about 6 to 8 dB.

B Aquarius piling sound model evaluation

In (de Jong, et al., 2019) the Aquarius 4 pile driving sound model was validated using data from 4 measurement stations located at various distances from the Gemini U8 pile. The focus was then on the ability to predict the unweighted broadband SELss. Since that validation, TNO has adopted various improvements to their North Sea sound propagation modelling approach:

- A grainsize and frequency dependent model for the seabed acoustical parameters (both sound speed and attenuation) originating from the JOMOPANS project.
- A refinement of the receiver depth resolution (from values at 1 m from the sea surface and 1 m from the seabed towards a receiver distribution over the full water depth with 1 m step size).
- Improvement of the stability of the mode lookup table generation code for the propagation loss calculations, by implementation of various checks:
    - Requiring the bottom loss to be negative (otherwise remove mode).
    - Requiring the grazing angle to decrease for increasing water depth.
    - Interpolating missing modal wavenumbers (for intermediate depths), using an analytic equation for the modal mode normalization.
    - Figure B.1 provides an example of calculated modes.

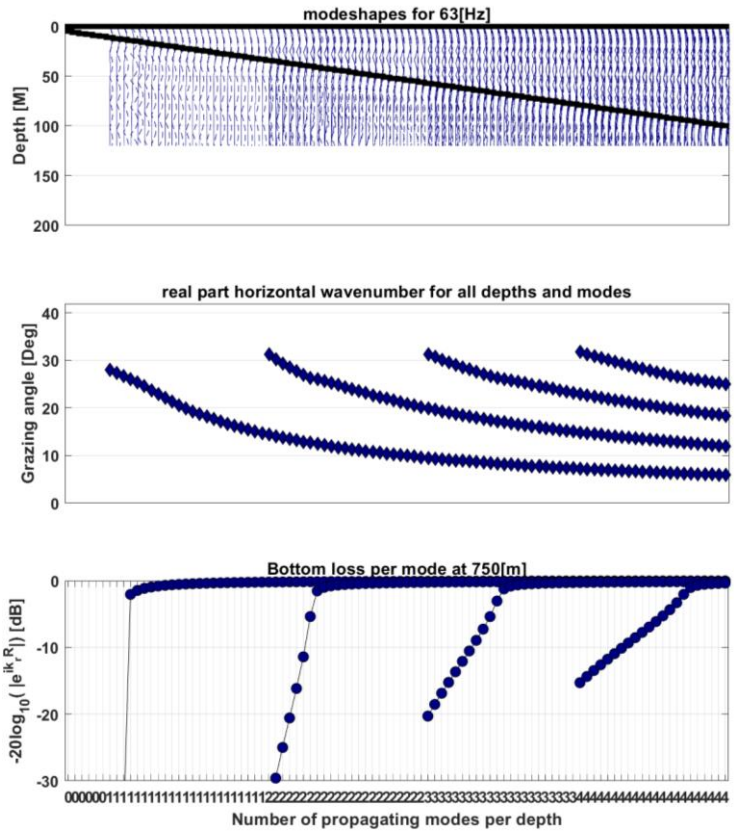

Figure B.1 phi =1.5 with Jomopans dispersion (V2: phi 1.5 model).

B.1 Model sensitivity studies

The sensitivity of the modelling results to the parameter and implementation variations have been tested in a series of model calculation versions for the Gemini U8 pile test case:

Model V0:
- Mode interpolation applied.
- Geo-acoustic parameters from table 4.18 (Ainslie, 2010).
- No dispersion in the sediment.
- SEL calculated at receiver grid, maximum over depth reported.

Model V1a:
- No mode interpolation applied.
- Geo-acoustic parameters from table 4.18 (Ainslie, 2010).
- Sediment dispersion model from (de Jong, et al., 2019).
- SEL at two receiver depths (near seabed and near sea surface).

Model V1b:
- Mode interpolation applied.
- Geo-acoustic parameters from table 4.18 (Ainslie, 2010).
- Sediment dispersion model from (de Jong, et al., 2019).
- SEL calculated at receiver grid, maximum over depth reported.

Model V2 (for phi =1 and phi=1.5):
- Mode interpolation applied.
- Geo-acoustic parameters from table 4.17 (Ainslie, 2010).
- Sediment dispersion model from (de Jong, Binnerts, Robinson, & Wang, 2021).
- SEL calculated at receiver grid, maximum over depth reported.

The geo-acoustic parameters used by the various model versions are shown in Figure B.2.

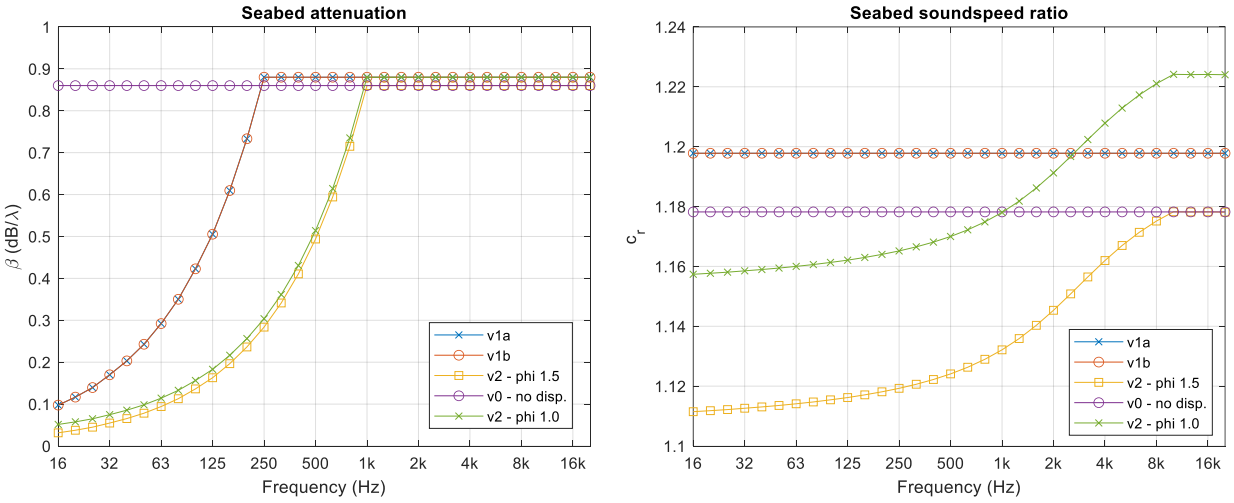

Figure B.2 Spectra of geo-acoustic parameters used by the various model versions.

Model configuration:
Aquarius 4 calculations are performed for a default hammer strike energy 2000 kJ. The measured SELss at the four receiver locations (MP1-4) is quantified in terms of the mean over the full series of piling strikes. A correction is applied for the variation of the hammer energy per strike during this series. In Aquarius 4, the SELss varies linearly with the logarithm of the hammer energy. Figure B.3 shows the correction per strike. In the model-data comparison, the mean measured SELss is compared with the modeled SELss for a hammer energy of 2000 kJ with a -3 dB correction.

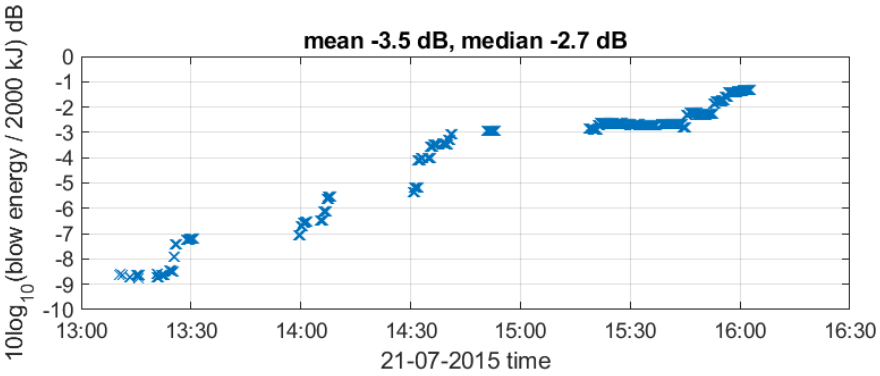

Figure B.3 Hammer strike energy correction per strike for the Gemini U8 pile.

Comparison with measurements:
- SELss spectra from (Remmers & Bellmann, 2016), measured at 2 m and 10 m above seabed.
- The mean + 1 standard deviation of the SELss of all strikes, measured at 10 m above the seabed, is assumed to be representative to assess the accuracy of the model predictions of the maximum SELss over the water depth.

The comparison of the measured SELss spectra at four distances from the U8 pile with predictions by from the various model versions is shown in Figure B.4 and Table B.1. Table B.2 gives the difference between the measured and modeled broadband SELss.

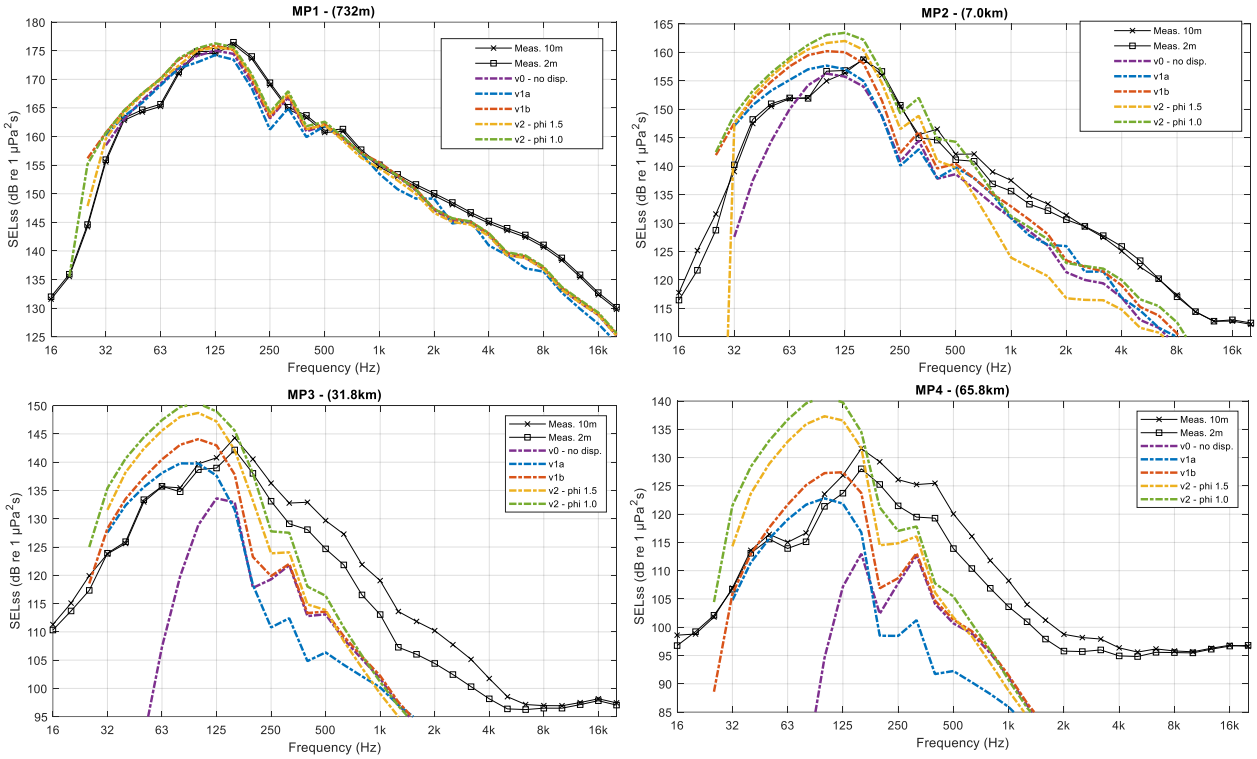

Figure B.4 Decidecade spectra of SELss at four distances of the U8 pile, measured (mean + 1 st.dev.) and calculated by the different model versions.

Table B.1 Broadband SELss values, measured and calculated by the different model versions.

	$SEL_{ss,unw}$ [dB re 1 µPa²s]				$SEL_{ss,VHF}$ [dB re 1 µPa²s]			
Location	MP1	MP2	MP3	MP4	MP1	MP2	MP3	MP4
Measured 10 m above seabed	182	164	149	136	141	120	102	100
Measured 2 m above seabed	182	165	147	133	141	120	101	100
Model V0	181	162	137	118	137	109	76	64
Model V1a	181	164	146	128	136	110	76	60
Model V1b	182	167	150	133	137	111	79	66
Model V2, ϕ=1.5	182	168	154	143	137	108	83	72
Model V2, ϕ=1.0	183	170	156	146	137	113	85	75

Table B.2 Difference between broadband SELss values calculated by the different model versions and measured 10 m above seabed.

	$\Delta SEL_{ss,unw}$ [dB]				$\Delta SEL_{ss,VHF}$ [dB]			
Location	MP1	MP2	MP3	MP4	MP1	MP2	MP3	MP4
Model V0	-1	-2	-12	-18	-4	-11	-26	-36
Model V1a	-1	0	-3	-8	-5	-10	-26	-40
Model V1b	0	3	1	-3	-4	-9	-23	-34
Model V2, ϕ=1.5	0	4	5	7	-4	-12	-19	-28
Model V2, ϕ=1.0	1	6	7	10	-4	-7	-17	-25

Observations, explanations and advise/discussion

- Not modelling dispersion (Model V0) results in a significant underestimation of SELss below ~250 Hz at MP2 and beyond. This was already observed in (de Jong, et al., 2019).
  - Dispersion is important for modelling the broadband SELss (where LF dominate) but may not be as essential when modelling the weighted SELss.
- The difference between the unweighted SELss at 1 m above the seabed (model V1a) and the max over depth (model V1b) is significant (up to 5 dB in the unweighted broadband SELss) at MP3 and MP4.
  - This is because at MP3 and MP4, less modes contribute resulting in a stronger depth pattern.
  - For this study, the max over depth is provided.
- The model results vary little at short range (MP1) and show good agreement with the measurements up to 2 kHz. Above 2 kHz there is 4 to 5 dB underestimation.
  - The low-frequency approximation in the hammer model, see (de Jong, et al., 2019), might explain this underestimation.
- A harder bottom Model V2 with (ϕ=1 compared to ϕ=1.5) leads to a 2 to 3 dB higher unweighted broadband SELss at MP2 and beyond.
  - This is caused by the increased sound speed of the seabed. At larger distances, smaller grazing angle modes dominate for which the sound speed of the seabed has a smaller effect.
- At larger ranges (MP3 and MP4) the models have a significant underestimation of SEL for higher frequencies relevant for porpoises.
- However, comparison of the model predictions with measurements at higher frequencies is complicated because the measurements at larger distance and higher frequencies are dominated by ambient noise, as visible by the flattening of the decidecade spectra in Figure B.4.
- Possible causes for the underestimation at high frequencies and large distances are:
  - The numerical model in the Aq4 model focusses the energy towards a ~16 degree grazing angle wrt to the seabed (so called Mach cone). Though this is expected to provide an accurate description for the bulk of the energy, at larger ranges sound radiating from the pile at smaller angles may start to dominant the energy in the Mach cone because sound propagates more effectively at smaller grazing angles. The amount of energy radiated by pile driving at smaller grazing angles is however not yet understood.
  - Another mechanism that could play a role is the possible redistribution of energy from higher grazing angles to smaller grazing angles because of interaction with the rough seabed and ocean surface. Including this effect in the model requires higher model complexity then currently supported.
- Aquarius 4 does not produce results for frequencies below 25 Hz, because Kraken does not find modes at these low frequencies. The energy at these frequencies however does not affect the broadband unweighted and weighted SELss.

B.2 Model sensitivity to geoacoustic properties

Reflections at the seabed are associated with a loss of acoustic energy, dependent on the grazing angle at which the sound waves interact with the seabed.

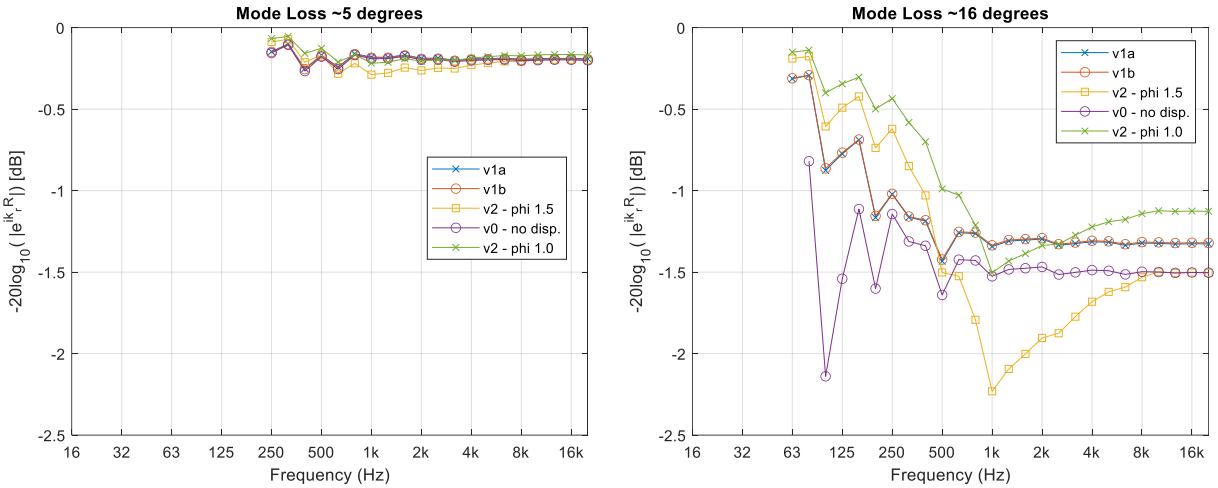

Figure B.5 Spectra of loss in dB per reflection at the seabed at two angles of incidence (left: 5°, right: 16°), for the different geo-acoustic parameters (Figure B.2).

These figures show that at 16 degrees, the loss is relatively large and very sensitive to the choice of the geoacoustic model. This means that the SELss at larger ranges from the pile, where propagating modes associated with smaller angles of incidence will dominate, is less sensitive to the choice of geo-acoustic parameters.

