Effects of offshore pile driving on harbour porpoise abundance in the German Bight

Assessment of Noise Effects

Final Report

Miriam J. Brandt, Anne-Cécile Dragon, Ansgar Diederichs, Alexander Schubert, Vladislav Kosarev, Georg Nehls

Veronika Wahl, Andreas Michalik, Alexander Braasch, Claus Hinz, Christian Ketzer, Dieter Todeskino

Marco Gauger, Martin Laczny, Werner Piper

Husum, June 2016

Prepared for Offshore Forum Windenergie

IBL Umweltplanung GmbH
Bahnhofstraße 14a
26122 Oldenburg
Tel. 0441 / 50 50 17-10
Fax. 0441 / 50 50 17-11
info@ibl-umweltplanung.de
www.ibl-umweltplanung.de

Institut für Angewandte Ökosystemforschung GmbH
Alte Dorfstraße 11
18184 Neu Broderstorf
Tel. 038204 / 618-0
Fax 038204 / 618-10
info@ifaoe.de
www.ifaoe.de

BioConsult SH GmbH & Co. KG
Schobüller Str. 36
25813 Husum
Tel. 04841 / 66 3 29-10
Fax 04841 / 66 3 29-19
info@bioconsult-sh.de
www.bioconsult-sh.de

Funding partners:

1. DONG Energy Wind Power A/S, Kraftværksvej 53, DK-7000 Fredericia
2. EnBW Energie Baden-Württemberg AG, Durlacher Allee 93 in 76131 Karlsruhe
3. Iberdrola Renovables Offshore Deutschland GmbH, Charlottenstraße 63 in 10117 Berlin
4. Offshore Forum Windenergie GbR, Kaiser-Wilhelm-Straße 93 in 20355 Hamburg;
5. OWP Butendiek GmbH & Co. KG, Stephanitorsbollwerk 3 in 28217 Bremen
6. PNE WIND AG, Peter-Henlein-Straße 2 - 4 in 27472 Cuxhaven
7. RWE Innogy GmbH, Gildehofstraße 1 in 45127 Essen
8. STRABAG OW EVS GmbH, Reeperbahn 1 in 20359 Hamburg
9. TenneT Offshore GmbH, Bernecker Straße 70 in 95448 Bayreuth
10. Vattenfall Europe Windkraft GmbH, Überseering 12 in 22297 Hamburg
11. wpd offshore GmbH, Stephanitorsbollwerk 3 in 28217 Bremen

Scientific partners:

12. Ocean Breeze Energy GmbH & Co. KG, Flughafenallee 11 in 28199 Bremen
13. E.ON Climate and Renewables GmbH, Brüsseler Platz 1 in 45131 Essen
14. Global Tech I Offshore Wind GmbH, Am Sandtorkai 62, Dock 4 in 20457 Hamburg;
15. Horizont II Renewable GmbH, Köpenicker Straße 154 in 10997 Berlin
16. Nordsee Offshore MEG I GmbH, Humboldtstraße 30/32 in 70771 Leinfelden-Echterdingen
17. Offshore Windpark RIFFGAT GmbH & Co. KG, Tirpitzstraße 39 in 26122 Oldenburg;
18. Stiftung Offshore Windenergie, Oldenburger Straße 65 in 26316 Varel
19. WindMW GmbH, Schleusenstraße 12 in 27568 Bremerhaven

Further scientific partners with special agreement:

20. Trianel Windkraftwerk Borkum GmbH & Co. KG, Krefelder Straße 203 in 52070 Aachen

IBL Umweltplanung GmbH
Bahnhofstraße 14a
26122 Oldenburg
Tel.: 0441 505017-10
www.ibl-umweltplanung.de

Institut für Angewandte Ökosystemforschung GmbH
Schulterblatt 120
20537 Hamburg
Tel.: 040 432139000
www.ifaö.de

BioConsult SH GmbH & Co KG
Schobüller Str. 36
25813 Husum
Tel.: 04841 66 32 9 -10
www.bioconsult-sh.de

Lead:	BioConsult SH GmbH & Co KG	Dr. Georg Nehls
Editing:	Hourly POD-data chapter: BioConsult SH GmbH & Co KG	Dr. Miriam J. Brandt Dr. Anne-Cécile Dragon Ansgar Diederichs Alexander Schubert Vladislav Kosarev
	Daily POD-data & PCoD chapters: IBL Umweltplanung GmbH	Veronika Wahl Andreas Michalik Dr. Alexander Braasch Dr. Claus Hinz Christian Ketzer Dieter Todeskino
	Aerial survey chapter: IfAÖ GmbH	Marco Gauger Martin Laczny Werner Piper

Acknowledgments:

We thank Dr. Michael Bellmann and Siegfried Gündert from itap for providing sound measurements and calculations. We are also grateful for statistical advice provided by Dr. Charles Paxton from CREEM. Furthermore, we thank Prof. Dr. John Harwood for answering questions on the interim PCoD model.

Contents

List of figures

List of tables

List of abbreviations

ABW: Offshore wind farm Amrumbank West

AIC: Akaike Information Criterion (statistical number, value for comparison of different models)

ARIMA: auto-regressive (AR)/ integrative (I)/ moving average (MA) process

AV: Offshore wind farm Alpha Ventus

BARD: Offshore wind farm BARD Offshore 1

BRT: Boosted Regression Tree

BD: Offshore wind farm Butendiek

BWII: Offshore wind farm Borkum West II

DT: Offshore wind farm DanTysk

DPH: Detection Positive Hours

DPM: Detection Positive Minutes

EEZ: Exclusive economic zone

ESW: Effective strip width

GAM: Generalised Additive Model

GAMM: Generalised Additive Mixed-effects Model

GLM: Generalised Linear Model

GTI: Offshore wind farm Global Tech I

HRW: Hour relative to piling work

MSO: Offshore wind farm Meerwind Süd/Ost

NSO: Offshore wind farm Nordsee Ost

NEY: Northern Energy

OWF: Offshore wind farm

(P)ACF: (Partial) AutoCorrelation Function

PCoD: Population Consequences of Disturbanse

POD: Porpoise detector

PTS: Permanent Threshold Shift

RG: Offshore wind farm Riffgat

SAC: Special area of conservation

SEL: Sound Exposure Level

SST: Sea Surface Temperature

SSTA: Sea Surface Temperature Anomaly

TTS: Temporary Threshold Shift

1 SUMMARY

This study analyses the effects of the construction of eight offshore wind farms within the German North Sea between 2009 and 2013 on harbour porpoises (*Phocoena phocoena*). It combines porpoise monitoring data from passive acoustic monitoring using Porpoise Detectors (POD data 2010-2013) and aerial surveys (2009-2013) with data on noise levels and other piling characteristics. These data were analysed in detail in connection to pile driving activities, most of which occurred with application of noise mitigation techniques in order to reduce disturbance effects.

Prior to investigating piling effects on porpoises, baseline analyses were conducted to identify the seasonal distribution of harbour porpoises in different geographic subareas. Daily POD data and aerial survey data highlighted similar seasonal patterns with higher densities in spring and summer. Highest porpoise occurrence was found next to the SAC Sylt Outer Reef in the northeast of the German Bight in early summer. Another high density area occurred near the SAC Borkum Reef Ground in the southwest almost year round, which is in line with previous findings.

In addition to porpoise monitoring data, noise measurements from the seven wind farms constructed between 2010 and 2013 were combined and noise levels extrapolated where measurements did not exist. Analyses of these measured noise levels revealed that there was high variability within each wind farm. Median noise levels during noise mitigated piling were about 10 dB lower than those measured during unmitigated piling. However, several noise levels measured during noise mitigated piling were as high as those during unmitigated piling and there was a high variability in these measurements within projects ranging over about 20 dB. This high variability probably results from differences in the combination of noise mitigation systems and how well a particular system worked at the time. It may also result from several environmental factors such as water depth, sediment and wind speed that all affect sound propagation. Probably due to these reasons, there was also no clear difference in noise levels between foundation types. The present study shows that noise mitigation systems used during this study were still under development and thus did not always work consistently well.

Establishing the relationship of noise levels to porpoise responses is crucial for environmental impact assessments based on noise prognosis for specific projects. Non-parametric analyses revealed a clear gradient in how much porpoise detections declined at different noise level classes: Compared to a baseline period 25-48 h before piling, porpoise detections declined by over 90 % at noise levels above 170 dB, but only by about 25 % at noise levels between 145 and 150 dB. Below 145 dB this decline was smaller than 20 % and may thus not clearly be related to noise emitted by the piling process. Based on the complete POD-dataset analysed at an hourly resolution using GAM techniques and controlling for other environmental variables, we also found a clear gradient in the decline of porpoise detections during piling with noise level. While a decline in porpoise detections was found at noise levels above 143 dB SEL_{05}, not all porpoises left the noise impacted area at that noise level.

In further analyses, distance from piling was used as a proxy for noise to analyse detailed effect ranges. This was done because it increased the sample size (noise data did not exist for each POD-position and each piling) and model fit using distance instead of noise improved. Analyses pooling all available POD-data yielded an effect range up to 17 km when analysed with General Additive Models (GAM). Non-parametric analyses revealed significant declines in porpoise detections dur-

ing piling when compared to 25-48 h before in up to 20-30 km, but only in up to 10-15 km was this decline at least 20 %. With increasing distances to the construction site, the magnitude of decline during piling clearly decreased.

When noise mitigation was considered within this GAM model, the estimated effect range of 14 km during noise mitigated piling was lower compared to the complete dataset (17 km) or unmitigated piling (between 20 and 34 km). Caution is required when interpreting these results because of the relatively low dataset for unmitigated piling events. Nevertheless, it shows that noise mitigation effectively reduced porpoise disturbance. This reduction in disturbance may be less than would usually be expected under properly working noise mitigation (when effects may be expected to only reach up to about 5 km). This is probably related to the high variability in noise level measurements due to the fact that noise mitigation systems were still under development and did not always work reliable at that time. Considerable improvement has happened since then. Our result that piling noise above 143 dB SEL_{05} led to disturbance effects in porpoises (even though not all porpoises were affected at these noise levels), supports earlier estimations by NEHLS ET AL. (2016) that properly working sound mitigation, under which 160 dB are not exceeded at a 750 m distance (as intended by the regulatory framework), would lead to a substantial reduction of the area in which porpoises are affected by about 90 %.

Project-specific models yielded large differences in effect ranges as well as effect magnitude. Declines in detection rates during piling in 0-5 km distance were smallest at the wind farm DT with 51 % and largest at BARD with 83 %. This also applies to effect ranges, which for DT were estimated to be 6 km based on GAM models and 0-5 km based on non-parametric statistics. During all other projects significant declines by at least 20 % were found in at least 5-10 km but occurred in up to 20-30 km distance. Such differences between projects cannot be explained by differences in noise levels alone as DT was not significantly quieter than several other projects. Instead it may be linked to a relatively high quality of feeding habitat and a lower motivation of porpoises to leave the noise impacted area, but exact reasons are currently not known.

From aerial survey data there was an indication for porpoise densities to be increased during and up to 12 h after piling at distances above 20 km. This effect could not be confirmed by POD-data, which could be related to the smaller spatial coverage of the latter. Elevated densities at distances above 20 km rely on little data, however, and need to be interpreted cautiously.

Effect duration after piling was about 20-31 h at the close vicinity of the construction site (up to 2 km) and decreased with increasing distance. Project-specific estimates ranged between 16 and 46 h (when defined as the first local maximum after an initial increase in detection rates), with the exception of DT where effect duration was difficult to define (no local maximum reached).

In all wind farm projects, we observed significant decreases in porpoise detections already prior to piling at distances of up to 10 km. This was independent of piling or deterrence measures. The most likely explanation for this are effects by increased shipping activity during preparation works in combination with increased sound propagation at low wind speed. It was found that deterrence effects prior to as well as during piling reached further at lower wind speed indicating that the effects of wind and sea state on sound propagation may be underestimated.

There was no indication for the presence of temporal cumulative effects. Only at BWII we found some indication for potential habituation of porpoises to piling. Neither analyses of hourly nor daily POD-data revealed any further indication for habituation. However, without any knowledge of porpoise residency patterns within the German Bight and individual responses to disturbance, this topic remains difficult to address.

From analyses of daily POD-data there was some indication for piling effects on porpoise detections to differ between seasons: Piling effects were longer lasting during winter and autumn than during spring and summer. As porpoise density tends to be lower in autumn and winter this effect may be related to longer lasting effects at lower porpoise densities. However, this could not be confirmed when looking at area-specific piling effects. Piling effects were not generally longer lasting in areas of lower porpoise densities.

Using results from aerial survey data and POD-data analyses, the PCoD model was applied to estimate disturbance consequences of wind farm construction on the population level. After exploration of the interim PCoD model, several limitations of the model were pointed out that may be improved before providing a realistic estimation for porpoise population trends as a result of disturbance. Applying the PCoD model using conservative input parameters for construction effects arising from the present study (increasing the chances for the model to predict a population decline), the risk of a decline of 1% of the population in the German Bight is estimated to be below 30 %. The predicted median decline is below the 1 % generally considered as critical for all chosen time periods and varies between 0.9 % for the piling period and 0.2 % for twelve years after piling had finished.

There were no indications for such a population decline of harbour porpoises over the five year study period arising from analyses of daily POD data and aerial survey data at a larger scale. Despite extensive construction activities over the study period and an increase in these over time, there was no negative trend in acoustic porpoise detections or densities within any of the subareas studied. In some areas, POD-data even detected a positive trend from 2010 to 2013.

On a regional scale, porpoise distribution patterns, as found by aerial survey data, differed between years. These regional changes could partly be related to wind farm construction sites but only within a radius of 20 km around piling. However, there was no evidence for an overall change in distribution patterns at a larger scale within the German Bight over the 5-year study period.

Even though clear negative short-term effects (1-2 days in duration) of offshore wind farm construction were found on acoustic porpoise detections and densities, there is currently no indication that harbour porpoises within the German Bight are presently negatively affected by wind farm construction at the population level. This is even though sound mitigation techniques were still under development and further improved after this study period.

2 INTRODUCTION AND AIMS OF THE PROJECT

In Europe, offshore wind energy is rapidly developing as an alternative energy source to nuclear power and fossil fuels. Specifically, the German offshore wind power production is supposed to expand to a nominal capacity of 15 GW until 2030 (BSH 2015). Since the offshore wind farm alpha ventus was installed in 2009, several research projects have been conducted to develop new noise mitigation methods, test noise mitigation measures and evaluate the regulatory framework for conducting environmental impact assessments.

During offshore wind farm construction, foundations are usually driven into the sea floor by means of noise-intense piling. Marine mammal species possess a very sensitive underwater hearing system, capable of detecting much higher frequencies than humans. Recent studies proved that pile-driving noise negatively affects seals' and cetaceans' hearing and disrupts the animals' natural behaviour (MADSEN et al. 2006; PUNT 2015; RUSSELL et al. 2015). A key species in this context is the harbour porpoise (*Phocoena phocoena*), the only resident cetacean species known to regularly reproduce in German waters (REID et al. 2003; SIEBERT et al. 2006) and listed as protected species in annex four of the council directive 92/43/EEC. Highly depending on echolocation for orientation and foraging, harbour porpoises are particularly vulnerable to noise-intense anthropogenic activities such as pile driving (MADSEN et al. 2006). Some of the energy exerted on the pile is transmitted into the water column as noise. Depending on how loud the noise is and how far animals are located from the source, it can affect their behaviour and/or induce physiological effects such as temporary or permanent increase in the hearing threshold (TTS = temporary threshold shift and PTS = permanent threshold shift). The combination of deterrence measures prior to piling and a soft-start in pile driving activities prevents the occurrence of such physiological effects. Hence, this study focuses on possible behavioural effects caused by noise emission during pile driving.

Previous studies on the effects of offshore wind farm construction on harbour porpoises used passive acoustic monitoring devices (e.g., C-PODs) that continuously record harbour porpoise echolocation, i.e. clicking activity. Passive acoustic devices allow comparing porpoise detections during the construction period to those of a preconstruction and/or post-construction period at high temporal resolution. Porpoise detections were shown to decrease significantly during piling up to 20 km around wind farm construction sites (TOUGAARD et al. 2009; BRANDT et al. 2011; DÄHNE et al. 2013A). In the absence of noise mitigation during pile driving, negative effects lasted up to two days within close vicinity of the foundations (TOUGAARD et al. 2009; BRANDT et al. 2011; BSH 2014; ROSE. et al. 2014). Furthermore, several studies analysed the distribution and behaviour of porpoises in relation to piling noise levels and tried to identify the noise level at which porpoise detections or abundance during piling significantly decreased compared to a given baseline period before or after piling. The onset of a behavioural reaction during pile driving (change in detection rates, density or observable behaviour) was estimated to occur at noise levels between 140 and 152 dB (BIOCONSULT SH 2009; DIEDERICHS et al. 2010; DEGRAER et al. 2012; DÄHNE et al. 2013A; BIOCONSULT SH et al. 2014; ROSE. et al. 2014). During an experimental study in captivity, KASTELEIN et al. (2013) observed a significant increase in jumping frequency of harbour porpoises when exposed to play back noise of pile driving of 145 dB re 1 µPa²s. The average of the lowest noise levels at which the studied animals started to jump out of the water was at 136 dB re 1 µPa²s. Hence, estimates as to what noise levels trigger a behavioural reaction in porpoises are quite variable, the reasons for which are not yet known. A recent review by TOUGAARD et al. (2015) confirms that behavioural reactions in porpoises highly depend on the frequency spectrum of the noise. They suggest that behavioural reactions are usually found at about 40-50 dB above the frequency-specific hearing threshold.

It may be expected that noise levels that trigger avoidance reactions in porpoises depend on the animals' return motivation, resulting from habitat quality, age, reproductive state, fitness, etc. Whether a statistically significant effect can be detected may also depend on prevailing porpoise densities. In addition, noise mitigation may also alter the broadband noise level that triggers a behavioural reaction because it alters the frequency spectrum of noise and diminishes the high-frequency component where porpoises have their most sensitive hearing capabilities. Despite a high number of studies dealing with impacts of offshore wind farms on harbour porpoises, trans-regional, long-term- and cumulative effects on porpoises as well as habituation of animals to piling have rarely been studied.

Between 2009 and 2013, eight offshore wind farms with a total of over 400 foundations were built in the German part of the North Sea. So far, Germany is the only country with mandatory noise mitigation measures for offshore wind farm construction with the aim not to exceed noise levels of 160 dB SEL_{05} at 750 m distance to the piling location as criterion for avoidance of physical injury of marine mammals. In the context of the "*Schallschutzkonzept*", published by the Federal Ministry for the Environment, Nature Conservancy, Building and Nuclear Safety in 2015 the Sound Exposure Level (SEL) of 140 dB was defined as precautionary criteria for disturbance effects. Consequently, the majority of the German North Sea wind farms were constructed under the use of various noise mitigation techniques. Much effort in terms of finances, research and planning was invested to design and plan their effective application. Therefore, it should be expected that the range and possibly also the duration of avoidance effects on porpoises was considerably reduced due to the application of noise mitigation. During these construction activities, extensive monitoring programs collected data on porpoise presence, which so far were only analysed for each single wind farm separately. Combining these data for a joint and cross-project analysis offers a unique opportunity to comprehensively study the pile driving impact on harbour porpoises within the whole German North Sea over a period of four to five years. For the present study, two datasets were used, one including abundance data from visual aerial surveys covering 2009 to 2013, and the other including relative abundance data from passive acoustic monitoring for the period 2010-2013. Furthermore, detailed information was gathered on deterrence and noise mitigation measures as well as several other piling characteristics like piling duration, applied energy and number of strokes. Measured noise levels were collected and extrapolated over the various distances where POD data existed. All these metadata, together with numerous environmental data, were prepared and merged with POD data and aerial survey data. This dataset allowed analysis of piling effects on porpoise acoustic detections and densities at both small and large spatio-temporal scales. Finally, a population dynamic model (Population Consequences of Disturbances = PCoD, HARWOOD et al. 2014) was used to better understand how strongly the entire regional porpoise population could be affected by construction activities in the long term.

Based on this large dataset, the main aim of the present study is to:

1) analyse small-scale avoidance behaviour of harbour porpoises in relation to piling, determine the noise levels that trigger a behavioural response and to estimate disturbance effect radii and duration (mainly chapter 4, but also 5-6)

2) test whether piling duration and consecutive piling events alter the behavioural reaction to piling which, depending on the direction of the effect, could indicate cumulative or habituation effects (mainly chapter 4, but also 5-6),

3) assess whether construction activities led to different distribution patterns and densities at a broader scale over the study period, which could indicate the presence of longer-term effects (chapters 5-7).

3 DESCRIPTION OF METADATA

The present study is looking at construction effects on harbour porpoises from eight wind farms that were built in the German North Sea between 2009 and 2013: Alpha Ventus (AV), BARD Offsore I (BARD), Borkum West II (BWII), DanTysk (DT), Global Tech I (GTI), Meerwind Süd/Ost (MSO), Nordsee Ost (NSO) and Riffgat (RG), whose geographic position can be seen in Figure 3-1. To investigate these effects on harbour porpoises, monitoring data of porpoise occurrence were available from passive acoustic monitoring (PODs) and from aerial transect surveys. Information on noise levels, noise mitigation measures and other piling characteristics were collected to investigate how these may affect behavioural reactions of porpoises to piling noise. Noise measurements of piling noise were available for many piling events consisting partly of measurements and partly of extrapolated/modelled data. Finally, environmental data were gathered from various open sources and extracted to match the spatiotemporal resolution of POD-data and aerial survey data on porpoise occurrence.

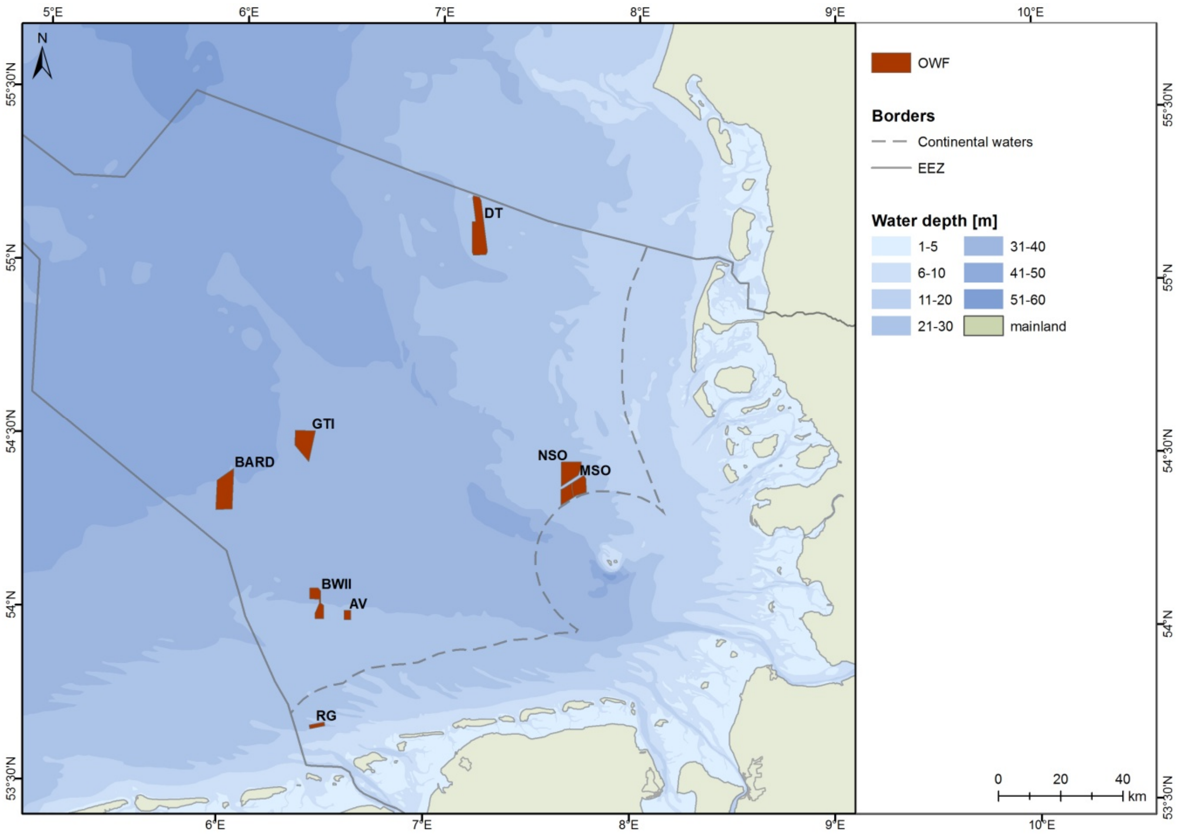


Figure 3-1 Map of the study area depicting all eight wind farms that were built between 2009 and 2013.

3.1 Pile driving data

While pile driving between 2009 and 2013 occurred during the construction of all eight wind farms only those constructed between 2010 and 2013 (all but AV) were considered for analysing POD-data. This is because a change in passive acoustic devices occurred in 2010 so that only POD-data from 2010 to 2013 were used. Construction of AV is still considered for analysing aerial survey data which last from 2009 to 2013. Analysing aerial survey data with respect to piling activities does not require fine scale information on piling characteristics. Therefore only the seven wind farms without AV will be looked at in more detail with respect to construction characteristics.

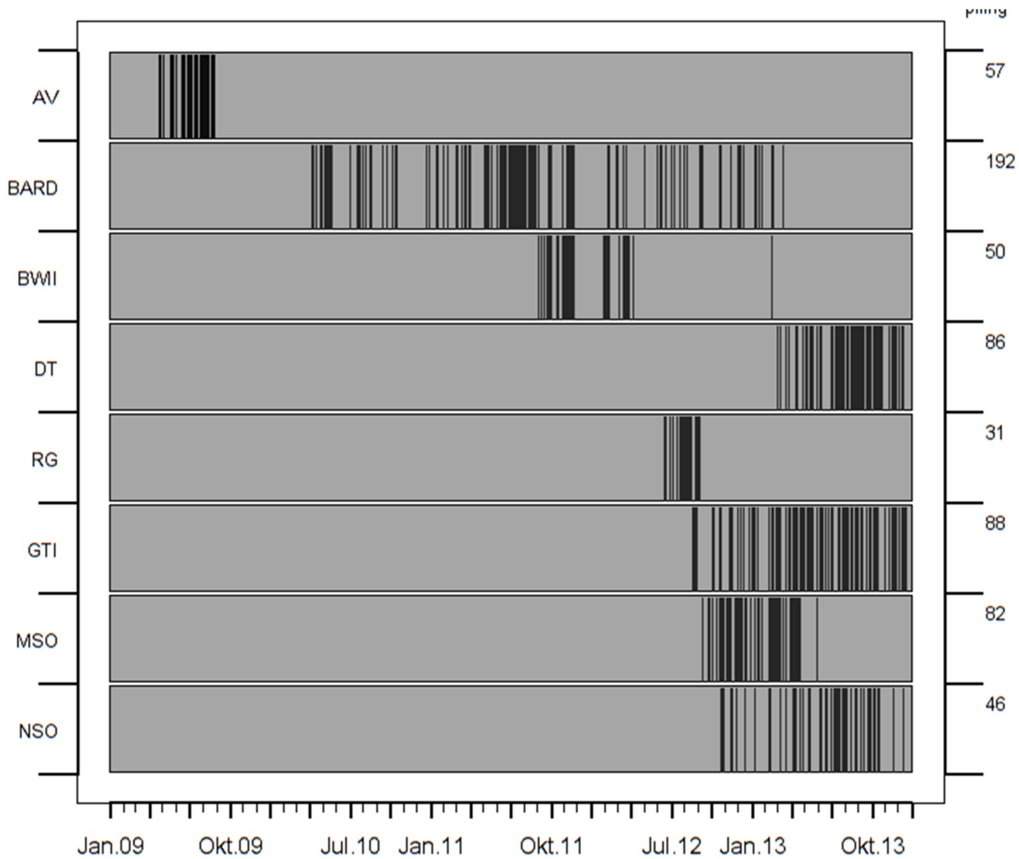


Figure 3-2 Pile driving periods for the eight wind farms that were built during the study period 2009-2013 and were included in this projects analyses.

Regarding foundation types, three wind farms were constructed using monopiles (DT, MSO and RG), three were constructed using tripods (BARD, BWII and GTI) and one was constructed using jackets (NSO). Figure 3-2 shows the piling periods for the wind farms. Construction periods overlap considerably and there are months during which construction took place in up to five wind farms. The dataset therefore includes numerous days when piling activities occurred on the same day. For instance, there were 139 days when construction took place in at least two different wind farms (Table 3-1).

Table 3-1 Number of days with simultaneous construction at up to five different wind farms between 2010 and 2013.

number of simultaneously active wind farm constructions	number of days
1	321
2	108
3	23
4	7
5	1

One piling event was defined as “a period over which piling is taking place without breaks longer than three hours”. For tripod and jacket foundations, several piling events could thus be defined per foundation. Table 3-2 presents the number of foundations erected, the number of piling events and how many piling events were necessary for the construction of one foundation. There were 574 separate piling events for the construction of a total of 434 foundations between 2010 and 2013.

Table 3-2 Number of foundations erected with multiple piling events and total number of piling events between 2010 and 2013. (Note that at GTI and NSO some foundations were constructed in 2014 that were not considered during this study.)

project	foundations with 5 piling events	foundations with 4 piling events	foundations with 3 piling events	foundations with 2 piling events	foundations with 1 piling events	total piling events	total foundations
BARD	3	7	33	14	24	194	81
BWII	0	0	0	10	31	51	41
DT	0	0	1	4	75	86	80
GTI	0	0	0	5	71	81	76
MSO	0	0	0	1	80	82	81
NSO	0	0	1	3	37	46	41
RG	0	0	0	0	30	30	30
total	3	7	35	38	350	574	434

Figure 3-3 gives an overview of how many piling events took place per month over the 2010-2013 study period. It shows that especially during 2013 many piling events occurred, and these often took place at the same day and sometimes even the same time.

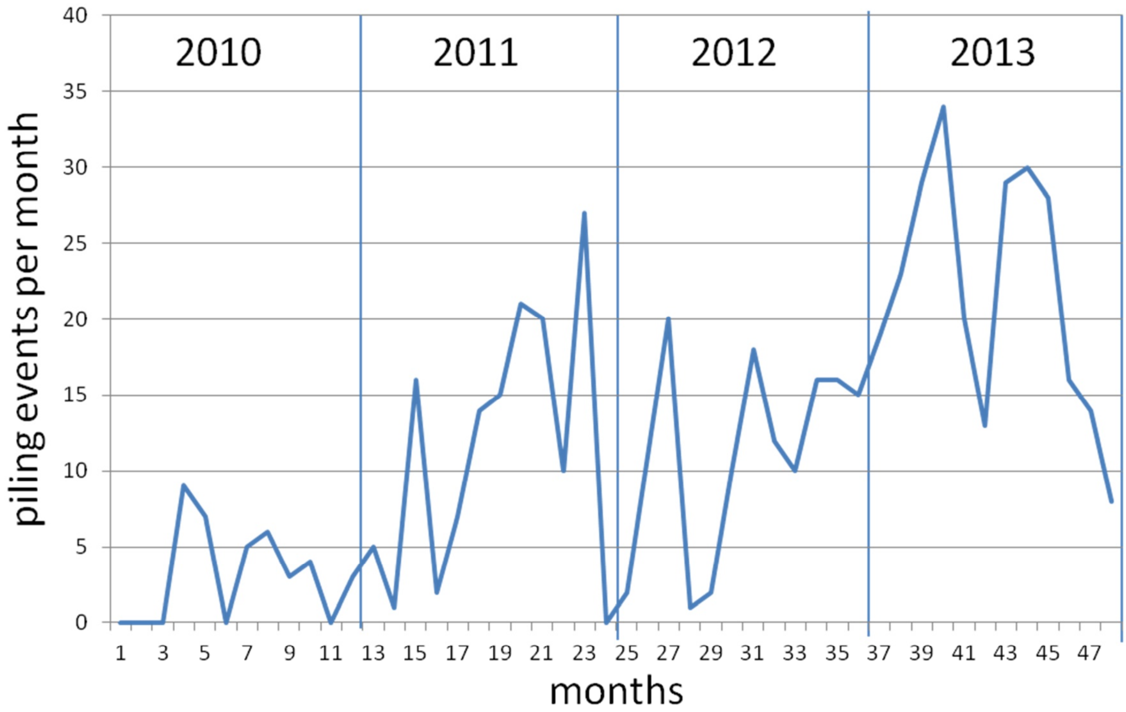


Figure 3-3 Number of piling events per month from 2010 to 2013.Piling activities increased between 2010 and 2013.

There are substantial differences between projects regarding the duration of piling events using monopile, tripod or jacket foundations. The median duration of a piling event ranges between 1.0 and 1.9 h for projects using monopile foundations (DT: 1.9 h, MSO: 1.5 h, RG: 1.0 h), whereas for projects using tripods or jackets the median piling duration ranged between 3.1 and 8.3 h (BARD: 3,1 h, BWII: 5.0 h, GTI: 8.3 h, NSO: 6.2 h, Figure 3-4). Note that the value for piling duration does not mean that there was continuous piling; per definition this could include breaks up to three hours. Furthermore, projects using tripod and jacket foundations were erected using more strokes per piling event than projects using monopiles. During these projects, there is also a much greater variation in piling duration and number of strokes. Previous studies have shown that piling effects can be detected from a few hours up to almost two days (E.G. TOUGAARD et al. 2009; BRANDT et al. 2011; ROSE. et al. 2014). Therefore, 48 hours between piling events have been chosen as an indicator for the number of piling events available to study effects up to a time when normalisation of porpoise activity may be expected. It is to be noted that there is a much greater proportion of piling events separated by over 48 h for the three projects using monopile foundations (DT, NSO and RG) than for the projects using tripod or jacket foundations (Figure 3-5). During projects using tripod and jacket foundations there is also a much greater variation in piling duration and number of strokes.

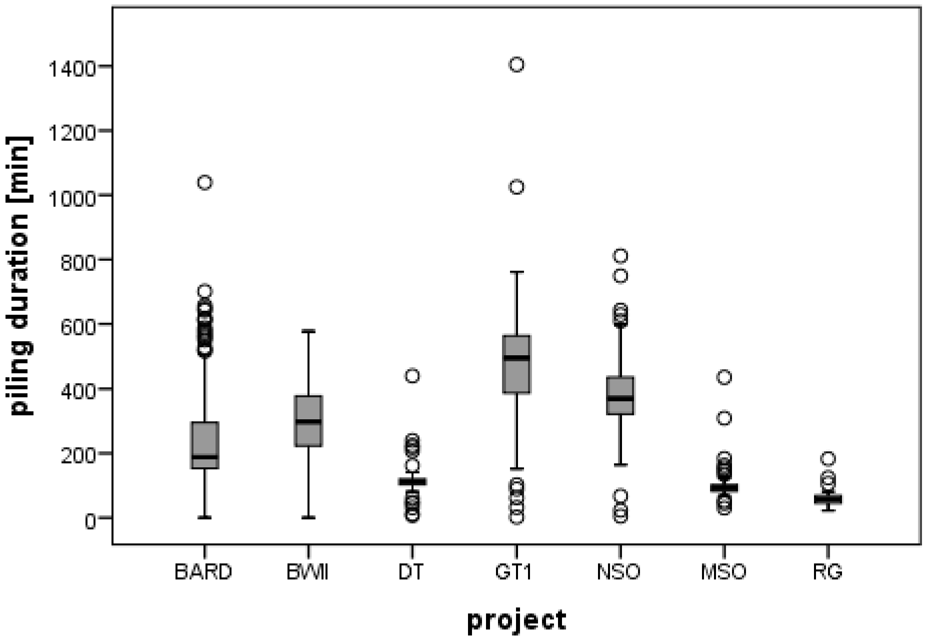


Figure 3-4 Boxplots showing median and quartiles of piling duration in minutes for the seven wind farm projects constructed between 2010 and 2013.

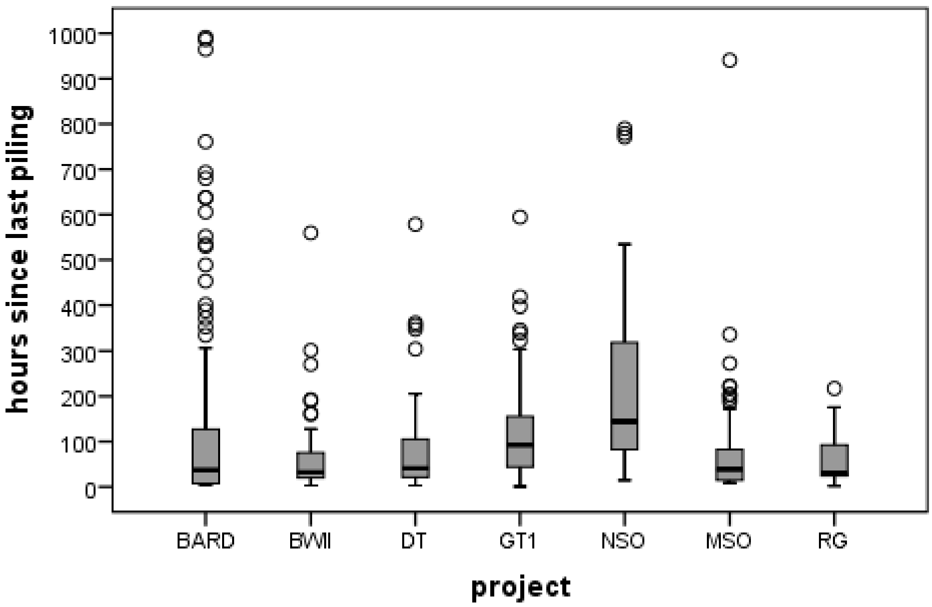


Figure 3-5 Boxplots showing median and quartiles of time between consecutive piling events in hours for the seven wind farm projects constructed between 2010 and 2013.

Noise mitigation characteristics vary between projects. Of all 574 piling events, noise mitigation was applied during 354 piling events (62 %), during 220 piling events (38 %) there was no mitigation. The great majority of piling events without noise mitigation occurred at BARD (190), where mitigation was only applied at four piling events. Apart from BWII, where 14 piling events (28%)

occurred without noise mitigation, during the great majority of piling events at all other projects noise mitigation was applied (Table 3-3).

Table 3-3 Number of foundations and piling events with and without noise mitigation per project and in total between 2010 and 2013.

project	foundations without noise mitigation	foundations with noise mitigation	piling events without noise mitigation	piling events with noise mitigation
BARD	79	2	190	4
BWII	11	30	14	37
DT	2	78	5	81
GTI	2	74	3	78
MSO	2	79	2	80
NSO	1	40	1	45
RG	0	30	0	30
total	97	330	220	354

For each piling event, the following variables were collected: Gross piling time in minutes, number of strokes, cumulative energy used, noise mitigation applied (yes/no) and time between consecutive piling events within one wind farm. It is to be noted that some piling variables correlate with one another. For instance, the number of strokes and the duration of piling correlate strongly.

3.2 Noise data

Direct noise measurements are available at some POD locations for at least two piling events in every wind farm. In contrast to other commonly used noise measurements, Sound Exposure Level (SEL) is not averaged over an a priori defined time interval. This is important as piling noise is an impulsive noise. Noise levels over time would strongly depend on the inter pulse duration and not only from the noise level of the single pulse. SEL is expressed in decibel (dB) and defined as:

$$SEL = 10\log\left(\frac{1}{T_0}\int_{T_1}^{T_2}\frac{p(t)^2}{p_0^2}dt\right)$$

T1 and T2 - start time and end time (the noise event has to lay between start and end time)

T0 - reference value of 1 second

p(t) - temporal varying noise level

p0 - reference noise level (under water: 1 µPa)

For each piling impulse, a single SEL value is available. In order to describe a piling event, percentile levels are given, among them the median, where 50% of values are louder (SEL_{50}). Similarly, SEL_{05} and SEL_{90} are defined as the noise values exceeded during 5% and 90% of values respectively. During this study only SEL_{05}) was used during the following analyses of POD-data.

Figure 3-6 illustrates the measured SEL_{05}- values for all wind farm projects and in presence, partial presence or absence of noise mitigation. The real measurement data do not show a marked difference between noise mitigation and no mitigation when pooled over all wind farm projects, mainly a result of considerable variation between wind farm projects. There also is considerable variance within and between project measurements.

As illustrated in Figure 3-6, noise measurements vary between the seven wind farm projects and are not available for the majority of POD-positions. A noise propagation model was thus used to determine the noise exposure at the distances were PODs were deployed. The used noise propagation model proposed by ITAP GmbH is suited to impulsive pile driving noise (details in Appendix). Noise levels were thus calculated for most piling events without direct noise measurements based on project-specific (or mostly pile-specific) noise measurements and the distance between POD and pile. It is to be noted that this calculation did not consider potential effects of water depth or sediment type. As the environment produces a general background level of noise, e.g. waves, we did not use noise levels of less than 110 dB, because these were masked by background noise and thus not meaningful. Noise data were then merged with the POD-dataset.

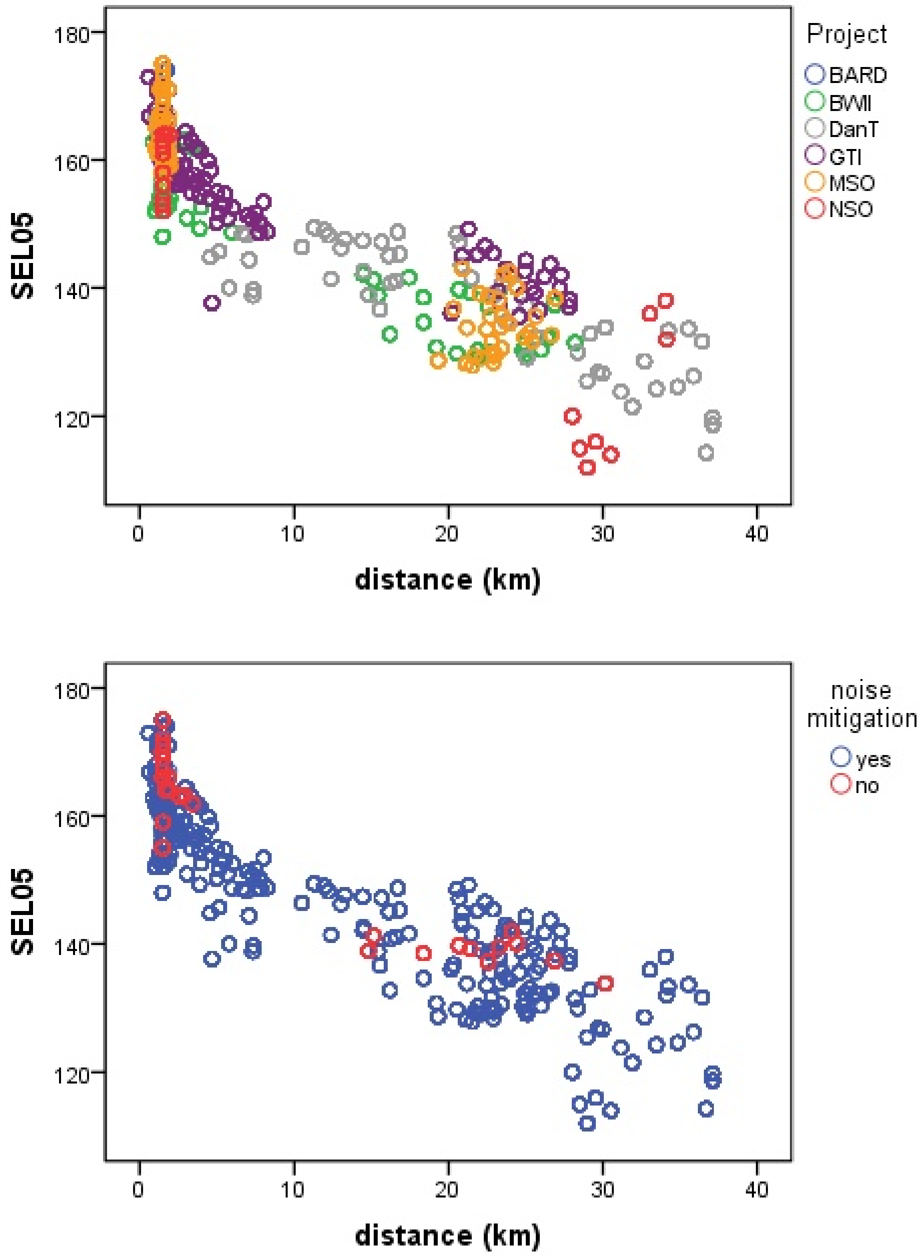


Figure 3-6 *Measured Noise Exposure Level exceeded during 5 % of the time of a piling event (SEL_{05} in dB) versus distance (in km) for six of the seven wind farm projects (no value available for RG). Top: point colours indicate the different wind farm projects. Bottom: point colours indicate the presence (blue) or absence (red) of noise mitigation during piling.*

In comparison to Figure 3-6, Figure 3-7 illustrates all available SEL_{05} levels for all wind farm projects. For a given distance, there is considerable variation in SEL_{05} between projects. For BARD and RG, SEL_{05} values have little variance, which is expected as those data were modelled based on measurements at only two and eight foundations respectively. For the other wind farm projects, measurements existed for almost every foundation. Therefore, it can be expected that noise levels will not reveal more information on the effects of porpoise detections than distances at BARD and RG but probably for the other projects.

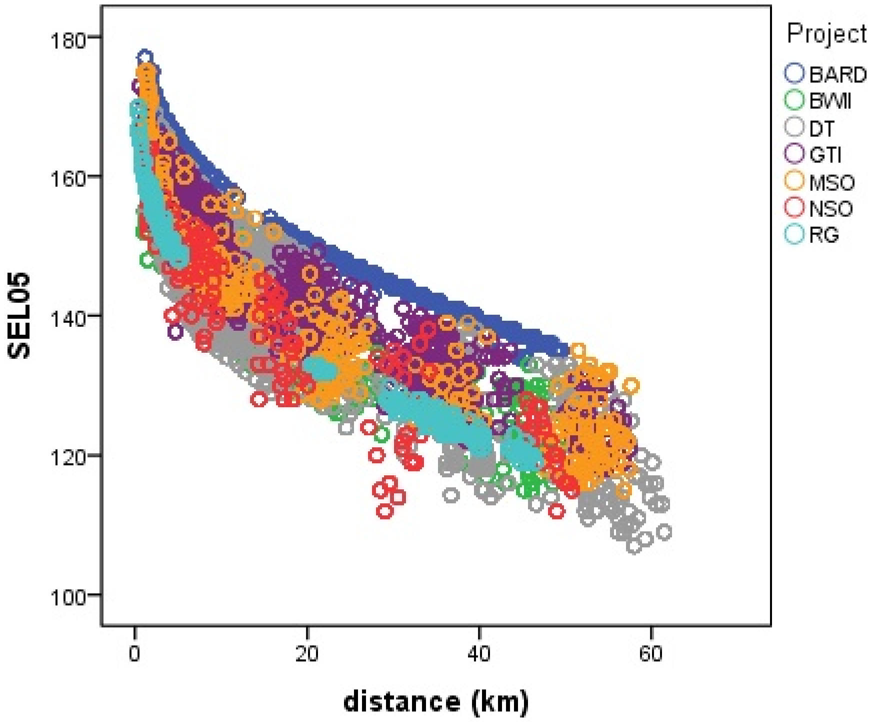


Figure 3-7 *Measured and modelled Noise Exposure Level exceeded during 5 % of time of a piling event (SEL_{05} in dB) versus distance (A_dist in km) for the seven wind farm projects. Note that most values are modelled. Especially, low variation for wind farms Riffgat and BARD originate from the high number of calculated values based on single reference measurements, excluding variability between piling events.*

Figure 3-8 illustrates the noise levels during piling events at 750 m distance and the impact of noise mitigation measures on those per wind farm project. Noise levels were either measured or calculated. For all wind farm projects that applied noise mitigation and also piled some piles without noise mitigation, median SEL_{05} values were lower when noise mitigation was applied than when it was not applied. However, values measured during noise mitigation show considerable variance by up to 20 dB and some values lie within the noise levels that were measured without noise mitigation. High variation in some projects is in part caused by different noise mitigation techniques and also varying effectiveness within the same noise mitigation system causing high variations in noise reduction. There were generally relatively few measurements for piling events without noise mitigation. Most of these stem from BARD, where all but one foundation were piled without noise mitigation. This makes a comparison of effects on porpoise detection between piling events with and without noise mitigation difficult as sample size for the latter is very small and largely based on only one project.

Finally, Table 3-4 summarises the project characteristics in terms of noise level, foundation types, piling duration, number of piling events and number of foundations constructed without noise mitigation measures. For instance, there were 51 separate piling events for the construction of BWII for a total of 41 foundations, of which 12 were not noise mitigated.

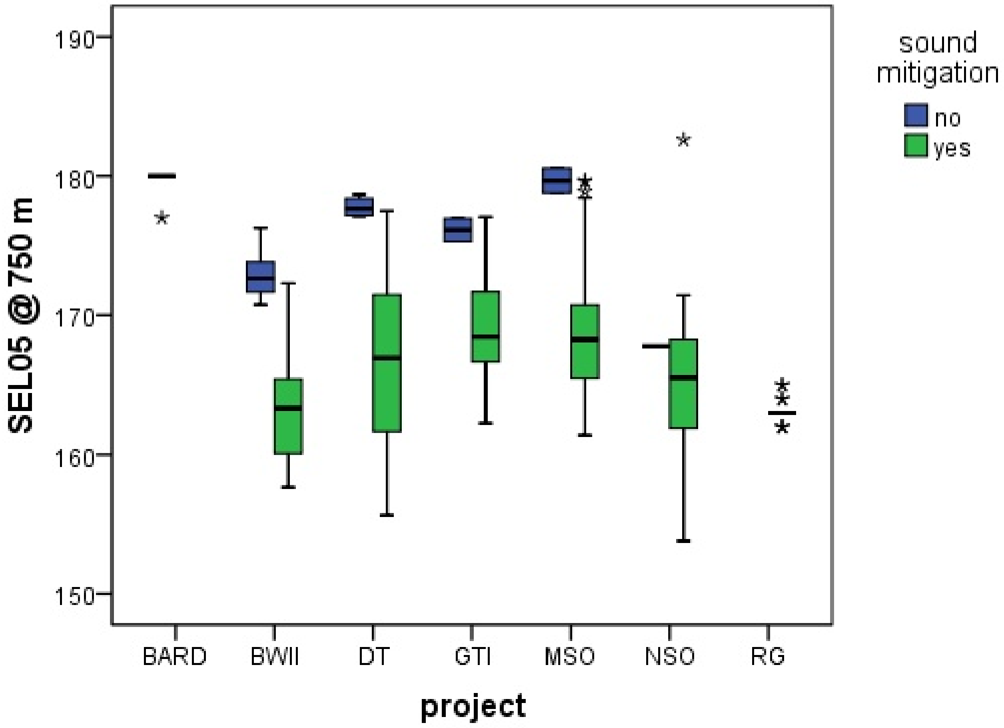


Figure 3-8 *Boxplot of SEL_{05} (in dB) at 750 m distance from piling with noise mitigation (green) and without noise mitigation (blue) per wind farm project. No noise levels with noise mitigation exist at this distance for BARD (where only two foundations were piled with noise mitigation) and no measurements without noise mitigation exist for RG (where noise mitigation was always applied). The bold black line presents the median value, boxplot ranges from 25% to 75% quantiles and whiskers indicate minimal and maximal values without outliers (these are shown as asterisks). Note that most values are modelled.*

Table 3-4 Project-specific characteristics of piling events occurring between 2010 and 2013.

project	SEL_{05}@750 m (mean value)	piling duration (min) (mean value)	founda-tion type	number of foundations (including platforms)	number of piling events	foundations without noise mitigation
BARD	180	310	tripod	81	194	80
BWII	165	300	tripod	41	51	11
DT	167	115	monopile	80	86	2
GTI	169	508	tripod	76	85	2
MSO	169	112	monopile	81	82	2
NSO	165	394	jacket	41	46	1
RG	163	70	monopile	30	30	0

3.3 Environmental data

Noise levels at various distances from the construction site mainly depend on the baseline source level (i.e. local background noise in addition to piling) but are also affected by several other factors. For example, noise propagation in water also depends on sediment, salinity, water depth, topography, sea state and obstructions. How these variables affect noise levels is further frequency-specific. Only environmental variables that were available for the whole study period were considered. These were day length, sea-surface temperature[1], sea-surface temperature anomalies[1], water depth[2], latitude and longitude, sea bed sediment[3], wind speed and wind direction[1]. Sea-surface temperature and sea-surface temperature anomalies were available at daily resolution, wind direction and wind speed at hourly resolution. Considering sea bed sediment, the variable was included as a 5-level factor. It is to be noted that there are few sediment variations in the study area, especially when looking at wind farms individually.

---

[1] *NOAA High Resolution SST data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at http://www.esrl.noaa.gov/psd/*

[2] http://portal.emodnet-hydrography.eu/

[3] *Information contained here has been derived from data that is made available under the European Marine Observation Data Network (EMODnet) Seabed Habitats project (http://www.emodnet-seabedhabitats.eu/ funded by the European Commission's Directorate-General for Maritime Affairs and Fisheries (DG MARE)*

4 HOURLY POD-DATA

4.1 Introduction

In this chapter, we analyse data from passive acoustic monitoring of harbour porpoises at the hourly resolution in order to study fine scale porpoise avoidance behaviour. We analyse the sound levels to which porpoises respond with avoidance behaviour and the range and duration of porpoise responses in relation to offshore wind farm construction activity. We address the following questions: a) At what noise levels do porpoises start to react to piling noise? b) What is the spatial correlation between piling noise emission and displacement effects? c) What is the duration of a displacement effect? d) How do effect ranges differ between piling events with and without applied noise mitigation? e) Is there evidence for temporal cumulative or habituation effects?

4.2 Methods

4.2.1 Data collection

POD-deployment and data processing

In this study, data from three different POD-deployment schemes are used: continuous monitoring positions (POD stations), project-specific stationary PODs (single stationary PODs) and mobile PODs. Although these differ in their settings or deployment design, the same technical device (the C-POD) was used at all of them to record porpoise echolocation clicks. Deployment design differs slightly between locations or responsible company, but the general principle is always the same: a POD is located in the water column 5-10 m above the sea floor. The POD position is fixed at the sea floor with a mooring system and kept in the water column by a buoy.

POD stations consist of three simultaneously deployed PODs in close vicinity. They are located within a square of four marker buoys that indicate the location of the POD station and prevent ships from accidentally crossing this area and causing equipment loss. Simultaneous deployment of multiple PODs at one location accounts for the occasional loss or malfunction of single PODs. POD stations are normally serviced once a month, when memory cards and batteries are exchanged and lost PODs replaced. In case of a noisy environment the memory cards data capacity could be exceeded. To avoid this, a maximum of 4,069 clicks/min was set to be recorded per minute. If that number was reached, the POD did not record for the remaining seconds of that minute. Per POD-station only data from one POD were included for analyses. Here it was always the POD with the most complete time series of recordings that was chosen. Single stationary PODs were deployed for specific wind farm projects and usually consist of only one POD using a similar mooring system and the same POD settings. Mobile PODs were deployed at close distances to the piling location (usually one at 750 m and one at 1500 m distance) for specific piling events with the aim to monitor the effectiveness of deterrence measures). Each specific POD is usually only deployed from a few hours before to a few hours after a specific piling event. For these PODs no scan limit was set due to their short deployment time and the need to maximise detection probability during that time. For this study POD-data were available from 20 POD-stations, 56 single stationary PODs and 49 mobile PODs. Figure 4-1 shows the positions of POD-stations and single stationary PODs. Figure 4-2 and Figure 4-3 illustrate data availability for these PODs. Data from all three types of deployments were merged for analysing the effects of pile driving on porpoise detections at the hourly scale.

POD data were merged with piling-related, noise and environmental variables so that spatial and temporal characteristics for each piling event had to be calculated. For calculating distances between PODs and foundations all three PODs from one POD station were assigned to one geographic location. The same was done for each pile from a single tripod or jacket foundation. The distances between the POD devices of a POD-station and between individual piles of one foundation are negligible in this context.

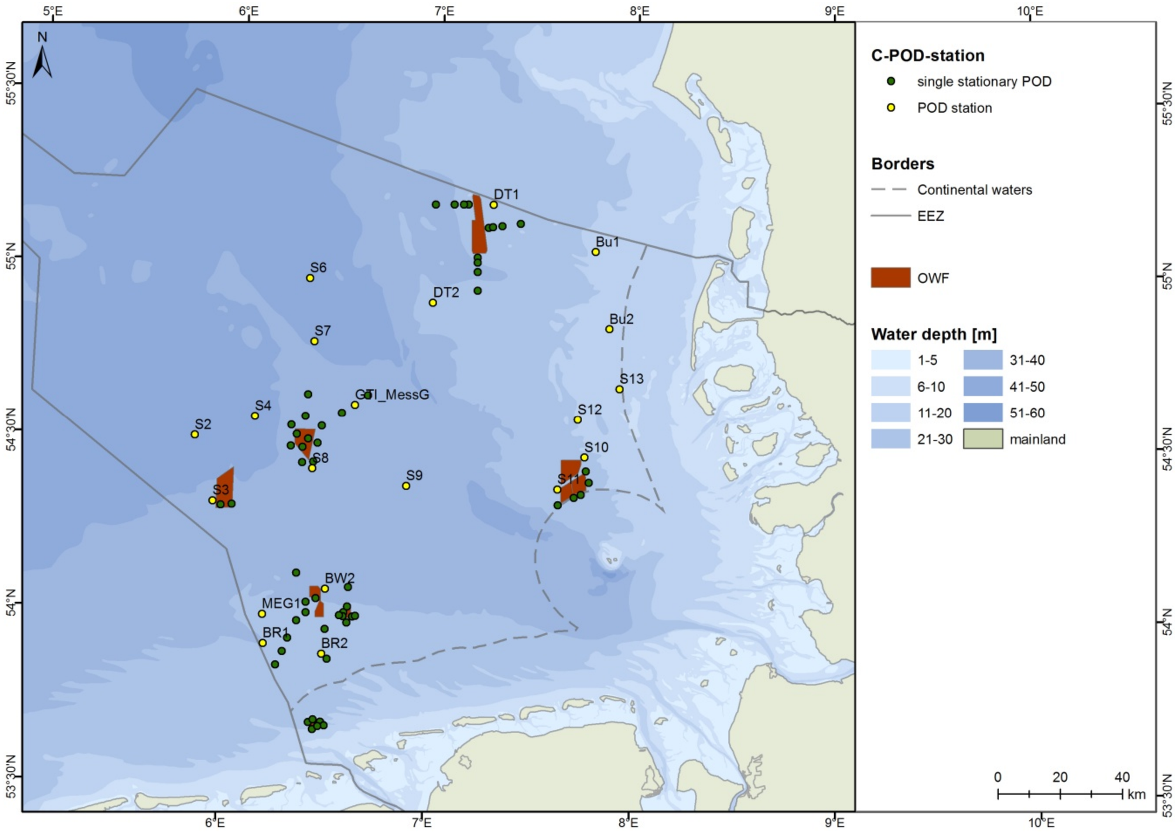


Figure 4-1 Stationary POD-positions from which data are available for this study. POD-stations are depicted as yellow points with name labels, single stationary PODs are depicted as green points. Offshore wind farms constructed between 2010 and 2013 are depicted as red areas.

Figure 4-2 Data availability between January 2010 and December 2013 at 55 single stationary POD-positions that were deployed for specific wind farm projects. Black bars show time periods when PODs recorded data at that position.

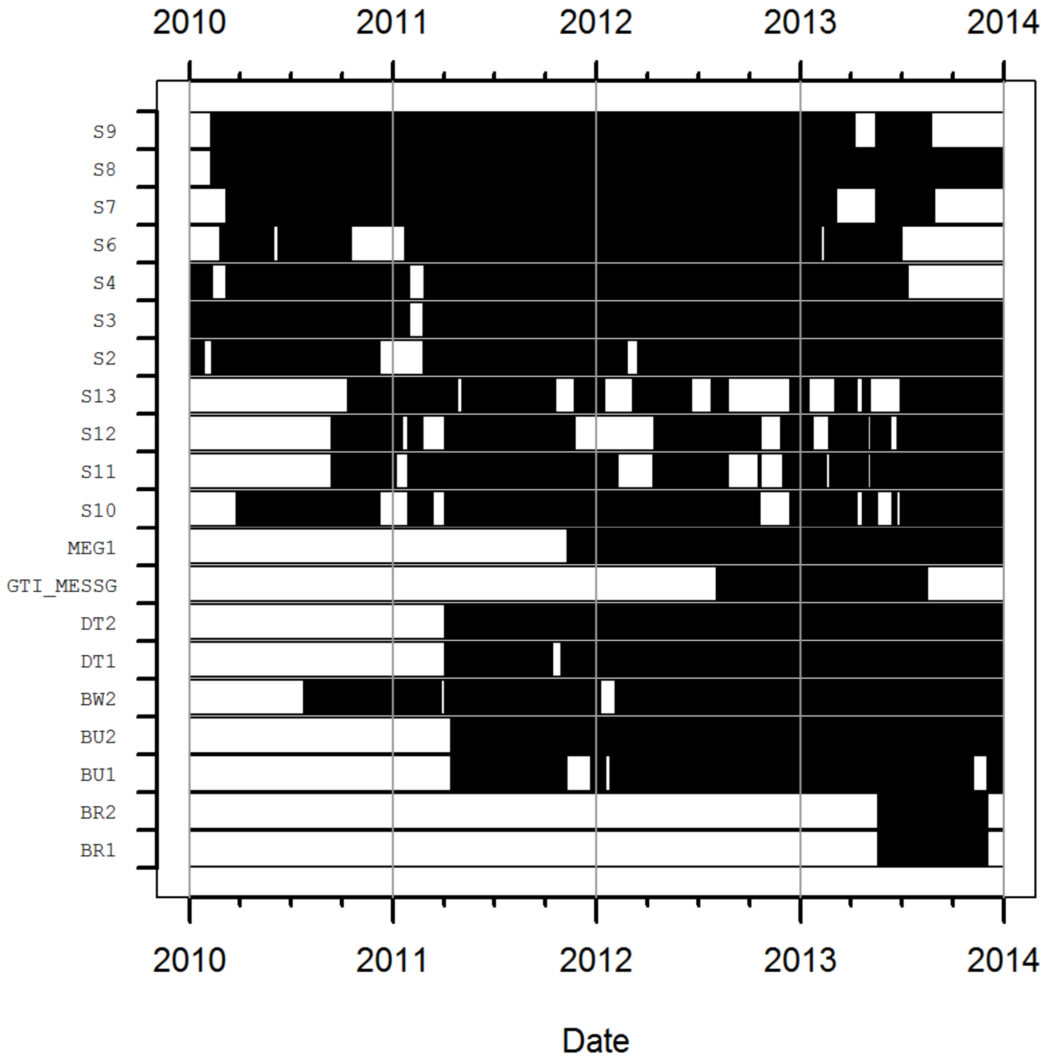


Figure 4-3 *Data availability between January 2010 and December 2013 at the 20 POD-stations. Black bars show time periods when PODs recorded data at that position.*

Data preparation

In order to test the short-term effects of pile driving on porpoise activity at a small spatial scale we used the parameter detection positive hours (DPH) as indicator for porpoise activity. This parameter was used as response variable during the following analyses. DPH describes whether or not a porpoise click-train was recorded and identified during a given hour and is thus a binary variable (with the values 0 or 1). These data were then merged with the environmental information based on geographic and time-related information, of which the ones used within the final models are listed in Table 4-1.

Table 4-1 List of all variables used within the final statistical models of hourly POD-data.

variable	type	description
piling related variables		
SEL_{05}	continuous	noise exposure level exceeded during 5 % of the piling period as measured at or extrapolated to the position of the POD
noise mitigation	factor (3 levels)	noise mitigation applied, not applied or functioning only part of the time
hour relative to piling	continuous	hour related to work (start of a piling event or deterrence) ranging from -48 to 120h
distance	continuous	distance to a piling event in metres
piling duration	continuous	duration of a piling event in minutes
piling order	continuous	consecutive number of a piling event within one wind farm
time related variables		
HH	continuous	hour of the day
day of year	circular and continuous	day of the year
Year	factor (4 levels)	year 2010 to 2013
environmental variables		
wind speed	continuous	wind speed in m/s
sediment	factor (5 levels)	sea bed sediment (1: coarse sand with <20 % mud, 2: medium coarse sand with <20 % mud, 3: medium sand, 4: fine sand with < 20 % mud, 5: fine sand with 21-50 % mud)
wind direction	circular and continuous	wind direction
noise clicks	continuous	number of clicks recorded by the POD in that hour (not including identified porpoise clicks)
SSTA	continuous	sea surface temperature anomaly
POD-Position	factor (many levels)	Position at which a POD was deployed

In order to merge POD data with variables that describe piling characteristics, a decision had to be made on the distances and times for which piling information is merged with POD data. Knowing from previous studies that piling effects (without noise mitigation) on porpoise detections occurred in up to about 20 km distance, we decided to set a precautionary 40 km boundary around each wind farm for considering and assigning POD-data to each wind farm. This was chosen to use a conservative limit in case effects reach further and to also capture distances at which no effect was assumed. Thus, all POD positions that were within a 40-km radius around a wind farm were included in the data subset specific for that wind farm. This means that single piling events could be as far as 60 km from a specific POD-location, as wind farm areas are up to 20 km in diameter. Then, the relative time of each hour to the next piling event within that particular wind farm is determined by counting 48 hours down from the start of deterrence and 48 hours up from the end of piling. Each hour during which deterrence before piling or whilst piling has taken place is counted as 0. Hours are only defined as being before a piling event (-48 to -1h) if at least 48 hours have passed since the end of the last piling event. If hours later than 48 hours after a piling event are not assigned to being before the next piling event, incrementing continues until 120 h after piling. All data outside this time window are excluded from analyses. Then, a separate variable “time relative to piling” was created for each wind farm that is within 60 km of the wind farm in question (so in and where pile driving occurred within this time window). This captures every other piling event at a minimum distance of 20 km. This results in up to three such variables (A-C) for the seven data-subsets.

In order to analyse the effects of specific piling events, we excluded data that were confounded by the effects of several piling events close in space and time. To achieve this, hourly data were excluded if piling took place within another wind farm in a 60 km radius of that POD-position during that hour or up to 24 hours before. Furthermore, we excluded hours when deterrence was active but no piling occurred. If there were times when deterrence took place without a piling event associated to it, these hours and up to 48 hours after were excluded from the dataset. Based on the assignment of a specific hour at a specific location to a particular piling event the other piling variables (noise level, piling duration etc.) were also merged with the POD-dataset.

Dealing with background noise

C-PODs do not only register porpoise clicks but all tonal signals i e. signals that have a characteristic peak within the power spectrum of porpoise clicks. Thus, “clicks” can originate from other sources such as sonar, noise from sediment suspension, surface noise from waves etc. Therefore, the quality of C-POD recordings has to be tested with respect to the effects that a noisy environment may have on the probability of recording porpoise clicks. Two problems emerge from high background noise:

1) In a noisy environment the memory card of a C-POD may quickly fill up. To prevent this, C-PODs can be programmed to contain a recording limit per minute, which means that during one minute only a maximum number of “clicks” is registered. If this click limit within one minute is reached, the POD stops recording for the remainder of this minute. This limits the amount of data that will maximally be stored per minute on the memory card and prevents the card from an overflow of data, which would result in no more data until the next recovery. After the click limit is reached nothing will be recorded for the remainder of that particular minute. If not controlled for this issue would lead to an underestimation of porpoise activity. The click limit for stationary PODs was set to be 4096 clicks per min, for mobile PODs no scan limit was set, as these PODs were only deployed for a couple of days at maximum and the memory card was unlikely to be filled up during this short time interval.

2) Substantial noise also affects the performance of the detection algorithm of the C-POD.exe software, as porpoise clicks will be harder to distinguish from background noise when noise is substantial (a phenomenon called masking). Thus, the likelihood that the algorithm identifies porpoise clicks during the recorded time interval decreases with increasing amount of background noise. This will then result in an underestimation of porpoise activity if background noise is not controlled for.

We addressed these issues by visually exploring 1) the relationship between porpoise detections and the number of minutes during an hour, when the scan limit was reached, and 2) between porpoise detections and the number of all clicks other than porpoise clicks that were recorded during that hour. Based on these relationships data with more than 100,000 clicks per hour and more than 2 min per hour when the scan limit was reached were excluded. This led to 10.7 % data exclusion. However, there remained a negative relationship between DPH and all clicks recorded but further data reduction seemed too drastic. We therefore always included the variable "noise clicks" (all clicks without identified porpoise clicks) into each model to control for its effect.

4.2.2 Statistical analyses

Explanatory variables and collinearity

Including variables with high collinearity into the final model should be avoided, as this can affect results of individual predictors, which is what we are interested in. Therefore, we examined collinearity of all continuous predictor variables in order to test which variables may not be used jointly in the final model. Variables with a correlation higher than r=0.5 were not used jointly in the same model. Variables highly correlated were day of year and sea surface temperature as well as piling duration, cumulative energy applied and number of strokes. Therefore, we only considered the effects of day of year and piling duration. Furthermore, due to the use of position as a random factor, static variables like water depth, latitude and longitude were not found to improve the global model and were therefore not considered further. A list explaining all the variables available for modelling the effects on porpoise detections that were used within this study can be seen in Table 4-1.

Screening for temporal autocorrelation

Dealing with biological processes, we expected the input (environmental covariates) and output (residuals) time series of statistical models to display temporal autocorrelation. Considering the model residuals, previous investigations showed that significant autocorrelation originated from the DPH response variable and not from environmental covariates (see details in WP2).

Considering the statistical model definition, we investigated different ways of taking autocorrelation into account and determined the most parsimonious autocorrelation patterns to be taken into account in further analyses. We specifically investigated the differences between available options within the gam function of mgcv package in R (MGCV 2015) and the definition of a differenced covariate (DPH at t-1) acting as an auto-regressive component of the first order (BESTLEY et al. 2010). The selection of the optimal model is based on the Akaike Information Criterion (AIC, AKAIKE 1974) and on a graphical investigation of the autocorrelation (ACF) and partial autocorrelation (PACF) functions of model residuals.

Based on the outcome of these analyses, we decided to use the differenced DPH(t-1) covariate as a factor within our GAM analyses. This covariate was found to significantly reduce the autocorrelation pattern in the global dataset as well as in the seven wind farm project-specific datasets and also allowed the use of the bam function, which has a faster computing time and is more flexible for statistical analyses of large datasets than the gam function (MGCV 2015). Therefore, we used the bam function from the mgcv package in R for all the following GAM analyses.

Random effect selection

Preliminary analyses were conducted to compare the inclusion of several random effects in terms of model goodness-of-fit and model outputs. POD position, POD campaign, POD ID and an interaction term between longitude and latitude were alternatively included in the model formula on the global dataset (details in WP2).

Based on the outcome of these analyses, we decided to include the POD position as a random effect within the following GAM analyses. This covariate was found (i) to improve the deviance explained by the model and decrease model AIC and (ii) to take into account the geographical location, hence geographical-related characteristics.

GAM model specifications

Following preliminary analyses on variable selection and model output investigations, we decided to use Detection Positive Hours (DPH) as a response variable and include smooth and factor covariates in the model. Day of year, hour of day (HH), wind speed, wind direction, noise clicks, sea surface temperature anomaly (SSTA) and piling duration were included as continuous smooth functions. Year and sediment category were included as factors. DPH(t-1) was included as a factor to correct for temporal autocorrelation. Position was included as a random effect. Hour relative to piling with distance from piling were included as an interaction term also specified as a smooth function. This gave the global model. As we also ran several different models on subsets of data according to the question to be addressed, these specifications had to be changed slightly within these specific models. This will be mentioned in the results part.

Statistical models always represent a compromise between model accuracy and data availability. The aim of the present study was to conduct a global analysis of all available data in order to yield global estimates on several different aspects of harbour porpoise avoidance behaviour during offshore pile driving. Global models were calculated but it also became apparent that in order to look at some specific aspects (e.g. avoidance radii in terms of distance, habituation, etc.) it is necessary to consider each wind farm separately, something that cannot be done within one global model. Ideally, there would be separate models for each season and wind farm to account for all these specific conditions that are likely to change the way porpoises respond to piling noise emission. This, however, would lead to data availability being too low to allow for meaningful analyses. Therefore, we chose two different approaches: First we combined all data from all seven wind farms to investigate general patterns of wind farm construction effects resulting in several models specifically aimed at addressing particular questions raised in the introduction. These models are called “global models”. Results of these models have to be seen as an average effect of all piling events in the German North Sea between 2010 and 2013, while geographic position and season were controlled for as much as possible within these models. Secondly, we ran “project-specific models” in order to look at project-specific differences, which are likely to occur due to different natural patterns (e.g. different prevailing porpoise densities and habitat usage) as well as different construction patterns (e.g. foundation type, piling duration). These project-specific models follow specification of a global model but are run separately for each wind farm dataset.

In total we ran nine different global models numbered from 1 to 9 and given the Prefix G (global) resulting in models G1 to G9. Table 4-2 provides an overview of the primary aim of each of these models and gives the chapter where results of these models are presented. It also provides the Table number were specifications of these models are shown. Of these nine models five were also run specifically for each wind farm project (unless data availability was not sufficient as in the case of MSO and RG for P2), resulting in a total of 33 models. These are given the prefix "P" (project-specific) and numbered according to the global models (1-9). When referring to a particular project-specific model, the wind farm in question is added as a further prefix (e.g. BWII_P1 referring to the project-specific model for BWII, which follows the structure of model G1). Table 4-2 indicates which global models were also run project-specifically. Aims and structure of all project-specific models follow that of the global model with the same number. However, due to differing data availability between wind farm projects, some variables could not be included within some of these project-specific models (e.g. sediment could not be considered for P1_GT1).

Table 4-3 provides an overview on the specifications of the four final "global" models (G1-G4), run on the "global" dataset (i.e. combined data of all seven wind farm projects), that were calculated to look at the range (in terms of sound and distance) and duration of piling effects on porpoise detections and how these may be altered by the use of noise mitigation.

Table 4-4 provides an overview on the specifications of the additional five "global" models (G5-G9), run to look at the effects of habituation, cumulative effects, effects before piling and the role of wind speed in altering piling effects.

As showing all model specifications for all 30 project-specific models would be extensive, we only show specifications for model P1 (Table 4-5).

Non-parametric test design

In addition to GAM analyses we also analysed porpoise detection rates with respect to piling effects using non-parametric tests. This allows to directly link detections during piling to a given baseline period specifically for each distance/noise level class. The disadvantage is that it cannot control for as many potentially confounding factors as the GAM, but has the advantage that patterns are not potentially blurred by smoothing functions and that analyses have less problems with data gaps at certain distance classes. However, this kind of statistics requires to create data classes, and thus it is not possible to obtain one specific value for when changes occur but only a given range. Whether or not significant effects can be demonstrated using non-parametric statistics highly depends on data availability as well as the general height of detection rates during the baseline and their natural variability. Rather than focusing on where significant effects are still found we therefore focus on where significant effects with a decline by at least 20 % during piling relative to the baseline occurred. This definition of 20 % decline was set as a subjective criterion based on the fact that changes in detection rates within 20 % often occur naturally without the effect of anthropogenic impacts.

The noise and distances classes we formed were based on obtaining roughly equal sample sizes within these classes and were as small as data availability allowed. Detection rates during piling were compared to detection rates 25-48 h before piling within the same class and Mann-Whitney

U tests were applied to check for significant differences between them. Results with significance levels <0.05 were considered to be significant

Table 4-2 Overview of the twelve global statistical models with respect to primary aim, localisation of results and specifications.

global model number	primary aim of the model	chapter	also run project-specifically?	model specifica-tions
G1	testing the effects of piling sound level and time to piling	4.3.1	n	Table 4-3
G2	testing the effects of distance and time to piling	4.3.2	y	Table 4-3
G3	testing the effects of noise mitigation on effect ranges	4.3.3	n	Table 4-3
G4	testing the effects of time to piling at distances < 2 km	4.3.4/4.3.6	y	Table 4-3
G5	testing for habituation and cumulative effects (piling duration)	4.3.6	y	Table 4-4
G6	testing for habituation and cumulative effects (consecutive piling number)	4.3.6	y	Table 4-4
G7	verifying effects before piling	4.3.7	y	Table 4-4
G8	investigating the importance of wind speed on decreases before piling	4.3.7	n	Table 4-4
G9	investigating the importance of wind speed on effect ranges during piling	4.3.7	n	Table 4-4

4.3 Results

In order to describe the spatio-temporal response of harbour porpoises to pile driving noise, we ran different models depending on the question to be addressed (see Table 4-2). Model outputs are organised in the different subsections corresponding to the questions outlined in the introduction. Model structure and results on significance levels of each variable for models G1-G4 that were run to look at the range (in terms of sound and distance) and duration of piling effects on porpoise detections and how these may be altered by the use of noise mitigation are presented in Table 4-3. Outputs of these models are shown in chapter 4.3.1 to 4.3.4. Table 4-4 provides an overview on the specifications and results of the additional models G5-G9, run to look at the effects of habituation, cumulative effects, effects before piling and the role of wind speed in altering piling effects. Outputs of these models are shown in chapters 4.3.6 and 4.3.7.

As explained in the methods and in Table 4-2, some of these models were also run project-specifically. Those looking at project-specific effect ranges and duration (P2 and 4) are shown in chapter 4.3.5, those looking at habituation, cumulative effects and effects before piling (P5-P7) are presented in chapters 4.3.6 and 4.3.7 together with the global models G5-7. Table 4-5 gives specifications and results of model P2 as an example for project-specific models.

*Table 4-3 Summary of specifications and results of the four global statistical models G1-G4. If a model was run on a data-subset, this is specified in the second row. For each variable, inclusion and significance levels are indicated ("-": variable not included in the model, "***": p<0.001,"**": p<0.01, "*": p<0.05 and "ns": not significant). Results of the variables of primary interest within each specific model are highlighted as grey cells.*

variable	G1	G2	G3	G4
data included	all data	all data	all data	distance <2 km
DPH(t-1) (factor)	***	***	***	ns
POD-Position (random factor)	***	***	***	***
year (factor)	***	***	***	***
day of year (smooth)	***	***	***	***
HH (smooth)	***	***	***	***
wind speed (smooth)	***	***	***	***
wind direction (smooth)	***	***	***	ns
SSTA (smooth)	***	***	***	***
noise clicks (smooth)	***	***	***	***
sediment (factor)	***	***	***	ns
hour relative to piling (smooth)	-	-	-	***
distance (smooth)	-	-	-	-
SEL_{05} (smooth)	-	-	-	-
piling duration (smooth)	***	***	***	***
hour relative to piling, distance (interaction)	-	***	-	-
hour relative to piling, SEL_{05} (interaction)	***	-	-	-
hour relative to piling, distance, for noise mitigation=no (interaction)	-	-	***	-
hour relative to piling, distance, for noise mitigation=yes (interaction)	-	-	***	-
piling order (smooth)	-	-	-	-
min since last piling (smooth)	-	-	-	-
distance, wind speed (interaction)	-	-	-	-
deviance explained	6.8 %	7.4 %	7.5 %	16.8 %

*Table 4-4 Summary of specifications and results of the five global models G5-G9 If a model was run on a data-subset, this is specified in the second row. For each variable, inclusion and significance levels are indicated ("-": variable not included in the model, "***": p<0.001,"**": p<0.01, "*": p<0.05 and "ns": not significant). Results of the variables of primary interest within each specific model are highlighted as grey cells.*

variable	G5	G6	G7	G8	G9
data included	hour relative to piling=1, distance <5 km	hour relative to piling =0, distance <5 km, piling_no < 100, min since last piling < 20,000 min	hour relative to piling <0	hour relative to piling =-5	hour relative to piling =0
DPH(t-1) (factor)	ns	ns	***	***	***
POD-Position (random factor)	**	***	***	***	***
year (factor)	ns	*	***	***	***
day of year (smooth)	***	***	***	***	***
HH (smooth)	ns	*	***	***	***
wind speed (smooth)	***	***	***	-	-
wind direction (smooth)	ns	ns	**	***	*
SSTA (smooth)	ns	*	***	***	***
noise clicks (smooth)	ns	*	***	***	***
sediment (factor)	ns	ns	***	***	***
hour relative to piling (smooth)	-	-	-	-	-
distance (smooth)	***	***	-	-	-
SEL_{05} (smooth)	-	-	-	-	-
piling duration (smooth)	ns	**	***	-	-
hour relative to piling, distance (interaction)	-	-	***	-	-
hour relative to piling, SEL_{05} (interaction)	-	-	-	-	-
hour relative to piling, distance, for noise mitigation=no (interaction)	-	-	-	-	-
hour relative to piling, distance, for noise mitigation=yes (interaction)	-	-	-	-	-
piling order (smooth)	-	**	-	-	-
min since last piling (smooth)	-	**	-	-	-
distance, wind speed (interaction)	-	-	-	***	***
deviance explained	15.2 %	18.4 %	7.1 %	7.8 %	13.8 %

*Table 4-5 Specifications for the model P2 run following model G1 but for the six specific wind farm projects (all apart from RG). For each variable, inclusion and significance level are indicated ("-": variable not included in the model, "***": p<0.001,"**": p<0.01, "*": p<0.05 and "n.s.": not significant). Results of the variables of primary interest within each specific model are highlighted as grey cells.*

variable	P2_ BARD	P2_ BWII	P2_ DT	P2_ GTI	P2_ NSO
data included	OFW= BARD	OFW= BWII	OFW= DT	OFW= GT	OFW= NSO
DPH(t-1) (factor)	***	*	ns	***	**
POD-Position (random factor)	***	***	***	***	***
year (factor)	-	-	-	-	-
day of year (smooth)	***	***	***	***	***
HH (smooth)	***	***	***	***	***
wind speed (smooth)	***	***	***	***	***
wind direction (smooth)	***	***	***	***	***
SSTA (smooth)	***	***	***	***	**
noise clicks (smooth)	***	***	***	***	***
sediment (factor)	ns	***	***	-	ns
hour relative to piling (smooth)	-	-	-	-	-
distance (smooth)	-	-	-	-	-
SEL_{05} (smooth)	-	-	-	-	-
pilingduration (smooth)	***	***	***	***	***
hour relative to piling, distance (interaction)	***	***	***	***	***
hour relative to piling, SEL_{05} (interaction)	-	-	-	-	-
hour relative to piling, distance, for noise mitigation=no (interaction)	-	-	-	-	-
hour relative to piling, distance, for noise mitigation=yes (interaction)	-	-	-	-	-
piling order (smooth)	-	-	-	-	-
min since last piling (smooth)	-	-	-	-	-
distance, wind speed (interaction)	-	-	-	-	-
deviance explained	8.8 %	8.4 %	7.8 %	10.1 %	5.4 %

4.3.1 Effects of noise levels

Modelling the effect of hour related to piling and SEL_{05}

Model G1 explains 6.8 % of deviance and the interaction of hour relative to piling with distance was highly significant (Table 4-3, Table A-1). As seen from the model output showing the deviation of DPH from the overall mean (Figure 4-4), DPH declined before piling, was lowest during the hour when piling occurred and then increased again afterwards. The lowest noise level when DPH reached the overall average during piling, so when porpoise detections were no longer different when compared to the overall average of all data, was at 143 dB SEL_{05}. As a rather surprising result, porpoise activity at the closest distances to piling started to decline already several hours before the start of piling, but this will be looked at in more detail in the next subchapter. Histograms above and to the right of Figure 4-4 indicate data availability within the different hours relative to piling and noise classes respectively. There are relatively few data available for hours before piling than during and after piling due to piling events often occurring within 48 hours of each other and our definition of baseline data requiring 48 hours to have passed before an hour can be defined as being before a piling event. Furthermore, data availability below 120 dB is also rather low. Because noise levels were not available for all POD-positions and piling events, the dataset looking at the effects of distance rather than SEL was larger and modelling effects of distance rather than sound yielded more robust results. Therefore, these models will be used during the following GAM analyses when looking at project-specific effects, effect duration and habituation effects.

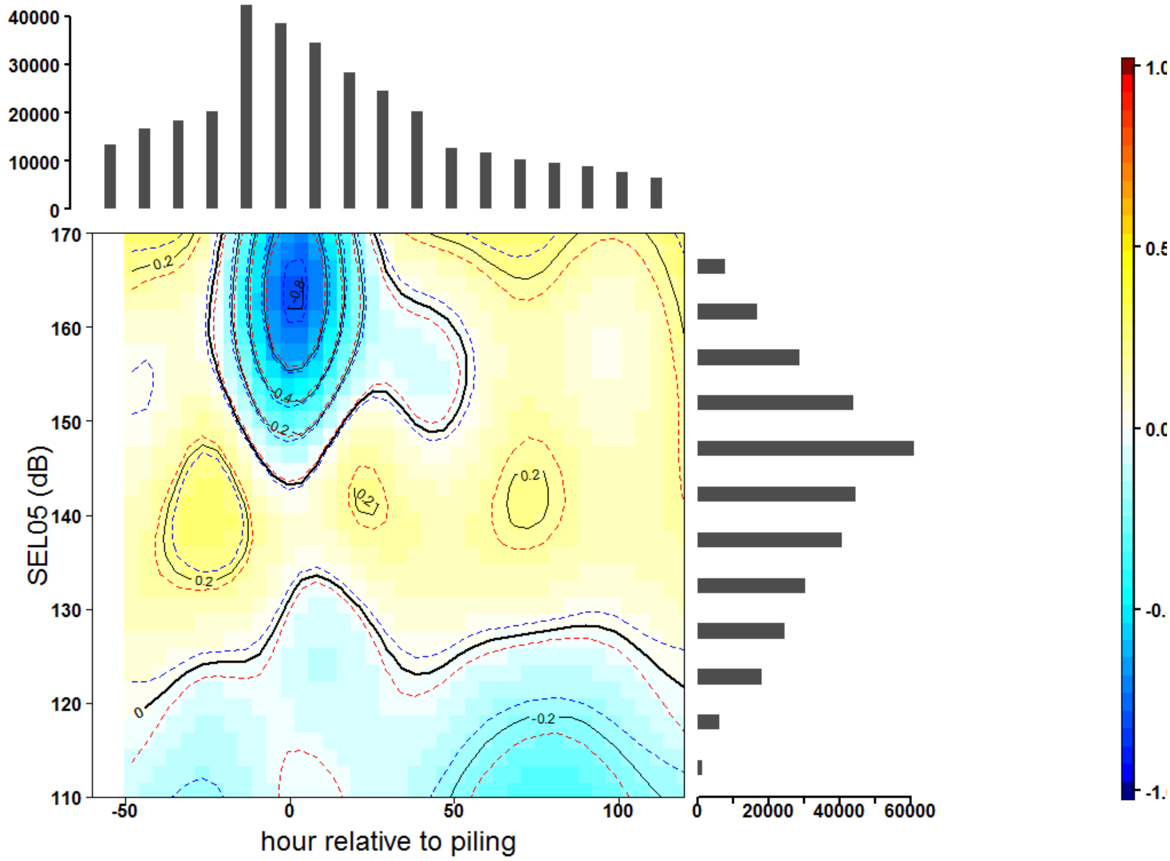


Figure 4-4 *Output from Model G1 showing the effects of the interaction of hour relative to piling (Hours Related to piling Work) with SEL_{05} on DPH. Shown is the deviance of DPH from the global mean (bold 0-line) with cold colours indicating a negative deviation and warm colours a positive one. Histograms indicate data availability at the different hours (-50 to -40h, -40 to -30h etc.) and SEL_{05}-classes (100 to 105dB, 105 to 110dB etc.).*

Investigating the effects of SEL_{05} using non-parametric tests

In order to test for the effects of noise level on porpoise detections, data were divided into different noise classes of 5 or 10 dB difference and DPH during piling was related to DPH during a baseline period (25-48 h before piling) to demonstrate absolute and relative changes in detection rates (Table 4-6). A statistically significant decrease of DPH during piling by at least 20 % occurred down to the noise class 145-150 dB SEL_{05}. Decreases were still significant at lower noise levels, but with a considerably smaller decrease in detection rates. Between 135 and 140 dB and between 140 and 145 dB there was a significant decrease by 14 %. No significant effect was detected at 130-135 dB. At 120-130 dB there was again a significant difference, but the decrease was only 8 %. No significant effect occurred below 120 dB (Table 4-6, Figure 4-7).

*Table 4-6 Average values for DPH calculated per project for nine different sound classes and two time classes Also given is the percentage by how much DPH declined at the hour of piling relative to 25-48 h before the start of piling and significance levels from MANN-Whitney U test for these differences (***: p<0.001, **: p<0.01, *: p<0.05, ns: p>0.05).*

pro-ject	time period	<120 dB	120-130 dB	130-135 dB	135-140 dB	140-145 dB	145-150 dB	150-160 dB	160-170 dB	>170 dB
all	hour relative to piling= -48 to -25	0.56	0.52	0.56	0.58	0.50	0.44	0.46	0.41	0.44
	hour relative to piling=0	0.58	0.48	0.43	0.50	0.43	0.33	0.24	0.09	0.03
	%decline	+ 4 % ns	8 % **	23 % ns	14 % ***	14 % ***	25 % ***	48 % ***	78 % ***	93 % ***

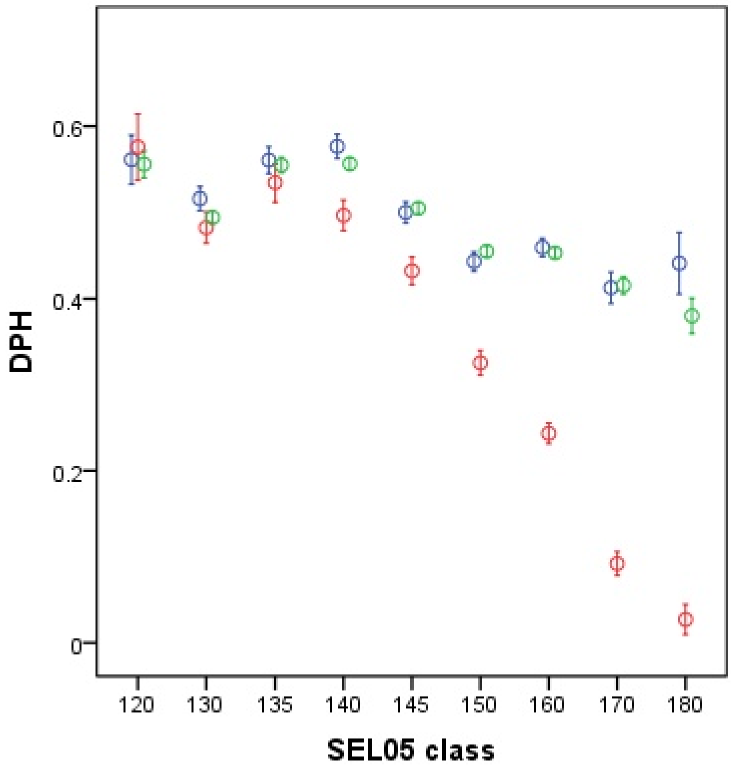


Figure 4-5 Error bars depicting mean and 95% confidence intervals of DPH at the different SEL $_{05}$ classes in dB (120:<120, 130:120-130, 135:135-140, etc.) and for different time classes relative to piling (blue: 25-48 h before piling, red: during piling, green: 25-48 h after piling) for all data combined.

4.3.2 Spatial range and duration of piling effects

Investigating the interaction of hour relative to piling with distance

Model G2 explains 7.4 % of deviance and the interaction of hour relative to piling with distance was highly significant (Table 4-3, Table A-1). As a rather surprising result, porpoise activity at the closest distances to piling started to decline already several hours before the start of piling. DPH remained at the overall average until about 24 h before piling and reached a minimum during the hours of piling (hour relative to piling=0). After piling, DPH increased until reaching the overall average at about 36 h after piling (Figure 4-6). The duration and magnitude of this effect decreased with distance from piling. As to be expected, effect duration was longest at the closest distance and appears to be present only during the hours of piling at about 17 km distance. Seventeen km is the largest distance where a decrease in DPH is still clearly present when compared to times before and after piling and where the overall average was reached. A smaller decrease during piling may still be present in up to 38 km when compared to times before and after piling, but this change is only minimal and less pronounced and cannot be confirmed statistically. It can also be seen that the change in DPH during the hour of piling was greatest at the closest distance and less pronounced with increasing distance from piling.

Histograms above and to the right of Figure 4-6 indicate data availability within the different hours and distance classes. There are less data available for hours before piling than during and after piling due to piling events often occurring within 48 h of each other and our definition of baseline data requiring 48 h to have passed before an hour can be defined as being before a piling event.

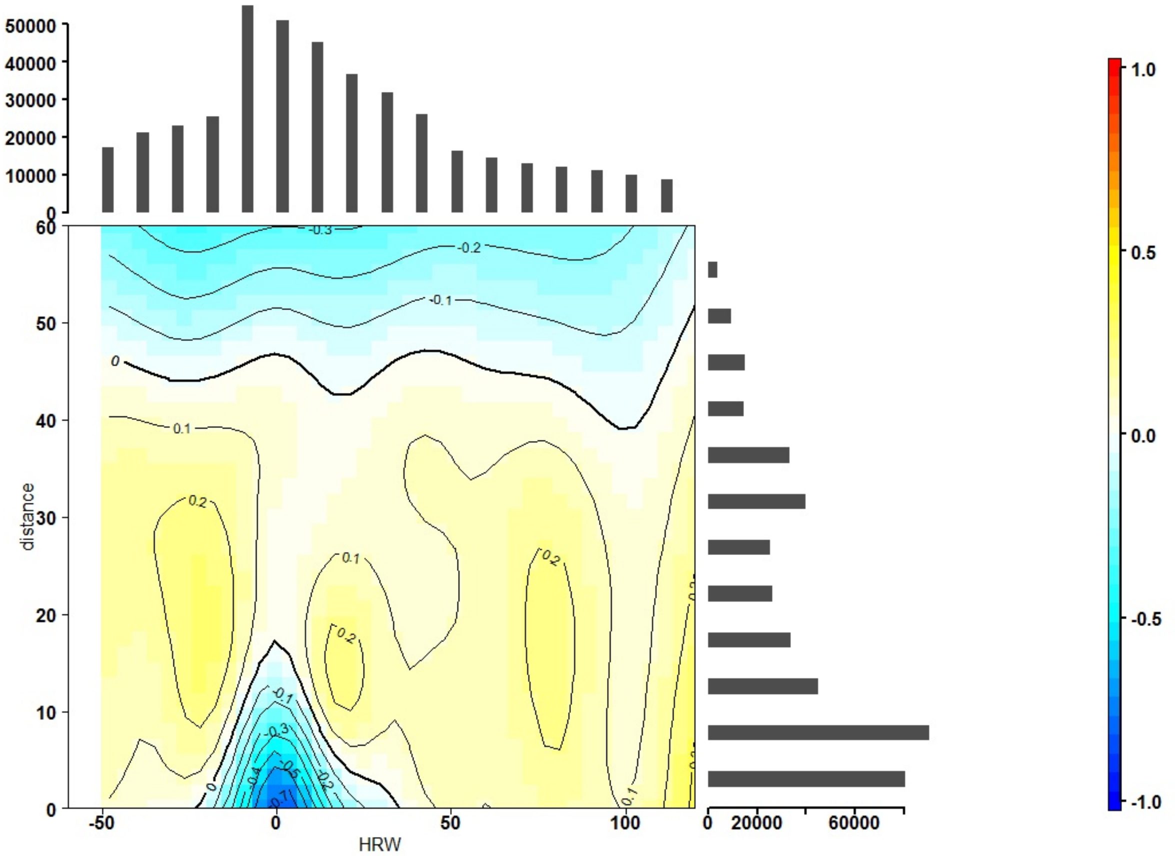


Figure 4-6 *Model output from Model G2 showing the effects of the interaction of hours relative to piling (-48h to +120h) with distance to the piling location on DPH. Shown is the deviance of DPH from the global mean (bold 0-line) with cold colours indicating a negative deviation and warm colours a positive one. Histograms indicate data availability at the different hour classes (-50 to -40h, -40 to -30h etc.) and distance-classes (0 to 5km, 5 to 10km etc.) respectively.*

Investigating effect ranges using non-parametric statistics

Figure 4-7 to Figure 4-9 present the raw data for DPH at the different hours related to piling and for three different distance classes. This proves the effects before piling do not result from a smoothing function within the GAM model. A clear decrease before piling can also be seen within distance categories 0-5 km and 5-10 km (Figure 4-7, Figure 4-8). During piling the decrease at 5-10 km is no longer as strong as in 0-5 km distance. At the distance category 30-40 km (Figure 4-9), a reduction in DPH seems no longer visible. Figures for intermediate distance classes (10-30 km) are shown in the appendix (Figure A-2-Figure A-4). Here the effect is mainly seen during the hours of piling if at all present. The reduced confidence intervals for DPH during the hours of piling relative to other hours are due to larger sample size for hour relative to piling=0 (as piling events often lasted longer than one hour).

Model outputs help to relate data around piling to generally prevailing patterns and take into account the effects of several other variables. However, comparing model outputs to raw data plots shows that smoothing functions can blur detailed patterns visible within the raw data, especially when it comes to the hour of piling. Within the raw data, one can see a gradual decrease in DPH before piling with a sudden further decrease during the hour of piling. After piling, a steep increase occurs that is more pronounced than the decrease before piling. This leads to the model overestimating DPH values predicted for the hour of piling, which has to be kept in mind when interpreting model outputs.

For additional information, we therefore included raw data plots as well as a summarising figure and table detailing average DPH values at different times before, during and after piling, separated for different distance classes as well as results from non-parametric tests (Table 4-7, Figure 4-10). Comparison of average DPH values during piling to values more 25-48 h before piling illustrates the significant decrease of DPH during piling. The magnitude of this effect decreases with increasing distance from piling (Table 4-7, Figure 4-10). Similar to the model output, one can see a spatial gradient in how much DPH is reduced during the hour of piling: While at distances below 5 km from piling there is a 68 % decrease in DPH during piling, it is only 15 % at 15-20 km and is below 10 % at distances above 30 km. Decreases in DPH were significant in up to the distance category 20-30 km, but significant declines by at least 20 % were only present in up to 10-15 km (Table 4-7).

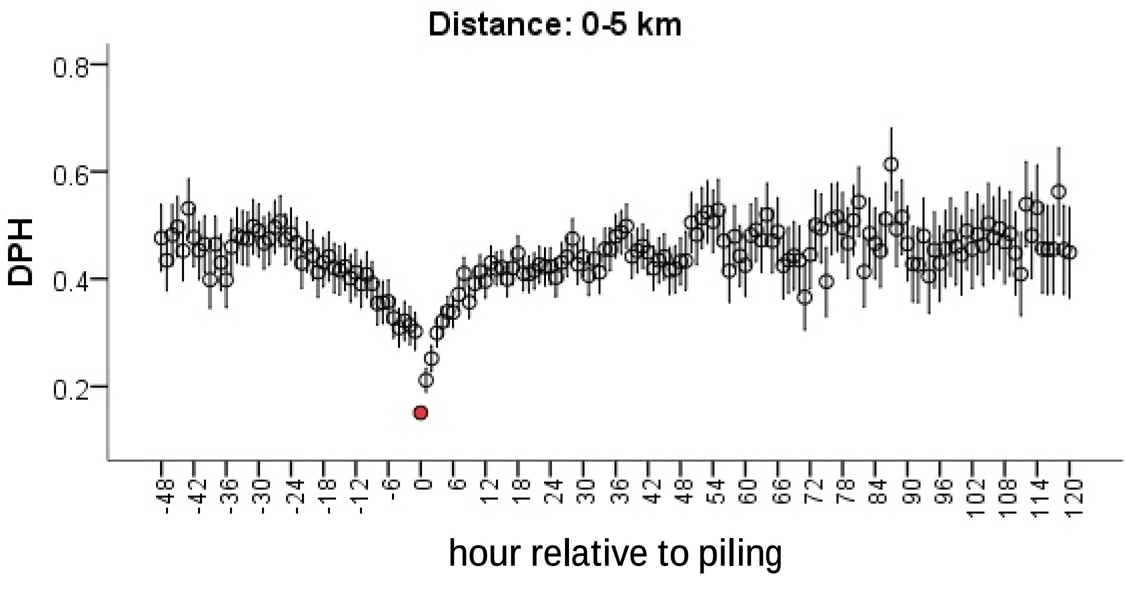


Figure 4-7 Error bars depicting mean and 95% confidence intervals of DPH at the different hours relative to piling for all wind farm projects and for distances between 0 and 5 km. The error bar for hour relative to piling=0 is shown in red.

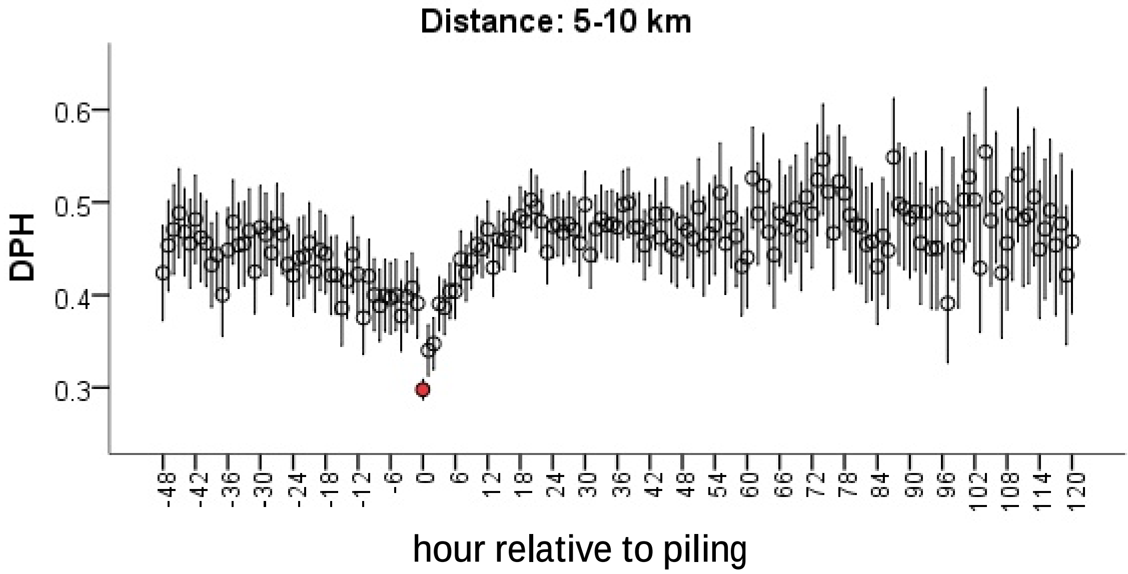


Figure 4-8 *Error bars depicting mean and 95% confidence intervals of DPH at the different hours relative to piling for all wind farm projects and for distances between 5 and 10 km. The error bar for hour relative to piling=0 is shown in red.*

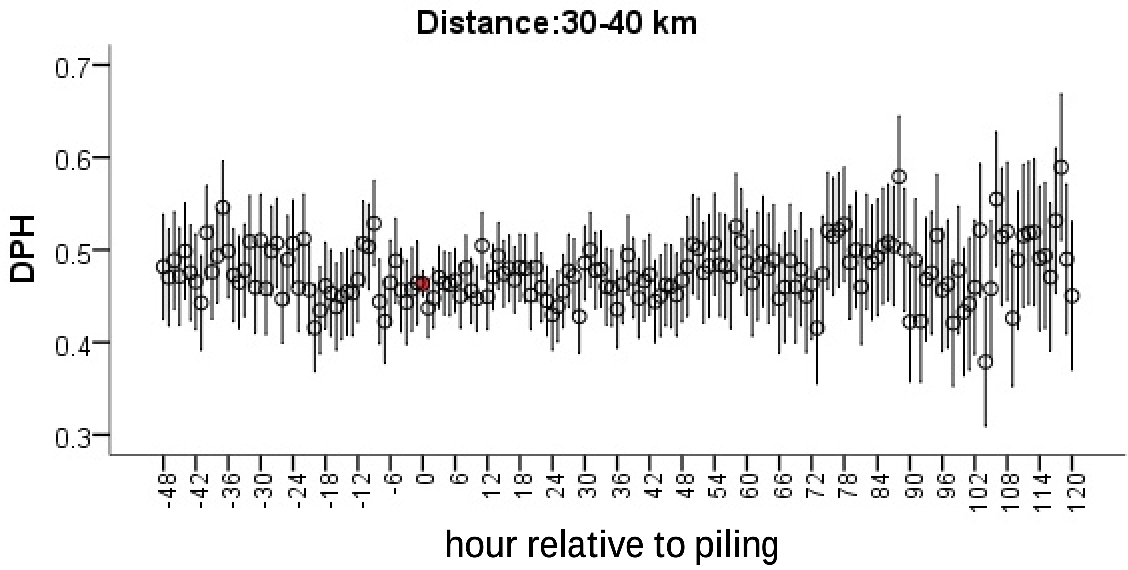


Figure 4-9 *Error bars depicting mean and 95% confidence intervals of DPH at the different hours relative to piling for all wind farm projects and for distances between 30 and 40 km. The error bar for hour relative to piling is shown in red.*

*Table 4-7 Average values for DPH calculated over the global dataset for five different time classes (HRP= Hour Relative to Piling) and six different distance classes Also given is the percentage by how much DPH declined at the hour of piling relative to between 25-48 h before the start of piling and significance levels from Mann-Whitney U test testing differences between DPH at 25-48 h before piling to DPH during piling (***: p<0.001, n.s.: p>0.05).*

variable	HRP= -48 to -25	HRP= -24 to -1	HRP= 0	HRP= 1 to 24	HRP= 25 to 48	DPH decline by	significance
distance: 0-5 km	0.47	0.39	0.15	0.37	0.44	68 %	***
distance: 5-10 km	0.45	0.41	0.30	0.44	0.47	33 %	***
distance: 10-15 km	0.55	0.48	0.41	0.51	0.52	26 %	***
distance: 15-20 km	0.54	0.51	0.46	0.52	0.52	15 %	***
distance: 20-30 km	0.50	0.47	0.44	0.48	0.49	12 %	***
distance: 30-40 km	0.48	0.46	0.46	0.47	0.48	4 %	n.s.
distance: 40-60 km	0.55	0.53	0.52	0.53	0.55	6 %	n.s.

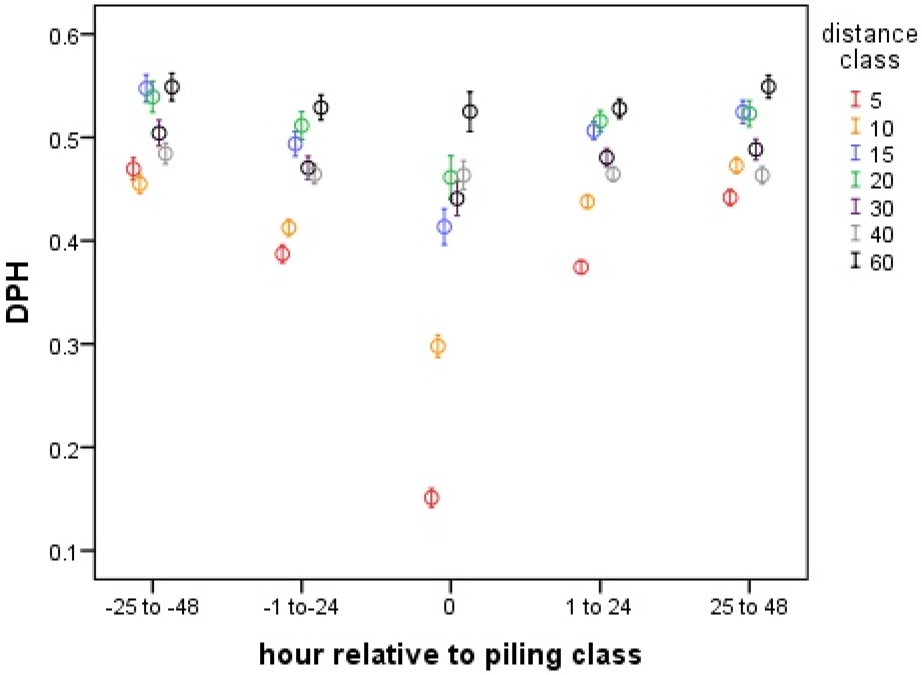


Figure 4-10 Error bars depicting mean and 95% confidence intervals of DPH at different time classes relative to piling (-48: 48-25 h before, -24: 24 to 1 h before, 0: during, 24: 1-24 h after, 48: 25-48 h after) and for different distances classes (shown in different colours, 5:0-5 km, 10:5-10 km, 15:10-15 km, etc.).

4.3.3 Effects of noise mitigation on effect ranges

In order to look at how noise mitigation may alter the effects of piling on harbour porpoise, an additional factor was be included within model G2 (noise mitigation) that had two levels ("yes" and "no") resulting in model G3. Dividing the factor noise mitigation into further categories based on the different mitigation techniques that were used would have resulted in too many levels to still adequately address this issue and would result in even more complications in terms of the variable being confounded with project than what we are dealing with when only using a two-level factor.

Including noise mitigation as a third variable into the interaction of hour relative to piling with distance slightly improved the model (ΔAIC= 380.2, 7.52 % as opposed to 7.44 % deviance explained) and the three-way-interaction term (hour relative to piling, distance, noise mitigation) was highly significant (model G3 in Table 4-3, Table A-2). Looking at predicted deviations of DPH-values from the overall average (Figure 4-11), DPH reached the overall average at about 14 km from piling for piling events with noise mitigation (Figure 4-11, upper figure). For piling events without noise mitigation estimated effect ranges were less clear due to a more complicated pattern of the 0-isocline (indicating the overall average). The global average during piling was reached at 33 km but the change relative to times before and after piling seems minor at that distance (Figure 4-11, lower figure). Clear changes are only visible until about 20 km, so true effect ranges may be somewhere between 20 and 33 km. A decline in DPH started about a day before piling irrespective of whether or not noise mitigation was applied.

When comparing these model outputs it has to be kept in mind that data from piling events without noise mitigation mainly originate from BARD and BWII and that only 1-2 foundations were erected without noise mitigation during the other projects. On the contrary, noise mitigation was only applied during the construction of two foundations at BARD. Thus, comparing results between piling events with and without noise mitigation is confounded by project-specific differences. Furthermore, there were considerably more data for piling events with noise mitigation than for piling events without noise mitigation and for the latter data availability was very uneven across different distances. This could also greatly result in differing model outputs that are unrelated to real noise effects. Thus, the present results could also be related to differences between specific wind farm projects, geographic areas and data availability in addition to different noise mitigation situations.

Nevertheless, it is clear that effect ranges during piling events with noise mitigation was less (about 14 km) when compared to piling events without noise mitigation (20-30 km) or when compared to results from the overall model (model G2 in chapter 4.3.2 yielding effect ranges of about 17 km).

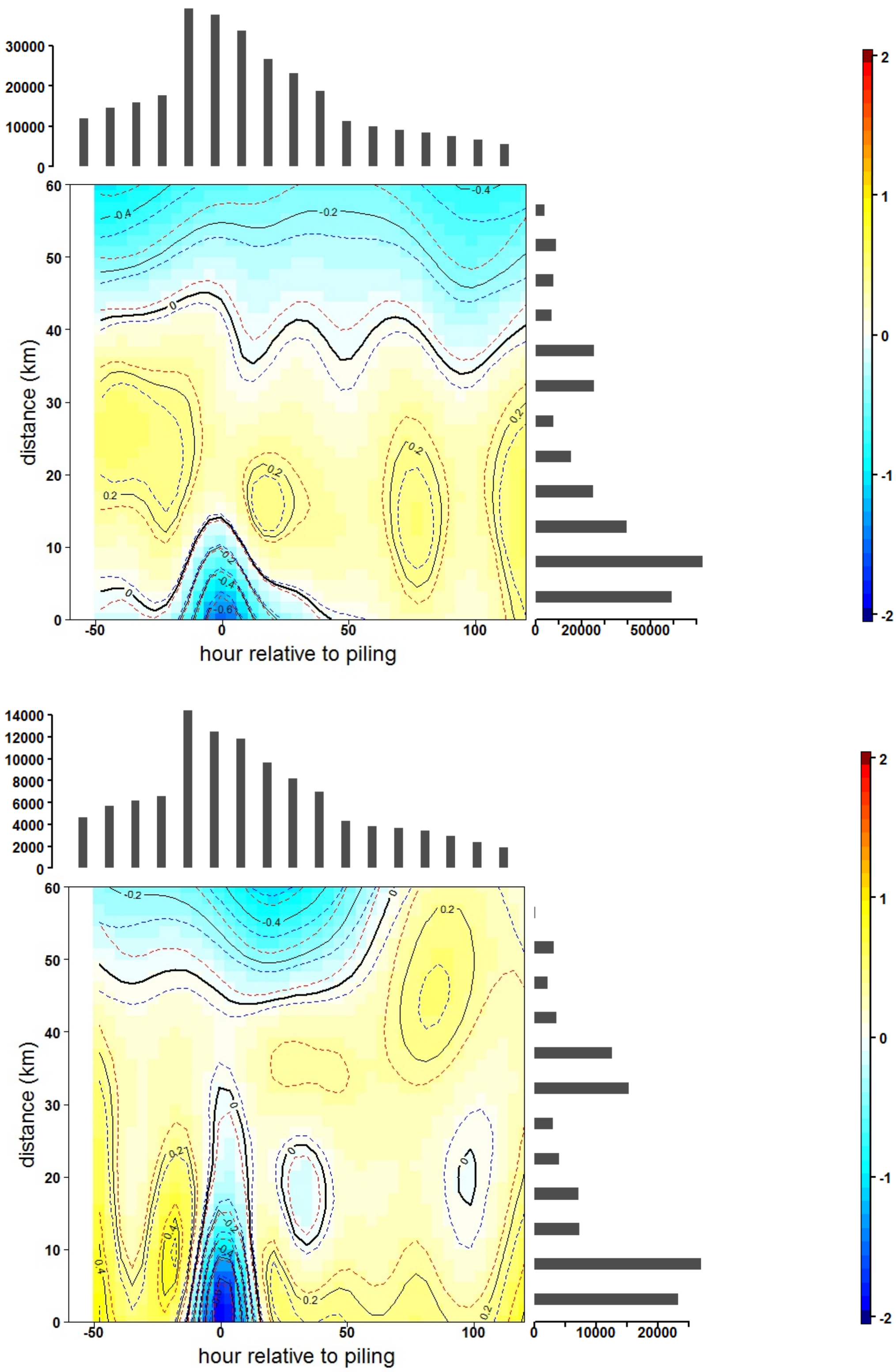


Figure 4-11 *Model outputs from model G3 showing the effects of the interaction of hour relative to piling with distance on DPH for piling events with noise mitigation (above) and without noise mitigation (below). Shown is the predicted deviation of DPH from the overall mean (cold colours: negative deviation, warm colours: positive deviation) in dependence of the relative hours to the time of piling and of the different distances of the specific POD-position.*

4.3.4 Duration of piling effects at close distance

Investigating the effect of hour relative to piling using data at less than 2 km distance

In order to more specifically look at the duration of piling effects at close distances from piling we ran an additional model that included only data within a radius of 2 km around piling (model G4, Table 4-3, Table A-2). Within this model, hour relative to piling had a highly significant effect on DPH, with porpoise detections declining before piling, reaching a minimum shortly after piling and then increasing again afterwards. Model G4 explains about 16.8 % of deviance. DPH started to decrease already about 29 h before the start of piling, reached the overall average at about 12 h before piling, continued to decline and reached the minimum shortly after piling (Figure 4-12). DPH then steeply increased until reaching the overall average at about 20 h after piling and the first local maximum at about 31 h after piling.

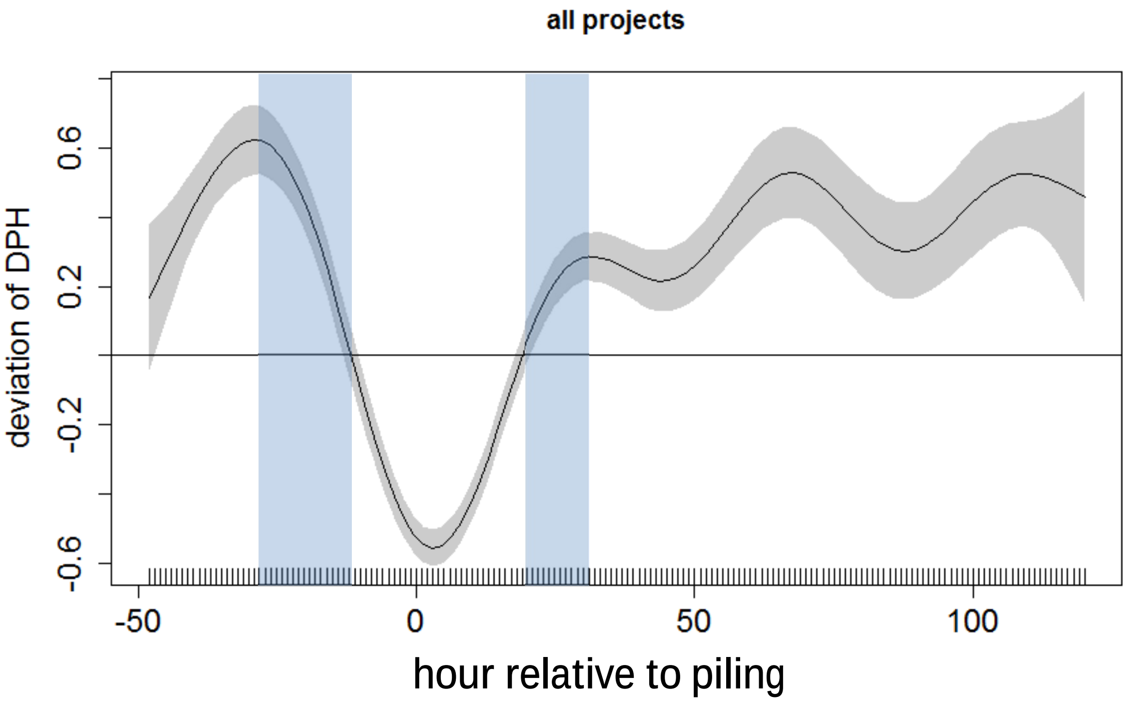


Figure 4-12 Model output from Model G4 showing the effect of hour relative to piling on DPH at distances <2 km from the piling location using the global dataset. Shown is the predicted deviance of DPH from the global mean (thin black horizontal line). Grey shaded areas indicate confidence intervals. The blue boxes indicate the likely onset and the end of piling effect from average values (zero) to last and first local maximum before and after piling. Black tick marks above x-axis indicate data availability.

4.3.5 Project-specific effect ranges and effect duration

Project-specific effect distances

In order to look at project-specific effect ranges, GAM models were run for each project separately (model P2, Table 4-3) but only including piling events when noise mitigation was applied (apart from BARD, which was constructed almost entirely without noise mitigation). Data availability was not sufficient for MSO and RG, so only five projects were considered. There was a significant effect of the interaction of hour relative to piling with distance for each of the five project-specific models. Porpoise detections, measured as DPH, were always lowest during piling at the nearest distance to the construction site. Porpoise detections increased with time and distance relative to piling and a decrease in DPH in the near vicinity to piling always started some hours before and lasted for several hours after piling (Figure 4-13). Project-specific models explain between 5.4 % and 10.1% of variance. For each project-specific model output, Figure 4-13 illustrates the deviation of DPH from the project average in dependence of hour relative to piling and distance.

One aspect that can be seen from these figures is that within most projects there seems to be a natural gradient in porpoise detections with distance from the construction site that is unlikely to be related to any construction effects. This relationship is negative at all projects but BARD indicating that the wind farms were built in areas of higher porpoise detections than surrounding areas in the far distances of up to 60 km distance. At BARD, detection rates are lower than in the vicinity between 10 and 30 km but increase markedly at further distances, indicating that this wind farm was built in an area of less porpoise activity than surrounding areas. These natural differences in detection rates at different distances complicate the detection of piling effects outside the near vicinity. This is especially evident for the areas around BARD where the likely limit for detecting piling effects (around 20-30 km) falls within an area of especially low detection rates. In general, detection rates during piling are always below the overall average at all distances: caused by piling in the near vicinity but likely due to natural patterns at further distances. Effect ranges for BARD can therefore only be assessed somewhat subjectively based on the pattern seen in Figure 4-13 and occur somewhere between 20 and 34 km. For the other projects averages during piling were reached at 6 km at DT, 9 km at GTI and NSO and 16 km at BWII (Table 4-9).

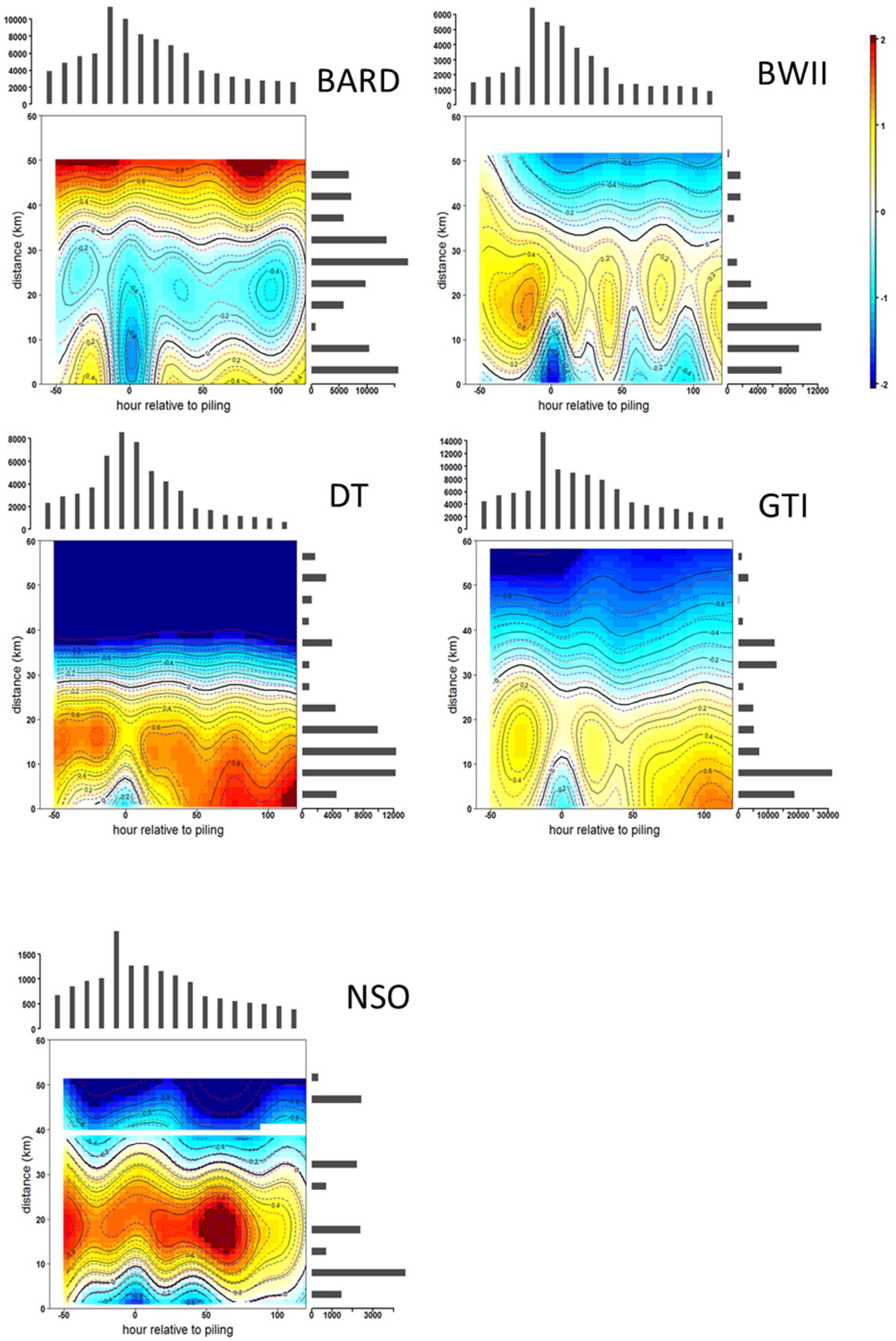


Figure continues on next page

Figure 4-13 Effects of hour relative to piling (from -48 h to +120 h) and distance to piling (in km) on DPH as predicted by the outputs from the six project-specific models (P1). Shown is the deviance of DPH from the project-specific mean (bold 0-line) with cold colours indicating a negative deviation and warm colours a positive one. Note that colour-codes are identical between the projects to ease comparison although absolute detection rates differ between projects. Histograms indicate data availability at the different hours (-50 to -40 h, -40 to -30 h etc.) and distance-classes (0 to 5 km, 5 to 10 km etc.).

All project-specific model outputs showed decreased detection rates prior, during and after piling at the near vicinity with project averages being reached about a day before and after piling, with the exception of NSO, where model outputs show a more complicated pattern (Figure 4-13). During all wind farm project declines in porpoise detections occur already before the start of piling. For all projects a spatial gradient in the magnitude of decreasing detection rates during piling occurs with stronger effects at the closer distances. A spatial gradient in effect duration can also be seen during all projects but BARD, where naturally occurring patterns in detection rates over distances probably complicate detecting piling effect.

Figure 4-14 further illustrates how the effect during the hours of piling differs between different distance categories when compared to 25-48 h before and after piling. Table 4-9 provides the mean values for DPH at the different distance classes during a baseline period before piling (25-48 h before piling) and during piling for each of the seven projects. There was a clear and distinct effect with a strong decrease in DPH during the hour of piling within 5 km for all projects but the magnitude with which detections declined differs markedly between them. With only 48 % it is smallest for DT and highest for BARD with 83 %. This decrease was reduced at greater distances during all projects, but the extent to which this occurred differed between them. DT was the only project, where a significant and clear decline during piling was only found in up to 5 km distance. At 5-10 km distance this decline during piling was already below 10 % and not significant. During all other projects there was still a significant decline in detection rates during piling by more than 20 % in up to 5-10 km distance, apart from RG where this decline was not significant due to low sample size and for which no data existed between 10 and 20 km. At MSO no significant effects occurred at greater distances but sample size at these distances was low. At NSO significant effects and declines by at least 20 % occurred in up to maximal distances of 10-15 km. AT BWII significant effects with a more than 20 % decline were found in up to 10-15 km distance, declines were still significant but only 14 % and 12 % between 15-20 and 20-30 km respectively, the 14 % decline in 30-40 km was not significant, but at 40-60 km a significant 20 % decline occurred. Given that this was not the case at three lower distance classes and that DPH during piling did not differ compared to 25-48 h after piling (Figure 4-14) this is unlikely to be an effect really related to piling activities. At GTI significant declines by at least 20 % occurred in up to 20-30 km, the decline in 30-40 km was still significant but only 12 %. Detection rates at BARD followed a somewhat strange pattern: Strong significant effects with declines by more than 60 % occurred in up to 10 km and these declines were the strongest of all projects despite of lowest detection rates next to GTI. Then effects were not significant at 10-20 km, significant again at 20-30 km with a 20 % decline, not significant at 30-40 km, but again significant at 40-60 km, however, only with a 16 % decline. Comparing the distances over which effects are being found and the duration over which effects are present, DT seems to be a clear exception with effects not reaching far beyond 5 km. Raw data plots further indicate that, at DT, DPH already decreased long before piling started but increased again at a faster rate than during the other projects (Figure A-7 in Appendix). As can be seen in Figure 4-14, DT is also the only project where DPH at 0-10 km distance is higher during hours more than 24 h after piling than during hours more than 24 h before piling and where DPH during more than 24 h after piling is greater at 5-10 km than at all other distance categories.

Table 4-8 summarises project-specific effect ranges based on GAM and non-parametric analyses as well as results found during analyses of the complete dataset.

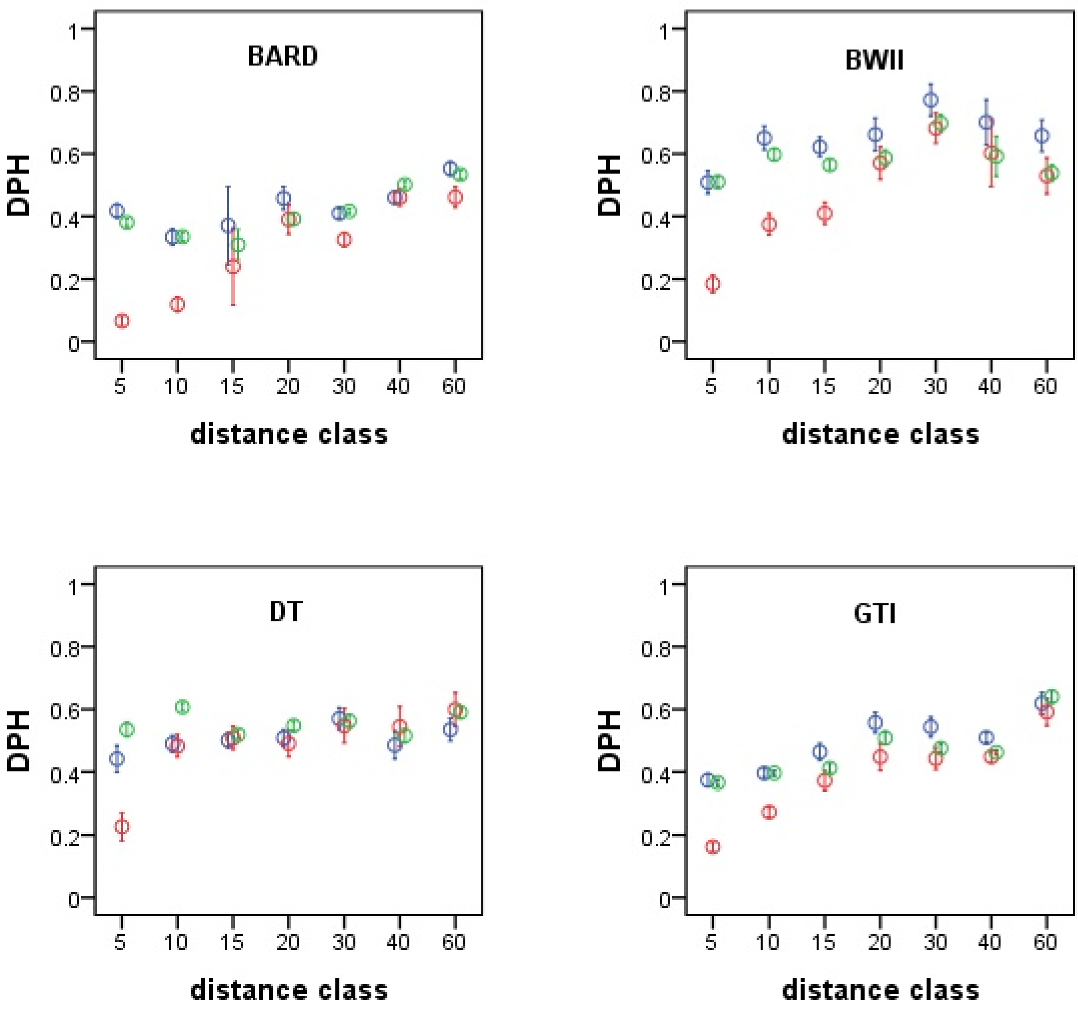


Figure continues on next page

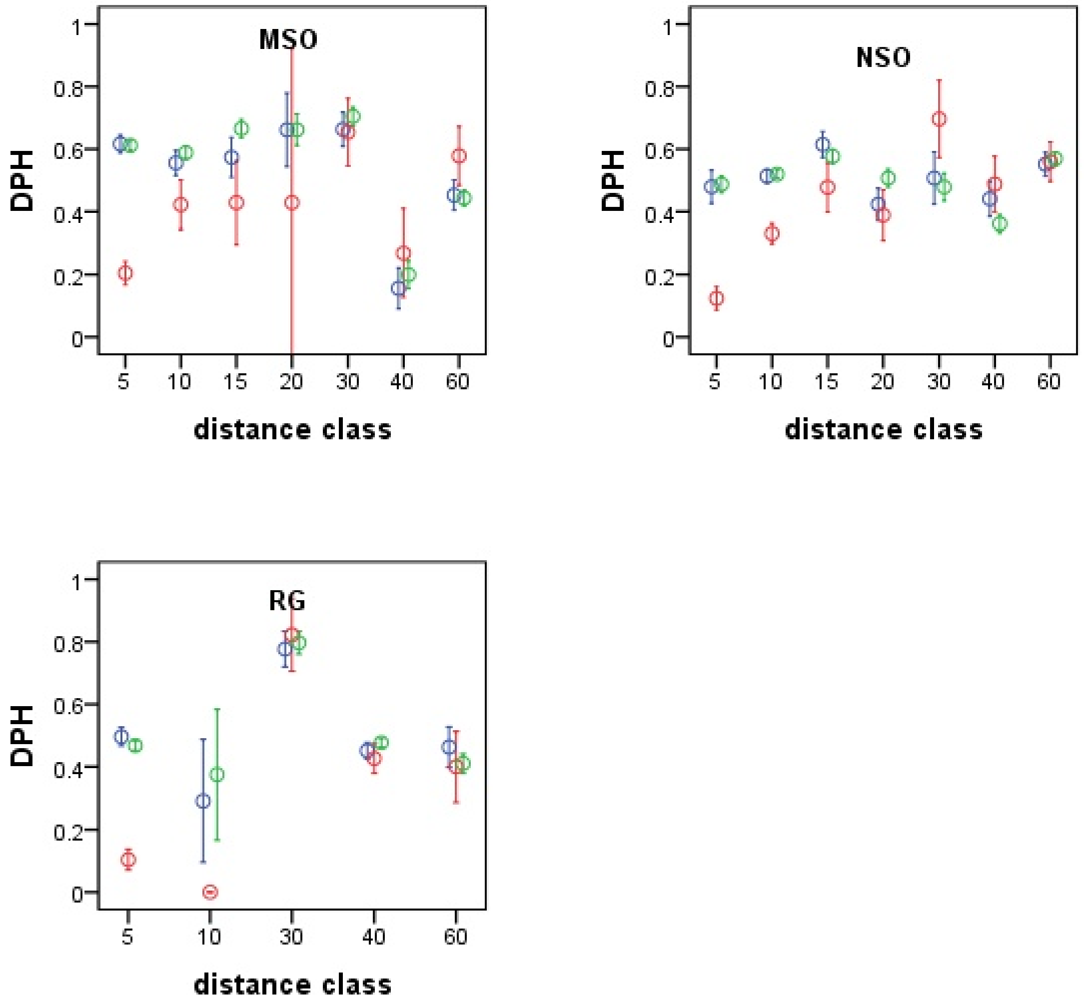


Figure 4-14 Error bars depicting mean and 95% confidence intervals of DPH at the different distance categories for 48-25 h before piling (blue bars), during piling (red bars) and 25-48 h after piling (green bars) for the different wind farm projects. Apart from BARD data only include piling events with noise mitigation. Distance categories: 5=0-5km, 10=5-10km, 15=10-15km, 20=15-20km, 40=20-40km, 60=40-60km.

*Table 4-8 Average values for DPH calculated per project for six different distance classes and two time classes Apart from BARD only piling events with noise mitigation are considered. Also given is the percentage by how much DPH declined at the hour of piling relative to 25-48 h before the start of piling. Also given are significance levels from MANN-Whitney U test testing differences between DPH at 48 to 25 h before piling to DPH during piling (***: p<0.001, *: p<0.05, n.s.: p>0.05). Numbers and tests based on low sample sizes (less than 100 values) are only shown in grey.*

project	time period	distance: 0-5 km	distance: 5-10 km	distance: 10-15 km	distance: 15-20 km	distance: 20-30 km	distance: 30-40 km	distance: 40-60 km
BARD	HRP= -48 to -25	0.42	0.33	0.37	0.46	0.41	0.46	0.55
	HRP=0	0.07	0.12	0.24	0.39	0.33	0.46	0.46
	%decline	83 % ***	64 %***	35 % ns	15 % ns	20 % ***	0 % ns	16 % ***
BWII	HRP = -48 to -25	0.51	0.65	0.62	0.66	0.77	0.70	0.66
	HRP=0	0.18	0.37	0.41	0.57	0.68	0.60	0.53
	%decline	65 % ***	43 % ***	34 % ***	14 % *	12 % *	14 % ns	20 % ***
DT	HRP = -48 to -25	0.44	0.49	0.50	0.51	0.57	0.49	0.54
	HRP=0	0.23	0.48	0.51	0.49	0.55	0.54	0.60
	%decline	48% ***	2 % ns	+2 % ns	4 % ns	4 % ns	+10 % ns	+11 % ns
GTI	HRP = -48 to -25	0.37	0.40	0.46	0.56	0.54	0.51	0.62
	HRP=0	0.16	0.27	0.37	0.45	0.44	0.45	0.59
	%decline	57 % ***	33 % ***	20 % ***	33 % ***	20 % ***	12 % ***	5 % ns
MSO	HRP= -48 to -25	0.62	0.56	0.57	0.66	0.66	0.16	0.45
	HRP=0	0.20	0.42	0.43	0.43	0.65	0.27	0.58
	%decline	68% ***	25 % *	25 % ns	35 % ns	2 % ns	+69 % ns	+29 % ns
NSO	HRP= -48 to -25	0.48	0.51	0.61	0.42	0.51	0.44	0.55
	HRP=0	0.12	0.33	0.48	0.40	0.70	0.49	0.56
	%decline	75 % ***	35 % ***	21 % *	5 % ns	+37 % ns	+11 % ns	+2 % ns
RG	HRP= -48 to -25	0.50	0.29	-	-	0.78	0.45	0.46
	HRP=0	0.10	0.00	-	-	0.82	0.43	0.40
	%decline	80 % ***	100 % ns	-	-	5 % ns	4 % ns	13 % ns

Table 4-9 Effect ranges based on non-parametric approaches (last distance class with significant at least 20 % decline during piling (relative to 25-48 before piling) before distance class with smaller declines) and based on GAM model output (when DPH reached the overall model average during piling). Values in grey and italics are difficult to interpret due to limited or no data availability at the next larger distance category. For comparison results from analyses of the complete dataset are also included.

OWF project	effect range based on significant 20 % decline	effect range based on overall average in GAM model output
all	10-15 km	17 (all), 14 (with noise mitigation)
BARD	5-10 km	20-34 (without noise mitigation)
BWII	10-15 km	16 (with noise mitigation)
DT	0-5 km	6 (with noise mitigation)
GTI	20-30 km	9 (with noise mitigation)
MSO	*5-10 km?*	-
NSO	10-15 km	9 (with noise mitigation)
RG	*0-5 km?*	-

Project-specific effect duration at close distance

Models P4 explained between 15.1 and 19.9 % of the overall deviance. Within all project-specific models looking at the effects of hour relative to piling only using distances < 2 km (P4), hour relative to piling had a highly significant effect (all p<0.001). This effect looks relatively similar within each wind farm specific model (Figure 4-15): DPH started to decrease about a day before piling, reached a minimum around piling and continued to increase over one to two days after piling. The estimated ranges of effect duration differed between projects, however (Figure 4-15, Table 4-10). With the exception of DT, DPH started to decrease between 19 and 32 h before piling and reached the overall average between 10 and 16 h before piling. After reaching a minimum around the hour of piling, DPH started to increase again until reaching the global average between 9 and 28 h after piling and continued to increase until about 16 to 46 h after piling. Effect duration at DT was more difficult to estimate. Here the overall average was reached already 19 h before piling, which is earlier than during any other project and again 23 h after piling, which is within the range of the other projects. However, there is no local maximum that was reached either before or after piling, instead the effect seems to start earlier than 42 h before and lasting longer than 80 h after piling. This, however, is not in line with outputs from model DT_P1, which predicts an effect duration of no longer than about 9 h. This issue may be related to difficulties with low data availability at long periods before and after piling, as piling events often occurred within short time windows.

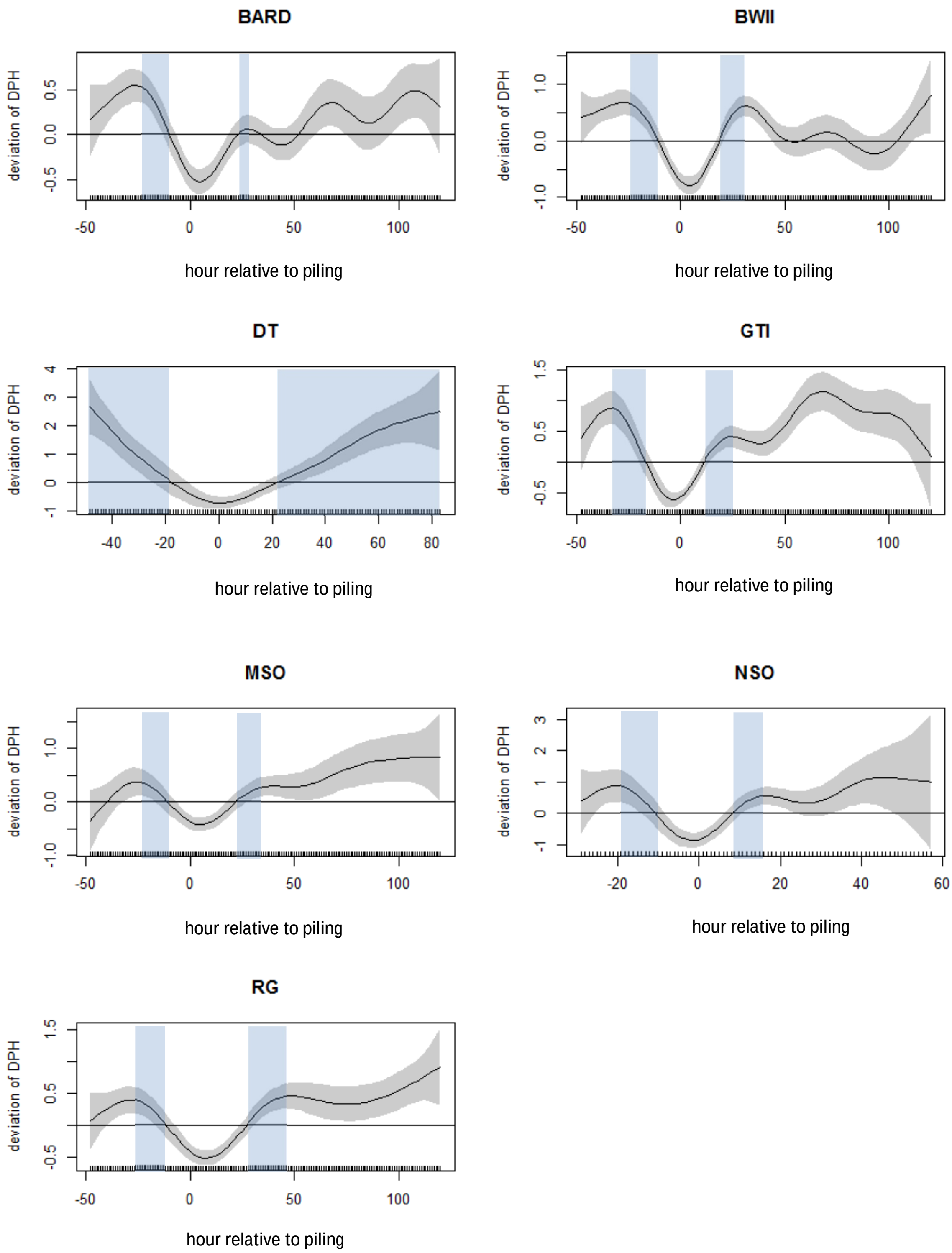


Figure 4-15 Model outputs from Models P4 showing the effects of hour relative to piling on DPH during at distances <2 km from the piling location for the seven different wind farm projects. Shown is the predicted deviance of DPH from the global mean (thin black horizontal line). Grey shaded areas indicate confidence intervals. Black tick marks above x-axis indicate data availability. Blue boxes indicate the likely onset and end of the effect of piling (between the global mean and the last or first local maximum).

Table 4-10 Summary of the effects of hour relative to piling on DPH at < 2 km distance from piling as predicted by the different models (G3, P3).

model	last local maximum before piling	global average reached before piling	global average reached after piling	first local maximum after piling
G5	29 h	12 h	20 h	31 h
BARD_P3	24 h	10 h	24 h	28 h
BWII_P3	24 h	10 h	19 h	31 h
DT_P3	>40 h?	19 h	23 h	>80 h?
GTI_P3	32 h	16 h	12 h	18 h
MSO_P3	24 h	10 h	22 h	34 h
NSO_P3	19 h	10 h	9 h	16 h
RG_P3	26 h	12 h	28 h	46 h

4.3.6 Temporal cumulative effects and habituation

Effects of piling duration

One may specifically expect the duration of a piling event to have a negative effect on DPH. In this case, we speak of temporal cumulative effect. This is because the longer a piling event lasts and the more it may be expected to cause porpoises to leave the area and/or show other behavioural reactions. Therefore, we looked at the effect that piling duration had on porpoise detection rates within the various models we ran. We also show results of one additional model specifically aiming at testing whether the effect of piling duration is apparent when focusing on the hour directly after a piling event (when such an effect, if present, should be most pronounced and clearest).

Piling duration had a significant effect on DPH within model G2 using data from all wind farm projects (Table 4-3, Table A-1). However, the predicted relationship (Figure A-12, lower right) does not show a clear pattern: DPH shows several positive and negative deviations from the global mean over the range of values for piling duration, which are mainly caused by several outliers in piling duration. This could be due to too much variance in piling duration between projects, such that the effect actually describes differences between projects rather than the effects of piling duration itself. Therefore, we also show outputs from the project-specific models P2 (Figure A-12, first five figures). For each wind farm project, piling duration had a significant effect but DPH also displayed several positive and negative deviations from the project mean without a clear discernible pattern.

As the effects of piling duration may highly depend on the distance to a piling event and a clear pattern may only emerge at close distances we also show outputs from models G5 and P5 (Table 4-4, Table A-3), where data were restricted to less than 2 km distance from the piling location. Piling duration had a significant effect within the global model and within all project-specific models apart from BWII. There was a tendency for decreasing DPH with increasing piling duration in the global dataset as well as in most project-specific datasets, especially GTI (Figure 4-16), but during all projects but GTI effects are mainly due to single extreme values for piling duration.

In another step we therefore ran one model specifically designed to test the effects of piling duration (G5), focusing on the hour directly after a piling event (hour relative to piling=1) and only including data at less than 5 km from the piling location. This model was run on the global data set as well as project-specifically (P5). The effect of piling duration within these models was never statistically significant, neither for the global dataset nor for any of the seven project-specific models (all $p>0.1$).

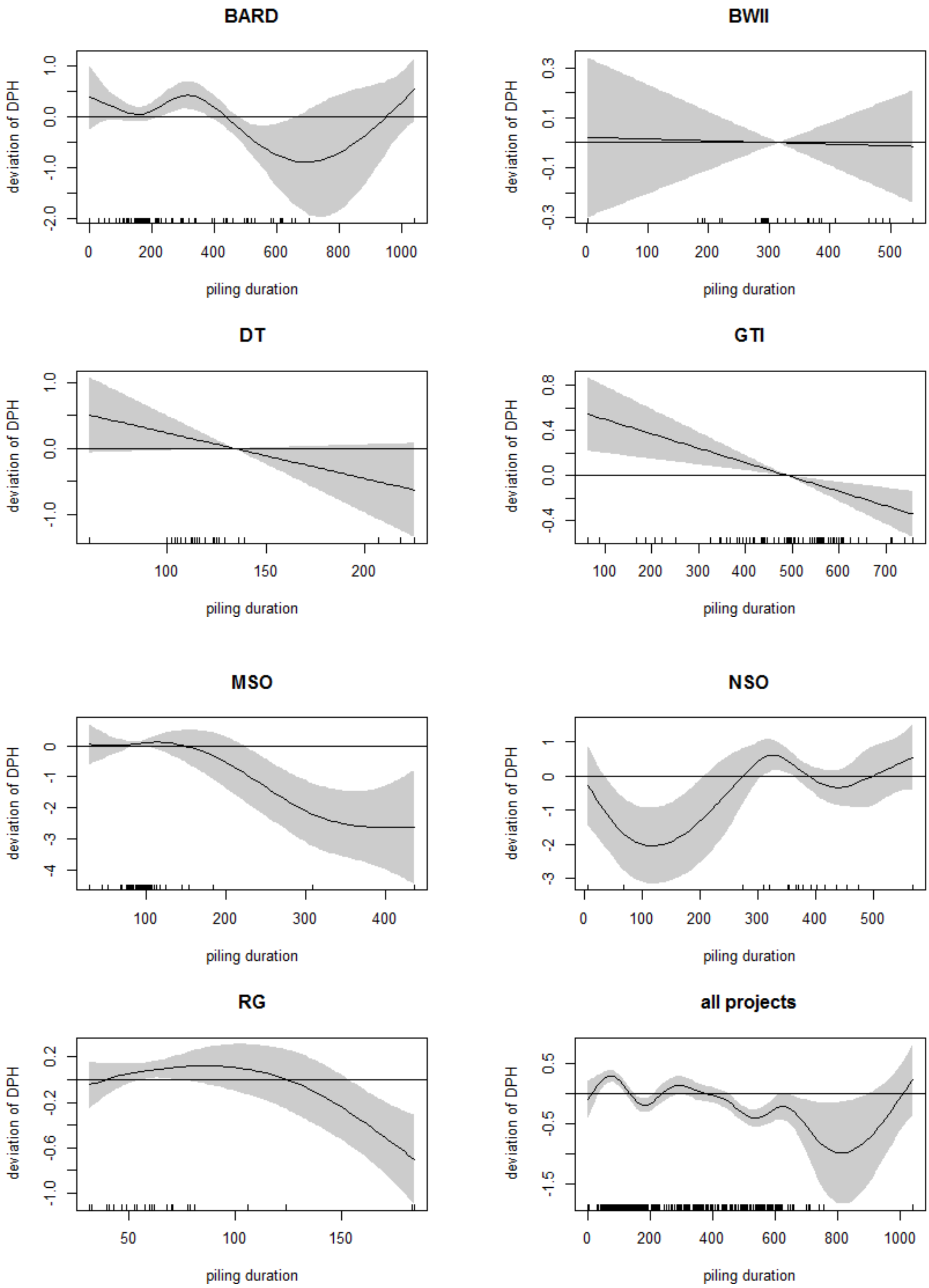


Figure 4-16 Model outputs from Models P5 (first seven figures) and G5 (lower right), in which only data with distance <2 km were included, showing the effects of piling duration on DPH. Shown is the predicted deviation from the overall mean including confidence intervals (grey shaded areas). Black tick marks indicate data availability.

Effects of consecutive piling number and time between piling events

In order to meaningfully test the effects of potential habituation and/or cumulative effects in time we tested the effects of two additional variables on DPH using a subset of the global dataset. The first variable whose effect we tested was piling order, describing the consecutive number of a piling event (numbered within a wind farm project). The second variable was "min since last piling", giving the minutes that had passed between the end of one piling event and the start of the next one. As we wanted to test if there was a habituation effect or cumulative effect in time, it made sense to focus on the times and distances when the effect of piling was greatest. Therefore we used all the data collected during the hour of piling (hour relative to piling=0) and at distances <5 km from the piling location. Due to outliers and extreme values coming from mainly only the project BARD, we further restricted the dataset to only using data with piling order lower than 100 and min since last piling under 20.000 min.

While both variables had a significant effect within this model (G6, Table 4-4, Table A-3) and the model explained 18.4 % deviance in the data, the predicted effect did not show any clear or meaningful pattern (Figure A-13).

When this model was run per wind farm (P6, Table 4-11), piling duration only had a significant effect at BWII and NSO but was only highly significant and showing a clear pattern at BWII. At BWII DPH clearly increased with piling number (Table 4-11, Figure 4-17). The effect at NSO was less clear (Figure A-14). Min since last piling only had a significant effect at GTI but the effect did not show a meaningful pattern (Table 4-11, Figure A-14).

Table 4-11 Summary of the effects of "piling number" and "min since last piling" on DPH within the different models.

model	effects of piling number	effects of min since last piling
G6 (all projects)	***	***
P6_BARD	n.s.	n.s.
P6_BWII	***	n.s.
P6_DT	n.s.	n.s.
P6_GTI	n.s.	*
P6_MSO	n.s.	n.s.
P6_NSO	*	n.s.
P6_RG	n.s.	n.s.

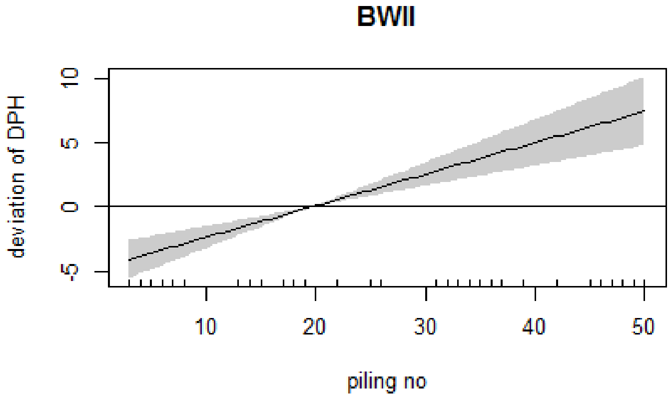


Figure 4-17 Model output from models P6_BWII looking at habituation or temporal cumulative effects illustrating the only highly significant effect found according to Table 4-11 and showing the effect of consecutive piling number on DPH. Shown is the predicted deviation of DPH from the overall mean including confidence intervals (grey shaded areas). Black tick marks indicate data availability.

4.3.7 Wind speed alters construction effects before and during piling

Model outputs on the effect of time relative to piling (4.3.2) showed that there was a decline in DPH already several hours before piling activity started. In order to investigate these effects more specifically and exclude the possibility that these effects are a result of smoothing functions within models G1 and P1, we reran this model only including data before the start of piling (models G7, Table A-4, and P7). There was still a significant effect of hour relative to piling in all these models, so the assumption that this effect only is a result of smoothing can be rejected. DPH started to decline between 15 and 30 hours before piling apart from DT, where the effect seemed to start much earlier (Figure 4-18).

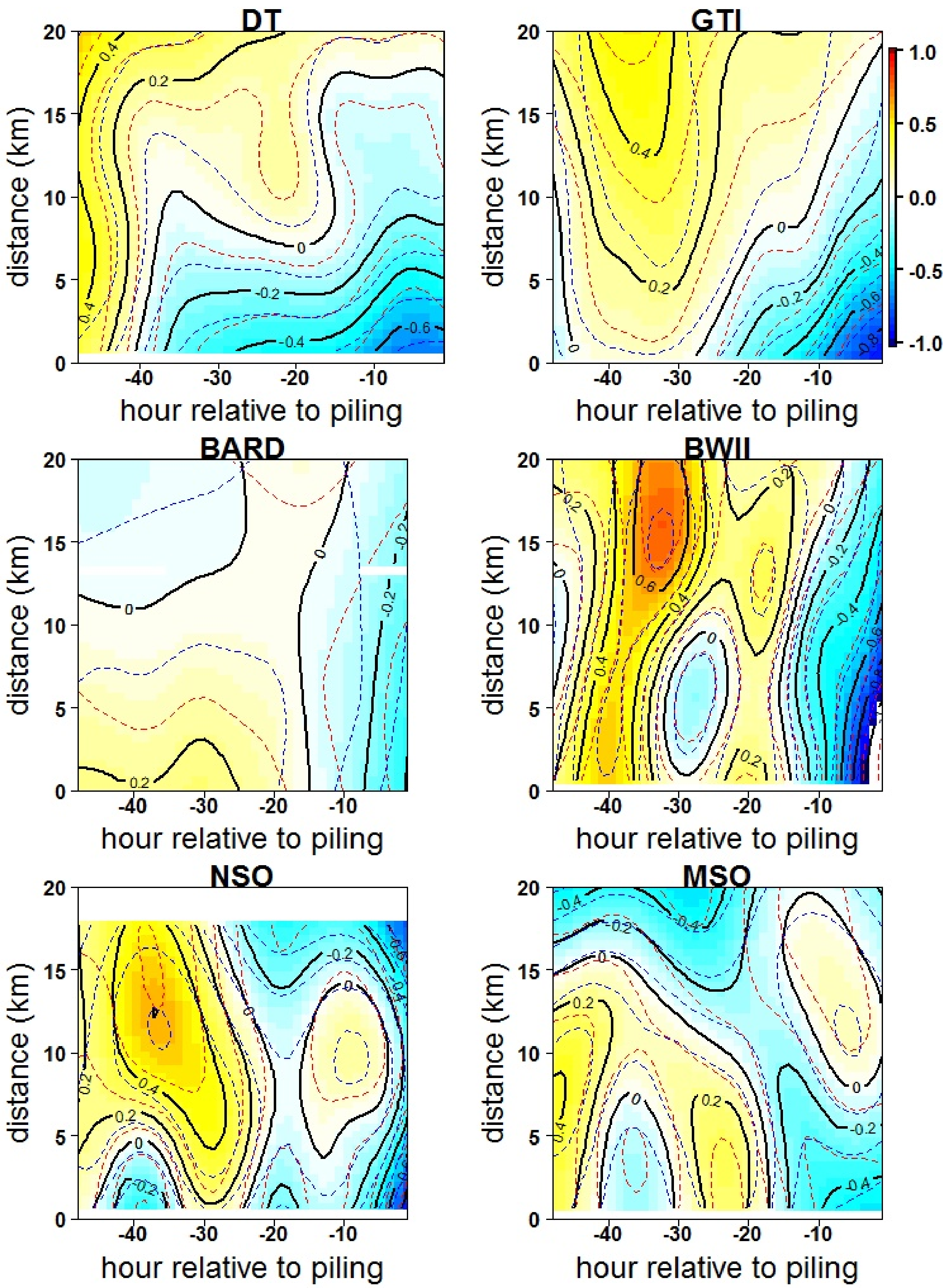


Figure 4-18 *Model output from model P7 illustrating the effect of hour relative to piling and distance to piling on DPH. Shown are the predicted absolute values for DPH. Blank areas indicate no data availability.*

While looking for causes for declining DPH before piling activities, we found that wind speed decreased several hours before piling and increased again after piling, resulting from the fact that construction activities are limited to relatively calm weather conditions. Within this context it may be asked if there is a general tendency of DPH to decline with declining wind speed (as initially seemed to be the case looking at Figure A-15 in chapter A.2.4), which could then cause the decline in DPH before piling. In this case effects before piling would simply be a "good weather phenomenon".

However, while we found that there was a positive effect of wind speed on porpoise detections (see Figure A-15 in chapter A.2.4), further investigations revealed that this was mainly the case within about 20 km from the piling site. So the effect of wind appears to be related to construction activities. We therefore ran a model including the interaction of distance with wind speed (model G8, Table 4-4, Table A-5) on only data five hours before piling (hour relative to piling =-5) and on data collected during piling (hour relative to piling =0, model G9, Table 4-4, Table A-5) to test if the distance where porpoise detections decline before and during piling are related to wind speed, which could affect noise propagation under water and how noise may be detected. The interaction of wind speed with distance was highly significant in both models (Table 4-4) showing that decreases in DPH occurred at larger distances from construction sites when wind speed was lower both five hours before piling as well as during piling (Figure 4-19, Figure 4-20).

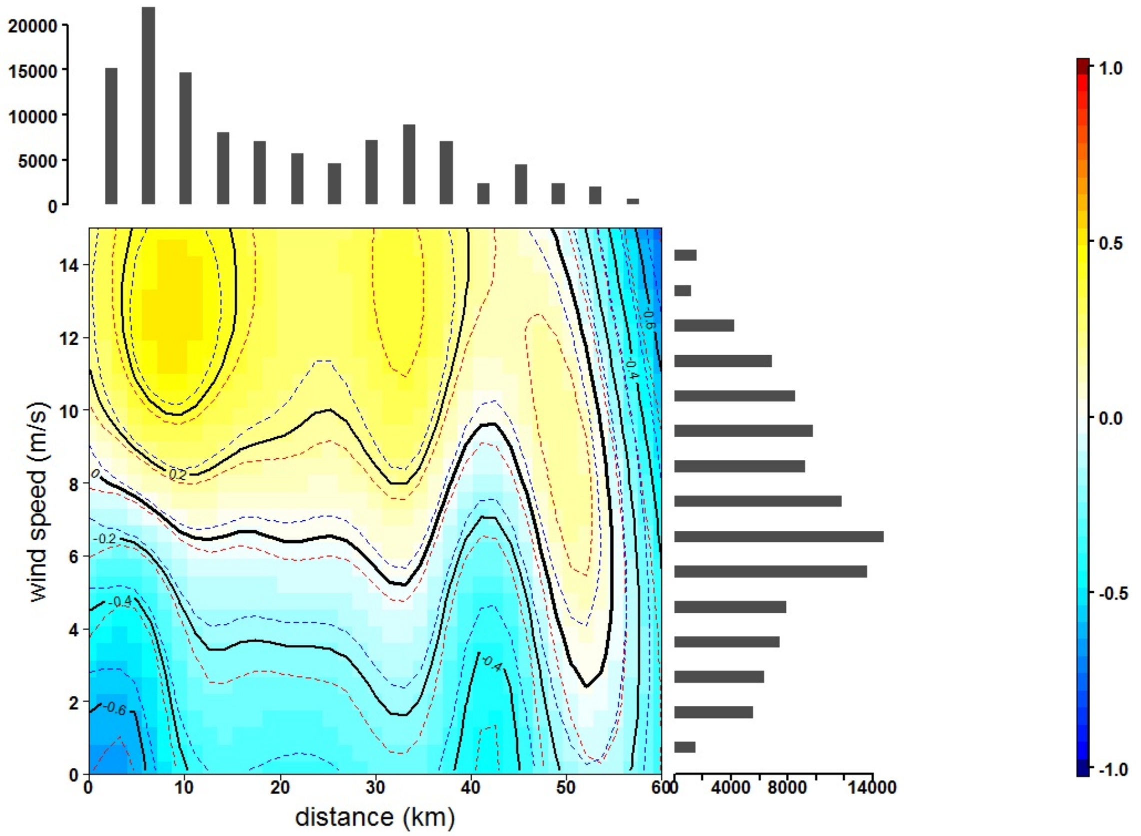


Figure 4-19 Model output from model G8 showing the effect of distance (in km) and wind speed (in m/s) on DPH during hour relative to piling=-5 (5th hour before piling). Shown are the predicted absolute values for DPH.

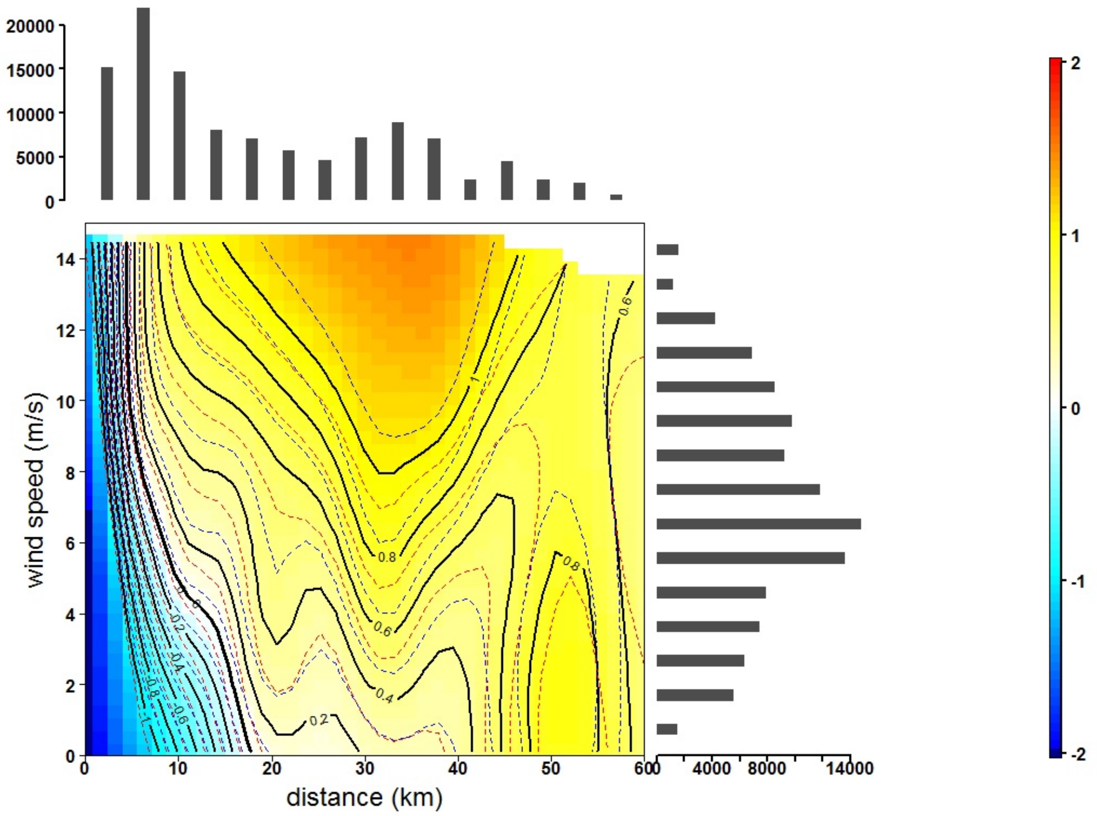


Figure 4-20 Model output from model G9 showing the effect of distance (in km) and wind speed (in m/s) on DPH during the hours of piling (hour relative to piling=0). Shown are the predicted absolute values for DPH.