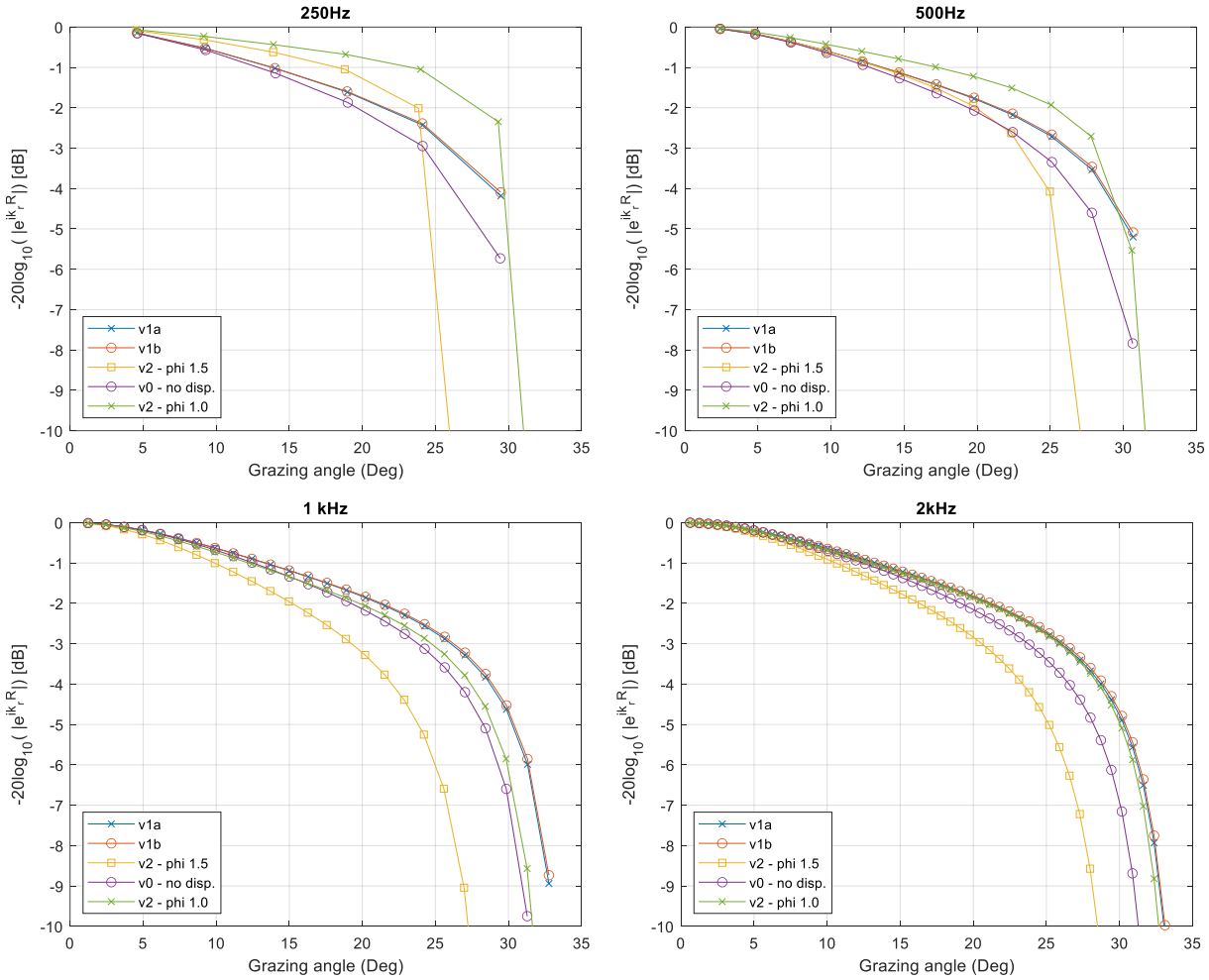

Figure B.6 Loss in dB per reflection at the seabed as a function of angles of incidence at four frequencies (250 Hz, 500 Hz, 1 kHz and 2 kHz) for the different geo-acoustic parameters (Figure B.2).

These figures show how the seabed model influences the reflection loss. At 1 kHz, the dispersive sound speed results in a significant increase in losses (attenuation is not just adjusted). At 500Hz this effect is much less pronounced, at 250Hz the attenuation becomes the dominant parameter influencing the loss mechanism.

C Statistical analysis – theory

C.1 Sensitivity, specificity, and 0/1 loss

- **Sensitivity**, or true positive probability, is the probability that the model predicts a 1 (=presence of PPM), conditional on the reality being also 1.
- **Specificity**, or true negative probability, is the probability that the model predicts a 0 (=absence of PPM), conditional on the reality being also 0.
- **0/1-loss** is a binary loss function, or the probability that the model makes a mistake.

C.2 Zero inflation

Zero inflation refers to there being too many zeros in the response variable of the model (in our case: too many zeros in the PPM). This is a problem. If a statistical model is given data where, say, 90% of the response values are zero, the model is inclined to (almost) always predict a zero, as that would necessarily give the high accuracy of about 90%. But this would result in more predicted zeros than would be actually be correct, and thus result in low sensitivity. It was attempted to fix this by using a complementary log-log link function (admittedly a difficult to interpret link function, but it works well for zero inflation), and by using the additional intercept term α.

C.3 Model fit and diagnostics

The fit for the presence/absence models were checked in 2 ways:

- The 0/1-loss was calculated between the observed presence/absence of fishing activity, and the predicted presence/absence of fishing activity. The lower the 0/1-loss, the better.
- Two complementary probabilities were calculated. One is the Sensitivity, which is the probability of correctly predicting a true 1 (presence), or P(prediction=1|observation=1). The other is the Specificity, which is the probability of correctly predicting a true 0 (absence), or P(prediction=0|observation=0). For both probabilities it holds that the higher (closer to 1), the better.

How well the non-zero models fit the data was checked in 2 ways:

- A scatter plot was produced with the observed response on the y-axis, and the predictions on the x-axis. A straight line was fitted through these points, and the slope of this line (the fit slope) was determined. The closer the fit slope is to 1, the better.
- The mean absolute deviation (“MAD”) was calculated between the observed response and the predictions. The closer MAD is to zero, the better.

Residual diagnostic plots were produced to check for violations of the model assumptions. Dunn-Smyth residuals (Dunn & Smyth, 1996) were used for all models. For the log-linked Gamma model, these residuals were also used, but the randomized jittering that is normally used for discrete distributions was also applied on the residuals of the Gamma model, since the Gamma model was fitted on integer values (the response values are not the result of actual count processes, but

rather measure the proportion of the time that events are happening). Thus, the residual value of each observation i for the Gamma model was defined as follows:

$$\text{residual}_i = F^{-1}_{\text{normal}}(u_i)$$

$$u_i \sim \text{Unif}(a_i, b_i)$$

$$a_i \sim F_{\text{gamma}}(y_i - 1|\mu_i, \phi) \quad \forall_{y>0}$$

$$b_i \sim F_{\text{gamma}}(y_i|\mu_i, \phi) \quad \forall_{y>0}$$

where F is a cumulative distribution function (the distribution being indicated in the subscript), and μ and θ being the mean and dispersion parameters of the gamma model in question, respectively.

C.4 Some math symbols briefly defined

A few math symbols that might be unknown to some readers are briefly defined here.

$$(x)_+ = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$(x)_- = \begin{cases} x & \text{if } x < 0 \\ 0 & \text{otherwise} \end{cases}$$

$$1_{\text{condition}} = \begin{cases} 1 & \text{if condition = TRUE} \\ 0 & \text{if condition = FALSE} \end{cases}$$

$$-1_{\text{condition}} = \begin{cases} -1 & \text{if condition = TRUE} \\ 0 & \text{if condition = FALSE} \end{cases}$$

$E(x|\text{condition})$ means "the expected value of x while `condition` is TRUE", or equivalently "the mean of x only when `condition` is TRUE".

C.5 Notes on R-INLA

For the Borssele analysis, the zero-truncated beta-binomial and zero-truncated negative binomial models were run using the `INLA` (Bakka, et al., 2018) (Lindgren & Rue, 2015) (Rue, Martino, & Chopin, 2009) R-package. R-INLA does not natively support zero-truncated model distributions, so for those models the zero-altered beta-binomial and zero-altered negative binomial distributions were used, with the probability of zero set to $\exp(-20)$.

We have also found some problems using R-INLA's "regular" offset methods, so any offset term was instead introduced as a fixed effect with a coefficient whose prior is defined as Norm($\mu = 1$, $\sigma^2 = \frac{1}{40000}$), which forces the coefficient to be estimated close to 1.

C.6 Other noteworthy R packages that were used

All R-functions we wrote ourselves were checked for scoping with the help of the `codetools` R-package (Tierney, 2018). This report was written with the Bookdown (Xie, 2016) extension of R-Markdown (Allaire, et al., 2019) of R-studio. Most of the figures shown in this report relied primarily on the `ggplot2` (Wickham, 2016)

R-package, accompanied by `viridis` (Garnier, 2018) (for plotting colour-blind friendly colours).

C.7 Non-zero models

C.7.1 Response and offset term

The proportion of the non-zero positive minutes can be expressed as follows:

$$y_{\text{prop}} = \frac{y_{y>0}}{60}$$

where $y_{y>0}$ are the non-zero positive minutes. This proportion is not a true probability. Although it is a proportion of integers, the integers are not true counts (they are a proportion of time, so no counting process is present). And its distribution is right skewed. For these reasons the gamma distribution was chosen for the model.

The response was re-defined to be a strictly positive ratio instead of a proportion, as follows:

$$y_{\text{ratio}} = \frac{y_{y>0}}{60 - y_{y>0}}$$

As a Gamma model with log-link function is used, one can formulate the relation between the linear predictors (the linear combination of all fixed and random effects), η, and the response as follows:

$$\log\left(\frac{E(y_{y>0})}{60 - y_{y>0}}\right) = \eta$$

This can then be re-formulated as:

$$\log\left(E(y_{y>0})\right) = \eta + \log(60 - y_{y>0})$$

Thus $\log(60 - y_{y>0})$ is the offset term.
Fitted values for use in the residuals and diagnostic plots are computed by making predictions while including the offset term. Predicting the ratio $\frac{y_{y>0}}{60-y_{y>0}}$ is computed by leaving out the offset term. Predicting the actual positive minutes, is done as follows:

$$E(y|y>0) = \frac{E(y_{\text{ratio}})}{1 + E(y_{\text{ratio}})} \times 60$$

where $E(y_{\text{ratio}})$) are the predictions from the gamma model when leaving out the offset term (i.e. replacing the offset term with 0).

C.7.2 Model formulations

The model formulations were very similar to that of the Bernoulli models.
for the regular SPL analysis:

$$\log\left(E\left(\frac{y}{60 - y}|y>0\right)\right) = \beta_{\text{Intercept}} + s(\text{SPL}) + \text{smoothers} + \text{RE}_{\text{(C-Squares)}}$$

for the SPL with mdtps analysis:

$$\log\left(E\left(\frac{y}{60-y}|y>0\right)\right) = \beta_{\text{Intercept}} + s(\text{SPL}) + \text{smoothers} + f_{\text{mdtps}}(\text{mdtps}) + \text{RE}_{\text{(C-Squares)}}$$

for the split SPL analysis:

$$\log\left(E\left(\frac{y}{60-y}|y>0\right)\right) = \beta_{\text{(split intercept)}} + s(\text{SPL, by=pilingactive}) + \text{smoothers} + \text{RE}_{\text{(C-Squares)}}$$

where “pilingactive” is a categorical covariate indicating whether there is piling activity present in the current time point or not.

In the above 3 model formulations, “SPL” refers to either unweighted or VHF weighted SPL (both types were modelled).

C.7.3 Results

Due to the extreme zero-inflation, there is little point to showing the results of the Gamma models in this report.

D Porpoise behavioural response – Borssele piling

The goal of this study is to investigate which acoustic metric (unweighted or weighted for the hearing sensitivity) provides the best prediction of behavioural response of marine mammals to piling sounds. Both metrics decrease with increasing distance from the pile. The rate of this decrease depends on environmental properties such as bathymetry and seabed properties. In this study the behaviour of harbour porpoises is expressed in terms of porpoise positive minutes (PPM), which is the number of minutes in each hour in which at least one porpoise click is detected.

Harbour porpoise acoustic activity was measured with CPODs at 16 locations in the study area (see Section 3.2). Details on identifying harbour porpoise clicks in the CPOD data are provided by (Brinkkemper, et al., 2021).

This chapter focusses on the porpoise behavioural response to piling sound. The analysis is based on the hourly median values of the modelled single strike sound exposure level (SEL_{SS}), expressed in dB re 1 μPa^2s due to the various piling events (see §3.6). SEL_{SS} can either be frequency-weighted, for the hearing of harbour porpoises ($SEL_{SS,VHF}$), or unweighted (SEL_{SS}). The distance from each CPOD to the piling locations is also known, and can also be used in the modelling.

The research questions are as follows:
- Is there a relationship between PPM and frequency-weighted $SEL_{SS,VHF}$?
- Is there a relationship between PPM and unweighted SEL_{SS}?
- Is there a relationship between PPM and the distance to the piling source?
- Which of these parameters predicts PPM best?

This chapter describes the statistical analyses performed to answer these questions. R version 3.6.3 (R Core Team, 2020) and R studio version 1.3.959 (RStudio Team, 2020) were used for the statistical analyses. The mgcv (Wood, 2017) R-package was used for fitting the Generalized Additive Models (Wood, 2011) with low-rank thin plate regression splines (Wood, 2003).

D.1 Materials-data

The detection of porpoise clicks was recorded in the period from 15 October 2019 to 31 October 2020 using 16 CPODs. The data are provided in terms of porpoise positive minutes per hour (PPM).

The calibrated acoustic model (see Section A.1) was applied to calculate the hourly average single-strike unweighted sound exposure level at the 16 CPOD locations, per piling hammer strike for every hour of the piling period:
- Unweighted broadband (50 Hz – 500 Hz) SEL_{SS} in dB re 1 μPa^2s.
- VHF-weighted broadband (50 Hz – 500 Hz) $SEL_{SS,VHF}$ in dB re 1 μPa^2s.

The following environmental datasets were provided:

- tides: contains tidal flow magnitude (m/s), tidal flow direction (°), and tidal height (m).
- wind: contains wind speed (m/s) and wind direction (°).
- wave: contains wave period (seconds per hour), wave direction (°), and significant wave height (m).
- temperature: contains air and water temperatures (°C).

The modelling data was constructed from the provided data. The constructed modelling data consisted of the variables listed below. Some of these environmental variables were highly correlated, consequently some of these environmental variables were not used in the models:

- air and water temperature were highly correlated; air temperature was therefore excluded.
- significant wave height was correlated with wave period and wind speed, and was therefore excluded.

Tidal flow magnitude and tidal height were measured at different locations, and not correlated. Since tidal flow magnitude was measured at a location closest to the study location, it was considered for the models, whereas tidal height was not used.

Environmental variables were unavailable for 15% to 25% of the observations; missing environmental variables were imputed instead. Imputation was done by first modelling the environmental variables using a generalized additive model (GAM) (Wood, 2011) with month of the year, day of the month, and hour of the day as covariates entered as low-rank thin-plate regression splines (Wood, 2003), and year as a categorical covariate. The missing values were replaced with the predictions from this model. The mgcv (Wood, 2017) R-package was used for this.

The conditional distribution chosen for the models differed per environmental variable:

- For water temperature, a gamma GAM with log link-function was used.
- For wind speed, a Negative Binomial GAM with log link-function was used.
- For tidal flow magnitude a beta GAM with logit link-function was used.

D.2 Methods-models

We modelled the following three relationships separately:

- The relation between the presence and number of porpoise positive minutes per hour, and the distance to piling. This will be referred to as the “distance-based analysis”.
- The relation between the presence and number of porpoise positive minutes per hour, and sound produced by piling, without weighting the sound for the hearing of porpoises. This will be referred to as the “unweighted SEL_{SS} based analysis”.
- The relation between the presence and number of porpoise positive minutes per hour, and sound produced by piling, with the sound being frequency-weighted for the hearing of porpoises. This will be referred to as the “$SEL_{SS,VHF}$ based analysis”.

Multiple modelling types were considered for modelling this data properly. Those considerations are detailed in Annex C. In this section, only the final choice is explained.

Since the CPOD data showed extreme zero-inflation (>90% of the response values consisted of zeros), for each of the three main relationships of interest, a hurdle model set was used consisting of two components:

- a Bernoulli model with complementary log-log link function, modelling the presence (PPM>0) or absence (PPM=0) of the porpoise positive minutes.
- a Gamma model with log-link function and with offset term, modelling the non-zero porpoise positive minutes (PPM>0), proportional to 60 (because one hour has 60 minutes).

The main effect in the models was either distance to the piling, the VHF-weighted $SEL_{SS,VHF}$, or the unweighted SEL_{SS}. These variables were only relevant during piling events, yet *some* number needs to be filled in the absence of piling events. The SEL variables were set to be 0 when no piling event was taking place (note that the lowest observed SEL was greater than 0). Then they were entered in the model as a (custom) spline, which is defined as follows. First, let x be the SEL effect in question, and let k be the lowest value x can have during a piling event (in the absence of piling, $x = 0$); so for the VHF-weighted $SEL_{SS,VHF}$ $k = 26$, and for the unweighted SEL_{SS} $k = 96$. And let β be the vector of coefficients. One can then define the spline with the following function:

$$f_{\text{SEL}}(x) = \beta_1 \times x + \beta_2 \times (x-k)_+ + \beta_3 \times ((x-k)_+)^2 + \beta_4 \times ((x-k)_+)^3$$

For numerical purposes, the terms of this function were scaled.

The mean distance to piling was set to be -1 when no piling event was taking place. Then they were entered in the model as a (custom) spline, which is defined as follows. Let x be the mean distance to piling sound (“mdtps”) effect, and let β be the vector of coefficients. One can then define the spline with the following function:

$$f_{\text{mdtps}}(x) = \beta_1 \times (x)_- + \beta_2 \times (x)_+ + \beta_3 \times ((x)_+)^2 + \beta_4 \times ((x)_+)^3$$

This function is equivalent to:

$$f_{\text{mdtps}}(x) = \beta_1 \times -1_{\text{piling inactive}} + \beta_2 \times (x)_+ + \beta_3 \times ((x)_+)^2 + \beta_4 \times ((x)_+)^3$$

Besides the main effect, the following covariates were entered in the model as low-rank thin-plate smoothers:

- month of the year (1 to 12), because seasonal changes may also affect porpoise behaviour;
- hour of the day, because the daily rhythms and day-night cycle may affect porpoise behaviour;
- water temperature, because that is an important aspect of seasonal change;
- tidal flow magnitude, because tidal flow, and the related moon cycles, may affect porpoise behaviour.

Wind speed was also an important covariate, as it affects both the ability to detect porpoise, and there also is less piling activity at high wind speeds due to weather restrictions. Data plots indicated that the relation between PPM and wind speed

differs between seasons. Therefore, wind speed was entered in the models as a covariate with a different low-rank thin-plate smoother for each season.
A Bernoulli generalized additive mixed model (GAMM) with a complementary log-log link function was used to model the presence and absence of porpoise positive minutes. Unfortunately, the model fitting process tries to get a fit slope (between the observed and predicted responses) as close to 1 as possible, but we are more interested in getting a good specificity and sensitivity. This was very difficult, due to the extreme zero-inflation. The specificity was found to be very high (nearly 1), but the sensitivity was very low. By adding a small positive number to the linear predictor (i.e. a second intercept), one can sacrifice a little bit of specificity to increase the sensitivity. This second intercept is referred to as α. For each Bernoulli model, values between 0 and 2, with increments of 0.025, were tried out for α.
In all Bernoulli models the α value of 1.8 was found to give the best balance between sensitivity, specificity, and 0/1-loss.

Let the term “smoothers” represent the following:

$$\text{smoothers} = s(\text{month}) + s(\text{hour}) + s(\text{tidal flow magnitude}) + s(\text{wind speed, by = season})$$

with s being the low-rank thin-plate smoothers with mgcv’s default parameters (see (Wood, 2003; Wood, 2011; Wood, 2017). And let $\text{RE}_{\text{(C-Squares)}}$ the random intercept for the 0.01 degrees C-Squares grid cells.