4.4 Discussion (hourly POD data)

This study analysed acoustic monitoring data on harbour porpoises gathered during several offshore wind farm projects in the German North Sea between 2010 and 2013. The study aimed at describing the effects of underwater noise from construction works on porpoises. In this study, “effects” are defined as a change is porpoise detection rates, which is supposed to be highly connected with porpoise presence (TOUGAARD et al. 2006; SIEBERT & RYE 2008; KYHN et al. 2012; HAELTERS et al. 2013). In order to address small-scale effects of piling activities during the construction of seven offshore wind farms, data were analysed at an hourly resolution with the aim to gain reliable estimates of the noise levels that trigger a change in detection rates and on effect range and duration. We found clear effects of piling niose on porpoise detections with the magnitude and duration of negative effects decreasing with increasing distance or decreasing noise level. When analysing these data separately for each of the seven wind farm projects, substantial differences in the effects occurred that could only partly be related to project-specific construction characteristics (such as noise mitigation). Some differences may be due to environmental characteristics and data availability also and the potential reasons for this are discussed.

4.4.1 Methodology

Acoustic detections: density or behaviour

Acoustic recordings of porpoises provide relative indices of porpoise activity but cannot at present be directly translated into porpoise density. However, previous studies have found these parameters to correlate broadly with porpoise densities obtained from porpoise sightings (TOUGAARD et al. 2006; SIEBERT & RYE 2008; KYHN et al. 2012; HAELTERS et al. 2013). More recent attempts are also being made to estimate densities from POD-data (MARQUES et al. 2009; KYHN et al. 2012). So acoustic porpoise detections seem to be linked to changes in porpoise densities at least to some extent. Nevertheless, behavioural patterns most likely also play a role. This is supported for instance by the occurrence of clear diurnal patterns in detection rates, the explanation for which is most likely linked to foraging behaviour (MIKKELSEN et al. 2013; BRANDT et al. 2014). It may be argued that declining acoustic detections during and after piling is a mere result of porpoises decreasing echolocation behaviour in response to piling rather than swimming away. Studies on other cetacean species such as pilot whales, sperm whales and Cuvier's beaked whales indeed reported reduced vocalising activity during noise exposure but animals also quickly returned to normal patterns after exposure (for review, see WEILGART 2007). Two studies also studied the echolocation activity of porpoises in response to noise: KOSCHINSKI et al. (2003) found no significant difference in the use of echolocation by porpoises when subjected to turbine noise; TEILMANN & TOUGAARD (2006) found echolocation activity of harbour porpoises to decrease in 3 out of 25 sessions when various frequency sounds with a source level of 153 dB re 1 µPa (rms) were played back to them. Although during the present study, porpoises near pile driving might have reduced echolocation activity as a response to the noise of pile driving; we see no convincing reason why animals that rely on their sonar for orientation and foraging should cease doing so for several hours after noise exposure. Most likely, however, decreases in detection rates during piling are a result of a combined effect on porpoise behaviour and abundance. If animals change from foraging behaviour to moving away from a noise source, this will probably lead to a decrease in vocalization. A more directional movement when fleeing from an area may reduce the possibility of a click being recorded by the POD. Finally, aerial survey data analysed during this study also show that porpoise sightings around wind farm construction decreased in up to 10-20 km distance and similar findings were reported by DEGRAER et al. (2012) AND DÄHNE et al. (2013A). BRANDT et al. (2013) also showed that porpoise density decreased following sealscarer exposure. It may be argued that these estimates are also subject to behavioural changes such as differences in diving behaviour. Studies on the effect of seismic surveys on several deep diving cetaceans actually also found the animals to surface in response to noise exposure rather than to decreased surfacing time (WEILGART 2007). This would result in an increased detectability of porpoises during piling and is opposite to what has been found so far. Therefore, porpoise sighting data in response to piling activities support the assumption that a decrease in acoustic detections is not only reflecting a behavioural change but also a change in porpoise abundance.

Methodological choices

Following thorough data exploration, we have chosen to analyse the dataset with a Generalised Additive Modelling (GAM) approach, as it is the most flexible statistical modelling methodology currently available. GAM can easily be combined with non-parametric approaches. Within the

GAM context, correlations were carefully taken into account in several ways (no significant spatial autocorrelation, short-term temporal autocorrelation and selection/exclusion of inter-correlated covariates). Although GAM models are quite sensitive to patchiness in data, the absence of a linear assumption, i.e. the flexibility of splines for smoothing, has helped in the context of the hourly data analyses. We also avoided the use of GAMM (the mixed-model version of GAM) or more complex statistical approaches (GEE for instance) since the hourly dataset is very large and the use of numerous covariates implied computing/convergence difficulties.

Even after carefully preparing the dataset to obtain a baseline reference, variations among wind farm projects occurred through naturally fluctuating detection rates across space. Comparisons between wind farms were thus limited, also partly due to differences in study design and naturally occurring patterns in detections (the problem with gradients across distance, etc.).

4.4.2 Effects of piling noise

From the analyses of the complete dataset and the ones pooled over all seven construction projects, clear negative effects of piling on porpoise detections occurred at noise levels exceeding 143 dB SEL_{05}. This estimate is based on the time when detection rates reached the overall average of the dataset. It may be an underestimation as this average is also partly based on data that were affected by piling. Due to the smoothing functions within the GAM models it is also difficult to draw conclusions on the absolute changes in detection rates at the various distances. Therefore, we also present some non-parametric statistics to show how strongly detections during piling declined relative to a period of 25-48 h before piling at different distance classes.

There was a clear spatial gradient in the amount by how much porpoise detection rates declined during piling depending on the noise level: The louder the noise the stronger the decline in detection rates during piling. Highly significant decreases in porpoise detections by more than 20 % were found in all noise level classes down to 145-150 dB SEL_{05}. Effects were still significant in 135-140 dB and 140-145 dB, but the decline during piling was only 14 %.

The estimate of 143 dB SEL_{05} is close to the estimate of 146 -148 dB SEL_{05} (transformed from 144-146 dB SEL_{50}) given by DIEDERICHS et al. (2014) for declines in porpoise detections during the construction of BWII. KASTELEIN et al. (2013) studied behavioural reactions of harbour porpoises to simulated piling noise in captivity. They found the mean onset of a reaction in terms of jumping out of the water at 136 dB, however, only at 154 dB was the number of jumps significantly different from a baseline. It is difficult, however, to relate findings from captivity to passive acoustic monitoring studies in the field. This is because animals in captivity are constrained in their avoidance behaviour and the motivation for avoidance may differ substantially. Furthermore, noise characteristics in a tank will differ substantially from a natural environment. Passive acoustic monitoring does not yield data on individual behaviour but on the general usage of an area by porpoises. As such 136 dB may be seen as a context specific value for when porpoises may be disturbed.

Comparing model fit, models including distance displayed better goodness of fit than models including noise. This is surprising at first, as noise is supposed to be the primary driver for porpoise avoidance reactions and should therefore have higher explanatory power than distance. Therefore, it raises questions about the quality and quantity of noise measurements. In the case of the

BARD offshore wind farm for instance, noise data were largely extrapolated and based on only two measurements at only one foundation and only at distances beyond 2 km. Noise data were in general largely extrapolated and due to the high variability in even measured data, their accuracy is probably limited. This is also the case as factors such as water depth and sediment type were not considered for noise modelling, but will surely affect noise propagation to a significant degree. An attempt was made to include only real noise measurements for modelling effects on porpoise detections, but it turned out that data were too few for meaningful analyses. Therefore, distance had to be used as a proxy for noise for more detailed analyses.

4.4.3 Spatial range and duration of piling effects

When the variable noise was replaced by distance within GAM models, it was found that clear negative effects of piling on porpoise detections occurred in up to 17 km distance for the complete dataset pooled over all seven wind farm projects and regardless of whether noise mitigation was applied. Non-parametric statistics revealed a significant decline by more than 20 % during piling up to distances of 10-15 km. Declined were still significant in up to 20-30 km but here the decline was only 12 %.

Project-specifically, effect ranges based on GAM analyses could be defined for four projects: 6 km at DT, 9 km at GTI and NSO and 16 km at BWII. At BARD, model outputs were not as clear: During piling the overall model average was reached at 34 km, but was close to the overall average without clear differences between 20 and 34 km. Non-parametric statistics revealed a significant decrease during piling by more than 20 % in up to 0-5 km at DT, 5-10 km at BARD, 10-15 km at BWII and NSO and in up to 20-30 km at GTI. At greater distance classes the decline may still be significant but below 20 %. There are cases at BARD (20-30 km) and BWII (40-60 km), however, when at greater distances classes there were again significant declines by over 20 % even though the effect had already declined at lower distance classes. This may not be a real piling effect as it is unlikely that there should not be effects at lower distance classes but again at larger distance classes. Furthermore, in the case of BWII DPH during piling did not differ from times 25-48 h after piling at 40-60 km distance, which was the case at distance classes below 20 km.

These results point towards considerably shorter effect ranges at DT than at all other wind farm projects. There is no clear distinction between DT and the other wind farms with respect to noise level, piling characteristics or applied noise mitigation that would explain such differences. One could argue that different effects are related to differences in porpoise densities. However, smaller effects should then be expected for BWII also, as it is located within a high density area next to the Natura 2000 area Borkum Reef Ground (see GILLES et al. (2014B) for general densities within these areas and see chapter 6 and for which detection rates in the present dataset were even higher than for DT). More likely, differences at DT are related to natural characteristics of the general habitat such as the presence of prey organisms, which may differ. In the absence of more detailed information on prey species and porpoise behaviour in these areas, this remains largely speculation. A unique effect specific to the area around the wind farm DT was that detection rates more than 24 h after piling were higher than at times more than 24 h before piling up to about 10 km distance from piling. Detection rates during this period were also higher than at greater distances, raising the question as to whether this may be related to increased prey availability after piling.

Effect duration at the close vicinity of the construction site (up to about 2 km distance) lasted up to between 20-31 h (overall average was reached at 20 h, the first local maximum at 31 h). Project-specific models yielded different estimates ranging from 9 to 28 h after piling for when the overall average was reached and 16 to 46 h for when the first local maximum occurred after increasing detection rates (though at DT a local maximum was never reached). It has to be kept in mind that the overall average includes piling affected data such that this estimate is probably an underestimation. This is more severe than for other models using the complete range of data because below 2 km distance all data are affected at least part of the time and based on the present estimates also for the majority of the time that was covered. It may therefore be a more realistic estimate to state when the first local maxima were reached (complete dataset: 31 h, project-specifically: 16-46 h).

A clear spatial gradient existed in effect duration with shorter lasting effects at greater distances and effects only being detected during piling at the largest distances. This is in line with DIEDERICHS et al. (2010) and BRANDT et al. (2011). TOUGAARD et al. (2009) could not show this for the wind farm Horns Rev 1, but this may be linked to limited data availability. Given that the magnitude of a decrease in detections also decreased with distance (indicating that a smaller proportion of porpoises left the area in response to piling noise), it is expected that this also applies to the duration of a negative piling effect. This is because if porpoises left the construction site in response to piling (and all data point towards this), it will naturally be expected that the outer areas will be revisited quicker than the vicinity to construction because porpoises have a smaller distance to cover to get there.

Despite a clear decrease in porpoise detection rates during piling for all wind farm projects, detection rates did not reach zero at any distance from piling. The present analyses on an hourly resolution and under the specific definition of a piling event we applied (including piling breaks up to 3 h) does not allow to conclude whether such detections occurred while piling was ongoing, shortly before or after piling or during a break. However, despite this limitation this could either indicate that even at a close range of a few kilometres distance to the piling location not all animals leave the area during piling or that due to high turnover rates there is a constant arrival of new animals that were not affected.

4.4.4 Effects of noise mitigation

This study is the first to combine data from several offshore wind farm projects applying noise mitigation measures during construction. So far, most previous studies have addressed unmitigated piling activities (e.g. TOUGAARD et al. 2009; BRANDT et al. 2011). Those studies showed significant decreases in acoustic porpoise detections in up to 20 km around wind farm construction sites. Estimates from our overall model resulted in an effect range of 17 km, estimates from the model only considering piling events with noise mitigation applied yielded an effect range of 14 km. Both are below previous estimates. Non-parametric approaches yielded similar results when effect range was defined as a significant decline by at least 20 % (10-15 km). Effect ranges partly depended on analysing techniques and especially on the definition used to identify effect ranges. This complicates comparisons to previous studies that used differing statistical approaches and possibly definitions. Assuming that noise mitigation worked perfectly, e.g. matching the criterion of 160 dB at 750 m, one may expect effect ranges not to exceed 5 km (DIEDERICHS et al. 2014; NEHLS et al. 2016). However, measured noise levels for piling events with noise mitigation displayed high variability within all wind farm projects. This high variance indicates large differences in the effectiveness of noise mitigation measures probably due to various reasons.

In order to look at the effects of noise mitigation a two-level factor (presence or absence) was introduced into the global model investigating the effects of distance from and time to piling on porpoise detections. The resulting model revealed a further reaching effect in terms of distance for piling events without noise mitigation than for piling events with noise mitigation. The global average was reached at about 14 km distance during piling for piling events with noise mitigation and between 20 and 33 km distance for piling events without noise mitigation. However, while including noise mitigation into the model slightly improved model fit, this improvement was minor. Furthermore, the model on piling events without noise mitigation is largely based on data from only BWII and BARD. Therefore, it is unclear whether differences are really due to the application of noise mitigation or caused by other project-specific differences. Due to a great variability in detection rates at the various distance classes at BARD, output estimates from the model on piling events without noise mitigation were less precise. Furthermore, data availability in terms of POD-data and noise levels was not sufficient to conclusively address porpoise reaction to noise levels when pile driving was not accompanied by noise mitigation systems.

Even an effect range reduced to 14 km under applied noise mitigation is larger than what would be expected if noise mitigation worked efficiently. This again indicates that noise mitigation was not always equally effective. Measured noise data indeed showed that there was great variability in noise levels during the operation of noise mitigation systems, indicating strong differences in their efficiency. Furthermore, various compositions of noise mitigation techniques were applied (linear bubble curtain in addition to a circular one etc.), tested and improved throughout construction, which certainly affected noise levels to a great degree. Although detailed information was available on the different combinations of noise mitigation techniques applied, there was little information available concerning the positioning of linear bubble curtains and their specific efficiencies in terms of noise reduction especially at further distances from the piling location. Under the difficult working conditions in the German North Sea, noise mitigation devices are expected to display varying degrees of effectiveness, caused by technical issues, weather-specific phenomena, sediment type and other reasons. Therefore, more detailed investigations on the different noise mitigation techniques applied were not possible.

Therefore, the only way to assess the success of effective noise mitigation at present is to theoretically apply the noise levels that trigger a change in detection rates. When a noise mitigation system is able to reduce the noise level equally into all directions to a certain degree, the area in which porpoises are affected by noise is also reduced. According to NEHLS & BELLMANN (2013) a noise level of 143 dB SEL_{05} was reached at about 20 km distance from piling when no noise mitigation was applied at BWII and at only about 6 km when a double bubble curtain was working effectively. This would reduce the affected area from 1,257 km² to only about 113 km² by over 90 %. The present results show, however, that noise mitigation did not always work sufficiently and therefore disturbance effects on porpoises were only mitigated to some extent.

4.4.5 Cumulative effects and habituation

The variable 'piling duration' was used as an indicator of temporal cumulative impacts of piling on porpoises. Within models restricted to 2 km distance, piling duration had a significant effect within most project-specific models, generally revealing a slightly negative relationship. This indicates a stronger effect of longer lasting piling activities, either in terms of magnitude, distance or duration. However, another distance-restricted model looking specifically at the hour directly after piling did not reveal significant effects. This means that effects could simply be due to detection rates decreasing throughout ongoing piling activities with smallest detection rates at the end of a long lasting piling event. It does not necessarily mean that the duration of an effect differs depending on piling duration. Nevertheless, it would still indicate a cumulative effect, as detection rates at the end of a long piling event may be lower than at the end of a short piling event. To our knowledge the only other study addressing the effect of piling duration so far was conducted by DÄHNE et al. (2013A). They showed that the first waiting time after piling was significantly longer with longer lasting piling events. These waiting times, however, spanned the piling period itself and therefore included the time when porpoises left the area during piling. It is thus expected that longer lasting piling events will cause longer waiting times, simply because porpoises left during piling, which does not prove a cumulative effect. In order to test if piling duration affects porpoise detections even after piling ended (in which case it would truly represent a cumulative effect) one would have to cut these waiting times at the end of piling and test if a significant effect remains.

Another approach to address temporal cumulative and habituation effects was to study the effect of the time between two piling events on porpoise detection rates. No clear pattern emerged from these analyses, indicating neither cumulative nor habituation effects. We also studied if the sequential number of a piling event during the construction period of a wind farm affected porpoise detection rates. If a temporal cumulative effect was present, a negative relationship would be expected. If there was habituation, one would expect a positive relationship. Within models specifically calculated to address this issue, we only found a clear significant effect within one out of seven wind farm projects. At BWII, a positive relationship was found, indicating habituation rather than cumulative effects. Given that this was not the case for the other project-specific models, however, we cannot entirely exclude the possibility that this result is confounded by other BWII characteristics. In conclusion, addressing potential habituation and cumulative effects is a complex issue as there are various confounding effects difficult to disentangle.

4.4.6 Effects before piling and environmental effects

One main result that appeared within all project-specific models was a decrease of porpoise detections already prior to the start of any piling and/or deterrence activities. Increased shipping and preparatory activities (e.g. bubble curtains lay-out) may well cause disturbance several hours before piling. In the absence of detailed information on the extent of industrial shipping and preparation activities, we focused on identifying whether this decrease in DPH prior to piling would be related to environmental characteristics. As wind farm construction is often limited to relatively calm weather conditions, there is a bias in piling occurring at relatively low wind speed. If there is a general pattern of porpoise detections to decrease during calm weather, a pattern, which emerged from our global and project-specific models, this decrease before piling may simply be a "good weather phenomenon" unrelated to any construction activities. Further investigations revealed that this was not the case as the positive relationship between wind speed and porpoise detections depended on the distance to the construction site. This indicates that noise from anthropogenic activities at the construction site travelled further at low wind speed (due to increased reflection at the sea surface and less noise mitigation by fewer air bubbles in the water) and that this led to a further reaching deterrence effect on porpoises at low wind speed.

Focusing on a specific hour before piling (5th hour before the start of piling and or deterrence) and using the global dataset, we found that a decrease in DPH occurred at further distances when wind speed was lower. This was also the case when we checked this for the hour during which piling occurred. This indicates that deterrence radii before, as well as during pile driving reached further at lower wind speeds. The issue may be related to differing noise propagation characteristics at different weather conditions. There is evidence for noise to travel further at lower wind speed due to fewer air bubbles in the water (decreasing natural noise mitigation) and stronger reflection of noise by a smooth water surface (HEINIS & DE JONG 2015). Furthermore, natural noise mitigation was found to play a greater role at the sea surface than just above the sea floor because of air bubbles mitigating the effect in the upper water layer (HEINIS & DE JONG 2015). It is similarly expected that a well-mixed water column (e.g., a still unstratified water column at the beginning of spring) might reduce noise propagation. The magnitude with which those physical processes may reduce piling noise and noise from preparation activities such as shipping traffic is unclear. It also remains unresolved whether this process influences the number of porpoise detections. Further investigations are needed, especially noise measurements of piling and shipping noise at various distances and at differing weather conditions, to shed light on this issue. Future consideration of AIS data, when studying effects of offshore construction activities on porpoise detections, may also be helpful in this respect.

Regarding the effects of other environmental variables on porpoise detections, it was found that, in addition to static environmental variables such as location and sediment type, there was an effect of sea-surface temperature that should be taken into account. Previous studies in the same area (HEINÄNEN et al. 2015) used different models according to season and highlighted the importance of salinity for porpoise distribution. Salinity data did not exist for the complete study period and for this reason were not considered.

4.4.7 Perspectives

Although this study involved a lot of analyses and results, it has also opened numerous perspectives on what to consider in future works. Further investigations would benefit from the calculation of absolute density maps directly from POD recordings, based on latest statistical methodologies (MARQUES et al. 2009; KYHN et al. 2012). Similar studies during the SAMBAH project allowed the estimation of porpoise density maps in the entire Baltic Sea, then to be compared to density maps from aerial flight data. Such an approach requires paying attention to the differences in spatio-temporal scales prior to any comparative interpretation. This would help in disentangling noise impact during the hours of piling from natural gradients occurring over distances.

The current project benefits from a relatively large dataset originating from standard monitoring activities during the construction of seven wind farms. Even though the standard program delivered a quite good and large dataset an improvement of data suitability during future (research) studies could better address questions of noise effects on harbour porpoises. This mainly depends on the possibility to combine real noise measurements with more balanced (in terms of covering different distances) porpoise data. Better coverage at the crucial noise levels between about 135 and 145 dB may also be beneficial when focussing on the exact effect range. Therefore, a more experimental approach with a higher number of regular POD locations at the crucial locations would increase our knowledge of baseline processes and greatly improve the outcomes of statistical models. Furthermore, experimental approaches focusing on the behaviour of individuals may also be of great value in this respect. Finally, further research on environmental characteristics (e.g., salinity, wind, etc.) and noise measurements at the appropriate frequencies conducted already before the start of piling could gain further insight into environment-related and piling-related processes that potentially cause porpoises disturbance before piling and deterrence activities.

5 DAILY POD-DATA

5.1 Introduction

In addition to POD-data on an hourly resolution, POD-data on a daily resolution provides insight into broader temporal and spatial trends and how these may be affected by construction activities. Using POD-data on a daily resolution, we specifically investigate whether it is possible to discern normal inter- and intra-annual variation of acoustic harbour porpoise detections and the effects of pile driving on these. This section is based on data collected with passive acoustic monitoring stations within the scope of STUK (BSH 2013). Here, we cover corresponding methodological issues as well as the results of statistical data analyses in favour of a better understanding of the impact of piling events in space and time. In addition to (a) general analyses concerning the relation of natural parameters to acoustic harbour porpoise detections, we will focus on four piling related questions: (b) Are long term effects visible over the course of the study period? (c) Is there any indication of habituation discernible from our data? (d) Is the effect size of piling context dependent on subarea and season representative for times and areas with different magnitude of porpoise detections? (e) Is there a mutual reinforcement effect of piling events during the same day or on successive days within the same wind farm?

5.2 Methods

5.2.1 Stationary PODs data

In this chapter, data from two different POD categories, POD stations and single stationary PODs, were used to record porpoise clicks. The general principle of POD deployment is covered in detail in subsection “POD-deployment and data processing” page 18.

Mobile POD data were excluded to ensure that data derived from a comparable study design only holding long term data. In total our dataset contained data from 76 different stationary POD positions. Figure 4-1 shows the deployment locations of stationary PODs from the beginning of 2010 until the end of 2013. Single stationary PODs and so called POD stations differ in the number of PODs deployed at one position, meaning that POD stations consist of three simultaneously deployed PODs in close vicinity. Multiple PODs simultaneously monitoring at one location accounts for occasional loss or malfunction of devices. To remove unwanted redundancy from our dataset only the POD with the longest time series of data available was chosen. Gaps, resulting from this procedure, were filled up with available data from another POD of the same station.

PODs record all kind of noise within a specific frequency band and with a specific click characteristic. Besides porpoises, clicks can result from environmental noise like clicking crustaceans and sediment movement as well as from human activities or the mooring of the POD. All clicks recorded are summed up per time in the variable “all clicks”. Later those clicks are categorised according to its most likely noise source (porpoise, sonar or noise) by evaluating the click characteristics and sequential arrangement of clicks in order to identify porpoise clicks. Furthermore, a likelihood (low, medium and high) on the goodness of the estimate is given. Only porpoise click data classified as medium or high are used for creating the response variable detection positive 10 minutes (dp10m) per day. Dp10m per day can be understood as every of the 144 possible ten minutes blocks per day getting assigned a 1 or a 0 (porpoise present yes or no) leading to a maximum value of 144 dp10m/day. It is therefore crucial to avoid any bias caused by diurnal activity patterns of porpoises and thus only complete days with 144 ten minutes blocks of data were used. Therefore, days where PODs had not recorded the entire day, e.g. because of deployment or recovery, where discarded.

5.2.2 Environmental and piling-related explanatory variables

Variables associated with Pile Driving

We defined a day to be a pile driving day if any portion of a pile driving event took place on that very day. The decision to do so is motivated by the definition of pile driving day for the flight data (see chapter 6 on Aerial survey data on page 103) and validated by the data exploration presented in Figure 5-1: The lower plot indicates that those piling events which took place over midnight were evenly distributed around midnight. The upper two plots show that there was no general limit of piling duration per day, which lead to a more rapid decline of harbour porpoise detections per day (Figure 5-1). On the contrary, they seem to be almost constant over this first period (Figure 5-1). If approximately 7 h per day were influenced by piling activities harbour porpoise detections decreased gradually but not rapidly (Figure 5-1). In our models we therefore included pile driving duration to account for different effect sizes due to this variable, as well as hour of piling start to account for diurnal differences in harbour porpoise behaviour. To correct for distance related effect sizes of piling events on acoustic harbour porpoise detections we included the distance to the nearest pile driving site as a numeric variable into our models. If required for the question at hand, we also used the number of piling events per day as an explanatory factor as well as the number of consecutive days with piling events in a row and a factor variable accounting for post effects of piling (days since piling day).

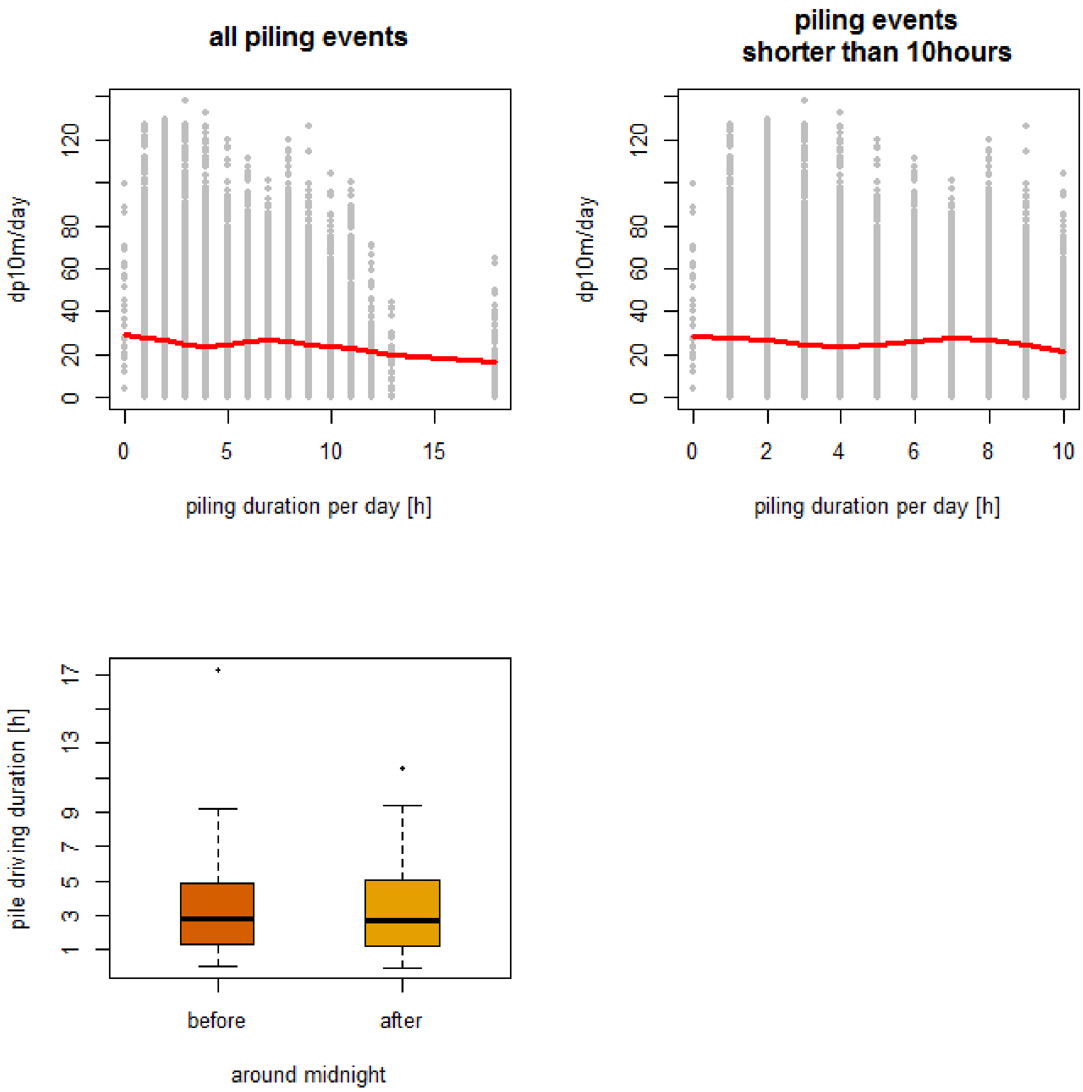


Figure 5-1 *Upper two plots: Piling duration and its effect on the decline of acoustic harbour porpoise detections. The grey dots mark the actual data and the red line is a trendline (fitted with smooth.spline() from the stats package). The lower left plot shows the distribution of piling hours for events which took place over midnight.*

Environmental, temporal, spatial and technical variables

Here we introduce all non-piling related explanatory variables that we finally selected for modelling the daily POD data. Natural parameters are crucial for separating effects caused by pile driving from natural variability and dependencies: Effects of piling can only be assessed when considering the biology of porpoises. The selection criteria for explanatory variables were potential biological relevance, goodness of fit of the model and no, or at least no strong intercorrelation with other variables used in the same model.

In the case of temporal, spatial and technical explanatory variables information was included in the data already. Environmental data (see subsection 3.3 on Environmental data on page 17) on the other hand were merged with porpoise data based on geographical positions and if needed on date (SST, SSTA).

The final choice of variables was day of year, an integer variable incorporated as a cyclic smooth into models to account for seasonal variation in acoustic harbour porpoise detections. This artificial variable describes inner-yearly fluctuations better than sea surface water temperature (SST). For our data sea surface temperature anomalies (SSTA) proved to be useful in explaining further variation in acoustic harbour porpoise detections. To account for inter-yearly differences in acoustic harbour porpoise detections the variable year was either included as a categorical (factor) or a continuous variable (smooth) into the respective models.

All clicks is a variable which holds any clicks recorded by the POD, regardless of their origin. Therefore, it gives an insight into the noise level surrounding the deployment position of the POD. The relation between wind speed and number of all clicks was analysed. We refrained from calculating noise clicks by subtracting identified porpoise clicks from all clicks because even then a possibly great amount of porpoise clicks (categorised as low likelihood or unidentified by the software) as well as other animal sounds or sonar would still be included in the resulting noise category. Moreover the proportion of all clicks categorised as identified porpoise clicks is evanescent. Figure 5-2 illustrates this dependency for the station S13 and S6 exemplarily. For both stations, a correlation between wind speed and the number of all clicks can be seen as it is known that ambient water noise increases with wind speed (WILLE & GEYER 1984). However, there is also an opposing trend. With increasing wind speed the numbers of air bubbles in the water increases which can attenuate the noise level again (similar to a bubble curtain). This cannot be seen in our data as the increase of clicks (noise) is especially high for wind speeds greater than 13 knots per hour. However, the impact of wind on station S6 is smaller than on station S13. Different noise levels of stations possibly can be caused by different water depth at the station, deployment depth of the C-POD or different anchoring systems. However, no clear trend was visible examining the effect of water depth alone (for further details see section A.3.1).

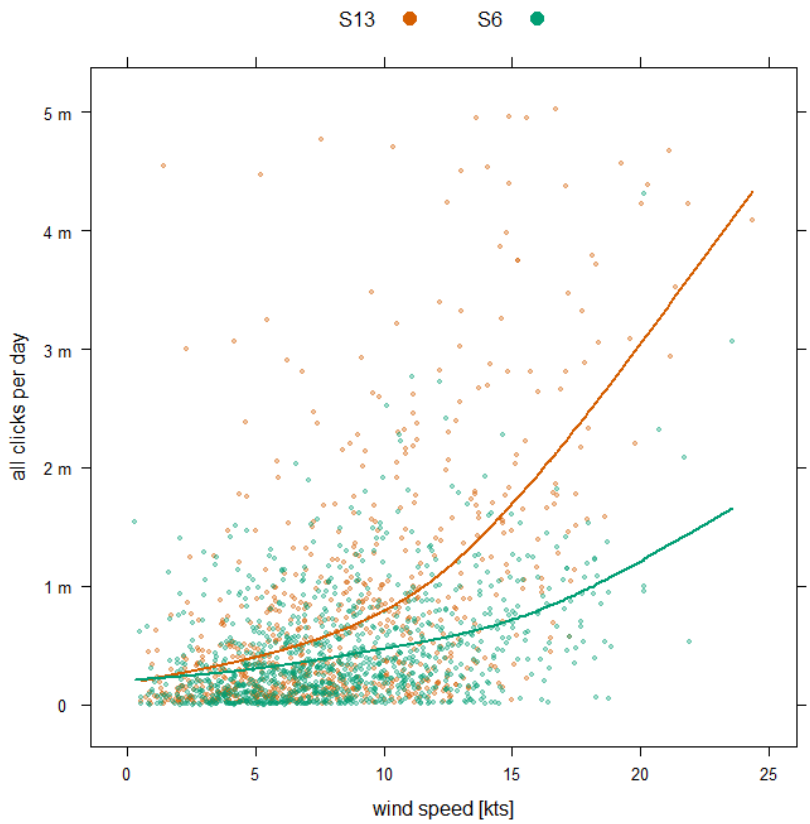


Figure 5-2 *Relation of wind speed and number noise of clicks recorded per day to average water depth illustrated with stations S13 (water depth of 14.34 m) and S6 (water depth of 43.64 m).*

The variable “all clicks” was identified as an important explanatory variable due to its strong correlation with the number of porpoise detections (see also chapter 4 on hourly data page 18). For stationary C-PODs, a limit was set for the number of clicks that can be detected per minute. This click limit was set at 4,069 clicks per minute. The maximum number of recorded clicks per day should therefore not exceed 5,859,360 total clicks per day (4,069*60 minutes*24 hours). This limit was exceeded in very few cases. If this limit was exceeded by more than one per cent, it indicates deviating settings (no click limit; resulted in a maximum 59 million of total clicks; smaller than 2 % of the data) and those data were excluded. Figure 5-3 illustrates the relationship between all clicks and the number of detected porpoises for the daily dataset. Two different effects act on the relationship “number of all clicks” and “number of porpoises detected”: 1) If no clicks were detected, then there was no porpoise detected. 2) If too much environmental noise is recorded, it becomes more difficult to detect porpoises. As we only wanted to correct for the technical limitations, the number of all clicks was set to a minimum value of 2.30*105 clicks per day (analogous to the hourly data). This was applied to 52.07 % of the data. If the number of all clicks became too large, however, then the technical limitations were predominant over the actual porpoise activity. Therefore, we excluded all data with a noise level of more than 5,160,000 clicks per day (right vertical line Figure 5-3). This is the case if the click limit was reached in nearly every minute of a day (maximum 5,898,240 clicks per day).

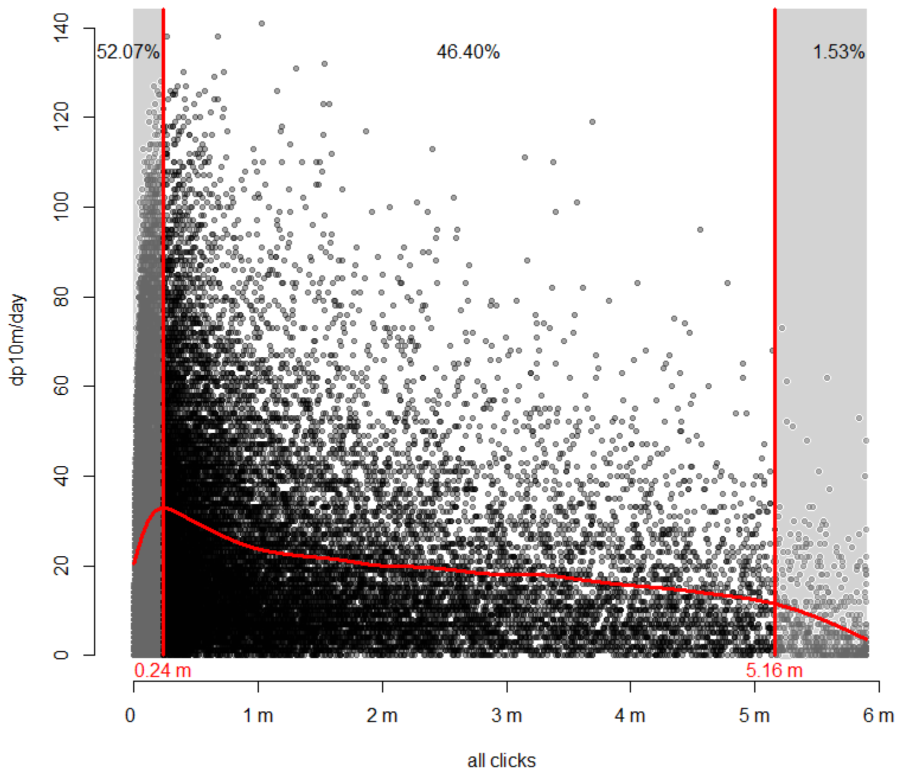


Figure 5-3 Detection positive ten minutes per day in relation to the total number of clicks detected. The red line visualises a smoothed spline fitted to the data whose characteristic were used to determine the limit with maximal porpoise detection rate(left vertical red line) and the limit above which data shall be discarded (right vertical red line) in the future. For data details of data treatment be-low and above red lines see text.

Geographic variables

Variables such as water depth, biozone or substrate are environmental variables that are – at least during the time frame of our study - constant over time. However, this does not make them less important explanatory variables as they describe the habitat found at the respective geographical position and thus influence e.g. the distribution of various fish species which form the diet of porpoises. These variables cannot be incorporated simultaneously into the same model as they strongly correlate with each other since, e.g. deeper waters often are categorised as circalittoral and characterised by fine sediment. Therefore these variables were used to cluster the German EEZ into subareas (see following section) rather than including them into the models. To account for monitoring position dependent variance we included the random factor station into our models. This random factor is either the name of a POD station or a single stationary POD.

Subareas

Exploring the raw data we found that those stations showed seasonal patterns in acoustic harbour porpoise detections (Figure 5-4). Some groups of stations, mostly neighbouring ones, clearly had a similar pattern (Figure 5-4) meaning that it was necessary to account for intra-annual differences in space in our models. Since grouping stations based on porpoise detection and then using the result to model variation in acoustic porpoise detections would have increased the risk of a type I error (i.e. falsely assuming a significant difference) we decided on clustering stations based on geographical properties (see above and section 5.2.4).

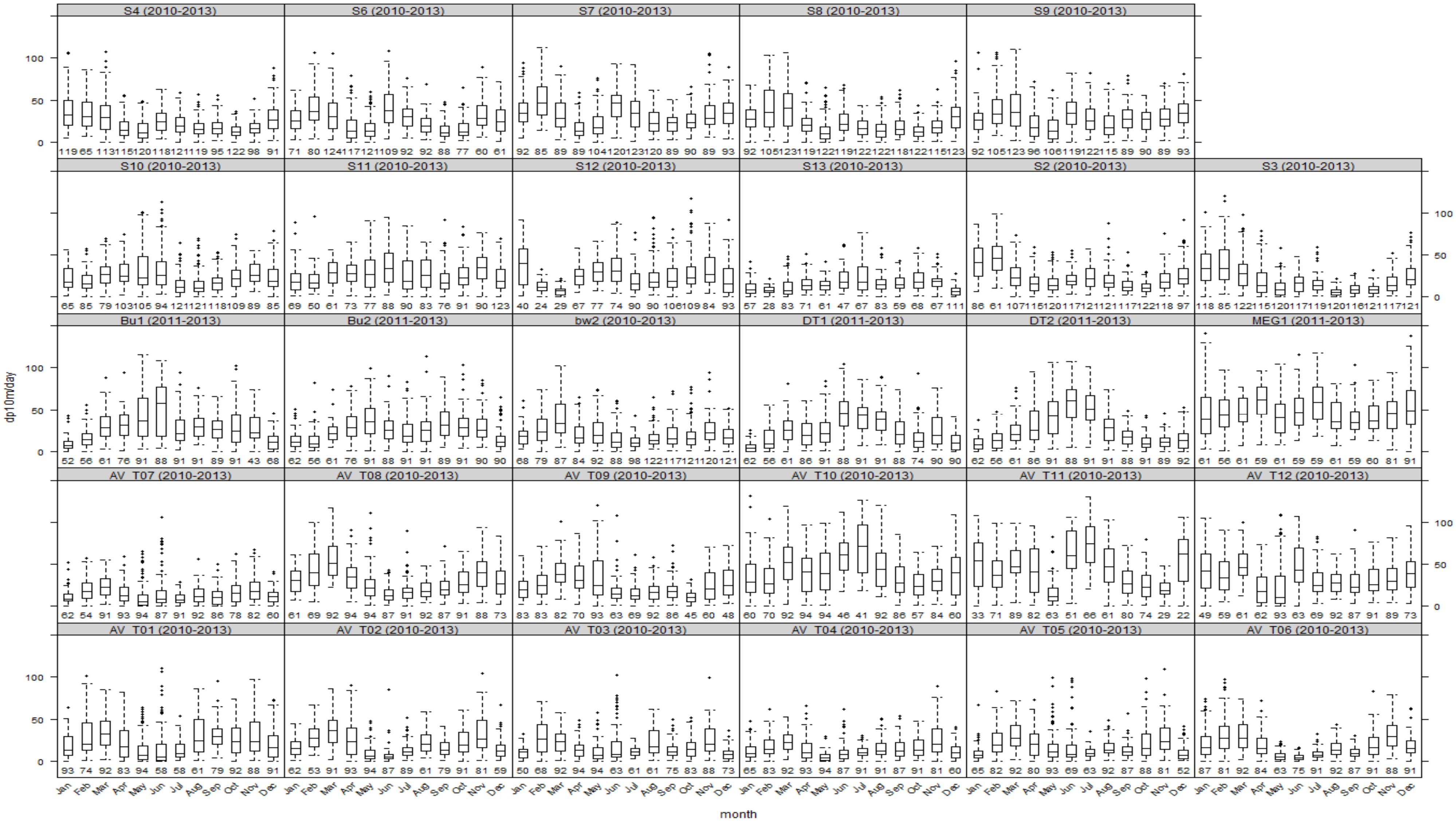


Figure 5-4 Monthly averages of dp10m/day (only POD positions are shown that recorded for at least three years). Numbers underneath each box indicate the sample size in days.

Used variables – an overview

This paragraph gives an overview (Table 5.1) over all variables used in our models. This includes all the above mentioned variables as well as custom made categorical variables designed to analyse the questions asked.

Table 5.1 Overview over variables used for modelling.

Variable	Type	Description
response		
dp10m		acoustic porpoise detections on a daily basis
piling related variables		
piling duration	continuous	gross piling duration per day including pauses without scarer application
pile driving day	categorical	days since pile driving: 0: pile driving day; 1: one day after; 2: two days after
min dist to pile	continuous	distance of POD to nearest active piling location
hour of piling start	discrete	hour of the day when piling started
piling year	categorical	combination of year and piling (yes/no) resulting in two levels per year: e.g. 2010_1: piling data in 2010; 2010_0: non piling data in 2010
time related variables		
year	discrete or categorical	monitoring year
day of year	discrete	the day of the year
environmental variables		
SSTA	continuous	anomaly from expected sea surface temperature
other variables		
subarea	categorical	subarea of German EEZ based (Figure 5-6)
station	categorical	monitoring location
POD ID	categorical	identification number of POD
all clicks	continuous	corrected number (Figure 5-3) of recorded all clicks per day
version	categorical	POD version: 0 or 1
subarea and season	categorical	combination of subarea and season: (12 levels; four seasons for each subarea)
consecutive	categorical	consecutive days of piling: 1: 1st day of piling; 2: 2nd day of piling; 3: 3rd day of piling
cumuCat	categorical	piling events per day: 1: one event; 2: two events
project	categorical	wind farm project

5.2.3 Question-related data subsets

For different questions a different data base was needed. Invalid or poor data, such as data where PODs had not recorded entire days or the maximum noise level was exceeded, were excluded from all datasets.

The origin of all other data-subsets mentioned below is the total dataset. It comprises 42,274 rows (each from one daily POD position).

The baseline dataset is the subset of the total dataset holding only data uninfluenced by any known short time effect of pile driving activities. This means, if any pile driving took place in any part of the German EEZ, the data from this day as well as from one day prior and up to two days post were discarded. This dataset was used to evaluate what the driving forces of acoustic harbour porpoise detections are, apart from pile driving.

The yearly-trends dataset holds all data except those where piling took place farther away than 20 km but closer than 60 km. All data derived from PODs farther away than 60 km from a piling site were considered as uninfluenced while all up to 20 km were considered as influenced.

The context-specific datasets are merely subsets of the yearly-trends dataset with respect to season (all data during winter (December, January, February), spring (March, April, May.), summer (June, July, August) and autumn (September, October, November)) or subarea (BWII-area, BARD-area, DanTysk-area and MSO-NSO-area) or only for the day of pile driving. This was done to avoid the statistical models becoming too complex to compute.

The accumulation datasets hold only pile driving days. All other data including one day prior to the piling day and up to two days after piling were discarded. Piling events which took place over midnight were also discarded. It was made sure that all piling events included in the datasets were uninfluenced by any piling event in neighbouring wind farms (Figure 5-5). Finally, the simultaneous dataset holds only data with one or two pile driving events per day and the consecutive dataset holds only data with one pile driving event per day.

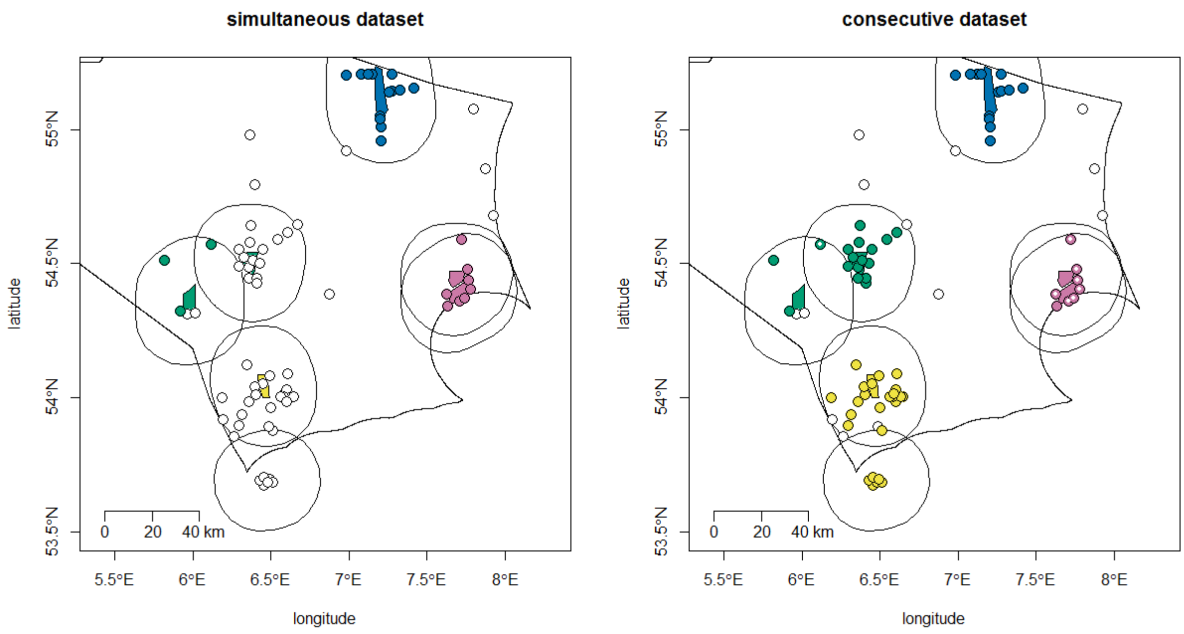


Figure 5-5 Visualisation of accumulation datasets. Wind farms which were considered to influence each other in terms of pile driving are indicated by the same colour. White circles indicate stationary POD positions not included in the respective dataset while coloured circles indicate their usage and the colour refers to the wind farm. Coloured dots with small white dots in them indicate that the POD position was used for both wind farms (though not at the same time).

5.2.4 Methodological choices for data analyses

For analysing and visualising data the script language R (R Development Core Team 2015) was used with the help of functions from several packages, like lattice (LATTICE 2008), latticeExtra (LATTICEEXTRA 2013), rgdal (RGDAL 2015), rgeos (RGEOS 2015), shapefiles (SHAPEFILES 2013), maps (MAPS 2014) and sp (SP 2015) for plotting and or handling geographical data. The libraries cluster (MAECHLER et al. 2015), mgcv (MGCV 2015), tseries (TSERIES 2015) and forecast (FORECAST 2015) were needed for statistical analyses. All model outputs were plotted using a self-made plotting function based on the plotting function provided by the library mgcv (MGCV 2015).

Subarea Definition

Clustering POD position is necessary to take into account differences in porpoise occurrence between regions within the North Sea. Clustering POD positions (stations) based on porpoise detection and then using the result to model variation in acoustic porpoise detections increases the risk of a type I error, assuming a significant difference when in fact there is none. In order to overcome this circular reasoning, we decided to cluster the stationary POD-positions based on their geographic properties, assumed to be independent from acoustic porpoise detections.

For clustering POD positions we used the average water depth, the sediment type, the biozone category as well as latitude and longitude (transformed into the relative variables easting and northing). We decided on the partition clustering method pam() from the package cluster (MAECHLER et al. 2015). Partition clustering of the geographical properties of stationary POD positions resulted in four different subareas of the German EEZ (Figure 5-6). The subarea around Riffgat and north of Borkum (subarea 1 in green), the Bard and GTI area (subarea 2 in yellow), the subarea around Helgoland (subarea 3 in blue) and the area of the DT wind farm project (subarea 4 in pink) (Figure 5-6). These results were used in many of our models to differentiate between the subareas.

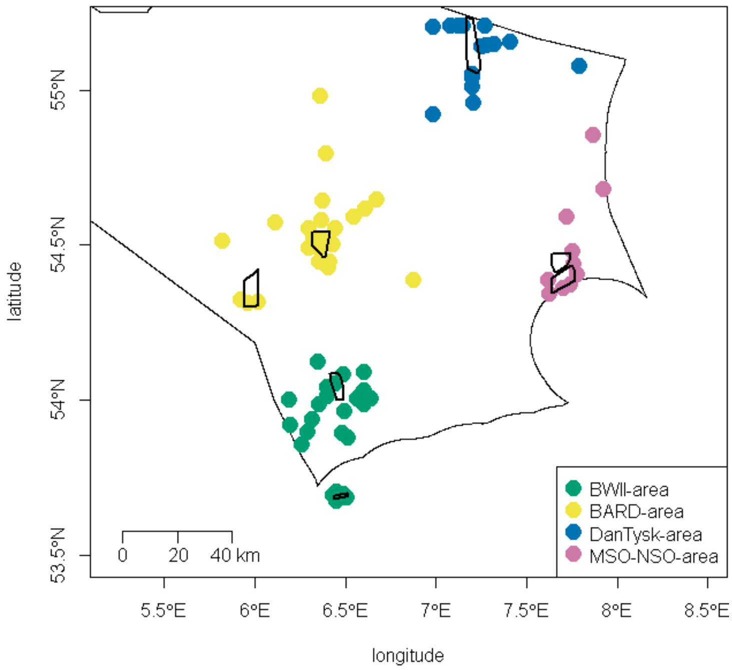


Figure 5-6 Subarea classification of stationary POD positions with wind farm outline.

Generalised Additive Mixed Models

For modelling our data we used the function gamm(), with a quasi corrected poisson error distribution to account for overdispersion, of the package mgcv (MGCV 2015). In contrast to analyses based on hourly POD-data we did not need the bam() function for very large datasets. The advantage of the gamm() function was the possibility to correct for temporal autocorrelation with an Auto Regressive Moving Average (ARMA) process implemented in the function corARMA(). Model residuals were tested for autocorrelation on population level as well as on the correct nested level. Then those residuals were tested for stationarity (Augmented Dicky-Fuller Test (ELLIOTT et al. 1996) and Phillips-Perron Unit Root Test (PHILLIPS & PERRON 1988); library tseries (TSERIES 2015)). All residuals tested were stationary and thus no order of differencing needed to be included into the correlation structure validating the use of an ARMA structure. The most parsimonious orders of p (AR-term) and q (MA-term) were estimated using the function auto.arima() of library forecast (FORECAST 2015). Adequate orders were then chosen on this estimation and their suitability in accounting for autocorrelation in the models evaluated by sight using the acf() and pacf() functions of package stats (R Development Core Team 2015).

Spatial autocorrelation can be understood as the correlation of properties of data with distance of data points (Figure 5-7). The magnitude as well as the direction of spatial autocorrelation may change with distance (Figure 5-7). In our data there was no need to correct for spatial autocorrelation which is shown in Figure 5-8 based on Moran's Index (MORAN 1950) analyses of the baseline model residuals. The analyses were done following the procedure presented in DORMANN et al. (2007).

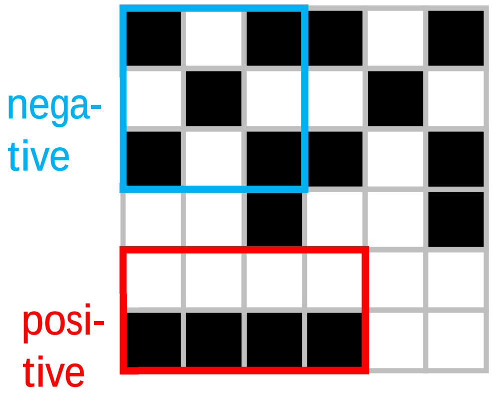


Figure 5-7 *Extremes of spatial autocorrelation with reference to Moran's Index (I). If properties of data are completely clustered then Moran's I takes on positive values, if they are wholly dispersed then Moran's I becomes negative and if the properties are randomly distributed Moran's I equals zero. Note that magnitude and direction of autocorrelation may change with distance.*

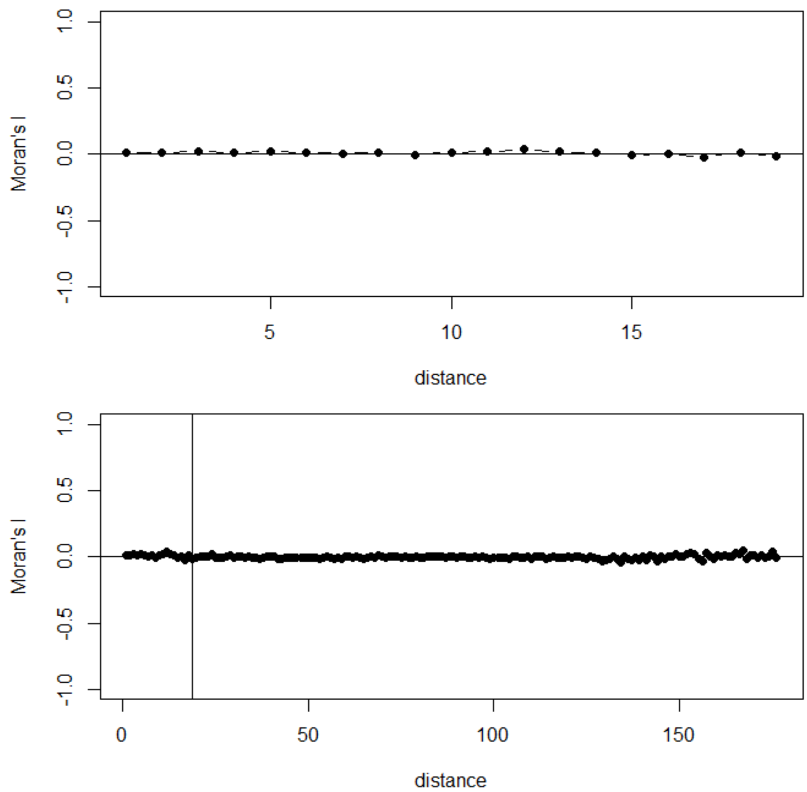


Figure 5-8 *Spatial autocorrelation of baseline model residuals. Moran's I can take on values between -1 and 1. If there is no strong autocorrelation present in the data, then Moran's I would be close to 0 which is the case here.*

The best model structure was chosen based on the lowest models AIC as well as on model diagnostic plots using the function gam.check() of library mgcv (MGCV 2015).

5.3 Results

In the following section, we present the results of the analyses run on the daily-POD-data.

5.3.1 Relation of natural parameters to acoustic harbour porpoise detections

Initially, a baseline model was designed to get an insight into the most important natural forces driving harbour porpoise acoustic activity throughout the German EEZ. Furthermore, this baseline analysis enabled us to make a first statement about possible long term effects of pile driving activities on acoustic harbour porpoise detections. For this purpose, only data not affected by known short term effects of piling were included. Finally, we decided on the model structure (presented in Table 5.2) based on the goodness of fit of the models (using the AIC as a criterion), whilst avoiding to use intercorrelated variables.

Table 5.2 Parameters of the baseline model.

Variable name	Variable type in model	Purpose
dp10m	Response	
year	Smooth by subarea	Yearly trends
day of year	Cyclic smooth by subarea	Seasonal patterns
all clicks	Smooth	Correct for technical POD artefact
SSTA	Smooth	Effect of temperature anomalies
subarea	Factor	Effect of region
station	Random factor	Effect of deployment position
POD ID	Random factor nested in station	Sensitivity differences in PODs
Error distribution	Quasi corrected Poisson	
Autocorrelation	ARMA on population level (p=1,q=0)	Ensure that requirements for running the model are fulfilled

Seasonal patterns and the magnitude of acoustic harbour porpoise detections differed among subareas (Figure 5-9). Data from PODs deployed in the DanTysk-area, showed a single and very defined peak during spring and summer (Figure 5-9). Acoustic harbour porpoise detections in the BWII-area and the BARD-area were very similar in pattern but intrinsically different in the magnitude of the fluctuations, which were much more defined in the BARD-area (Figure 5-9), and in the average level of detections (Figure 5-10). With increased pile driving activities from 2010 to 2013, acoustic porpoise detections either increased or remained constant over this period (Figure 5-11). The raw data also revealed these patterns, although there they were not as defined (Figure 5-12). Common to all four subareas was that the main proportion of acoustic harbour porpoise detections per day were beneath 50 dp10m (coloured boxes in Figure 5-12), which was less than 36 % of possible dp10m per day. The four subareas differed, however, in their estimated overall acoustic porpoise detections, with the BARD-area showing the least acoustic detections and DanTysk-area and MSO-NSO-area showing the most acoustic detections (Figure 5-10).

In addition, the model showed that harbour porpoise acoustic detections were highly dependent on the noise-level with click characteristics in the surrounding water column (upper left plot in Figure 5-10). This is, however, not a biological but mostly a technical effect since too many noise clicks confound porpoise clicks. Even in the raw data this strong relation could be seen (section

5.2.2). Biologically interesting is that acoustic harbour porpoise detections decrease with a negative SSTA (negative anomaly from the expected sea surface temperature; upper right plot in Figure 5-10). Although temperature dependent behaviour of unknown magnitude must be assumed (www.chelonia.co.uk and personal communication) this effect cannot be assumed to be technical but must predominantly originate from a biological cause: First of all, we only considered an anomaly from the expected temperature and not the actual sea surface temperature (SST) and secondly, even with SST and acoustic harbour porpoise detections we did not find any suspicious correlation in our data (Figure 5-13). In summary it can be said that acoustic harbour porpoise detections were rather low and variations therein could be explained by porpoise detections greatly depending on water temperature anomalies, the area and year looked at, season and thus related parameters like water temperature, day length, wind speed etc.

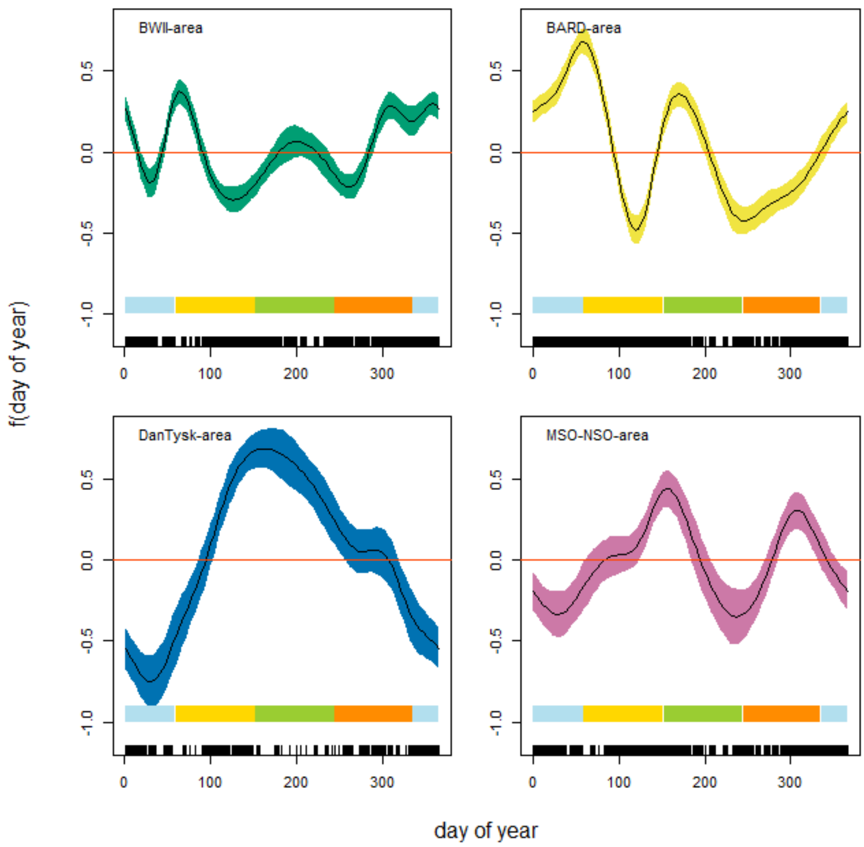


Figure 5-9 *Subarea specific seasonal patterns of acoustic porpoise detections from 2010 - 2013. The colours of the confidence intervals correspond to the specific subarea (section 5.2.4). Coloured horizontal bars on the bottom of each plot correspond to the four seasons: blue: winter; yellow: spring; green: summer; orange: autumn. Ticks at the bottom of the plots indicate data availability.*

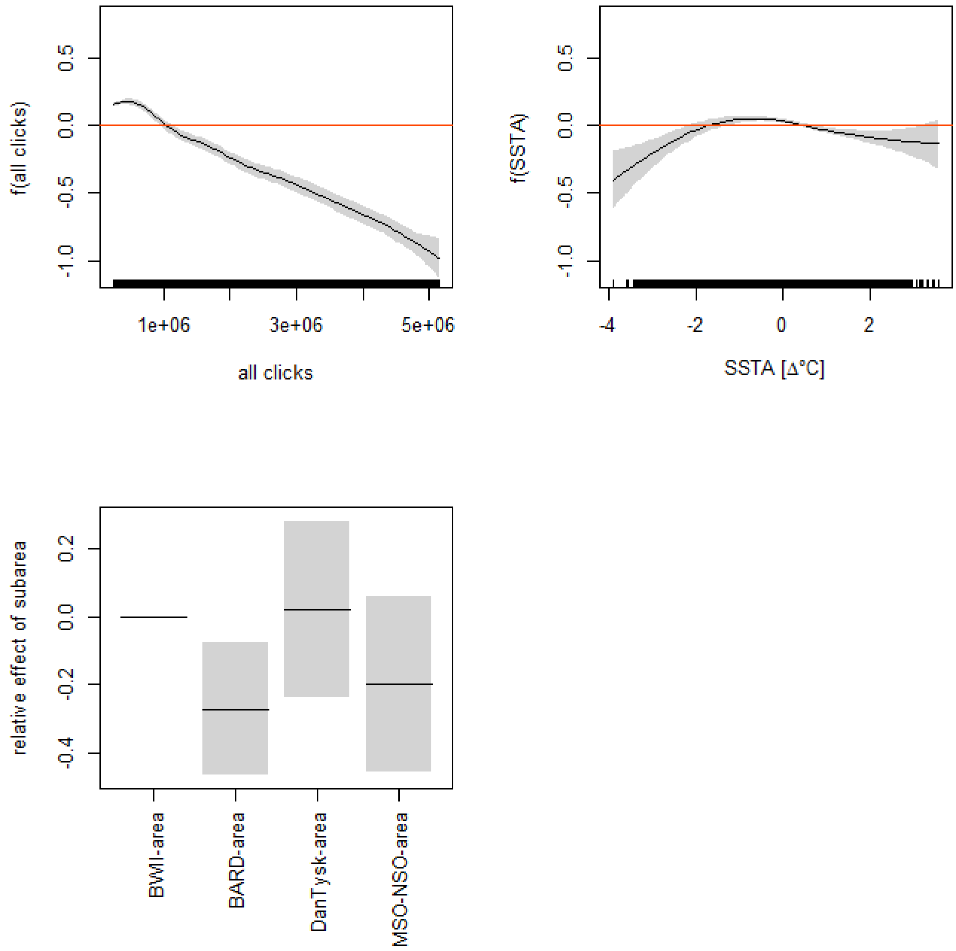


Figure 5-10 Effects of noise, SSTA and subarea on acoustic porpoise detections predicted from the baseline model. Grey shaded areas indicate the confidence intervals. It shows that the louder the environment was the fewer porpoise detections were made, which can mostly be considered a technical artefact of the PODs not being able to distinguish noise clicks from porpoise clicks. When deviation from the expected average sea surface temperature SST occurred, a decline in porpoise detections was observed. This effect was stronger if the water was colder than usual. The four subareas differed in the acoustical porpoise detections made, with the BARD-area having the fewest detections. All other subareas are considered statistically equal.

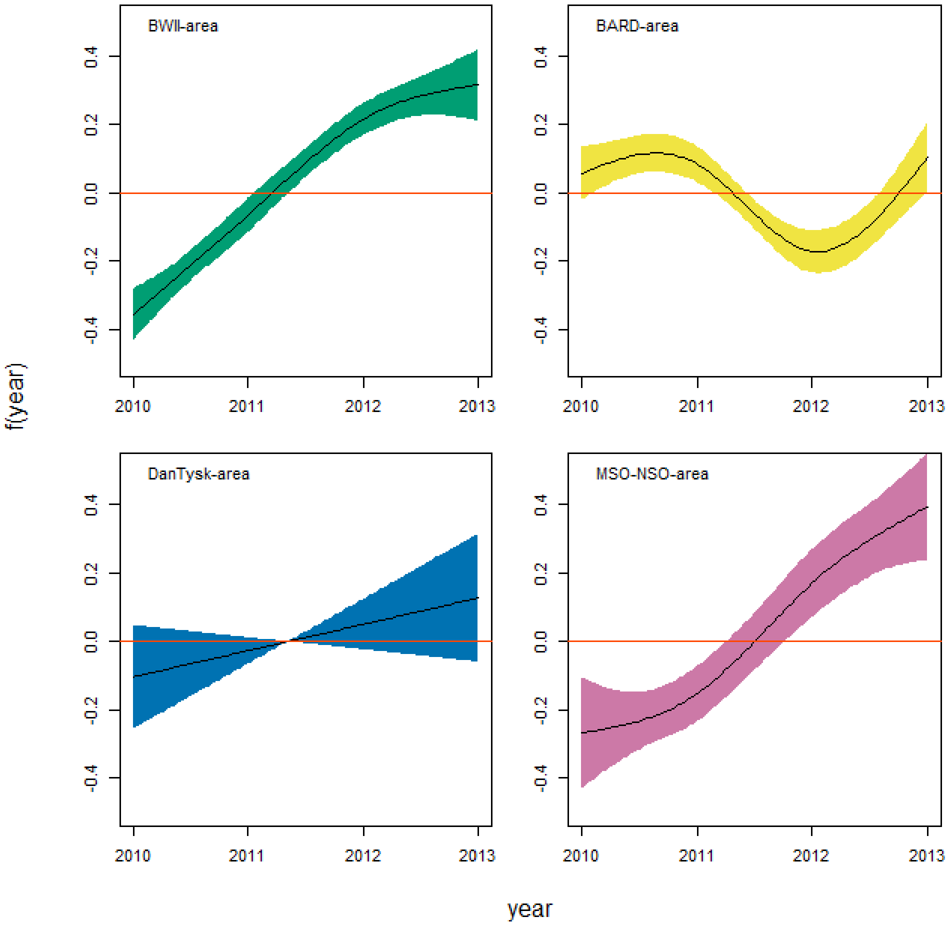


Figure 5-11 *Subarea-specific yearly trends of acoustic porpoise detection. The colours of the confidence intervals correspond to the specific subarea (section 5.2.4). The graph shows, that over the years acoustical porpoise detections either increased or remained constant.*

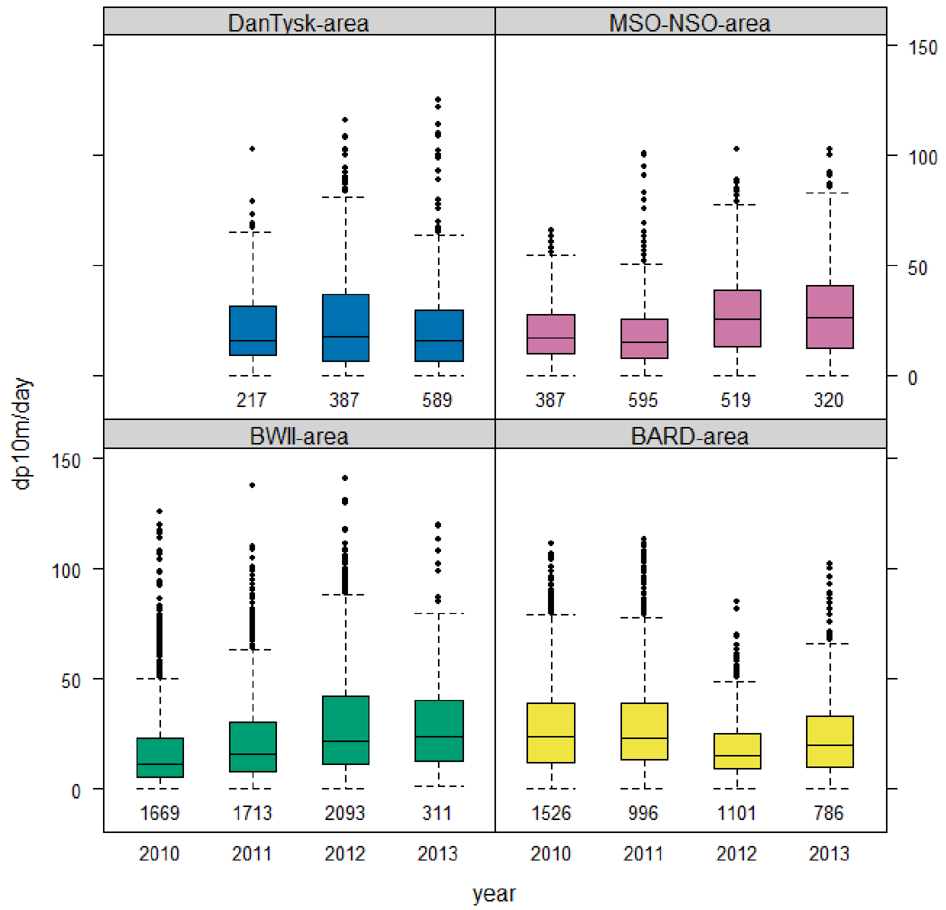


Figure 5-12 *Boxplot of acoustic harbour porpoise detections pooled by year and subarea. The median is represented by the black horizontal line within the respective box, dashed lines indicate the 25 % and 75 % quantiles and the box indicates the 50 % range around the median.* Outliers are marked by dots. *Numbers underneath each box indicate the sample size in days.*

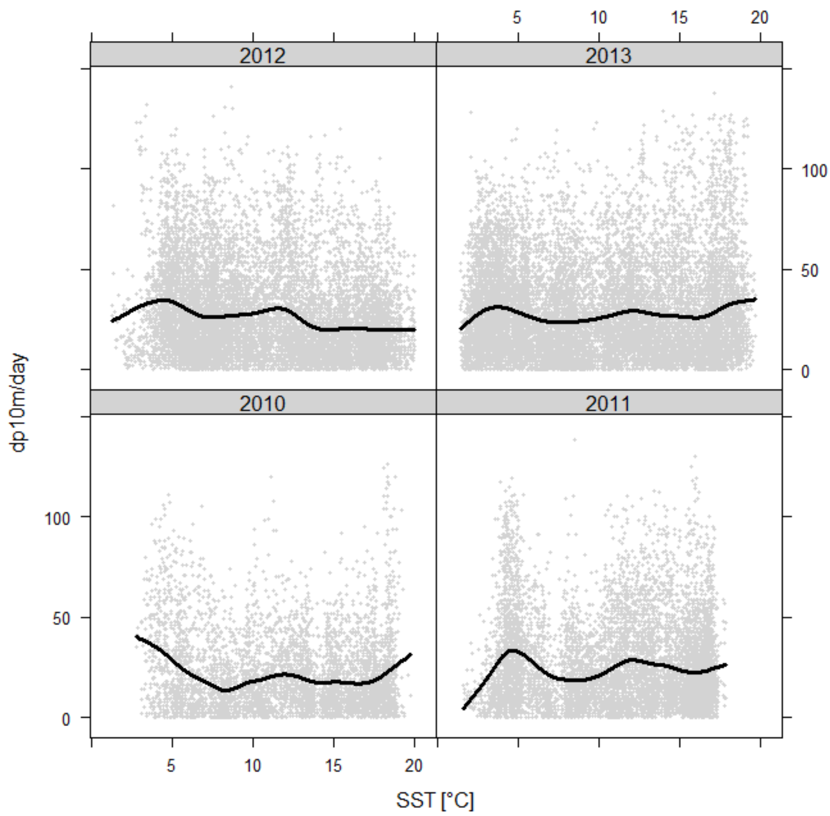


Figure 5-13 There were no signs of severe temperature dependence in acoustical porpoise detections. If there were a POD-related temperature dependency of acoustical porpoise detections stronger than any biological or random effect in our data it would be shown by plotting dp10m/day against SST.

5.3.2 Detectable effects of pile driving activities on acoustic porpoise detections

Using PODs, harbour porpoise acoustic activity can be monitored in the surrounding water and thus effects of pile driving are potentially detectable through a change in acoustic porpoise detections. In all our analyses we found that data influenced by piling activities showed fewer acoustic harbour porpoise detections than data which was assumed to be uninfluenced.

Long term effects and habituation

Apart from short term effects, like animals temporarily leaving an area due to pile driving activities, also long term effects may be expected. Possible long term effects, detectable with our kind of data, would be the altered use of an area by harbour porpoises, e.g. using an area to a lesser extent or habituating to the increased noise levels. Fundamental changes in their clicking behaviour might be less likely. With our data these long term effects may only be seen by altered acoustical porpoise detections.

Assuming that porpoises did not change their clicking behaviour nor were forced to migrate to the German EEZ for some reason, no negative long term effects occurred since the beginning of piling events in the North Sea, as acoustic detections either increased or remained constant (Figure 5-11 and Figure 5-14). If harbour porpoises habituated to short term effects of piling this would be seen in a more rapid increase than with unaffected data and if they became more sensitive to piling it would be seen in a more reluctant increase or even a decrease in acoustic harbour porpoise detections during pile driving days over the years. We thus formulated a model (Table 5.3) where a factor variable was included to account for piling (yes if piling took place up to 20 km from the monitoring site or no if no piling occurred within a distance of 60 km to the monitoring site; data within less than 60 km but further away than 20 km to an active pile driving site was excluded) and the year where data had been recorded. For DanTysk-area and MSO-NSO-area, it was impossible to identify an effect over the years or to characterise its direction (Figure 5-14) due to reduced data availability (Figure 5-15). For the other two subareas, the results showed only a weak indication of habituation and no sign of increased sensitivity, as the yearly trend was almost similar between piling and non piling days (Figure 5-14).

Table 5.3 Parameters of the yearly trends model. One model for each subarea.

Variable name	Variable type in model	Purpose
dp10m	Response	
day of year	Cyclic smooth	Seasonal patterns
all clicks	Smooth	Correct for technical POD artefact
SSTA	Smooth	Effect of temperature anomalies
piling year	Factor	Piling related yearly trends
station	Random factor	Effect of deployment position
POD ID	Random factor nested in station	Sensitivity differences in PODs
Error distribution	Quasi corrected Poisson	
Autocorrelation	ARMA on population level (p=1,q=0)	Ensure that requirements for running the model are fulfilled

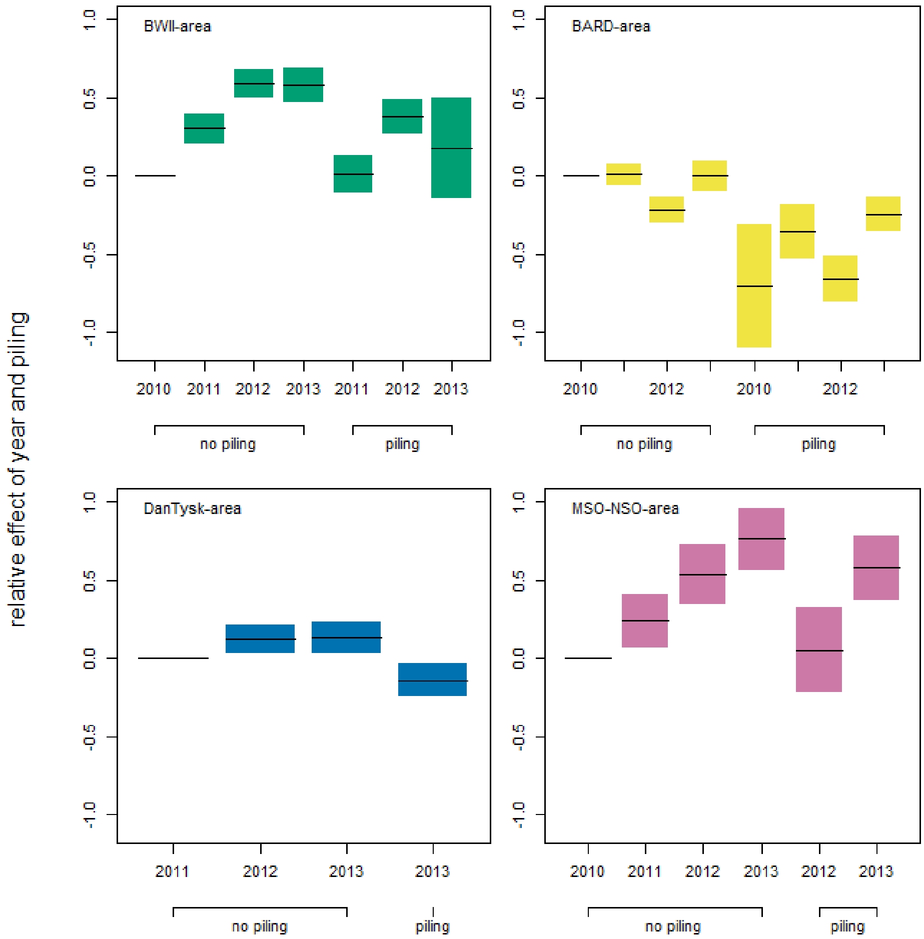


Figure 5-14 Relative effect of year and piling on acoustic porpoise detections for each subarea. Coloured areas indicate the confidence intervals while the colours refer to the subarea (section 5.2.4). Data was considered to be influenced if a piling event took place in a radius of 60 km around the POD. The observed patterns for piling and non piling data were similar within each subarea throughout the years.

When interpreting the model output in detail it has to be kept in mind that the amount of data, especially with respect to piling influenced data, was not comparable among subareas and years (Figure 5-14 and Figure 5-15). Only piling influenced data with a distance between monitoring sites and piling events of less than 20 km were used for modelling, whilst data further away than 20 km but closer than 60 km was discarded due to the uncertainty how to categorise them. Thus in some years the data basis was poor (e.g. BWII-area year 2013; Figure 5-15). Nothing can be said about long term effects in the DanTysk-area since piling data was available only for the year 2013 when piling took place in DT itself. The sample sizes of piling data were rather small in 2013 (BWII-area), 2010 (BARD-area) and 2012 (MSO-NSO-area) (Figure 5-15), which led to large confidence intervals (Figure 5-14). In total the increase in acoustic harbour porpoise detections can be considered steeper for piling than non-piling days for the years 2011 and 2012 in the BWII-area and the years 2012 and 2013 in the BARD-area (Figure 5-14). This may be an indication for habituation of porpoises to piling activities (Figure 5-15).

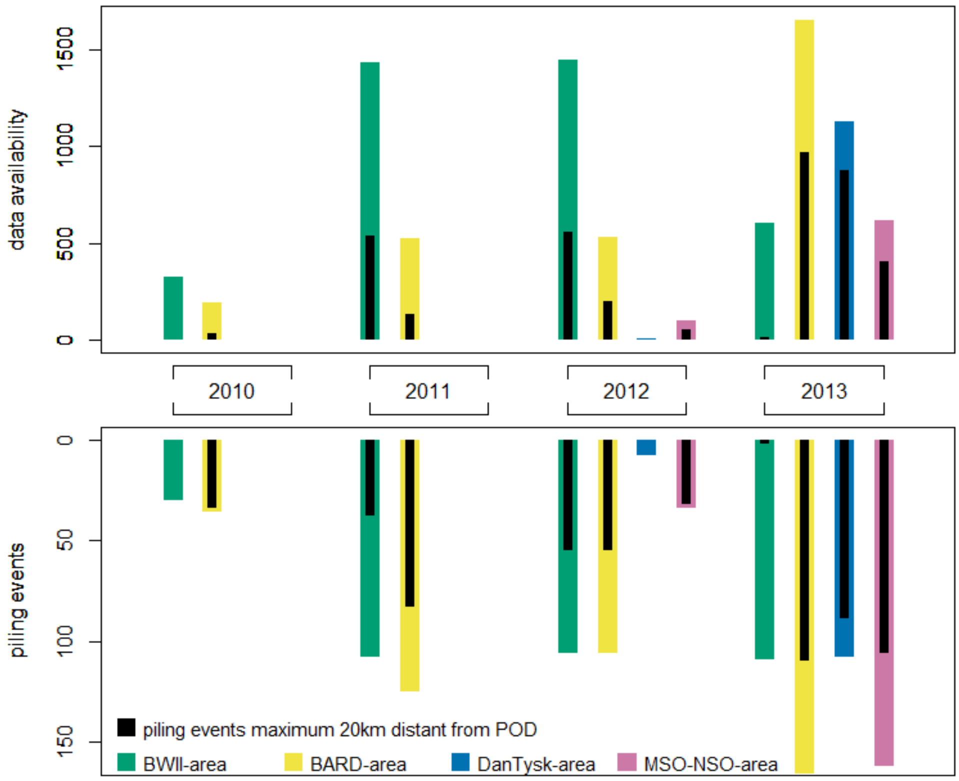


Figure 5-15 Piling events and data availability per year and subarea. Coloured bars indicate pilings (lower graph) or data availability (upper graph) in a 60 km radius around the PODs. The overlying black bars provide the same information for the respective subarea for a radius of 20 km around the PODs.

Context specific effects: is the effect of piling related to the baseline level of acoustic porpoise detections?

With the baseline model, we showed that acoustic porpoise detections varied over the course of a year, between years and between subareas. These findings are consistent with the work by VERFUSS et al. (2007). Since harbour porpoises are not evenly distributed over the German EEZ, we wanted to evaluate whether the short term effect of pile driving events on the day of piling and the recovery time were context-dependent. We therefore had a closer look at piling effects on a seasonal basis with respect to the four subareas we defined and vice versa. Due to intrinsic properties of the dataset, like the clumped distribution of piling events throughout the seasons (e.g. piling activity is generally higher during summer than during winter), we were not able to run a comprehensive analysis on the overall dataset but generated subsets either with respect to season or subarea. Again for the same reasons it was not possible to run exactly the same model on every subset-family (season or subarea) (Table 5.4 and Table 5.5). However, we incorporated the key elements into every one of those models (Table 5.4 and Table 5.5).

Season dependent effects of pile driving

To check whether the recovery time after the end of piling might be dependent on season, we ran a model on a data subset comprising the day of piling and up to two days after, including a factor pile driving day with three levels: 0 for the day of piling, 1 for one day after and 2 for two days after (Table 5.4). The pattern we found was almost identical throughout the seasons (Figure 5-16): Throughout the year days with piling events showed significantly lower acoustic porpoise detections than the following two days (Figure 5-16). From the first to the second day after piling the magnitude of detections did only change significantly for winter and autumn (estimated smooth terms do not overlap with confidence intervals of the other levels) but not for spring and summer (Figure 5-16).

Table 5.4 Parameters of seasonal models on piling and two days after data with min dist to pile. One model for each season.

<table>
<tr><th>Variable name</th><th>winter</th><th>spring</th><th>summer</th><th>autumn</th><th>Purpose</th></tr>
<tr><td>dp10m</td><td colspan="4">Response</td><td></td></tr>
<tr><td>all clicks</td><td colspan="4">Smooth</td><td>Correct for technical POD artefact</td></tr>
<tr><td>SSTA</td><td colspan="4">Smooth</td><td>Effect of temperature anomalies</td></tr>
<tr><td>year</td><td colspan="4">Factor</td><td>Piling related yearly trends</td></tr>
<tr><td>min dist to pile</td><td colspan="4">Smooth by subarea</td><td>Difference in distance effect</td></tr>
<tr><td>piling duration</td><td colspan="2">Smooth by subarea</td><td>Smooth</td><td>Smooth by subar-ea</td><td>Estimate effect of pile driving time</td></tr>
<tr><td>pile driving day</td><td colspan="4">Factor</td><td>Days since pile driving (three lev-els: 0: pile driving day; 1: one day after; 2: two days after)</td></tr>
<tr><td>subarea</td><td colspan="4">Factor</td><td>Effect of region</td></tr>
<tr><td>version</td><td>-</td><td colspan="3">Factor</td><td>Correct for different POD versions</td></tr>
<tr><td>station</td><td colspan="4">Random factor</td><td>Effect of deployment position</td></tr>
<tr><td>Error distribution</td><td colspan="4">Quasi corrected Poisson</td><td></td></tr>
<tr><td rowspan="2">Autocorrelation</td><td colspan="4">ARMA on piling level</td><td rowspan="2">Ensure that requirements for run-ning the model are fulfilled</td></tr>
<tr><td>p=0, q=1</td><td>p=1, q=0</td><td>p=1, q=1</td><td>p=1, q=1</td></tr>
</table>

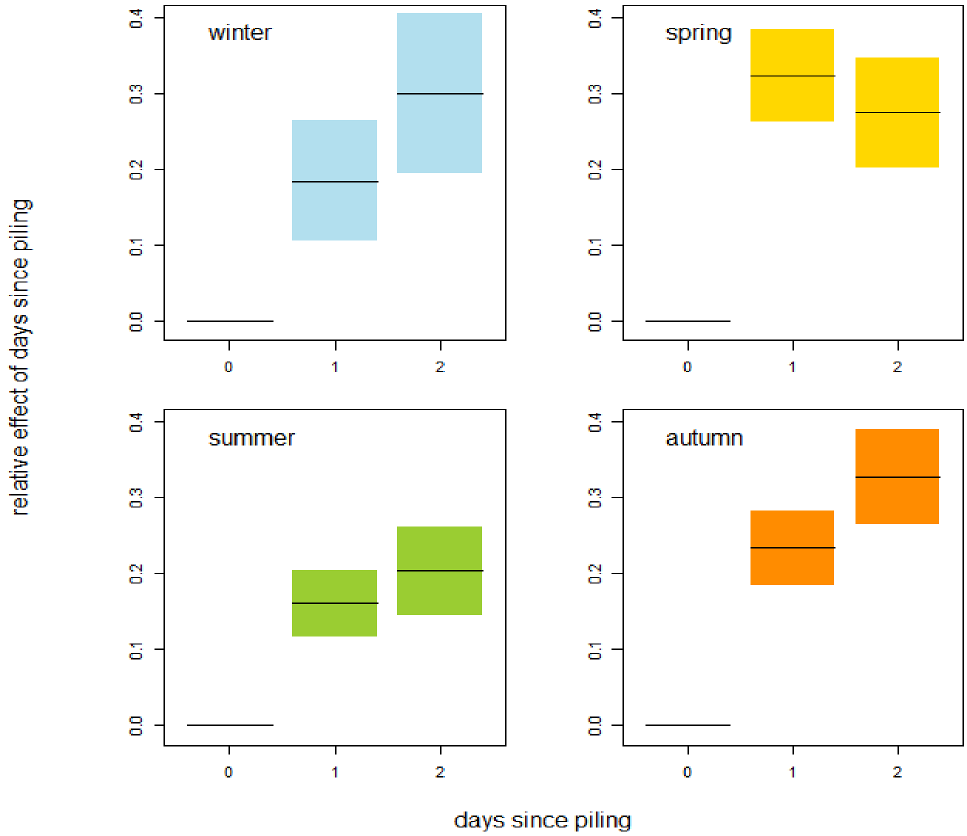


Figure 5-16 *Seasonal differences in the effect of pile driving on piling day and up to two days after. Pile driving day itself (factor level 0) is seen as reference value for the two other levels, meaning, that the relative magnitude of these only relates to the respective reference level per season. Note: Take care not to compare the actual magnitude among seasons.*

Subarea dependent effects of pile driving

Investigating matters further for differences in effect size of piling between subareas and seasons we ran a second family of models on subsets for subarea. Here again, it was not possible to run the same model on all four subarea-subsets due to statistical reasons (Table 5.5).

For all subareas we found significantly lower acoustic harbour porpoise detections for the day of piling (level 0 in Figure 5-17) than for the two following days. Solely for the BWII-area a significant rise in acoustic harbour porpoise detections was found from the first to the second day after piling (Figure 5-17). A trend in the same direction was detected for the MSO-NSO-area but nothing of this kind for the DanTysk-area and the BARD-area (Figure 5-17).

Table 5.5 Parameters of subarea models on piling and two days after data with min dist to pile. One model for each subarea.

Variable name	BWII-area	BARD-area	DanTysk-area	MSO-NSO-area	Purpose
dp10m	Response				
all clicks	Smooth				Correct for technical POD artefact
SSTA	Smooth				Effect of temperature anomalies
year	-	Factor	-		Yearly differences
min dist to pile	Smooth by season				Difference in distance effect
piling duration	Smooth by season		Smooth		Estimate effect of pile driving time
pile driving day	Factor				Days since pile driving (three levels: 0: pile driving day; 1: one day after; 2: two days after)
season	Factor				Effect of season
version	-	Factor	-		Correct for different POD versions
station	Random factor				Effect of deployment position
Error distribution	Quasi corrected Poisson				
Autocorrelation	ARMA on piling level				Ensure that requirements for running the model are fulfilled
	p=1, q=0	p=1, q=0	p=1, q=1	p=1, q=0	

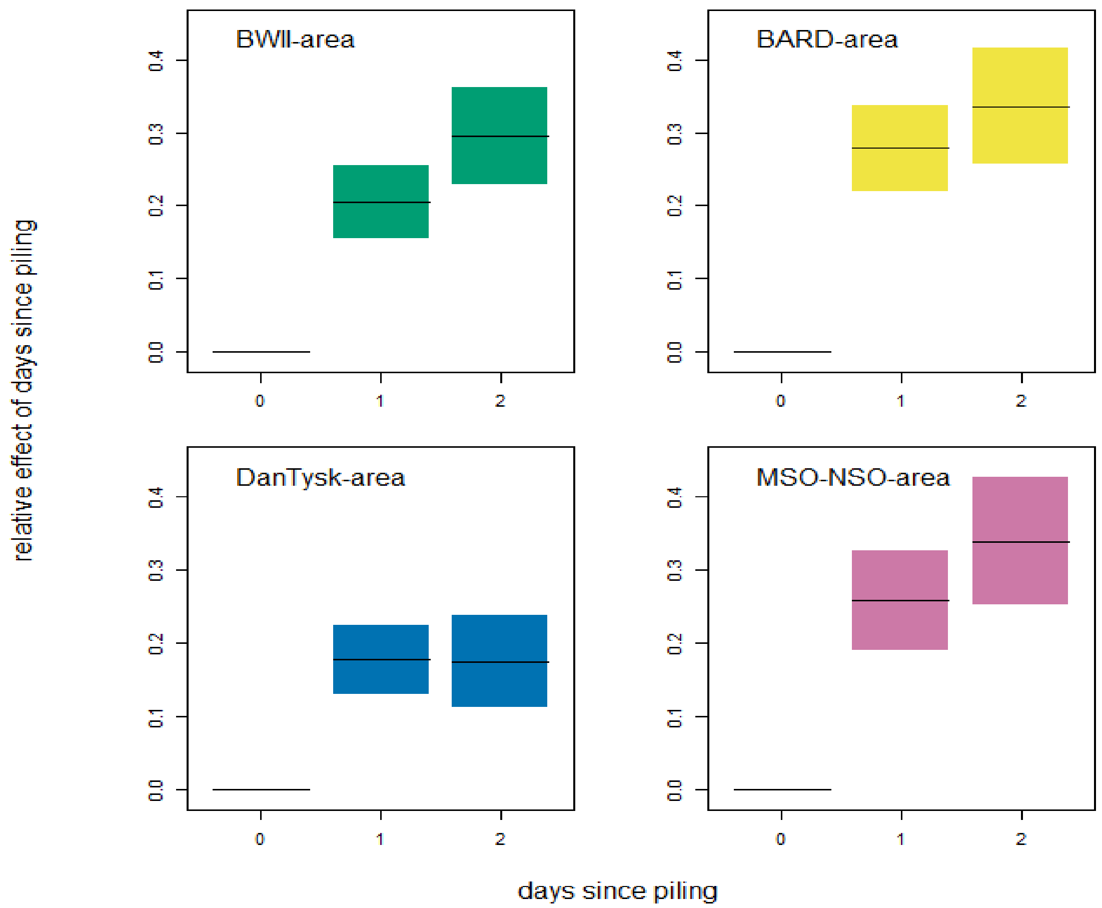


Figure 5-17 Difference among subareas in the effect of pile driving on piling day and up to two days after. Pile driving day itself (factor level 0) is seen as reference value for the two other levels, meaning, that the relative magnitude of these only relates to the respective reference level per subarea. We took care not to compare the actual magnitude among subareas.

Putting in context: Season and Subarea

More knowledge on why recovery time was estimated by the models run on season- and subarea-subsets can be gained by looking at the raw data with respect to subarea and season including reference data (Figure 5-18). The patterns seen in Figure 5-18 may not be generalised since we were not able to account for different stations and PODs whose contribution to the sample size in either levels of pile driving influence is likely to be biased. It is interesting that uninfluenced reference data in the MSO-NSO-area showed a tendency for higher acoustic porpoise detections than piling influenced data during the warmer months (Figure 5-18). This is in contrast to the other three subareas where uninfluenced data is not different from the second day after piling ceased (Figure 5-18).

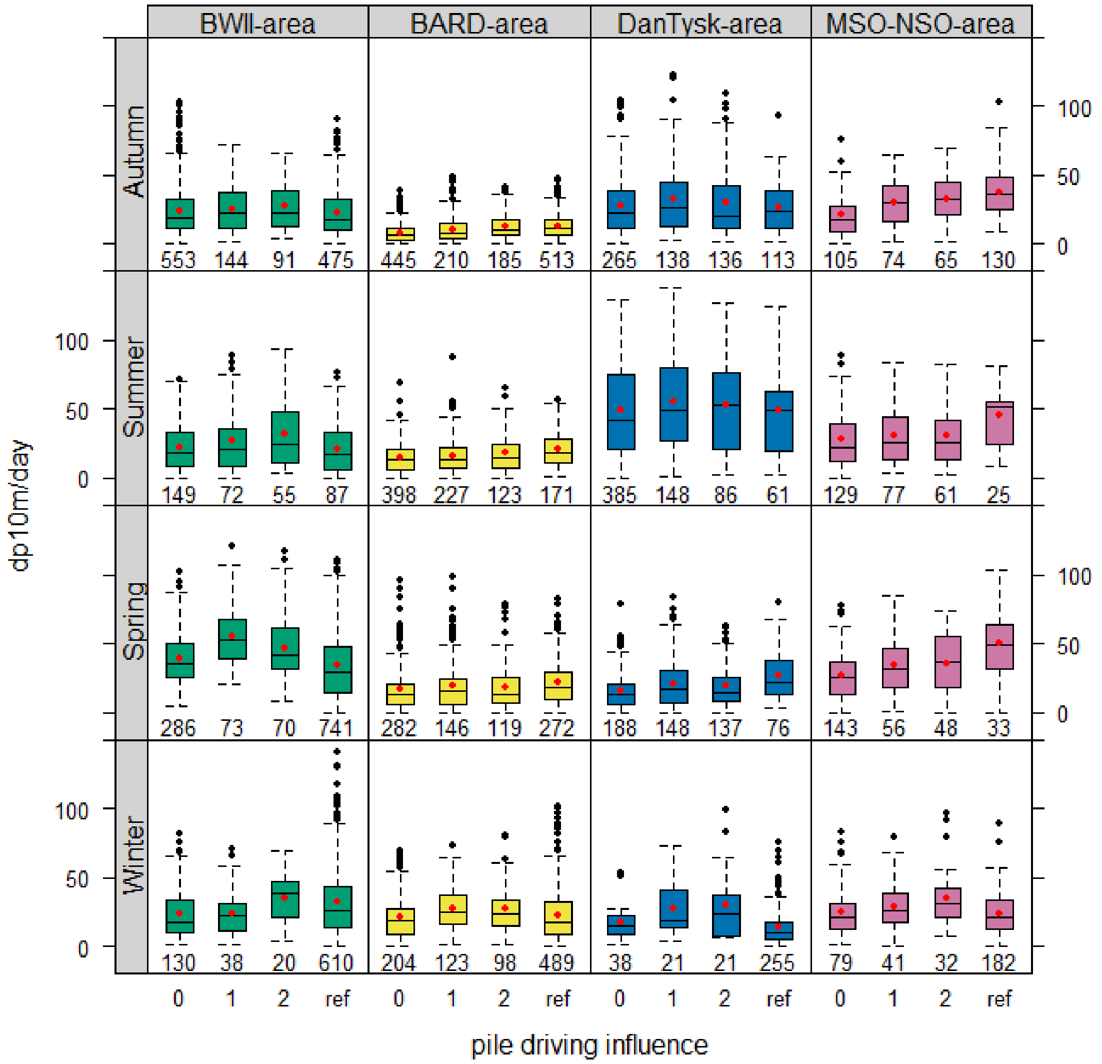


Figure 5-18 Acoustic harbour porpoise detections with respect to subarea and season under the influence of piling. The detections are shown with respect to pile driving influence: day of pile driving (level 0); one and two days post pile driving (levels 1 and 2) and uninfluenced reference data (level ref). The median is represented by the black horizontal line within the respective box and the mean by a red dot, dashed lines indicate the 25 % and 75 % quantiles and the box indicates the 50 % range around the median. Outliers are marked by black dots. *Numbers underneath each box indicate the sample size in days.*

Further knowledge may be gained by comparing acoustic harbour porpoise detections among subareas and seasons for the day of pile driving. We thus ran a model on the overall context-specific dataset for the day of pile driving only (Table 5.6). Very different patterns can be observed for the four subareas (Figure 5-19). Throughout the year acoustic harbour porpoise detections held its level, which was slightly higher than autumn in BWII-area, in the MSO-NSO-area (Figure 5-19). For all other subareas at least one season was significantly different than the others in that subarea (Figure 5-19). The highest numbers of detections during pile driving were found in spring and summer for the BWII-area and the DanTysk-area (Figure 5-19). The lowest numbers of detections were found in autumn in the BARD-area (Figure 5-19).

Table 5.6 Parameters of the context-specific model comparing acoustic harbour porpoise detections on pile driving day.

Variable name	Variable type in model	Purpose
dp10m	Response	
all clicks	Smooth	Correct for technical POD artefact
SSTA	Smooth	Effect of temperature anomalies
piling duration	Smooth	Duration of nearest piling event
min dist to pile	Smooth	Distance to nearest piling event
subarea and season	factor	Effect size with respect to subarea and season (12 levels; four seasons for each subarea)
station	Random factor	Effect of deployment position
POD ID	Random factor nested in station	Sensitivity differences in PODs
Error distribution	Quasi corrected Poisson	
Autocorrelation	ARMA on piling level (p=1,q=0)	Ensure that requirements for running the model are fulfilled

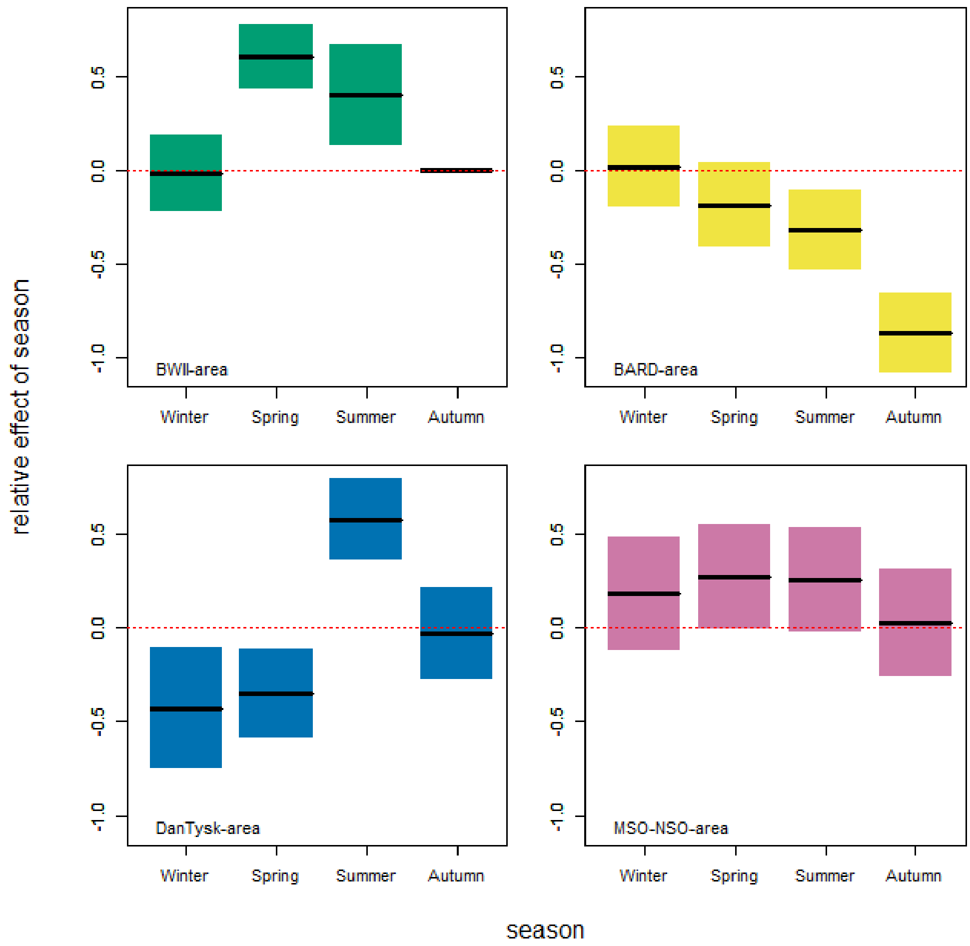


Figure 5-19 Modelled relative acoustic harbour porpoise detections on pile driving days with respect to subarea and season. Autumn in the BWII-area is seen as reference value (and thus is zero with no confidence intervals) for all levels of all subareas, meaning, that the relative magnitude of these only relates to the autumn in the BWII-area. Here it is requested to compare the magnitude of the various levels among each other.

Accumulation

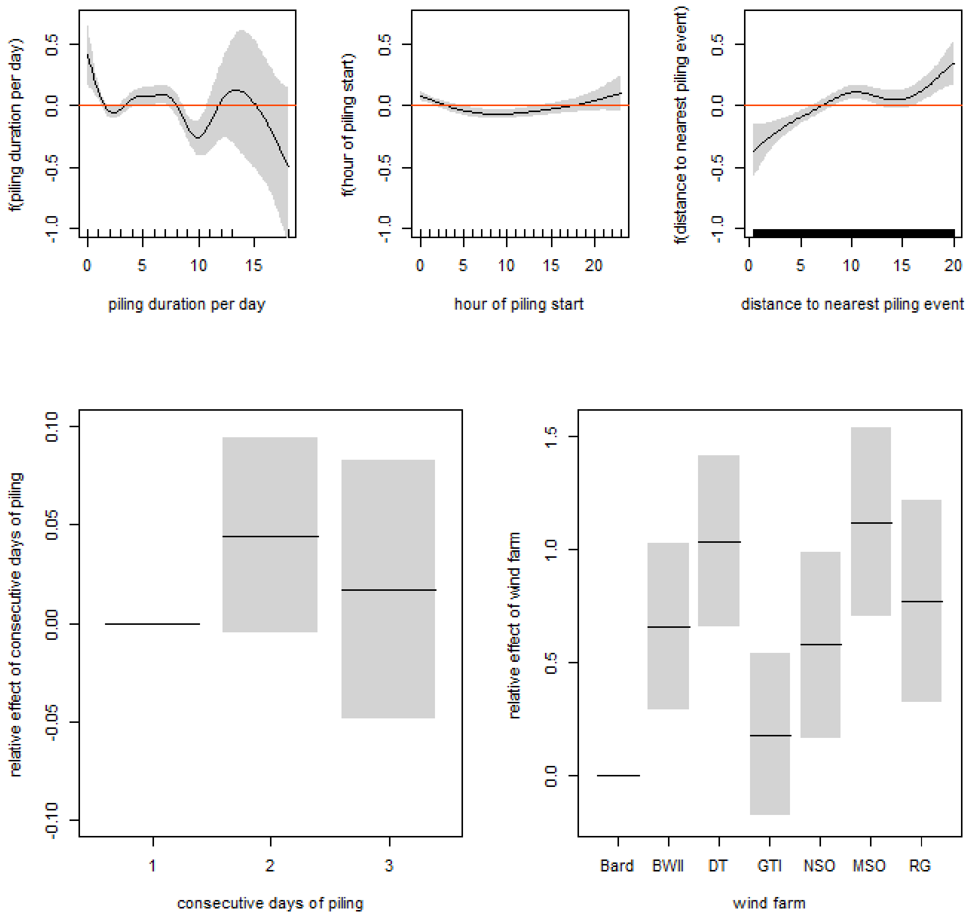
Most studies are based on data from single wind farms and thus are only able to evaluate the general effect of pile driving events for that wind farm project. In reality, however, it happens at times that two wind farm project pile simultaneously and such accumulations of piling events may lead to a considerable difference in effect size. The impact difference between single and accumulated pile driving events is crucial to know for coordination of pile driving events in order to minimise the disturbance for harbour porpoises. Thanks to our vast dataset, we were able to have a look into two different accumulation scenarios: Simultaneous and consecutive pile driving events on a daily basis. All models corrected for the technical artefact of pile driving start time per day and for the pile driving duration (Table 5.7 and Table 5.8).

Consecutive days with piling

Consecutive pile driving events on a daily basis need to be understood as consecutive pile driving days uninfluenced by any other piling event of a nearby wind farm. Due to sample size, we evaluated the difference between one single day of piling and two or three consecutive days with piling events. This was implemented with a three level factor consecutive (Table 5.7). We found that especially short pile driving events up to approximately two hours of length had higher porpoise detection rates than longer ones and a defined distance effect (the closer to the pile the fewer detections) was observed (Figure 5-20). Interestingly, there was a considerable difference in acoustic detection rates between wind farm projects, with Bard having the lowest and MSO the highest rates (Figure 5-20). Consecutive piling events did not lead to a statistically significant trend of decrease or increase in acoustic harbour porpoise detections (Figure 5-20). This however could be an artefact of sample size (Figure 5-21) meaning that the effect might be too small to be proven with the available data.

Table 5.7 Parameters of the model on the effects on consecutive piling events.

Variable name	Variable type in model	Purpose
dp10m	Response	
all clicks	Smooth	Correct for technical POD artefact
SSTA	Smooth	Effect of temperature anomalies
min dist to pile	Smooth	Estimate distance effect on porpoises
piling duration	Smooth	Estimate effect of pile driving time
hour of piling start	Smooth	Correct for different effects of piling on daily basis
consecutive	factor	Estimate effect of consecutive days of piling (1: 1st day of piling; 2: 2nd day of piling; 3: 3rd day of piling)
project	factor	Correct for wind farm related differences
station	Random factor	Effect of deployment position
Error distribution	Quasi corrected Poisson	
Autocorrelation	ARMA on consecutive days level (p=2,q=0)	Ensure that requirements for running the model are fulfilled



-

Figure 5-20 Effect of consecutive days with piling. Grey areas indicate confidence intervals. For the lower two factor plots it must be kept in mind that the first factor level is considered as reference level, and thus it is without confidence interval, and all other levels have to be seen relative to it.

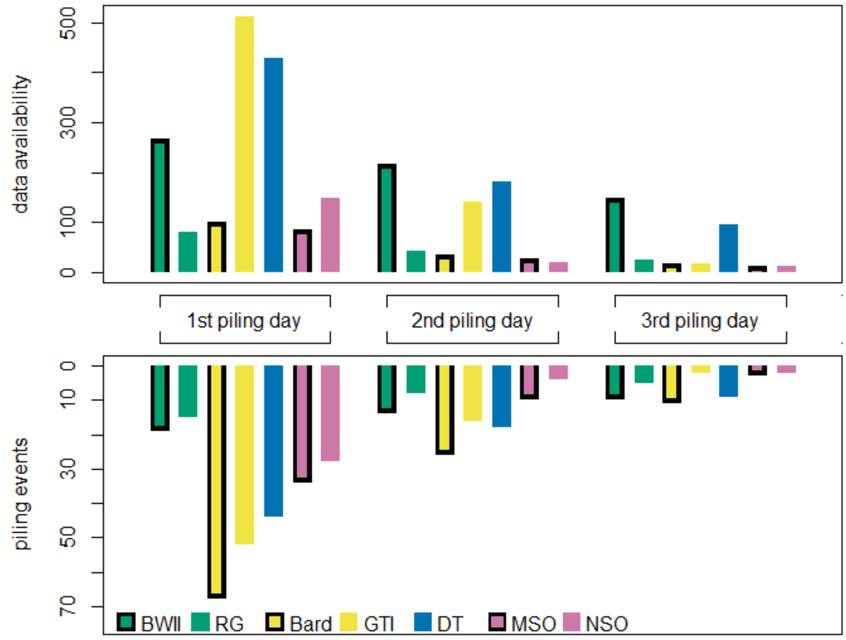


Figure 5-21 *Piling events and data availability per wind farm project and consecutive days with piling events. The colours refer to the subareas the wind farms are located in.*

Multiple piling events per day

We compared the effect of multiple piling events per day, namely two since sample size did not allow for more, to single piling events per day. Only pile driving events uninfluenced by any other piling effects of a nearby wind farm projects as well as only data from stationary PODs which were located in a distance up to 20 km from piling site. Due to sample size this led to the exclusion of the wind farm projects BWII, NSO, GTI and RG (Figure 5-22). For the three remaining wind farms Bard, DT and MSO sample size was still biased towards singular pile driving events per day with considerably less data for two piling events per day (Figure 5-22). In order to evaluate the effect of multiple piling events, we included a factor (cumuCat) with two levels, one for a single piling event and the other for two piling events per day (Table 5.8). Although our results show that on days with two piling events there were statistically significantly fewer harbour porpoise detections (Figure 5-23, left panel), one has to bear in mind the sample size (Figure 5-22) to understand the broad confidence interval. For this dataset, no distinct difference could be found in acoustic harbour porpoise detections among wind farm projects, which had statistically similar detection levels (Figure 5-23, right panel).

Table 5.8 *Parameters of the model on the effects of simultaneous piling events.*

Variable name	Variable type in model	Purpose

Variable name	Variable type in model	Purpose
dp10m	Response	
all clicks	Smooth	Correct for technical POD artefact
SSTA	Smooth	Effect of temperature anomalies
min dist to pile, by = cumuCat	Smooth	Estimate distance effect on porpoises
piling duration, by = cumuCat	Smooth	Estimate effect of pile driving time
hour of piling start, by = cu-muCat	Smooth	Correct for different effects of piling on daily basis
cumuCat	factor	Estimate effect of several piling events per day
project	factor	Correct for wind farm related differ-ences
station	Random factor	Effect of deployment position
Error distribution	Quasi corrected Poisson	

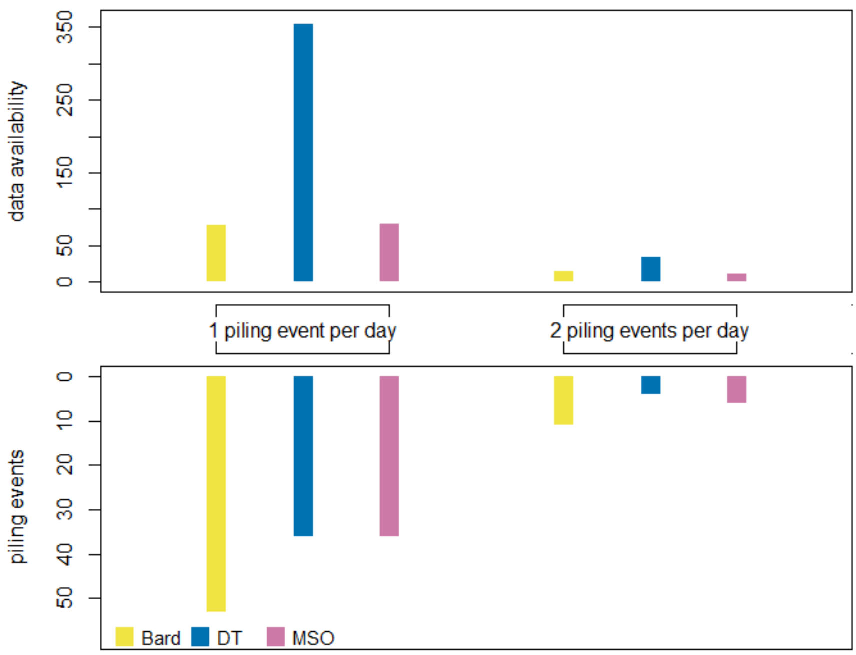


Figure 5-22 Piling events and data availability per wind farm project and piling events per day. Only the wind farm projects shown here were used in the analyses due to sample size. The colours refer to the subareas the wind farms are located in.

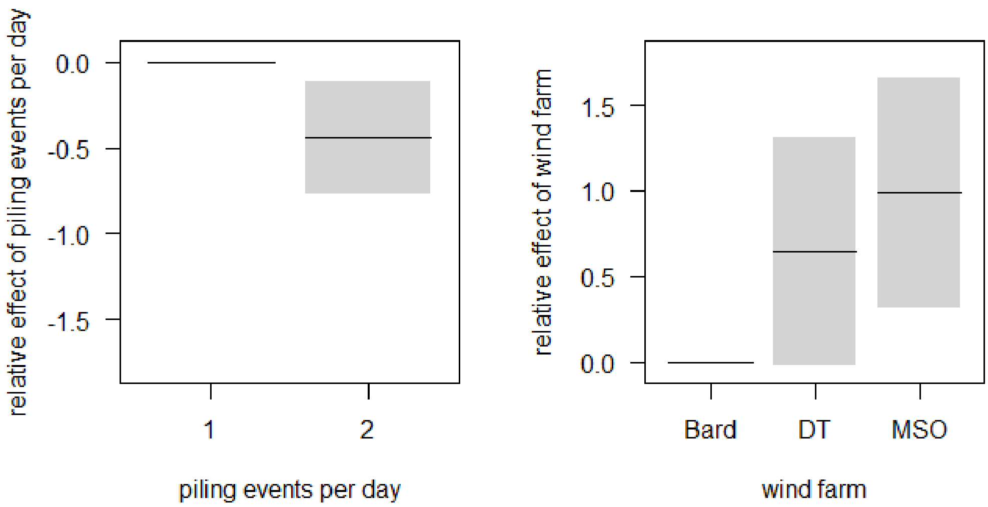


Figure 5-23 Effects of up to two piling events per day. Factor levels with grey confidence intervals have to be considered relative to the first factor level in each plot.

5.4 Discussion (daily POD-data)

The daily POD-dataset was designed to investigate long lasting and broader effects of piling on acoustic harbour porpoise detections. With our analyses we therefore specifically addressed issues concerning long term effects over the course of our study, context related effect size of piling events and changes in effect size with accumulated effects of piling on acoustic harbour porpoise detections. When interpreting our results we must be careful not to forget that although acoustic detection rates are closely intertwined with density (e.g. SVEEGAARD et al. 2011) behavioural aspects could also play a role in altering detection rates. Although the effects should become propagated to a broader resolution this is especially the case when analysing POD-data at a finer resolution such as detection positive minutes per hour. Here, however, we analysed detections on a daily level (dp10m/day). We did not find any indication of a negative long term effect nor any sign of habituation of harbour porpoises to piling noise. Moreover we were able to show that the effect size of piling events on porpoises is dependent on season and subarea. Greater effects of accumulated piling events could only be shown for pile driving events during the same day and not for those on consecutive days.

5.4.1 Natural parameters and their importance for acoustic harbour porpoise detections

Evaluating the effects caused by pile driving activities on harbour porpoise detections requires some knowledge on naturally occurring patterns of porpoise detections in the study area. Only few explanatory parameters were available and amongst these none containing direct information on food availability (probably the most important parameter). We found that sea surface temperature had less explanatory power than day of year, a time-depending variable which correlates not only with intra-annual water temperature fluctuations (colder in winter and autumn and warmer in summer and spring) but also with day length, hours of sunshine per day, properties of wind etc.. The reason for this correlation might be a migratory behaviour of harbour porpoises in the German North Seas which is indicated by our findings that intra-annual acoustic harbour porpoise detection patterns differ between subareas. Anomalies of the expected water temperature caused a change in acoustic harbour porpoise detections, however, indicating that the animals are sensitive to changes in temperature.

In general, some areas, like the BWII-area and MSO-NSO-area, show higher acoustic harbour porpoise detections than for example the BARD-area which might be an indication for their importance to the harbour porpoise population. We found inter-annual differences in harbour porpoise detections, which most likely is related to natural population fluctuations.

Wind speed correlates highly with background noise recorded by C-PODs. On calm days or days with light winds background noise, with click characteristics to become recorded by PODs, is generally lower as well. Simultaneously a negative correlation between acoustic porpoise detections and wind speed can be seen. This effect, however, cannot be considered biological but is merely a technical artefact: porpoise clicks are confounded by background noise.

5.4.2 Long term effects and habituation

Piling activities increased from 2010 to 2013 and acoustic harbour porpoise detections remained constant or also increased during this period – depending on the subarea. We were not able to detect any negative long-term effect over the course of the four study-years within our data. Provided that construction activities did not countervail an otherwise increasing trend we therefore may be reassured to a certain degree that no severe visible effects were caused by piling activities so far.

Harbour porpoises are known to live up to 10 years on average, but there are cases where animals grew to be older than 20 years (BJØRGE & TOLLEY 2009). Females start reproducing from the age of three and then give birth to one calf every year (BJØRGE & TOLLEY 2009). Therefore, a study-duration of four years can hardly be considered a long-term study in terms of the harbour porpoise population life-cycle. Based on our four year dataset, conclusions on long-term effects are somewhat limited, as a reduction in reproductive success during the years of 2011-2013 (when most wind farms were constructed) will only manifest itself in terms of population size changes within the next few years.

Habituation is a highly complex effect to assess especially with data acquired from studies not designed for this purpose. Nevertheless, we approached the question by analysing the change in acoustic harbour porpoise detections over the years for days with and without pile driving. This analysis did not provide any indication that habituation or sensitisation occurred on an annual basis.

5.4.3 Context specific short term effects of pile driving

In order to reconcile the requirements of advancing development in sustainable energy in the German EEZ with the conservation of harbour porpoises, it is of great interest to get information on whether the vulnerability of these mammals to disturbance by pile driving varies with condition, i.e. abundance of harbour porpoises/ amount of detectable harbour porpoise vocalisations. It therefore has been proposed before that seasonal variations in animal density might be important for conservation (GILLES et al. 2011). However, to our understanding so far no effort was made to investigate into effects of piling within different contexts. Based on the knowledge that harbour porpoise detections show seasonal differences within the German Bight (GILLES et al. 2011), which was also validated by our findings, we chose to investigate these matters further with respect to subarea and season. For statistical reasons, however, it was not possible to include both parameters into the models, why we ran the models with respect to season and had a closer look at the raw data. We found that in autumn and winter acoustic harbour porpoise detections were lower on the 1st day after pile driving than on the 2nd day, whereas for spring and summer no differences between the 1st and 2nd day occurred. From this we conclude that the recovery from piling effects is longer during autumn and winter. With regard to subarea we found significantly longer recovery times for the BWII-area only. Since with these analyses we were not able to compare the actual effect size among seasons and subareas, we had a closer look at the effect size of piling on the pile driving day itself: The lowest detection rates were found for the BARD-area and the highest in spring and summer for the BWII-area and the DanTysk-area. We thus cannot generally conclude that lower or higher detection rates in an area or time period lead to longer lasting effects.

5.4.4 Accumulation of piling events

To coordinate piling activities in favour of harbour porpoise conservation within wind farms and among neighbouring projects it is important to be able to pre-estimate the magnitude of accumulated pile driving effects. It was not possible to do any analyses of effects resulting from different wind farm projects since the distance between active sites was beyond the effect range. We thus had a closer look at piling activities within wind farms, which were accumulated with respect to time and space; namely piling events which occurred during the same day and on consecutive days. Due to sample size, we were only able to investigate up to two piling events occurring at one day and up to three days in a row with piling activities. A significantly greater effect size was only observed for piling events during the same day suggesting a mere additive effect meaning that simply a greater proportion of the day was influenced by piling. If there would have been a cumulative effect of successive piling events then it should have been measurable for events on consecutive days of piling as well. These results have to be considered with great care, however, since sample size had been rather small for both analyses and the assignment of pile driving events to days has its shortcomings.

5.4.5 Data basis and methods

Generalised Additive Mixed Models are an advanced frequentist statistics method, which is powerful enough to be used widely for analyses of complex ecological data (e.g. see various examples in ZUUR 2012). Although our dataset is the most comprehensive of its kind for the German EEZ, covering four years of monitoring in 76 different locations, it has intrinsic shortcomings: Study design and POD deployment was not equal among wind farms and thus the dataset is patchy with respect to time and deployment position. This affects analyses of long term effects and habituation especially strongly and makes it difficult to carve out the differences between wind farms. Nevertheless, with our data analyses we are able to add to the understanding of acoustic harbour porpoise detections in the German EEZ and to the effects of wind farm construction.

Our findings are consistent with prior analyses done on C-POD-data in the German EEZ (TOUGAARD et al. 2009; BRANDT et al. 2011; DÄHNE et al. 2013A). When interpreting our results it is crucial to bear in mind that no study using static passive acoustic monitoring can account for the high mobility of harbour porpoises.

As mentioned in subsection 4.4.7, acoustic recordings of porpoises provide relative indices of porpoise activity but cannot at present be directly translated into porpoise density. Previous studies have however found these parameters to correlate broadly with porpoise densities obtained from porpoise sightings (TOUGAARD et al. 2006; SIEBERT & RYE 2008; KYHN et al. 2012; HAELTERS et al. 2013) and even more recent attempts are being made to estimate densities from POD-data (MARQUES et al. 2009; KYHN et al. 2012). Nevertheless, behavioural patterns are most likely also playing a role. Decreases in detection rates during piling are thought to be a result of a combined effect on porpoise behaviour and abundance. We do not know to what exact degree a change in acoustic detections indicates a change in the presence of animals and how much could be explained by an altered clicking behaviour. More specifically, this means that we may not assume acoustic harbour porpoise detections to be straight proportional to harbour porpoise density. However, behavioural aspects should especially play a role when detections are analysed at a finer resolution such as detection positive minutes per hour. Here, however, we focus on a broader resolution (dp10m/day) where those effects can assumed to be smaller.

6 AERIAL SURVEY DATA

6.1 Introduction

Estimating densities based on aerial transect line surveys is one of the standard techniques used to study birds and marine mammals at sea (ICES 2014). With the aid of distance sampling methodology (BUCKLAND et al. 2001; THOMSEN et al. 2005; TEILMANN et al. 2013), the estimation of marine mammals densities from observations made during transect line flights became state of the art methodology and enabled reliable statements on species densities and distribution patterns. In comparison to acoustic data, aerial survey data cover a larger area, but flights are conducted only during favourable weather conditions (sea state < 3; Beaufort scale) and during daylight conditions (THOMSEN et al. 2004). Furthermore, passive acoustic monitoring (PAM) stations record continuously but have restricted spatial coverage (TREGENZA 2011; MCGOVERN et al. 2012; BRUNDIERS et al. 2014). Density estimates from aerial count surveys have been recommended to study the effects of offshore wind farm construction and other anthropogenic influences on marine mammals (GILLES et al. 2009; ICES 2014). Although acoustic data can also be used to calculate density estimates (MCDONALD & FOX 1999; MARQUES et al. 2009; SVEEGAARD et al. 2011; KYHN et al. 2012), this study was not designed to use this methodology, thus only aerial surveys were used to obtain absolute porpoise densities.

This study uses a dataset combining aerial transect surveys from 13 different monitoring projects designed for wind farm development. In addition to surveys conducted in the vicinity of the eight wind farm projects constructed between 2009 and 2013, numerous aerial surveys for baseline investigations were conducted next to these construction sites. As a result, the total number of flights included in this study increased with respect to the area covered during construction and the surveys conducted during or shortly after a piling event. Thus, in comparison to environmental impact assessments of individual wind farm construction, this study incorporates an unprecedented number of surveys related to piling events.

This following section summarizes the influence of the construction of wind farms in the German Bight between 2009 and 2013 on density and distribution of harbour porpoises. More specifically, the aim of this chapter is to address whether the effects of temporally overlapping or successive piling events within different wind farms reinforce each other. To investigate the spatio-temporal disturbance effects of pile driving and estimate whether they have an influence on seasonal and spatial distribution patterns of harbour porpoises. Finally to determine whether the presence, distribution and density of harbour porpoises observed in 2013 have been influenced by the previous piling activities.

6.2 Methods

6.2.1 Study area

The analysed dataset consisted of data obtained from 458 aerial flights conducted in 13 different areas, all located in the German Bight (Figure 6-1). The names of the areas correspond to the names of the respective environmental impact assessment studies of constructed wind farm projects or baseline studies.

Generally, the sampling methodology (devices used and parameters applied) followed the standard used for the investigation of environmental impacts of offshore wind farms on the marine environment (StUK 3, BSH 2007) and thus was similar for all three companies (BioConsult SH, IBL Umweltplanung and IfAÖ). However, some differences between project areas occurred due to specifications agreed to with clients, consultancies and the approving authority (BSH).

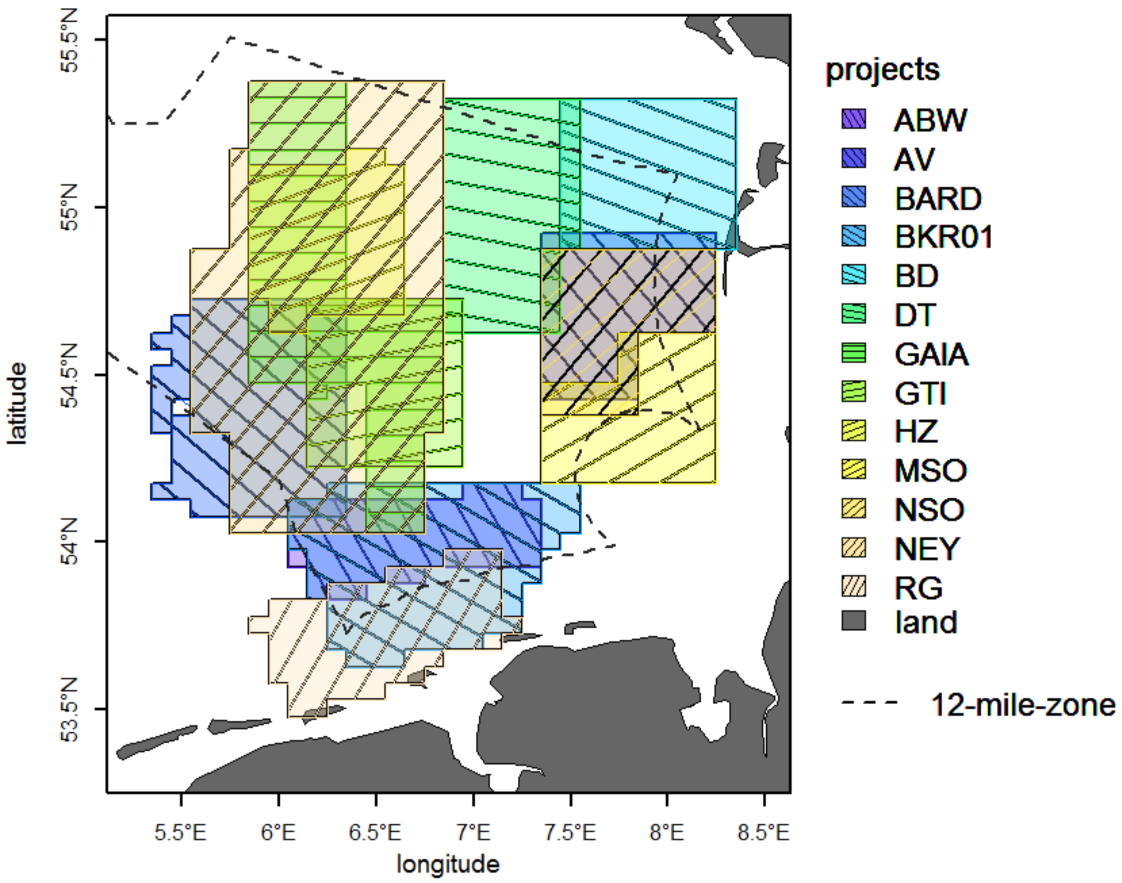


Figure 6-1 Location of the 13 survey areas covered by aerial surveys in the German Bight between 2009 and 2013.

Each survey was carried out as a "total count" (THOMSEN et al. 2004; LAURSEN et al. 2008) and it was classified as a "combined survey" (for birds and marine mammals) or as a "marine mammals survey". For this study, only data collected on harbour porpoises at both flight classes were analysed.

The aerial surveys were conducted by means of twin-engine, top-winged airplanes equipped with bubble windows to allow observations directly underneath the plane at 0°. To minimize the possibility of missing animals due to the crest and spume of breaking waves, the surveys were conducted only during sea state conditions ranging from 0 to 2, according to the Beaufort scale. The flight speed was 100 knots (180 km/ hr), at an altitude of 250 feet (76 m) for the “combined survey” and at 600 feet (183 m) for the “marine mammal survey”.

For each survey, three experienced observers collected the required data. Two of them were considered as “main observers” and sat next to the bubble windows. The third observer, called the “control observer”, sat behind them, close to a normal window (no observations directly underneath the plane were possible) and his/her data were used to estimate observation biases (THOMSEN et al. 2004, 2005).

At the beginning of each transect, parameters related to weather conditions (sea state on the Beaufort scale, turbidity, maximal visibility range in km, cloud cover, cloud reflexion and glare) were recorded in UTC time on a digital tape recorder. Whenever conditions changed, the observers stored that information. These parameters were used to estimate the valid observation effort (THOMSEN et al. 2004, 2005; KLÜVER & IFAÖ 2011).

Along each transect, when a sighting event occurred, the following information was recorded: UTC time, number of animals, age (adult or calf), behaviour (porpoising, foraging, diving, resting, etc.), swimming direction and perpendicular angle using a clinometer. The distance of the animals from the transect line was calculated using the perpendicular angle and the height of the plane (THOMSEN et al., 2004).

By using an on-board GPS, it was possible to geo-reference any single observation. For further details concerning this methodology, please see THOMSEN et al. (2004, 2005).

6.2.2 Data treatment and density estimation

The key assumption of line-transect sampling is that all animals on the transect line are detected with certainty (i.e. the probability of detecting animals directly below the plane is 1). For marine mammals surveys, this assumption is almost certainly violated because the probability of detecting all animals at distance 0 from the line is small (THOMSEN et al. 2005).

There are two sources of bias that need to be taken into consideration: "perception" and "availability". The "perception bias" occurs when animals are missed by the observers, even when they are available to be spotted, while "availability bias" occurs when porpoises are not visually detected due to a diving behaviour and not because they are not present in the sampling area (LAAKE et al. 1997; THOMSEN et al. 2005). The "perception bias" was corrected using a double platform approach with two observers or the cycle back method (HIBY & HAMMOND 1989; THOMSEN et al. 2005), the "availability bias" was addressed by using correction factors for diving behaviour from TEILMANN et al. (2013). Both factors were combined to calculate the probability g(0) that an object that is on the line was detected (details, see THOMSEN et al. 2006A, 2007).

The density of harbour porpoises was estimated for all 458 flights using the program DISTANCE, Version 6.0 (BUCKLAND et al. 1993, 2001). The program selects the model with the best fit, calculated on the basis of the lowest Akaike`s Information Criterion (AIC; BURNHAM & ANDERSON 2002). DISTANCE uses this function to determine the effective strip width (ESW). All the detections that are missed inside the ESW equal the number of observed detections outside the ESW. The actual density, consequently, is represented by an estimation that incorporates all the observed data and areas based on the ESW.

In this study, the ESW values from the original dataset were maintained. However, the detection probability was recalculated due to more recent information that diving behaviour of harbour porpoises changes geographically (TEILMANN et al. 2013). As a result, actual density estimates might deviate from those presented in the previous environmental impact assessment studies that based the "availability bias" on older results (TEILMANN 2001).

The density was finally estimated by considering the valid observed area, consisting of the weather-corrected length of the transect line and the ESW multiplied by g(0).

After the integration of all data, densities were calculated for grid cells of comparable size and effort (3 x 6 arc minutes, approx. 6.0 x 6.0 km). Cells with sightings and small spatial coverage might show unreliable above-average densities (for example, one animal sighted in an area of 0.125 km² results in a density of 8.0 ind./km², when ignoring g(0) and diving time). To reduce artefacts, density estimates were only derived for cells with at least a 2.0 km² effort.

Survey effort and the number of porpoises observed are shown in (Table 6-1).

Table 6-1 Number of harbour porpoises sighted during surveys conducted between 2009 and 2013 in relation to number of flights and valid two-sided effort (km) that were included in this report.

Year	2009	2010	2011	2012	2013	Total
Number of harbour porpoises	2,174	3,670	6,031	2,264	3,084	17,223
Number of flights	66	93	128	73	98	458
Valid two-sided flight effort [km]	37,260.1	47,601.8	52,850.5	30,393.2	33,212.8	201,318.3

Valid visual observations were obtained during 201,318 km of flights that covered an area of approx. 24,500 km², representing 2/3 of the entire German Bight area. Splitting the dataset into 768 grid cells resulted in 32,054 values for all aerial surveys. For 7,520 (24%) of those cells, a density > 0 was calculated. In the remaining grid cells, no porpoises were detected (Figure 6-2). Transects differed between the 13 projects. Generally, the direction was from north to south, but those of RG deviated by 20 degrees, and those of Butendiek, ABW, MSO and NSO were flown in an east-west direction. Furthermore, distances between transects also deviated. The grid size was chosen independently of the distance between parallel transects, because these deviated, as well. Per grid cell and individual flight, one or more parallel transect lines were merged in one grid. For this reason, the temporal resolution of the grids has been smoothed and, consequently, autocorrelation of neighbouring grids had to be considered more on the spatial than on the temporal level (Figure 6-3).

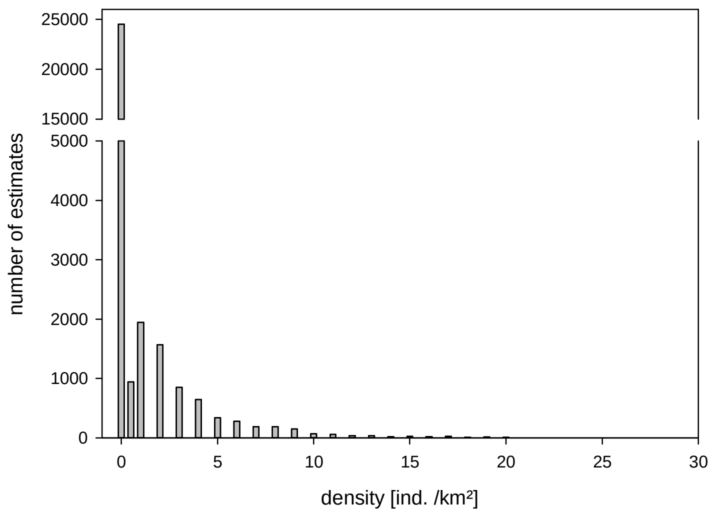


Figure 6-2 Histogram showing the frequencies of density estimates from the grid cells of 6 by 6 km cells estimated from the aerial survey dataset (n = 30,054).

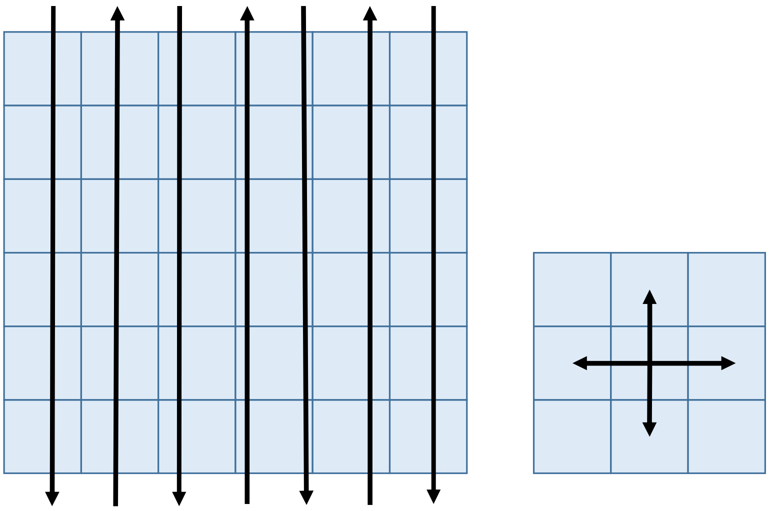


Figure 6-3 The grid size of 6 by 6 km resulted in varying numbers of transect lines per grid cell, which resulted in a temporally incoherent sequence of neighbouring grid cells (left). Autocorrelation of data is therefore more a function of space than of time (right).

6.2.3 Distribution of aerial survey effort

The flights were conducted on 261 different days, which corresponds to 11.9 % of the span of the studied period (2009-2013). For information regarding the number of flights per day, see also Table 6-2.

Table 6-2 Number of flights conducted per day in total and per year in the German North Sea between 2009 and 2013.

No. of flights conducted per day	Total	2009	2010	2011	2012	2013
1	140	38	25	26	26	25
2	72	11	22	15	12	12
3	29	2	4	9	5	9
4	14	0	3	6	2	3
5	5	0	0	3	0	2
6	1	0	0	1	0	0
Total	458	66	93	128	73	98

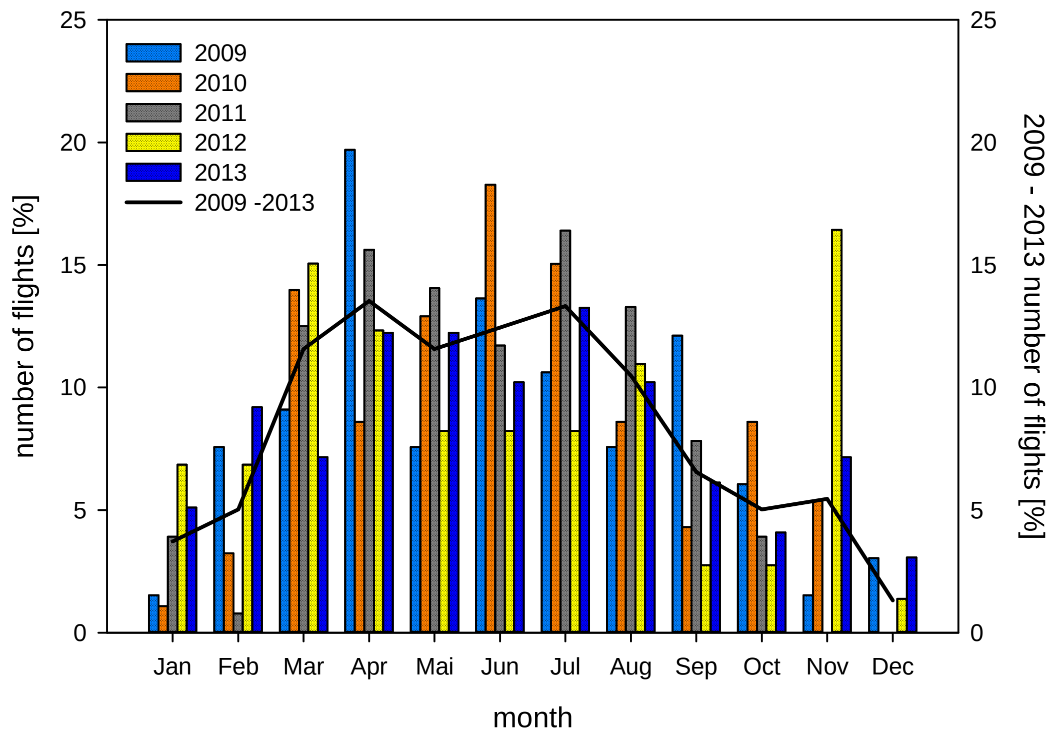


Figure 6-4 Relative monthly coverage of survey flights in calendar years 2009 to 2013.

For this study, only data collected during valid, two-sided observation effort were analysed. When taking the variables “spatial coverage” or “year” into account, the number of observations were not equally distributed (Figure 6-5).

In 2009, the surveys conducted were restricted to the area around AV and to the area north of it, between 53.75° and 55.25° N. In 2010, the survey coverage was expanded into the western part of the study area, from the East Frisian coast up to 55.4° N. The eastern part was included only in the survey area of ABW. In 2011, the area west of the North Frisian coast was completely covered, while the western part was included only in the survey area of BARD, GAIA Nord, AV and RG. Finally, the coverage in 2012 and 2013 was similar to that of 2011, but included GTI, as well. At this time, the survey of the GAIA Nord area was terminated.

To analyse flight coverage for the whole study in more detail, the data were subsequently compared based on the association of the variables "year" and "season" (Figure 6-5).

The areas that were covered to the greatest extent with regard to seasonal and inter-annual effort were the areas north of the East Frisian Islands and around the BARD and GTI wind farms. Here, with the exception of autumn, 2013, and winters, 2013-2014, all seasons were covered with at least one flight each.

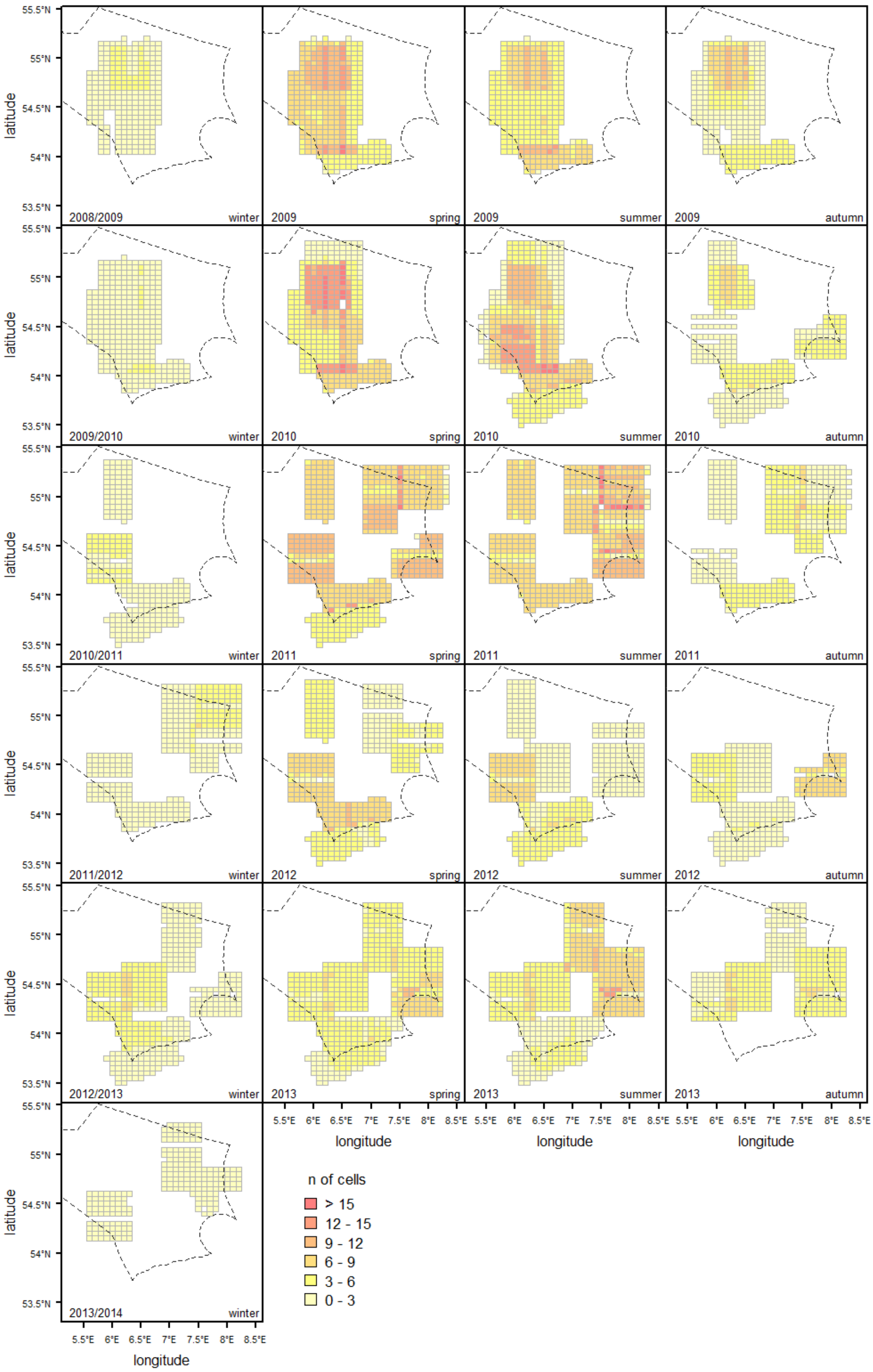


Figure 6-5 Spatial coverage and number of surveys conducted in the German Bight per grid cell between January 2009 and December 2013 illustrated per year and season.

Conversely, the area within the Sylt Outer Reef Special Area of Conservation (SAC) was covered relatively little: the survey flights between Helgoland and the SAC did not start until autumn, 2010. Thereafter, however, the time coverage was considered as relatively good.

Similarly, the area located in the north-western part, west of the Sylt Outer Reef and north of BARD, was covered continuously only between the winters of 2008-2009 and autumn, 2011. It has to be taken into consideration, however, that the analysis of this area was associated with the analysis of GTI, BARD and DT, due to the high construction effort undertaken during the study period.

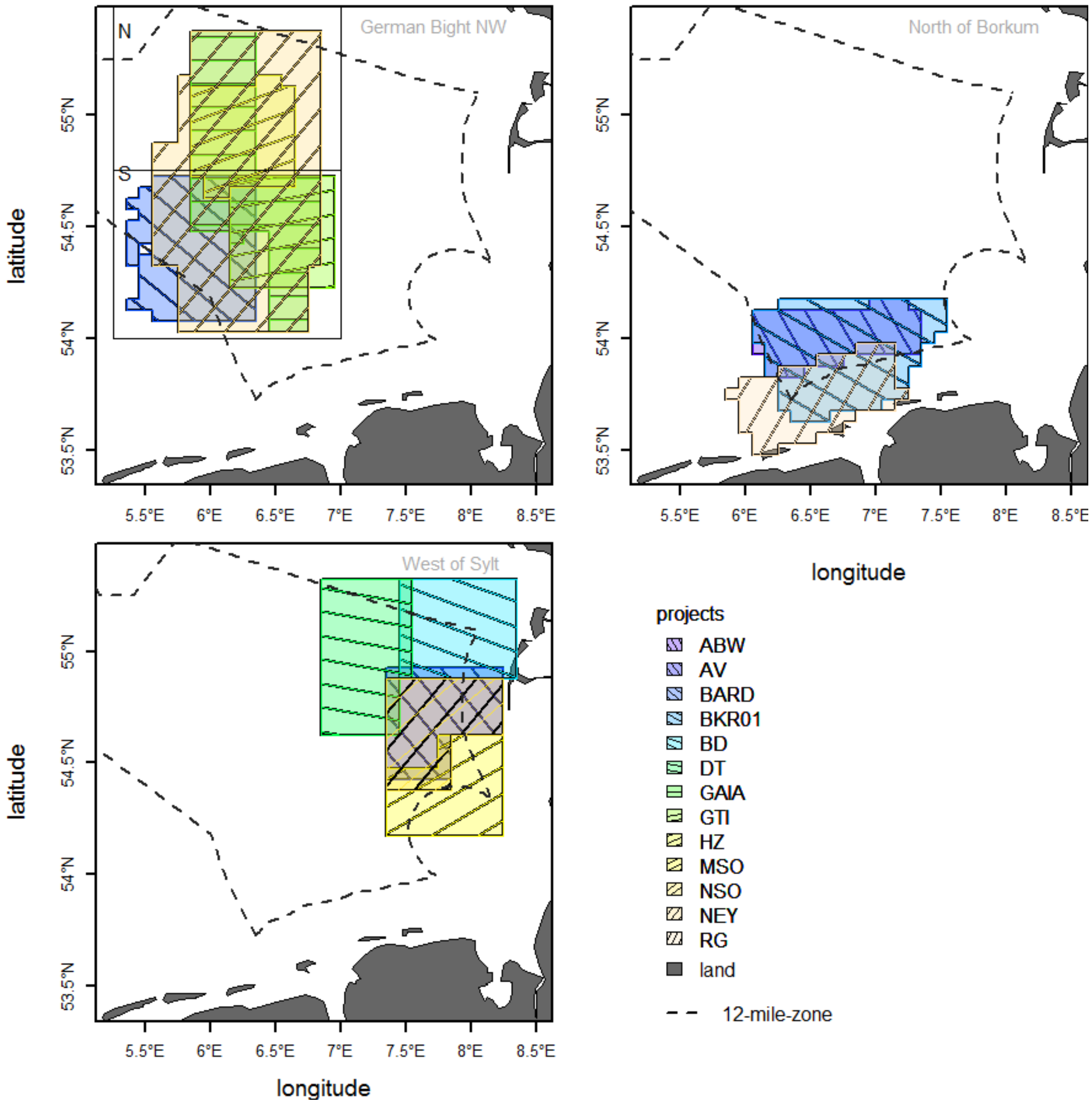


Figure 6-6 *Division of the study area into three separate subareas according to temporal coverage and overlapping surveys (see Figure 6-1).*

In light of the seasonally- and annually-changing effort, the dataset was screened in more detail to investigate how the flights were spatially and seasonally dispersed per wind farm project. Furthermore, the comparability of the datasets was analysed based on affected and unaffected values collected during aerial surveys conducted within the 60-km radius around all eight construction sites, before and after the construction phase. If piling events occurred during or shortly after flights, the surveys during the construction phase were separated into "piling" and "no piling" flights. Based on data availability, the study area was divided into three different subareas: "West of Sylt", "German Bight NW" and "North of Borkum". Construction in the area “West of Sylt” was undertaken at DT from the end of February to November 2013, at NSO from October 2012 to December 2013 and at MSO between September 2012 and May 2013. Construction at ABW started in 2014 and was not considered in this report. In the subarea “German Bight NW” construction was undertaken at BARD from April 2010 until March 2013 and at GTI from August 2012 until December 2013 (some construction continued in 2014). In the subarea “North of Borkum”, construction was undertaken at AV between April and August 2009, at BWII from September 2011 until March 2012 and finally at RG from June until September 2012.

6.2.4 Assignment of environmental data to flight data

For a joint analysis, environmental data were integrated into a database and data quality was tested (WOLLERT IT & IFAÖ 2011). All environmental data (see subsection 3.3 Environmental data on page 17) were acquired as "static" or "dynamic" variables, were assigned per grid cell and, if possible, included as daily or weekly values. Table 6-3 provides a list of environmental and other variables assigned per grid cell.

Table 6-3 List of all variables associated per grid cell.

Variables	Type	Scale	Description
Piling-related variables			
deterrence	factor – 3 levels	Minutes	Noise mitigation applied, not applied or functioning only part of the time
distance	continuous	Decimal km	Distance to nearest piling event in km or nearest piling event, if two or more piling events occurred within the last 60 hours before the flight
hour relative to piling	continuous	Decimal hrs	Time passed since last piling ceased
strokes	continuous	Counts	Number of strokes applied during a piling event
duration	continuous	Minutes	Duration of a piling event in min
cumEnergy	continuous	kJ	Cumulative energy applied during a piling event
SEL_{05}	continuous	dB	Noise Exposure Level exceeded during 5 % of time of a piling event
Environmental variables			
substrate_cat	factor – 5 levels	categorical	Seabed sediment (0: fine seabed <0.3 %, 1: mud <3.3 %, 2: sand to muddy sand ~ 89 %, 3: coarse sediment < 7.5%, 4: mixed sediment: < 0.1 %)
biozone	factor – 2	categorical	biozone (Circalittoral 77.8 %; infralittoral 22.2 %)

Variables	Type	Scale	Description
	levels		
day of year	continuous	digit	day of the year (1 – 365)
year	factor – 5 levels	digit	year (from 2009 to 2013)
HH	continuous	digit	hour of the day (0-24)
flight time	continuous	decimal hrs	length of flight
SST	continuous	weekly	sea surface temperature in °C
SSTa	continuous	weekly	anomaly of the sea surface temperature in °C
day_length	continuous	decimal	duration from sunrise to sunset in h
latitude N	continuous	Decimal degree	latitude (WGS 1984)
longitude E	continuous	Decimal degree	longitude (WGS 1984)
moon illumination	continuous	daily	moon illumination
water depth	continuous	decimal	average water depth in m
cumulative variables			
fPiling	factor – 4 levels	categorical	categorical factor describing piling events during flight, no piling, piling ended 0 to 24 hrs before flight, 24 – 48 hours before flight and 48 – 60 before flight
day and distance	factor – 10 levels	categorical	Combination of three time periods and three distance classes 0: unaffected area or timeframe; Time periods: 1, 2 and 3 piling refers to the time in days passed since piling ceased) Distance < 20 km: piling occurred within a radius of 0 to 20 km; < 40 km: piling occurred within a radius of 20 to 40 km, < 60 km: piling occurred within a radius of 40 to 60 km
cum_0.5month_60km	continuous	count	Piling events within a 60-km radius around the centroid of the grid cell during the 15 days before the flight
cum_1month_60km	continuous	count	Piling events within a 60-km radius around the centroid of the grid cell during the 30 days before the flight
cum_0.5month_40km	continuous	count	Piling events within a 40-km radius around the centroid of the grid cell during the 15 days before the flight
cum_1month_40km	continuous	count	Piling events within a 40-km radius around the centroid of the grid cell during the 30 days before the flight

6.2.5 Aggregation of piling events

To specifically study the effect of wind farm construction, grid cells were assigned to piling events if the distance between the centroid of the grid cell and the piling event was smaller than 60 km (Figure 6-7, left) and if piling ceased less than 60 hours before the flight. These dimensions were chosen due to the negative correlation between wind farm construction and distances ranging from 6.0 to 24 km described in the literature (TOUGAARD et al. 2006; DIEDERICHS et al. 2010; DÄHNE et al. 2013B; DIEDERICHS et al. 2014). The specific setting of the radius permitted the designation of "reference areas" that were beyond 24 km and beyond 24 hours. The distance was extended to also cover the study areas adjacent to the construction site.

To test for cumulative effects, the number of piling events occurring within a 40- and 60-km radius from the centroid in the 15 and 30 days preceding the flight were calculated as well. All variables concerning piling events, plus oceanographic and bathymetric information, were assigned to each grid cell (for details, see Table 6-3).

Individual measurements of noise pressure levels were undertaken for up to 50 % of the piling events. Preliminary analysis indicated that the distance from piling contained relevant information on piling events and thus could be used as a describing variable to replace noise pressure level and increase the number of piling events analysed in this study. Nevertheless, the dimension of the grid cells (approx. 6.0 x 6.0 km) was so large that it could not be accurately represented by a single noise pressure value. Incorporation of the source level at 750m distance from piling, to model the effects of noise pressure levels on the density and distribution of harbour porpoise, was considered too imprecise because of the logarithmic scale for further analysis. This reasoning is based on the geometric dispersion of noise in water, which results in increasing uncertainty with increasing distances (RICHARDSON et al. 1995; THIELE 2002; SHAPIRO et al. 2009).

For this reason, the minimum distance from piling sites (distance) and the minimum time since piling ceased (hour relative to piling) were used as principle describing variables and were included as "interacting variables".

In conclusion, three datasets were generated to study the spatial and temporal effects of pile driving on the distribution and density of harbour porpoises in the German Bight:

- The entire dataset containing all grid cells (n = 32,054). A factor was included to indicate if piling occurred or not. Further, a variable was added to provide information on cumulative effects by calculating the moving average of the number of days with piling events within a 60-km radius during the previous 15 days.

- The unaffected dataset comprising unaffected grid cells from aerial surveys beyond a 60-km radius around piling events, either before deterrence measures started or 60 hours after piling ceased (n = 24,324).

- The affected dataset containing only grid cells associated with flights conducted within a 60-km radius around a piling event that occurred within the 60 hours before the flight (n = 7,730).

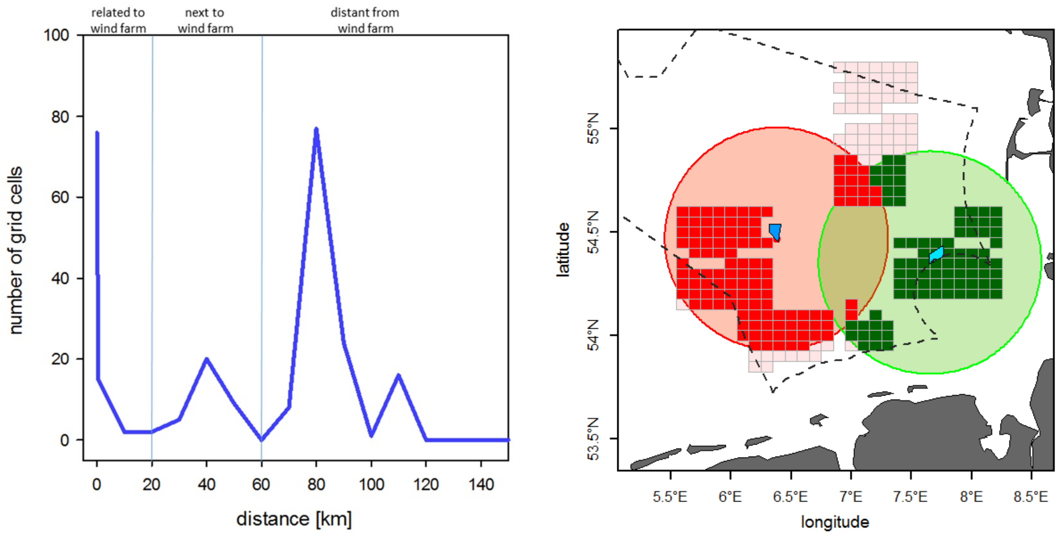


Figure 6-7 *The distance (km) between piling events and cells is multi-modal; the green line shows the distance threshold chosen for this project (left). The right panel plot presents example affiliations to piling events. If piling occurred on more than one foundation on the same day, the cells were associated according to the minimal duration since piling ceased (right green circle). Only cells within the two circles were assigned to one of the two piling events. Cells within the intersecting circles were assigned to the red circle*

If construction was occurring for two projects simultaneously, more than one foundation could have been piled within a 60 km radius around the grid cells centroid and 60 hours before the flight. To avoid replication in such a circumstance, the cell was associated to the last piling event, according to the minimal duration that had passed since the end of the piling event and the time at which the plane surveyed the grid cell (Figure 6-7, right). According to this approach, 7,730 out of 32,054 cells (24.1 %), belonging to 179 different flights (39.1 % of all flights) and conducted on 104 discrete days (39.8 % of all days), were associated to 133 piling events (n = 207). These piling events represent approximately 19 % of all piling events occurring in the study area between 2009 and 2013. 83 of these piling events occurred between 0 and 24 hours before the aerial survey, 38 between 24 and 48 hours and 11 between 48 and 59 hours.

In total, 743 piling events occurred on 515 days. On 96 days, piling was repeated once for the same foundation; while on 16 days, it was repeated 2 or 3 times. These were not cumulative but rather consecutive pilings and they were considered as variable in the modelling process (chapters 6.3.3 and 6.3.4). During the entire study period, two different foundations were piled on 108 occasions on the same days, three foundations on 25 days and 4 foundations on 3 days. Not all of the considered pilings events occurred simultaneously (Table 6-4).

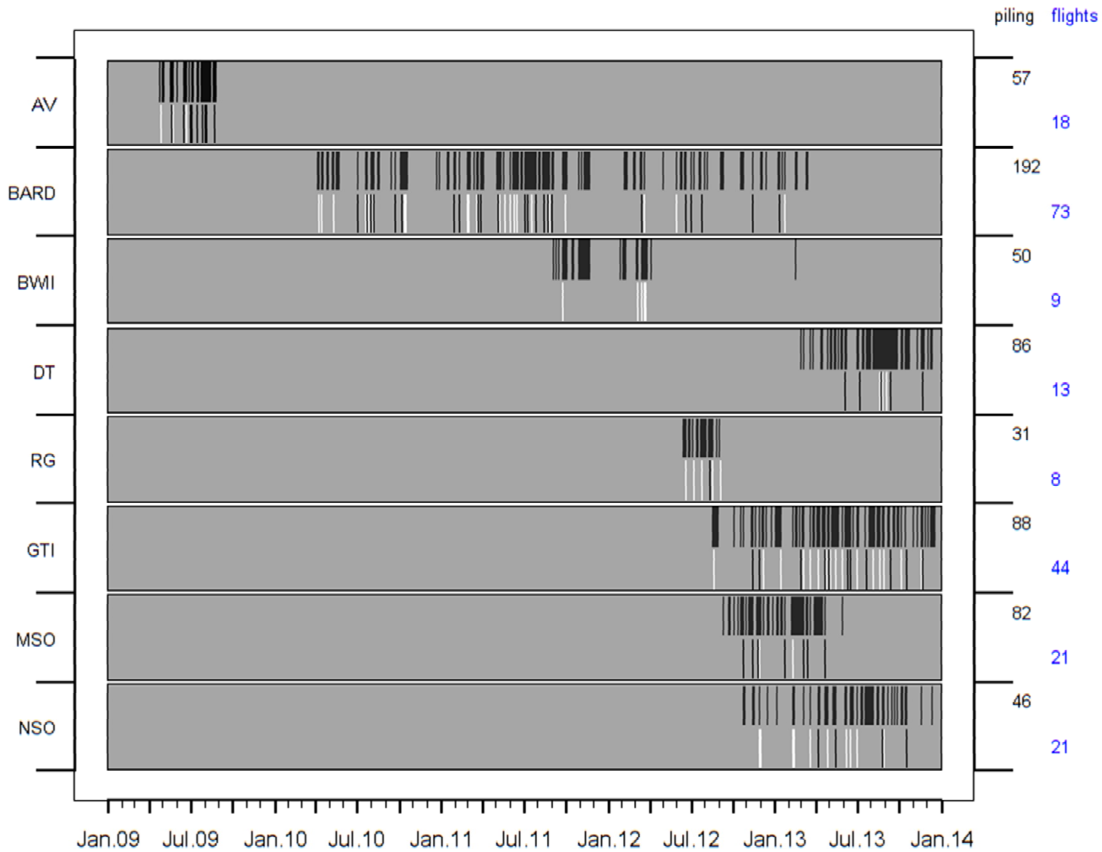


Figure 6-8 *Pile driving periods for wind farm projects constructed during the 2009-2013 study period and the corresponding number of flights carried out in a 60-km radius around each wind farm. The upper portion of the bar related to each project indicates when piling events occurred; the lower part of the bar indicates when aerial surveys were conducted during pile driving. Black lines indicate surveys conducted by the project that was executing the construction; white lines indicate all other surveys in the same area.*

These surveys covered 133 piling events, by means of 179 flights conducted over 121 days. Grid cells were associated to each wind farm project and the density curves were set according to the distance between piling events and the centroid of the grid cell (Figure 6-9). The number of aerial surveys assigned to piling events differed between construction projects due to the length of the construction phase, surveys conducted in relation to the eight construction sites and other surveys within 60-km distance that were conducted for other projects (Figure 6-8).