The model formulation for the Bernoulli models were as follows:
for the VHF-weighted $\text{SEL}_{\text{SS,VHF}}$:

$$\text{cloglog}(P(y > 0)) = \beta_{\text{Intercept}} + \alpha + f_{\text{SEL}}(\text{SEL}_{SS,\text{VHF}}) + \text{smoothers} + \text{RE}_{\text{(C-Squares)}}$$

for the unweighted SEL_{SS}:

$$\text{cloglog}(P(y > 0)) = \beta_{\text{Intercept}} + \alpha + f_{\text{SEL}}(\text{SEL}_{SS}) + \text{smoothers} + \text{RE}_{\text{(C-Squares)}}$$

for the distance based model:

$$\text{cloglog}(P(y > 0)) = \beta_{\text{Intercept}} + \alpha + f_{\text{mdtps}}(\text{mdtps}) + \text{smoothers} + \text{RE}_{\text{(C-Squares)}}$$

The model formulation for the non-zero models (not used) were as follows:
for the VHF-weighted $\text{SEL}_{\text{SS,VHF}}$:

$$\log(E(\frac{y}{60 - y}|y > 0)) = \beta_{\text{Intercept}} + f_{\text{SEL}}(\text{SEL}_{SS,\text{VHF}}) + \text{smoothers} + \text{RE}_{\text{(C-Squares)}}$$

for the unweighted SEL_{SS}:

$$\log(E(\frac{y}{60 - y}|y > 0)) = \beta_{\text{Intercept}} + f_{\text{SEL}}(\text{SEL}_{SS}) + \text{smoothers} + \text{RE}_{\text{(C-Squares)}}$$

for the distance based model:

$$\log(E(\frac{y}{60 - y}|y > 0)) = \beta_{\text{Intercept}} + f_{\text{mdtps}}(\text{mdtps}) + \text{smoothers} + \text{RE}_{\text{(C-Squares)}}$$

The Akaike information criterion (AIC) (Sakamoto, Ishiguro, & Kitagawa, 1986) and the Bayesian information criterion (BIC) (Schwarz, 1978), the model fit, and the confidence interval width of the main effects were used to compare the three models.

D.3 Prelude to the results

Due to the complexity of the models, results are shown in terms of predictor effect plots. The following sections show results of the Bernoulli models. The model diagnostic plots are available at WMR upon request.
The predictor effect plots for any given covariate effect $f(x)$ are made by varying x, and keeping all other covariates fixed at reasonable values. The shape of the graphed relation - and therefore the **relative** effect of x on the predicted value of P(PPM>0) - will not change if other fixed values are chosen, except that the graphed relationship will become flatter as the predicted values comes close to 0 or 1, as probabilities are necessarily bounded between 0 and 1. This has (of course) been considered also when choosing values to fix the other covariates at.

D.4 Results - VHF weighted Bernoulli model

The main effect of the VHF weighted $SEL_{SS,VHF}$ based Bernoulli model is shown in Figure D.1. The probability of PPM decreases significantly above a piling $SEL_{SS,VHF}$ of about 55 dB re 1 μPa^2s.

The effects of the other covariates are shown in Figure D.2 to Figure D.4:

- There is a seasonal pattern in the probability of PPM, with an increase from December-January to November.
- The daily pattern shows a night-day cycle, with lower probability of PPM during the day and higher probability of PPM during the night.
- There is a higher probability of PPM at lower water temperatures.
- Generally speaking, the higher the tidal flow magnitude, the lower the probability of PPM, though the effect is not very strong.
- Irrespective of the season, the higher the wind speed, the lower the probability of PPM (Figure D.3).
- The probability of detecting a PPM is a bit higher to the South than in the North (Figure D.4).

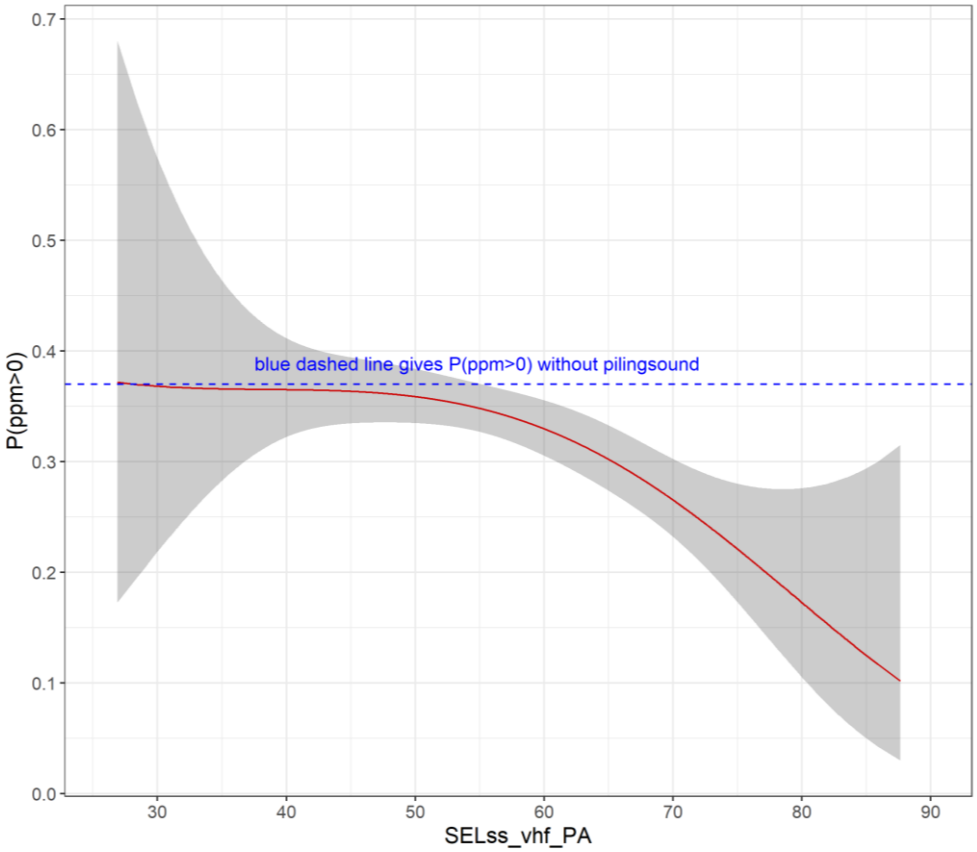

Figure D.1 Predictor effect plot of the main effect of the SELSS,VHF based Bernoulli model. The plots show the predicted probability of a porpoise positive minute (PPM), while ignoring all random effects/smoothers presented in Figure D.2, Figure D.3 and Figure D.4 (which average out to 0). The shaded ribbon gives the 95% confidence interval.

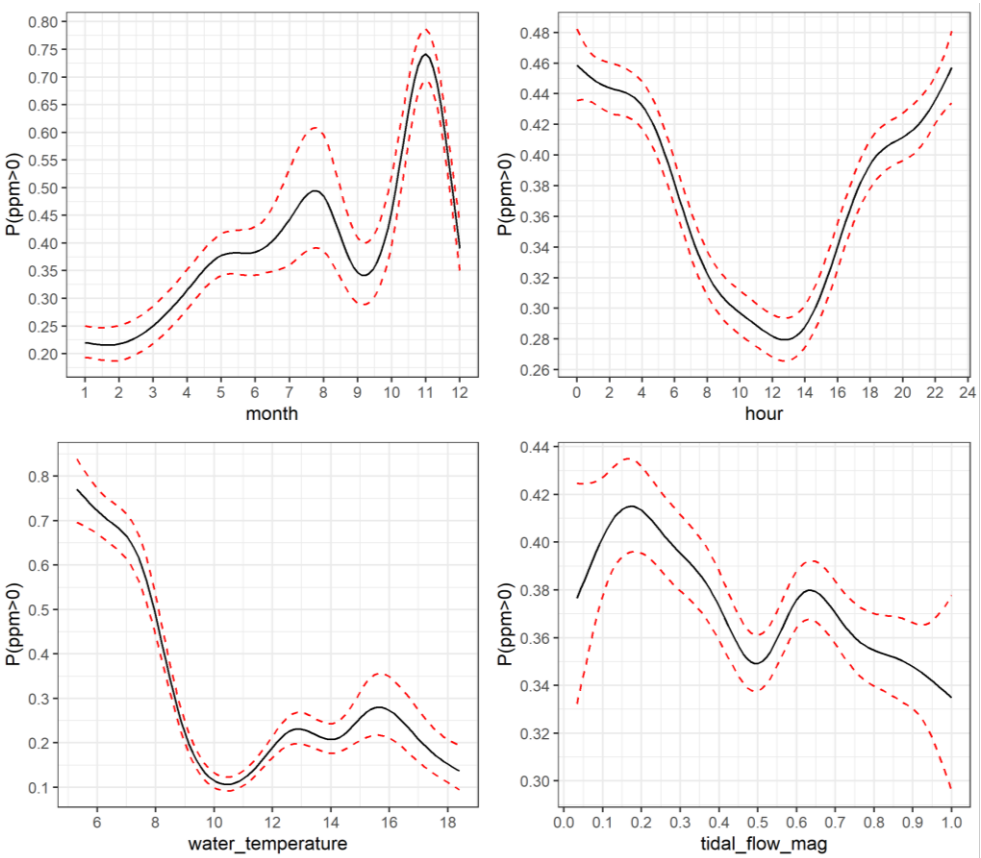

Figure D.2 Predictor effect plot of the regular smoothers for the SELSS,VHF based Bernoulli model. The plots show the predicted probability of a porpoise positive minute (PPM), while SEL to zero, and ignoring the other random effects/smoothers (which average out to 0). The red dashed lines indicate the 95% confidence interval.

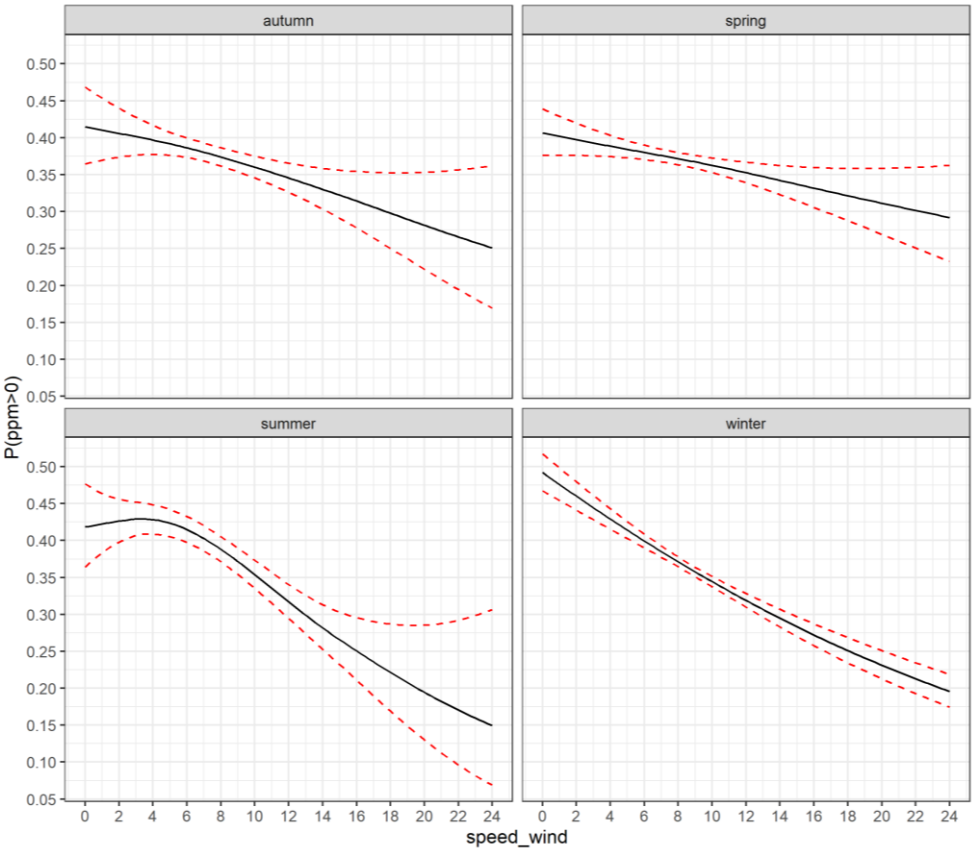

Figure D.3 Predictor effect plot of the interaction smoothers for the SELSS,VHF based Bernoulli model. The plots show the predicted probability of a porpoise positive minute (PPM), while setting SEL to zero, and ignoring the other random effects/smoothers (which average out to 0). The red dashed lines indicate the 95% confidence interval.

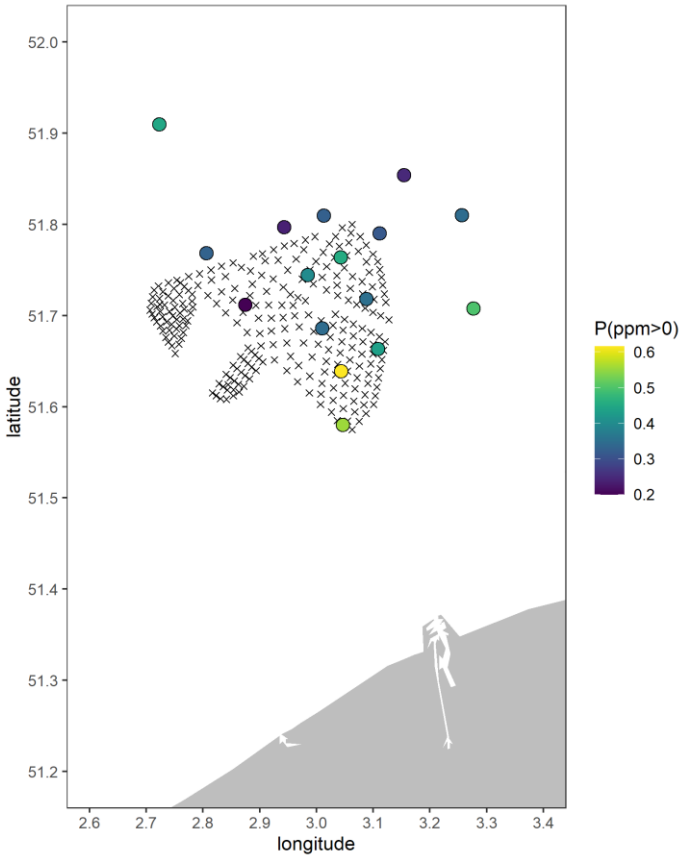

Figure D.4 Predictor effect plot of the random intercept for location for the SELSS,VHF based Bernoulli model. The plots show the predicted probability of a porpoise positive minute (PPM), while setting SEL to zero, and ignoring the other random effects/smoothers (which average out to 0).

D.5 Results - unweighted Bernoulli model

The main effect of the unweighted SEL based Bernoulli model is shown in Figure D.5. The probability of PPM decreases significantly from piling unweighted SEL_{SS} of about 130 dB re 1 µPa^{2}s. The effects of the other covariates are shown in Figure D.6 to Figure D.8. They show the same patterns as the effects of VHF weighted $SEL_{SS,VHF}$ based Bernoulli model (Figure D.2 to Figure D.4).

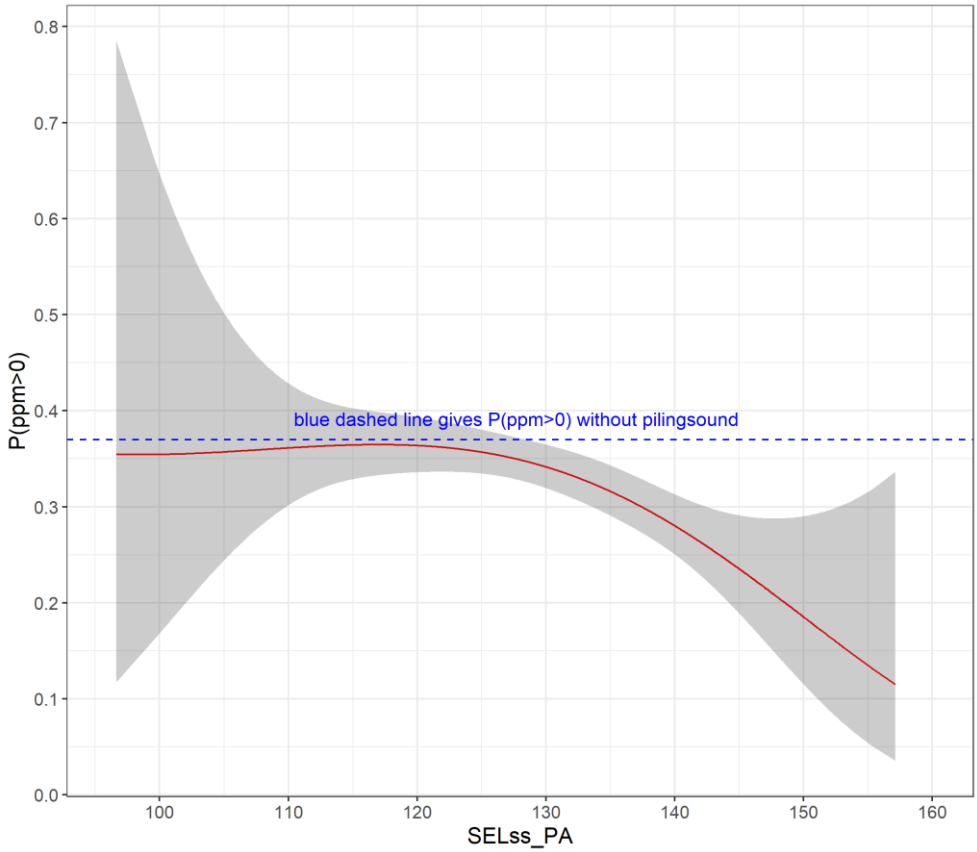

Figure D.5 Predictor effect plot of the main effect of the unweighted SELSS based Bernoulli model. The plots show the predicted probability of a porpoise positive minute (PPM), while ignoring all random effects/smoothers presented in Figure D.6 to Figure D.8 (which average out to 0). The shaded ribbon gives the 95% confidence interval.

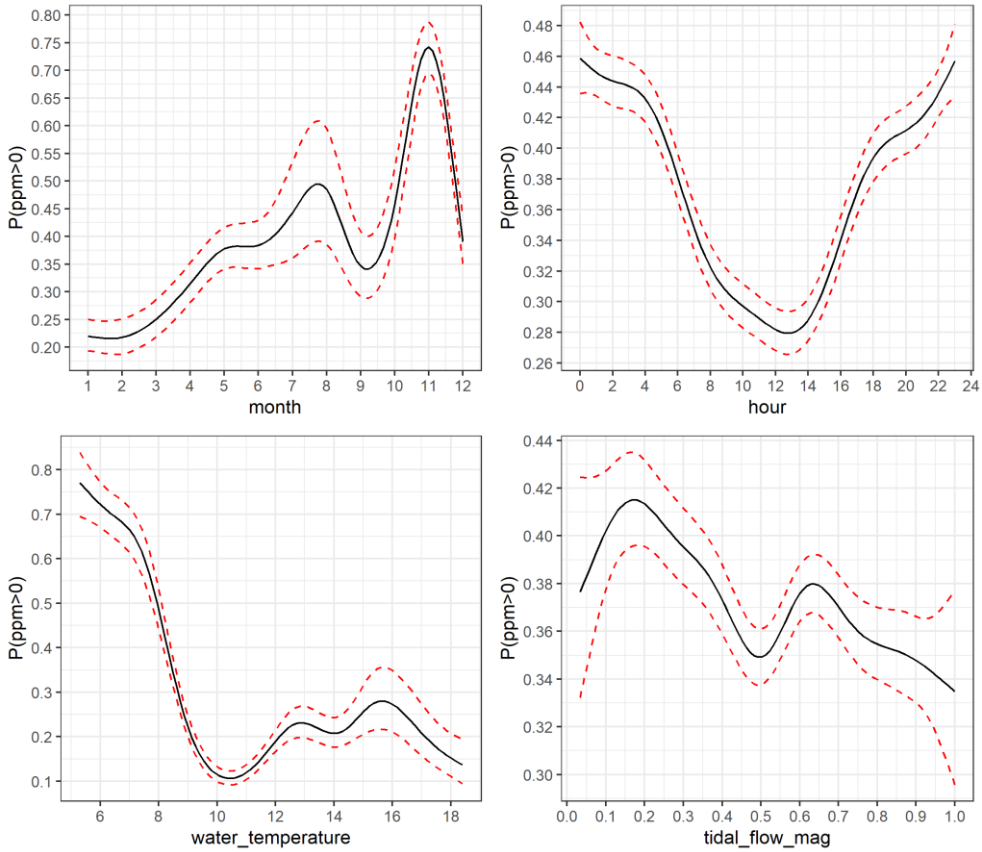

Figure D.6 Predictor effect plot of the regular smoothers for the unweighted SELSS based Bernoulli model. The plots show the predicted probability of a porpoise positive minute (PPM), while setting SEL to zero, and ignoring the other random effects/smoothers (which average out to 0). The red dashed lines indicate the 95% confidence interval.