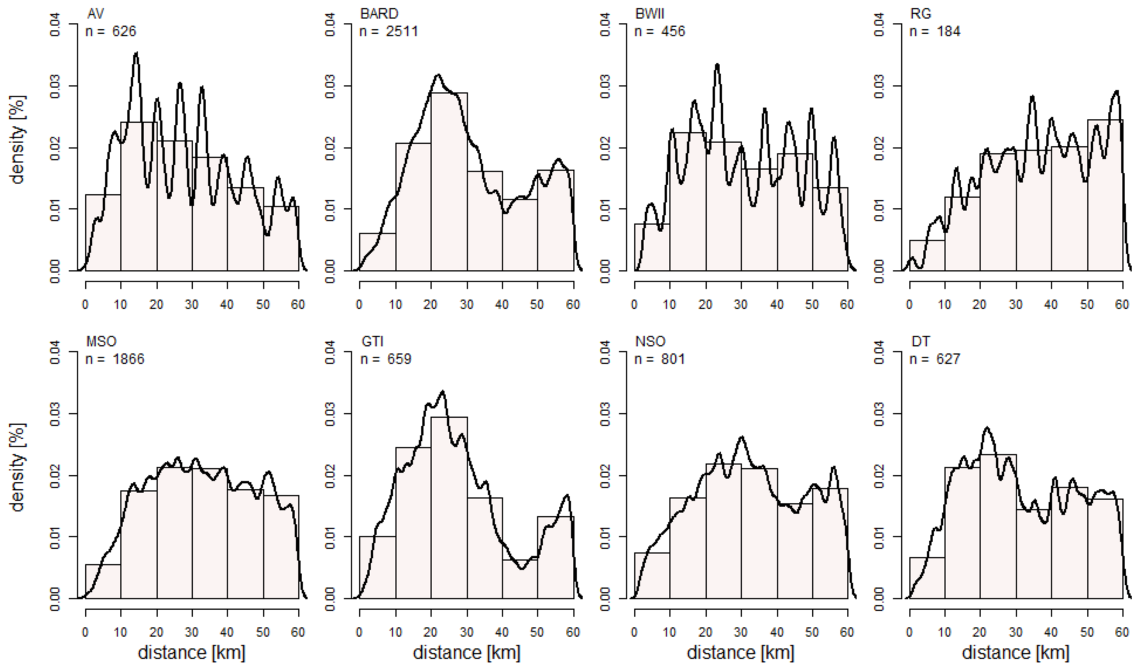


Figure 6-9 Distribution of distances between piling events and cells in kilometres. Each plot represents a subset per wind farm (n: number of grid cells assigned to piling events).

Generally, the number of piling events rose during the study period and the data collected in 2012 and 2013 were more likely to be excluded from the unaffected dataset than those sampled between 2009 and 2011 (Figure 6-10).

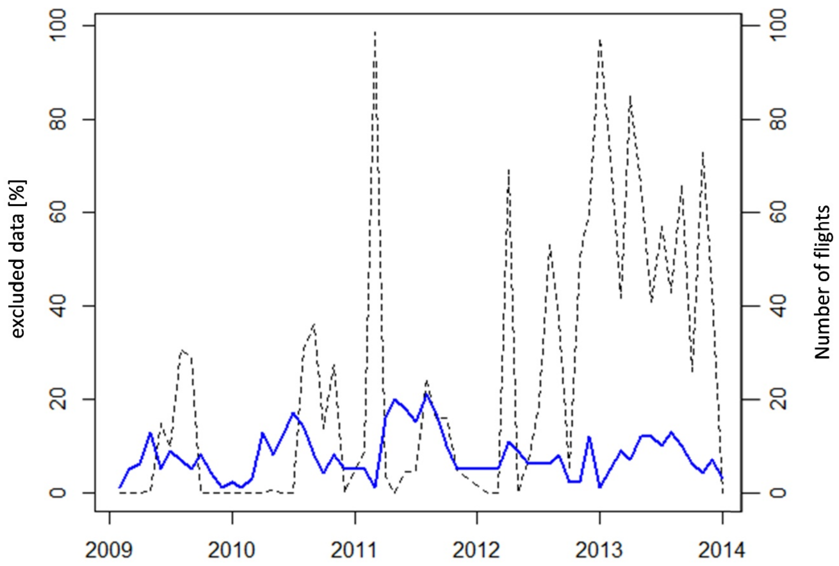


Figure 6-10 Data reduction by excluding all grid cells from the dataset (dashed line) and flights conducted in total (blue line). Excluded were only grid cells within a 60-km radius around piling events, if these events did not cease more than 60 hours before the flight.

Before the construction phase, 4 of 8 wind farms (BWII, DT, GTI and RG) were surveyed by at least seven flights. During pile driving activities, between 3 and 52 aerial surveys were conducted (Table 6-4). The seasonal coverage was complete at three wind farms (BARD, GTI and NSO), while it was reduced to only one season (summer) at one (RG). During the other projects, the surveys were carried out only in two or three successive seasons. During this phase, the highest number of flights (n=52 during piling; n=17 for no piling) was registered at the wind farm BARD (Table 6-4).

Table 6-4 Number of flights conducted in the vicinity of the eight wind farm projects in total and per season (grid cells nearer than 60 km to project site; season: spring/ summer/ autumn/ winter, for more details see Figure A-23 to Figure A-30).

Wind farm	Aggregated	Before	Construction		After
			Piling	No piling	
AV	total	0	11	1	22
	per season	0/0/0/0	2/9/0/0	1/0/0/0	5/11/3/3
BARD	total	2	52	17	3
	per season	1/1/0/0	8/29/10/5	4/6/3/4	1/2/0/0
BWII	total	13	8	1	14
	per season	3/7/2/1	7/0/1/0	1/0/0/0	4/5/1/4
DT	total	7	13	4	3
	per season	5/0/0/2	0/9/4/0	0/3/1/0	2/1/0/0
GTI	total	8	40	8	5
	per season	3/3/1/1	14/11/10/5	1/3/0/4	1/2/1/1
MSO	total	1	18	2	4
	per season	0/1/0/0	7/0/7/4	0/0/0/2	2/2/0/0
NSO	total	0	21	8	0
	per season	0/0/0/0	7/6/4/4	3/4/1/0	0/0/0/0
RG	total	11	3	1	7
	per season	3/6/2/0	0/3/0/0	0/1/0/0	2/1/2/2

6.2.6 Statistical analysis

The statistical program R 3.2.0 (R Development Core Team 2015) was used to create maps, to parameterise models used to describe the density and distribution of harbour porpoises and to perform statistical tests.

Testing of differences was accomplished using the Kruskal Wallis test (base package), because density data are not normally distributed. When multiple comparisons between dataset were necessary, multiple comparisons were accomplished with a Nemenyi test (PMCMR package), a post hoc test applied after performing a Kruskal Wallis test. Probability values smaller than 0.05 were considered significant.

Modelling of density and distribution of harbour porpoises:

Due to the non-normal distribution of the aerial flight data, the large amount of zero values and the Poisson-like distribution of the values greater than zero (Figure 6-2), Generalised Additive Models (GAMs) (HASTIE & TIBSHIRANI 1990; WOOD & AUGUSTIN 2002) were compiled using the function gam() of the package mgcv (MGCV 2015).

A negative binomial distribution was considered after Poisson and quasi-Poisson distributions had been ruled out in the interim report; a gamma-level of 1.4 was chosen in order to prevent oversmoothing due to the dispersion parameter value of the models (above 2.2) obtained despite using quasi-Poisson distributions (ZUUR et al. 2009; ZUUR 2012b). The significance level was set to α = 0.1 in order to prevent the exclusion of important predictors, as the considered study area was very heterogeneous (GILLES et al. 2011).

Model selection was performed using a stepwise exclusion of variables based on AIC. Variables were excluded from the model based on the most parsimonious model (δAIC > 2). If the p-value of a variable was smaller than 0.1, it was tested to determine whether exclusion resulted in a model with a lower AIC than the model including this variable and the simpler model was consequently preferred. This process was initiated using static variables that were included as random factors (substrate, observation height and biozone) and was continued with all the other variables (continuous and static). During model selection, geographic cell location and time of year appeared as influential variables, reducing the models AIC values.

If the count data originate from different sampling areas, Poisson distributed regression requires an offset in order to adjust for the different sizes. Thus, the area covered per grid cell was included as an offset. Considering that the probability of observations did not increase linearly with increasing size of the area, the area was converted using a logarithmic transformation (ZUUR 2012b).

Geographic distribution models of the entire study area:

Three different datasets were used to model how the distribution patterns of harbour porpoises changed in the German Bight over time. The entire dataset included both affected and unaffected grid cells (n = 32,054 lines of data, model A1). Those two types of data were distinguished by assigning a factor that indicated whether piling occurred within the last 60 hours before the flight (n = 7,730, model A3). Further, an unaffected dataset containing only unaffected data (n = 24,324 lines of data, model A2) was analysed to distinguish general pattern differences within the unaffected data.

The density pattern of harbour porpoise was modelled considering the spatial coordinates as tensor spline (te), because latitude and longitude have different dimensions. Due to the size of the study area, the temporal extent of the dataset and the distribution patterns that were described in the literature, a plasticity was granted to these variables by grouping the distribution per season (as an interaction term) (GILLES 2003; FFH 2007-2012; GILLES & SIEBERT 2009; HAELTERS et al. 2011; GEELHOED et al. 2013; VIQUERAT et al. 2015). Further parameters were day of year, anomaly of the sea surface temperature, water depth, moon illumination, flight time, hour and cumulative effect. These were parametrised as thin plate splines (s). Position ID of the grid cell was considered as Markov random smooth - mrf (KNEIB et al. 2006; WOOD 2006).

Geographic distribution models of subareas:

When limiting the modelling of the regional effect of piling effects to the three subareas (model A6, A7 and A8: “North of Borkum”, “West of Sylt” and “German Bight NW”), the spatial distribution of harbour porpoises was studied for spring and summer on the basis of the model described above. Model parametrisation followed the methodology described above (ANDERSON 2002; ZUUR et al. 2009).

In the course of the study, the subarea “German Bight NW” was divided in a northern and southern part (south and north of 54.7 ° N, model A4 and A5) because the northern part was not covered in the year 2013 and no piling events occurred within this region. Differing distribution patterns of harbour porpoise and densities in spring and summer were studied to analyse how those changed in a region not directly affected by piling events.

Spatio-temporal effect of piling events:

The effect of piling on harbour porpoise densities was modelled using the same variables used during the preliminary study. A Markov random field (mrf) smooth predictor (WOOD 2006) was applied. This smooth predictor is a built-in function of the mgcv package that smooths over a set of discrete areas (Position ID of the grid cell) based on a neighbouring structure that is calculated per area and that was included as a covariate of the mrf smooth. Each grid cell was coded as a polygon and a neighbourhood relation was assigned if the grid cells shared one, of two, vertices.

Additionally, the distance between centroid of the grid cells and piling position (distance), as well as the time since piling ceased (hours related to piling), were modelled as a two-dimensional tensor spline, because of the different nature of the variables and the discontinuous data availability.

Using this model as a reference, several other models composed of a reduced amount of data (different time and distance ranges) were studied in depth to understand the relationship of distance and time since piling ceased, in total or per project.

In this regard, the effect of mitigation measures was also tested. The only project with a sufficient amount of trials utilizing different noise mitigation systems was the construction of the Trianel Borkum West II (BWII). The incorporation of such information was not considered in this study.

In order to determine which variables should not be used jointly in the final model, a Spearman rank correlation coefficient was calculated for all pairs of continuous predictor variables.

The model presented here (model ST1) contained 6,857 lines of data, with the response variable (porpoise density) being described by 10 predictor variables. This number is smaller, because some of the variables were not present for the length of deterrence. In the first model, the following variables describing the density of harbour porpoises were investigated per season: day of the year (day of year), sea surface temperature anomaly (SSTA) and moon illumination as a proxy for tidal currents and strength. Inter-annual variation was included by grouping the variable day of year by the year and by including year as a further smooth. The geographical location variables (latitude N, longitude E) were included, as were the variables related to data quality and flight conditions (flight duration and flight time, in hours). Variables related to piling events included duration of piling and duration of deterrence measures. Further aspects related to piling events included the amount of time elapsed since the last piling event stopped and the distance from the last piling event within a 60-km radius (distance in km). These last two variables were described as two-dimensional smooths using a tensor spline. A second set of models was set with the same variables as model ST1, but subsets of the data were used that contained a minor distance (model ST2 and 3) or time range (model ST4 and 5). Last, model ST1 ran per wind farm project to individualise the outcomes (model ST6 - ST14).

6.3 Results

6.3.1 Density and distribution of harbour porpoise sightings

Between January 2009 and December 2013, 458 aerial surveys covering 201,318 km of valid survey effort were conducted, during which 17,223 harbour porpoises were sighted (Table 6-1). These data originated from 13 different monitoring studies, each with a distinct spatio-temporal coverage of the study area (Figure 6-1, Figure 6-4). In this section, the available density estimates were visualised per month and season to show the variability of the entire dataset. To see how density estimates differed spatially, the representation per season was illustrated for the entire study area and the three subareas defined a priori (Figure 6-5). A monthly illustration per subarea was biased due to a limited number of surveys conducted in autumn and winter. Furthermore, distribution maps were generated per season and year (Figure 6-13). Lastly, the seasonal distribution patterns of all years were pooled together.

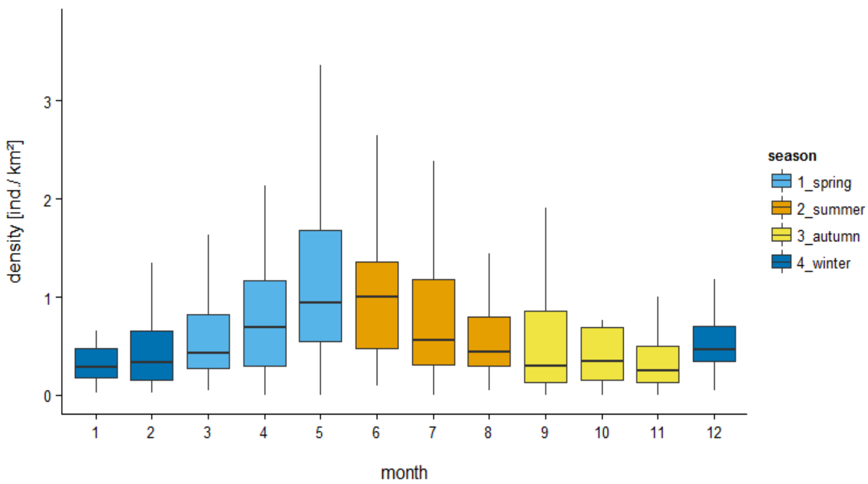


Figure 6-11 Monthly mean density estimates of harbour porpoise per flight in the German Bight, years 2009-2013 pooled. 95 % quantile: lower bar; 25 % to 75 % quantiles mark borders of upper and lower box, 5% quantiles are marked by upper error bar.

Density estimates pooled over the complete study area and years were found to differ significantly between months (Kruskal Test: Chi^2= 61.5242, df = 11, $p < 0.0001$) and seasons (Kruskal Test: Chi^2= 37.182, df = 3, $p < 0.0001$, Figure 6-11). The lowest densities were found in the autumn and winter months. The highest densities were typically found from April to June (0.7 – 1.0 ind./km²).

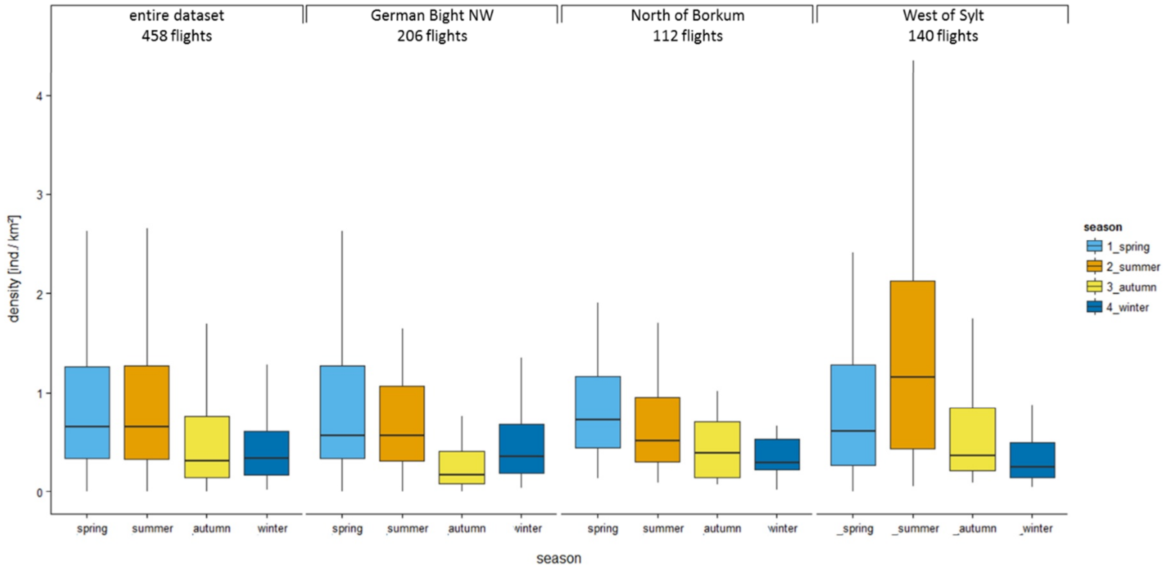


Figure 6-12 Box-Whisker plots showing seasonal density estimates of flight surveys conducted to count harbour porpoises in the German Bight years 2009-2013 pooled (95 % quantile: lower bar; 25 % to 75 % quantiles mark borders of upper and lower box, 5% quantiles are marked by upper error bar, compare to Table 6-5).

On a seasonal scale (entire dataset, Figure 6-12), densities during spring and summer were more similar to each other (median: 0.7 ind./km² each) than those of autumn and winter (median: 0.3 ind./km² each; Nemenyi test: spring-autumn: $p < 0.001$; spring-winter: $p < 0.001$; summer-autumn: $p < 0.001$, summer winter: $p < 0.001$). No differences were detected between spring and summer as well as between autumn and winter (Nemenyi test: p = ns). The effort per location was different during the five years and further, winter and autumn were surveyed more seldom than spring and summer (Figure A-31). Nevertheless, densities differed between years only in summer (Kruskal Test Summer: $Chi^2 = 16.186$, df = 4, $p < 0.0028$). Densities during summer in 2012 were lower than in 2011 and those of 2013 were higher than 2009 and 2012 (Nemenyi test: $p_{2009-2013}=0.039$, $p_{2011-2012}=0.034$, $p_{2012-2013}=0.016$). In 2012, relatively few surveys were conducted in summer because of bad weather conditions, especially between August and October. To address the lack of conducted flights, more surveys were conducted in November 2012 (Figure 6-2). To test whether different densities resulted from patchy spatial coverage or averaging; decreased seasonal differences, density estimates were studied hereafter per subarea.

Significant seasonal differences were present in all three subareas (Figure 6-11; Kruskal test: North of Borkum: $p < 0.01$, $Chi^2 = 11.45$, df = 3; German Bight NW $p < 0.001$, $Chi^2 = 19.396$, df = 3; West of Sylt $p < 0.001$, $Chi^2 = 20.665$, df = 3). Within the subarea “German Bight NW”, densities were higher in spring and summer than in autumn, but there was not a clear distinction between winter and the other seasons. Densities from that subarea were unusual, because only there were densities higher in winter than in autumn. In the subarea “North of Borkum”, densities in spring were higher than densities in autumn and winter. In comparison, at “West of Sylt”, the highest densities were detected in summer, with similar densities in spring and autumn and lower densities in winter, which deviated only from summer densities (Figure 6-11). Significance values of performed tests were summarised hereafter (Table 6-5). Furthermore, possible annual differences of seasonal density estimates were also tested, demonstrating that estimates from “North of Borkum" sampled in summer and those from "West of Sylt" sampled during autumn differed significantly between years (Figure A-31, Kruskal test: North of Borkum – summer: $p < 0.008$, $Chi^2 = 13.562$, df=4; West of Sylt – autumn: $p < 0.049$; $Chi^2 = 7.8422$, df=3). Significant annual differences were apparent only at "West of Sylt", with lower densities detected in 2012 than in 2011. As piling events started in that subarea in 2013, this difference might have resulted rather from different flight effort or location of the flight area within the subarea.

*Table 6-5 Nemenyi test of seasonal difference of harbour porpoise densities in different subareas (p-values corrected according to Tukey; significance level: ** p < 0.01; * p < 0.05).*

	German Bight NW			**North of Borkum**			**West of Sylt**		
	Spring	Summer	Autumn	Spring	Summer	Autumn	Spring	Summer	Autumn
Summer	0.960	-	-	0.045	-	-	<0.001	-	-
Autumn	<0.001	<0.001	-	0.001	0.306	-	0.479	<0.001	-
Winter	0.035	0.094	0.201	0.001	0.180	0.963	0.040	<0.001	0.449

Seasonal differences among the three subareas were only detected during summer (Kruskal test: $Chi^2= 17.093$, df = 2, $p < 0.001$), demonstrating that densities in the subarea "West of Sylt" were higher than in the other two subareas. These results were in line with the outcome of the seasonal distribution maps discussed in the following paragraphs (Figure 6-13).

In light of seasonally changing effort, weighted mean density values were calculated per year and season as a first step (Figure 6-13). Thereafter, seasonal values pooled over years were also illustrated (Figure 6-14). To provide a comprehensible but detailed view of the spatial distribution of harbour porpoises in the study area, the seasonal values are hereafter described considering the distribution shown in the first illustration.

In spring, high densities were recorded throughout the study area. Animals accumulated in and next to the SAC Sylt Outer Reef and next to Borkum Reef Ground. The accumulation increased in summer, exhibiting even higher densities in the SAC Sylt Outer Reef and fewer animals in Borkum Reef Ground, although those densities were still higher than in areas directly to the north or east. In autumn, a concentration of animals around Sylt Outer Reef was still evident, but the densities had decreased considerably. During winter, harbour porpoises were dispersed evenly throughout the German Bight. Higher densities were detected primarily at the north-western edge of the study area and more to the south, in the vicinity of Borkum Reef Ground.

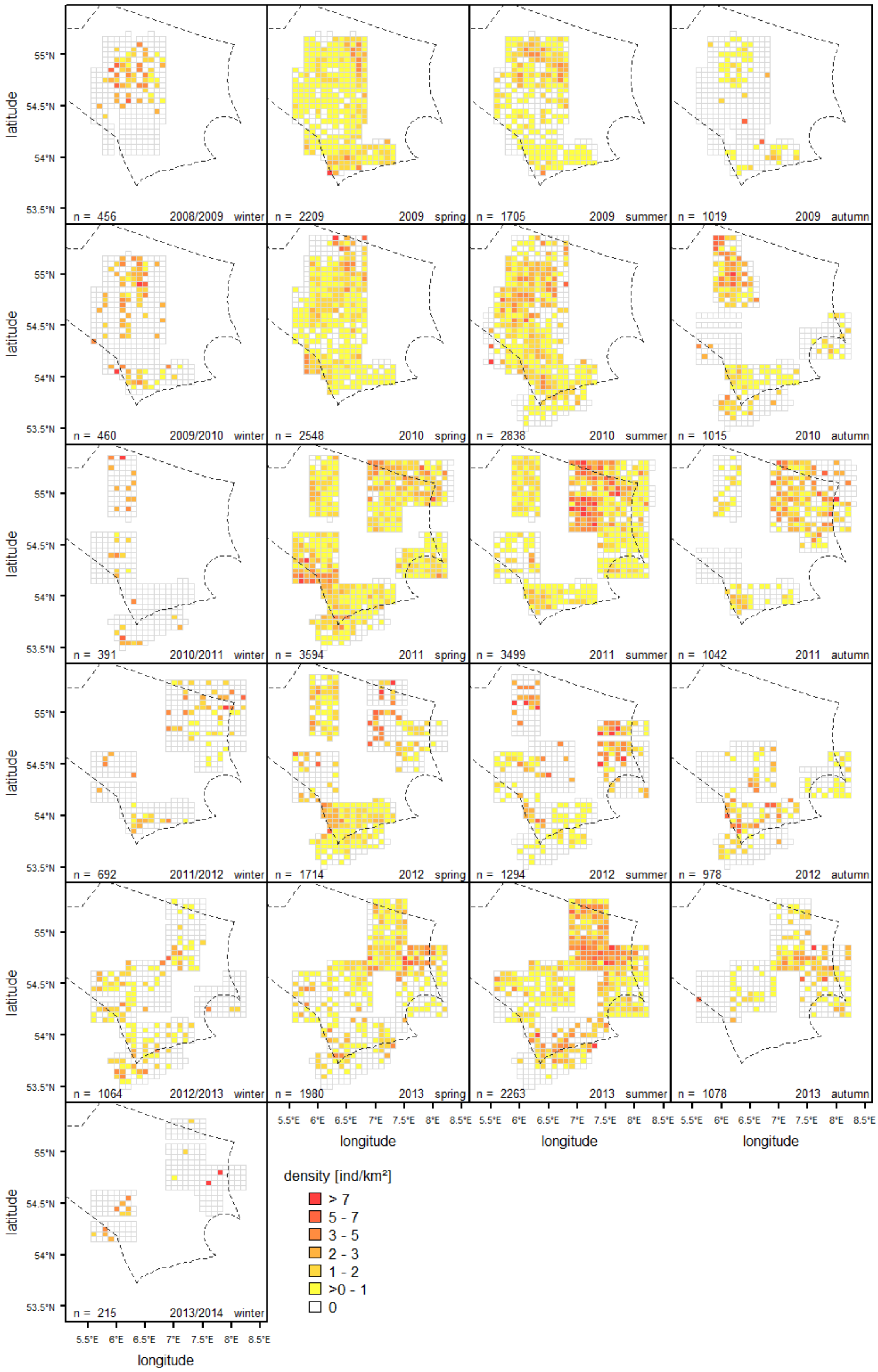


Figure 6-13 Density distribution of harbour porpoises in the German Bight between winter 2008/2009 and winter 2013/2014, illustrated separately per year and season. Aggregated data per grid cell (n) as weighted mean (see Figure A-31).

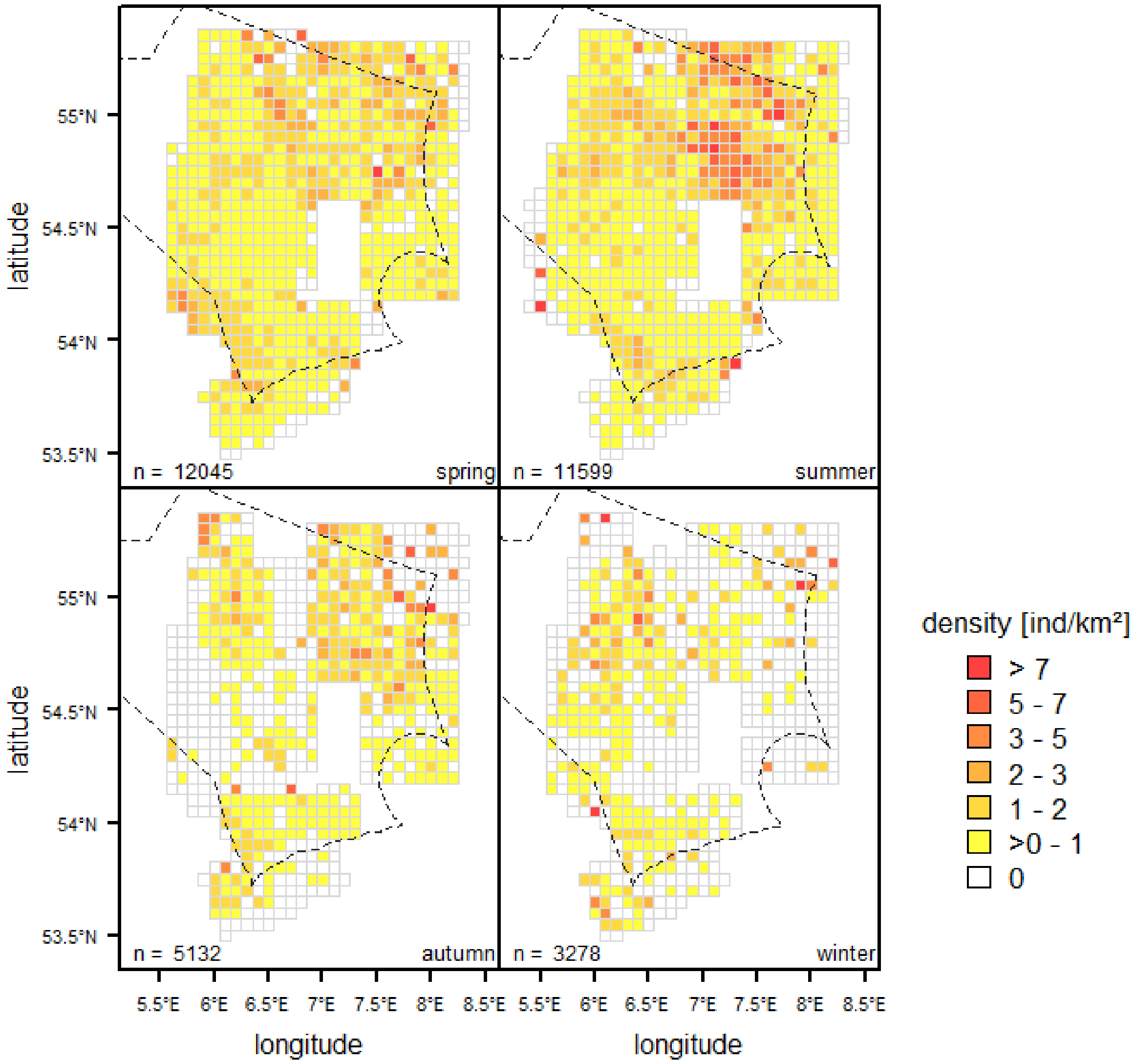


Figure 6-14 Seasonal density distribution of harbour porpoises in the German Bight between 2009 and 2013 illustrated separately per season. Aggregated data per grid cell (n) as weighted mean.

It should be considered that densities were pooled over several years, which could result in an over interpretation of the graphs. Consequently, the impression of an decreasing density gradient from the west to east in the northern part of the study area should be interpreted with care, because these areas were studied simultaneously only in spring, summer and autumn of 2010 as well as in spring 2011. In all other seasons, surveys were conducted only in one of the two areas (Figure 6-13).

6.3.2 Geographic distribution models of harbour porpoises

In this chapter, six different models were discussed that were numbered from 1 to 8 and given the prefix A (aerial) resulting in models A1 to A8. Table 6-6 summarises the objective of these models, the used data, area specifications and location of the results. The first three models show the general distribution patterns of harbour porpoises. These three models are almost identical but differ in the underlying data. A1 considered all data, while A2 considered all those data that were defined as been not affected by piling events, while A3 considered all data that have been affected by piling events (i.e. 0-60 hrs relative to piling and 0-60 km distant from piling events). The comparison of seasonal distribution map was illustrated in the subsequent text. The models A4 – A8 considered all data independent of piling events from spring and summer, but were limited to regions of consistent data availability (Table 6-6).

Table 6-6 Overview of the global aerial statistical models with respect to primary aim, data usage, position of results and area specifications.

global aerial model	primary aim	Data usage	specific area	model specifications
A1	testing annual density trend	n = 32,054 (all) all seasons, 2009-2013	n	Table A-13, Figure 6-15 - Figure 6-18
A2	testing annual density trend in unaffected data	n = 24,324 (unaffected) all seasons, 2009-2013	n	Table A-14, Figure 6-15 - Figure 6-18
A3	testing piling effect on distribution patterns and annual trend	n = 7,730 (affected) all seasons, 2009-2013	n	Table A-15, Figure 6-15 - Figure 6-18
A4	German Bight NW south: regional effect of piling activity on porpoise distribution	spring and summer, 2009-2013	y	Table A-16, Figure 6-19, Figure 6-22, Figure 6-23
A5	German Bight NW north: regional effect of piling activity on porpoise distribution	spring and summer, 2009-2012	Y	Table A-17, Figure 6-20 and Figure 6-19, Figure 6-22, Figure 6-23
A6	North of Borkum: regional effect of piling activity on porpoise distribution	spring and summer, 2009-2013	y	Table A-18, Figure 6-19, Figure 6-21, Figure 6-23
A7	West of Sylt: regional effect of piling activity on porpoise distribution	spring and summer, 2011-2013	y	Table A-19, Figure 6-19, Figure 6-22, Figure 6-23
A8	German Bight NW: regional effect of piling activity on porpoise distribution	spring and summer, 2009-2013	y	Only in appendix, Table A-20

Geographic distribution models of the entire study area

After depicting the general distribution of harbour porpoises in the last chapter, the successive aim of the study was to analyse how piling events affected the observed distribution patterns of harbour porpoises. This analysis was carried out using Generalised Additive Models (GAMs), a modelling technique also used in other studies of this kind (AGENCY & SAFETY 2014; GILLES et al. 2014A; VIQUERAT et al. 2015). Three different datasets were selected for this purpose. The entire dataset (n = 32,054), the unaffected dataset (n = 24,324; 75 % of the entire dataset) and the affected dataset (consisting of the remaining 25 % of the data, n = 7,730). The temporal coverage and the spatial extent of the three datasets differed, especially between the affected and the two other datasets.

The following variables were considered in the three models: the day of year grouped per year, coordinates grouped per season, year and season as a factor, the Position ID of the grid cells with a neighbouring matrix to compensate for spatial autocorrelation, the flight time, the hour of the day, the moon illumination as an indication for the oceanography, the anomaly of the sea surface temperature and the water depth. Furthermore, in model A1 that used the entire dataset, the potential impact of piling events was parameterised using the categorical variable “day and distance”, a combination of distance to piling (0-20, 20-40 and 40-60 km) and the time passed since piling ended (1, 2 or three days). Other variables tested but not included in the final model were substrate, biozone and observation height of the plane (250 or 600 feet) as random factors. Furthermore, the variables indicating cumulative piling events in the previous 15 days in 20, 40 and 60 km were tested, too. The latter variables and the random factors previously described were excluded to gain more parsimonious models (delta AIS < 2).

Tables summarising the model output of the models A1, A2 and A3 are included in the Appendix (Table A-13, Table A-14 and Table A-15). The variables used in the model A1 explained 12.6 % of the deviation in the data, model A2 that used only unaffected data explained 9.0 % and model A3 that considered only affected data explained 9.0 %. The effects of the seasonality, geographic distribution as well as the partial effect of the year and the categorical variable “day and distance” are illustrated in this chapter (Figure 6-15, Figure 6-16, Figure 6-17 and Figure 6-18). Further information and illustration of variables are included in the Appendix (A.6).

The distribution patterns modelled using the entire dataset (A1; Figure 6-15, left) demonstrated that the distribution of harbour porpoises changed seasonally during the sampled years. Specifically, in spring the density was higher in the western edge of the study area. From there, an increased probability of higher densities was registered towards the northern edge of the study area, west of the SAC Sylt Outer Reef. In comparison, the modelled density estimates were lower in the eastern and the southern parts of the German Bight. In summer, the northern part of the study area exhibited higher densities than the southern part (south of 54.5° N). The highest estimates were registered along the northern edge, close to the border between the German and the Danish Exclusive Economic Zones (EEZ), and in the north-western part, the area in the direction of the Dogger Bank. In comparison, the lowest estimates were located in coastal areas and near the Dutch EEZ. In autumn, the estimates in the north and north-eastern part were still higher. In the area close to the SAC Borkum Reef Ground, in the south-western part of the study area, however, the density estimates increased and were comparable to those of Sylt Outer Reef. In winter, the lowest estimates were located in the area north of Helgoland, almost entering Danish waters. The higher values were, conversely, associated with the north-western part of the German Bight and the north-western part of the Borkum area.

The graphical representation of model A2 (unaffected data only; Figure 6-15, middle) exhibited, in general, a similar seasonal distribution of harbour porpoises when compared to those of model A1 (entire dataset). For example, in spring the area from the Dutch-German border of the EEZ to the northern edge of the study area showed significantly higher densities in the reduced dataset, while the model A2 did not exhibit confidence bands in this area.

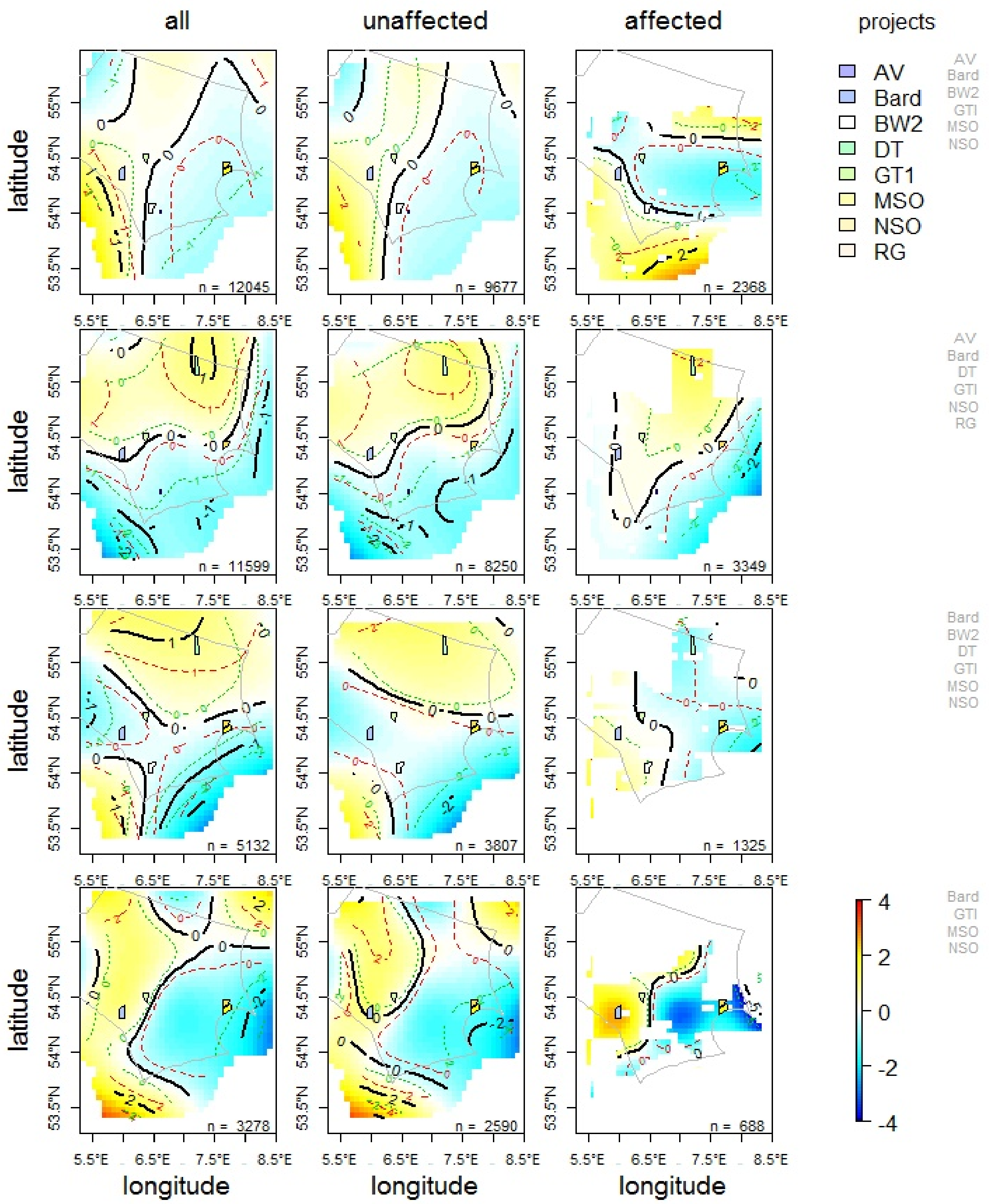


Figure 6-155 Distribution patterns of harbour porpoises from model A1, A2 and A3 (left, middle and right columns). GAM-plots illustrating the two-dimensional smooth of longitude and latitude per season (from spring in the upper line to winter in the lower line); wind farm locations were included, if piling occurred during that season (black lines: mean density, dotted lines confidence band of mean; red: lower green: upper, i.e. densities significantly above or below mean; green solid lines: isoclines <> 0; higher densities yellow, lower densities red).

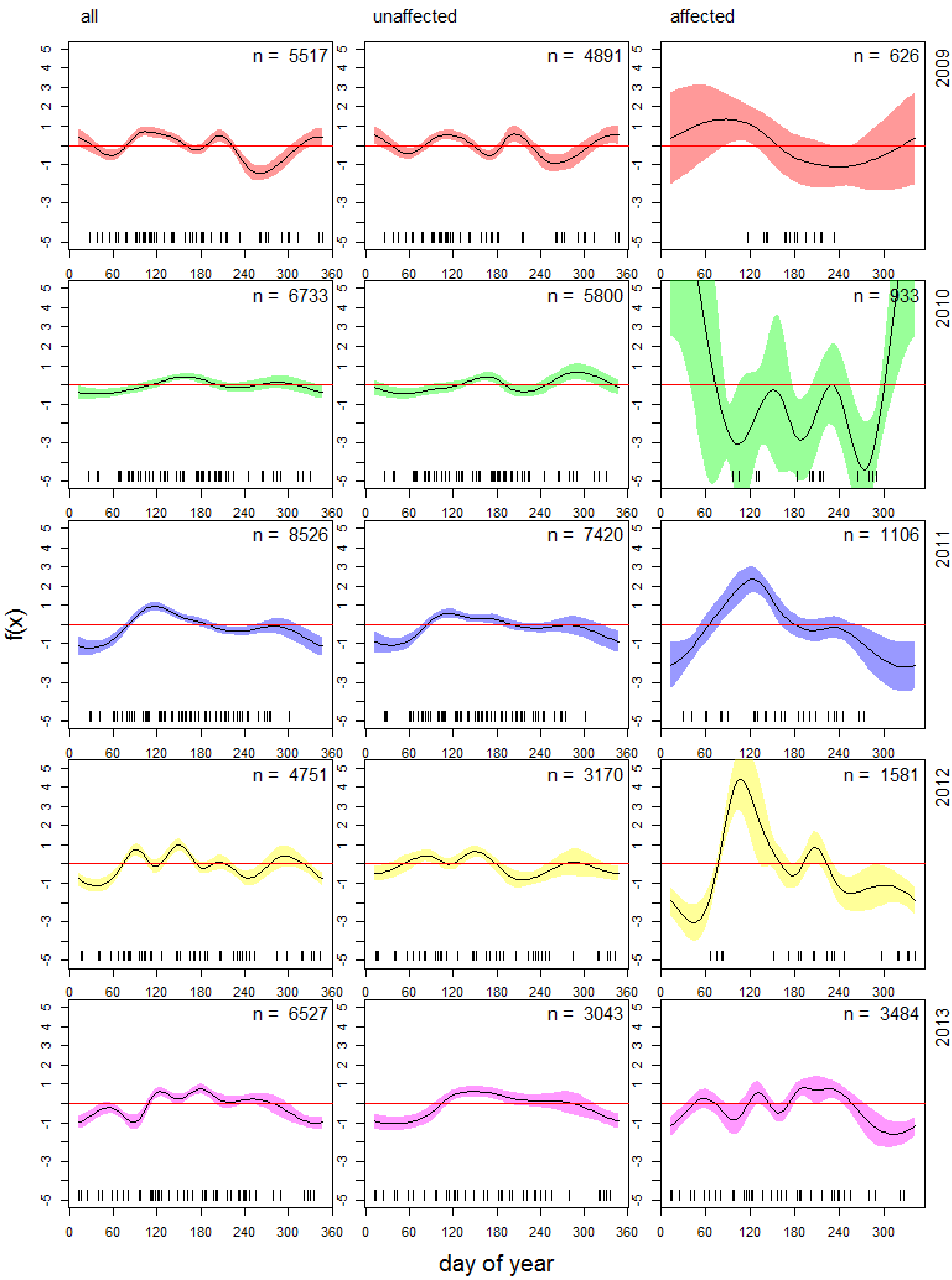


Figure 6-16 Seasonality of porpoise densities – smooth splines of day of year per year of model A1, A2 and A3 (upper and lower graphs; Red lines: mean density estimate - if the shaded confidence intervals surpass the red line, the effect of the value on the x-axis is significantly positive or negative, resulting in higher or lower densities than the mean, n: number of considered grid cells).

The gradients in spring and summer followed slightly different patterns, but were generally very similar. Conversely, the distributions of autumn and winter were more heterogeneous. Furthermore, the positions of GTI and BWII were within regions with significantly lower densities when considering only unaffected data.

The graphical representation of model A3 (Figure 6-15, right), which used the affected dataset, showed a different seasonal distribution of harbour porpoises than the other two models. The data coverage in spring was complete, but the area north of 55.0° N was not covered. Consequently, only parts of the SAC Sylt Outer Reef known for high densities in spring were included. Nevertheless, the illustration indicates high densities north of 54.75° N. As in the other two models, higher densities were apparent west of AV and BARD, next to the SAC Borkum Reef Grund. The best spatial coverage of this model was in summer. In this season, highest densities were estimated in the SAC Sylt Outer Reef showing a gradient from north to south. This representation was more similar to the other models from all other seasons. The visual representation of autumn was very different from that of the entire and unaffected datasets. Densities in the SAC Sylt Outer Reef were low, especially at DT and to the west of it; however, densities at BARD and GTI were high. Both were in contradiction to results of the entire and unaffected datasets (A1 and A2). Estimates in winter were based on piling events at GTI, BARD, MSO and NSO. The visual representation was different from the other two models but showed a similar gradient from high densities in the west to low densities in the east.

The modelled smooths representing the seasonal density pattern of harbour porpoises (day of year) were illustrated per year. In the left column (model A1), the smooths of the entire dataset were represented (Figure 6-16, left). Densities of that model show generally that the density in all years exhibited highest values in spring and summer and the lowest in autumn and winter. Densities were significantly higher during winter only in 2009. The smooths of 2011 and 2012 displayed the highest similarity. Further similarities were visible concerning the intersection with the mean (red line in Figure 6-16). In 2009, 2011 and 2012 the mean value was reached simultaneously during the second and third week of March, while in 2010 and 2013 it occurred for values registered during the second week of April. Comparing the model A2 (Figure 6-16, middle) with A1 revealed that the smooths of each year differed between the two models as was expected, due to a 25 % reduction of data. The seasonal curves were generally similar to those of A1, but in 2012 and 2013, which were the years with the highest number of piling events, a higher reduction of data occurred (Figure 6-10). In 2012, the local maxima with higher densities in the end of the summer and in autumn (day of year 210 and 310) were not visible in model A3. Similarly, the local maximum in late winter (day of year day 60) was not present and furthermore the curve was smoother than in the model A1. In contrast, the seasonal smooths of the model A3 exhibited higher deviations from smooths than model A1, as was the case for model A2. The number of days on which the affected dataset was based on was very low (2009: 17, 2010: 16, 2011: 25; 2012: 22, 2013: 41). Both 2009 and 2010 had the lowest number of days with flights, which in turn resulted in the widest confidence limits.

Inter-annual differences were visualised in partial effect plots. Model A1 showed that in 2009 and 2012, densities were low, and no differences were apparent between 2010, 2011 and 2012, while densities in 2013 were highest (Figure 6-17, left). Model A3 showed a stronger difference between 2009 and 2013 in comparison to intermediate values of 2010, 2011 and 2012. The difference between 2013 and the other years was stronger in the model of the unaffected dataset (A2,

Figure 6-17, middle) than in the entire dataset (A1). As stated in the last paragraph, it has to be taken into account that the highest proportion of flights related to piling events was in 2012 and 2013. Therefore, these years had the highest reduction of data (Figure 6-10). In contrast to the first two models, the affected data showed no annual differences (A3; Figure 6-17, right).

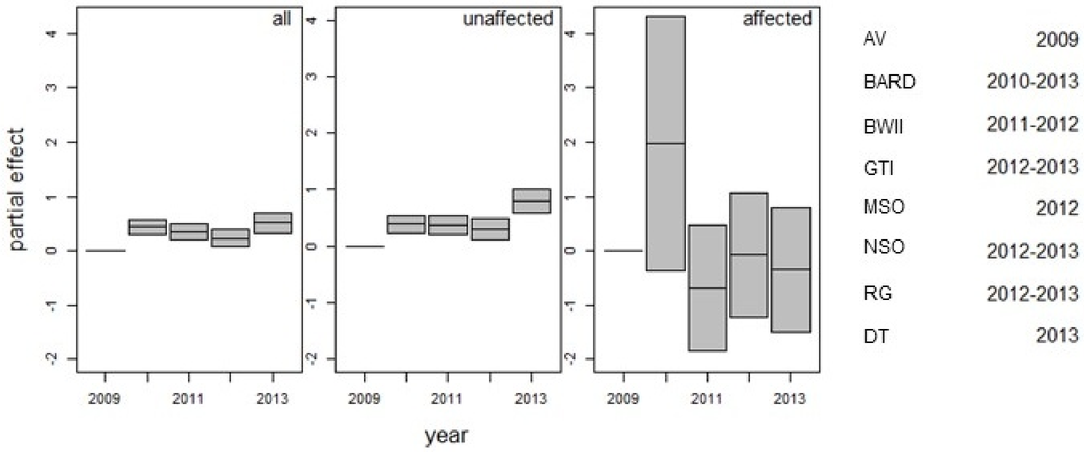


Figure 6-17 Inter-annual trend in porpoise densities - comparison of model A1, A2 and A3 (left, middle and right columns). GAM-plot illustrating the partial effect of year.

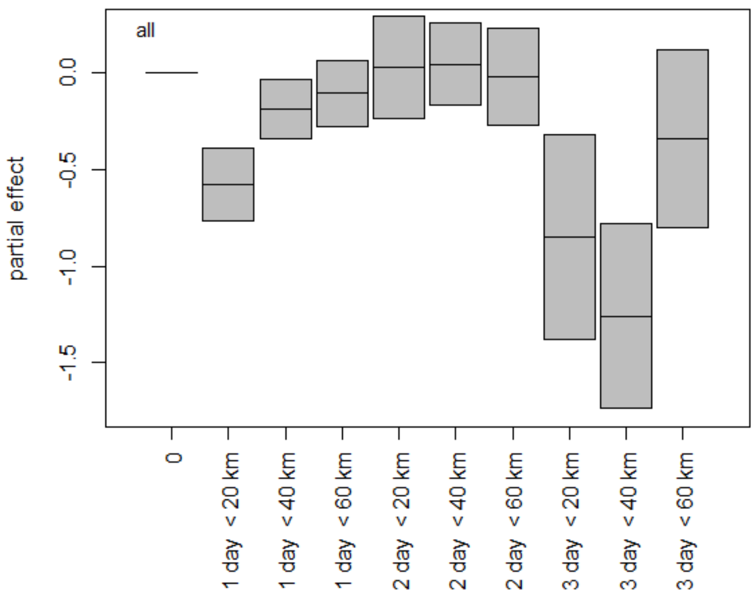


Figure 6-18 Model A1- partial effect of variable "day and distance" (0: unaffected area or timeframe; < 20 km: piling occurred within a radius of 0 to 20 km; < 40 km: piling occurred within a radius of 20 to 40 km, < 60 km: piling occurred within a radius of 40 to 60 km, 1, 2 and 3 piling refers to the time in days passed since piling ceased, black boxes represents data availability).

The effect of piling distance by day (day and distance) was only incorporated in model of the entire dataset (A1). It revealed that the density estimates were likely to be lower on the first and third day after piling in distances of up to 40 km, whereas no significant effects could be found in distances above 40 km and on the second day after piling (Figure 6-18). There were few data available from the third day, which most likely affected this result. Those data were pooled from grid cells sampled 48 to 60 h after piling ended. Effects of piling typically last 24 to 48 h (TOUGAARD et al. 2006; BRANDT et al. 2011; DÄHNE et al. 2013a). To have a sufficient amount of data as reference at the beginning of the study, a timeframe of 2.5 days after piling ended was chosen.

Geographic distribution models of the three subareas

The study area was split in this section of the chapter into the three subareas as described above (Figure 6-6), because some of the areas were not covered in all years. A preliminary analysis comparing the entire, the unaffected and the affected dataset per season indicated that in those seasons with wind farm construction, the distribution patterns of harbour porpoises was reduced around piling events (chap. 6.3.2, Figure A-32 to Figure A-34), with the exceptions of GTI and BARD, where harbour porpoises showed higher densities in winter and autumn. It was concluded that either the models were not complex enough to analyse the nature of the data or effects of piling might have been less important than large-scale density gradients. Due to data availability and time resolution (3 vs. 5 years) the model outputs differ (Figure A-23 to Figure A-30).

Hereafter, the analysis focused only on data from spring and summer when densities were highest, due to the scarcity of data from autumn and winter. Furthermore, due to incomplete coverage in 2013, the subarea "German Bight NW" was split into a northern and southern part (but see model A8, Table A-20). Density patterns in spring and summer were analysed per year and illustrated for each of the four subareas (Figure 6-19, Figure 6-20 and Figure 6-21). Model outputs were summarised in the Appendix (Table A-16, Table A-17 and Table A-19 Table A-19).

Within the northern part of subareas German Bight NW (model A5), where no wind farm construction occurred, porpoise densities did not significantly change between 2009 and 2012 (Figure 6-22). They were slightly lower in 2012, but large confidence intervals indicate that this was based on a low sample size. As well, the variable „day and distance“ did not have a significant effect on harbour porpoises (Figure 6-23). The distribution of porpoises changed between the four years, but the distribution patterns in 2010, 2011 and 2012 seemed not to be affected by the piling events (Figure 6-19).

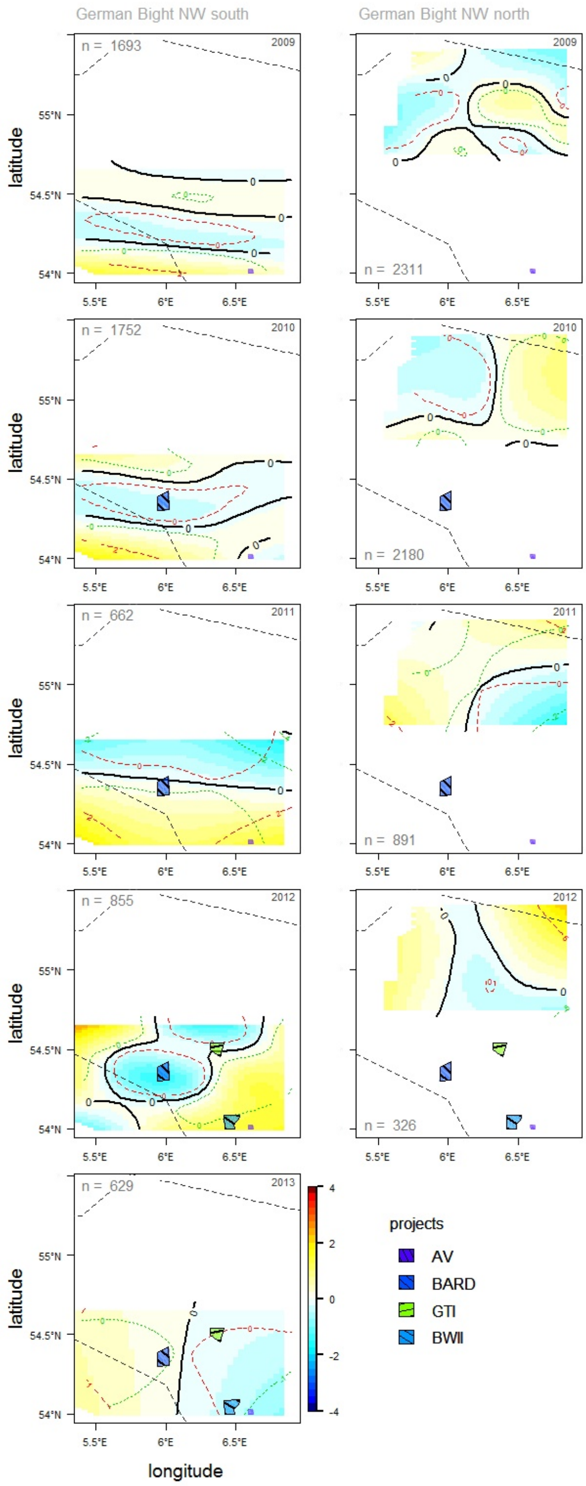


Figure 6-19 Annually changing distribution patterns of harbour porpoises in the subarea "German Bight NW" model A4 and A5 spring and summer pooled. GAM-plot showing tensor spline smooths of latitude and longitude per year for the southern and northern part separately (black lines: mean density, dotted lines confidence band of mean; red: lower green: upper, i.e. densities significantly above or below mean; green solid lines: isoclines <> 0). Considered piling events: 2009: AV 14; 2010 BARD: 10; 2011: BARD: 18; 2012: BARD: 6; GT1: 1; BWII: 4; 2013: GTI: 17.

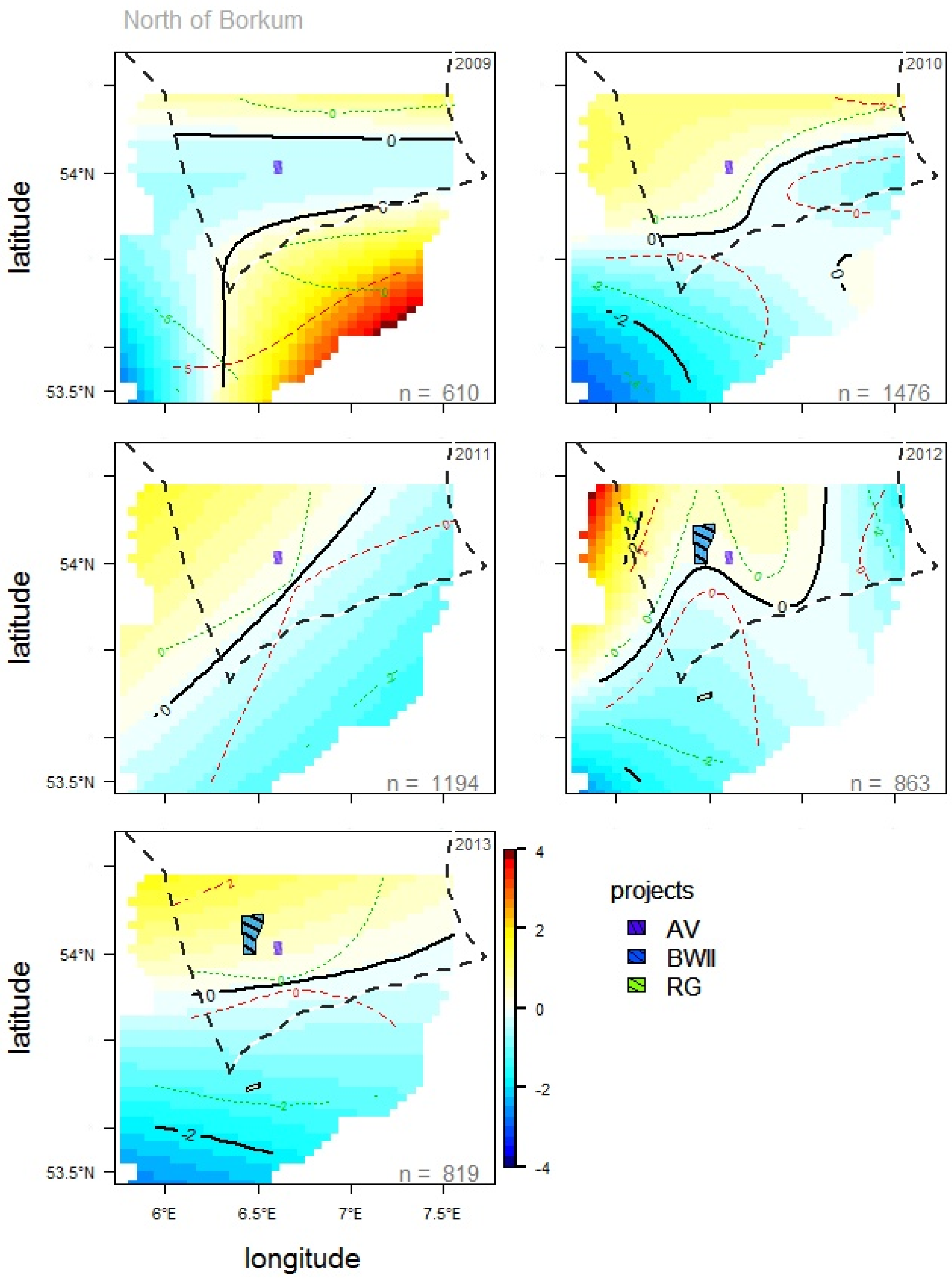


Figure 6-20 Annually changing distribution patterns of harbour porpoises in the subarea "North of Borkum" model A6 spring and summer pooled. GAM-plot showing tensor spline smooths of latitude and longitude per year (black lines: mean density, dotted lines confidence band of mean; red: lower green: upper, i.e. densities significantly above or below mean; green solid lines: isoclines <> 0); Considered piling events: 2009: AV 14; 2012: BWII: 4; RG: 5.

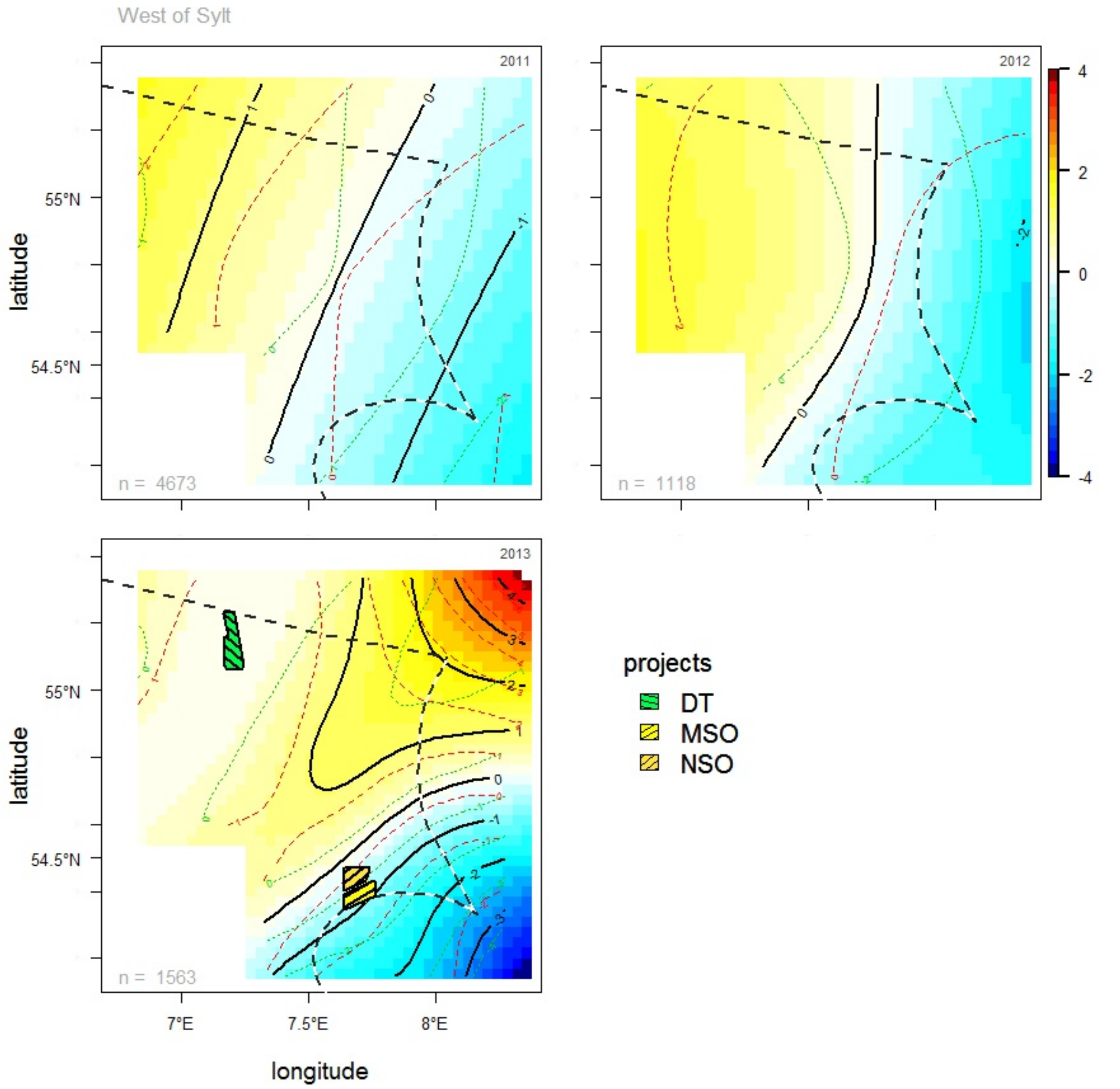


Figure 6-21 Annually changing distribution patterns of harbour porpoises in the subarea "West of Sylt" model A7 spring and summer pooled. GAM-plot showing tensor spline smooths of latitude and longitude per year (black lines: mean density, dotted lines confidence band of mean; red: lower green: upper, i.e. densities significantly above or below mean; green solid lines: isoclines <> 0); Considered piling events: 2013: DT: 5; NSO: 9; MSO: 5.

Within German Bight NW (model A4), the only significant inter-annual differences occurred between 2009 and three other years (2010, 2011 and 2013), with 2009 showing lower densities; During 2012, when GTI and BARD were under construction in this area, densities tended to be lower than in other years apart from 2009, but the differences were not significant. In the previous years, piling events occurred only at BARD (Figure 6-22). In the first three years, there was a similar density pattern exhibiting high densities in the southwestern edge next to the SAC Borkum Reef Ground and in the north-eastern edge of the subarea at the German Danish boarder (Figure 6-19). Conversely, in 2012 and 2013 there was a pattern exhibiting lower densities in the south-east and higher densities in the northwest. Densities were significantly lower around BARD in 2012 and around GTI in 2013, but not in 2009, 2010 and 2013. Interestingly in 2012 activity next to GTI and BWII were not as low as at Bard. However, data from 2012 included only one flight at GTI but four at BWII and six at BARD.

Densities at the SAC Borkum Reef Ground are typically high in spring and might have been more important than effects of pilings. Looking at the effect of piling (Figure 6-23) densities were significantly lower on the day of piling at distances between 0 and 20 km and between 40 and 60 km. On the second day, no effect was present in any of the three distance classes and on the third day, densities were only lower in 20 to 40 km distance. In the northern part of the study area, no effect was visible.

At "North of Borkum" (model A6), densities were not different among the five years (Figure 6-22). In contrast, the distribution differed in all five years (Figure 6-20). Generally, there was a gradient with lower densities along the coast and higher densities to the north, with the exception of 2009. In that year, densities were significantly lower in the west and significantly higher along the northern and southern edge of the subarea. Around AV, densities were not significantly different from the mean. The isocline splitting the southern and northern part differed during the other four years. In 2010, 2011 and 2013 densities were higher north of AV and declined towards the coast. In 2012, when construction of RG and BWII occurred, densities were higher north of BWII. Looking at the effect of piling (Figure 6-23), densities were significantly higher on the day of piling at distances between 0 and 60 km, on the second day no effect was present at any of the three distance classes, and on the third day densities were lower at the distance of 20 to 40 km.

At "West of Sylt" (model A7), densities were not different between the three years (Figure 6-22), however, the distribution pattern in spring and summer differed between 2013 and the other two years (Figure 6-21). Densities were higher in spring than in summer and including this variable improved the model. In 2011 and 2012, there was a density gradient from the coast of Schleswig-Holstein to the German Bight, with higher densities out at sea than along the coast. In 2013, when construction occurred at NSO, MSO and DT, a band with significantly higher densities stretched from the north-eastern corner of the subarea in a south-easterly direction between DT and the construction sites of MSO and NSO, while densities around the construction sites were either below the mean (MSO and NSO) or not different from the mean (DT). Looking at the effect of piling (Figure 6-23), densities were significantly higher on the day of piling at distances between 20 and 60 km, on the second day densities were lower at a distance of 40 to 60 km and on the third at a distance of 20 to 40 km.

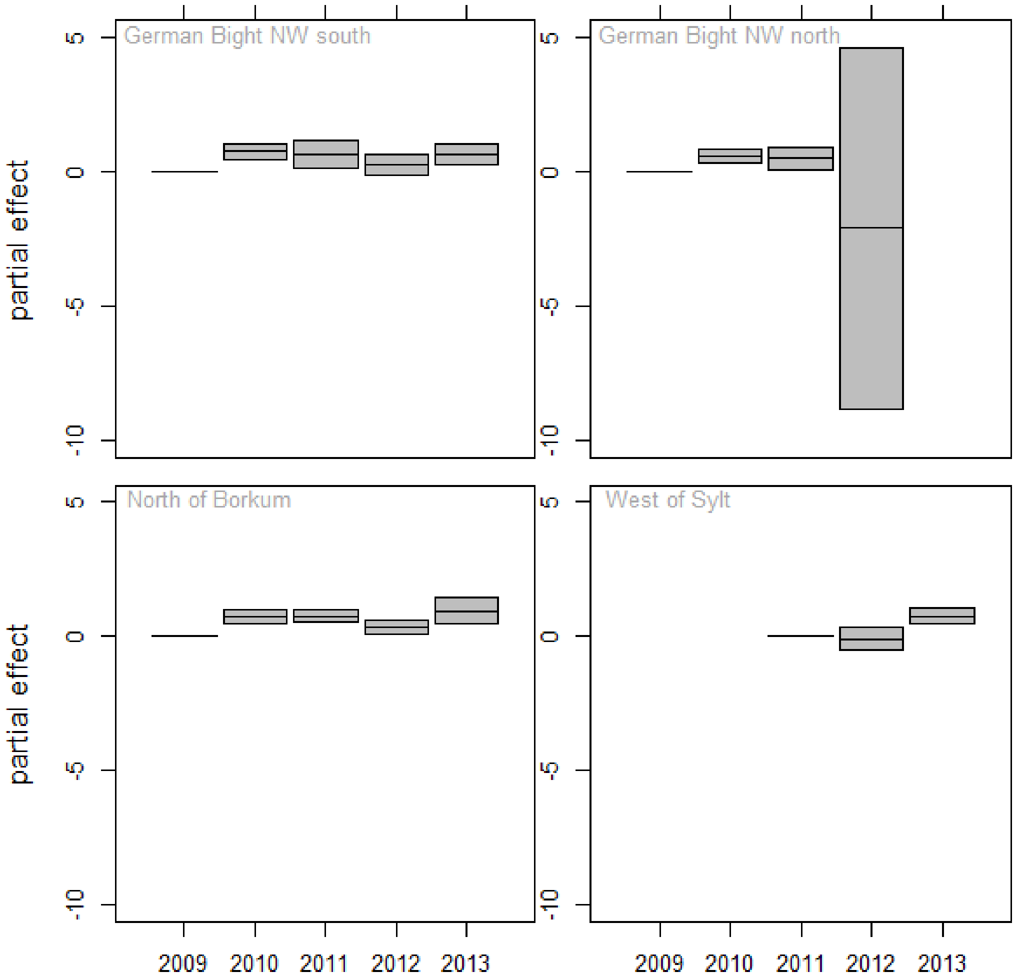


Figure 6-22 Inter-annual trend in porpoise densities comparison of model A6 – A8 (Considered piling events – model A4 and A5 - German Bight NW south and north: 2009: AV 14; 2010 BARD: 10; 2011: BARD: 18; 2012: BARD: 6; GT1: 1; BWII: 4; 2013: GTI: 17/ model A6 - North of Borkum: 2009: AV 14; 2012: BWII: 4; RG: 5/ model A7 - West of Sylt: 2013: DT: 5; NSO: 9; MSO: 5).

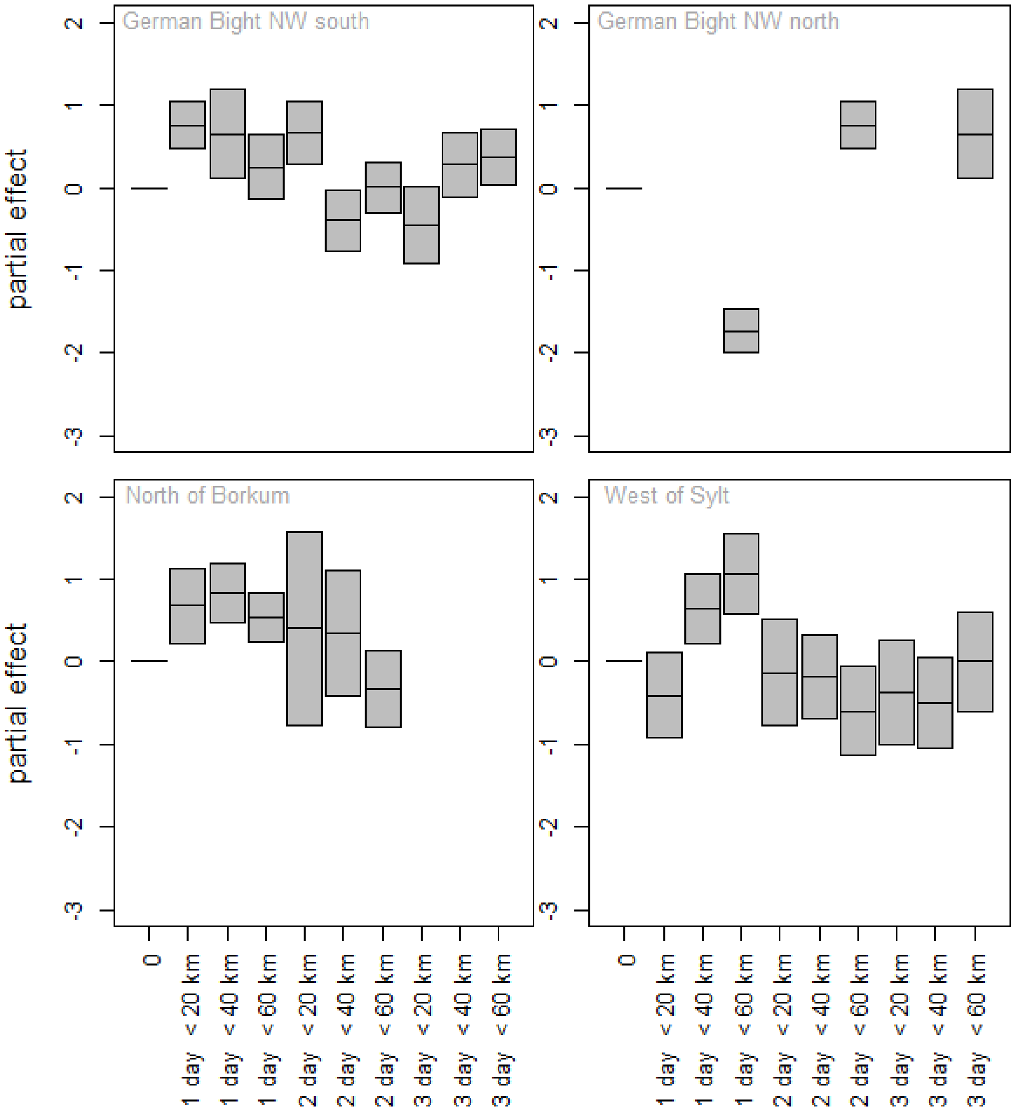


Figure 6-23 Partial effect of variable „day and distance" in models A5 to A8 (0: unaffected area or timeframe; < 20 km: piling occurred within a radius of 0 to 20 km; < 40 km: piling occurred within a radius of 20 to 40 km; < 60 km: piling occurred within a radius of 40 to 60 km. The numbers 1, 2 and 3 refer to the time in days since piling ceased. Considered piling events – model A3 and A4 - German Bight NW south and north: 2009: AV 14; 2010 BARD: 10; 2011: BARD: 18; 2012: BARD: 6; GT1: 1; BWII: 4; 2013: GTI: 17/ model A6 - North of Borkum: 2009: AV 14; 2012: BWII: 4; RG: 5/ model A7 - West of Sylt: 2013: DT: 5; NSO: 9; MSO: 5).

6.3.3 Short term effects of piling events on harbour porpoise densities

After analysing the data in the previous chapter, which relates to large scale and long-term effects of piling events on the distribution patterns of harbour porpoises, the emphasis of this chapter was placed on short-term effects. As described above, the categorical variable „day and distance" has been included in the models that used both affected and unaffected data but not in those models that used affected data only (Figure 6-18, Figure 6-23). This variable encodes the spatio-temporal association of grid cells to piling events by blending three distance (0-20, 20-40 and 40-60 km) with three temporal classes (0-24, 24-48 and 48-60 h after piling ended) and using “0” as a reference for no piling. Presumably, the definition of this variable was defined too broad to provide the GAMs enough flexibility in the modelling process, since this variable remained in the most parsimonious models only if both affected and unaffected grid cells were used in the models, but not when only affected data were considered (delta AIS <2).

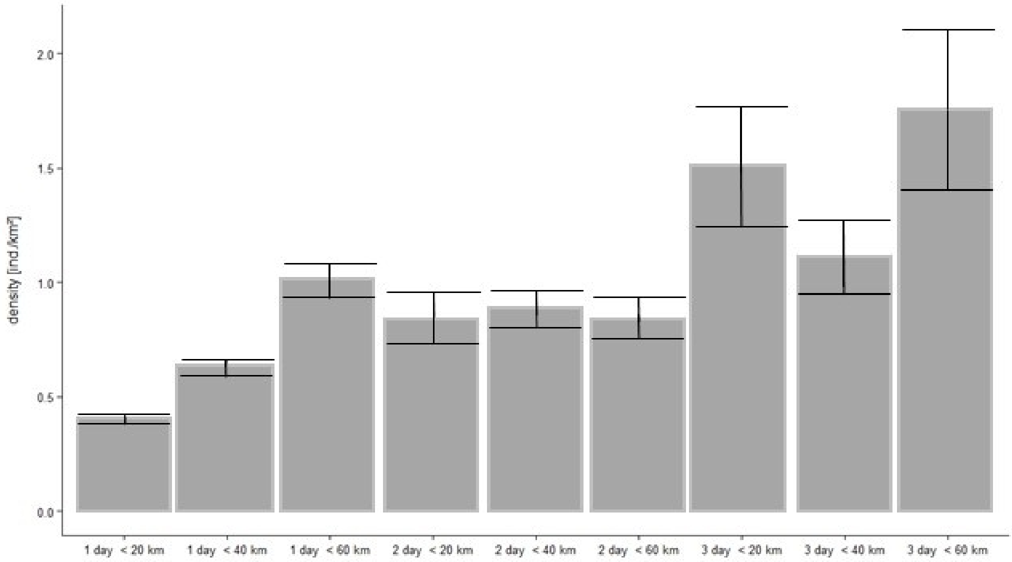


Figure 6-24: Barplot of harbour porpoise densities grouped by “day and distance” (illustration of mean and standard deviation).

*Table 6-7: Significance values of Nemenyi test of variable “day and distance” (p-values corrected according to Tukey; significance level: ** p < 0.01; * p < 0.05).*

Time [h]		**0-24 h**			**24-48 h**			**48-60 h**	
	Distance [km]	**0 - 20 km**	**20 - 40 km**	**40 - 60 km**	**0 - 20 km**	**20 - 40 km**	**40 - 60 km**	**0 - 20 km**	**20 - 40 km**
0-24 h	20 - 40 km	0.294	-	-	-	-	-	-	-
	40 - 60 km	<0.001 **	0.010 *	-	-	-	-	-	-
24-48 h	0 - 20 km	0.614	1.000	0.484	-	-	-	-	-
	20 - 40 km	0.173	0.997	0.506	1.000	-	-	-	-
	40 - 60 km	0.002 **	0.250	1.000	0.766	0.837	-	-	-
48-60 h	0 - 20 km	0.012 *	0.151	0.909	0.318	0.375	0.918	-	-
	20 - 40 km	0.007 **	0.133	0.949	0.332	0.390	0.956	1.000	-
	40 - 60 km	0.003 **	0.054	0.760	0.160	0.191	0.790	1.000	1.000

In advance of a more detailed model-based analysis, it was tested whether harbour porpoise densities of affected grid cells differed when separately analysing densities per distance class and time passed since piling ended (Figure 6-24). Densities differed significantly (Kruskal test: Chi² = 104.63, df = 8, $p < 0.001$). In the first 24 h after piling ended, mean densities were significantly higher at distances of 40-60 km than at 0-20 and 20-40 km. Furthermore, densities at the 0-20 km distance on the first day after piling were significantly lower than the densities at a distance of 40-60 km on the second day after piling and all distances on the third day after piling (Table 6-7).

Splitting the estimates according to season revealed that densities were more likely to differ between the three distance classes in spring and summer (Figure A-35; Kruskal test: spring: p = 0.026, Chi² = 7.28, df = 2; summer: $p < 0.001$, Chi² = 54.29, df = 2; autumn: p = ns; winter: p = ns). However, only in summer were these differences also significant in a post-hoc test (Nemenyi: psummer_1 = 0.26; psummer_2 = < 0.001; psummer_3 < 0.001). Densities in summer were generally higher and the data availability was also higher. Therefore, the potential to detect differences was correspondingly higher, as well.

6.3.4 Modelling the effect of pile driving on porpoise densities

In this chapter, the spatial- temporal effect of piling events was analysed. Seven GAM-models were discussed that were numbered from 1 to 7 and given the prefix ST (spatio-temporal) resulting in models ST1 to ST7 (Table 6-8). The first six models tested the effect of piling events pooled for all wind farms. For the reason that generalised additive models are sensitive to data availability when estimating the mean and the deviation from it, respectively, in different spatial and temporal windows where selected to test the sensitivity of the model (Table 6-8).

Despite the high number of aerial surveys considered in the current study, data availability proved to be not high enough to run project-specific models. As an example, the results of model ST7 were illustrated in the annexe. In model ST7 distance from piling, hour relative to piling and wind farm were modelled as a three-way interaction term.

In the previous chapters, it was demonstrated that piling events affected the density and the distribution of harbour porpoises in the study area. Their density was reduced after piling stopped, which might have been a result of animals swimming away from the construction site. The effect of noise pressure levels on the densities of harbour porpoises was modelled as part of a preliminary analysis. However, due to the size of the grid cells, the noise values were not available at the needed resolution. The analysis presented hereafter was adopted, and instead of noise pressure level, the distance from piling events was considered as the principal variable. The noise pressure level decreases with increasing distance from its source and can adequately replace noise as a variable (THIELE 2002).