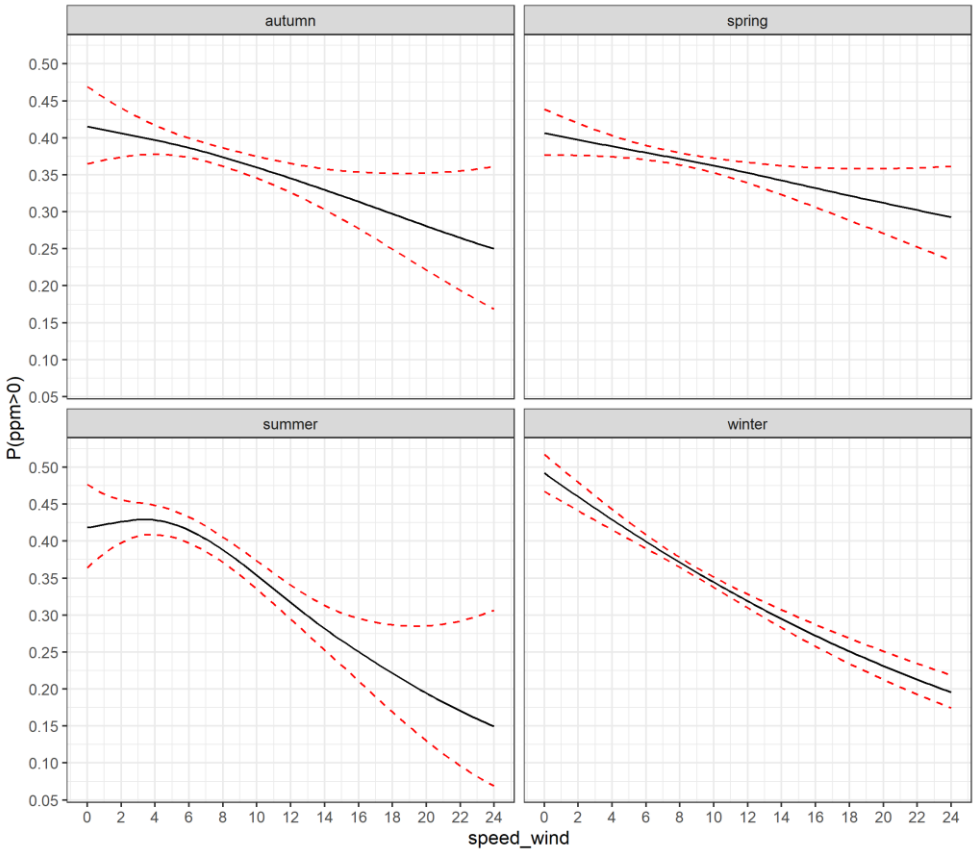

Figure D.7 Predictor effect plot of the interaction smoothers for the unweighted SELSS based Bernoulli model. The plots show the predicted probability of a porpoise positive minute (PPM), while setting SEL to zero, and ignoring the other random effects/smoothers (which average out to 0). The red dashed lines indicate the 95% confidence interval.

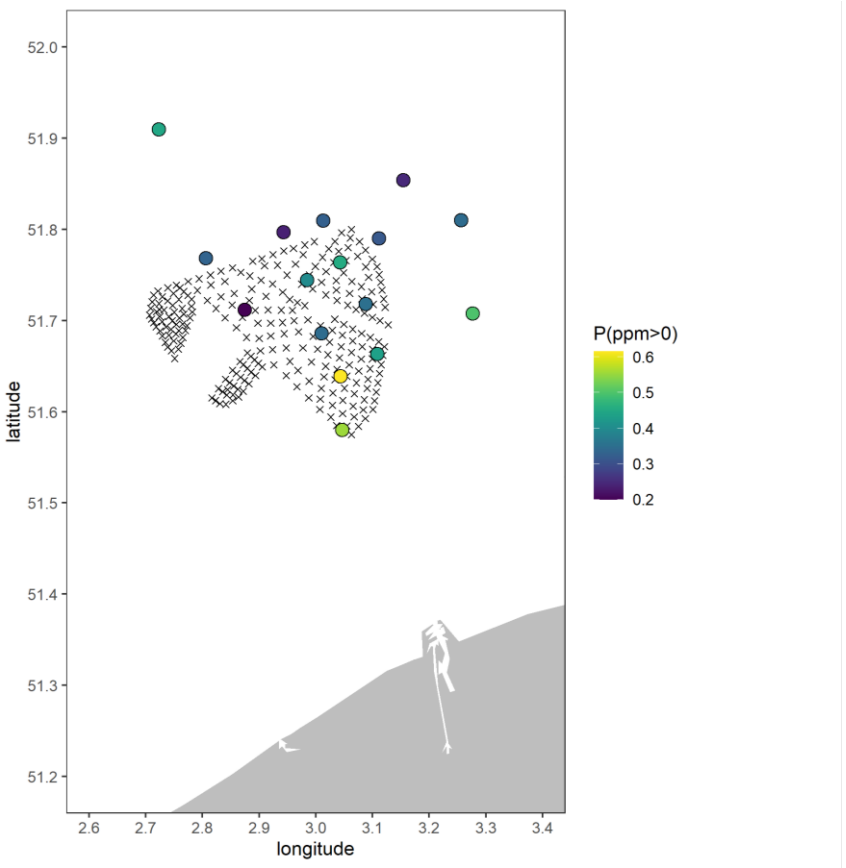

Figure D.8 Predictor effect plot of the random intercept for location for the unweighted SELSS based Bernoulli model. The plots show the predicted probability of a porpoise positive minute (PPM), while setting SEL to zero, and ignoring the other random effects/smoothers (which average out to 0).

D.6 Results – distance-based Bernoulli model

The main effect of the distance-based Bernoulli model is shown in Figure D.9. The probability of PPM increases up to a distance of 15 km from piling, although the statistical significance of the distance is already lost at about 7 km. Further away than 15 km piling sound does not affect the probability of PPM.
The effects of the other covariates are shown in Figure D.10 to Figure D.12. They show the same patterns as in both SEL based Bernoulli models (Figure D.2 to Figure D.4 and Figure D.6 to Figure D.8).

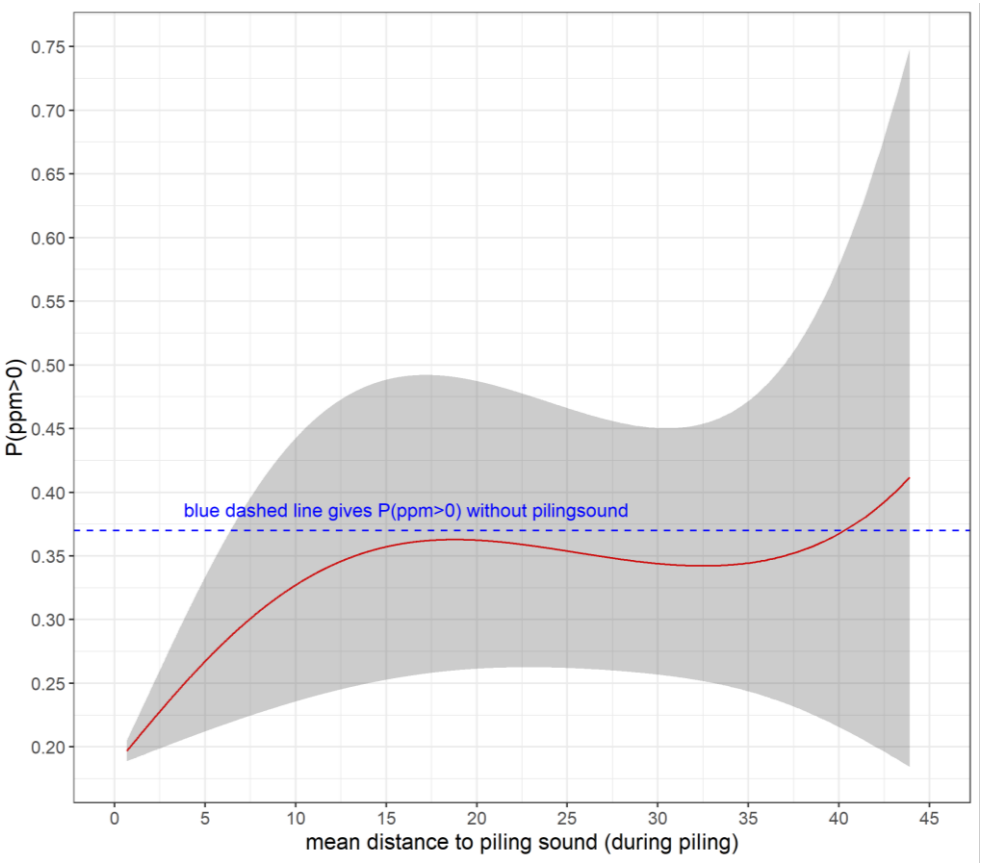

Figure D.9 Predictor effect plot of the main effect of the distance based Bernoulli model. The plots show the predicted probability of a porpoise positive minute (PPM), while ignoring all random effects/smoothers presented in Figure D.10 to Figure D.12 (which average out to 0). The shaded ribbon gives the 95% confidence interval.

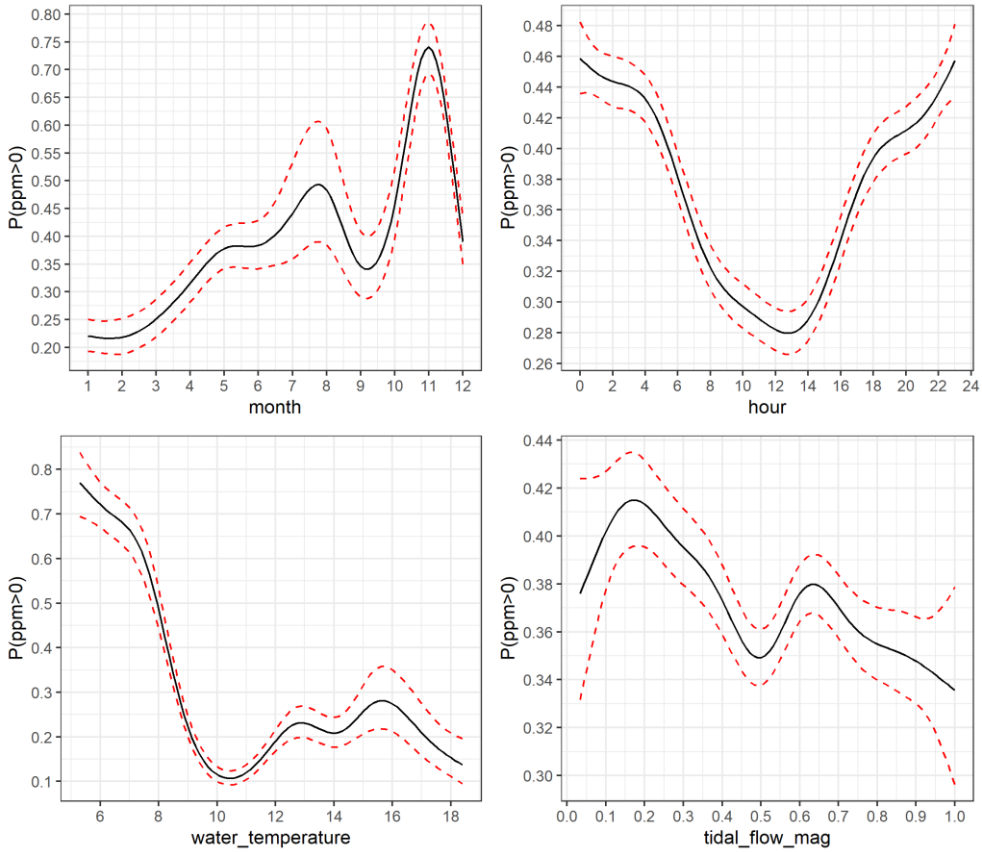

Figure D.10 Predictor effect plot of the regular smoothers for the distance based Bernoulli model. The plots show the predicted probability of a porpoise positive minute (PPM), while setting mdtps to -1 (meaning no piling), and ignoring the other random effects/smoothers (which average out to 0). The red dashed lines indicate the 95% confidence interval.

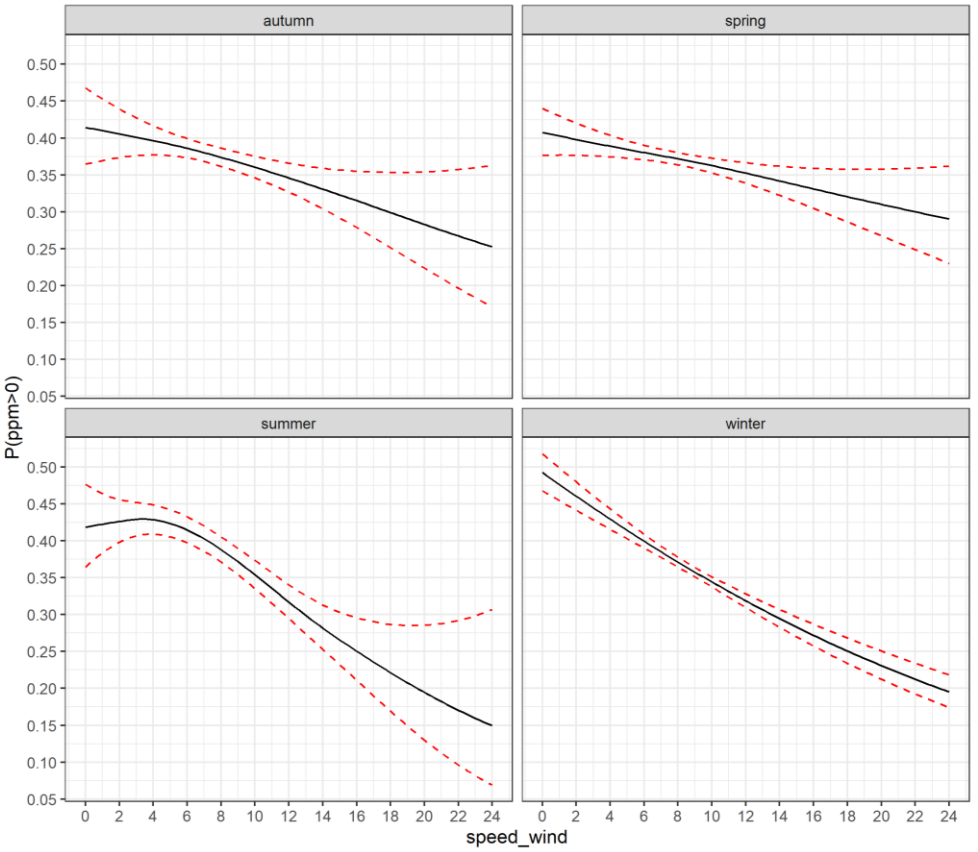

Figure D.11 Predictor effect plot of the interaction smoothers for the unweighted SEL based Bernoulli model. The plots show the predicted probability of a porpoise positive minute (PPM), while setting mdtps to -1 (meaning no piling), and ignoring the other random effects/smoothers (which average out to 0). The red dashed lines indicate the 95% confidence interval.

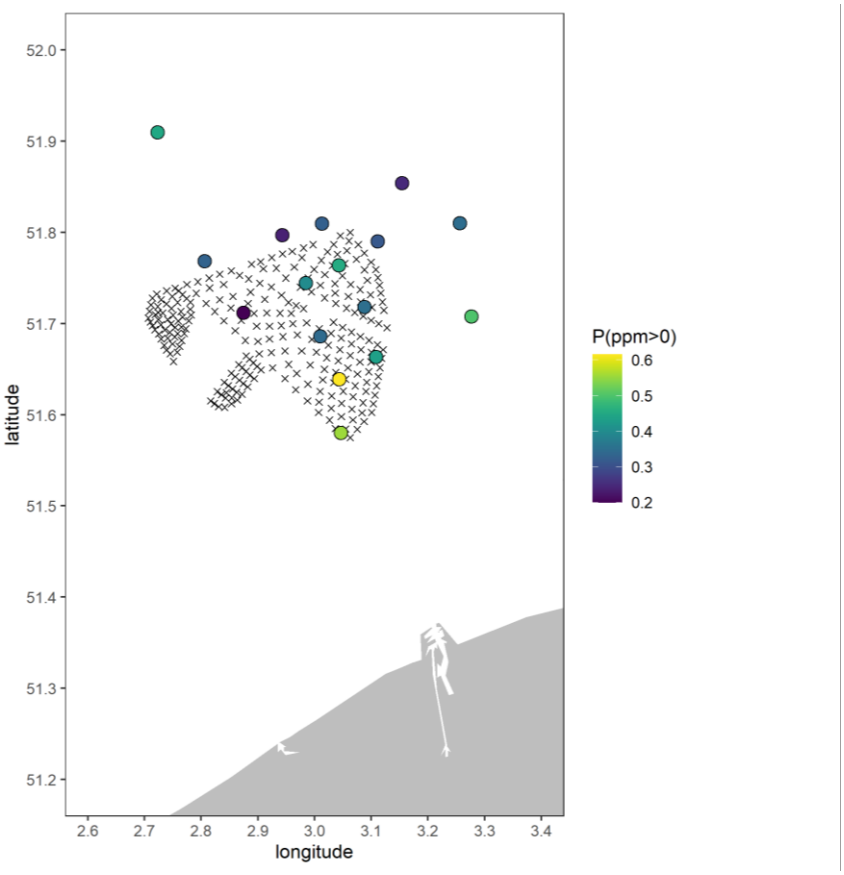

Figure D.12 Predictor effect plot of the random intercept for location for the unweighted SEL based Bernoulli model. The plots show the predicted probability of a porpoise positive minute (PPM), while setting mdtps to -1 (meaning no piling), and ignoring the other random effects/smoothers (which average out to 0).

D.7 Results - Model comparison

Three measures were used for model comparisons:

1 the fit of the models (expressed in specificity, sensitivity, and 0/1-loss; see Table D.1),
2 the information criterion of the models (expressed in AIC and BIC; see Table D.2), and
3 the main effect splines (see Figure D.13).

The VHF SEL based model has the lowest BIC, but the distance based model has the lowest AIC (Table D.2).