Table 6-8: Overview of the spatio-temporal models with respect to primary aim, data usage, position of results and area specifications.

spatio-temprol model	primary aim	data usage	project-specific	model specifications
ST1	testing spatio-temporal effect of piling events on harbour porpoises	n = 6,871 (affected), 0 – 60 km & 0 – 60 h, all wind parks, 2009 – 2013	n	Table A-21; Figure 6-25
ST2	testing model ST1 for data reduction	n = 4,675 (affected), 0 – 40 km & 0 – 60 h, all wind parks, 2009 – 2013	n	Figure A-36; Table A-22
ST3	testing model ST1 for data reduction	n = 1,819 (affected), 0 – 20 km & 0 – 60 h, all wind parks, 2009 – 2013	n	Figure A-36; Table A-23
ST4	testing model ST1 for data reduction	n = 1,819 (affected), 0 – 60 km & 0 – 48 h, all wind parks, 2009 – 2013	n	Figure A-36; Table A-24
ST5	testing model ST1 for data reduction	n = 1,819 (affected), 0 – 60 km & 0 –24 h, all wind parks, 2009 – 2013	n	Figure A-36; Table A-25
ST6	testing model ST1 for data reduction	n = 3,151 (affected), 0 – 40 km & 0 – 40 h, all wind parks, 2009 – 2013	n	Figure A-36; Table A-25
ST7	testing spatio-temporal effect of piling events on harbour porpoises per windfarm project	n = 6,871 (affected), 0 – 60 km & 0 – 60 h, all wind parks, 2009 – 2013	y	Figure A-38; Table A-26

The spatio-temporal effect of piling events on harbour porpoises was considered with the two variables, “distance” and “hour relative to piling”, that were modelled as a two-dimensional smooth (model ST1). Further variables considered in the model were the day of year grouped per year, coordinates grouped per season, year and season as a factor, the Position ID of the grid cells with a neighbouring matrix to compensate for spatial autocorrelation, the flight time, the moon illumination as an indication for the oceanography, the water depth, the length of deterrence and piling. Furthermore, the cumulative effects of piling events in the previous 15 and 30 days at 20, 40 and 60 km were also tested, but they were not included, as a more parsimonious model without these variables was preferred (delta AIC < 2). Tables summarising the model output are included in the Appendix (Table A-21). The variables considered in the model explained 25.9 % of the deviation of the affected dataset. The effects of distance and time since piling ended are illustrated in this chapter (Figure 6-25). Further information and illustration of variables are included in the Appendix (Chap. A.6.2).

During piling, the densities of harbour porpoises were lower at a distance of up to 19 km and higher at 22 km (Figure 6-25). Within a distance of 0 to 6 km from the piling site, the density was likely to be low for the first 26 hours after piling ceased. At distances of 30 to 50 km, the increased density was apparent only for a short time (0-10 h). Thereafter, this effect turned negative and stayed negative for almost one day at 30 km distance and even longer with increasing distances. Before interpretation of these results the histograms illustrating data availability should be taken into consideration. The spatial coverage increased with increasing distance from the piling event. Furthermore, temporally coverage was only high during piling with more than 1,000 grid cells covered and decreased strongly during the following 48 h. Already 20 h after stop of pile driving only less than 50 grid cells could be considered, so that all interpretation beyond this time should be taken with care.

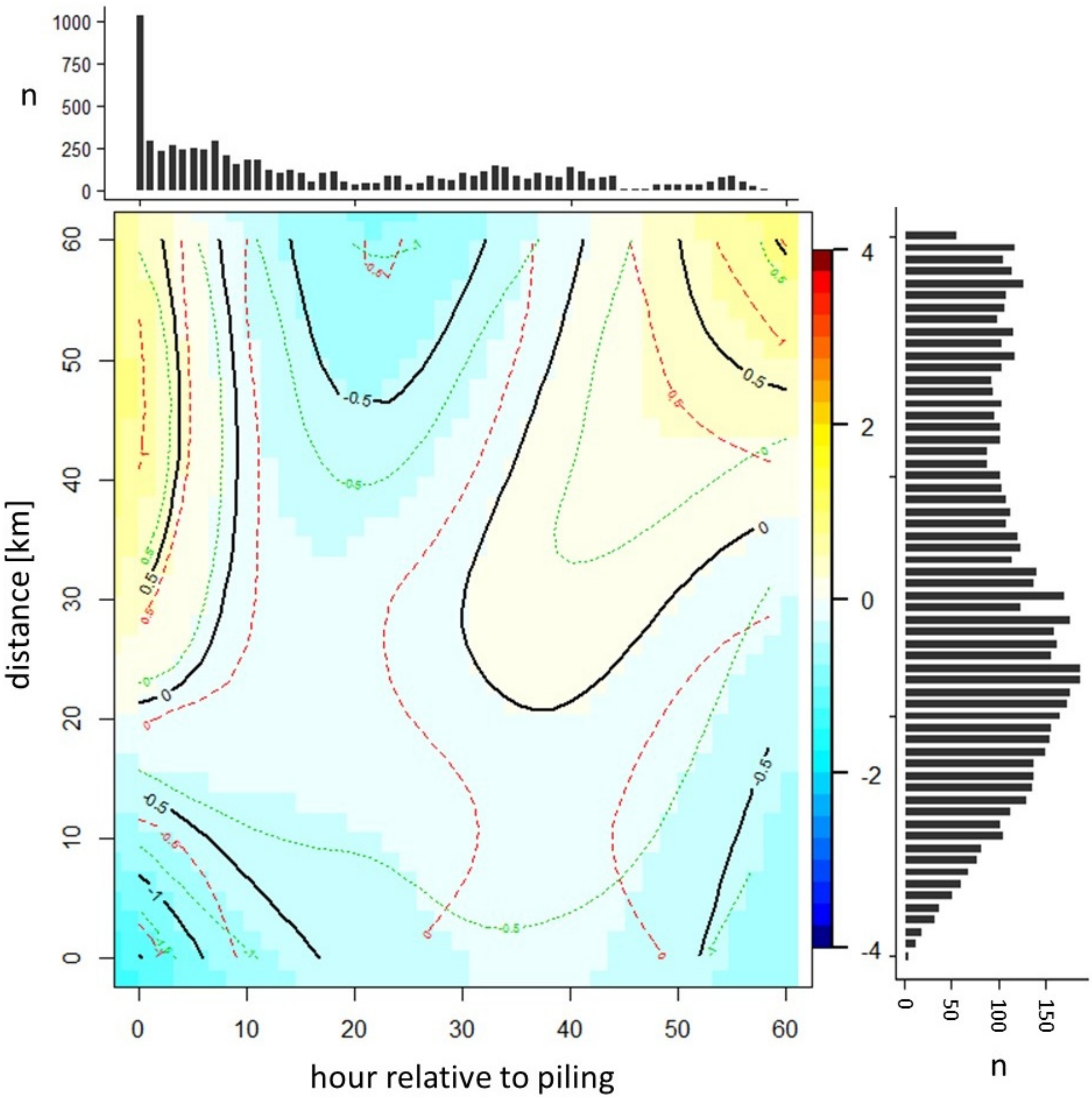


Figure 6-25: Spatio-temporal effect of piling activities on porpoise densities of model ST1 (data coverage given as histogram per variable; isoclines with confidence bands; please compare Figure 6-9 and Figure A-37, Table A-21).

The illustration of GAMs and the deviation from the mean depends mainly on the amount of data considered in the dataset. Unlike POD data, this dataset lacks a reference dataset that could be used to ground the data. Flights did not occur frequently enough, neither directly before piling events nor at a reference location before and after construction. Consequently, the dimension in time and space included in the affected dataset was chosen in a manner large enough to enable the inclusion of a reference in the analysis. Distances chosen for this analysis were considerably larger and the time after piling ended longer than the expected effect ranges (9 to 24 hours and 6 to 24 km) based on literature (TOUGAARD et al. 2006; BRANDT et al. 2011; DÄHNE et al. 2013B; DIEDERICHS et al. 2014). To test the robustness of this assumption, models that used the same variables were parameterised. Five different sets were compared (ST1: 0-60 h & 0-60 km, ST2: 0-48 h & 0-60 km, ST3: 0-24 h & 0-60 km, ST4: 0-60 h & 0-40 km and ST5: 0-60 h & 0-20 km). The reduction in the variable "distance" from the piling site changed the smooth and its confidence bands more drastically than the reduction in the variable "hour relative to piling ". The smooths, in fact, remained almost constant when only time was reduced (Figure A-36). The reduction of “hour relative to piling” resulted in the exclusion of complete flights, while the reduction of “distance” reduced the number of grid cells per flight. Most flights were conducted in summer and spring and a reduction of the temporal scale led primarily to a reduction of flights during those seasons.

Models ST1 through ST6 pooled all aerial surveys conducted during the construction of eight wind farm projects between 2009 and 2013. In the next step, it was investigated whether the resolution of this outcome could be depicted more precisely by grouping the analysis by wind farm project (model ST7). The variables used to model the entire dataset explained 12.6 %, tables and figures summarising the model output are included in the Appendix (Table A-26, chap. A.6.1). Figure A-38 (chap. A.6.2) summarises the tensor products per wind farm. For most wind farm projects, the data were very limited. The graphical illustration for missing data was based on interpolation from other wind farms and were unreliable. In this light, the resulting wind farm specific effect radii should be considered with care (Figure A-37). Particularly, the data from BWII, RG, NSO and MSO were not sufficient. In the case of GTI, despite the fact that the available number of flights was the second highest, the modelling of piling effects was not significant (Table A-26, chap. A.6.1). The same process was applied for each wind farm project. None of those subset models maintained all model variables. They resulted, in fact, in over-parameterised models and a reduction in parsimony, with few informative variables. Some of these models even missed the variables relevant to estimate the effect radii of piling events, resulting ultimately in non-informative models from the perspective of the question posed in this report.

6.3.5 Effects of cumulative piling events on harbour porpoise densities

Addressing effects of simultaneously occurring piling events on porpoise densities was difficult due to a very limited database: aerial surveys were conducted during 121 of the 515 days with piling events. Theoretically, 27 of those days occurred on days with activities at more than one foundation. However, the number of pilings installed on most of those days occurred either after the flight or beyond the 60 km range, defined as the maximal effect zone.

Regarding the distances and the areas covered, there would have been four days with piling events occurring at the same time before the flight, at a distance of less than 60 km from each other (14.03.2012: BARD and BWII; 15.11.2012: BARD and GTI; 21.03.2013 and 21.04.2013: MSO and NSO). These days could be analysed in more detail, but the number of pilings per location was too small and the number of affected days too diverse to run a thorough statistical analysis.

Therefore, the focused was placed on days with consecutive pilings at the same foundation of the same wind farm. The variables investigated were the number of piling events occurring on the same day, the day before and two days before the flights. Preliminary results indicated that including these variables in the modelling process did not result in models that were more informative in light of parsimonious modelling and, consequently, they were excluded from the analysis.

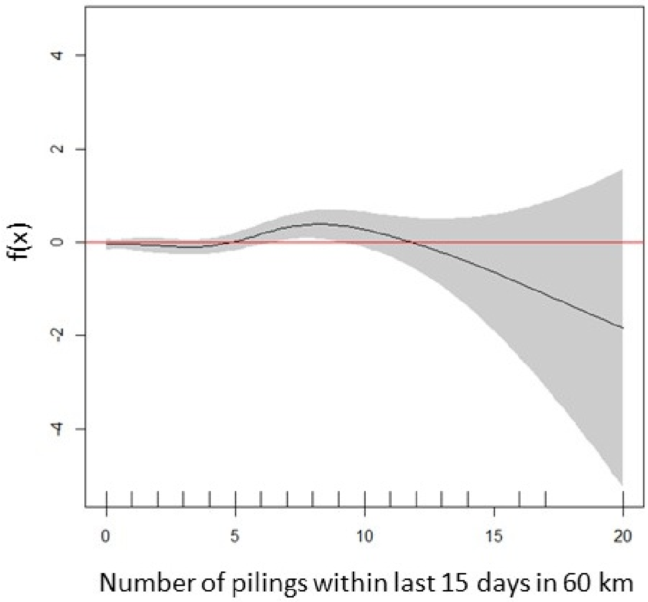


Figure 6.26: GAM-plot illustration effect of cumulative piling events on the distribution of harbour porpoise densities in the subarea "North of Borkum" within the last 15 days at a radius of 60 km. (Red lines represent the mean density estimate. If the shaded confidence intervals surpass the red line, the effect of the value on the x-axis is significantly positive or negative, resulting in higher or lower densities than the mean. Ticks along the x-axis indicate data availability).

Another approach considered in this study was the inclusion of the number of piling events occurring within a certain period (15 or 30 days) at different distances from each grid cell (20, 40 and 60 km). Six different variables were tested on all models. Following the approach of parsimonious modelling, the variable summarising discrete piling events over the 15 days before the flight within a 60 km radius resulted in the most informative variable. The variable was considered only in the model A6 of the subarea "North of Borkum", in all other models this variable was excluded, as the variable was also excluded from the most parsimonious models (delta AIC < 2). This model describing the effect of piling events on the spring and summer distribution of harbour porpoises (Figure 6.26, model A6, subarea "North of Borkum") showed significantly higher density estimates when 7 to 9 piling events occurred within the last 15 days. At higher numbers, there was a trend to lower densities (however not significantly). This numeric variable did not differentiate, if piling was in a distance of 60 km or less, so whether this is an indication for habituation or an influence of animals swimming away from another piling event is not clear.

6.3.6 Comparison of harbour porpoise distribution and activity rates

Figure 6-27 illustrates per season the mean daily acoustic detection rates in comparison with average porpoise densities estimated from aerial surveys. Although, aerial surveys and PAM-stations did not have the same spatial and temporal coverage in this study area, both datasets show a discrete pattern that corresponds to each other. Higher densities in spring and summer both next to the SAC Borkum Reef Ground and Sylt Outer Reef correspond to high values of detections in these areas. Similarly, low detections of porpoises in areas with low densities in the central German Bight and south of Sylt Outer Reef are also in agreement. Low detection rates in Sylt Outer Reef in autumn and winter also confirm low densities from aerial surveys conducted within this area in winter.

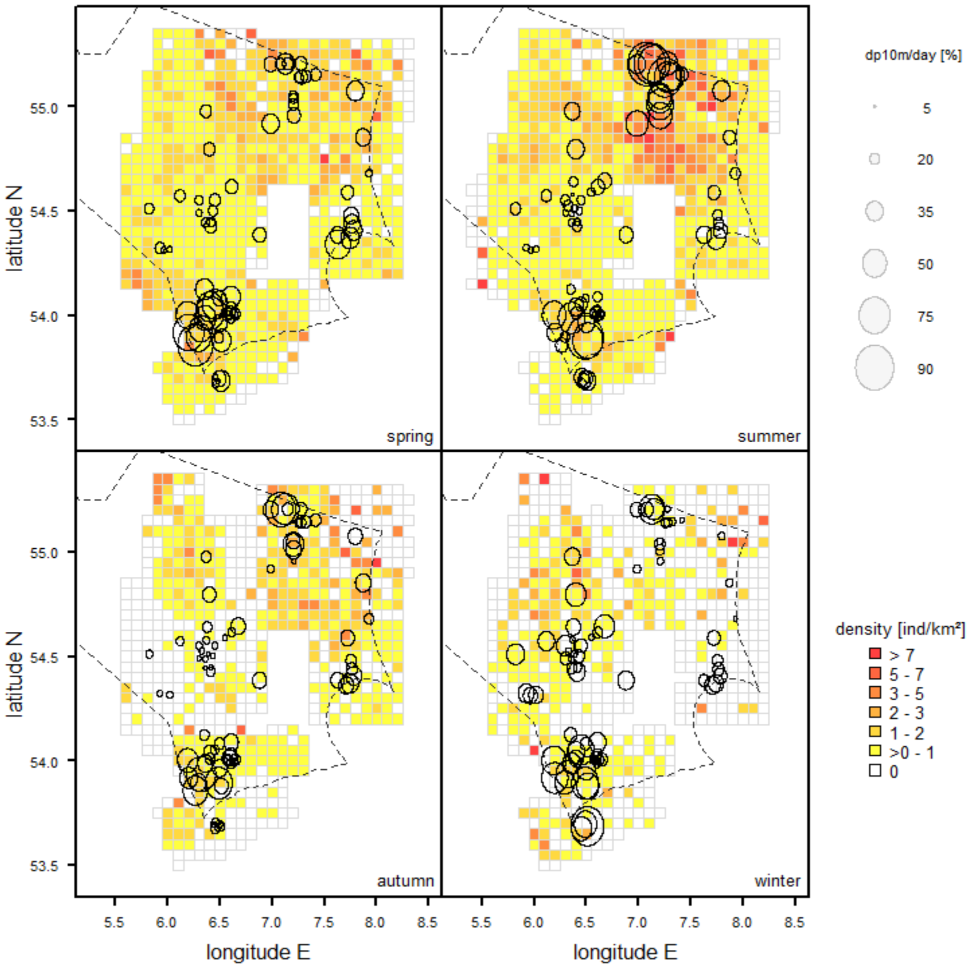


Figure 6-27 Comparison of average seasonal acoustic porpoise detections (dp10m/day [%] 2010-2013) and mean seasonal porpoise densities from aerial surveys (2009 and December 2013, aggregated data per grid cell as weighted means).

6.4 Discussion (Aerial survey data)

This dataset based on aerial surveys was compiled to evaluate specifically large scale and inter-annual effects of piling events on the density and distribution of harbour porpoises. In comparison to singular environmental impact assessments (EIA's) compiled for individual projects, this meta-study benefits from merging data sampled in adjacent study areas that not only originated from the monitoring of the eight construction projects, but also from adjacent baseline investigations for other planned wind farm projects in the German Bight. Consequently, the total area surveyed increased in general and the number of surveys conducted during construction increased, in particular. Even though the construction of wind farms increased between 2009 and 2013, inter-annual density estimates in the study area were not significantly different (model A1-A3). There was a tendency to higher densities in 2013 and to lower densities in 2009 compared to 2010 to 2012. Contrarily to the density of harbour porpoises, their distribution patterns differed between years. Additionally, short-term effects of piling events were analysed by two means. First, comparing density data grouped by the categorical variable „day and distance“. An analysis of the entire dataset showed that densities were lowest 0-24 h after piling ceased in 0-20 km distance from piling and increased with increasing distance from the construction sites (chap. 6.3.3). Second, the effect of piling events was modelled considering the distance from piling and time since piling ceased as a two-dimensional smooth (model ST1, chap. 6.3.4). In the first 12 h after piling ended, densities were reduced in distances of 20 km around wind farms in comparison to distances between 20 and 60 km. Although the resolution of the models was narrowed down, the results of both analyses were not enough explicit and contradict one another, which could stem from the uneven spatio-temporal distribution of the surveys.

6.4.1 Distribution of harbour porpoises in the German Bight

Since the 1990s, marine mammals have been intensively studied in the North Sea. Based on two extensive porpoise surveys within the whole North Sea and adjacent waters in 1994 and 2005, a shift of the porpoise population from the northern to the southern North Sea has become apparent, while the overall densities remained unchanged (HAMMOND & MACLEOD 2006; SCANS 2008; HAMMOND et al. 2013). Several studies confirmed an on-going increase of harbour porpoise densities since SCANS II within the southern North Sea when compared to previous decades (THOMSEN et al. 2006; HAELTERS et al. 2013; PELTIER et al. 2013). Results from this study did not show any annual trend in porpoise densities over the entire study area. Dividing the data into different subareas and seasons only brought a few significant differences between years that occurred without showing a clear trend, although for the area North of Borkum, 2009 was the year with the lowest and 2013 the year with the highest densities. However, since only five years were covered during this study and most subareas had only data from less years, it does not contradict results given by (PESCHKO et al. 2016), who proved a statistically significant increase of porpoise density during the last ten years in a similar area of the southern German Bight. If this regional shift in occurrence might have masked a possible negative effect induced by pile-driving activities of offshore wind farms remains unlikely, as a slight increase is also noticeable within the present study. No negative trend could be shown in different sub areas of the German Bight, which in turn gives no indication for a long-term negative effect of increased pile-driving activities on harbour porpoises.

Furthermore, in this study, geographic gradients that changed in the course of the year were identified (model A1-A3). In spring and summer, densities were highest in general and showed strong geographic gradients from the north-eastern part of the study area (at the German-Danish boarder) to the southwest (German-Dutch boarder). In autumn and winter, porpoises were distributed more evenly than in spring and summer, and densities were generally lower. All year round, high densities of harbour porpoises superior to 1 Ind/km² were apparent in the western part of the southern German Bight next to the SAC Borkum Reef Ground except for summer. However, still densities were generally lower than those found adjacent to the SAC Sylt Outer Reef especially during spring and summer. Great parts of the animals most likely left the study area in autumn and winter except for the central area of the German Bight (Figure 6-14). Due to limited available information, it remains unknown where these animals stay during autumn/winter. However, they might have swum in direction of Dutch and Belgian waters, where in autumn and winter porpoise densities typically increased (HAELTERS et al. 2011; SCHEIDAT et al. 2012).

6.4.2 Effects of piling

Constructions in 2012 and 2013 occurred in six different areas compared to one in 2009 and 2010 and two in 2011. The year with the highest number of piling events was 2013, but densities in that year were not lower, but higher than in the years before (chap. 6.3.1). In fact, the years with the lowest densities were 2009 and 2012. Yearly estimates in this study were within the natural range documented for the period between 2002 and 2014 (VIQUERAT et al. 2015; PESCHKO et al. 2016). Since 2008 onwards, densities in parts of the southern North Sea were higher than in the period between 2002 and 2008 (PESCHKO et al. 2016). This potential regional shift in occurrence might have reduced a negative effect induced by piling events. Three datasets (entire, unaffected and affected from pile driving) were analysed to evaluate how distribution patterns changed between years (chap. 6.3.2). Generally, outcomes indicate that harbour porpoises were less abundant in a distant below 20 km from wind farms after piling activities ceased (model A1-A3).

In a spatially more restricted analysis, differences between three subareas were investigated (model A4 - A7), which cannot be attributed solely to wind farm constructions. In the subarea North of Borkum, on the first day after piling ceased densities were higher compared to the unaffected data. In the area German Bight W and West of Sylt densities, were significantly lower on that day (Figure 6-24). West of Sylt densities were also higher on the second and third days after pile driving stopped. Significant yearly differences in the four models occurred only between 2009 and the other years (apart from 2012) in the southern part of German Bight NW. No yearly difference was detected in the northern part of German Bight NW, in North of Borkum and West of Sylt. Variations of intra-annual distribution patterns were bigger than within inter-annual densities, also indicating a likely stable porpoise population during the study period. An analysis based on data from the StUK-plus project at AV showed that densities in summer 2009 were lower than in the following four years and that densities at Borkum Riffgrund were higher than in surrounding areas both during the years of piling as in the reference years (ROSE. et al. 2014). At a wind farm in Belgian waters, the distribution patterns based on four flights (two before and two after piling activities) showed decreased densities within a 20 km radius, but an inter-annual comparison is missing to confirm different distributional patterns (DEGRAER et al. 2012; HAELTERS et al. 2013).

In a further step, the spatio-temporal short-term effect of piling events on harbour porpoises based on aerial surveys was considered with the two variables, “distance” and “hour relative to piling”, that were modelled as a two-dimensional smooth (model ST1). The analysis demonstrated that during piling, porpoise densities were reduced at distances of up to 19 km and higher than the mean beyond 23 km. Porpoise densities became more evenly distributed with passing time. Adjacent to the construction site (0 – 6 km), densities remained low for up to 27 h, but data availability beyond 20 h after stop pile driving was poor. At distances of 15 km to the wind farm site, it took only about a few hours before densities reached nearly mean values. Further effects at distances beyond 20 km and time over 20 h are probably related to lower sample size and other co-variables not considered within this two-dimensional plot. However, a reduction of the temporal and spatial range of included data showed a consistent result (model ST2 – ST6). These results correspond well with the few impact ranges based on aerial survey data from other projects. HAELTERS et al. (2012) found an impact range of 22 km in Belgian water. Off the Belgian coast, the effect range decreased to 13 km within 24 h (DEGRAER et al. 2012). At AV, an effect range based on one aerial survey is suggested to reach up to 20 km. However, a statistical proof for that statement is not given as it is based on one survey only (DÄHNE et al. 2013A). The results of the StUK-plus study (GILLES et al. 2014A) emphasise that piling events at AV, BARD and BWII primarily had effects on the annual density estimates rather than on seasonal densities. Due to low sample size and few flights connected with pile driving, the study could not come up with effect ranges (GILLES et al. 2014A). The same applies for the AV-StUK3-investigation (ROSE. et al. 2014).

Habituation to and cumulative effects of anthropogenic activities on marine mammals is a growing concern amongst scientists (MACLEAN et al. 2012; ICES 2014; KING et al. 2015), especially in light of the growing number of construction sites in the North Sea (GILLES et al. 2009; MACLEOD et al. 2011; BSH 2015A). Cumulative piling events and their effect on porpoise densities were analysed to estimate habituation and / or effects of consecutive piling events (chap. 6.3.5). The number of piling events occurring during the two days before the flight was tested but excluded from the modelling process, because it did not improve model fit and was therefore of only minor importance. Furthermore, the number of piling events during a specific period (15 and 30 days) prior and within a certain radius (20, 40 and 60 km) to the survey was tested, too. Different combinations of these variables were tested to estimate which variable was the most informative and used thereafter in the models. The model describing the spatial effect of piling on spring and summer data from the area North of Borkum was the only analysis, in which the cumulative variable (number of pilings within a 60 km radius in the previous 15 days) was considered in the final model. The smooth of this variable indicates that density increased with the number of pile driving within the previous 15 days until it reaches a certain number. Due to large confidence intervals because of low amount of data, no clear statement can be given on the effect with a higher number of pile driving within the previous 15 days. This result, that density became slightly higher with increasing number of pile driving within the previous 15 days might be an indication towards habituation. Additional focus needs to be placed on assessing the importance of consecutive and cumulative effects of anthropogenic activities on the abundance of porpoises in any given area.

6.4.3 Methodological issues & Perspectives

Data from aerial surveys are well suited to investigate changes in overall distribution and densities between years and seasons, but are less adequate to address small-scale effects. The main restriction of the dataset was the uneven distribution of flight areas over the years and especially seasonally during autumn and winter in this study. Therefore, part of the analysis had to be restricted to data from spring and summer. To take uneven coverage of data into account, subsets of the entire dataset were taken with the aim of analysing specific areas of concern such as annual differences, while also focusing on changes in distribution due to piling events.

In this study, it was difficult to account for autocorrelation. The data from aerial surveys were aggregated in grid cells (6.0 by 6.0 km) with one or two transects passing through them, what suspended the temporal timestamps of the grid cells. Consequently, time interval between grid cells deviated and the functions did not contain the necessary information to calculate the temporal autocorrelation precisely. Therefore, autocorrelation was considered based on the neighbouring matrix using a Markov random field (Wood 2006; Gilles et al. 2014b). Due to weather limitations and the uneven coverage of the area, the temporal gap between surveys was uneven. It was difficult to consider autocorrelation within one survey, because estimates of two parallel transect lines were condensed per grid cell, thus averaging values and timestamps at one time.

7 APPLYING THE INTERIM PCOD MODEL – AN ASSESSMENT

Combining passive acoustic and conventional flight data provides a broader and more detailed view on the impact of piling activities on harbour porpoises. However, it remains difficult to estimate how strongly the entire regional porpoise population could be affected in the long term, as changes in demographic factors (e.g. lifespan, fertility) are unknown. To assess the impact on abundance and distribution of the harbour porpoise population in the German Bight in relation to construction activities, one option is to develop and apply demography-based modelling approaches.

Concerning impacts on harbour porpoise populations exposed to piling activities only few such models are suitable. The most popular approaches are the interim PCoD model (Population Consequences of Disturbances = PCoD, HARWOOD et al. 2014) and the DEPONS model (Disturbance Effects on the Harbour Porpoise Population in the North Sea, VAN BEEST et al. 2015). At this point in time the DEPONS model is still under development, while the interim PCoD model is ready to use. A thorough documentation is available for the interim PCoD model, including a detailed step by step explanation for the main part of the interim PCoD code (HARWOOD et al. 2014, HARWOOD & KING 2014, SCHICK et al. 2014, SMRU 2014, KING et al. 2015). It has already has been applied for estimating the cumulative effects of piling activities in the North Sea (HEINIS & DE JONG 2015). Therefore, we chose to apply and evaluate the interim PCoD model.

This section firstly presents the interim PCoD model structure and requirements. Secondly, we evaluate and illustrate the sensitivity of modelled impact of pile-driving on porpoise-populations to selected input parameters with theoretical/test data. Thirdly, we run the interim PCoD model with specifications based on results from the present project (offshore wind farms in the German North Sea built between 2009 and 2013). Finally, we evaluate the model as well as its results and give suggestions that might improve it.

7.1 Interim PCoD model - structure and requirements

The interim PCoD model is based on a theory called the Population Consequences of Acoustic Disturbance (PCAD). Figure 7-1 illustrates the conceptual PCAD model, which highlights the cascade from any noise event to individual consequences and finally to population effects. It summarises the interaction of noise and behaviour as a critical first step in the cascade leading to (possible) population effects.

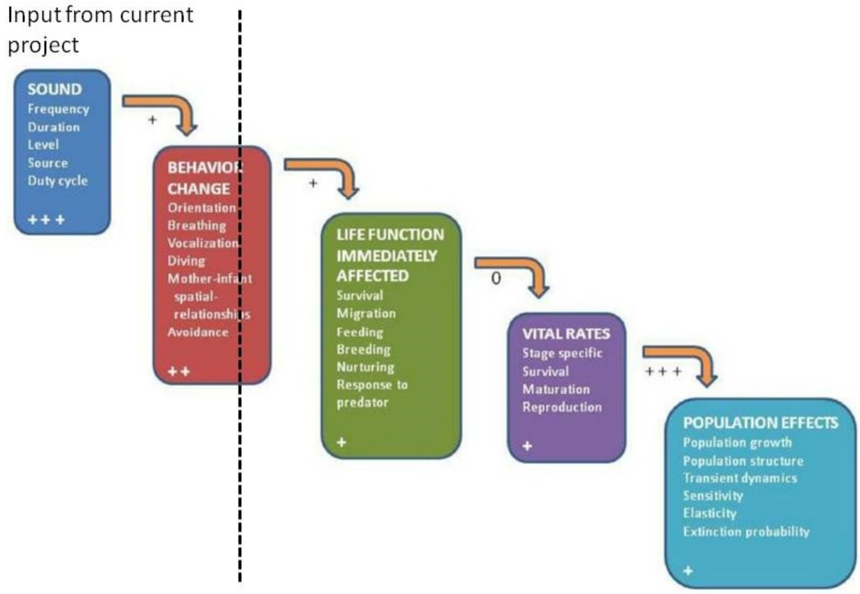


Figure 7-1 *The conceptual Population Consequences of Acoustic Disturbance (PCAD) model (NATIONAL RESEARCH COUNCIL 2005). The relative level of scientific knowledge about the links between boxes is given by the number of "+" signs, "0" indicates no knowledge.*

The interim PCoD model addresses population effects of marine renewable energy developments on chosen marine mammal populations, e.g. grey seals, harbour seals, bottlenose dolphins, minke whale or harbour porpoise, and has been published in a first interim version (HARWOOD et al. 2014; HARWOOD & KING 2014). It is currently the most comprehensive approach to address effects on the population level. However, most factors/estimates are based on expert judgement and hence the authors of interim PCoD model stress the importance of empirical data to reduce the large degree of uncertainty associated with the estimated population changes (HARWOOD et al. 2014, compare Figure 7-1).

The principal theoretical outline of the model is given in Figure 7-2. A detailed description of the interim PCoD code and a tutorial can be found in (SCHICK et al. 2014) and (SMRU 2014). The model compares simulated populations affected by underwater noise with identical undisturbed populations to evaluate population consequences of underwater noise at regular intervals (e.g. years).

The number of individuals affected by piling noise is calculated from two input parameters determined by the user: 1) the number of animals that suffer from PTS (not necessarily considered lethal but reducing survival probabilities) and 2) disturbance that is likely to impair an individual's ability to survive or to reproduce. These two parameters coincide with behavioural responses with a score of at least 5 on the severity scale (SOUTHALL et al. 2007). "extensive changes in speed, direction and/or dive profile; shifts in group distribution; the aggregation or separation of groups of animals; changes in vocal behaviour; active avoidance of the noise source; separation of females and dependent offspring; visible startle response; cessation of reproductive behaviour; and aggressive behaviour" (HARWOOD & KING 2014, p. 10). Thus affected individuals can be inferred indirectly from C-POD data as a decline in detection rates.

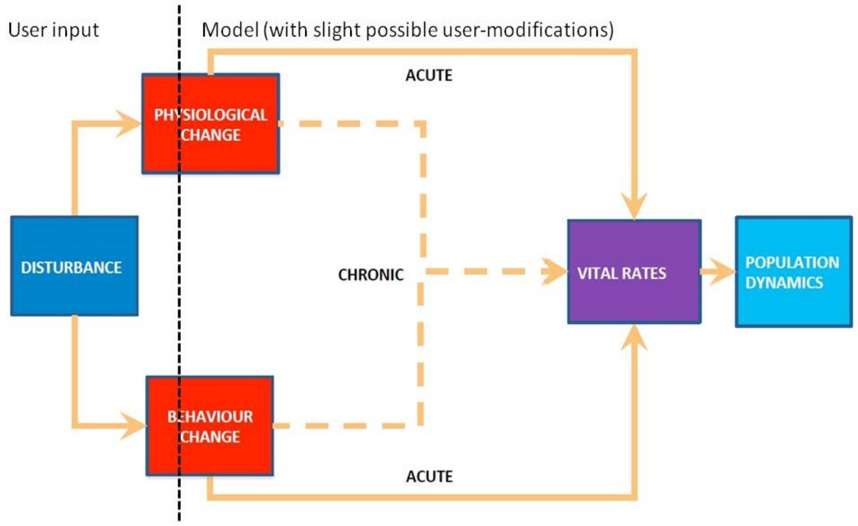


Figure 7-2 The interim Population Consequences of Disturbance (PCoD) model (modified from Harwood et al. 2014). Note the vertical dotted black line that separates the user input part from the model part.

7.1.1 Parameters based on expert judgment as given by Harwood et al. (2014)

Table 7-1 lists the main categories of parameters used in the interim PCoD model that are based on expert judgment. All parameters concerning demographic rates (survival and fertility for pups, immature and mature individuals) of any species included, as well as the consequences of experiencing disturbance or PTS for these demographic rates, are based on a rigorous, up-to-date elicitation of experts and are modelled on a daily time step. Note that the interim PCoD model is a birth-pulse model, which does not model changes over the course of a year, and with all offspring occurring at the start of the year (set to be the 1st of June for the harbour porpoise).

Table 7-1 Parameters of the interim PCoD model based on expert judgment.

Parameter	Source
Demographic rates (survival and fertility for three age classes)	Experts (Harwood & King 2014 or Harwood et al. 2014)**
Consequences	Expert judgement Effect of the impact on demographic parameters
Carrying capacity, density effects	Not incorporated! (No well-founded data available)

**Two rather simple sets, e.g. survival rates are constant for adults, birth pulse at the first day of the year

The effect of disturbance is modelled as being tolerable for several days before survival is affected (Figure 7-3). Animals experiencing less than B days of disturbance (B and C are illustrated in Figure 7-3) are categorised as undisturbed. With more than B days but less than C days of disturbance simulated animals are moderately disturbed and their survival or fertility is reduced by the midpoint between disturbance with no effect and maximum effect. Those with more than C days of disturbance experience the maximum effect. It is assumed that for calves/pups and juveniles the vital rate most likely to be affected by disturbance is survival, while for adults it is fertility. The uncertainty regarding the points B and C is covered in the code by sampling in each iteration from the statistical distribution derived from the results of different experts asked during the expert elicitation (Appendix, Figure A-17).

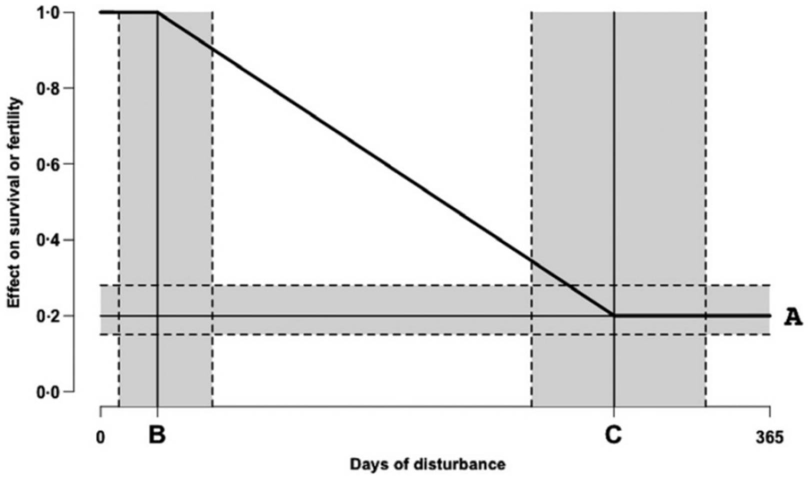


Figure 7-3 A hypothetical relationship between the number of days of disturbance experienced by an individual and its survival was used as the basis for the expert elicitation process. B is the number of days of disturbance an individual can tolerate before its probability of survival is affected, A is the maximum effect of disturbance on this probability, and C is the number of days of disturbance required to cause this maximum effect. Shaded areas indicate likely ranges around the best estimates of A, B and C provided by each expert (from KING et al. (2015) – figure 2, compare with Appendix, Figure A-17).

For the impact of PTS, experts were asked to give estimates of the likely impact on survival (for pups/calves and juveniles) or the probability of giving birth (for mature adults). Using the likely range for each estimate, uncertainty levels were calculated to obtain corresponding probability distributions. In the questionnaire, experts were asked to estimate the potential effect of an exposure to a SEL that is 20dB or more above the threshold for TTS (Temporary Threshold Shift) onset (i.e. the level recommended by FINNERAN & JENKINS (2012) (p.21) for calculating the likely onset of PTS = SEL of 172 dB re 1 µPa2·s). This neglects differences in the amount of resulting hearing loss or the impacted frequencies, which was criticised by the consulted experts.

Uncertainty not in the effect itself, but in the number of affected individuals is furthermore covered by the interim PCoD model. The standard to do this (+/- 50%) can be modified as well.

Considering the long-term effects of piling on the population(s) as modelled in the interim PCoD model, it is important to note that within the model's framework, experiencing PTS influences reproduction and survival permanently, while disturbance affects individuals only temporarily for one PCoD year (HARWOOD et al. 2014, p. 75). The interim PCoD model is a burst population model, e.g. all harbour porpoise calves are born on the 1st of June and thus the PCoD year starts at the 1st of June and ends in May.

7.1.2 Parameters requiring user input

Table 7-2 lists the main categories of parameters in the interim PCoD model, which have to be defined by the user.

Table 7-2 Parameters of the interim PCoD model to be chosen by the user.

Parameter	Source
Population size and proportion of females	(for example based on IAMMWG 2015, ICES 2014)
Vulnerable sub-population	User
Demographic stochasticity	User or Standard values
Operation	User: Days with or without piling, different operations possible
Impact	User: Number of affected Individuals 1) disturbed or 2) experiencing PTS

The model user has to define how large the whole population is and which part of this population is vulnerable to construction, with the possibility of differentiating between piling operations. Animals that are not part of vulnerable subpopulation(s) are unaffected. Each operation is included as days with or without piling. The temporal aspect is important in the context of the possible choice of days of residual disturbance during which PTS is not possible. In order to model the effect of the days with piling, the user has to provide the number of individuals that are disturbed or experience PTS during one day of piling, with the possibility of discerning between two (winter and summer) or four seasons and different operations. The model itself does not evaluate these numbers (Figure 7-2). Additionally, two choices can be made that influence the consequences of the impact: 1) Residual days of disturbance after one day of actual disturbance by piling activities (no PTS possible during this time), e.g. the time needed by an individual to reach a suitable habitat and to behave normally again and 2) if individuals are at risk of experiencing PTS each time they get disturbed or only on the first day they experience disturbance, with the latter reducing the population level effects considerably (HARWOOD et al. 2014 – figure 18).

7.1.3 Methods used

We used the Interim PCoD model version 1.1 and ran it with R version 3.1.1 (R 2014). The usage of the code was inferred from two instructional guides (SCHICK et al. 2014; SMRU 2014). The existing PCoD code was used and parameters as presented in chapter 7.1.2 were adapted to meet our needs. The PCoD code itself was not changed.

The exploration of the effect of selected parameters on the results of the interim PCoD model was automated with R-scripts. Data points shown are each the result of 2000 model iterations. Starting population size for the sensitivity analysis was always set to 10,000. Note that the results are based on PCoD years (June-Mai, with birth pulse on the 1st of June). One residual day of disturbance was included each time and depicted are the results for one PCoD year after piling if not stated otherwise. We investigated the sensitivity of model outputs (e.g. median population decline in comparison to identical unaffected populations, see 7.2) to:

- The number of individuals impacted by piling activities
- The number of individuals experiencing PTS among the individuals impacted by piling activities

- The size of the vulnerable subpopulation (ranging from 1% to 100% of the whole population)
- Piling in different seasons
- Seasonally differing densities
- Cumulative effects of two construction sites
- Continuous piling versus piling interrupted by breaks

Finally (see 7.3), the interim PCoD model was run with specifications based on results from the present project (each model with 2000 iterations). Due to the use of mitigation measures in many cases resulting in noise levels below 172 dB re 1 µPa2·s or harbour porpoise not being that near to the piling all affected individuals were classified as disturbed only. The number of affected individuals was based on the mean of the density for the subareas "German Bight NW", "West of Sylt" and "North of Borkum" as in chapter 6 split into data for spring, summer, autumn and winter. The average effect range was estimated to be 17 km for all projects (see chapter 4 Hourly POD-data, from p 18 on). The actual number of affected individuals per piling day is obtained by multiplying density per km² with the circular area with a radius of 17 km. Due to reduction of detection rates up to 24h after as well as before piling (see chapter 4 Hourly POD-data, page 40) two residual days of disturbance were chosen. This choice is rather conservative as the effect radius before piling is limited to 5-10 km and thus smaller compared to the time during and after piling (see chapter 4 Hourly POD-data, from p 18 on). Therefore, a second model was run where the effect radius was two third 8.5 km for the day before and after piling and one third 17 km for the piling day itself, resulting in a total effect radius of 11.3 km with again 2 residual days of disturbance.

The piling schedule was based on data of the current project summarised in the subareas as defined in the aerial survey data (see chapter 6 on page 103). The 29.02.12 had to be deleted as each year needs to consist of 365 days for the interim PCoD. Piling up to two days before the 31.05 was brought forward for the years 2012 and 2013 to prevent residual days of disturbance to reach into the next PCoD year, as the interim PCoD model produces warnings in that case. In the subarea "West of Sylt" piling occurred in more than one wind farm at the same day in some cases. Possible additive effects, like two or more pile driving activities within the same day and subarea were not especially accounted for in our model (27 days out of 541 total days with piling distributed over 6 (PCoD) years). It is theoretically possible to model additive effects with the interim PCoD, either by adding a proxy operation or by using the values in the piling.csv file as multipliers of the basic values of affected individuals per day, rather than just as an index of the presence (=1)/absence(=0) of piling (personal communication with J. Harwood). However, the potential increase due to additive effects is already included in the estimated effect range (see chapter 4 Hourly POD-data), derived from a comprehensive analysis over all piling events.

In line 62 of PopDyn.R the interim PCoD code generates a single value from a distribution with a mean a little bit larger than 1.0 and a given standard deviation for each iteration. The appropriate values of affected individuals are multiplied by this value to simulate uncertainty in these estimates. In order to incorporate the uncertainty of the density estimates into our models, we used +/- 110 %, corresponding to the uncertainty of our density estimates, instead of the standard value of +/- 50 % for the uncertainty. Usage of this uncertainty feature as well as the use of seasonality was inferred from a consultation with John Harwood.

We restricted population size to 54,227 individuals (AWZ plus 12 sm zone, (GILLES et al. 2010) as choosing a larger population would likely underestimate the effects of piling due to neglecting wind farms under construction in the remainder of the area (HEINIS & DE JONG 2015). We assumed that 50 % of the population were females as suggested by SCHICK et al. (2014).

7.2 Evaluation/Sensitivity of the interim PCoD model using test data (to selected parameters)

Number of disturbed individuals and/or individuals experiencing PTS

The effect of increasing numbers of individuals per day disturbed by piling is depicted in Figure 7-4. The modelled population decline increases a little slower at low and high numbers of affected individuals. This is a consequence of disturbance being modelled as being tolerable for some days with no effect, a maximum effect at a relatively high number of days of disturbance and a medium effect between these two points. The latter is irrespective of the exact number of days of disturbance within these two extremes, e.g. disturbance is modelled with three effect levels: no effect (number of days of disturbance between 0 and B), intermediate effect (number of days of disturbance between B and C) and maximal effect (number of days of disturbance above C, see Figure 7-3, Figure 7-5 and discussion in chapter 7.4.1).

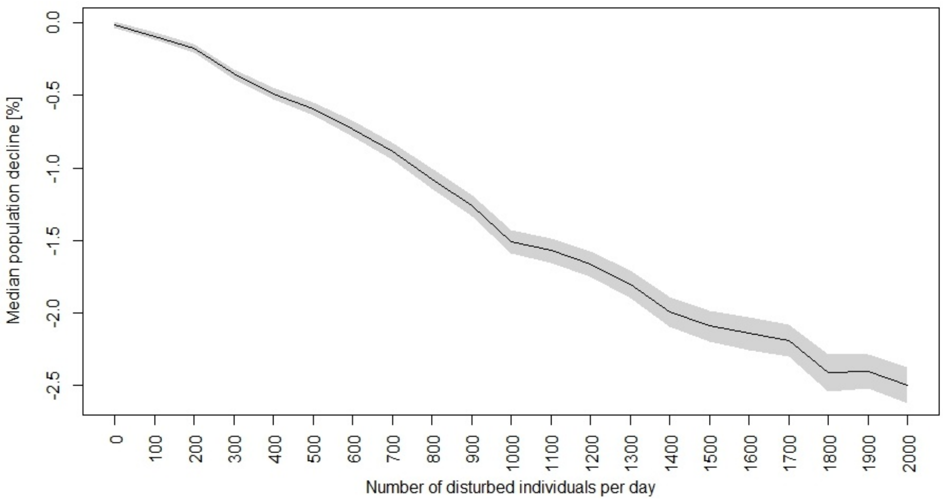


Figure 7-4 Median population decline depending on number of individuals disturbed per piling day. Depicted are the results for 31 days of piling.

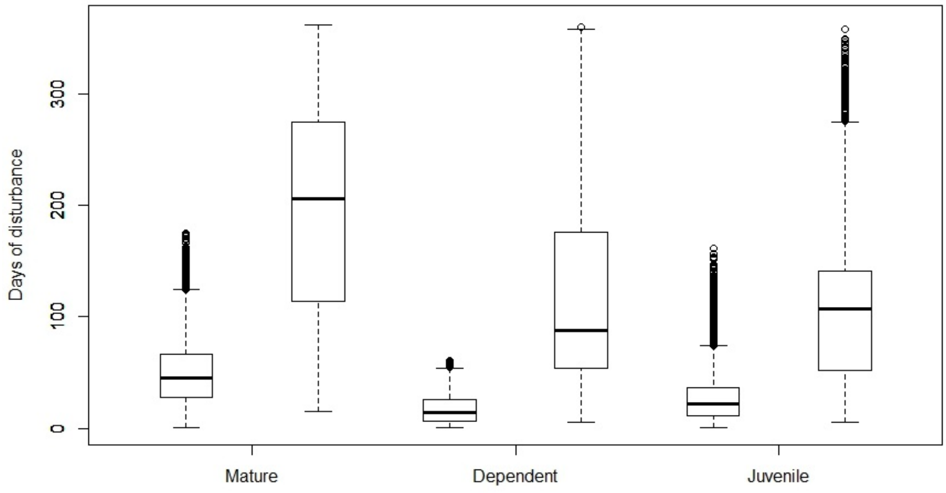


Figure 7-5 *Days of disturbance required before the survival/fertility of mature, dependent or juvenile harbour porpoises is affected by disturbance in the interim PCoD model (each left box plot) and before it is maximally affected (each right box plot).*

Affected individuals could also suffer from PTS instead of only being disturbed. Keeping the total number of affected individuals per day constant, modelled population decline strongly increases if the proportion of animals suffering from PTS rises as shown from model results in Figure 7-6. The confidence intervals arise from the uncertainty incorporated into the interim PCoD model (chapter 7.1.1).

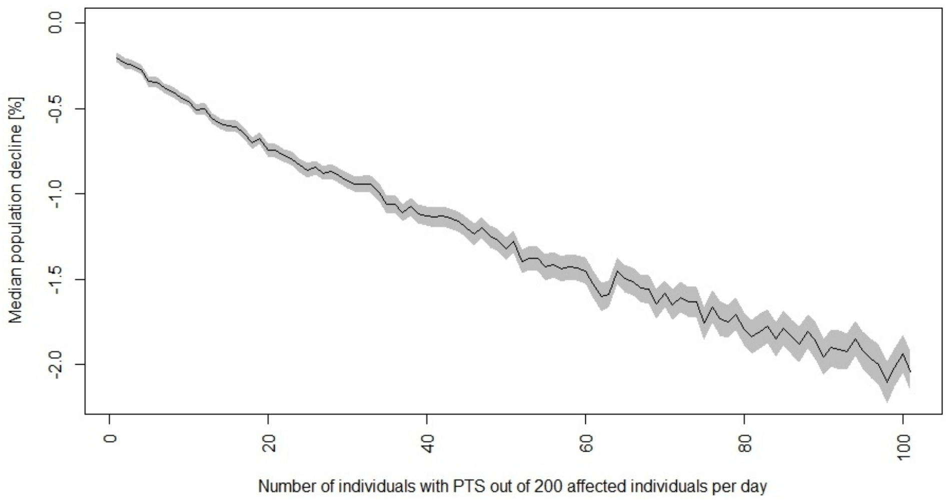


Figure 7-6 *Median population decline depending on number of individuals experiencing PTS out of 200 individuals affected by pile driving per day. Depicted are the results of 31 days of piling.*

Vulnerable Subpopulations

The interim PCoD model offers the possibility to define only a part of the modelled population as affected by certain piling activities, as only few individuals might be in the vicinity of the construction site and thus be affected by noise. These individuals constitute a vulnerable subpopulation, while the rest of the population remains unaffected by piling (of a certain operation).

At a starting population of 10,000 individuals the effect on the population is greater when only 10 % of the population is defined vulnerable at up to about 600 disturbed individuals per day (Figure 7-7). This is because the days of disturbance an individual experiences accumulate faster and consequently the level resulting in an effect is reached faster (compare Figure 7-3) than when all individuals are vulnerable. Above 600 disturbed individuals per day, however, population decrease slows down, until all 1,000 individuals (10 % of 10,000) are affected from each piling day and thus each individual of the vulnerable subpopulation, the remaining 90% of the total population cannot be affected, accumulated the maximum number of days of disturbance possible in the used piling schedule.

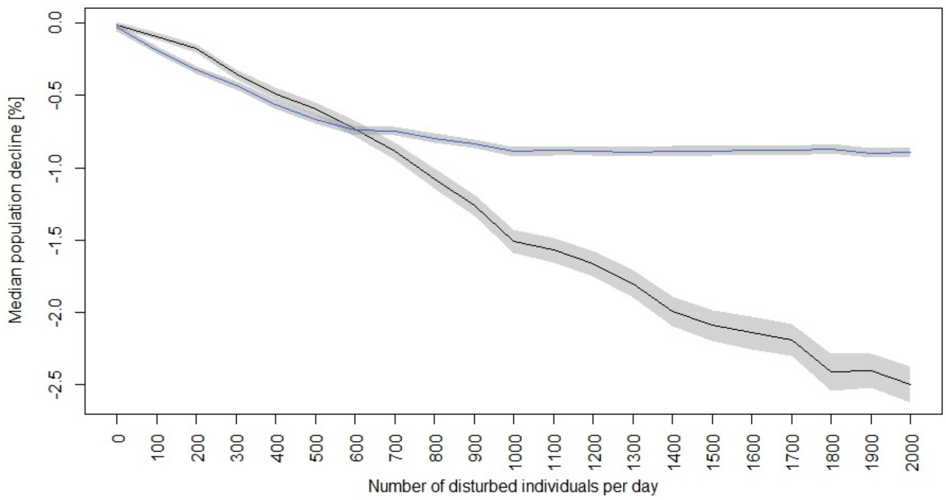


Figure 7-7: Effect of vulnerable subpopulations on median population decline depending on the number of individuals per day experiencing disturbance of 31 days of pile driving. The median in black is with the whole population vulnerable, while the median in blue is with only 10% of the whole population affected by pile driving.

Figure A-18 (Appendix) further illustrates how the effects of defining the vulnerable subpopulation alter the relationship between disturbed individuals per day and population consequences: With a smaller vulnerable subpopulation a maximum effect is reached faster but is limited compared to when the complete population is vulnerable, e.g. the point at which all individuals experience so many days of disturbance that survival and/or fertility is affected (more than Point B in Figure 7-3). These results highlight that the proportion of the population defined as vulnerable has strong implications on how the number of disturbed individuals per day affects the population and this has to be kept in mind when simulating different scenarios.

Using vulnerable subpopulation sizes and number of affected individuals per day did not result in the effects as shown in Figure 7-7 and Figure A-18. Thus, no different effect was found due to the assignment of separate vulnerable subpopulations to two construction sites compared to assigning the combination of both as one vulnerable subpopulation to both construction sites. In other words, there is no cumulative effect in the interim PCoD model if populations are affected by more than one construction site (Appendix, Figure A-19).

Temporal aspects

The interim PCoD model is a burst population model, e.g. all calves are born on the 1st of June. We investigated how the interim PCoD model handles disturbance at different time points. Here we tested whether piling had different effects if conducted in a different season and whether there are model generated differences if piling takes place before or after the burst pulse (1st of June).

Figure 7-8 shows no change in the effect of piling in different seasons. Piling around birth pulse has a lesser effect one year after piling has started (upper panel of Figure 7-8), as the population was only affected by half of the total piling days that occurred until that point. Two years after piling has started, e.g. piling activities are also finished in the case of piling around birth pulse, this difference is diminished to some degree (lower panel of Figure 7-8, in comparison with Figure 7-9). This highlights that the days of disturbance are counted only for the PCoD year in which they occur (June-Mai, with birth pulse on the 1st of June) while its effect determined as shown in Figure 7-3 and Appendix, Figure A-17 (HARWOOD et al. 2014, p. 75). Temporal aspects within the year, e.g. potentially increased sensitivity around birth of calves, are not modelled.

Furthermore, the negative effect of residual days of disturbance, the continuance of disturbance after one day with actual disturbance by piling and thus an increase in the days of disturbance (Figure 7-3), can be seen (Figure 7-8, yellow box plots with 0, blue with 1 and green with 2 residuals days of disturbance). However, although the days of disturbance are doubled or tripled, the increase is small as the disturbance, as stated before, is modelled in three levels only (no, medium or maximum effect). It has to be noted, that the interim PCoD model could not handle residual days of disturbance reaching beyond the 31st of May, resulting in warnings during modelling or interruption of the modelling.

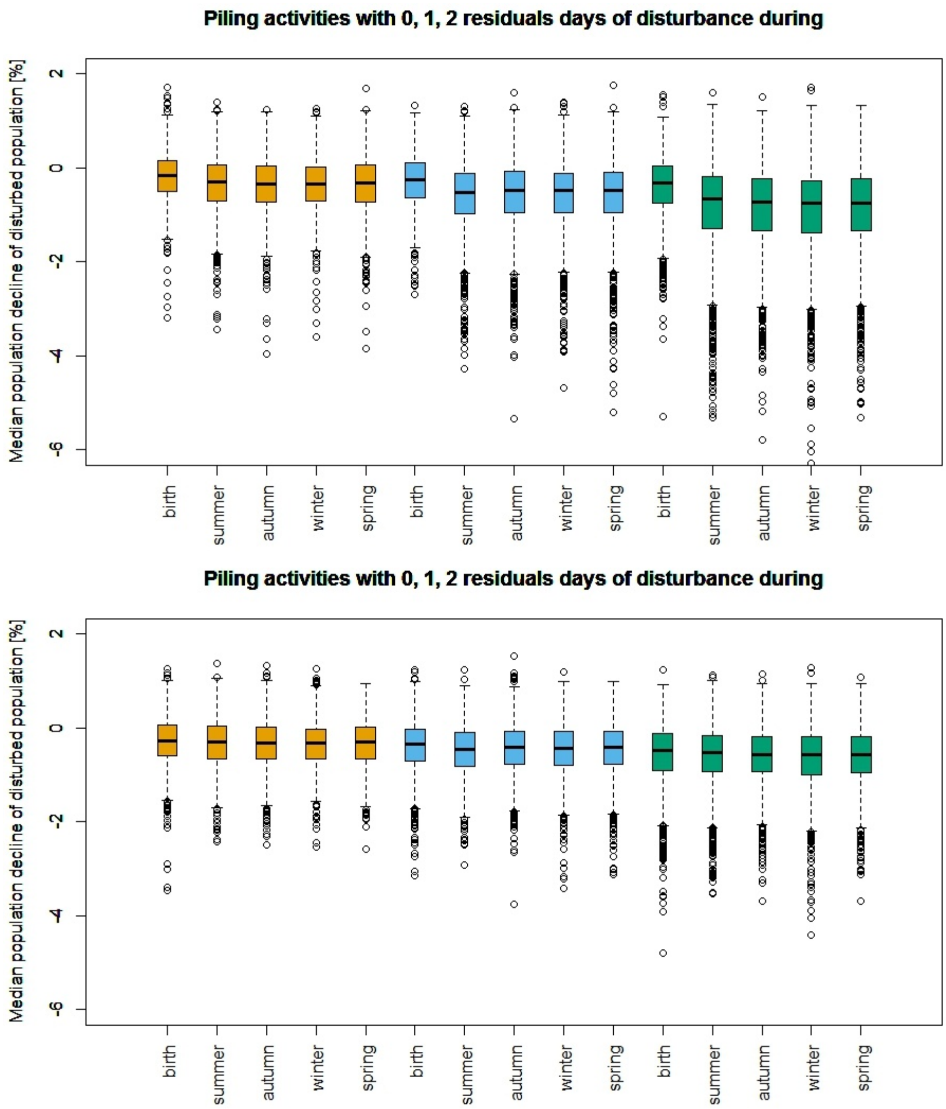


Figure 7-8 *Median population decline depending on the season of piling. Depicted are the results one (upper panel) and two (lower panel) year(s) after the start of 31 days of piling during different seasons and around birth pulse with 0 (yellow), 1 (blue) and 2 (green) residual days of disturbance. 190 individuals were disturbed per piling day and 10 individuals experienced PTS per piling day.*

The importance of the total amount of disturbance but not of temporal aspects is underscored if comparing PCoD model results including seasonal differences. We compared a piling schedule with averaged numbers of affected individuals with a piling schedule with one half in a high density season and the other half in a low density season. There was no difference between them (Appendix, Figure A-20).

The degree of temporal overlap of piling activities between two operations does not influence the total effect on the population modelled by the interim PCoD model either. In other words, simultaneous work at different sites has no cumulative effect (Appendix, Figure A-21). This results from disturbance/PTS being considered as a binary response in the model, that is an individual can be categorised as either having experienced PTS/disturbance or not. Consequently, there is thus no mean of differentiating the degree of PTS or disturbance.

Piling schedules should also specify whether construction activities will have regular breaks or whether they are (rather) continuous (HARWOOD et al. 2014). This would be important in the context of residual days of disturbance after the initial disturbance caused by the piling. During these residual days of disturbance the individuals are assumed not to return to the construction site and thus are not at risk to experience PTS while still negatively affected by the experienced disturbance. However, there was no effect to be seen in our comparison of continuous piling versus regular breaks (Appendix, Figure A-22). This might only be important if the number of affected individuals per day is high compared to the size of the vulnerable population.

Considering the long-term effects of piling on population(s) as modelled by interim PCoD, it is important to discern between disturbed individuals and those experiencing PTS. Only the latter permanently influences reproduction and survival, while disturbance affects individuals only temporarily for one year (HARWOOD et al. 2014, p. 75). In consequence, a negative effect on a population diminishes faster in the years following the piling activity if only disturbance is considered (Figure 7-9). Note that due to the lack of density dependence, the population will not be able to compensate the effects of piling even after a long period of time (Figure 7-9). In consequence the once disturbed population remains smaller than an equal undisturbed population even after 12 years without any disturbance.

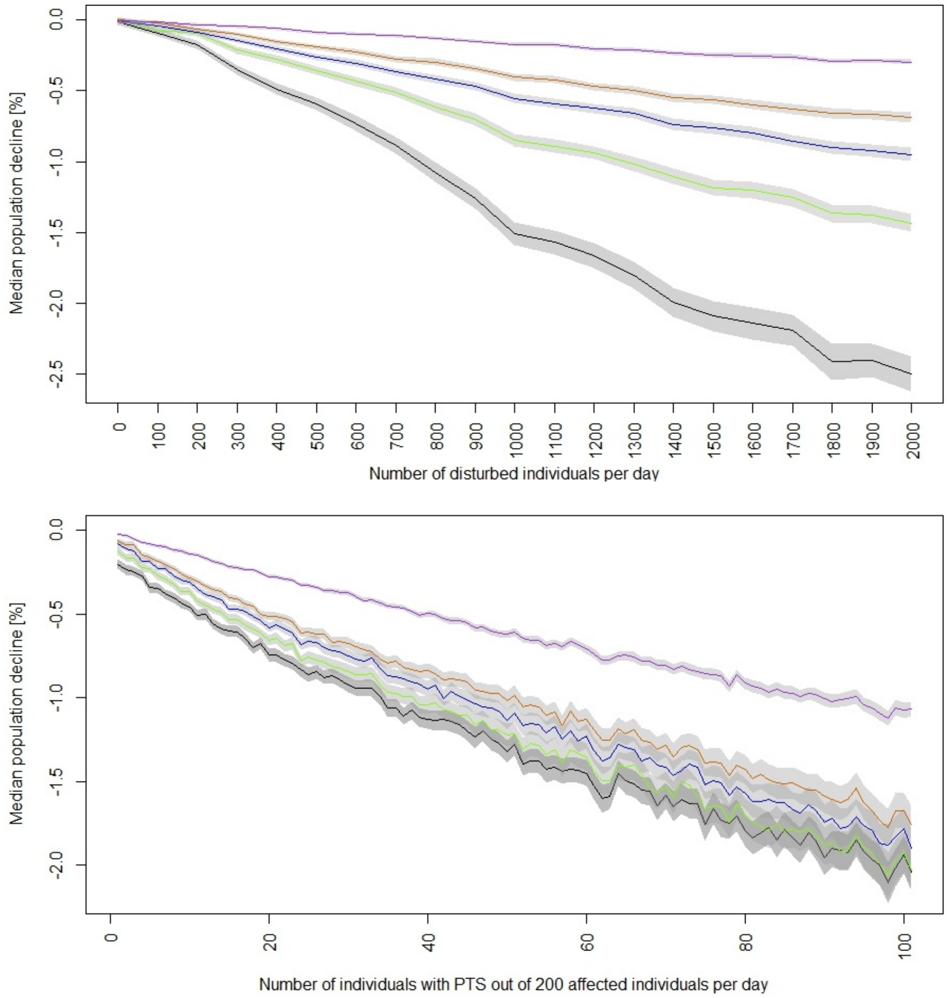


Figure 7-9 Median population decline 1 (in black), 2 (in green), 3 (in blue), 4 (in orange) and 12 (in purple) years after construction with 31 piling days and either causing disturbance only (upper panel) or disturbance and PTS as modelled by the interim PCoD. In the upper panel the number of individuals disturbed per day increase from 0 to 2,000 in 100 steps, in the lower panel the number of individuals experiencing PTS out of 200 totally affected individuals was increased from 0 to 100 in single steps.

7.3 Interim PCoD model with specifications based on results from the present project

Due to the use of mitigation measures in many cases resulting in noise levels below 172 dB re 1 μPa^2 s or harbour porpoise not being that near to the piling, all affected individuals were classified as disturbed only. The more severe effects of PTS are not expected to occur (chapter 7.2). We focus on the effects of disturbance as modelled by the interim PCoD model.

In order to capture the modelled effects three different approaches are applied. First, the impact of piling activities was assessed after modelled piling finished as construction activities will continue beyond this point and the model does not describe any (potential) recovery of an affected population (HEINIS & DE JONG 2015). Secondly, following the suggestion of KING et al. (2015), we looked at whether the model results in a population decline of more than 1% over a 12-year period and thus indicate an unfavourable conservation status. Thirdly, the model results after piling occurred in 2009-2012 was compared to trends in the density estimates of the aerial survey data.

For the three different time points the model predicts a median decline between 0.2 and 0.9 % for the specified scenario of two residual days and 11.3 km effect range (resulting from an assumed 17.3 km effect range during the day of piling and 8.5 km effect range for the two residual days). The predicted decline is 0.9 % during the piling period (2013), 0.6 % directly after the piling period (2015) and 0.2 % twelve years after piling had finished (2027, Table 7-3). For all three time points the model thus predicts a median decline less than 1 % (Figure 7-10). The median decline in percentages as well as the probability of a decline of 1 % (between 0 and 1) were both compared to identical undisturbed populations and results are given in Table 7-3. The probability for a 1 % decline is 27 % in 2013, 17 % in 2015 and 9 % in 2027 and thus always below 30 %. This includes 2013, which could be compared to the aerial survey estimates of the densities showing no clear trend but a high standard deviation of 110 % overall (see chapter 6 Aerial survey data on page 103).

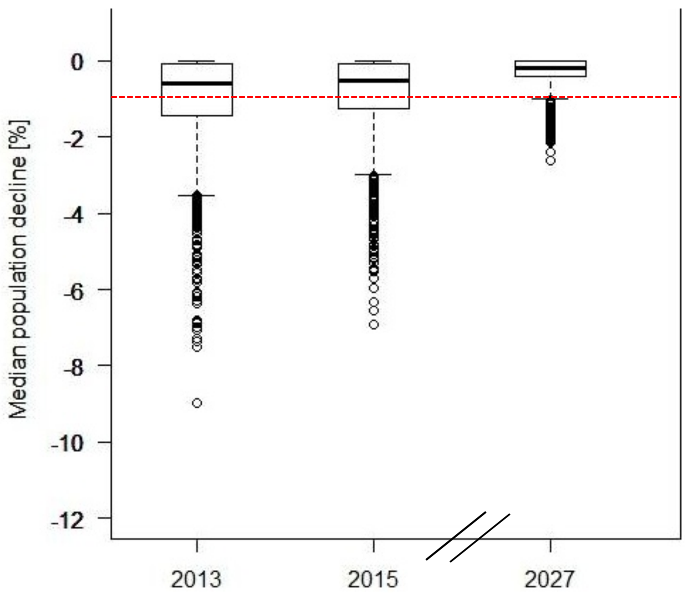


Figure 7-10 Median population decline at different years for all wind farms of the present project. under the specified szenario of a 11.3 km disturbance radius (17 km for the piling day and 8.5 km for the day before and after piling) and two additional residual days.

Table 7-3 Median % decline as well as the probability of a decline of 1 % at different years for all wind farms of the present project assuming a disturbance radius of 11.3 km and two residual days..

Year	Median % decline (11,3 km)	Probability of a decline of 1 % (11,3 km)
2013	-0,9	0,27
2015	-0,6	0,17
2027	-0,2	0,087

7.4 Conclusion and suggestion

7.4.1 Model properties and input parameters

Number of disturbed individuals and/or individuals experiencing PTS

Disturbance is modelled in the interim PCoD model as having no effect until a certain number of days of disturbance are reached and a maximum effect at a certain, often considerably higher number of days of disturbance (Figure 7-5). For all numbers of days of disturbance within these two extremes, the effect is determined as the mid-point between disturbance with no effect and maximum effect, e.g. there are only three levels to describe the large range of days of disturbance considered. This certainly leads to overestimating the effects on the population level if the number of days disturbance experienced by an individual is near the limit of either no effect at all or effects, with the latter modelled as medium effects as soon as the limit is exceeded. It leads to underestimating the effects when getting closer to the maximal effect limit (Figure 7-4). This should be changed in the PCoD code as soon as more background knowledge is available.

Our results furthermore highlight how important PTS is within the scope of the interim PCoD model (Figure 7-6). The strong impact of PTS in the interim PCoD model highlights the importance of a critical evaluation of numbers used for individuals affected by PTS and the need of more research to differentiate the effect PTS based on impacted frequencies and how severe the hearing loss is to critically evaluate the expert judgement incorporated in interim PCoD model.

Vulnerable Subpopulations

The interim PCoD model offers the possibility of defining vulnerable subpopulations. The rationale is that not all individuals in the whole region, e.g. the German North Sea, will spend time in the impacted/affected vicinity of wind farm construction and thus are at risk of being affected. However, the exact (genetically) harbour porpoise population boundaries, like for the North Sea population, are still under investigation (e.g. SVEEGAARD et al. 2011). Furthermore, there is considerable variation within seasons and years in the distribution of harbour porpoise within the North Sea (see chapter 6 Aerial survey data on page 103) and thus migration has to be considered when looking at which individuals are at risk to be impacted by a certain piling activities spanning different seasons/years (VAN BEEST et al. 2015). Finally, the movement of individual harbour porpoise disturbed by pile driving still has to be recorded in order to get knowledge of the individual reactions to pile driving (VAN BEEST et al. 2015). Will individuals actually return to the pile driving site or will other individuals swim into the vacated area, and if they return how often will they return and get disturbed again? This information cannot be obtained from the aerial and POD survey data of the current project as single individuals cannot be recognized. For this it would be necessary to mark and follow single individuals with GPS devices. Information on these behaviours is urgently needed to improve the assessment of the potential long-term effects of pile-driving.

This is important as our sensitivity analysis has shown, as expected, that the definition of the vulnerable subpopulations within the interim PCoD model has strong implications for the predicted effects. However, due to the missing background information needed to define vulnerable subpopulations for projects spanning different seasons and years we refrained from defining vulnerable subpopulations within the model runs we tried.

Temporal aspects

The number of affected individuals can be adjusted for different seasons in the interim PCoD model. However, the effects of disturbance might be stronger during certain times of the year (e.g. during the calving period) and there could also be cumulative effects over several years. Furthermore, it still has to be investigated if an interacting effect of pile driving with more or less simultaneous temporal occurrence at different sites exists, as more of the harbour porpoise habitat is disturbed at the same time. This is not considered within the current interim PCoD model, instead days with disturbance are simply counted, irrespective of the time of the year and with a reset after every PCoD year.

7.4.2 Interim PCoD model with specifications based on results from the present project

The interim PCoD model on the data of the current project and our specifications (chapter 7.1.3) resulted in only a slight risk for possible effects of piling disturbance on the population level. The predicted median population decline was lower than 1 % and the risk for a 1% decline was below 30 %. This result is based on, firstly, a high variability in the estimates of affected individuals leading to a broader effect range. Secondly, we neglected the result that detection rates within the estimated effect range did not drop to zero (see chapter 4 Hourly POD-data, Table 4-7). This may indicate that not all individuals within the radius are affected. The piling noise certainly dilutes towards the edge. Some individuals might show habituation or individual differences in the sensitivity towards the piling noise. However, individuals might experience disturbance but stay in the area for other reasons, e.g. food availability. Until further background knowledge is available, we refrained from considering this information for the estimation of affected individuals. Thirdly, the

PCoD model assumes medium effects on fertility and/or survival as soon as individuals cannot tolerate disturbance anymore (compare previous chapter on page 164). This all might lead to an overestimation effect on the population level.

Considering this, application of the interim PCoD model run using a specified szenario of a 11.3 km effect radius during three days (piling day and two residual days) resulted in a very small risk for a population decline that is unlikely to be equal to or above the 1% generally considered as critical. While a reduction of population size cannot be completely excluded based on the interim PCoD model results, harbour porpoise density was estimated over the project period to be rather constant despite construction activities (see chapter 6 Aerial survey data on page 103). Nevertheless, it must be stressed that such a small predicted decline below 1 %, if present, would be hard to detect given the uncertainty of density estimates by aerial surveys. However, POD-data did also not show any negative trends in detection rates.

Generally, it has to be kept in mind that the model results might be limited since we were not able to define which animals out of the German Bight were disturbed by a certain piling operation(s). If the effects of piling operations were restricted to small subpopulations within the German Bight, days of disturbance would accumulate faster (Figure 7-7) leading to more individuals with negatively affected survival/fertility (Figure 7-3, Figure 7-5). However, it is still unknown how often certain individuals will be affected by piling activities and how severe individuals might be affected at all (Figure A-17). Even if we had all this data, it has to kept in mind that the PCoD model is an interim version. Consequences of disturbance are based on expert judgement that might require modifications as soon as more background data becomes available (KING et al. 2015).

Additionally, harbour porpoises in the German Bight are affected by factors other than wind farm construction. Therefore, further stress courses such as shipping noise, noise from existing wind farms (MORRETTI et al. 2014; NABE-NIELSEN et al. 2014; CHRISTIANSEN & LUSSEAU 2015) and food availability (NABE-NIELSEN et al. 2013; VAN BEEST et al. 2015) should also be taken into account when assessing population effects. Effects like habituation or compensation of population effects by unaffected or less affected individuals need to be considered as well in the future. Together with the uncertainty of future piling events especially the long term prediction of population effects still is a difficult issue.

Finally, it is noted that these results should be interpreted with caution and not yet used to assess population consequences in the “real world”. As outlined in the following sections it is an important spadework on the way to predict population consequences but more background knowledge is needed to improve the PCoD model.

7.4.3 Suggestion to improve the model

It still has to be considered how piling duration / noise characteristics modify the effects of piling. E.g. do harbour porpoise leave a larger area and/or for longer time if piling last longer or based on differences in frequencies and noise level caused by piling?

So far we did not find an effect of piling duration per se (see chapter hourly POD-data from page 18 on). This might be explained by the finding that harbour porpoise already react to certain clues before piling started and thus without a direct interaction with the piling event itself, which still needs to be investigated in detail (see chapter Hourly POD-data from page 18 on). This might indicate that modelling if piling occurred on a day or not is detailed enough to model population consequences.

Sound levels are still important to predict the effect range (see chapter hourly POD-data from page 18 on), while seasonal aspects have to be considered (see chapter daily POD-data from page 66 on). This might be true for more detailed noise characteristics as well, which were not investigated in this study.

However, clearly further systematic studies are warranted to be able to predict effect ranges and duration of planned wind farms in order to assess possible population consequences (see chapter hourly POD-data from page 18 on and chapter daily POD-data from page 66 on).

How to achieve improvement

Currently, modifiable parameters are distributed all over the code in different files, which could be better arranged to make it more user friendly. Considering a better understanding of population effects, it would help to include in the model the assessment of the affected individuals per day.

From a biological point of view, we also feel that modelling disturbance with three levels (one being no effect, two being medium effect and three being maximum effect) is a bit coarse, especially when keeping in mind the large range over which medium effects are modelled, e.g. as soon as the number of tolerable days of disturbance is reached fertility and/or survival are affected at a medium level.

Furthermore, more background knowledge is required for the parameters based on expert judgement, for instance, the effects of pile driving on survival and fertility rates. Some of the data in question will probably remain sparse, as the effects of stress and the different levels of permanent shift in hearing threshold (Permanent Threshold Shift, hereafter referred as PTS) are derived from experiments with few captive individuals (see TEILMANN & TOUGAARD 2006). The individual risk of getting disturbed multiple times, resulting in effects on survival and/or fertility, also depends on how individuals generally migrate and react once disturbed by pile driving. Parameters that likely depend on several factors, for example food availability and migration routes, and thus might be site specific and hard to grasp. Temporal aspects such as habituation, density dependent compensation of disturbance effects and interaction of multiple construction sites are biological processes not clearly taken into account in the current version of the PCoD model.

Finally, further aspects which could affect harbour porpoise negatively, like food availability/habitat suitability in undisturbed sites (NABE-NIELSEN et al. 2013; VAN BEEST et al. 2015) and noise from other sources (MORRETTI et al. 2014, NABE-NIELSEN et al. 2014, CHRISTIANSEN & LUSSEAU 2015), might have to be considered.

7.4.4 Conclusion

The interim PCoD model it is an important spadework on the way to predict population consequences. However, the current version of the model is based mainly on expert judgment as background information is lacking. The following information is needed to improve/ assist in forecasting any effects of pile driving on the population level:

- Empirical data on the effects of disturbance on the individual level is needed as illustrated by the different expert opinions on its implications (Figure A-17) and its current implementation in three levels (no, medium and maximum effect)
- Empirical data on the effects of PTS and the risk thereof
- Empirical data on density dependent compensation of effects and on individual habituation
- Emperical data on the interaction with other confounding factors affecting fitness
- Validation of model results with empirical data

8 GENERAL DISCUSSION & CONCLUSIONS

In order to evaluate different aspects of pile driving impacts on harbour porpoises, the present study combines monitoring data on harbour porpoises, noise measurements during piling and other available information on piling activities collected in connection with offshore wind farm construction in the German Bight between 2009 and 2013. Porpoise monitoring data consisted of aerial survey data gathered between 2009 and 2013 and passive acoustic monitoring data (C-POD data) collected between 2010 and 2013. Aerial survey data cover the construction period of eight wind farms within the German Bight (AV, BARD, BWII, DT, GTI, MSO, NSO and RG), while POD-data cover the construction period of seven wind farms (not AV, which was constructed in 2009). Six of these eight wind farms applied noise mitigation systems for the majority of piling events. At BWII noise mitigation was applied during about ¾ of piling events, BARD and AV were constructed almost entirely without noise mitigation. At all other wind farms more than 90 % of piling events were noise mitigated. Effects of construction on harbour porpoise detections and densities were analysed in order to draw general conclusions about the small-scale and large-scale effects on porpoises and to help in assessing potential population level consequences.

Small-scale and short term effects of piling activities on porpoises were mainly addressed by analysing acoustic porpoise detections on an hourly basis. An effort was also made to address these effects by analysing aerial survey data at an hourly resolution, but data availability turned out to be too patchy in space and time relative to piling activities to yield clear results. Hereinafter, we therefore focus on results from hourly POD-data when discussing findings on small-scale effects, as these yielded more robust and meaningful results on this topic.

Large-scale effects, on the other hand, are impossible to address with POD-data at an hourly resolution, because small-scale resolution is not meaningful when looking at trends over several years. More suitable for this purpose were aerial survey data and POD-data at a daily resolution: Aerial survey data have a greater spatial coverage such that potential redistribution of porpoises in response to piling can be captured more easily. POD-data at a daily resolution have the advantage of covering small areas over an almost continuous time period.

Clear negative short-term effects of piling activities were found on acoustic porpoise detections from POD-data as well as on densities calculated from aerial survey data. These effects decreased in magnitude and duration with decreasing noise level and increasing distance. Considerable variation was found among wind farms that could not be related to construction characteristics (e.g. noise level and noise mitigation) alone. Below we discuss the present findings in light of the various factors that may explain the large differences in short term effects between wind farm projects. Several more wind farms are planned in the German EEZ and beyond. Therefore, it is important to try and address long term effects of offshore construction. We discuss yearly trends that were found in porpoise detections and densities and their relation to construction activities in the German Bight over the study period. Moreover, we critically review the outcomes of the PCoD (Population Consequences of Disturbance) model, which was applied using specifications based on results from the present study.

8.1 Short-term effects of offshore piling

8.1.1 Noise levels during construction

The most important aspect of disturbance effects on porpoises from offshore construction is noise emission. Therefore, we collected all available noise measurements of the studied wind farms and extrapolated noise levels where needed.

Earlier wind farm projects, where construction effects on porpoises were studied, were mainly constructed without applying any noise mitigation measures (e.g. Horns Rev I, Horns Rev II, Nystedt, AV). During the present study, six out of eight wind farms were constructed while applying noise mitigation during the majority of piling events with the aim to eliminate the risk of hearing impairments in porpoises and to greatly reduce disturbance effects. According to BIOCONSULT SH et al. (2014) and NEHLS et al. (2016) such mitigation systems should lead to a noise reduction of between 9 and 13 dB SEL_{50} and therefore reduce the disturbance radius by about 10 km, leading to a reduction in the disturbed area by up to 90 %. Exploring the available noise data, it became obvious that there was strong variance in measured as well as extrapolated noise levels especially at further distances. Within projects, noise levels during noise mitigated piling varied by more than 15 dB even at 750 m distance, and loudest noise levels during mitigated piling were almost as loud as those during unmitigated piling. This indicates that the efficiency of noise mitigation was very variable and probably dependent on weather related phenomena as well as technical difficulties. In fact, throughout the construction period, several different configurations of noise mitigation systems were tested, developed and improved, leading to strong variance in the reduction of noise levels. Furthermore, the noise level of piling depends on several other factors, such as the diameter of the pile, sediment, water depth etc. Noise levels during mitigated piling varied so considerably, that it blurred project-specific differences. Median noise levels during noise mitigated piling at distances beyond 5 km were highest for GTI and lowest for RG and BWII, even though GTI used tripods. This may be related to the greater water depth at GTI (around 40 m) than at all other projects apart from BARD. BARD was also constructed at water depths of about 40 km, while water depth at all other projects ranged between 20 and 35 m. In deeper water sound will travel further due to less reflection and thus less propagation loss (PORTER & SCHMIDT 2000). At greater distances sound propagation becomes even more complex and variable due to a greater effect of several environmental variables affecting noise propagation.

The high variability in noise levels shows that the noise mitigation systems applied during these six wind farm projects were still under development and did not always work equally well. It also demonstrates that extrapolating noise levels comes with high uncertainties about the accuracy of these values and it raises the question whether it may be better to use distance from the noise source as a proxy for noise when assessing disturbance effects rather than extrapolated noise levels. This is especially the case for the projects BARD and RG, where extrapolated noise levels where based on only two and eight measurements, respectively. On the other hand, if values for noise levels exist, these might be more informative than simply using distance. Therefore, we tried both approaches.

8.1.2 Effects of noise levels

For environmental impact assessments based on noise prognosis for specific projects it is of special interest to establish the relationship between noise levels from offshore pile driving and porpoise responses. Therefore, one aim of the present study was to establish at what noise levels changes were found in hourly porpoise detection rates.