The AIC and BIC values of all the models are relatively close to each other. Moreover, the three Bernoulli models had very similar fit and model diagnostics (Table D.1). So, it appears that none of the three Bernoulli models performs better than the other. However, as shown in the comparison in Figure D.13, the main effect spline in the distance based Bernoulli model has a larger confidence interval than the main effect splines in the SEL based models, and thus less precision. Although the two SEL based Bernoulli models are slightly better than the distance based Bernoulli model, the differences between the models are small.

Table D.1 Sensitivity, specificity, and the 0/1 loss of each of the three Bernoulli models. The closer specificity and sensitivity are to 1, the better. The closer the 0/1-loss is to 0, the better.

Model	chosen alpha	Sensitivity	Specificity	0/1 loss
Vhf SEL	1.8	0.5304	0.7372	0.2811
Unweighted SEL	1.8	0.5303	0.7372	0.2811
Distance	1.8	0.5303	0.7371	0.2812

Table D.2 AIC and BIC of each of the three Bernoulli models. The lower AIC and BIC, the better.

Model	AIC	BIC
Vhf SEL	72,435.01	73,038.45
Unweighted SEL	72,437.31	73,040.45
Distance	72,434.96	73,038.47

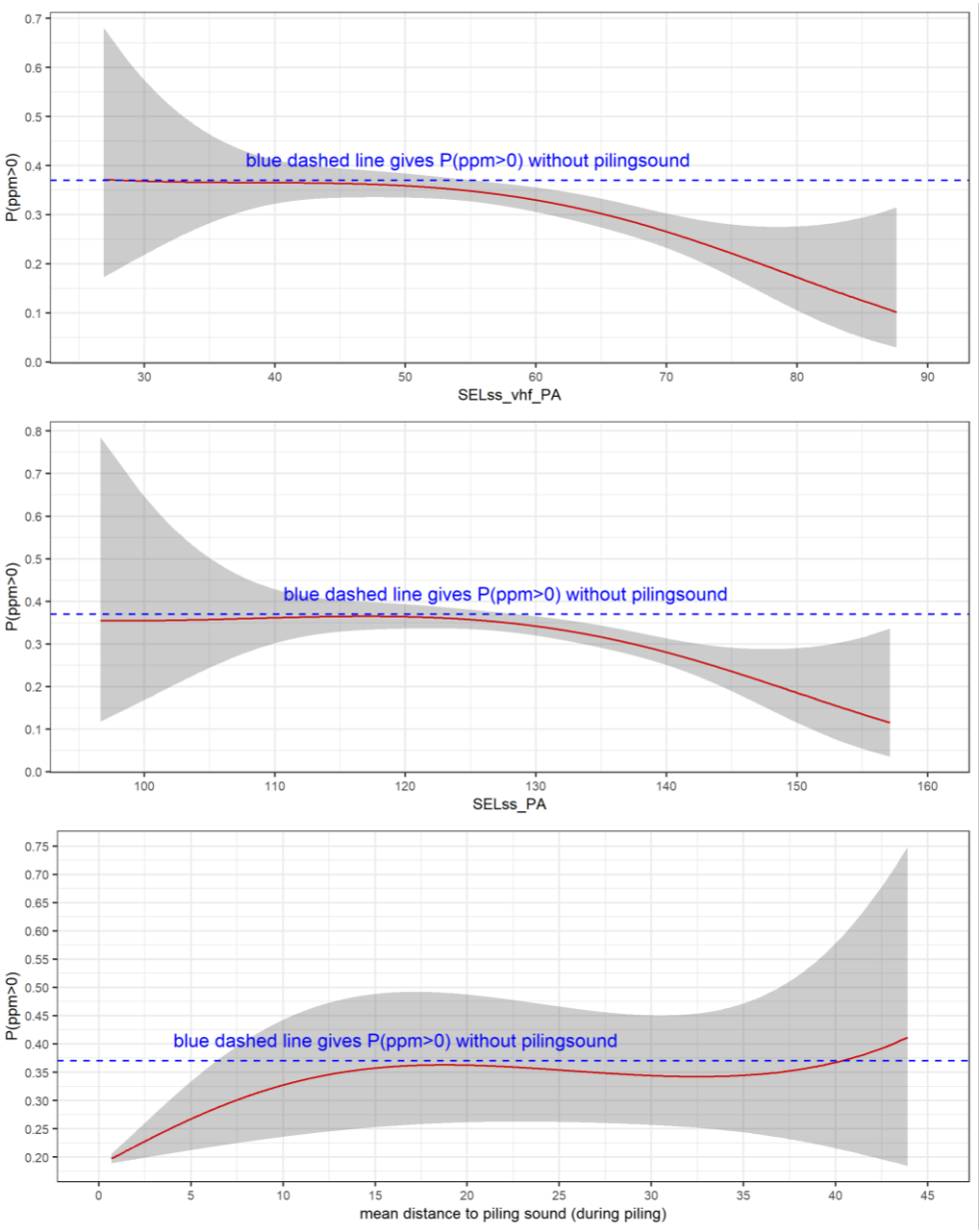

Figure D.13 Comparison of the main effects of the three Bernoulli models.

E Porpoise behavioural response – Gemini piling

E.1 Methods: Statistical models

Table E.1 gives an overview of the general model structure maintained through all Gemini statistical analyses. For the response variable, the results of the binomial model (Porpoise activity per hour) are given in the report. The results of the quasi-Poisson model (Porpoise positive minutes per hour) are provided in the appendix (E.5). Then, predictor variables are shown, which were modelled in separate models. Then, control variables are shown, which are consistent throughout all models. Day in the year is crossed out, as it could not be included as a control variable due to collinearity issues with water temperature. The column' shape' gives statistical details of how the variable was included in the model, the response variable, and which error structure the model used. The last column, 'comparison to Borssele analysis,' highlights where the Gemini model formulation differs from the Borssele analysis.

Variable type	*Category*	*Variable*	*Shape*	*Comparison to Borssele analysis*
Response	Porpoise presence	Porpoise positive minutes per hour Porpoise activity per hour (yes/no)	Quassi-Poisson Binomial	Non-zero not used Bernouilli (Very similar)
Predictor	Disturbance	Distance to disturbance	Smooth, bs = cr (cubic spline)	
		SELss or SELss, vhf	Smooth, bs = cr (cubic spline)	Borsele assumed values of 0 when no piling
Control	Weather	Tidal height	Smooth, bs = cr (cubic spline)	Tidal flow magnitude
		Wind speed	Smooth, bs = cr (cubic spline)	Wind speed, but by season
		Water temperature	Smooth, bs = cr (cubic spline)	
	Time	~~Day in the year~~	~~Smooth, bs = cc (cyclic spline)~~	Gemini had to exclude
		Hour	Smooth, bs = cc (cyclic spline)	Non-cyclic spline
	Location	CPOD	factorial	C-square, but that results in same effect as both are treated as factor

Table E.1 Overview model structure, variable names, shapes in model and comparison to model in Borssele analysis.

E.2 Methods: data

An overview of the data that was used for the statistical models in the Gemini analysis is shown in Table E.2. The detection of porpoise clicks was recorded in the period from 23 June 2015 to 17 February 2016 using 15 CPODs. Extra control (‘zero’) data was added from earlier recordings in the same area to increase the observation period outside piling activity from 9% to 96% of the total number of cumulative observational hours (n=168160). The table shows the variable names, followed by descriptive statistics, including mean, standard deviation, minimum and maximum observed values, and quantiles.

Table E.2 Overview of data set completeness and descriptive statistics of data used in the analyses.

Variable	Mean	Sd	Min	25th pct	Median	75th pct	Max
Hour	11.50	6.93	0.00	5.00	12.00	18.00	23.00
PPM/h	1.78	4.60	0.00	0.00	0.00	1.00	60.00
Distance to piling	16.21	13.44	0.04	6.27	11.72	21.79	55.39
SELss	102.11	11.02	100.00	100.00	100.00	100.00	191.40
SELss_vhf	80.71	4.85	80.00	80.00	80.00	80.00	148.56
Wind speed	61.57	24.66	11.00	42.00	57.00	77.00	156.00
Water temperature	12.41	4.72	0.90	8.90	12.60	16.90	19.80
Tidal height	-1.61	70.54	-142.90	-67.63	6.94	62.90	126.95
Day in the year	182.58	91.90	1.00	108.00	182.00	257.00	365.00
Porpoise Activity (1/0)	0.24	0.42	0.00	0.00	0.00	0.00	1.00

E.3 Results: Descriptive results

When considering the effect for the CPOD locations separately, there was a spatial effect of the harbour porpoises moving away from the piling sites. The location of the CPODs and piling events is shown in Figure E.1. The change in porpoise presence at the CPOD locations during piling compared to no piling is shown in Table E.3. For the CPODs closest to the piling locations, relative porpoise presence decreases, while at the CPOD locations just outside of the core location, relative porpoise presence increases. This testifies to the porpoises moving outward the core area when piling.

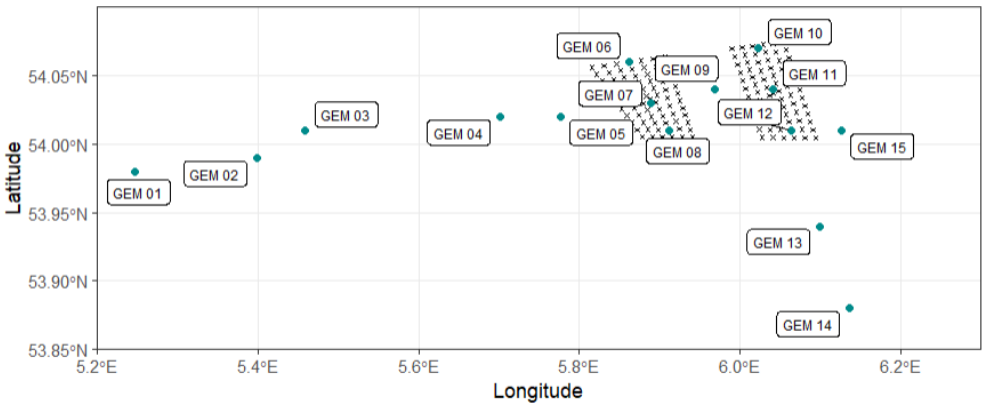

Figure E.1 The locations of the pile driving events and CPOD recorders in Gemini wind farm construction. Pile driving locations are shown in a grey 'x', CPOD locations are shown as blue points and their names in a tag.

Table E.3 Overview of mean porpoise presence values (p PPM/h>0) and activity values (PPM/h) per CPOD locations. The green shaded values highlight the CPOD location for which the presence and activity values increased during piling activity.

	Mean chance of porpoise presence			**Mean porpoise positive minutes per hour**		
CPOD	*No piling*	*Piling*	*Difference*	*No piling*	*piling*	*Difference*
1	0.179	0.156	0.023	1.109	0.810	0.299
2	0.166	0.222	-0.056	0.931	1.343	-0.413
3	0.214	0.232	-0.018	1.231	1.142	0.088
4	0.260	0.444	-0.184	1.890	3.536	-1.646
5	0.237	0.224	0.012	1.752	1.478	0.273
6	0.261	0.124	0.137	2.078	0.863	1.215
7	0.202	0.106	0.096	1.644	0.607	1.037
8	0.236	0.170	0.066	1.682	1.134	0.549
9	0.227	0.140	0.087	1.767	1.078	0.689
10	0.167	0.121	0.046	1.073	0.840	0.234
11	0.272	0.094	0.179	2.051	0.691	1.360
12	0.315	0.183	0.132	2.826	1.169	1.657
13	0.284	0.412	-0.128	2.210	3.945	-1.735
14	0.231	0.204	0.028	1.670	1.030	0.640
15	0.250	0.309	-0.059	2.514	2.632	-0.118

E.4 Results: plots control variables

For all models, having either porpoise presence (yes/no) or porpoise positive minutes as the response variable and either distance to piling, unweighted SEL & SPL, and weighted SEL & SPL showed very similar results for the control variables. Here, the results of the probability of porpoise presence in response to porpoise-weighted SELss,vhf values are presented.

The probability of porpoise presence slowly increased with temperature, with an optimum of around 9 degrees, then stabilizing. It was higher during the day and lowered during the night. It decreased somewhat with increasing tidal height. Lastly, it dropped with increasing wind speed. Explanations of the observed relations are given in the main text (5.3).

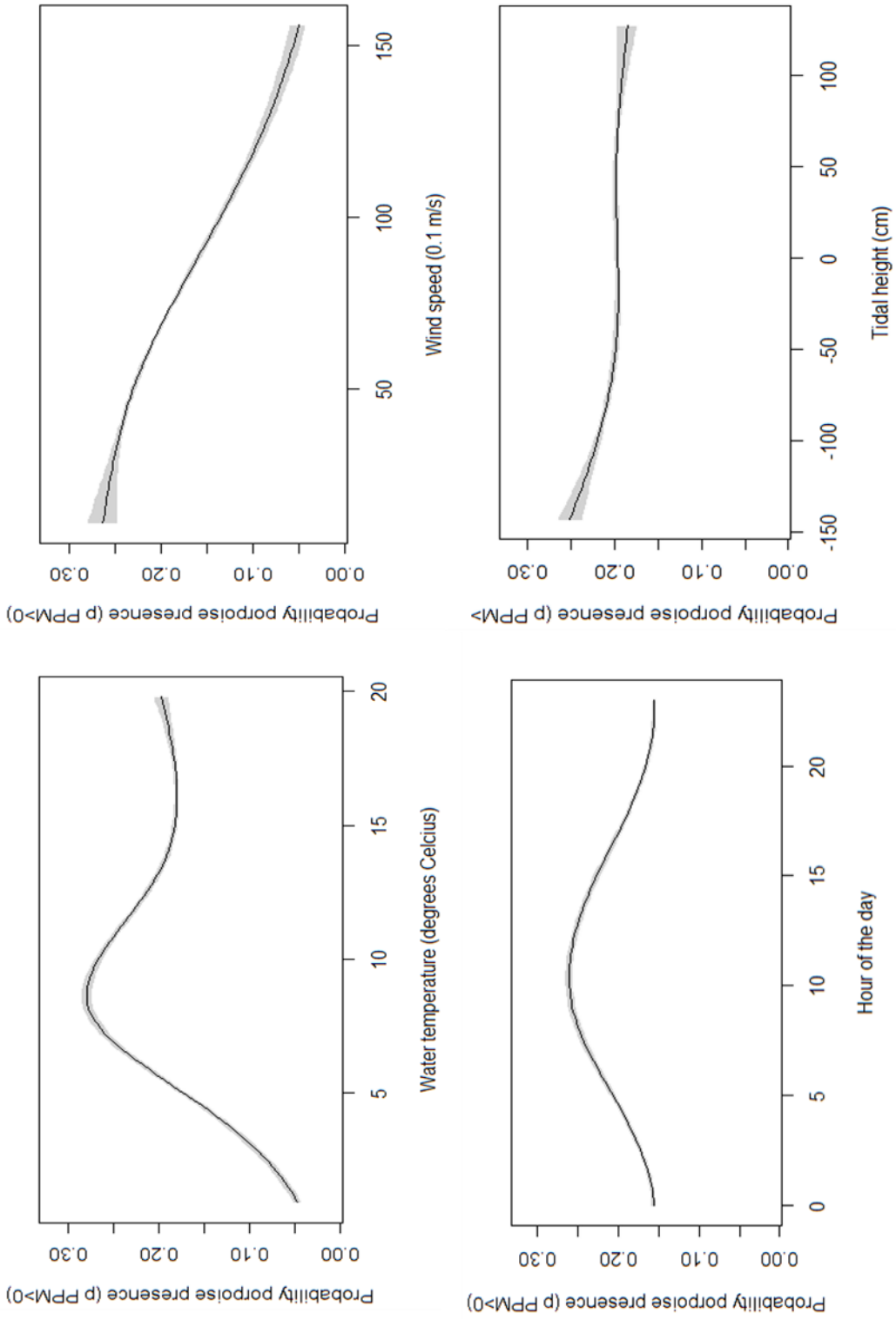

Figure E.2 Predictor effect plot of the random intercept for location for the unweighted SEL based Bernoulli model. The plots show the predicted probability of a porpoise positive minute (PPM).

E.5 Results: Non-zero porpoise click count analysis

This section describes the results of the quasi-Poisson analysis that studies the relation of the SEL variables of interest to the number of porpoise-positive minutes per hour. In contrast to the binomial analysis, where only presence (PPM/h>0) and absence (PPM/h = 0) are considered, we here get a finer understanding of the quantity of porpoise-positive minutes that is detected in an hour. It must be noted that there is currently no way to translate the number of positive porpoise minutes in an hour to the number of animals detected in an hour, as there is no way to know which clicks come from what animal. However, it tells us something about the general activity of present porpoises that hour.

The porpoise-positive minutes (PPM) relations observed here are similar to the binomial analysis of porpoise presence. Figure E.3 shows the modelled relations between PPM/h and the variables of our interest. For unweighted SELss, there is a small increase in PPM/h, after which it steadily decreases close to zero. The peak is around 130 dB but only drops below the reference level at 154 dB. This would mean that with disturbance, there first seems to be an increase in the number of clicks produced. For weighted $SEL_{ss,VHF}$ the relation is again very similar to the binomial analysis. Very early in the SEL scale, there is a linear-like reduction in the probability of porpoise presence. The effect of distance is also consistent with the binomial analysis, where there is an increase in PPM/h with increasing distance. PPM/h seems to be lower in general when piling activity occurs compared to mean levels when no piling occurs (blue dotted reference line).

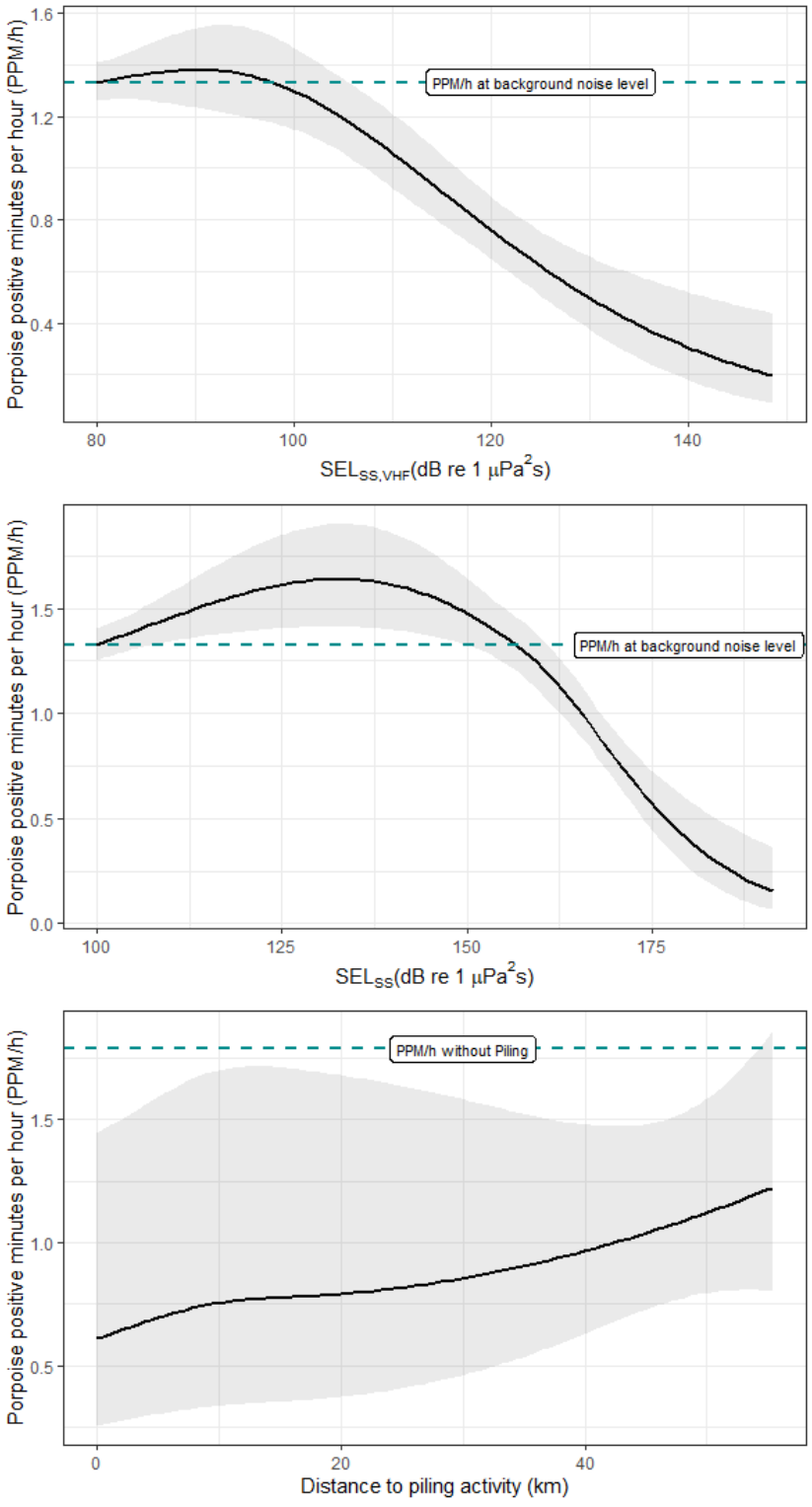

Figure E.3 Predictor effect plot of the random intercept for location for the unweighted SEL based Bernoulli model. The plots show the predicted probability of a porpoise positive minute (PPM).

F Porpoise behavioural response – SPL – Borssele

The analysis in this appendix concerns the relation between porpoise click detections and the underwater ambient sound that was measured by the seven SoundTrap recorders. This analysis is not limited to the effects of the piling sounds, but includes all geophonic, biophonic and anthropogenic contributions that the porpoises were potentially exposed to in the vicinity of these seven locations.

As explained in Chapter 6, the ambient sound exposure is quantified in terms of the unweighted and VHF-weighted hourly sound pressure level ($L_{p,1\mathrm{h}}$). The distance from each CPOD to the piling locations is also known, and can also be used in the modelling.

We have two research questions, with each research question having three sub-questions. For the answering of each question, a statistical model was made, thus six models were made.

The research questions are as follows:
1) Is there a relationship between the PPM, and the VHF weighted SPL?
2) Is there a relationship between the PPM, and the unweighted SPL?

The sub-questions for each of these two research questions are as follows:
a) Does the distance to the piling sound matter?
b) Is the relationship between PPM and SPL different when there is piling versus when there is no piling?
c) Will leaving out piling-related covariates result in a worse (or better) model compared to the models for questions a and b.

This appendix describes the statistical analyses performed to answer the above research questions.

F.1 Materials & Methods

The materials for and the setup of the analysis was completely similar to that of the SEL analysis (see §5.2).

F.1.1 Modelling choice

There are three main relations of interests that we attempt to model:
- The relation between the presence and number of positive minutes per hour, and SPL, taking also into account the mean distance to piling sound (mdtps). This will be referred to as the "**analysis A**."
- The relation between the presence and number of positive minutes per hour, and the SPL. But here we create 2 separate PPM ~ SPL relationships: one during the moments when there is piling, and one during the moments when there is no piling. This will be referred to as the "**analysis B**."
- The relation between the presence and number of positive minutes per hour, and the SPL, without any piling-related covariates. This will be referred to as the "**analysis C**."

Each of these three analyses requires a statistical model. The data showed extreme zero-inflation (about 90% of the response values consisted of zeros). For each of the three main relationships of interest (PPM ~ VHF SEL, PPM ~ unweighted SEL, and PPM ~ distance to piling), a hurdle model set was used consisting of two models:

- a Bernoulli model with complementary log-log link function, modelling the presence (PPM>0) or absence (PPM=0) of the Porpoise positive minutes. A complementary log-log link function was used instead of the more common logit link function to handle the zero inflation better.
- a Gamma model with log-link function and with offset term, modelling the non-zero porpoise positive minutes (PPM>0), proportional to 60 (because one hour has 60 minutes).

Due to the extreme zero-inflation, the Gamma model is, unfortunately, statistically not significant nor interesting. The main documentation of this study therefore focusses on the Bernoulli model only.

F.1.1.1 The main effects

The main effects were all implemented in the statistical models as low-rank thin-plate smoothers.

F.1.1.2 Piling effects

For analysis B

For analysis B, we need the distance to piling sound, but only when there is piling. We also need a coefficient to be active when there is no piling. To accomplish both , the following was done. The mean distance to piling was set to be -1 when no piling event was taking place. Then they were entered in the model as a (custom) spline, which one can define as follows.

Let x be the mean distance to piling sound ("mdtps") effect, and let β be the vector of coefficients. One can then define the spline with the following function:

$$f_{\text{mdtps}}(x) = \beta_1 \times (x)_- + \beta_2 \times (x)_+ + \beta_3 \times ((x)_+)^2 + \beta_4 \times ((x)_+)^3$$

The above function is equivalent to:

$$f_{\text{mdtps}}(x) = \beta_1 \times -1_{\text{piling inactive}} + \beta_2 \times (x)_+ + \beta_3 \times ((x)_+)^2 + \beta_4 \times ((x)_+)^3$$

For analysis C

For analysis C, a two-part or split intercept term was made: One part to be active during piling, and one part when there is no piling. We define this "split intercept" as follows:

$$\beta_{(\text{split intercept})} = \beta_{\text{piling=T}} \times 1_{\text{piling=T}} + \beta_{\text{piling=F}} \times 1_{\text{piling=F}}$$

Here, "piling=T" refers to the presence of piling activity, and "piling=F" refers to the absence of piling activity, in any given time point.

F.1.1.3 Covariates

Regarding the other covariates in the models, additional intercepts in the Bernoulli models, model formulations, and model comparison methods, the same approach was taken as in §5.3.

F.1.2 Results - analysis A

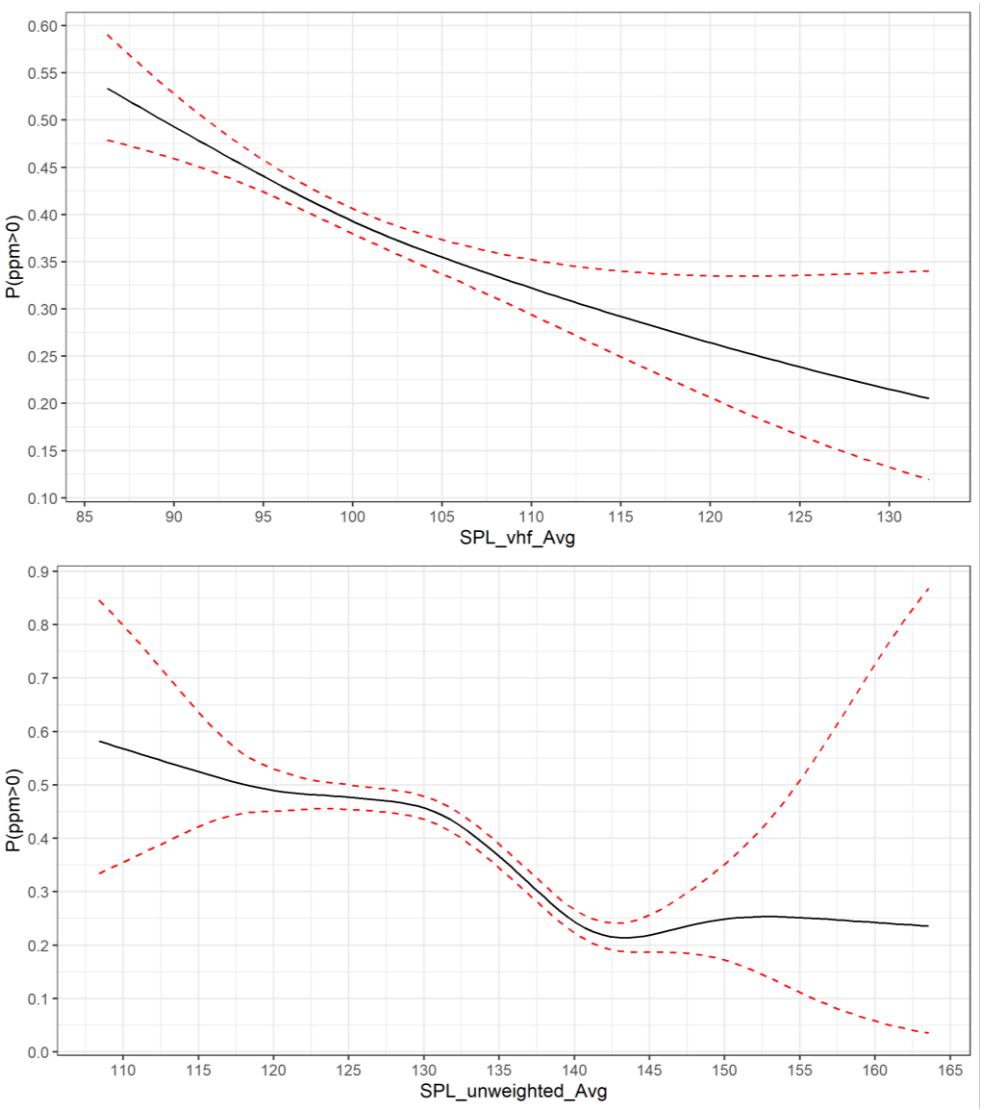

Figure F.1 Predictor effect plot of the main effect of the VHF weighted SPL and unweighted SPL, for the Bernoulli models of analysis A. The plots show the predicted probability of a porpoise positive minute (PPM) on the y-axis, while ignoring all random effects/smoothers (which average out to 0), and setting mdtps to -1 (so no piling events). The red dashed lines indicate the 95% confidence interval.

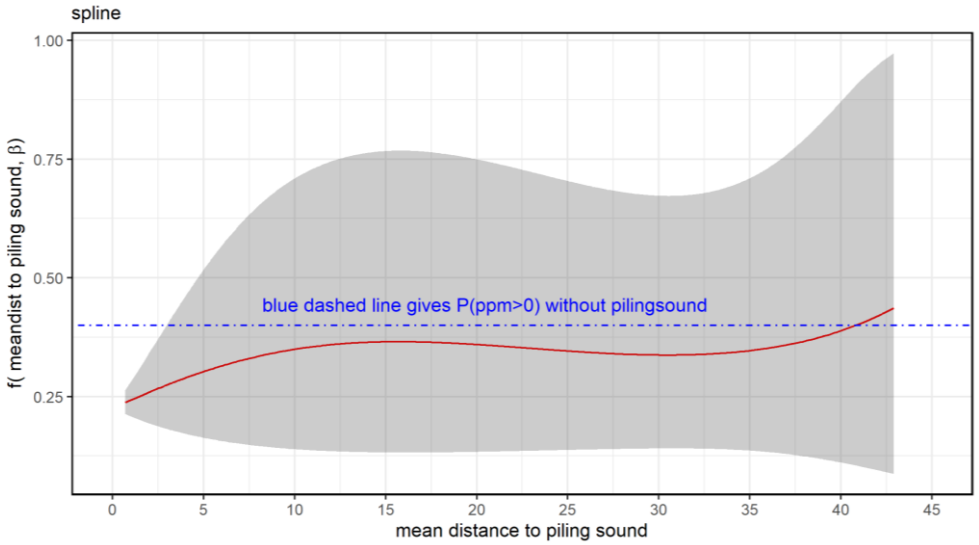

Figure F.2 Predictor effect plots of the mean distance to piling sound (mdtps) effect from the Bernoulli models of analysis A. The plot shows the predicted probability of a porpoise positive minute (PPM) on the y-axis, while ignoring all random effects/smoothers (which average out to 0). The shaded ribbon indicates the 95% confidence interval. The blue dashed lines indicate P(PPM>0) if there is no piling sound.

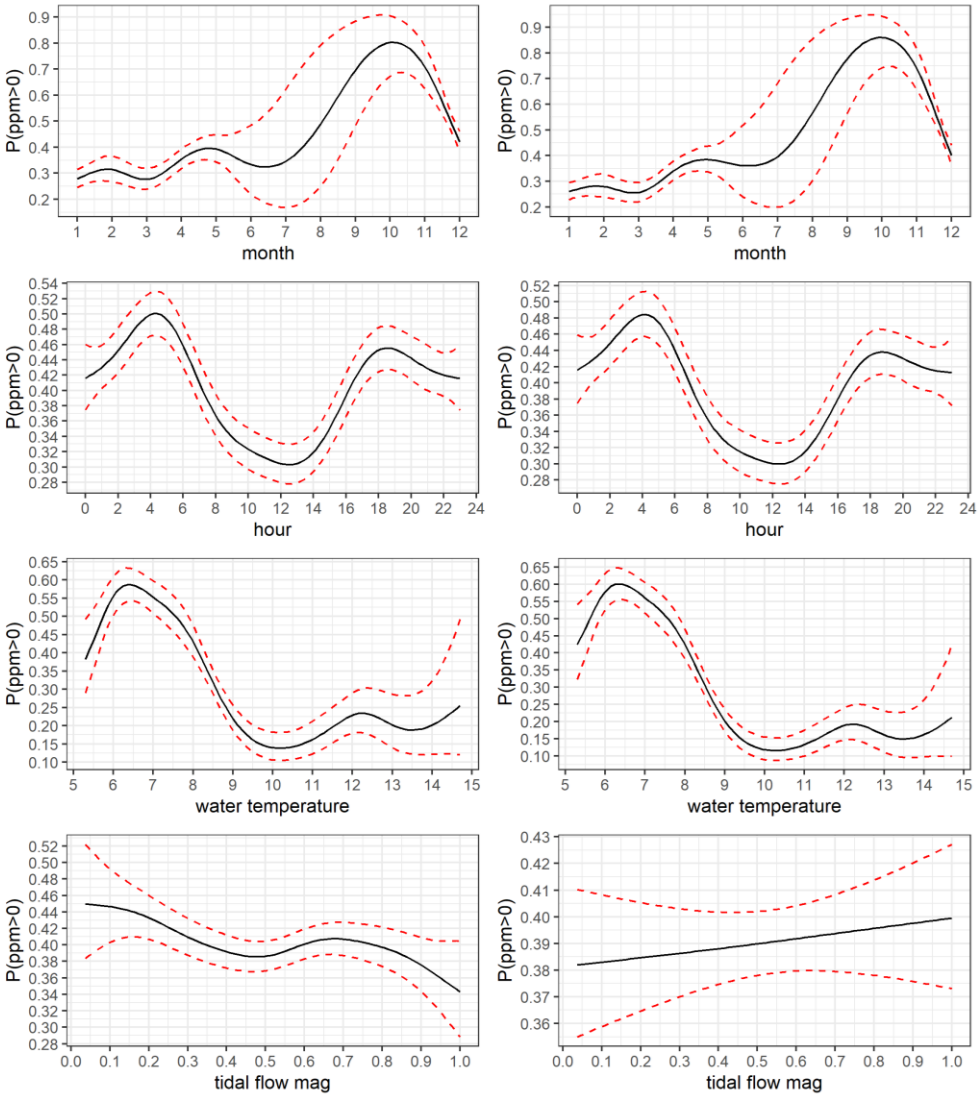

Figure F.3 Predictor effect plot of the regular smoothers from the Bernoulli models of analysis A. The plots on the left are from the VHF-weighted SPL model, the plots on the right from the unweighted SPL model. The plots show the predicted probability of a porpoise positive minute (PPM), while ignoring the other random effects/smoothers (which average out to 0), and setting mdtps to -1 (so no piling). The red dashed lines indicate the 95% confidence interval.

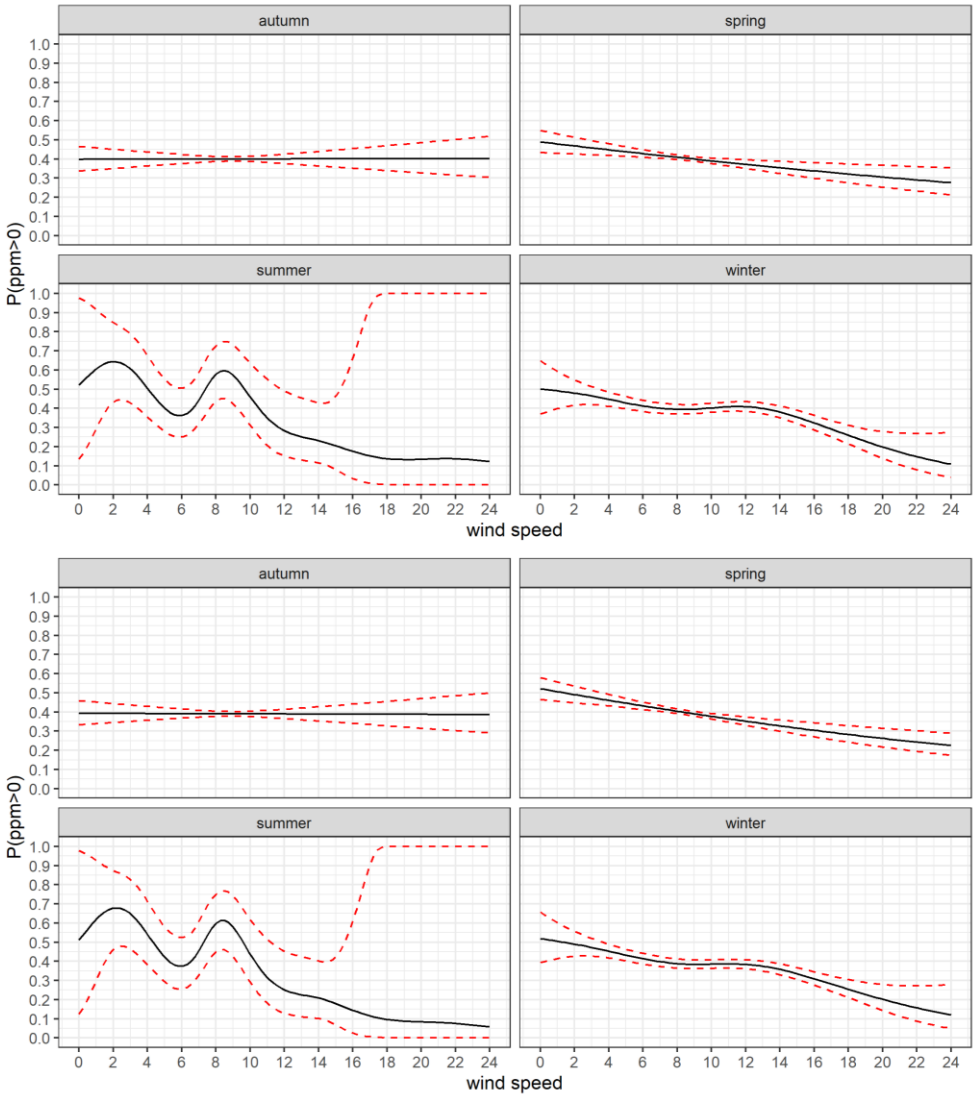

Figure F.4 Predictor effect plot of the wind and season interaction smoothers from the Bernoulli models for analysis A. The 4 plots on the top are from the VHF-weighted SPL model, the 4 plots on the bottom from the unweighted SPL model. a weighted SEL based Bernoulli model. The plots show the predicted probability of a porpoise positive minute (PPM), while ignoring the other random effects/smoothers (which average out to 0), and setting mdtps to -1 (so no piling). The red dashed lines indicate the 95% confidence interval.