GAM models revealed clear declines in acoustic porpoise detections at noise levels exceeding 143 dB SEL_{05} (Sound Exposure Level). This estimate is based on the noise level where porpoise detections during piling reached the overall average of all data. Estimates may vary depending on the statistical definition of effect range and on the statistical approach applied. We also used non-parametric analyses, which allow for relating detection rates during piling directly to a baseline period before piling for a specific noise level class, but cannot control for as many potentially confounding effects as a GAM. Also estimates of effect ranges will be based on noise classes rather than yielding a specific estimate. Furthermore, whether or not effects are significant strongly depends on the height and variability of baseline detections and the amount of available data (with increasing availability and decreasing variability of baseline data chances increase to detect small changes). Therefore, we chose to define the effect range based on non-parametric statistics by identifying down to what noise class significant declines by at least 20 % occur, instead of just focusing on statistical significance. This resulted in an effect range down to 145-150 dB SEL_{05} (using noise level classes of 5 or 10 dB, depending on data availability within these). There was a clear gradient in the amount by which porpoise detections declined at the different noise classes. While the decline was 93 % at noise levels above 170 dB, this decline became gradually smaller within the next quieter noise level class until it was only 25 % at 145-150 dB and below 20 % or not significant at noise classes below that. This clearly shows that not all porpoises respond equally to the same noise level or that the type of response changes with noise level. The lowest noise level class with a decline by over 50 % was found to be at 150-160 dB.

This estimate of 143 dB SEL_{05} is also close to the estimate of 146 -148 dB SEL_{05} (transformed from 144-146 dB SEL_{50}) given by DIEDERICHS et al. (2014) for the construction of BWII. KASTELEIN et al. (2013) studied behavioural reactions of harbour porpoises to simulated piling noise in captivity. They found the mean onset of a reaction in terms of porpoises jumping out of the water at 136 dB, however, only at 154 dB was the number of jumps significantly different from a baseline. It is difficult to relate findings from captivity to passive acoustic monitoring studies in the field. This is because animals in captivity are constrained in their avoidance behaviour and the motivation for avoidance may differ substantially. Furthermore, noise characteristics in a tank will differ substantially from a natural environment. As such 136 dB may be seen as a context specific value for when porpoises may be disturbed.

The above broadband noise estimate for the onset of porpoise behavioural reactions has to be interpreted only within the context of piling noise from offshore wind farm construction activities. This noise has the greatest energy at relatively low frequencies below 1 kHz. Noise from other activities with different frequency spectra will naturally yield different estimates for the onset of behavioural reactions. Seal scarer noise for example is emitted at higher frequencies of about 15 kHz where porpoise hearing is more sensitive (KASTELEIN et al. 2002). Accordingly, avoidance behaviour by harbour porpoises to seal scarer noise was found at noise levels of about 119 dB SEL (BRANDT et al. 2013B). TOUGAARD et al. (2015) reviewed the available literature to assess frequency-specific responses of harbour porpoises to noise. Their results suggest that behavioural reactions of porpoises are usually found at about 40-50 dB above the frequency-specific hearing threshold.

Project-specific analyses of the effects of noise and addressing the role of noise mitigation were limited by few noise measurements during some projects especially at distances beyond 2 km. In order to look at project-specific differences in porpoise reactions and the effects of noise mitigation, we therefore focused on the effects of distance from the noise source (which correlates with noise levels) during more detailed analyses of piling effects on porpoises.

8.1.3 Effect ranges

In order to further understand small-scale disturbance effects of piling on porpoises, we studied the distances at which porpoise detections and densities changed during and after piling. The joint analyses of all passive acoustic data at an hourly resolution with respect to the construction of seven offshore wind farms, regardless of whether or not noise mitigation was applied, revealed that acoustic porpoise detections clearly declined in up to 17 km during piling. Non-parametric analyses resulted in an effect range of 10-15 km for the complete dataset. As found in the results from analysing declines at different noise classes, a clear gradient in the decline during piling was also found depending on the distance class. Detection rates declined by about 68 % at 0-5 km distance, but only by about 26 % at 10-15 km distance. At further distance classes, the decline during piling was always below 20 % and declines by more than 50 % were only found at the nearest distance class of 0-5 km. Analyses of daily POD-data and aerial survey data also found negative effects of piling mainly at distances up to about 20 km and thus are largely in line with results from analyses of hourly POD-data.

Pooling all project data for analyses of effect ranges in terms of distance is meaningful when aiming to assess the overall effect of all wind farm construction activities within the four study years. Pooling has its limitations, however, when providing specific estimates on effect radii, because the seven wind farm projects vary extensively with regards to several parameters that all play a role in determining at what distances porpoises will react to piling noise. The main determining factor among these one would expect to be noise level. However, as discussed above, differences between projects were largely blurred by differences in noise mitigation efficiency due to malfunction and different noise mitigation systems within each wind farm. Wind farm projects also differ with respect to habitat characteristics and porpoise density, and these differences surely play a role in determining porpoise response to piling. Therefore, project-specific models were calculated for assessing effect radii for each of the seven wind farm projects.

When comparing wind farm projects, estimates for effect ranges differed extensively. Project-specific GAM models could not be calculated for MSO and RG due to generally low data availability and several gaps with respect to distance from piling. Effect ranges based on GAM-model outputs were difficult to establish for BARD due to high detection variability at various distances. Therefore, a possible range is given based on visual inspection of the model output and where detection rates during piling reached the overall average (20-34 km). For the other projects, effect ranges (based on where the overall detection average was reached during piling) ranged from only 6 km at DT to 16 km at BWII, with intermediate ranges at GTI and NSO (9 km). We also applied non-parametric statistics to address project-specific differences. These yielded differing results for two projects: Significant declines by over 20 % at BARD were found at distances of only up to 5-10 km but at GTI they were found at up to 20-30 km. For the other projects, estimates are close to results from GAM analyses: 0-5 km for DT, 10-15 km for NSO and BWII. Such differences may be related to difficulties with data availability in the case of BARD, where substantial differences in detection rates were found between POD-positions and into different directions from the wind farm. It is not possible at the moment to say which estimate is more reliable. Why effect ranges at GTI differed so substantially between the two methods remains unclear. Reasons for the differences between projects are diverse. One would expect them to mainly originate from differences in noise levels between projects. However, while BARD and GTI were louder, the other projects were quite similar with respect to the estimated noise levels, such that this could not explain the differences in effect ranges between them. Most noise levels at greater distances are based on extrapolation. This means that there may have been differences in noise propagation between wind farm projects that we could not capture with the available data and that may explain different effect radii between wind farms.

Another factor that so far may be underestimated when assessing the effect of noise on porpoises is the effect that different weather conditions may have on noise propagation. We found indications for disturbance effects before and during piling to reach further at lower wind speeds. This may be linked to natural noise mitigation at higher wind speed due to more air bubbles in the water and decreased reflection from the sea surface. Such assumptions are supported by earlier findings that found sea state to have clear effects on noise propagation (e.g. THIELE & SCHELLSTEDE 1980; JONES et al. 2009) especially mitigating frequencies above 1 kHz (where porpoise hearing is more sensitive) to a relatively larger degree (JONES et al. 2009). THIELE & SCHELLSTEDE (1980) point out, however, that the strongest effects of sea state were found during winter at wind speed above 15 m/s, when water layering was substantially less than in summer. These wind speeds are above the conditions when piling usually occurs and we found differing effects on porpoises at wind speeds below this. Similarly, a recent report by HEINIS & DE JONG (2015) demonstrated that substantial noise mitigation occurred at the sea surface due to a mitigating effect of air bubbles only in the upper water layer. Differences in wind speed between construction periods of the seven wind farm projects may have led to differences in noise propagation. This could be another source of variation, we could probably not sufficiently control for. This also applies to the extrapolated noise levels used during the present study as extrapolation did not take weather into account.

From the present analyses there are also indications that noise level is not the only factor determining porpoise reactions to piling noise. Smaller effects in terms of spatial range and magnitude found at DT for example, cannot be explained by noise level alone, as piling here was not clearly quieter than at BWII or NSO. DT lies in a relatively high density area next to the Natura 2000 area Sylt Outer Reef. BWII is also located in a high density area, according to results from aerial survey data and daily POD data, even though PODs at far distances extend into lower density areas. Even though porpoise densities at DT are not generally higher than at BWII, the surrounding high density area at DT is larger than at BWII. This could mean that there was a larger number of animals in the vicinity that were not affected by piling. This in turn could lead to the construction area at DT being revisited by porpoises at a quicker rate than at other wind farms. However, effect duration at the close vicinity (< 2 km) was not necessarily shorter at DT than at other wind farms. When looking at distance ranges of 0-5 and 5-10 km, however, detection rates 25-48 h after piling were increased compared to 25-48 h before piling. As mentioned abote at DT effects during piling effects did not reach as far as in the other wind farms and in the vicinity the decline was of considerably smaller magnitude. This indicates that fewer animals left the construction area than at other wind farms. Smaller effects in this area may be related to an especially high quality feeding habitat. This may lead to a lower motivation of porpoises to leave the noise-intense areas around construction. This may be related to different prey organisms occurring between the two areas and maybe differences in their vulnerability to piling. Increased porpoise detection rates after piling at DT lead to speculations about increased attractiveness of an area as foraging habitat after piling within a 10 km radius. Before more detailed data are available on food availability and porpoise behaviour within the different areas, this remains speculative.

In some cases, differences between projects may partly be a result of patchy data availability and of naturally occurring gradients in porpoise detection rates, which complicate the estimation of detailed maximum effect ranges and limits comparability between projects. It seems relatively difficult to establish maximum effect ranges of piling activities based on proving statistical significance. This is because there are several difficulties with natural gradients in detection rates, data availability at the various distances and probably also a much greater variance in noise levels at further distances (which are affected to a much greater degree by environmental characteristics than those in close vicinity). Therefore, more emphasis could be put on the overall rate with which detections decline and stating the distance at which a decline is below a certain limit. In our opinion 20 % decline is a relatively good indicator for effects that are not simply a result of natural fluctuations and usually declines by more than 20 % were also statistically significant. Within the present study we found no declines in porpoise detections by more than 20 % at any 10 km distance classes beyond 30 km, showing that, even if significant effects were present during some projects (BARD, GTI, BWII), these were only of minor magnitude.

From aerial survey data there were indications for porpoise densities to be increased shortly after piling at distances above 20 km. This may indicate that animals leaving the vicinity of the construction site accumulate at greater distances. This effect could not be confirmed by POD-data, but this may be due to the smaller spatial coverage of these. Since avoidance effects decrease with distance while at the same time the circular area at further distances to the piling location increases exponentially, such effects are probably rather difficult to prove.

8.1.4 Effects of noise mitigation

Of the seven wind farms where porpoise reactions were studied with passive acoustic monitoring, only BARD was constructed entirely without noise mitigation. BWII used no noise mitigation during piling of 12 foundations. During all other projects either none (RG) or only 1-2 foundations were piled without noise mitigation. The noise mitigation systems that were applied consisted of bubble curtains or the Noise Mitigation Screen (NMS). In some cases several bubble curtains (circular ones and linear ones) were combined. However, as demonstrated above, noise mitigation does not seem to have worked equally well at all times, and it can be assumed that it will also differ depending on direction, water depth, sediment etc. In some cases noise levels were almost as loud as when no noise mitigation was applied. Given this, it is not surprising that estimates for effect ranges arising from the present study are not that different from previous studies on the effects of unmitigated piling. All these studies found negative effects of piling on porpoise detections to reach up to at least 15-20 km (Carstensen et al. 2006; Tougaard et al. 2009; Diederichs et al. 2010; BRANDT et al. 2011; DÄHNE et al. 2013B). Some of the previous projects (e.g. Horns Rev II) were constructed at shallower water depth than the wind farms subject to this study, which will probably have increased noise transmission loss (BRANDT et al. 2011). In one case (TOUGAARD et al. 2009), maximum effect range could not be established due to no PODs deployed beyond the 20 km were effects were still found, so theoretically effects could have reached further.

Comparing effect differences between piling events with and without noise mitigation based on the present dataset was unfortunately quite limited, not only because the efficiency of noise mitigation varied. Little data existed from piling events without noise mitigation and these originated mainly from only two projects (BARD and BWII). BARD posed substantial difficulties for testing effect ranges because of patchy data availability with respect to distance from piling, a relatively high natural variation in porpoise detections and generally low detection rates at the crucial distances, resulting in uncertain estimates from GAM models and non-parametric tests. Comparing effect ranges between different wind farms for addressing effects of noise mitigation also has its limitations, because of other parameters also playing a role (e.g. potentially attractive habitat at DT). Nevertheless, we found effect ranges to be slightly reduced from 17 km including all data to 14 km when only noise mitigated piling events were considered, demonstrating that noise mitigation reduced disturbance to at least some extent, only not as much as may be expected if noise mitigation always worked reliably. This is in line with results from DIEDERICHS et al. (2014) on the effects of noise mitigation at BWII. They showed that theoretically, disturbance effects should only reach up to 5 km from the noise source when sound mitigation worked best. When comparing effect ranges based on porpoise monitoring data from this wind farm between piling events with and without noise mitigation, there were no clear differences. However, given that noise levels with sound mitigation were so variable and that sound mitigation was not always sufficiently effective, clear differences in porpoise reactions may not be expected. Overall, we expect that high variability in noise mitigation success, relatively low data availability for piling events without noise mitigation, high natural fluctuations in detection rates and the various factors affecting noise propagation to have confounded the single effect of noise mitigation.

Finally, during all piling activities a seal scarer was deployed prior to piling as a deterrence measure. It may be possible that the seal scarer has a further reaching effect than piling under sufficiently working noise mitigation (when 160 dB at 750 m are not exceeded). Observations of porpoise behaviour in conjunction with noise measurements revealed that porpoises started to avoid seal scarer noise at noise levels of about 119 dB (BRANDT et al. 2013A). It is not clear at what distances this noise level was reached at the wind farm sites in question. A study looking at porpoise detection rates in response to seal scarer deployment in the North Sea found that these were decreased at 7.5 km, which was the maximum distance studied (BRANDT et al. 2013B). Whether the effect reached further, remains unclear. However, noise modelling based on noise measurements of seal scarer noise predicted levels to be below 119 dB at 10 km distance, so it is unlikely that piling effects found in over 10 km distance are due to animals avoiding the seal scarer.

Nevertheless, the present study confirmed that porpoise detections showed marked decreases at noise levels above 143 dB SEL_{05}. Under the condition that noise mitigation is further improved so that 160 dB are not exceeded at a 750 m radius, as intended by the regulatory framework, the application of noise mitigation would indeed lead to a substantial reduction of the significantly affected area as pointed out by e.g. DIEDERICHS et al. (2014) and NEHLS et al. (2016).

8.1.5 Effect duration

Effect duration in close vicinity of the construction site (up to about 2 km distance) analysed by means of hourly data lasted up to between 20-31 h (overall average was reached at 20 h, the first local maximum at 31 h). Project-specific models yielded different estimates, ranging from 9 to 28 h after piling when looking at the time when detection rates reached the overall average and ranging from 16 to 46 h when looking at where detection rates reached the first local maximum after piling (though at DT a local maximum was never reached). It has to be kept in mind that the overall average includes piling affected data so that this estimate is probably an underestimation. This is more severe than in other models using the complete range of data, because below 2 km distance all positions are affected by piling at least some of the time and based on the present estimates also for the majority of the time that was covered. It may therefore be a more realistic estimate to state when the first local maxima were reached. Results from the analyses of daily POD-data also support the assumption that effects did not last beyond the first day after piling, as detection rates no longer increased from the second to the third day after piling. It cannot be completely ruled out that detection rates would have increased further several days later, but that seems unlikely. If it did, one would have expected porpoise detections to be generally decreased during the years with piling activities, which was not the case. Analyses of all available aerial survey data also mainly found effects lasted about a day after piling. However, due to a limited temporal coverage of aerial survey data with respect to piling activities, conclusions based on these analyses are limited. The present results are largely in line with previous studies that also found clear effects of piling on porpoise detections to last less than two days (TOUGAARD et al. 2009; DIEDERICHS et al. 2010; THOMPSON et al. 2010; BRANDT et al. 2011; HAELTERS et al. 2012; DÄHNE et al. 2013B).

A clear spatial gradient existed in effect duration with shorter lasting effects at greater distances and effects at the largest distances only being detected during piling. This is in line with DIEDERICHS et al. (2010) and BRANDT et al. (2011). TOUGAARD et al. (2009) could not show this for the wind farm Horns Rev 1, but this may be linked to limited data availability. Given that the magnitude of a decrease in detection rates also decreased with distance (indicating that a smaller proportion of porpoises left the area in response to piling noise), it is expected that this also applies to the duration of a negative piling effect. This is because if porpoises left the construction site in response to piling, it will take less time for porpoises to return to the outer areas simply because of the shorter distance to get there. Finally, effects lasting beyond the piling time may not only be a result of piling activities, but also of other construction activities resuming after the end of piling, such as demounting noise mitigation systems and the increased shipping activity that goes with it. One factor that points towards this is that detection rates were already decreased for some time before piling during all seven projects that were investigated during the present study (see below).

8.1.6 Decreased porpoise detections before piling

Decreased detection rates before the start of piling were found from about 24 h before the start of piling from model outputs on the complete dataset. Estimates varied between projects (7-33 h), but a generally similar effect was found in all projects. As this decrease was also strongest at ranges below 5 km distance but detectable until about 10 km, these effects are most likely related to activities in and around the construction site. We do not have detailed information on the activities that occurred prior to piling and on how long before piling these started. Therefore, we can only discuss which of these are likely to cause porpoise deterrence at such distances.

During some projects vibrating occurred prior to piling (GTI, RG and NSO). This is not very loud and noise is not expected to be detectable by porpoises in up to 10 km (BELLMANN, pers. communication), however. Vibrating is also only a short-lasting procedure of a few minutes. Furthermore, vibrating of piles was only used during three projects, but declines in porpoise detections prior to piling occurred in all seven projects. Therefore, vibrating is unlikely to be a significant factor to explain this decline. Loudness of jack-up procedures is also not expected to cause disturbance in up to 10 km, however, the exact noise characteristics of this have not yet been studied.

Something that all projects have in common is an increase in shipping activity around the construction site, but we currently lack information as to the distances at which this occurred. A recent study found behavioural reactions of porpoises to shipping noise at noise levels far below of what had previously been assumed (DYNDO et al. 2015), so such effects in relation to wind farm construction may also have been underestimated so far. On the other hand this assumption is in contrast to findings from an EIA carried out within the Fehmarn Belt, Baltic Sea (MATUSCHEK et al. 2011), which could not relate porpoise detections to differences in shipping activity and to results by a study on the same topic within the Great Belt, Baltic Sea (MORTENSEN et al. 2011). What should be kept in mind is that effects of shipping activity on porpoises may depend on the general level of shipping within a given area. In areas with generally high shipping activity (such as in the Fehmarn Belt) porpoises may be used to it and not react as sensitively to shipping as in generally quieter areas (such as most wind farm areas in the German Bight). But even following DYNDO et al. (2015) effects from shipping activity should not reach further than 2 km, whereas we found effects in up to 5-10 km before piling. It is not known over how big an area shipping traffic increases before piling. There are guard vessels and a ship responsible for carrying out deterrence measures as well as the vessels that set up the noise mitigation systems. They are not expected to operate at distances beyond 2 km from the construction site, so would still not explain an over 5 km deterrence radius. One aspect that may so far be underestimated is the effects that weather has on noise propagation. As already discussed, the present study found that disturbance effects during as well as before piling reached further during low wind speed. Wind farm construction activities are generally limited to calm weather conditions. Therefore, disturbance of porpoises by shipping activity may also reach substantially further prior to and after piling simply because noise travels further at low wind speed due to decreasing natural noise mitigation by air bubbles in the water and increased reflection on the sea surface.

Finally, it is possible that porpoises learn to associate a decrease in wind speed combined with increased noise from ship traffic to upcoming piling activities, so that some animals already leave the construction area several hours before piling occurs in expectation of piling. We found no indication of effects before piling increasing during the course of construction within one wind farm. However, studying such conditioning effects in a meaningful way would necessitate prior knowledge of individual residency patterns within a given area and thus over what time period and within what percentage of animals such learning patterns are expected to develop. In the absence of any studies on individual behaviour of porpoises in the North Sea such information is currently not available.

8.1.7 Cumulative effects and habituation

Frequently re-occurring piling activities within one wind farm could potentially cause habituation of harbour porpoises. After having experienced piling noise several times, they may show a lower degree of avoidance behaviour as they grow used to this noise. On the other hand, it is also possible that animals get sensitised and, after having experienced piling one time, will already leave at a lower stimulus the second time they are exposed to it. The latter would represent a temporal cumulative effect. For future planning of offshore construction activities it would be of great value to gain more insight into the occurrence of such effects when harbour porpoises are repeatedly exposed to piling activities over different lengths of time. However, in order to address habituation or sensitisation effects, one needs to have an idea over what time period and area such effects are to be expected. This is only possible if information on residency patterns of porpoises within local areas of the North Sea are known and some information on individual responses to disturbance exist. As this is currently not the case, it remains difficult to design analyses appropriately. Nevertheless, we addressed this issue by assuming that habituation and sensitisation would increase over the entire construction period within one wind farm.

We found little indication for the occurrence of habituation or sensitisation effects within this study. Only at BWII there was a slight indication for decreases to become smaller during the course of construction, when analysing hourly POD-data. This could indicate that porpoises grew used to piling noise over the construction period. This finding was supported when analysing yearly trends with the daily POD-dataset. It seems surprising, however, that no such effect was found at the other wind farms. The area at BWII was characterised by relatively high detection rates and densities of porpoises that were comparable to the area around DT. A difference to DT is that detection rates were relatively high throughout the year and did not show a distinct peak in early summer. It is possible, therefore, that animals around BWII show a higher degree of site fidelity, which should lead to a higher potential for learning effects to play a role.

Analyses of daily POD-data found a stronger decrease of detection rates during days with two successive piling events within one wind farm but no difference in detection rates between successive piling days with one piling event each. This also gives no indication for cumulative or habituation effects as otherwise detection rates should be lower or higher during the second day of piling than during the first day of piling. Lower detection rates during days with more than just one piling event may simply be a case of additive effects as a greater proportion of the day is affected by piling itself as the time during which piling occurs and therefore also the time that negative clear effects are present accumulate.

Similarly, an effect of piling duration on hourly detection rates was minor when the complete time period was analysed and not detectable when only considering the hour after piling. This also gave no indication for any cumulative effects or habituation.

8.1.8 Context-specific effects

Disturbances often result in different reactions depending on the context under which organisms experience them. Individual fitness, behaviour at the time of exposure (e.g. resting or feeding) and habitat quality for example could affect the way in which porpoises react to piling noise. Foraging animals may be less likely to leave a disturbed site than animals only travelling through.

Porpoises may also be less likely to leave high quality feeding areas than low quality ones. To our knowledge, this issue has not yet been investigated with respect to the effects of offshore construction on porpoises.

We thus approached this challenging question by testing for different piling effects depending on season and subarea based on the daily POD dataset. Statistical reasons made it impossible, however, to combine both contexts in the same model. Therefore, we analysed season and subarea specific effects separately. We found that in autumn and winter acoustic harbour porpoise detections were lower on the first day after piling than on the second day, whereas for spring and summer no differences between the first and second day occurred. We therefore conclude that the recovery from piling effects was longer during autumn and winter. With regard to subarea we found significantly longer recovery times for the BWII-area only. However, within these analyses we were not able to compare the actual effect size between seasons and subareas. We therefore had a closer look at the effect size of piling on the piling day itself: The lowest detection rates were found for the BARD-area and the highest in spring and summer for the BWII-area and the DanTysk-area. We thus cannot make a general claim that lower or higher detection rates in an area or time period lead to longer lasting effects. The duration of the piling effect is related to a certain degree to the magnitude of acoustic harbour porpoise detections on the day of piling and mean detection rate, but matters are too complex to be generalised on this level.

As already discussed, effects of piling as found by analysing hourly POD-data revealed lower effect magnitude and effect range at DT than at any other wind farm. This may be linked to high porpoise detections and densities in this area in summer. However, effects were greater at BWII, where porpoise detections are comparably high even over a longer time period. Therefore, these results also do not point at general differences between high-density and low-density areas. Possibly, specific habitat characteristics, such as prey availability, could play a greater role.

8.2 Comparability of POD- and aerial survey data

Acoustic recordings of porpoises provide relative indices of porpoise activity but cannot at present be directly translated into porpoise density. However, previous studies have found acoustic detections to correlate broadly with porpoise densities obtained from porpoise sightings (Tougaard et al. 2006; Siebert & Rye 2008; Kyhn et al. 2012; Haelters et al. 2013), and attempts are being made to estimate densities from POD-data (Marques et al. 2009; Kyhn et al. 2012). Thus, changes in acoustic porpoise detections seem to be linked to changes in porpoise densities at least to some extent. This is more likely with POD-data at a broader resolution such as at a daily scale (e.g. PP10M/day). This is supported when comparing mean daily acoustic detection rates to average porpoise densities estimated from aerial surveys. Although spatial and temporal coverage differs substantially between aerial survey data and POD-data, both datasets analysed during this study revealed general patterns that correspond well: Daily POD data and aerial survey data highlighted similar seasonal patterns, with highest porpoise occurrence next to the SAC Sylt Outer Reef in the northeast of the German Bight in early summer and another high density area near the SAC Borkum Reef Ground in the southwest almost year round. This is in line with previous findings (Gilles et al. 2014). Furthermore, results from aerial survey data on small-scale effects did not contradict findings from POD-data. Even though aerial survey data were not well suited to address small-scale effects, general findings are largely in line with results from POD-data. The same applies to results from addressing larger-scale effects discussed below. Therefore, it can be concluded that both data reveal similar patterns and generally correspond to each other.

8.3 Larger-scale effects of offshore piling

8.3.1 Predictions from the PCoD model

We applied the interim PCoD model in order to predict population level consequences for harbour porpoises within the German Bight resulting from offshore windfarm construction between 2009 and 2013. Using specifications on disturbance based on results from this study lead to a predicted slight possible effect on the population level. However, these results should be interpreted with caution and not yet used to assess population consequences in the "real world". Although the interim PCoD model is an important spadework on the way to predict population consequences, it clearly requires more background knowledge and subsequent improvement.

Porpoise density was estimated by aerial survey data to be rather constant over the project period despite construction activities. Although the German part of the North Sea population was actually monitored from 2009 to 2013, it is possible that the rather small effects of piling activities modelled by the interim PCoD for 2013 may not yet be visible within the present dataset. This is because most piling events actually occurred in 2012 to 2013. If construction led to reduced fertility, population level consequences may only appear after a few years. Nevertheless, results from application of the interim PCoD model did not contradict the absence of negative yearly trends within the daily POD-data and aerial survey data. This is because even with conservative choices for input parameters considering piling effects the PCoD model predicted only a 30 % risk for a decline of 1 % of the German Bight population of harbour porpoises and a predicted median decline that was always below the 1 % generally considered as critical. Input parameters more closely representing the project data predicted a decline to be even less likely.

Several shortcomings of the PCoD model in relation to biological veracity of the model in terms of process equations and input parameters were found and pointed out. For instance, more background knowledge is required for the parameters currently based on expert judgement (e.g. effects of pile driving on survival and fertility rates, habituation effects, and cumulative effects of multiple construction sites). More information is also needed on the individual risk of getting disturbed multiple times, which will depend on individual movement patterns and individual responses to disturbance. Likewise, potential population resilience (due to high number, high migratory patterns etc.) has to be examined. Further aspects which could affect the harbour porpoise negatively, like food availability/habitat suitability in undisturbed sites (NABE-NIELSEN et al. 2013; VAN BEEST et al. 2015) and noise from other sources (MORRETTI et al. 2014; NABE-NIELSEN et al. 2014; CHRISTIANSEN & LUSSEAU 2015) should also be considered.

So far, results from the PCoD model support our findings that porpoise densities and acoustic detections stayed relatively constant over the four to five year study period despite numerous offshore wind farm construction activities. However, such predictions need to be interpreted with caution due to several shortcomings of the PCoD model and difficulties to predict long-term consequences.

8.3.2 Annual trends

From the present study, there is no indication for a distinct shift in porpoise presence during the five-year study period that could be attributed to offshore wind farm construction, neither from analyses of daily POD-data nor from aerial survey data. Porpoise densities estimated from aerial surveys partly tended to be lower within the vicinity of wind farm construction sites during construction years, but this was no longer apparent within the following years, so this probably only represents short-term effects.

Daily detection rates of porpoises did not significantly decrease from 2010 to 2013 neither within the complete study area, nor when studied separately for the four subareas (BWII-area, MSO/NSO-area, BARD-area, DT-area). When only days without piling activity were considered, there was even a positive trend with highest detection rates occurring in 2013 within the BWII-area and MSO/NSO-area. Similarly, densities estimated from aerial survey data did also not show negative trends from 2009 to 2013, neither over the complete study area nor within any of the four subareas (German Bight N, German Bight W, North of Borkum and West of Sylt). Analyses of aerial survey data did also not reveal any positive trend within any subarea. This is in contrast to the positive trend detected by means of POD-data for some subareas, but note that the chosen subareas were not identical between analyses of POD-data and aerial survey data. In western parts of the German Bight there was a tendency for reduced densities during 2012 when piling occurred at three wind farms. However, densities in 2013 did not significantly differ from any year apart from 2009, when densities were generally lower. This is in line with findings by GILLES et al. (2014B). It has to be borne in mind, however, that confidence intervals for density estimates were rather large due to substantial variance between days even within the same area and in the same month. The magnitude of a change in densities that can be statistically detected depends on the accuracy of density estimates. The wider the confidence intervals the stronger a change in densities needs to be in order to be detected. A positive trend found by POD-data may correspond to a density increase that was too small to be detectable by aerial surveys, and this could also explain this discrepancy. This also raises the question as to the power of detecting potential slight decreases in densities caused by wind farm construction. Nevertheless, as no negative trend was identified, with POD-data or with aerial survey data, it is unlikely that the harbour porpoise population is presently negatively affected by wind farm construction within the German Bight.

Annual aerial surveys carried out within the German Bight for monitoring the NATURA 2000 areas (GILLES et al. 2014A; VIQUERAT et al. 2015) did also not find a decline in porpoise densities in any of the studied subareas from 2009 to 2013. There was a positive trend for porpoise densities in the area around the Natura 2000 area Borkum Reef Ground from 2002 to 2013 (PESCHKO et al. 2016), which is supported by results from POD-data during the present study also showing an increase from 2010 to 2013 within this area.

Altogether, analyses of daily POD-data and aerial survey data did not yield any indication for larger-scale effects of wind farm construction on porpoises during the studied time period 2009 to 2013. Despite clear negative effects of piling on porpoise detections and densities with effect ranges up to about 17 km on average and effect duration in the near vicinity of about 1-2 days, there are no indications for negative effects on porpoises at the population level that arise from the present study. Since 2013 much improvement in noise mitigation techniques occurred so that it can be assumed that avoidance effects of porpoises due to piling were also further reduced.

Thus potential negative effects on the population level due to wind farm construction since then become even less likely.

9 REFERENCES

AGENCY, F. M. AND H. & SAFETY, F. M. FOR THE E. NATURE CONSERVATION AND NUCLEAR (Hrsg.) (2014): Marine habitat modelling for harbour porpoises in the German Bight. In: *Ecological Research at the Offshore Windfarm alpha ventus* Springer Fachmedien Wiesbaden, DOI: 10.1007/978-3-658-02462-8_14, 151–169.ISBN: 978-3-658-02461-1.

AKAIKE, H. (1974): A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19/6, S: 716–723. DOI: 10.1109/TAC.1974.1100705, ISSN: 0018-9286.

MAPS (2014) BY BECKER, R. A., WILKS, A. R., BROWNRIGG, R. & MINKA, T. P. - MAPS (2014): maps: Draw Geographical Maps.

VAN BEEST, F. M., NABE-NIELSEN, J., CARSTENSEN, J., TEILMANN, J. & TOUGAARD, J. (2015): Disturbance Effects on the Harbour Porpoise Population in the North Sea (DEPONS): Status report on model development, No. 140. *Scientific Report from DCE – Danish Centre for Environment and Energy*, Aarhus University, DCE – Danish Centre for Environment and Energy, 1–43.

BESTLEY, S., PATTERSON, T. A., HINDELL, M. A. & GUNN, J. S. (2010): Predicting feeding success in a migratory predator: integrating telemetry, environment, and modeling techniques. *Ecology* 91/8, 2373–2384.

BIOCONSULT SH (Hrsg.) (2009): Responses of harbour porpoises to pile driving at the Horns Rev II offshore wind farm in the Danish North Sea, (Autor: M. J. BRANDT, A. DIEDERICHS & A. NEHLS). BioConsult SH/Husum (DEU), Final report to DONG Energy, 70 pp.

BIOCONSULT SH, HYDROTECHNIK LÜBECK & ITAP (Hrsg.) (2014): Entwicklung und Erprobung des Großen Blasenschleiers zur Minderung der Hydroschallemissionen bei Offshore-Rammarbeiten - HYDROSCHALL-OFF BW II, (Autor: A. DIEDERICHS, H. PEHLKE, G. NEHLS, M. BELLMANN, P. GERKE, J. OLDELAND, C. GRUNAU, S. WITTE & A. ROSE), Schlussbericht. BioConsult SH, HYDROTECHNICK Lübeck & itap/Husum (DEU), Forschungsprojekt gefördert mit Mitteln des Bundesministeriums für Umwelt, Naturschutz und Reaktorsicherheit (BMU), Förderkennzeichen 0325309A/B/C, 247 pp.

RGDAL (2015) BY BIVAND, R., KEITT, T., ROWLINGSON, B., PEBESMA, E., SUMNER, M., HIJMANS, R. & ROUAULT, E. - RGDAL (2015): rgdal: Bindings for the Geospatial Data Abstraction Library.

RGEOS (2015) BY BIVAND, R., RUNDEL, C., PEBESMA, E. & HUFTHAMMER, K. O. - RGEOS (2015): rgeos: Interface to Geometry Engine - Open Source (GEOS).

BJØRGE, A. & TOLLEY, K. A. (2009): Harbour Porpoise Phocoena phocoena. In: *Encyclopedia of Marine Mammals* (2. Auflage). Elsevier.

BRANDT, M. J., DIEDERICHS, A., BETKE, K. & G. NEHLS (2011): Responses of harbour porpoises to pile driving at the Horns Rev II offshore wind farm in the Danish North Sea. *Marine Ecology Progress Series* 421, 205–216.

BRANDT, M. J., HANSEN, S., DIEDERICHS, A. & NEHLS, G. (2014): Do man-made structures and water depth affect the diel rhythms in click recordings of harbor porpoises (*Phocoena phocoena*)? *Marine Mammal Science* 30/3, 1109–1121. DOI: 10.1111/mms.12112, ISSN: 08240469.

BRANDT, M. J., HÖSCHLE, C., DIEDERICHS, A., BETKE, K., MATUSCHEK, R. & NEHLS, G. (2013a): Seal scarers as a tool to deter harbour porpoises from offshore construction sites. *Marine ecology. Progress series* 475, 291–302.

BRANDT, M. J., HÖSCHLE, C., DIEDERICHS, A., BETKE, K., MATUSCHEK, R., WITTE, S. & NEHLS, G. (2013b): Far-reaching effects of a seal scarer on harbour porpoises, *Phocoena phocoena* . *Aquatic Conservation: Marine and Freshwater Ecosystems* 23/2, 222–232. DOI: 10.1002/aqc.2311, ISSN: 1099-0755.

BRUNDIERS, K., KOST, M., BURT, L., LEN, T., MAC AULAY, J. & TIBERI LINQUIST, J. (2014): Detection functions of C-PODs: determining the probability of detecting harbor porpoises using a large hydrophone array. 28th annual conference of the ECS, Liege, Belgique, 00000.

BSH (2013): Standard: Untersuchung der Auswirkungen von Offshore-Windenergieanlagen auf die Meeresumwelt (StUK4), BSH-Nr. 7003. Bundesamt für Seeschifffahrt und Hydrographie/Hamburg, Rostock, pp.

BSH (2015a): BSH. URL: „http://www.bsh.de/en/Marine_uses/Industry/CONTIS_maps/NorthSeaCompleteUsesAndNatureConservation.pdf".

BUCKLAND, S. T., ANDERSON, D. R., BURNHAM, K. P., LAAKE, J. L., BORCHERS, D. L. & THOMAS, L. (2001): Introduction to distance sampling. Oxford University Press/New York, 432 pp.

BUNDESAMT FÜR SEESCHIFFFAHRT UND HYDROGRAPHIE (BSH) (Hrsg.) - BSH (2015b): Bundesfachplan Offshore für die deutsche ausschließliche Wirtschaftszone der Nordsee 2013/2014 und Umweltbericht. Nr. 7603, Hamburg & Rostock (DEU), 195 pp.

BURNHAM, K. P. & ANDERSON, D. R. - MODEL SELECTION AND MULTI-MODEL INFERENCE (2002): Model selection and multi-model inference: a practical information-theoretic approach. Springer/New York.

CARSTENSEN, J., HENRIKSEN, O. & TEILMANN, J. (2006): Impacts of offshore wind farm construction on harbour porpoises: acoustic monitoring of echolocation activity using porpoise detectors (T-PODs). *Marine Ecology Progress Series* 321, 295–308. DOI: 10.3354/meps321295, ISSN: 0171-8630, 1616-1599.

CHRISTIANSEN, F. & LUSSEAU, D. (2015): Linking Behavior to Vital Rates to Measure the Effects of Non-Lethal Disturbance on Wildlife. *Conservation Letters*, 1–8. DOI: 10.1111/conl.12166, ISSN: 1755-263X.

DÄHNE, M., GILLES, A., LUCKE, K., PESCHKO, V., ADLER, S., KRÜGEL, K., SUNDERMEYER, J. & SIEBERT, U. (2013a): Effects of pile-driving on harbour porpoises (*Phocoena phocoena*) at the first offshore wind farm in Germany. *Environmental Research Letters* 8/2, 00000, S: 025002. DOI: 10.1088/1748-9326/8/2/025002, ISSN: 1748-9326.

DÄHNE, M., GILLES, A., LUCKE, K., PESCHKO, V., ADLER, S., KRÜGEL, K., SUNDERMEYER, J. & SIEBERT, U. (2013b): Effects of pile-driving on harbour porpoises (*Phocoena phocoena*) at the first offshore wind farm in Germany. *Environmental Research Letters* 8/2, 00000, S: 025002. DOI: 10.1088/1748-9326/8/2/025002, ISSN: 1748-9326.

DEGRAER, S., BRABANT, R. & RUMES, B. (Hrsg.) (2012): The effect of pile driving on harbour porpoises in Belgian waters, (Autor: J. HAELTERS, W. VAN ROY & S. DREGRAER). *Offshore wind farms in the Belgian part of the North Sea*, Royal Belgian Institute of Natural Sciences (RBINS) and Management Unit of the North Sea Mathematical Models (MUMM), 00005, 127–143.

DIEDERICHS, A., BRANDT, M. J., NEHLS, G., LACZNY, M., HILL, A. & PIPER, W. (2010): Auswirkungen des Baus des Offshore-Testfelds „alpha vetus" auf marine Säugetiere, Bericht, Juli 2010, 120 pp.

DIEDERICHS, A., PEHLKE, H., NEHLS, G., BELLMANN, M., GERKE, P., OLDELAND, J., GRUNAU, C., WITTE, S. & ROSE, A. (2014): Entwicklung und Erprobung des Großen Blasenschleiers zur Minderung der Hydroschallemissionen bei Offshore-Rammarbeiten - Projektkurztitel: HYDROSCHALL-OFF BW II, Schlussbericht. 250pp.

DORMANN, C. F., MCPHERSON, J. M., ARAÚJO, M. B., BIVAND, R., BOLLIGER, J., CARL, G., DAVIES, R. G., HIRZEL, A., JETZ, W., KISSLING, W., KÜHN, I., OHLEMÜLLER, R., PERES-NETO, P. R., REINEKING, B., SCHRÖDER, B., SCHURR, F. M. & WILSON, R. - METHODS TO ACCOUNT FOR SPATIAL AUTOCORRELATION IN THE ANALYSIS OF SPECIES DISTRIBUTIONAL DATA (2007): Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography* 30/5, 609–628. DOI: 10.1111/j.2007.0906-7590.05171.x, ISSN: 1600-0587.

DYNDO, M., WIŚNIEWSKA, D. M., ROJANO-DOÑATE, L. & MADSEN, P. T. (2015): Harbour porpoises react to low levels of high frequency vessel noise. *Nature - Scientific Reports* 5. DOI: 10.1038/srep11083.

ELLIOTT, G., ROTHENBERG, T. J. & STOCK, J. H. (1996): Efficient Tests for an Autoregressive Unit Root. *Econometrica* 64/4, S: 813. DOI: 10.2307/2171846, ISSN: 00129682.

FINNERAN, J. J. & JENKINS, A. K. (2012): Criteria and Thresholds for U.S. Navy Acoustic and Explosive Effects Analysis. SPAWAR Marine Mammal Program, 1–60.

GEELHOED, S. C., SCHEIDAT, M., VAN BEMMELEN, R. S. & AARTS, G. (2013): Abundance of harbour porpoises (*Phocoena phocoena*) on the Dutch Continental Shelf, aerial surveys in July 2010-March 2011. 00000.

GILLES, A. (2003): Verbreitungsmuster von Schweinswalen (*Phocoena phocoena*) in deutschen Gewässern. 153 pp.

GILLES, A., ADLER, S., KASCHNER, K., SCHEIDAT, M. & SIEBERT, U. (2011): Modelling harbour porpoise seasonal density as a function of the German Bight environment: implications for management. *Endangered species research* 14/2, 157–169.

GILLES, A., DÄHNE, M., RONNENBERG, K., VIQUERAT, S., ADLER, S., MEYER-KLAEDEN, O., PESCHKO, V. & SIEBERT, U. (2014a): Ergänzende Untersuchungen zum Effekt der Bau- und Betriebsphase im Offshore-Testfeld „alpha ventus“ auf marine Säugetiere Schlussbericht zum Projekt Ökologische Begleitforschung am Offshore-Testfeldvorhaben alpha ventus zur Evaluierung des Standarduntersuchungskonzeptes des BSH (StUKplus). Stiftung Tierärztliche Hochschule Hannover und Deutsches Meeresmuseum Stralsund/Büsum, 116 pp.

GILLES, A., SCHEIDAT, M. & SIEBERT, U. (2009): Seasonal distribution of harbour porpoises and possible interference of offshore wind farms in the German North Sea. *Marine Ecology Progress Series* 383, 295–307.

GILLES, A. & SIEBERT, U. - Erprobung eines Bund/Länder-Fachvorschlags für das Deutsche Meeresmonitoring von Seevögeln und Schweinswalen als Grundlage für die Erfüllung der Natura 2000-Berichtspflichten mit einem Schwerpunkt in der deutschen AWZ von Nord- und Ostsee (FFH-Berichtsperiode 2007-2012) Teilbericht (2009): Visuelle Erfassung von Schweinswalen. Forschungs- und Technologiezentrum Westküste (FTZ) Christian-Albrechts-Universität zu Kiel und Deutsches Meeresmuseum Stralsund im Auftrag des BfN./Vilm, Mai 2009, 5–30.

GILLES, A., SIEBERT, U., GALLUS, A., DÄHNE, M. & BENKE, H. (2010): Monitoringbericht 2009-2010 Marine Säugetiere und Seevögel in der deutschen AWZ von Nord- und Ostsee - Teilbericht marine Säugetiere -, Endbericht. FTZ, DMM Auftraggeber Bundesamt für Naturschutz/Büsum, Stralsund, 56 pp.

GILLES, A., VIQUERAT, S., SIEBERT, U., GALLUS, A. & BENKE, H. (2014b): Monitoring von marinen Säugetieren 2013 in der deutschen Nord-und Ostsee. 00000, 78 pp.

HAELTERS, J., DEBUSSCHERE, E., BOTTELDOOREN, D., DULIÈRE, V., HOSTENS, K., NORRO, A. M. J., VANDENDRIESSCHE, S., VIGIN, L., VINCX, M. & DEGRAER, S. (2013): The Effects Of Pile Driving On Marine Mammals And Fish In Belgian Waters. Haelters, J., Debusschere, E., Botteldooren, D., Dulière, V., Hostens, K., Norro, A., Vandendriessche, S., Vigin, L., Vincx, M. & S. Degraer (2013) The Effects Of Pile Driving On Marine Mammals And Fish In Belgian Waters. In: Degraer, S., Brabant, R. & B. Rumes (Eds.): Environmental Impacts Of Offshore Wind Farms In The Belgian Part Of The North Sea. Royal Belgian Institute Of Natural Scien¬ces (RBINS), Brussels, Belgium. Pp. 71–77., 71–77.

HAELTERS, J., KERCKHOF, F., JACQUES, T. G., JAUNIAUX, T. & DEGRAER, S. (2011): The return of the harbour porpoise *Phocoena phocoena* to the Belgian part of the North Sea.

HAELTERS, J., VIGIN, L., DEGRAER, S. & OTHERS - ATTRACTION OF HARBOUR PORPOISES TO OFFSHORE WIND FARMS (2013): Attraction of harbour porpoises to offshore wind farms: what can be expected. *Degraer, S, Brabant, R, Rumes, B eds*, 167–171.

HAMMOND, P. S. & MACLEOD, K. (2006): Progress report on the SCANS-II project. Tampere (FIN), Paper prepared for the 13th Advisory Committee to ASCOBANS, 25-27 April, 6 pp.

HAMMOND, P. S., MACLEOD, K., BERGGREN, P., BORCHERS, D. L., BURT, L., CAÑADAS, A., DESPORTES, G., DONOVAN, G. P., GILLES, A., GILLESPIE, D., GORDON, J., HIBY, L., KUKLIK, I., LEAPER, R., LEHNERT, K., LEOPOLD, M., LOVELL, P., ØIEN, N., PAXTON, C. G. M., RIDOUX, V., ROGAN, E., SAMARRA, F., SCHEIDAT, M., SEQUEIRA, M., SIEBERT, U., SKOV, H., SWIFT, R., TASKER, M. L., TEILMANN, J., VAN

CANNEYT, O. & VÁZQUEZ, J. A. (2013): Cetacean abundance and distribution in European Atlantic shelf waters to inform conservation and management. *Biological Conservation* 164, 107–122. DOI: 10.1016/j.biocon.2013.04.010, ISSN: 0006-3207.

HARWOOD, J. & KING, S. (2014): The Sensitivity of UK Marine Mammal Populations to Marine Renewables Developments. REPORT NUMBER SMRUL-NER-2012-027. 1–53.

HARWOOD, J., KING, S., SCHICK, R., DONOVAN, C. & BOOTH, C. (2014): A Protocol for Implementing the Intern Population Consequences of Disturbance (PCOD) Approach: Quantifying and asessing the Effects of UK Offshore Renewble Energy Developments on Marine Mammal Populations. REPORT NUMBER SMRUL-TCE-2013-014. Scottish Marine and Freshwater Science 5/2, 1–90.

HASTIE, T. J. & TIBSHIRANI, R. J. (1990): Generalized additive models. (43), Chapman & Hall/CRC.

HEINÄNEN, S., SKOV, H., ARREBORG HANSEN, D., LADAGE, F., SCHLENZ, S., ŽYDELIS, R., THOMSEN, F. & HAMMRICH, A. (2015): Analyse von Langzeitdaten und Modellierung der Verteilung von Schweinswalen im Testfeld *alpha ventus* als Grundlage von Entscheidungshilfen für die maritime Raumordnung; Schlussbericht zum Projekt; Ökologische Begleitforschung am Offshore-Testfeldvorhaben *alpha ventus* zur Evaluierung des Standarduntersuchungskonzeptes des BSH(STUKplus). DHI, DHI-WASY GmbH/Hamburg.

HEINIS, F. & DE JONG, C. A. F. (2015): Cumulative effects of impulsive underwater sound on marine mammals. *TNO report*, Nr. TNO 2015 R10335-A, TNO, 86 pp.

FPC (2015) BY HENNIG, C. - FPC (2015): fpc: Flexible Procedures for Clustering.

HIBY, A. R. & HAMMOND, P. S. (1989): Survey techniques for estimating abundance of cetaceans. *Report of the International Whaling Commission* 11/Special Issue.

FORECAST (2015) BY HYNDMAN, R. J., ATHANASOPOULOS, G., BERGMEIR, C., CINELLI, C., KHAN, Y., MAYER, Z., RAZBASH, S., SCHMIDT, D., SHAUB, D., TANG, Y., WANG, E. & ZHOU, Z. - FORECAST (2015): forecast: Forecasting Functions for Time Series and Linear Models.

ICES (2014): Report of the Working Group on Marine Mammal Ecology (WGMME), 10–13 March 2014, Woods Hole, Massachusetts, USA. International Council for the Exploration of the Sea Conseil International pour l'Exploration de la Mer/Copenhagen, 234 pp.

INGENITO, F. & WOLF, S. N. (1989): Site dependence of wind-dominated ambient noise in shallow water. *The Journal of the Acoustical Society of America* 85/1, 141. DOI: 10.1121/1.397722, ISSN: 00014966.

JONES, A., SENDT, J., DUNCAN, A. J., CLARKE, PAUL A & MAGGI, AMOS (2009): Modelling the acoustic reflection loss at the rough ocean surface. Konf.: *Acoustics*. Australian Acoustical Society.

KASTELEIN, R. A., BUNSKOEK, P., HAGEDOORN, M., AU, W. W. L. & DE HAAN, D. (2002): Audiogram of a harbor porpoise (Phocoena phocoena) measured with narrow-band frequency-modulated signals. *The Journal of the Acoustical Society of America* 112/1, 334-344. DOI: 10.1121/1.1480835, ISSN: 00014966.

KASTELEIN, R. A., VAN HEERDEN, D., GRANSIER, R. & HOEK, L. (2013): Behavioural responses of a harbour porpoise (Phocoena phocoena) to playbacks of broadband pile driving sounds. *Marine Environmental Research* 92, 206-2014.

KING, S. L., SCHICK, R. S., DONOVAN, C., BOOTH, C., BURGMAN, M., THOMAS, L. & HARWOOD, J. (2015): An interimframework for assessing the population consequences of disturbance. Methods in Ecology and Evolution, 6, 1150-1158. DOI: doi: 10.1111/2041-210X.12411.

KLÜVER, R. & IFAÖ (2011): FULMAR. Wollert Information Technology/Hamburg.

KNEIB, T. & FAHRMEIR, L. - STRUCTURED ADDITIVE REGRESSION FOR CATEGORICAL SPACE–TIME DATA (2006): Structured additive regression for categorical space–time data: A mixed model approach. *Biometrics* 62/1, 109–118.

KOSCHINSKI, S., CULIK, B., DAMSGAARD HENRIKSEN, O., TREGENZA, N., ELLIS, G., JANSEN, C. & KATHE, G. (2003): Behavioural reactions of free-ranging porpoises and seals to the noise of a simulated 2 MW windpower generator. *Marine Ecology Progress Series* 265, 263–273. DOI: 10.3354/meps265263, ISSN: 0171-8630, 1616-1599.

KYHN, L. A., TOUGAARD, J., THOMAS, L., DUVE, L. R., STENBACK, J., AMUNDIN, M., DESPORTES, G. & TEILMANN, J. (2012): From echolocation clicks to animal density—Acoustic sampling of harbor porpoises with static dataloggers. *The Journal of the Acoustical Society of America* 131/1, 550–560.

MRDS (2015) BY LAAKE, J., BORCHERS, D., THOMAS, L., MILLER, D. & BISHOP, J. - MRDS (2015): mrds: Mark-Recapture Distance Sampling.

LAAKE, J. L., CALAMBOKIDIS, J., OSMEK, S. D. & RUGH, D. J. - PROBABILITY OF DETECTING HARBOR PORPOISE FROM AERIAL SURVEYS (1997): Probability of detecting harbor porpoise from aerial surveys: estimating g (0). *The Journal of wildlife management*, 63–75.

LAURSEN, K., FRIKKE, J. & KAHLERT, J. (2008): Accuracy of'total counts' of waterbirds from aircraft in coastal waters. *Wildlife Biology* 14/2, 165–175.

MACLEAN, I. M. ., REHFISCH, M., SKOV, H. & THAXTER, C. B. (2012): Evaluating the statistical power of detecting changes in the abundance of seabirds at sea. *IBIS*, 155, 113-126. DOI: 10.1111/j.1474-919X.2012.01272.x.

MACLEOD, K., LACEY, C., QUICK, N., HASTIE, G. & WILSON, J. (2011): Guidance on survey and monitoring in relation to marine renewables deployments in Scotland. Volume 2. Cetaceans and Basking Sharks. *Unpublished draft report to Scottish Natural Heritage and Marine Scotland*.

MADSEN, P. T., WAHLBERG, M., TOUGAARD, J., LUCKE, K. & TYACK, P. L. (2006): Wind turbine underwater noise and marine mammals: implications of current knowledge and data needs. *Marine Ecology Progress Series* 309, 279–295.

MAECHLER, M., ROUSSEEUW, P., STRUYF, A., HUBERT, M. & HORNIK, K. (2015): cluster: Cluster Analysis Basics and Extensions.

MARQUES, T. A., THOMAS, L., WARD, J., DIMARZIO, N. & TYACK, P. L. - MARQUES et al. (2009): Estimating cetacean population density using fixed passive acoustic sensors: An example with Blainville's beaked whales. *The Journal of the Acoustical Society of America* 125/4, 1982–1994.

MATUSCHEK, R., BRANDT, M., DIEDERICHS, ANSGAR, HÖSCHLE, C., NEHLS, G., THOMSEN, F. & MACKENZIE, M. (2011): Fehmarnbelt Marine Mammal Studies - Ambient noise in the Fehmarnbelt area. Cefas, Bioconsult-SH, Biola, DHI, SMRU, ITAP.

MCDONALD, M. A. & FOX, C. G. (1999): Passive acoustic methods applied to fin whale population density estimation. *The Journal of the Acoustical Society of America* 105, 2643-2651.

MCGOVERN, B., BERROW, S. & MCKEOGH, E. (Hrsg.) (2012): Testing POD detection range under optimal field conditions. Konf.: *26th European Ceatcean Society Meeting*. Galway, Ireland, 26th-28th March 2012.

MIKKELSEN, L., MOURITSEN, K., DAHL, K., TEILMANN, J. & TOUGAARD, J. (2013): Re-established stony reef attracts harbour porpoises Phocoena phocoena. *Marine Ecology Progress Series* 481, 239–248. DOI: 10.3354/meps10260, ISSN: 0171-8630, 1616-1599.

DSM (2015) BY MILLER, D. L., REXSTAD, E., BURT, L., BRAVINGTON, M. V. & HEDLEY, S. - DSM (2015): dsm: Density Surface Modelling of Distance Sampling Data.

MORAN, P. A. P. (1950): Notes on Continuous Stochastic Phenomena. *Biometrika* 37/1/2, 17. DOI: 10.2307/2332142, ISSN: 00063444.

MORRETTI, D., THOMAS, L., MARQUES, T. A., HARWOOD, J., DILLEY, A., NEALES, B., SHAFER, J., MCCARTHY, E., NEW, L. F., JARVIS, S. & MORRISSEY, R. (2014): A Risk Function for Behavioral Disruption of Blainville's Beaked Whales (Mesoplodon densirostris) from Mid- Frequency Active Sonar. PLoS One 9/1, S: e85064. DOI: doi:10.1371/journal.pone.0085064.

MORTENSEN, L. O., TOUGAARD, J. & TEILMANN, J. (2011): Effects of underwater noise on harbour porpoises around major shipping lanes. Aarhus University, DCE – Danish Centre for Environment and Energy.

NABE-NIELSEN, J., SIBLY, R. M., TOUGAARD, J., TEILMANN, J. & SVEEGAARD, S. (2014): Effects of noise and by-catch on a Danish harbour porpoise population. *Ecological Modelling* 272, 242–251. DOI: 10.1016/j.ecolmodel.2013.09.025, ISSN: 03043800.

NABE-NIELSEN, J., TOUGAARD, J., TEILMANN, J., LUCKE, K. & FORCHHAMMER, M. C. (2013): How a simple adaptive foraging strategy can lead to emergent home ranges and increased food intake. *Oikos* 122/9, 1307–1316. DOI: 10.1111/j.1600-0706.2013.00069.x, ISSN: 1600-0706.

NATIONAL RESEARCH COUNCIL (2005): Marine Mammal Populations and Ocean Noise: Determining when noise causes biologically significant effects. Washington D.C., 1–126.

NEHLS, G. & BELLMANN, M. (2013): Auswirkungen von Schallemissionen durch Offshore-Rammarbeiten auf Schweinswale. ITAP & Bioconsult-SH.

NEHLS, G., ROSE, A., DIEDERICHS, A., BELLMANN, M. & PEHLKE H. (2016): Noise Mitigation During Pile Driving Effi ciently Reduces Disturbance of Marine Mammals. In: A.N. POPPER, A. HAWKINS (EDS.), The Effects of Noise on Aquatic Life II, Advances in Experimental Medicine and Biology 875, Chapter 92. DOI 10.1007/978-1-4939-2981-8_92

RCOLORBREWER (2014) by NEUWIRTH, E. - RCOLORBREWER (2014): RColorBrewer: ColorBrewer Palettes.

OWP BUTENDIEK GMBH (Hrsg.) (2014): Offshore Windpark Butendiek. Ökologisches Monitoring. Untersuchungskonzept für die Bauphase, (Autor: BIOCONSULT SH).

SP (2015) BY PEBESMA, E., BIVAND, R., ROWLINGSON, B., GOMEZ-RUBIO, V., HIJMANS, R., SUMNER, M., MACQUEEN, D., LEMON, J. & O'BRIEN, J. - SP (2015): sp: Classes and Methods for Spatial Data.

PELTIER, H., BAAGØE, H. J., CAMPHUYSEN, K. C., CZECK, R., DABIN, W., DANIEL, P., DEAVILLE, R., HAELTERS, J., JAUNIAUX, T. & JENSEN, L. F. (2013): The stranding anomaly as population indicator: the case of harbour porpoise Phocoena phocoena in North-Western Europe. *PLoS One* 8/4, e62180.

PESCHKO, V., RONNENBERG, K., SIEBERT, U. & GILLES, A. (2016): Trends of harbour porpoise (*Phocoena phocoena*) density in the southern North Sea. *Ecological Indicators* 60, 174–183.

PHILLIPS, P. C. B. & PERRON, P. (1988): Testing for a unit root in time series regression. *Biometrika* 75/2, 335–346. DOI: 10.1093/biomet/75.2.335, ISSN: 0006-3444, 1464-3510.

NLME (2015) by PINHEIRO, J., BATES, D., DEBROY, S., SARKAR, D., EISPACK AUTHORS & R-CORE - NLME (2015): nlme: Linear and Nonlinear Mixed Effects Models.

POPPER, A. N. & HAWKINS, A. (Hrsg.) - NEHLS et al. (2016): Noise Mitigation During Pile Driving Efficiently Reduces Disturbance of Marine Mammals. In: *The Effects of Noise on Aquatic Life II*(875), Springer New York/New York, NY, S. 755–762.ISBN: 978-1-4939-2980-1.

PORTER, F. J.-W. K.-M. & SCHMIDT, H. (2000): Computational ocean acoustics. Springer-Verlag: New York, NY, USA.

PUNT, A. (Hrsg.) (2015): Sound exposure in harbour seals during the installation of an offshore wind farm: predictions of auditory damage. *Journal of Applied Ecology* 52/3, 631–640. DOI: 10.1111/1365-2664.12403, ISSN: 00218901.

R CORE TEAM - R (2014): R: A language and environment for statistical computing. R Foundation for Statistical Computing/Vienna, Austria.

FOREIGN (2015) by R. CORE TEAM, BIVAND, R., CAREY, V. J., DEBROY, S., EGLEN, S., GUHA, R., LEWIN-KOH, N., MYATT, M., PFAFF, B., QUISTORFF, B., WARMERDAM, F., WEIGAND, S. & FREE SOFTWARE FOUNDATION, INC. - FOREIGN (2015): foreign: Read Data Stored by Minitab, S, SAS, SPSS, Stata, Systat, Weka, dBase, ...

R DEVELOPMENT CORE TEAM (2015): R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria.

REID, J. B., EVANS, P. G. & NORTHRIDGE, S. P. (2003): Atlas of cetacean distribution in north-west European waters. Joint Nature Conservation Committee.

RICHARDSON, W. J., GREENE, C. R., MALME, C. I. & THOMSON, D. H. (1995): Marine Mammals and Noise Academic Press. *San Diego*.

TIMEDATE (2015) by RMETRICS CORE TEAM, WUERTZ, D., SETZ, T., CHALABI, Y., MAECHLER, M. & BYERS, J. W. - TIMEDATE (2015): timeDate: Rmetrics - Chronological and Calendar Objects.

ROSE., A., DIEDERICHS, A., NEHLS, G., BRANDT, M., WITTE, S., HÖSCHLE, C., DORSCH, M., LIESENJOHANN, T., SCHUBERT, A., KOSAREV, V., LACZNY, M., HILL., A. & PIPER, W. (2014): Offshore Test Site Alpha

Ventus – Expert Report: Marine Mammals – Final Report: From baseline to wind farm operation. Im Auftrag der DOTI. BioConsult SH, Husum, IfAÖ, Hamburg., 163 pp.

RUSSELL, D. J. F., MCCLINTOCK, B. T., MATTHIOPOULOS, J., THOMPSON, P. M., THOMPSON, D., HAMMOND, P. S., JONES, E. L., MACKENZIE, M. L., MOSS, S. & MCCONNELL, B. J. (2015): Intrinsic and extrinsic drivers of activity budgets in sympatric grey and harbour seals. *Oikos* 124/11, 1462–1472. DOI: 10.1111/oik.01810, ISSN: 00301299.

LATTICE (2008) by SARKAR, D. - LATTICE (2008): Lattice: multivariate data visualization with R. (1. Auflage). Springer New York, 268 pp. ISBN: 0-387-75968-9.

LATTICEEXTRA (2013) by SARKAR, D. & ANDREWS, F. - LATTICEEXTRA (2013): latticeExtra: Extra Graphical Utilities Based on Lattice.

SCANS, I. I. (2008): Small Cetaceans in the European Atlantic and North Sea (SCANS II). *Final Report to the European Commission under project LIFE04NAT/GB/000245. Available from SMRU, Gatty Marine Laboratory, University of St Andrews, St Andrews, Fife, KY16 8LB, UK*.

SCHEIDAT, M., VERDAAT, H. & AARTS, G. (2012): Using aerial surveys to estimate density and distribution of harbour porpoises in Dutch waters. *Journal of Sea Research* 69, 1–7. DOI: 10.1016/j.seares.2011.12.004, ISSN: 1385-1101.

SCHICK, R., HARWOOD, J., KING, S., BOOTH, C. & DONOVAN, C. (2014): A Tutorial for the Interim PCOD Code. 1–43.

SHAPIRO, A. D., TOUGAARD, J., JØRGENSEN, P. B., KYHN, L. A., BALLE, J. D., BERNARDEZ, C., FJÄLLING, A., KARLSEN, J. & WAHLBERG, M. (2009): Transmission loss patterns from acoustic harassment and deterrent devices do not always follow geometrical spreading predictions. *Marine Mammal Science* 25/1, 53–67. DOI: 10.1111/j.1748-7692.2008.00243.x, ISSN: 08240469.

SIEBERT, U., GILLES, A., LUCKE, K., LUDWIG, M., BENKE, H., KOCK, K.-H. & SCHEIDAT, M. (2006): A decade of harbour porpoise occurrence in German waters—Analyses of aerial surveys, incidental sightings and strandings. *Journal of Sea Research* 56/1, 65–80. DOI: 10.1016/j.seares.2006.01.003, ISSN: 13851101.

SIEBERT, U. & RYE, J. H. (2008): Correlation between aerial surveys and acoustic monitoring. *Marine mammals and seabirds in front of offshore wind energy. Teubner Verlag, Wiebaden*, 37–39.

SMRU (2014): Interim PCoD v.1.1: A „How To“ Guide. SMRU Consulting, 28 pp.

SOUTHALL, B. L., BOWLES, A. E., ELLISON, W. T., FINNERAN, J. J., GENTRY, R. L., GREENE, C. R., KASTAK, D., KETTEN, D. R., MILLER, J. H., NACHTIGALL, P. E., RICHARDSON, W. J., THOMAS, J. A. & TYACK, P. L. (2007): Marine Mammal Noise Exposure Criteria: Initial Scientific Recommendations. *Aquatic Mammals* 33/4, 411–509. ISSN: ISSN 0167-5427.

SHAPEFILES (2013) by STABLER, B. - SHAPEFILES (2013): shapefiles: Read and Write ESRI Shapefiles.

SVEEGAARD, S., TEILMANN, J., BERGGREN, P., MOURITSEN, K. N., GILLESPIE, D. & TOUGAARD, J. (2011): Acoustic surveys confirm the high-density areas of harbour porpoises found by satellite tracking. *ICES Journal of Marine Science* 68/5, 929–936.

TEILMANN, J., CHRISTIANSEN, C. T., KJELLERUP, S., DIETZ, R. & NACHMAN, G. (2013): Geographic, seasonal, and diurnal surface behavior of harbor porpoises. *Marine Mammal Science* 29/2, E60–E76. DOI: 10.1111/j.1748-7692.2012.00597.x, ISSN: 1748-7692.

TEILMANN, J. & TOUGAARD, J. (2006): Reactions of captive Harbor porpoises (Phocoena phocoena) to pinger-like sounds. *Marine Mammal Science* 22/2, 240–260. DOI: 10.1111/j.1748-7692.2006.00031.x.

THIELE, R. (2002): Propagation loss values for the North Sea. Handout zum Fachgespräch: Offshore-Windmills-sound-emissions and marine mammals. FTZ-Büsum, 15.01.2002.

THIELE, R. & SCHELLSTEDE, G. (1980): Standardwerte zur Ausbreitungsdämpfung in der Nordsee. Nr. FWG Bericht 1980-7, Forschungsanstalt der Bundeswehr für Wasserschall und Geophysik.

THOMPSON, P. M., LUSSEAU, D., BARTON, T., SIMMONS, D., RUSIN, J. & BAILEY, H. (2010): Assessing the responses of coastal cetaceans to the construction of offshore wind turbines. *Marine Pollution Bulletin* 60/8, 1200–1208. DOI: 10.1016/j.marpolbul.2010.03.030, ISSN: 0025326X.

THOMSEN, F., LACZNY, M. & PIPER, W. (2004): Methodik zur Erfassung von Schweinswalen (*Phocoena phocoena*) und anderen marinen Säugern mittels Flugtransekt-Zählungen. *Seevögel* 25 (1), 3–12.

THOMSEN, F., LACZNY, M. & PIPER, W. (2006): A recovery of harbour porpoises (Phocoena phocoena) in the southern North Sea? A case study off Eastern Frisia, Germany. *Helgoland marine research* 60/3, 189–195.

THOMSEN, F., UGARTE, F. & EVANS, P. G. . (2005): Estimation of g (0) in line-transect surveys of cetaceans. *European Cetacean Society Newsletter No* 44 - Special Issue, 46 pp.

TOUGAARD, J., CARSTENSEN, J., WISZ, M., JESPERSEN, M., TEILMANN, J., BECH, N. & SKOV, H. (2006): Harbour porpoises on Horns Reef. Effects of the Horns Reef Wind Farm. Final Report to Vattenfall A/S., Commissioned report. NERI/Roskilde, Denmark, 111 pp.

TOUGAARD, J., HENRIKSEN, O. D. & MILLER, L. A. (2009): Underwater noise from three types of offshore wind turbines: Estimation of impact zones for harbor porpoises and harbor seals. *The Journal of the Acoustical Society of America* 125/6, 3766–3773. DOI: 10.1121/1.3117444, ISSN: 0001-4966.

TOUGAARD, J., WRIGHT, A. J. & MADSEN, P. T. (2015): Cetacean noise criteria revisited in the light of proposed exposure limits for harbour porpoises. *Marine pollution bulletin* 90/1, 196–208.

TSERIES (2015) by TRAPLETTI, A., HORNIK, K. & BLAKE, L. - TSERIES (2015): tseries: Time Series Analysis and Computational Finance.

TREGENZA, N. (2011): CPOD.exe: A guide for users.

FOSSIL (2012) by VAVREK, M. J. - FOSSIL (2012): fossil: Palaeoecological and Palaeogeographical Analysis Tools.

VERFUSS, U. K., HONNEF, C. G., MEDING, A., DÄHNE, M., MUNDRY, R. & BENKE, H. (2007): Geographical and seasonal variation of harbour porpoise (Phocoena phocoena) presence in the German Baltic Sea revealed by passive acoustic monitoring. *Journal of the Marine Biological Association of the UK* 87/01, 165. DOI: 10.1017/S0025315407054938, ISSN: 0025-3154, 1469-7769.

VIQUERAT, S., GILLES, A., HERR, H., SIEBERT, U., GALLUS, A., BENKE, H. & KRÜGEL, K. (2015): Monitoring von marinen Säugetieren 2014 in der deutschen Nord-und Ostsee. 00000, 83 pp.

WEILGART, L. S. (2007): The impacts of anthropogenic ocean noise on cetaceans and implications for management. *Canadian Journal of Zoology* 85/11, 1091–1116. DOI: 10.1139/Z07-101, ISSN: 0008-4301, 1480-3283.

WILLE, P. C. & GEYER, D. (1984): Measurements on the origin of the wind-dependent ambient noise variability in shallow water. *The Journal of the Acoustical Society of America* 75/1, 173. DOI: 10.1121/1.390411, ISSN: 00014966.

MGCV (2015) BY WOOD, S. - MGCV (2015): mgcv: Mixed GAM Computation Vehicle with GCV/AIC/REML Smoothness Estimation.

WOOD, S. N. (2006): On confidence intervals for generalized additive models based on penalized regression splines. *Australian & New Zealand Journal of Statistics* 48/4, 00048, 445–464.

WOOD, S. N. & AUGUSTIN, N. H. (2002): GAMs with integrated model selection using penalized regression splines and applications to environmental modelling. *Ecological modelling* 157/2, 157–177.

ZOO (2015) by ZEILEIS, A., GROTHENDIECK, G., RYAN, J. A. & ANDREWS, F. - ZOO (2015): zoo: S3 Infrastructure for Regular and Irregular Time Series (Z's Ordered Observations).

ZUUR, A. F. (2012): A Beginner's Guide to Generalized Additive Models with R. Highland Statistics Ltd, 206 Seiten. ISBN: 0-9571741-2-8.

ZUUR, A., IENO, E. N., WALKER, N., SAVELIEV, A. A. & SMITH, G. M. (2009): Mixed effects models and extensions in ecology with R. Springer/New York, 03313, 574 pp.

A. APPENDIX

A.1 Noise-data

The used noise propagation model proposed by ITAP GmbH is suited to impulsive pile driving noise and leads to a transmission loss curve shown in Figure A-1. In this figure, three different transmission models are plotted as lines, where the used model is denoted by the black dashed line. As comparison a formula introduced by Thiele & Schellstede (1980) and a simple logarithmic transmission (15 log R) are plotted. The simple logarithmic transmission assumes an inverse proportionality between noise pressure and the logarithm of the distance to the source. Red crosses are median values of the noise exposure level of pile driving noise measurements from 70 meters up to approximately 25 km distance to the noise source. As can be seen, the ITAP formula fits best to the measured data.

When noise measurement was not available, the computed values were derived by calculating the transmission loss of the distance between the measured value and the desired immission point. The obtained transmission loss was then subtracted to the measured value. This leads to an approximation of the SEL at the desired immission point. At BARD, only one noise measurement was available and all other calculated SEL were thus based on this one value.

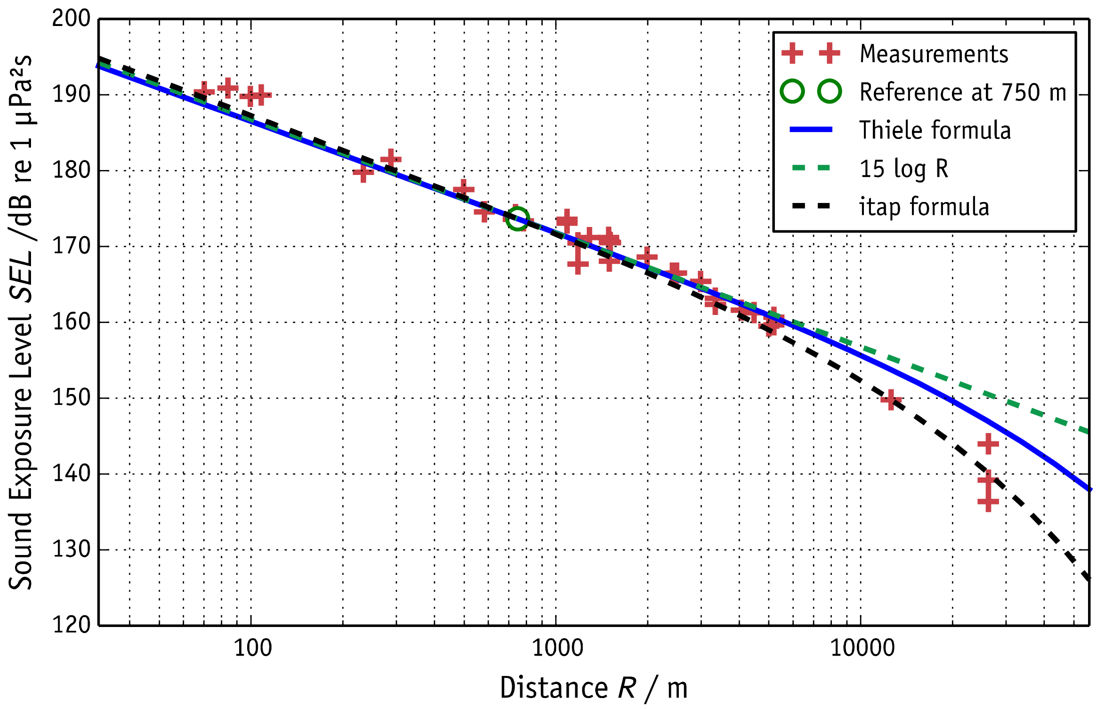


Figure A-1 *Comparison of transmission models and measured Noise Exposure Levels. The red crosses indicate noise measurements. The green circle is a reference value at 750 m distance. The blue line is a model proposed by Thiele & Schellstede (1980). A simple propagation model (green dashed line) assumes an inverse proportionality of noise pressure with distance logarithm to noise source. The model proposed by ITAP GmbH is plotted as a black dashed line that best fits noise measurements.*

A.2 Hourly POD-data

A.2.1 Additional tables and figures for chapter 4.3.1-4.3.5 (spatial range and duration of piling effects)

Investigating the interaction of hour relative to piling with distance combining all data

In order to test whether effects before piling are not only a result from smoothing functions within the model, we also present the raw data for DPH at the different hours related to piling and for three different distance categories. Here, Figure A-2-Figure A-4 illustrate DPH at intermediate distance classes (10-30 km). In addition, Table A-1 presents the results from models G1 and G2 that were run on the global dataset (i.e. all seven wind farm projects).

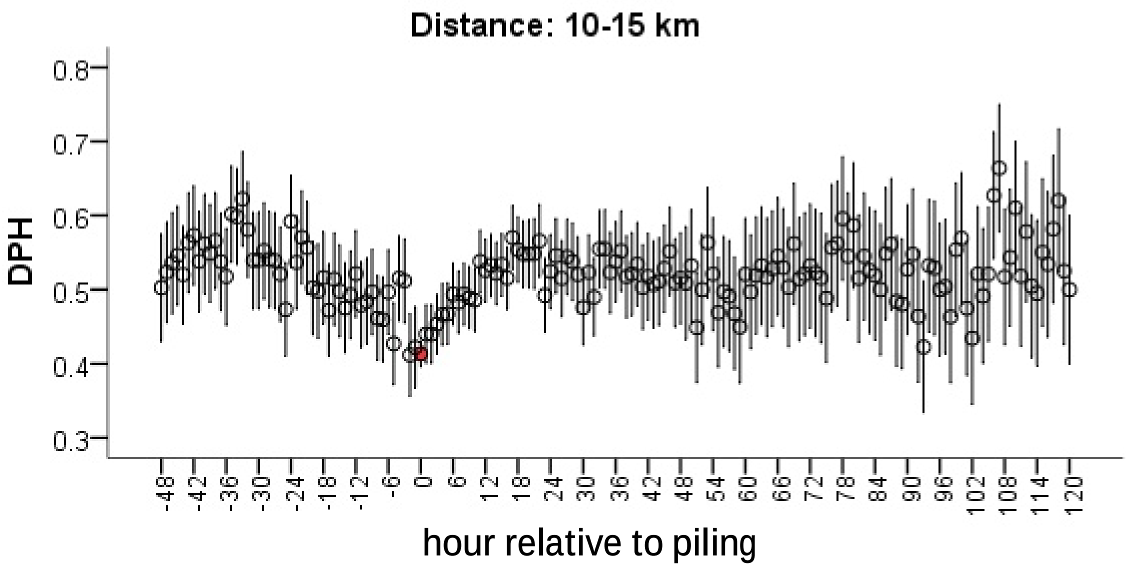


Figure A-2 Error bars depicting mean and 95% confidence intervals of DPH at the different hours relative to piling for all wind farm projects and for distances between 10 and 15 km. The point and error bar for HRW=0 are shown in red.

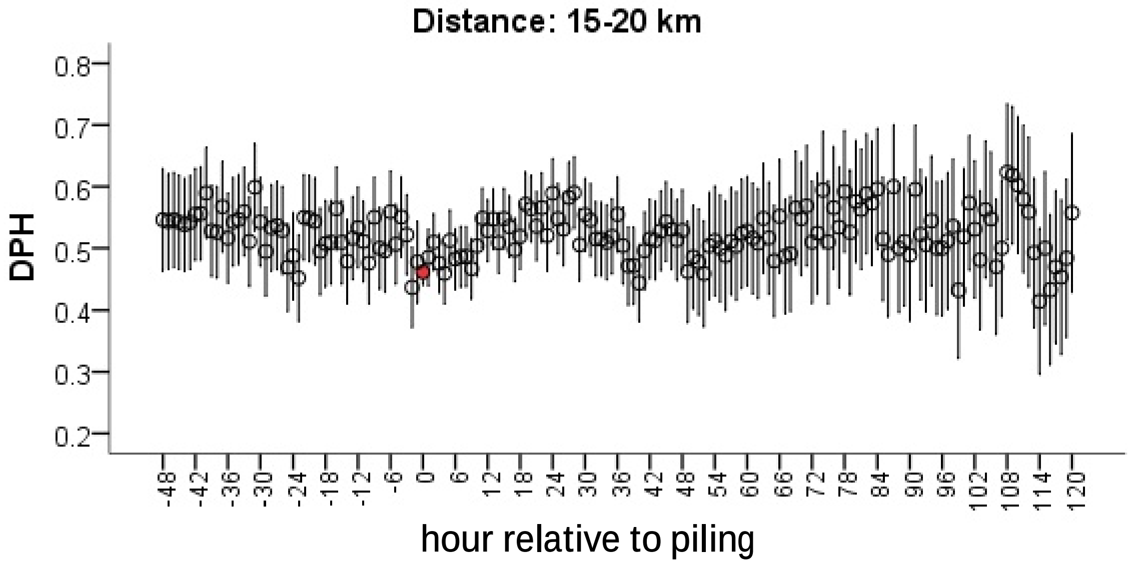


Figure A-3 *Error bars depicting mean and 95% confidence intervals of DPH at the different hours relative to piling for all wind farm projects and for distances between 15 and 20 km. The point and error bar for HRW=0 are shown in red.*

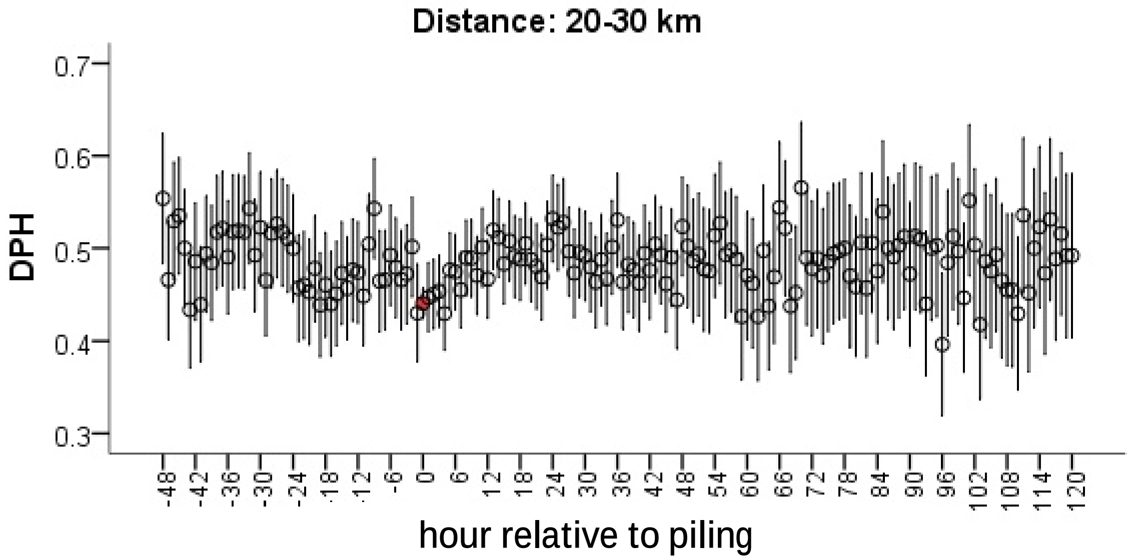


Figure A-4 *Error bars depicting mean and 95% confidence intervals of DPH at the different hours relative to piling for all wind farm projects and for distances between 20 and 30 km. The point and error bar for hour relative to piling=0 are shown in red.*

Table A-1 Results from Model G1 and G2 run on the global dataset including data from all seven wind farm projects. Deviance explained: 6.8 % for G1 and 7.4 % for G2.

Model:	G1 (all data)			G2 (all data)		
Variable	(e)df	Chi²	p	(e)df	Chi²	p
DPH(t-1)	1	337	***	1	499	***
POD-Position (random factor)	219123	7891	***	445.8	12393	***
Year (factor)	3	779	***	3	1134	***
day of year (smooth)	8.0	8586	***	8.0	15882	***
HH (smooth)	7.0	909	***	7.1	1015	***
wind speed (smooth)	8.0	2053	***	8.3	3007	***
wind direction (smooth)	6.5	131	***	7.2	448	***
SSTA (smooth)	8.7	764	***	8.7	874	***
noise clicks (smooth)	6.1	1687	***	7.7	2118	***
sediment (factor)	4	21	***	4	36	***
pilingduration (smooth)	8.9	583	***	8.9	625	***
hour relative to piling, distance (interaction)	-	-	-	28.4	2535	***
hour relative to piling, SEO_{05} (interaction)	28.4	2204	***	-	-	-

Table A-2 Results from models G3 run to check for the effects of noise mitigation on effect ranges of piling (explained deviance: 7.5 %) and model G4 run to look at effect duration at close distances (data <2 km distance from piling, explained deviance: 16.8 %).

model:	G3 (noise mitigation yes or no)			G4 (distance < 2 km)		
variable	(e)df	Chi²	p	(e)df	Chi²	p
DPH(t-1) (factor)	1	483.4	***	1	0.4	ns
POD-Position (random factor)	448.8	12246	***	257.4	399	***
Year (factor)	3	1136	***	3	42	***
day of year (smooth)	8.0	14541	***	7.7	9105	***
HH (smooth)	7.1	980	***	6.1	127	***
wind speed (smooth)	8.3	2878	***	3.4	298	***
wind direction (smooth)	7.3	370	***	4.0	60	**
SSTA (smooth)	8.7	892	***	7.9	57	***
noise clicks (smooth)	7.7	2043	***	2.1	95	***
sediment (factor)	4	35	***	4	14	**
Pilingduration (smooth)	8.9	664	***	8.0	507	***
hour relative to piling, distance, for noise mitigation=no (interaction)	28.1	1153	***	-	-	-
hour relative to piling, distance, for noise mitigation=yes (interaction)	28.0	1701	***	-	-	-
hour relative to piling (smooth)	-	-	-	8.0	63	***

Raw data plots for DPH at the different hours relative to piling activity and at distances below 5 km from piling are shown for each wind farm project Figure A-5-to Figure A-11.