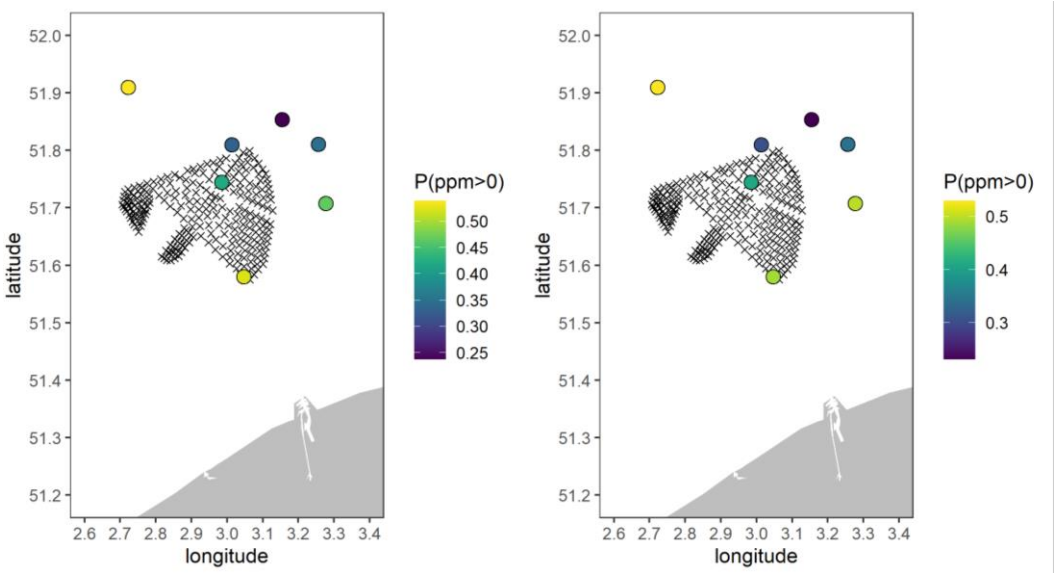

Figure F.5 Predictor effect plot of the random intercept for location for the regular SPL Bernoulli models. The left plot is from the VHF-weighted SPL Bernoulli model, and the right plot from the unweighted SPL Bernoulli model. The large points indicate the random effects, the crosses indicate the positions of the piling locations, and a small part of the land mass is visible to the south. The aspect ratio of the figure has been corrected for the longitude and latitude locations. The plots show the predicted probability of a porpoise positive minute (PPM), while ignoring the other smoothers (which average out to 0). The red dashed lines indicate the 95% confidence interval.

F.1.3 Results - analysis B

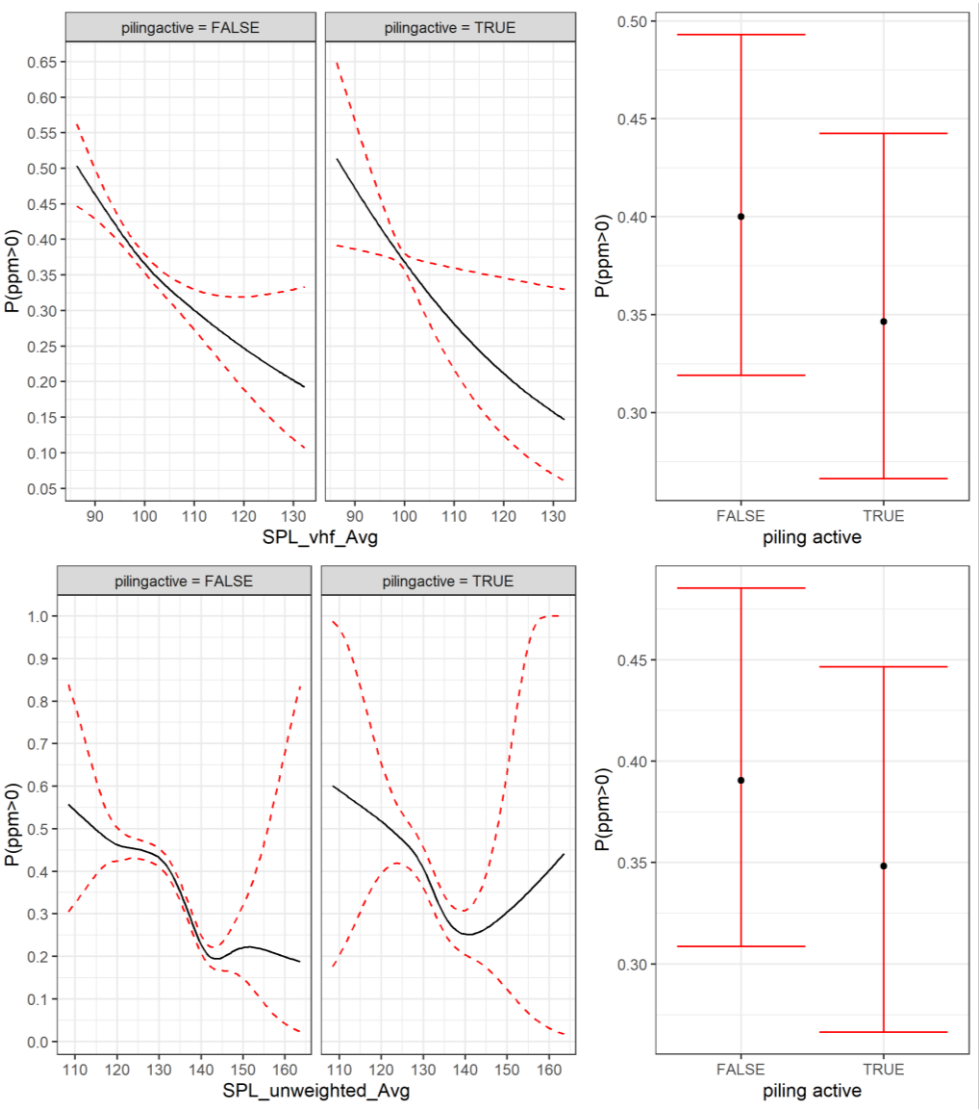

Figure F.6 Predictor effect plots of the effects of the VHF weighted and unweighted SPL, of the covariate piling active, and of their interaction, from the Bernoulli models of analysis B. The plots show the predicted probability of a porpoise positive minute (PPM) on the y-axis, while ignoring all random effects/smoothers (which average out to 0). The red dashed lines indicate the 95% confidence interval. The plots on the top are from the VHF-weighted SPL, and the plots on the bottom from the unweighted SPL.

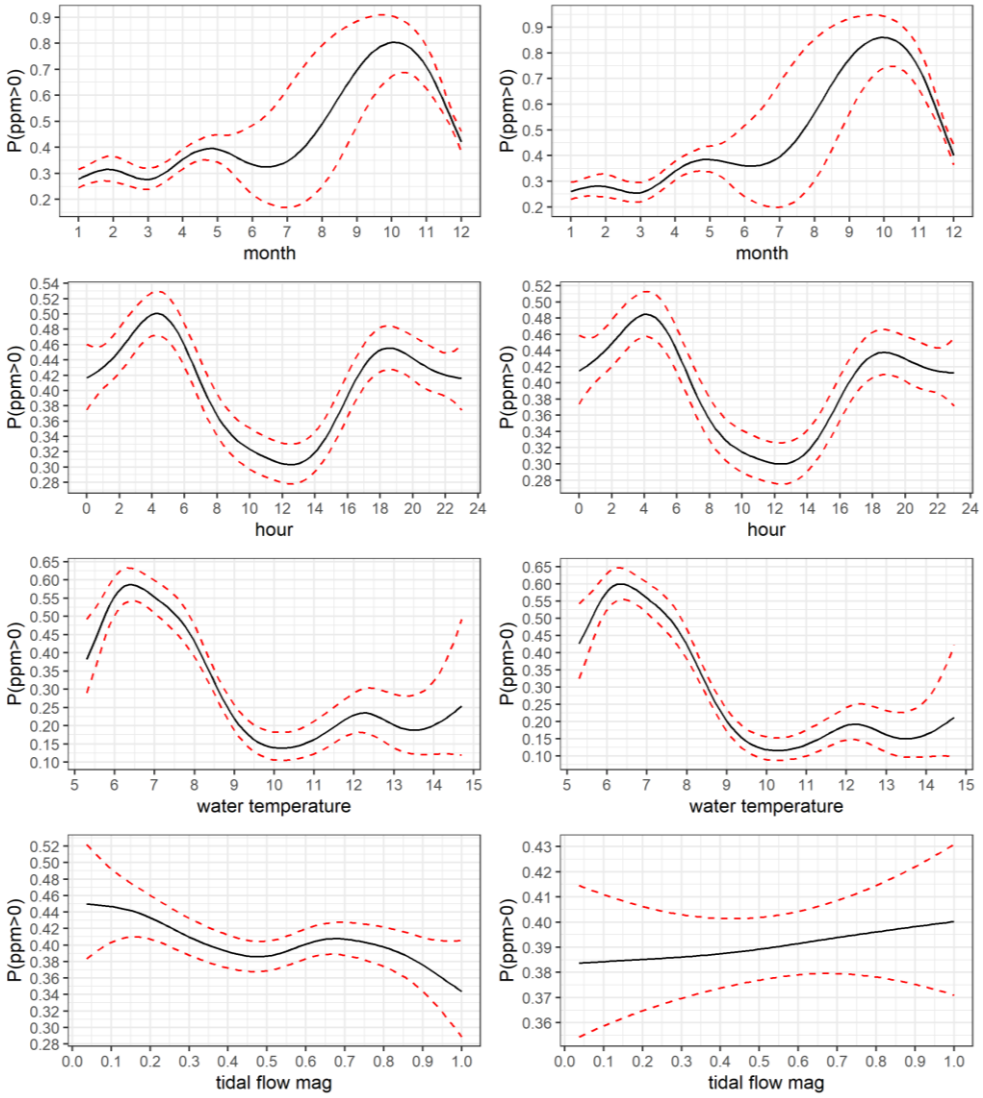

Figure F.7 Predictor effect plot of the regular smoothers from the Bernoulli models of analysis B. The plots on the left are from the VHF-weighted SPL model, the plots on the right from the unweighted SPL model. The plots show the predicted probability of a porpoise positive minute (PPM), while ignoring the other random effects/smoothers (which average out to 0), and while piling active is set to False. The red dashed lines indicate the 95% confidence interval.

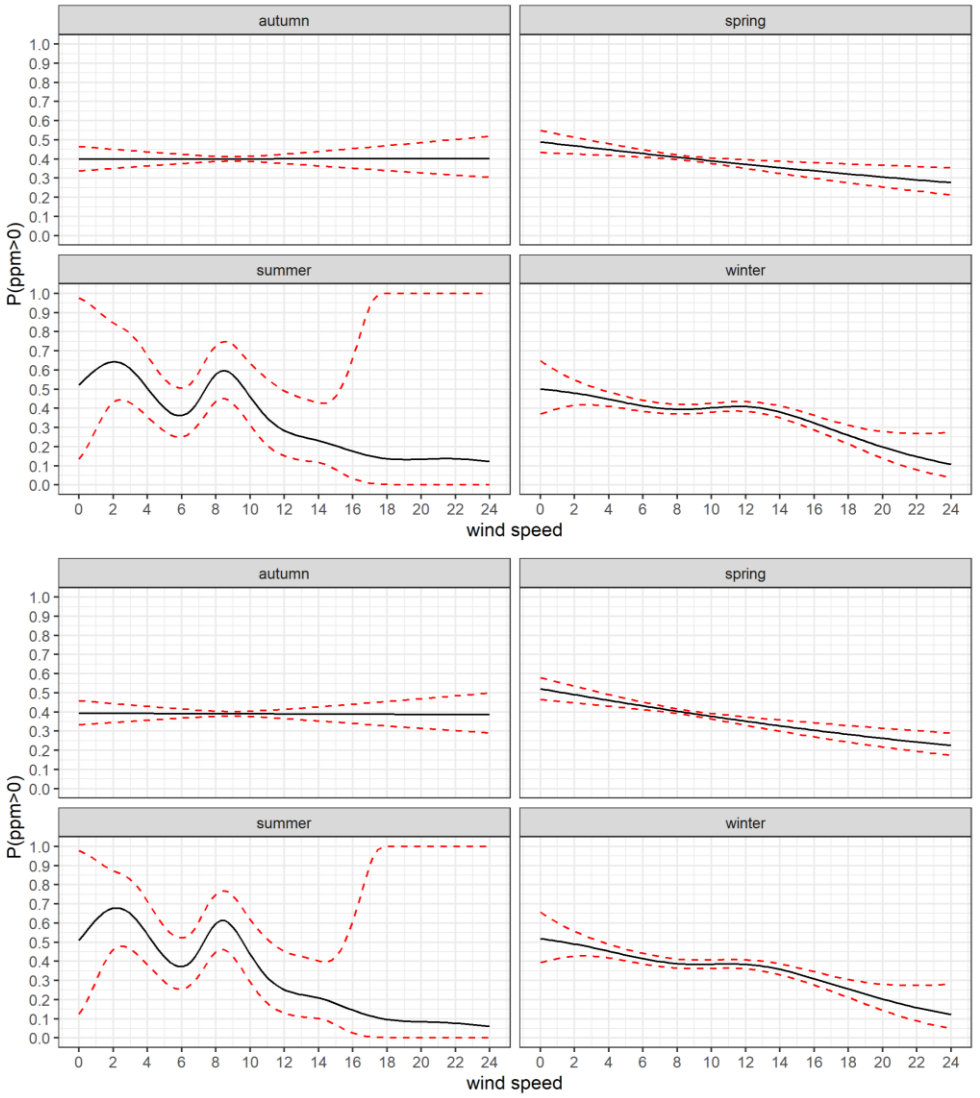

Figure F.8 Predictor effect plot of the wind and season interaction smoothers from the Bernoulli models for analysis B. The 4 plots on the top are from the VHF-weighted SPL model, the 4 plots on the bottom from the unweighted SPL model. a weighted SEL based Bernoulli model. The plots show the predicted probability of a porpoise positive minute (PPM), while ignoring the other random effects/smoothers (which average out to 0), and while piling active is set to False. The red dashed lines indicate the 95% confidence interval.

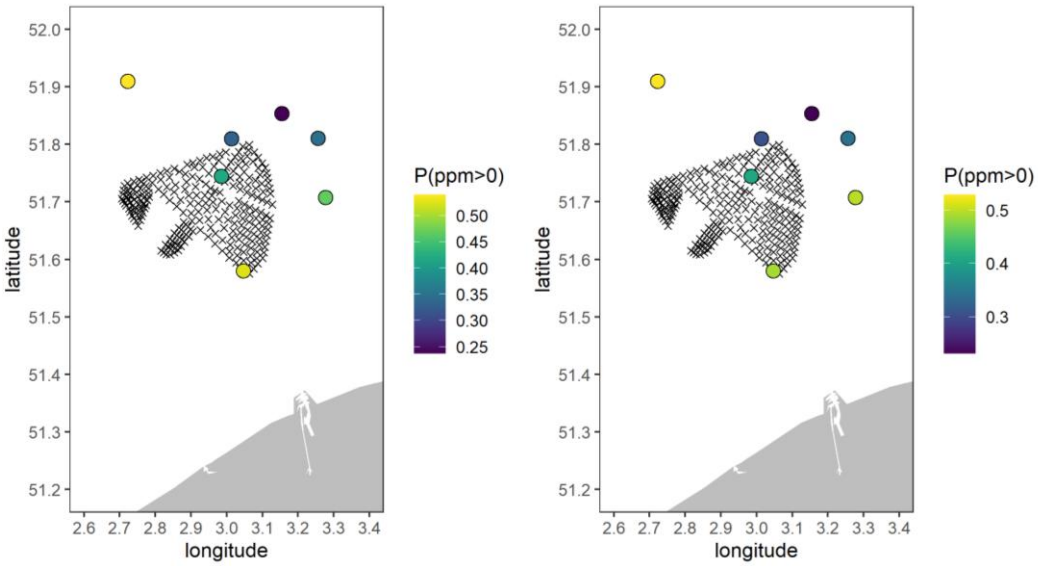

Figure F.9 Predictor effect plot of the random intercept for location from the Bernoulli models of analysis B. The left plot is from the VHF-weighted SPL Bernoulli model, and the right plot from the unweighted SPL Bernoulli model. The large points indicate the random effects, the crosses indicate the positions of the piling locations, and a small part of the land mass is visible to the south. The aspect ratio of the figure has been corrected for the longitude and latitude locations. The plots show the predicted probability of a porpoise positive minute (PPM), while ignoring the other smoothers (which average out to 0). The red dashed lines indicate the 95% confidence interval.

F.1.4 Results - analysis C

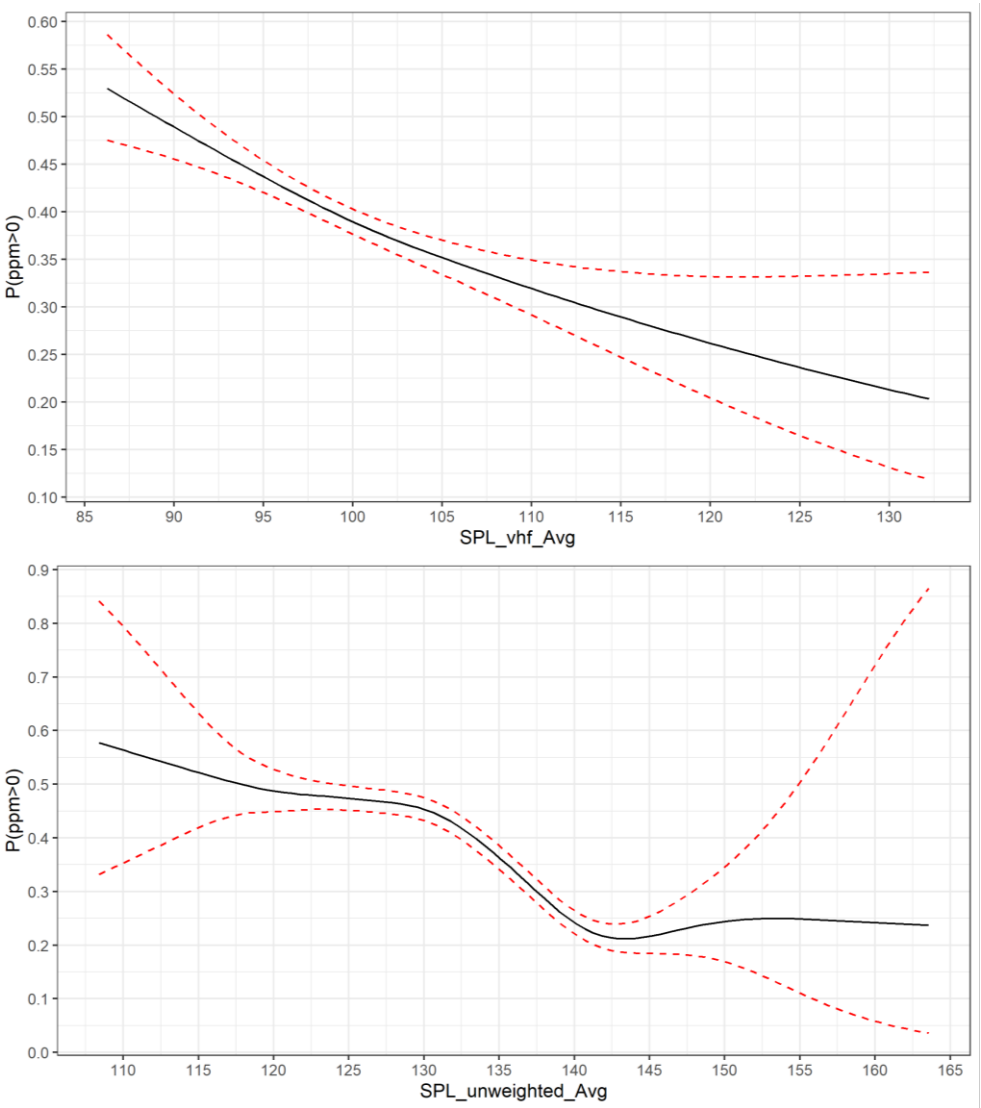

Figure F.10 Predictor effect plot of the main effect of the VHF weighted SPL and unweighted SPL, from the Bernoulli models of analysis C. The plots show the predicted probability of a porpoise positive minute (PPM) on the y-axis, while ignoring all random effects/smoothers (which average out to 0). The red dashed lines indicate the 95% confidence interval.

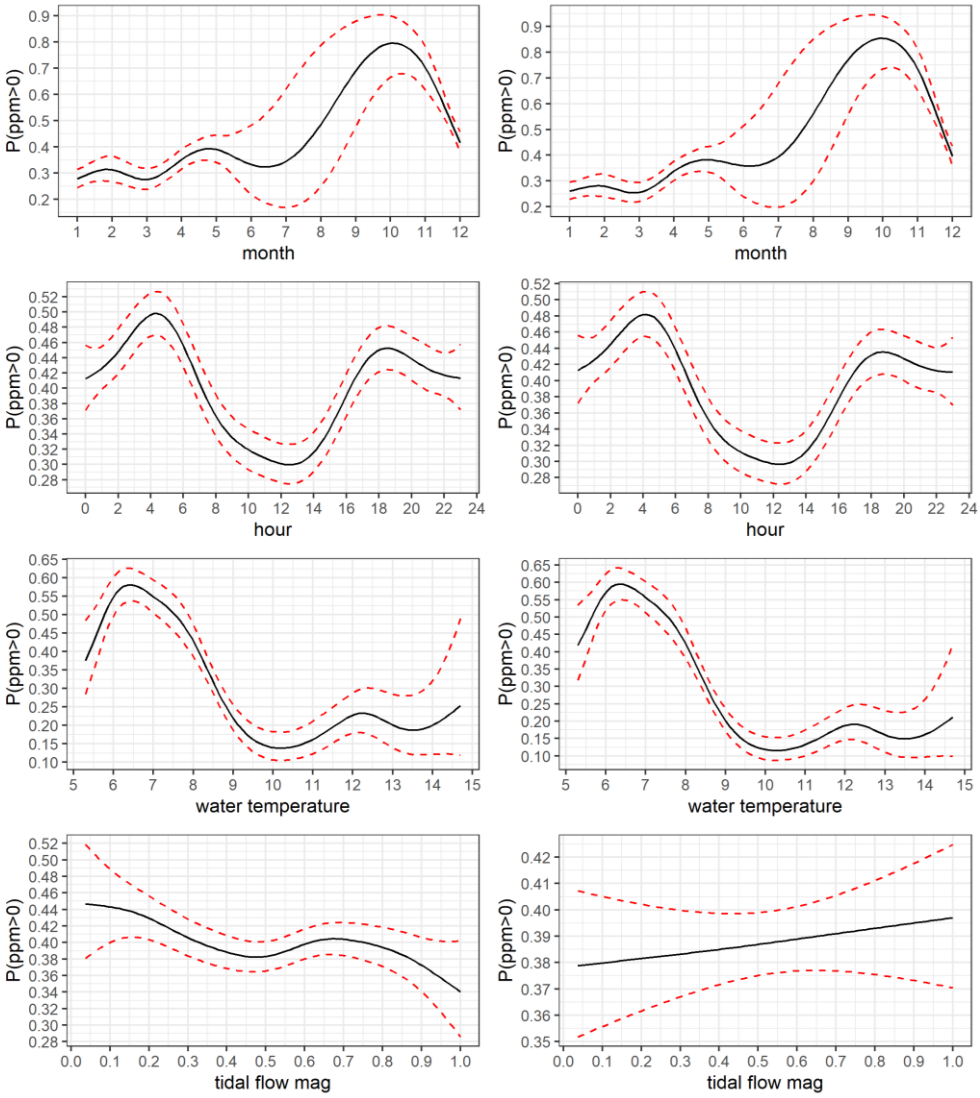

Figure F.11 Predictor effect plot of the regular smoothers from Bernoulli models of analysis C. The plots on the left are from the VHF-weighted SPL model, the plots on the right from the unweighted SPL model. The plots show the predicted probability of a porpoise positive minute (PPM), while ignoring the other random effects/smoothers (which average out to 0). The red dashed lines indicate the 95% confidence interval.

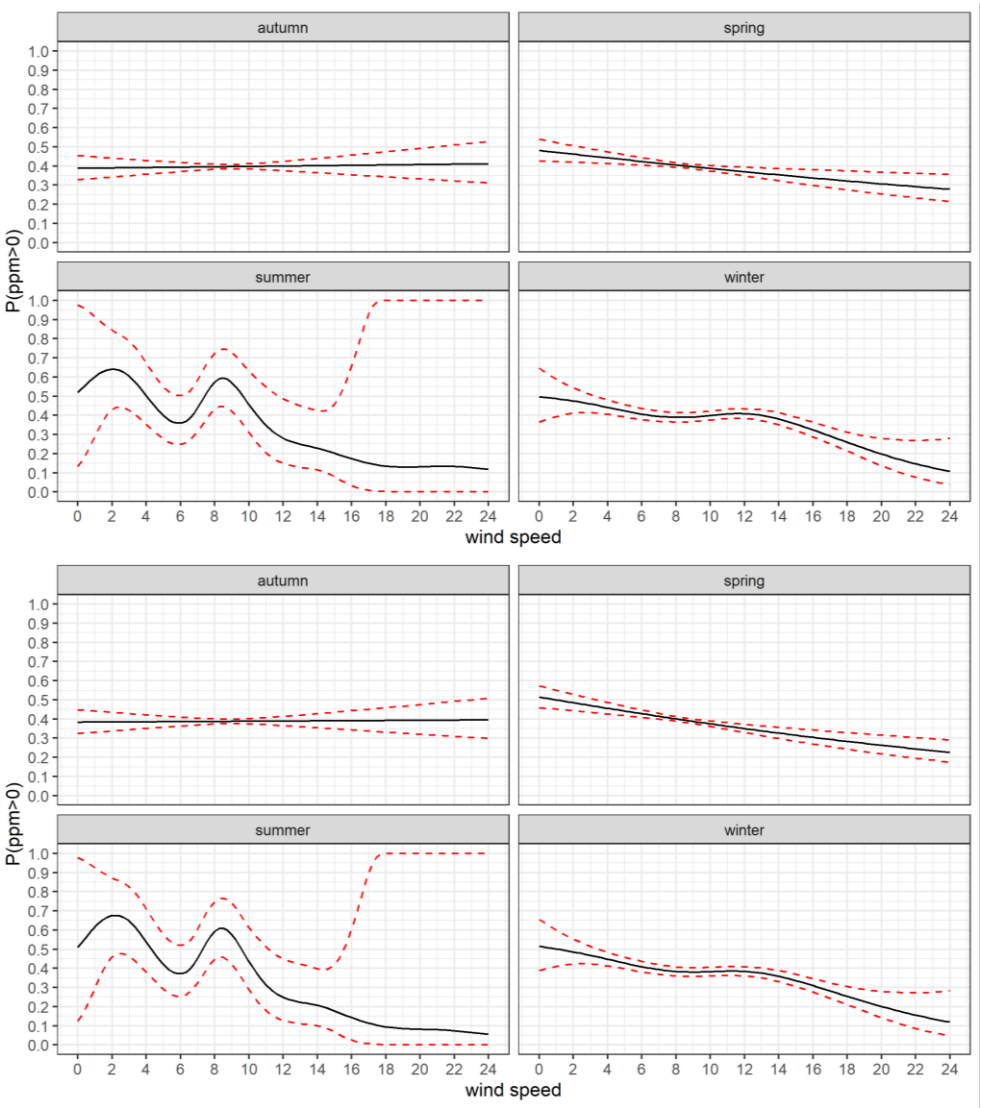

Figure F.12 Predictor effect plot of the wind and season interaction smoothers from the Bernoulli models of analysis C. The 4 plots on the top are from the VHF-weighted SPL model, the 4 plots on the bottom from the unweighted SPL model. a weighted SEL based Bernoulli model. The plots show the predicted probability of a porpoise positive minute (PPM), while ignoring the other random effects/smoothers (which average out to 0). The red dashed lines indicate the 95% confidence interval.

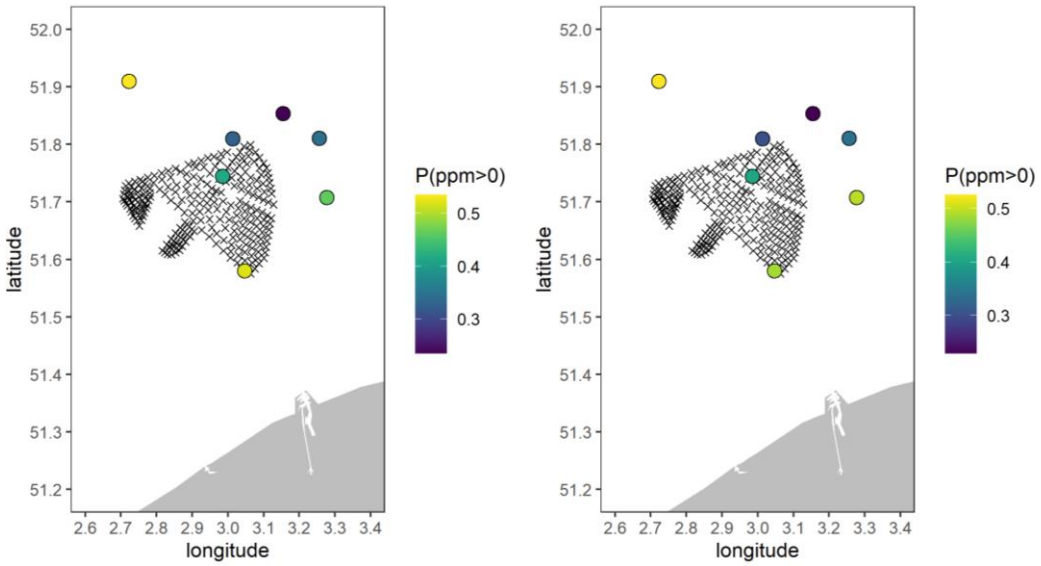

Figure F.13 Predictor effect plot of the random intercept for location from the Bernoulli models of analysis C. The left plot is from the VHF-weighted SPL Bernoulli model, and the right plot from the unweighted SPL Bernoulli model. The large points indicate the random effects, the crosses indicate the positions of the piling locations, and a small part of the land mass is visible to the south. The aspect ratio of the figure has been corrected for the longitude and latitude locations. The plots show the predicted probability of a porpoise positive minute (PPM), while ignoring the other smoothers (which average out to 0). The red dashed lines indicate the 95% confidence interval.

F.1.5 Results - model comparisons

For the model comparisons two measures are used: The fit of the models (expressed in specificity, sensitivity, and 0/1-loss; see Table F.1), and the information criterion of the models (expressed in AIC and BIC; see Table F.2). All models had very similar AIC and BIC values and they all fit comparably well. No convincing indication of the superiority of some models over other models was found.

Table F.1 Sensitivity, specificity, and the 0/1 loss of each of the 6 Bernoulli models. The closer specificity and sensitivity are to 1, the better. The closer the 0/1-loss is to 0, the better.

Analysis	SPL type	chosen alpha	sensitivity	specificity	0/1 loss
A	VHF	1.55	0.5126	0.7321	0.2928
A	Unweighted	1.55	0.5278	0.7302	0.2928
B	VHF	1.55	0.5126	0.7321	0.2928
B	Unweighted	1.55	0.5284	0.7296	0.2933
C	VHF	1.55	0.5121	0.7320	0.2930
C	Unweighted	1.55	0.5255	0.7296	0.2936

Table F.2 AIC and BIC of each of the 6 Bernoulli models. The lower AIC and BIC, the better.

Analysis	SPL type	AIC	BIC
A	VHF	21,145.87	21,599.24
A	Unweighted	20,921.60	21,381.56
B	VHF	21,142.48	21,578.61
B	Unweighted	20,919.43	21,393.59
C	VHF	21,144.89	21,565.25
C	Unweighted	20,920.91	21,353.69

F.2 Conclusions and Discussion

Just like in the SEL analysis, the models for SPL suffered from the extreme zero-inflation, which is particularly noticeable in the low true positive rate (sensitivity). The model diagnostics did not reveal any violation of the model assumptions. The results found can therefore be considered trustworthy. The same problem was present in the SEL analysis (see Appendix D).

The VHF-weighted SPL shows generally the same effect in all three analyses: the higher the SPL, the lower the probability of a Porpoise positive minute. This relationship is quite linear, and reasonably significant. The unweighted SPL is less certain. There is an approximately linear negative relation between PPM and unweighted SPL between an SPL of 130 dB to 140 dB re 1 μPa^2, but outside of that range the relationship appears to become flat and insignificant.

The piling covariates (whether it be the mean distance to piling sound effect in analysis A or the SPL $\times$ piling active interaction in analysis B) were not found to be relevant neither in analysis A nor in analysis B.

The effects of the other covariates (wind speed, tidal flow magnitude, and the temporal smoothers) are very similar between the six Bernoulli models. We can see the following in these covariates:

- There is a small dip in the probability of a PPM around midday (12 o'clock).
- There is a higher probability of PPM at lower water temperatures.
- Tidal flow magnitude shows a different effect, depending on whether the model uses VHF-weighted or unweighted SPL. For the VHF-weighted SPL models, it holds that the higher the tidal flow magnitude, the lower the probability of PPM, though the effect is not very strong. For the unweighted SPL models, it has an opposite effect, though again not very strong. Tidal flow magnitude covariate was found to be slightly significant in the models with VHF-weighted SPL, and insignificant in the models with unweighted SPL.
- During winter and summer, a mostly linear, negative relationship was found between wind speed and the probability of a PPM. During the summer and autumn, this relationship was found to be mostly insignificant.

G Porpoise behavioural response – SPL – Gemini

This section describes the results of the quasi-Poisson analysis that studies the relation of unweighted and porpoise-weighted SPL to the number of porpoise-positive minutes per hour. In contrast to the binomial analysis, where only presence (PPM/h>0) and absence (PPM/h = 0) are considered, we here get a finer understanding of the quantity of porpoise-positive minutes that is detected in an hour. It must be noted that there is currently no way to translate the number of positive porpoise minutes in an hour to the number of animals detected in an hour, as there is no way to know which clicks come from what animal. However, it tells us something about the general activity of present porpoises that hour.

The porpoise-positive minutes (PPM) relations observed here are similar to the binomial analysis of porpoise presence. Figure G.1 shows the modelled relations between PPM/h and the variables of interest. For unweighted SPL, there is a sigmoid-like relation, where ambient sound has no influence on porpoise activity for the first part. Around 105 dB, PPM steadily decreases close to zero.
This would mean that from 105 dB, there seems to be disturbance, after which harbour porpoise activity in the area decreases. For weighted SPL, the relation is very similar to the binomial analysis. There first seems to be a slight increase for the lower SPL values, after which it steadily decreases. The peak value is around 78 dB.

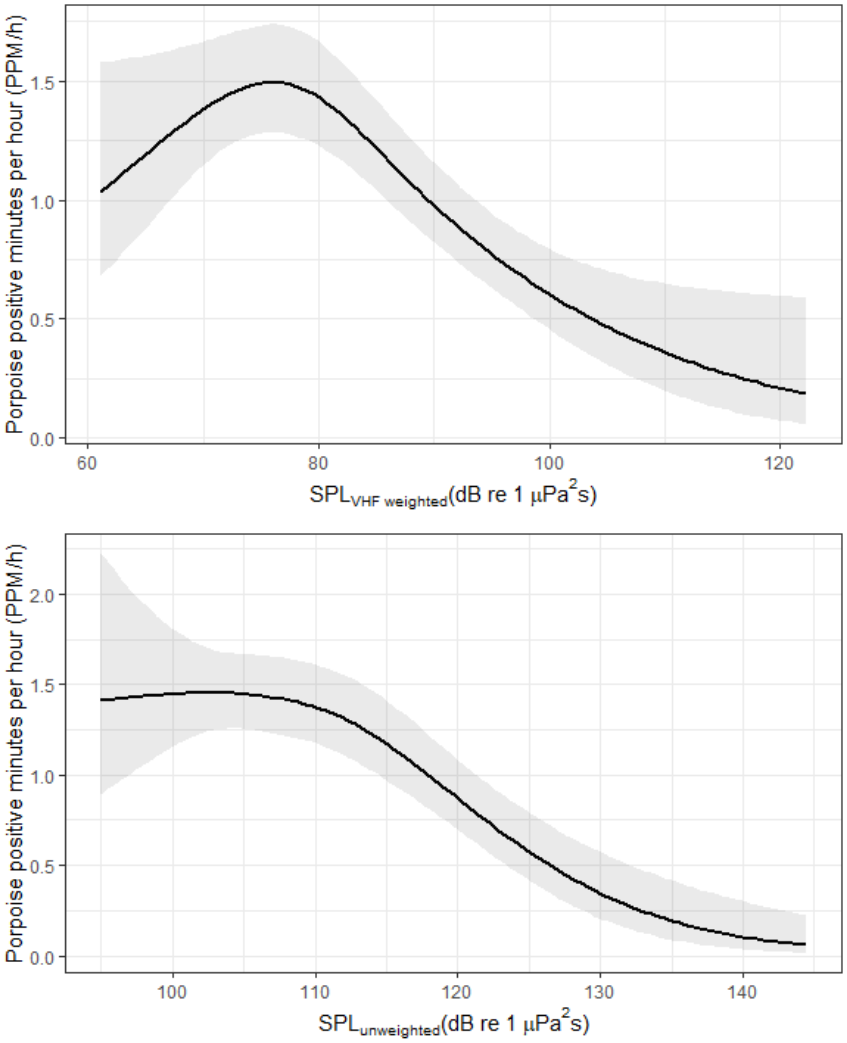

Figure G.1 Comparison of unweighted and weighted SPL in relation to porpoise positive minutes per hour (PPM/h) of the two quasi-Poisson models for the Gemini data. Upper: unweighted SPL (10Hz-20kHz); Lower: VHF-weighted SPL$_{VHF}$(10Hz-20kHz); The shaded area indicates the 95% confidence interval of the model predictions.