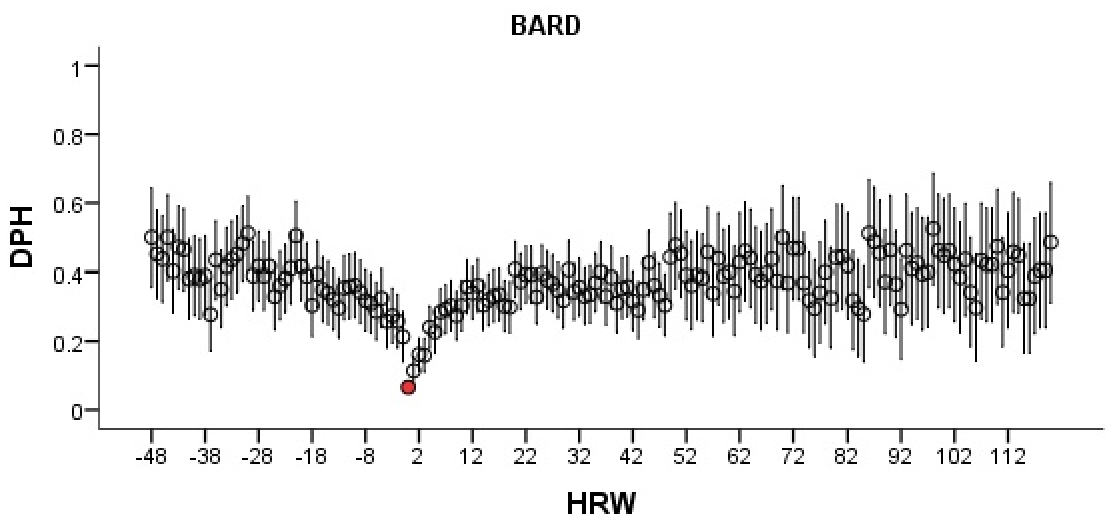


Figure A-5 Error bars depicting mean and 95% confidence intervals of DPH at the different hours relative to piling at distances below 5 km for the wind farm project BARD. The error bar for hour relative to piling=0 is shown in red.

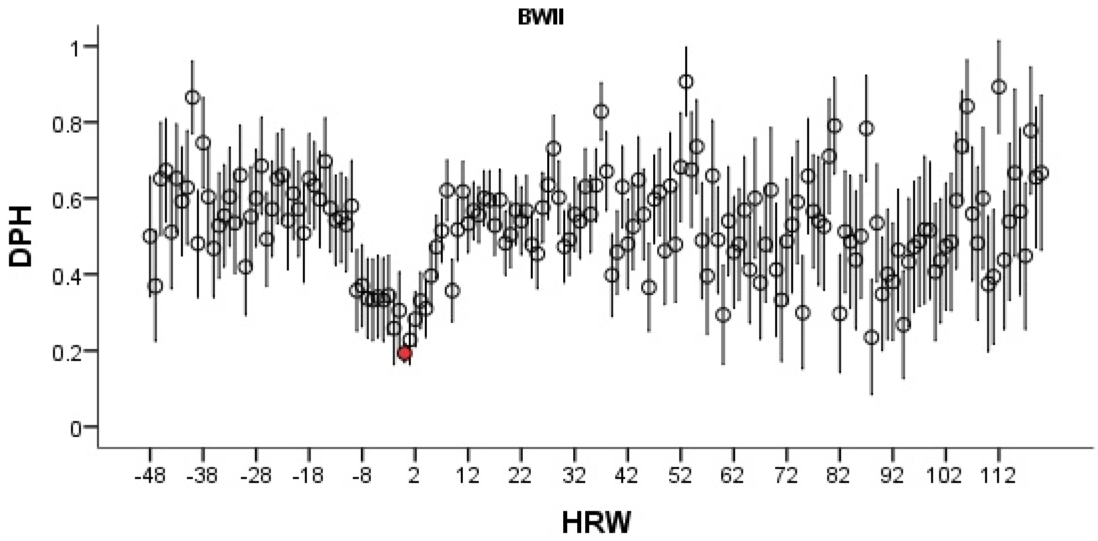


Figure A-6 Error bars depicting mean and 95% confidence intervals of DPH at the different hours relative to piling at distances below 5 km for the wind farm project BWII. The error bar for hour relative to piling=0 is shown in red.

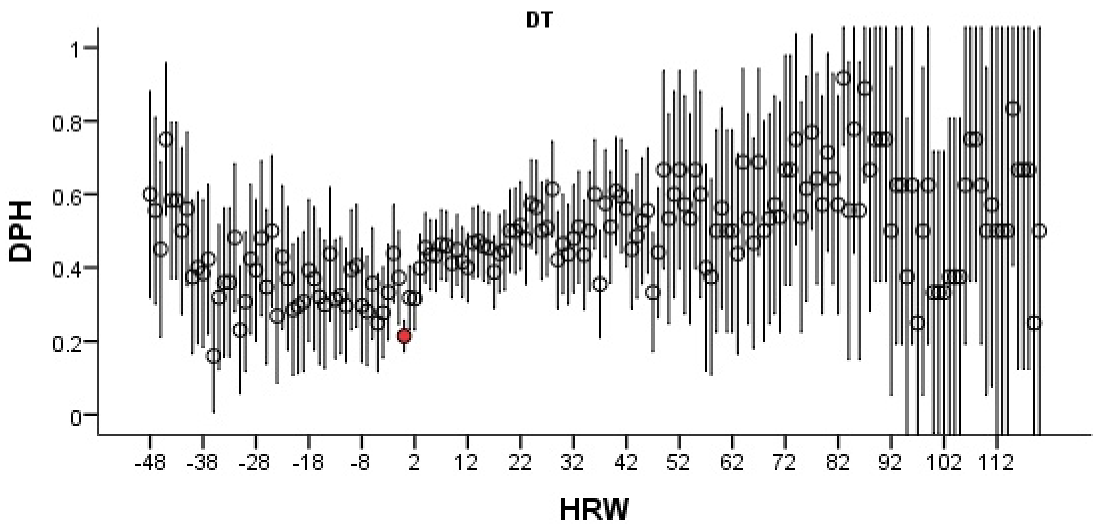


Figure A-7 Error bars depicting mean and 95% confidence intervals of DPH at the different hours relative to piling at distances below 5 km for the wind farm project DT. The error bar for hour relative to piling=0 is shown in red.

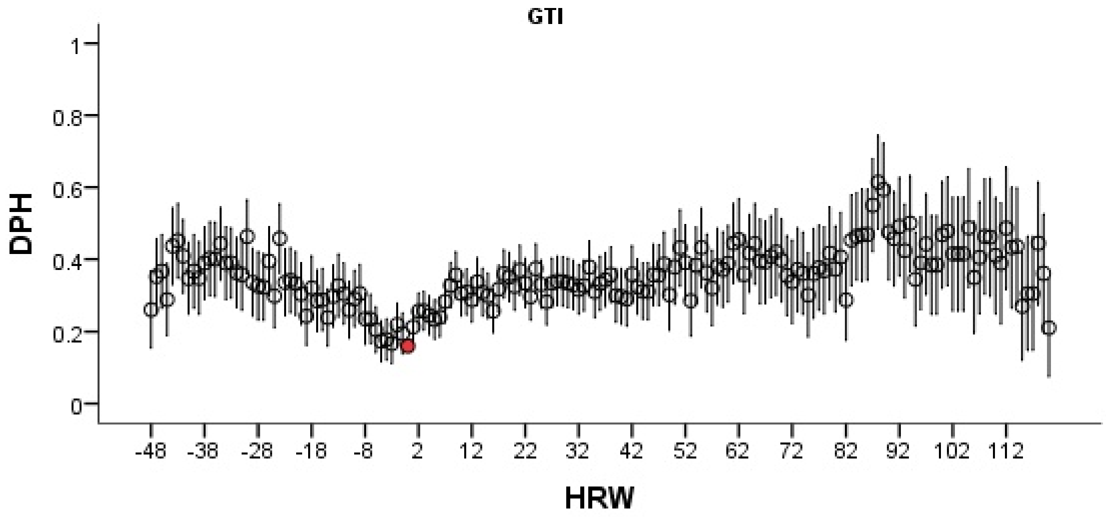


Figure A-8 *Error bars depicting mean and 95% confidence intervals of DPH at the different hours relative to piling at distances below 5 km for the wind farm project GTI. The error bar for hour relative to piling=0 is shown in red.*

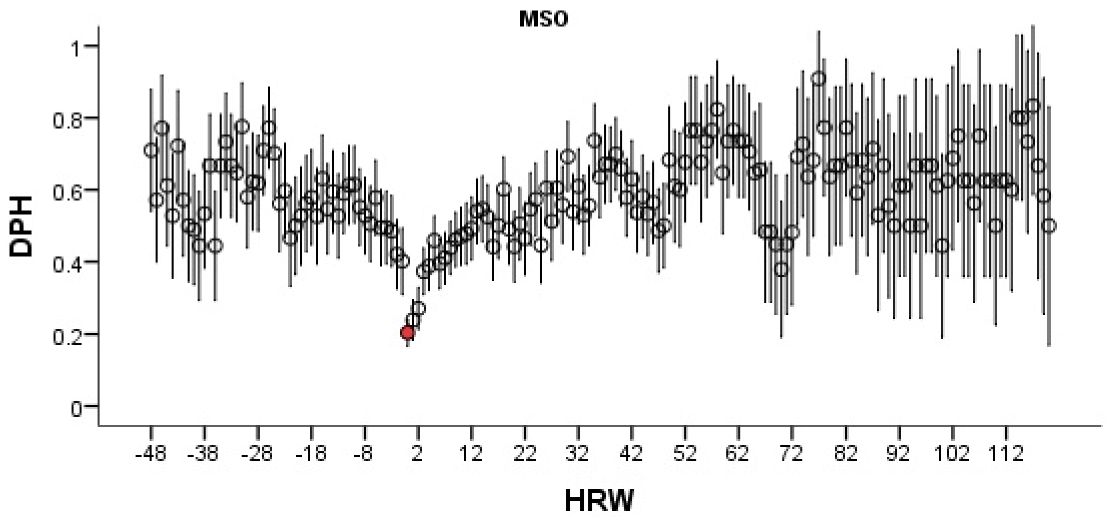


Figure A-9 *Error bars depicting mean and 95% confidence intervals of DPH at the different hours relative to piling work (HRW) at distances below 5 km for the wind farm project MSO. The error bar for hour relative to piling=0 is shown in red.*

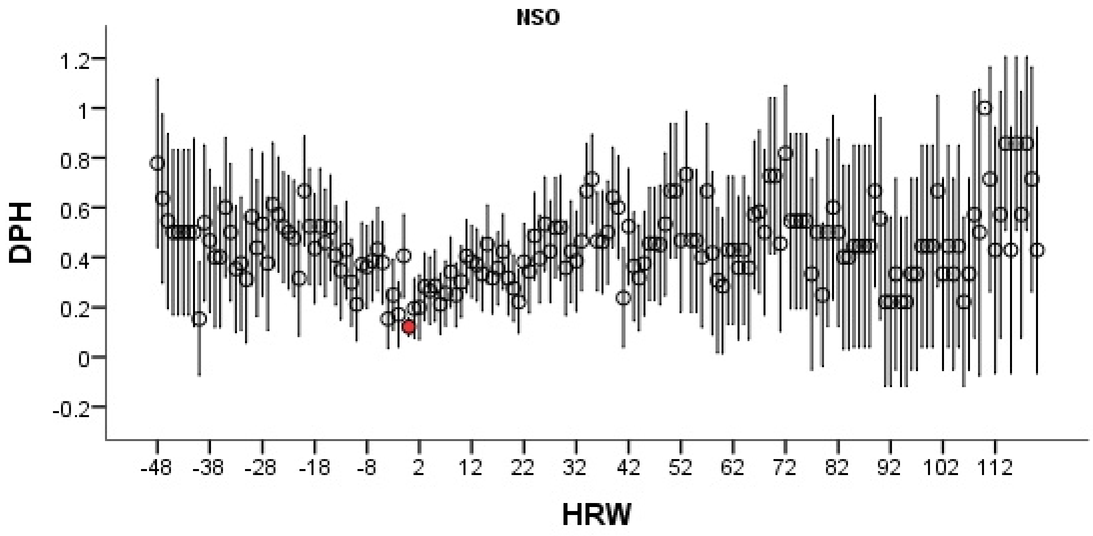


Figure A-10 Error bars depicting mean and 95% confidence intervals of DPH at the different hours relative to piling at distances below 5 km for the wind farm project NSO. The error bar for hour relative to piling=0 is shown in red.

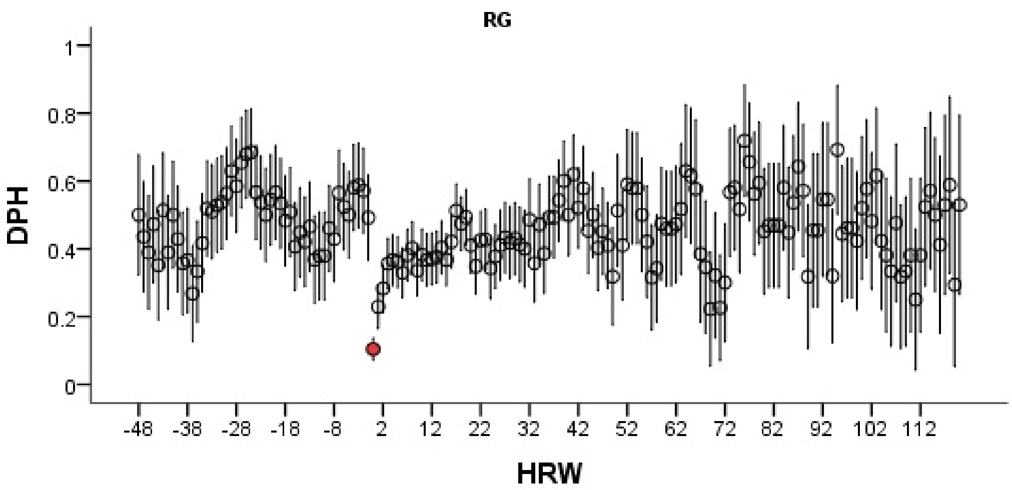


Figure A-11 Error bars depicting mean and 95% confidence intervals of DPH at the different hours relative to piling work (HRW) at distances below 5 km for the wind farm project RG. The error bar for hour relative to piling=0 is shown in red.

A.2.2 Additional tables and figures for chapter 4.3.6 (cumulative effects and habituation)

Table A-3 Results from models G5 run to check for the effects of piling duration (explained deviance: 15.2 %) and model G6 run to look at effects of piling order and min since last (explained deviance: 18.4 %).

model:	G5 (hour relative to piling =1, distance < 5 km)			G6 (hour relative to piling =1, distance < 5 km, piling order < 100, min since last piling <20,000 min)		
variable	(e)df	Chi²	p	(e)df	Chi²	p
DPH(t-1) (factor)	1	0.1	ns	1	4	ns
POD-Position (random factor)	<1	43	**	45	77	***
Year (factor)	3	8	ns	3	8	*
day of year (smooth)	6	43	***	7	76	***
HH (smooth)	<1	0	ns	2	6	*
wind speed (smooth)	1	15	***	6	60	***
wind direction (smooth)	<1	0	ns	1	1	ns
SSTA (smooth)	1	3	ns	5	16	*
noise clicks (smooth)	2	6	ns	2	10	*
sediment (factor)	4	4	ns	4	6	ns
distance (smooth)	1	15	***	1	84	***
pilingduration (smooth)	2	1	ns	4	15	**
piling order (smooth)	-	-	-	7	25	**
min since last piling (smooth)	-	-	-	6	22	**

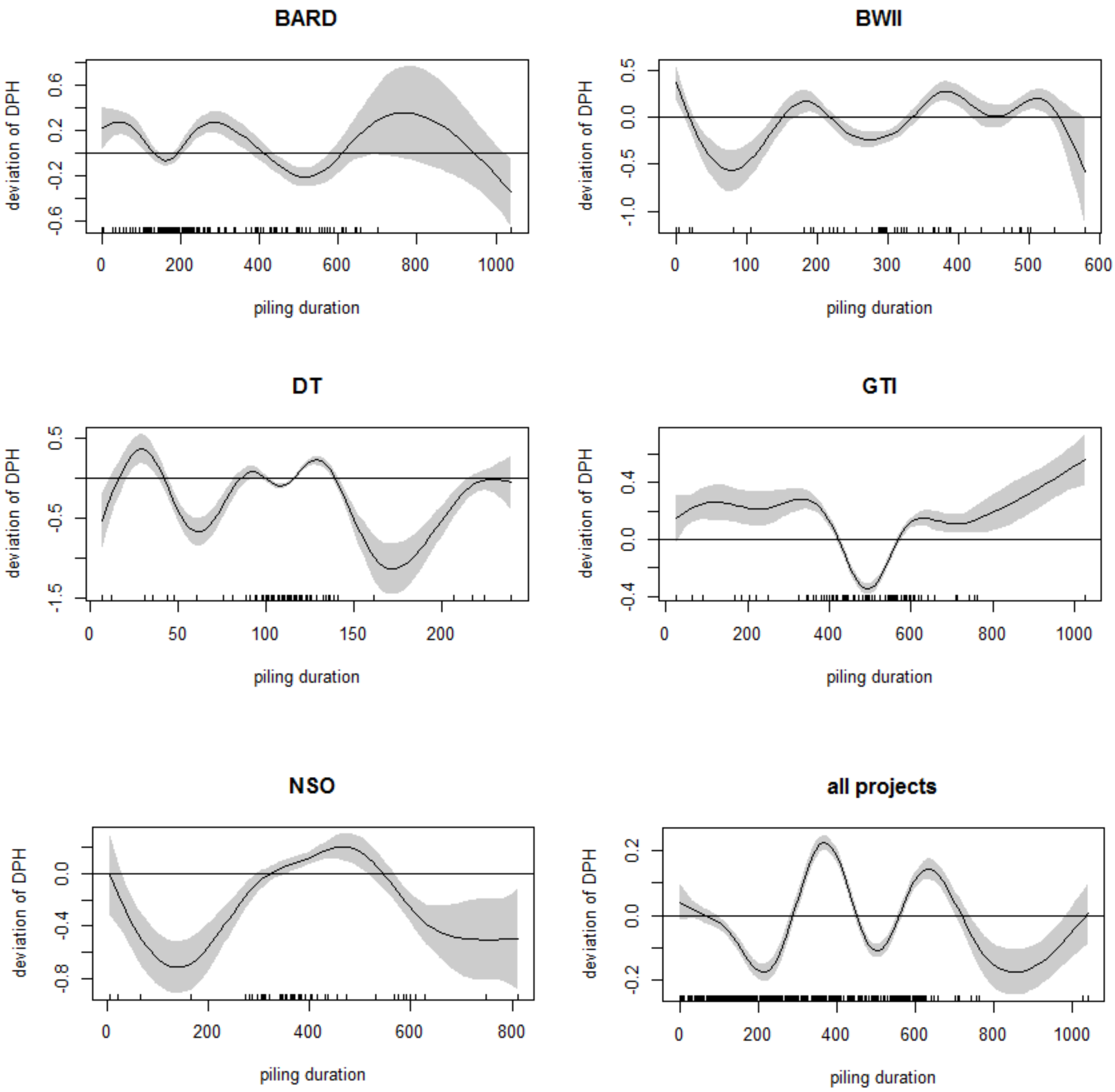


Figure A-12 Model output from Models P2 (first five figures) and G2 (lower right) showing the effects of piling duration on DPH. Shown is the predicted deviation from the overall mean including confidence intervals (grey shaded areas). Black tick marks indicate data availability.

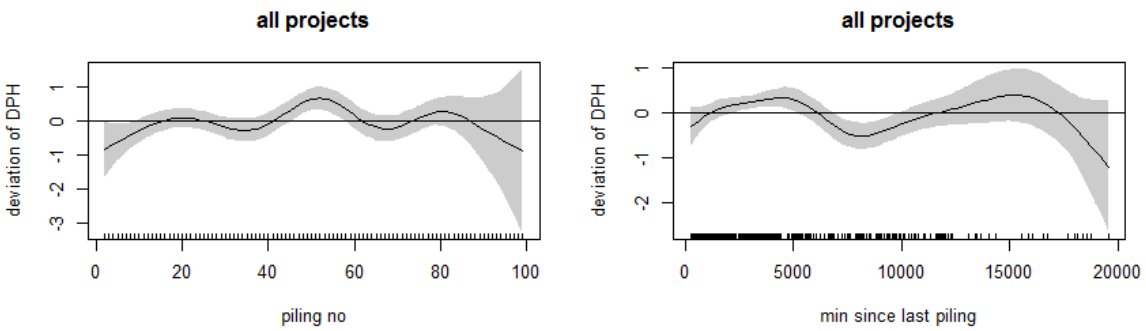


Figure A-13 Model output from Model G6 looking at habituation or temporal cumulative effects showing the effect of consecutive piling number (left) and minutes since the last piling event (right) on DPH. Shown is the predicted deviation of DPH from the overall mean including confidence intervals (grey shaded areas). Black tick marks indicate data availability.

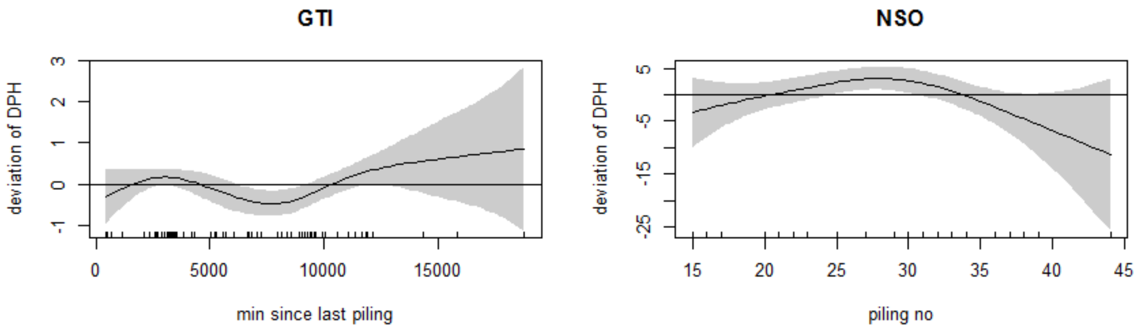


Figure A-14 Model output from models P6 looking at habituation or temporal cumulative effects illustrating the only significant effects found apart from the one shown in Figure 4-4 according to Table 4-11 and showing the effect of minutes since the last piling event (left) and consecutive piling number (right) on DPH. Shown is the predicted deviation of DPH from the overall mean including confidence intervals (grey shaded areas). Black tick marks indicate data availability.

A.2.3 Additional tables and figures for chapter 4.3.7 (wind speed)

Table A-4 Results from models G7 (explained deviance: 7.1 %) run to verify construction effects before piling.

model:	G5 (hour relative to piling < 0)		
variable	(e)df	Chi²	p
DPH(t-1) (factor)	1	43	***
POD-Position (random factor)	280	3359	***
Year (factor)	3	280	***
day of year (smooth)	8	2717	***
HH (smooth)	6	374	***
wind speed (smooth)	8	983	***
wind direction (smooth)	7	129	**
SSTA (smooth)	8	180	***
noise clicks (smooth)	1	453	***
sediment (factor)	4	29	***
pilingduration (smooth)	9	242	***
hour relative to piling, distance (interaction)	22	210	***

Table A-5 Results from models G8 and G9 run to look at effects of wind speed on construction effect ranges (explained deviance: 7.8 % (G8) and 13.8 % (G9)).

model:	G8 (hour relative to piling = -5)			G9 (hour relative to piling = 0)		
variable	(e)df	Chi²	p	(e)df	Chi²	p
DPH(t-1) (factor)	1	68	***	1	33	***
POD-Position (random factor)	292	4200	***	122	871	***
Year (factor)	3	341	***	3	194	***
day of year (smooth)	8	8	***	8	491	***
HH (smooth)	6	8	***	3	24	***
wind direction (smooth)	8	8	***	2	9	*
SSTA (smooth)	8	9	***	8	9	***
noise clicks (smooth)	2	527	***	1	60	***
sediment (factor)	4	34	***	4	45	***
distance, wind speed (interaction)	26	28	***	24	1115	***

A.2.4 Effects of environmental and time-related variables

We briefly describe here the effects of time-related and environmental variables based on model outputs from model G1 (shown in Figure A-15). The effects shown are calculated over the global dataset and do not differentiate between different geographical regions nor between different seasons. It has to be considered, that effects are highly likely to differ between geographic subareas and seasons, and that addressing this is a project in itself, but not the aim of the present study to address. The present study only included these variables into data modelling in order to control for their influences and to describe more precisely the effect of wind farm construction on porpoise detections.

Wind speed had a mainly positive effect on DPH. DPH almost linearly increase with increasing wind speed until reaching a maximum at about 14 m/s. At higher wind speeds, there is a negative relationship.

With respect to wind direction, DPH were the highest at easterly and south-easterly winds and the lowest with northerly winds.

The number of noise clicks negatively affected DPH. DPH almost linearly decrease with increasing noise clicks recorded by the POD.

DPH are lower with negative values for sea surface temperature anomaly, so fewer porpoise detections occur when sea surface temperature is below the generally prevailing temperature for this time and place, while at positive values DPH are increased.

DPH decrease from sediment category 1 (mainly coarse sand) to category 5 (finest sediment, i.e. mainly mud). DPH therefore seem to be the highest with the proportions of coarsest sediment.

DPH are decreased during mostly dark hours between 16:00 and 5:00 UTC, while during mostly daylight hours DPH seem relatively constant.

The yearly cycle in DPH seems characterised by a maximum in spring around March and a second peak during the summer around June/ July. There is a very clear pattern of differing porpoise detections within a year. DPH also differ greatly between years. Between 2010 and 2013, DPH were the highest in 2013 and the lowest in 2010. It has to be noted again here, however, that the whole dataset is chosen based on 2 days before and 4 days after piling events and therefore the geographic region covered within each season and year depends on where piling occurred. Therefore, differences between seasons and years should not be interpreted as real effects over the entire study area. They are partly an artefact of data selection aimed at studying short-term effects of wind farm construction.

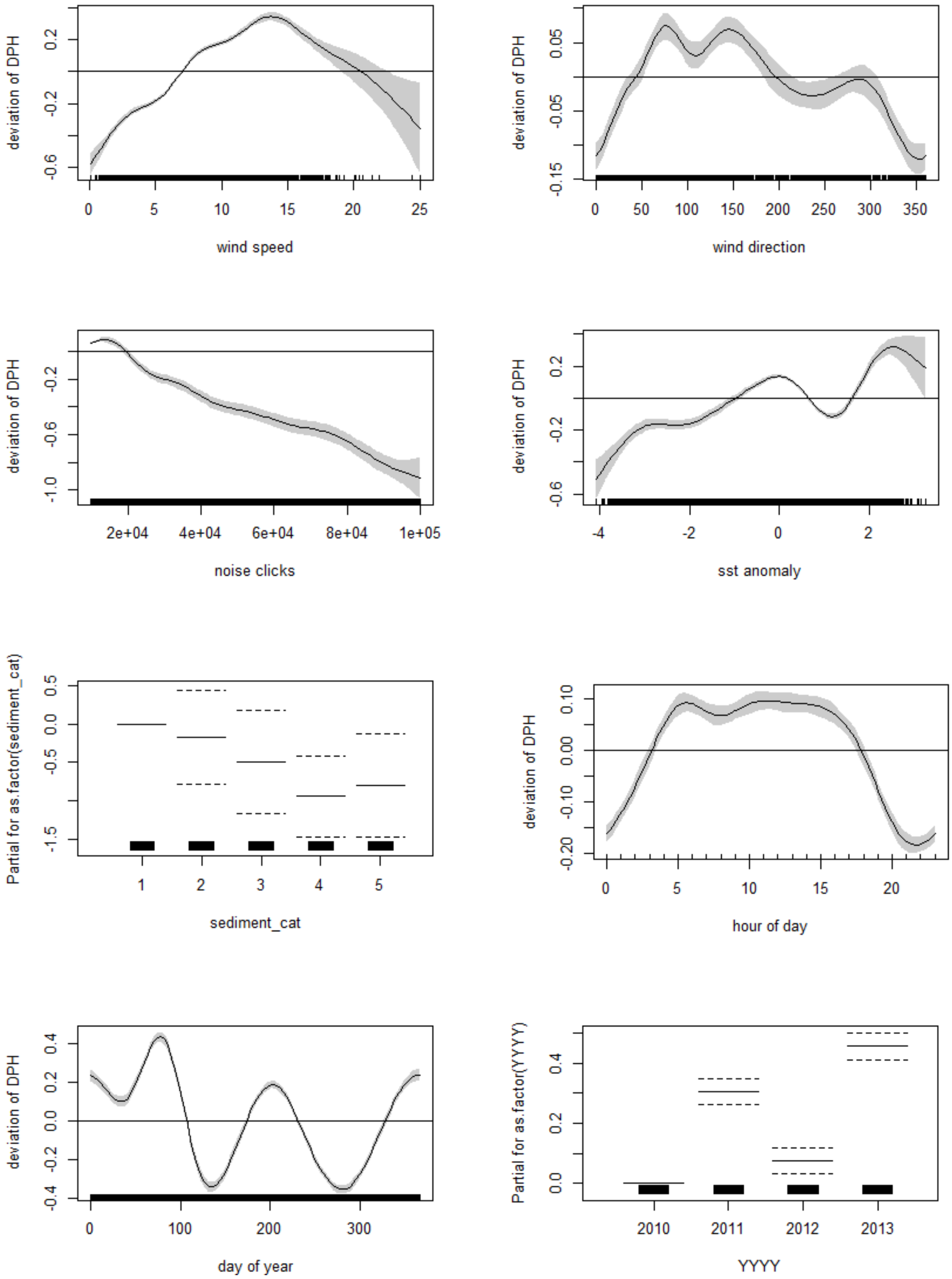


Figure A-15 *Model outputs from model G2 showing the effects of environmental and time-related variables on DPH. The predicted deviation of DPH from the overall mean are shown with 95% confidence intervals (grey shaded areas) or in the case of categorical variables (year and sediment) errorbars.*

A.3 Daily POD-data

A.3.1 Relation of all clicks to wind speed

A detailed analysis of water depth and number of all clicks showed a tendency to fewer clicks with increasing water depth. However, the relationship was not distinct (Chap. 5.3). Data from Riffgat had to be excluded from this analysis, since recordings in this area were so noisy that C-POD settings had to be changed to a low sensitivity setting. The noise levels are still very different. This could be the case as C-PODs are mounted in different distances to the sea surface and ambient noise caused by wind is also dependent on other factors, especially in shallow waters (INGENITO & WOLF 1989). Additionally, it has to be kept in mind that C-PODs only record noise in a specific frequency range with a click characteristic. A known factor creating those clicks is sand in suspension which mainly depends on the turbidity in the water that is created by currents and on the sediment in that region. Another noise source could be the movement of the anchorage of the C-POD. This would also be dependent on the wind speed.

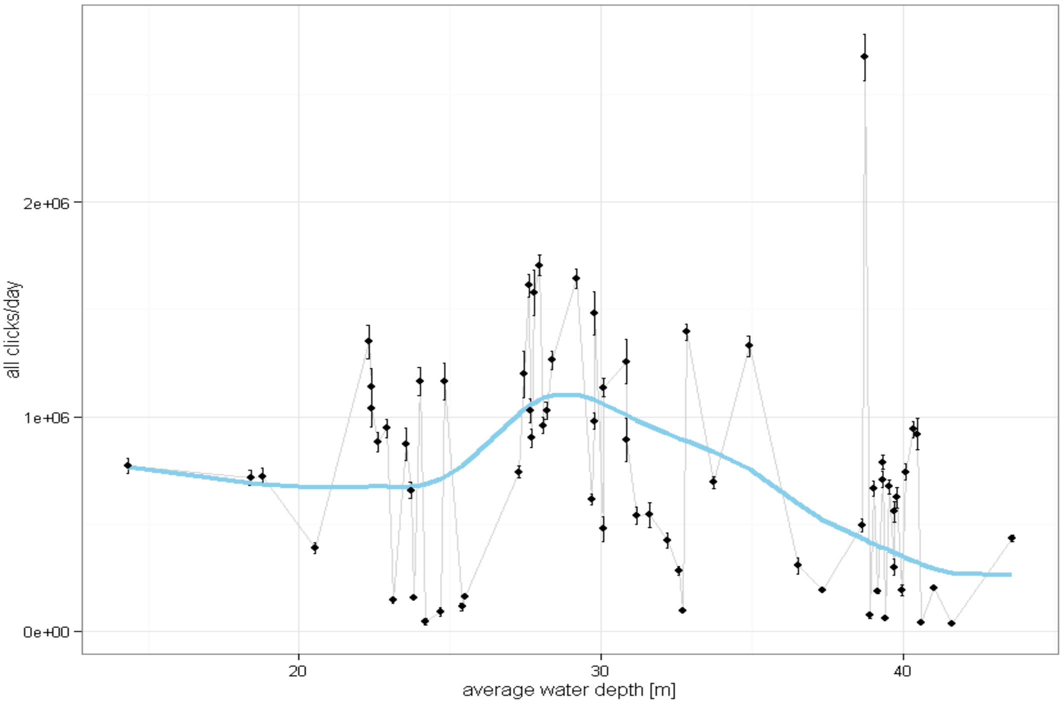


Figure A-16 Relationship between water depth and number of all click.

A.3.2 R packages

Four our data preparation and exploration, analyses and plotting of the results, we used the following R-packages: forecast (FORECAST 2015), timeDate (TIMEDATE 2015), zoo (ZOO 2015), lattice (LATTICE 2008), latticeExtra (LATTICEEXTRA 2013), rgdal (RGDAL 2015), rgeos (RGEOS 2015), RColorBrewer (RCOLORBREWER 2014), fossil (FOSSIL 2012), shapefiles (SHAPEFILES 2013), foreign (FOREIGN

2015), maps (MAPS 2014), sp (SP 2015), fpc (FPC 2015), cluster (MAECHLER et al. 2015), dsm (DSM 2015), mrds (MRDS 2015), mgcv (MGCV 2015), nlme (NLME 2015), tseries (TSERIES 2015). Most model outputs were plotted using a self-made plotting function based on the plotting function provided by the library mgcv (MGCV 2015).

A.3.3 Baseline model

*Table A.6 Model statistics of baseline model. Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 'n.s.' 1. Terms for whom no significance estimates are available are marked n.a. and terms not included in the model are marked with '-'.*

Name	Property	Baseline model
AIC	Goodness of fit	25,712.8
r-squared (adjusted)	Coefficient of determination	0.228
n (number of data records)	Sample size	13,208
dp10m	Response	-
all clicks	Smooth term	***
year by subarea	Factor	***
subarea	Factor	**
SSTA	Smooth term	***
random=list(station=~1, podident=~1)	Nested random factors	-
corARMA	Autocorrelation on population level	p=1, q=0

A.3.4 Detectable Effects of Pile Driving Activities on Acoustic Porpoise Detections

Yearly trends

*Table A.7 Model statistics of yearly trends model. Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 'n.s.' 1. Terms for whom no significance estimates are available are marked n.a. and terms not included in the model are marked with '-'.*

Name	Property	BWII-area	BARD-area	DanTysik-area	MSO-NSO-area
AIC	Goodness of fit	25,541.3	23,417.1	12,033.2	10,499.7
r-squared (adjusted)	Coefficient of determination	0.167	0.263	0.228	0.138
n (number of data records)	Sample size	12,950	11,911	6,525	5,450
dp10m	Response	n.a.	n.a.	n.a.	n.a.
day of year	Smooth term	***	***	***	***
all clicks	Smooth term	***	***	***	***
pilingYear	Factor	***	***	**	***
SSTA	Smooth term	***	***	***	**
random=list(station=~1, podident=~1)	Nested random factors	n.a.	n.a.	n.a.	n.a.
corARMA	Autocorrelation on population level	p=1, q=0	p=1, q=0	p=1, q=0	p=1, q=0

Context specific

*Table A.8 Model statistics for season models. Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 'n.s.' 1. Terms for whom no significance estimates are available are marked n.a. and terms not included in the model are marked with '-'.*

Name	Property	Winter	Spring	Summer	Autumn
AIC	Goodness of fit	1559.0	3211.8	3352.3	4628.6
r-squared (adjusted)	Coefficient of determination	0.237	0.400	0.349	0.281
n (number of data records)	Sample size	858	1,696	1,910	2,411
dp10m	Response	n.a.	n.a.	n.a.	n.a.
pileDrivingDay	Factor	***	***	***	***
minDist by subarea	Smooth term	***	***	***	***
all clicks	Smooth term	***	***	***	***
pilingDuration by subarea	Smooth term	***	***	-	***
pilingDuration	Smooth term	-	-	ns	-
version	Factor	-	ns	***	***
year	Factor	**	***	ns	-
subarea	Factor	ns	***	ns	***
SSTA	Smooth term	***	***	***	*
random=list(station=~1)	Random factor	n.a.	n.a.	n.a.	n.a.
corARMA	Autocorrelation	p=0, q=1	p=1, q=0	p=1, q=1	p=1, q=1

*Table A.9 Model statistics for subarea models. Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 'n.s.' 1. Terms for whom no significance estimates are available are marked n.a. and terms not included in the model are marked with '-'.*

Name	Property	BWII-area	BARD-area	DanTysk-area	MSO-NSO-area
AIC	Goodness of fit	2,842.996	5,356.871	3,056.264	1,533.455
r-squared (adjusted)	Coefficient of determination	0.247	0.282	0.329	0.157
n (number of data records)	Sample size	1,681	2,573	1,711	910
dp10m	Response	n.a.	n.a.	n.a.	n.a.
pileDrivingDay	Factor	***	***	***	***
minDist by season	Smooth term	***	***	***	***
all clicks	Smooth term	**	***	***	***
pilingDuration by season	Smooth term	***	***	-	-
pilingDuration	Smooth term	-	-	n.s.	**
version	Factor	-	.	-	-
year	Factor	-	**	-	-
season	Factor	**	***	***	*
SSTA	Smooth term	**	***	***	***
random=list(station=~1)	Random factor	n.a.	n.a.	n.a.	n.a.
corARMA	Autocorrelation on piling level	p=1,q=0	p=1,q=0	p=1,q=1	p=1,q=0

*Table A.10 Model statistics for season and subarea combined on the day of piling. Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 'n.s.' 1. Terms for whom no significance estimates are available are marked n.a. and terms not included in the model are marked with '-'.*

Name	Property	Model_Season-Subarea
AIC	Goodness of fit	7,747.552
r-squared (adjusted)	Coefficient of determination	0.263
n (number of data records)	Sample size	3,784
dp10m	Response	n.a.
clusterSeason	Factor	***
minDist	Smooth term	***
all clicks	Smooth term	***
pilingDuration	Smooth term	***
SSTA	Smooth term	n.s.
random=list(station=~1,podident=~1)	Nested random factors	n.a.

Accumulation

Consecutive piling events

*Table A.11 Model statistics for consecutive pile driving impact model. Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 'n.s.' 1. Terms for whom no significance estimates are available are marked n.a. and terms not included in the model are marked with '-'.*

Name	Property	Model for consecutive piling events
AIC	Goodness of fit	5,630.9
r-squared (adjusted)	Coefficient of determination	0.205
n (number of data records)	Sample size	2,622
dp10m	Response	n.a.
pilingDuration	Smooth term	***
HourOfPilingStart	Smooth term	***
consesc	Factor	ns
project	Factor	***
minDist	Smooth term	***
all clicks	Smooth term	***

Simultaneous piling events

*Table A.12 Model statistics for multiple piling events per day model. Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 'n.s.' 1. Terms for whom no significance estimates are available are marked n.a. and terms not included in the model are marked with '-'.*

Name	Property	Model for simultaneous piling events
AIC	Goodness of fit	1,389.1
r-squared (adjusted)	Coefficient of determination	0.191
n (number of data records)	Sample size	569
dp10m	Response	n.a.
pilingDuration by cumuCat	Smooth term	**
HourOfPilingStart by cumuCat	Smooth term	ns
cumuCat	Factor	**
project	Factor	**
minDist by cumuCat	Smooth term	***
all clicks	Smooth term	***
SSTA	Smooth term	***
year	Smooth term	*
random=list(station=~1)	Random factor	n.a.

A.4 Interim PCoD-model - general structure

Parameters based on expert judgment as given by Harwood et al. (2014)

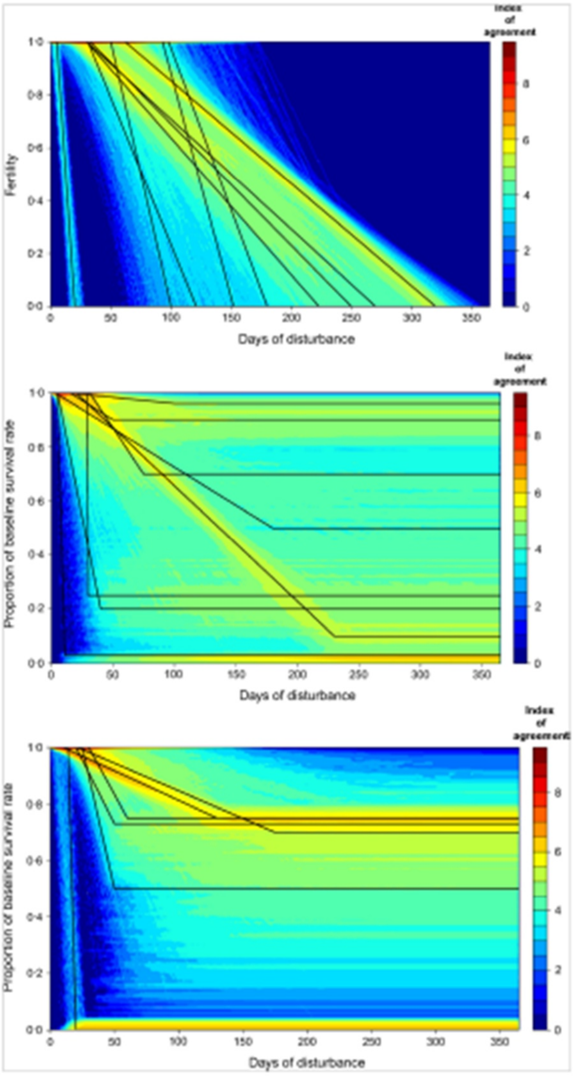


Figure A-17 The degree of support among experts for different relationships linking days of disturbance to (a) fertility, (b) calf survival and (c) juvenile survival. Individual experts' suggestions (black lines) are superimposed on heat maps showing overall support among experts for particular combinations of values. Reds and yellows indicate well-supported combinations that are more likely to be sampled in simulations. Shades of blue indicate poorly supported combinations (Fig. 3 taken from KING *et al. (2015).*

A.4.1 Results: Sensitivity of the interim PCoD Model to selected parameters

Vulnerable Subpopulations

Figure A-18 illustrates how the effects of defining the vulnerable subpopulation alter the relationship between disturbed and population consequences. Generally, the more numerous the individuals are that are disturbed, the stronger the global population decline. With a smaller vulnerable subpopulation a maximum effect is reached faster, but is limited compared to when the complete population is vulnerable.

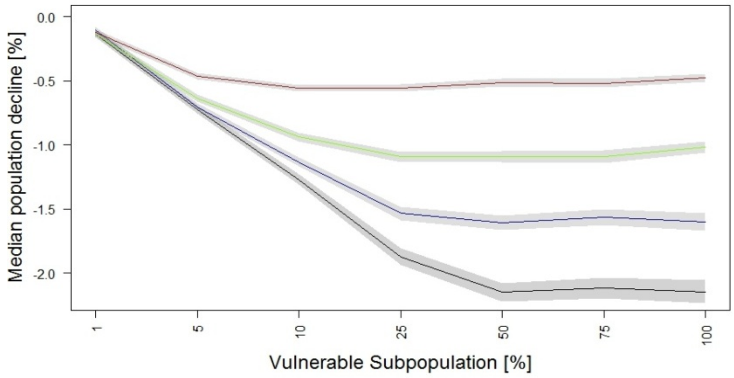


Figure A-18 Effect of including vulnerable subpopulations on median population decline depending on the number of individuals experiencing disturbance and PTS by pile driving per day. Depicted are the results of 31 days of piling with: 190 individuals disturbed and 10 with PTS per day in red, 380 individuals disturbed and 20 with PTS per day in green, 570 individuals disturbed and 30 with PTS per day in blue and 760 individuals disturbed and 40 with PTS per day in black.

There is no cumulative effect in the interim PCoD model if populations are affected by more than one construction site (Appendix, Figure A-19)

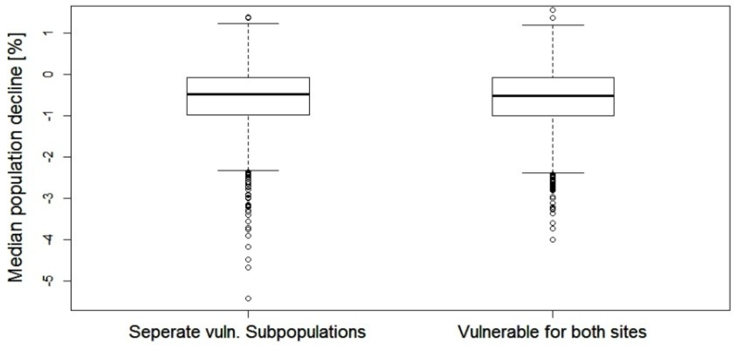


Figure A-19 Effect of the usage of vulnerable subpopulations on median population decline for two construction sites with each 31 days of piling. With 95 individuals being disturbed and 5 experiencing PTS per piling day and two vulnerable subpopulations (each 45%) only vulnerable to one of these (left site) or one combined vulnerable subpopulation (90%) vulnerable subpopulation to both sites (right site).

Temporal aspects

The importance of the total amount of disturbance but not of temporal aspects is underscored if comparing PCoD model results including seasonal differences (Figure A-20).

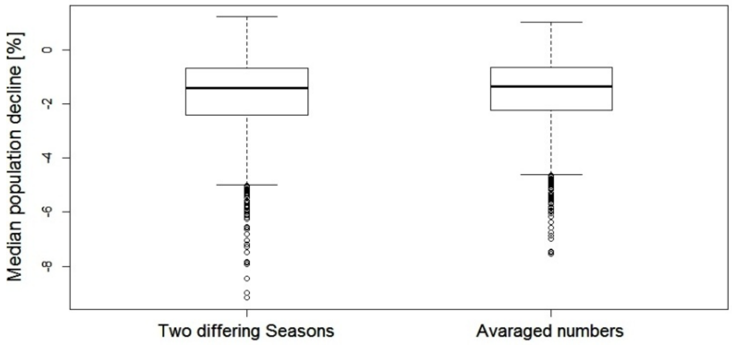


Figure A-20 Median population decline depending on seasonal differences in affected individuals. Depicted are the results for 82 days of piling. Either with 41 piling days with 285 disturbed individuals and 15 with PTS and 41 piling days with 95 disturbed individuals and 5 with PTS (left side) or 82 piling days with an average of 190 disturbed individuals and 10 with PTS (right side).

The degree of temporal overlap of piling activities between two operations does not influence the total effect on the population modelled by the interim PCoD model either. In other words, simultaneous work at different sites has no cumulative effect (Figure A-21).

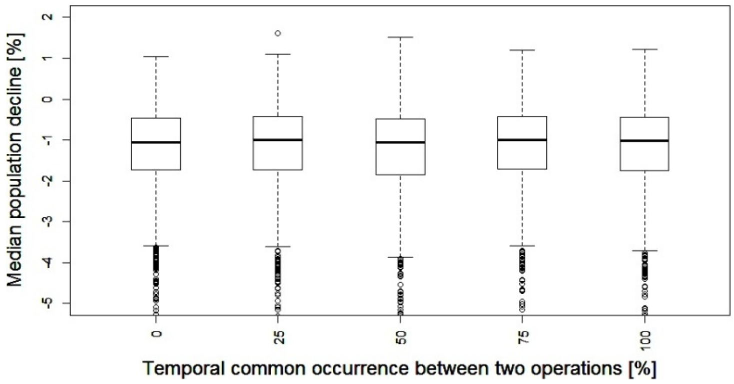


Figure A-21 Median population decline depending on the concurrent activity of two construction sites with 31 piling days each, resulting in 190 disturbed individuals and 10 with PTS per day.

Piling schedules should also specify whether construction activities will have regular breaks or whether they are (rather) continuous (HARWOOD et al. 2014). However, there was no effect to be seen in our comparison of continuous piling versus regular breaks (Figure A-22).

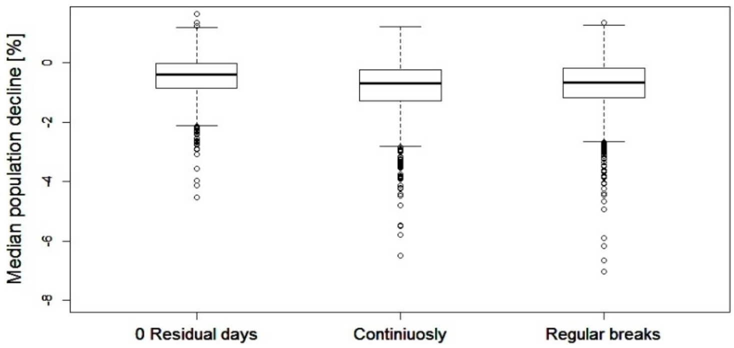


Figure A-22 Median population decline depending on residual days of disturbance and either 81 continuous construction days or 81 construction days with breaks after each day as modelled by the interim PCoD. The left one is modelled with 0 residual days of disturbance and the other with each 1 residual days of disturbance. In each case 95 individuals are disturbed and 5 additional individuals experience PTS per day of construction.

A.6 Aerial Survey data predictors

A.6.1 Spatial and temporal distribution of grid cell densities from aerial surveys relative to piling events

The spatial temporal coverage of the data is not enough to compare all wind farm projects. The best coverage is found at BARD, DT, GT1 and NSO.

At the wind farm BARD, densities differed between the construction phases in spring and autumn (Figure A-24). Before and after construction, two and three flights, respectively, were realised in these seasons. In spring (Kruskal test: $p < 0.001$, $Chi^2 = 15.976$, $df = 3$), densities were lower before the construction phase than during the piling (Nemenyi test: $p = 0.044$). However, it has to be considered that these estimates were based on the data from only one flight and might have occurred by chance. No differences were shown for the other periods. Density estimates in autumn originated only from the construction phase. The density estimates during piling were lower than those being not related to piling events (Kruskal test: $p < 0.001$, $Chi^2 = 12.60$, $df = 1$).

At the wind farm DT, summer densities related to piling events were higher than in periods with no piling (Figure A-26; Kruskal test: $p = 0.028$, $Chi^2 = 4.837$, $df = 1$). However, there were no values before and after the construction phase. Estimates from spring, autumn and winter were not comparable, or covered only once during the construction phases of the wind farm.

At the wind farm GTI, densities did not exhibit any significant differences during the construction phases, during any season (Figure A-27). Seasonal boxplots showed higher densities in spring and autumn after piling. However, these results came from seven grid cells originating from only one flight each and, in comparison, the number of grid cells and realised surveys was much higher during the flights associated with pilings.

At the wind farm NSO, flights were realised in all four seasons; however, no flight data were sampled before or after the construction, making comparisons to time periods without affected periods impossible. Nevertheless, flights were realised in spring, summer and autumn and they were not associated to any piling event (Figure A-29). Estimates in summer showed significant differences between piling and no piling flights (Kruskal test: $p = 0.0219$, $chi = 5.254$, $df = 1$).

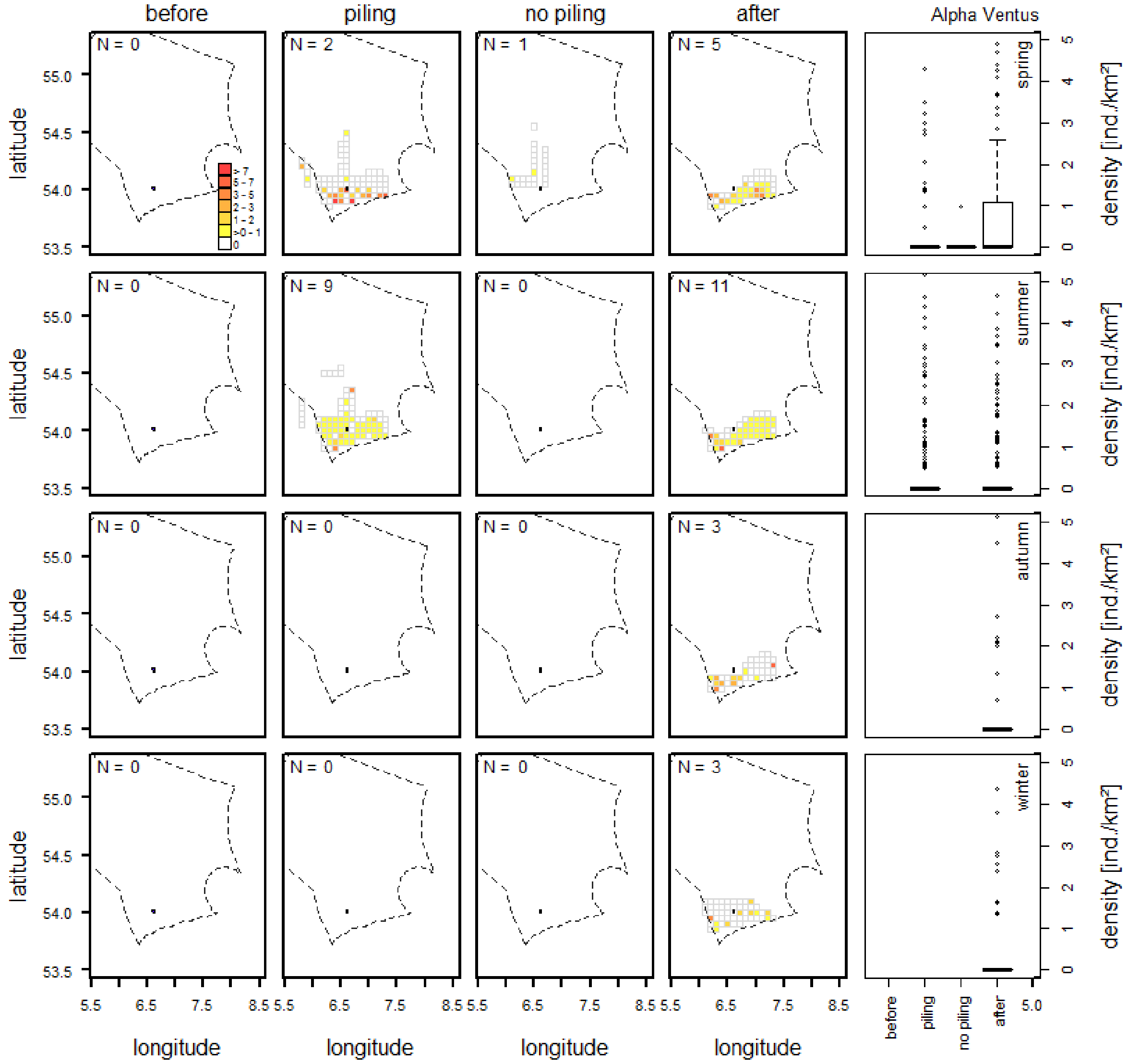


Figure A-23 *Seasonal mean density estimates for harbour porpoises in the vicinity of the wind farm AV before, during and after the construction of the wind farm (before: flight was conducted before the first piling activity was realised and piling occurred at no nearby place; the period during the construction phase was divided in flights with piling and no piling activity; piling: piling ended within a 60 km radius and at least 60 hours before the flight; no piling: flight occurred at least 60 hours after the last piling and any other piling event at another wind farm occurred at a distance of 60 km; after: flight occurred after the end of any piling activity in the wind farm and piling occurred at no nearby place; N: number flights used per map).*

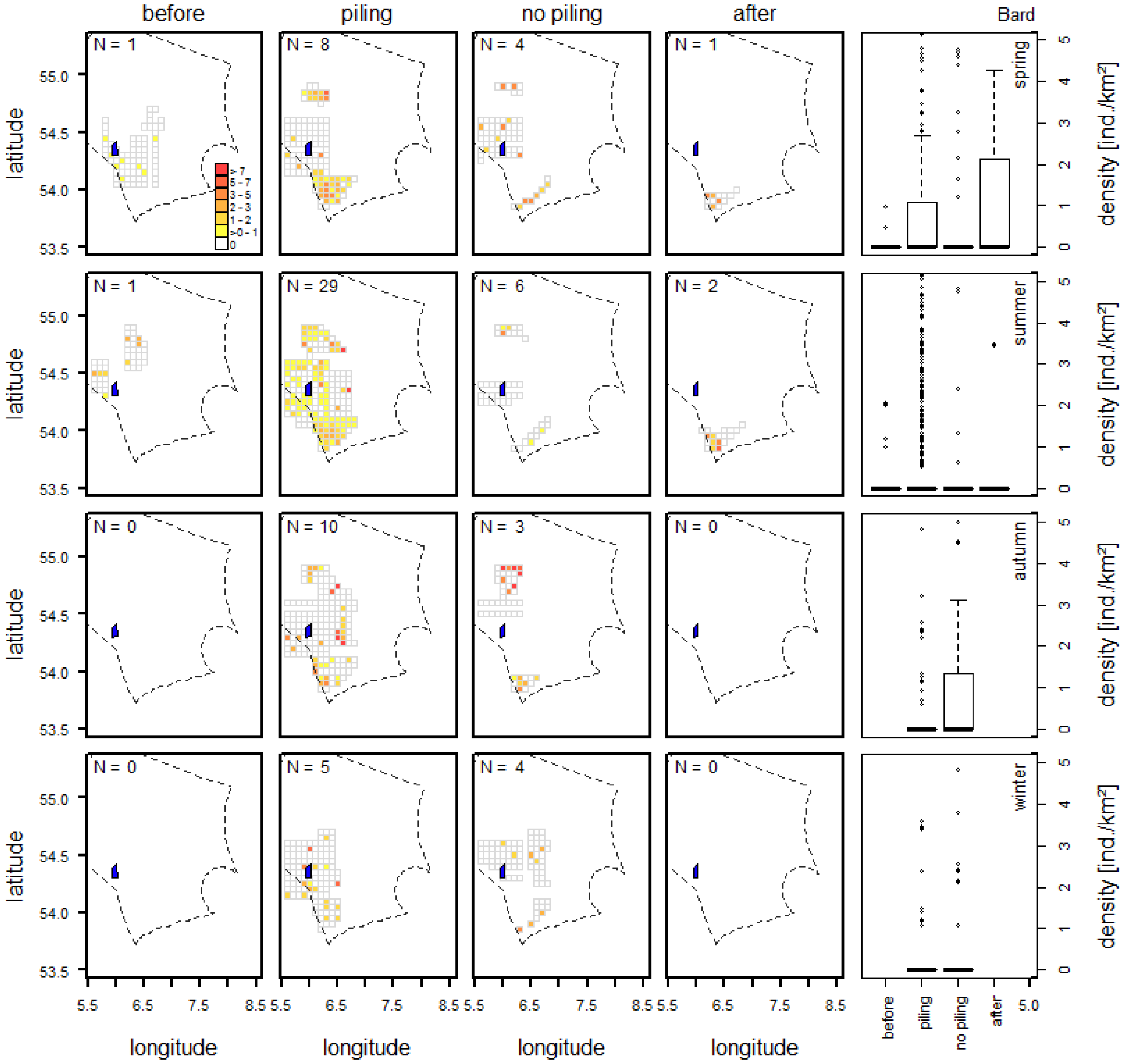


Figure A-24 Seasonal mean density estimates for harbour porpoises in the vicinity of the wind farm BARD before, during and after the construction of the wind farm (before: flight was conducted before the first piling activity was realised and piling occurred at no nearby place; the period during the construction phase was divided in flights with piling and no piling activity; piling: piling ended within a 60 km radius and at least 60 hours before the flight; no piling: flight occurred at least 60 hours after the last piling and any other piling event at another wind farm occurred at a distance of 60 km; after: flight occurred after the end of any piling activity in the wind farm and piling occurred at no nearby place; N: number flights used per map).

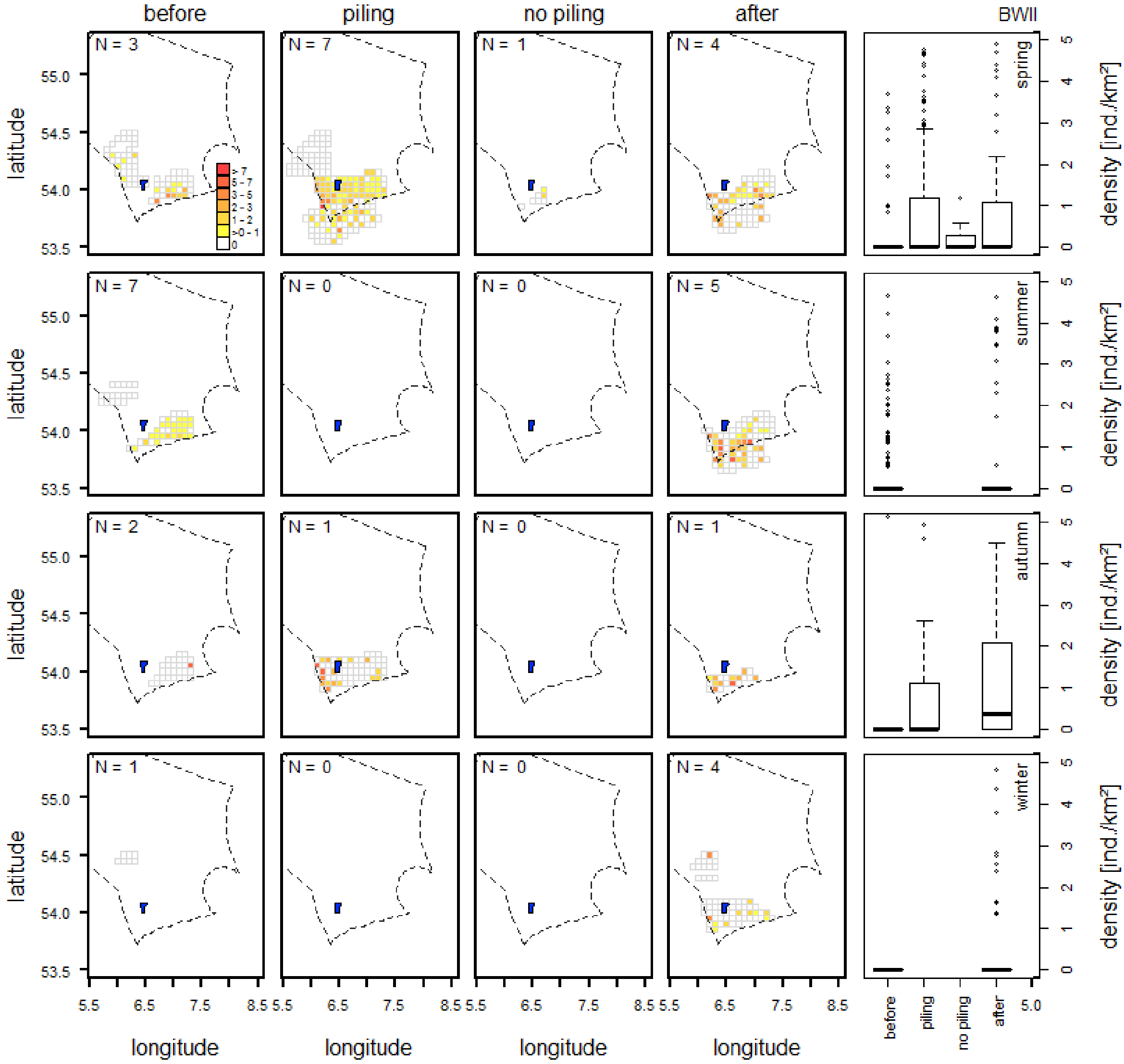


Figure A-25 Seasonal mean density estimates for harbour porpoises in the vicinity of the wind farm BWII before, during and after the construction of the wind farm (before: flight was conducted before the first piling activity was realised and piling occurred at no nearby place; the period during the construction phase was divided in flights with piling and no piling activity; piling: piling ended within a 60 km radius and at least 60 hours before the flight; no piling: flight occurred at least 60 hours after the last piling and any other piling event at another wind farm occurred at a distance of 60 km; after: flight occurred after the end of any piling activity in the wind farm and piling occurred at no nearby place; N: number flights used per map).

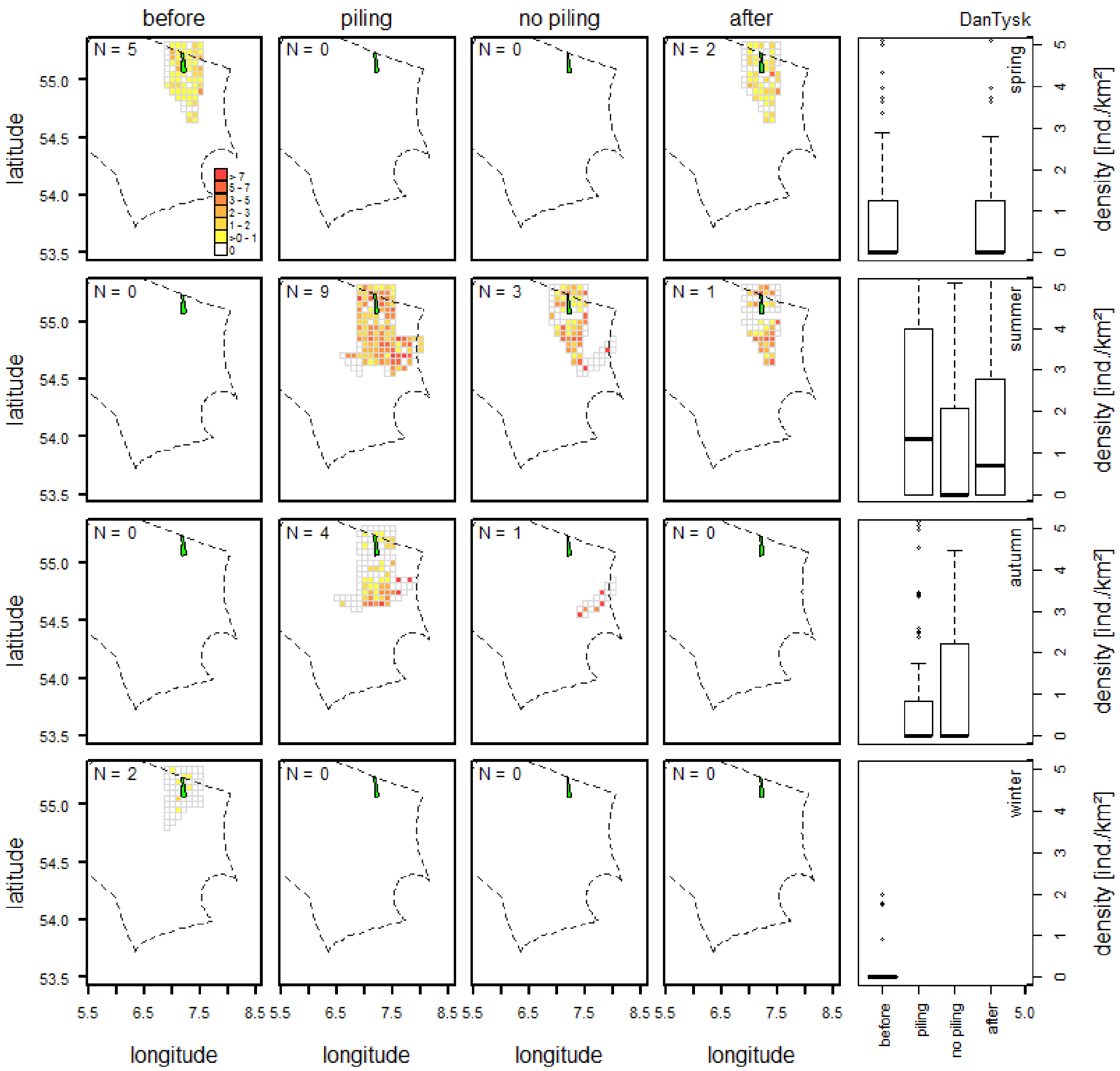


Figure A-26 *Seasonal mean density estimates for harbour porpoises in the vicinity of the wind farm DT before, during and after the construction of the wind farm (before: flight was conducted before the first piling activity was realised and piling occurred at no nearby place; the period during the construction phase was divided in flights with piling and no piling activity; piling: piling ended within a 60 km radius and at least 60 hours before the flight; no piling: flight occurred at least 60 hours after the last piling and any other piling event at another wind farm occurred at a distance of 60 km; after: flight occurred after the end of any piling activity in the wind farm and piling occurred at no nearby place; N: number flights used per map).*

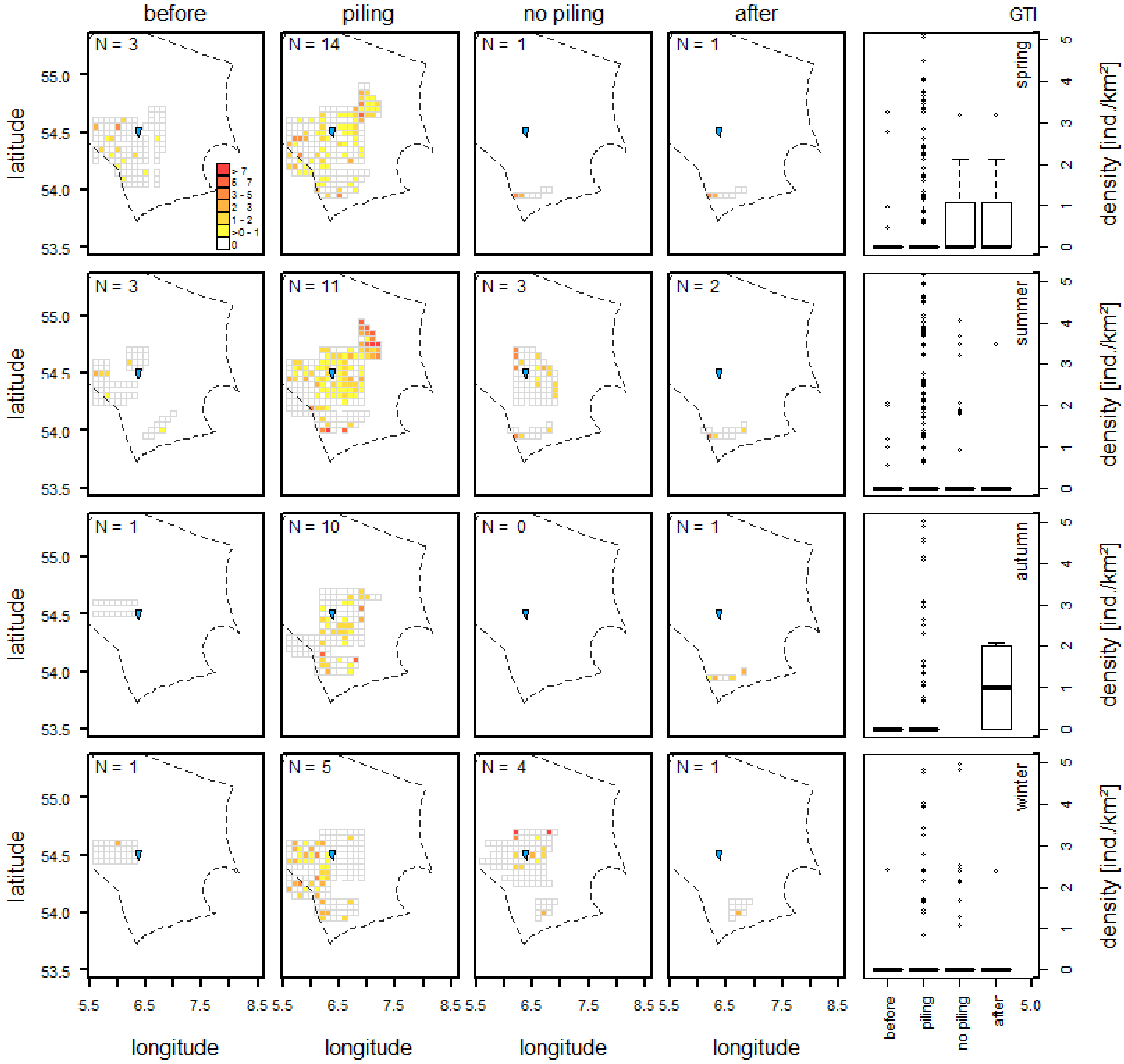


Figure A-27 *Seasonal mean density estimates for harbour porpoises in the vicinity of the wind farm GTI before, during and after the construction of the wind farm (before: flight was conducted before the first piling activity was realised and piling occurred at no nearby place; the period during the construction phase was divided in flights with piling and no piling activity; piling: piling ended within a 60 km radius and at least 60 hours before the flight; no piling: flight occurred at least 60 hours after the last piling and any other piling event at another wind farm occurred at a distance of 60 km; after: flight occurred after the end of any piling activity in the wind farm and piling occurred at no nearby place; N: number flights used per map).*

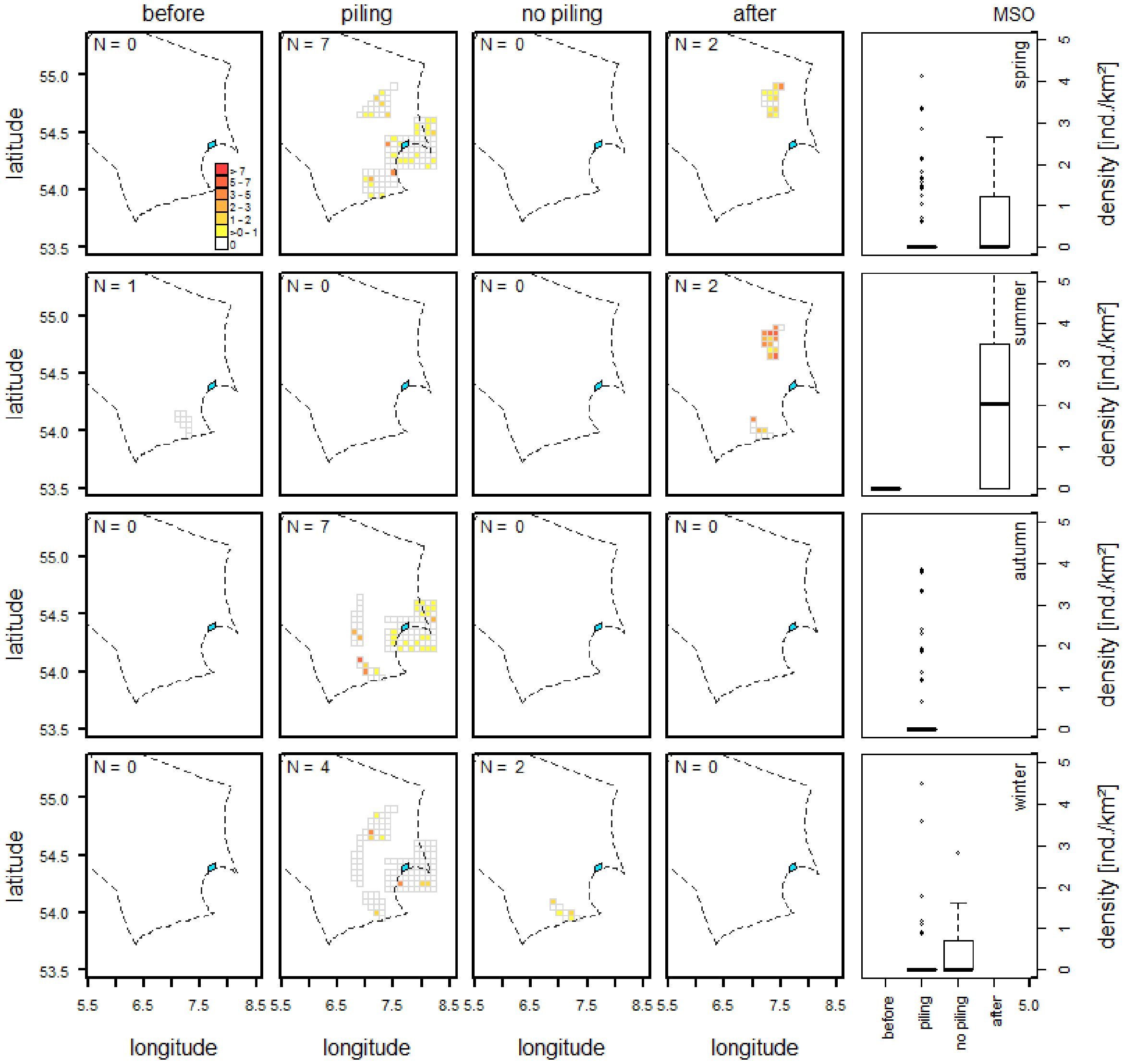


Figure A-28 Seasonal mean density estimates for harbour porpoises in the vicinity of the wind farm MSO before, during and after the construction of the wind farm (before: flight was conducted before the first piling activity was realised and piling occurred at no nearby place; the period during the construction phase was divided in flights with piling and no piling activity; piling: piling ended within a 60 km radius and at least 60 hours before the flight; no piling: flight occurred at least 60 hours after the last piling and any other piling event at another wind farm occurred at a distance of 60 km; after: flight occurred after the end of any piling activity in the wind farm and piling occurred at no nearby place; N: number flights used per map).

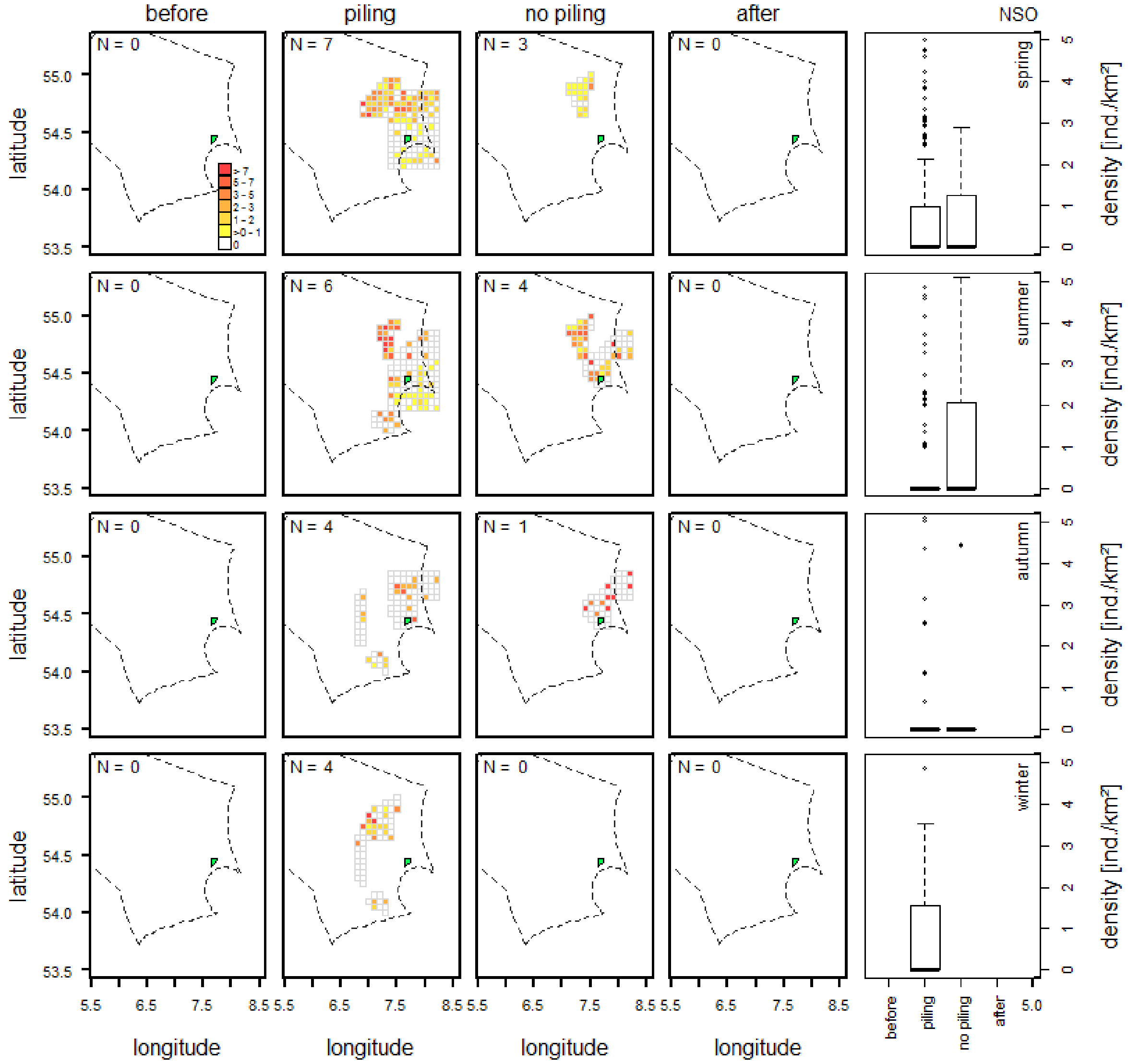


Figure A-29 *Seasonal mean density estimates for harbour porpoises in vicinity of the wind farm NSO before, during and after the construction of the wind farm (before: flight was conducted before the first piling activity was realised and piling occurred at no nearby place; the period during the construction phase was divided in flights with piling and no piling activity; piling: piling ended within a 60 km radius and at least 60 hours before the flight; no piling: flight occurred at least 60 hours after the last piling and any other piling event at another wind farm occurred at a distance of 60 km; after: flight occurred after the end of any piling activity in the wind farm and piling occurred at no nearby place; N: number flights used per map).*

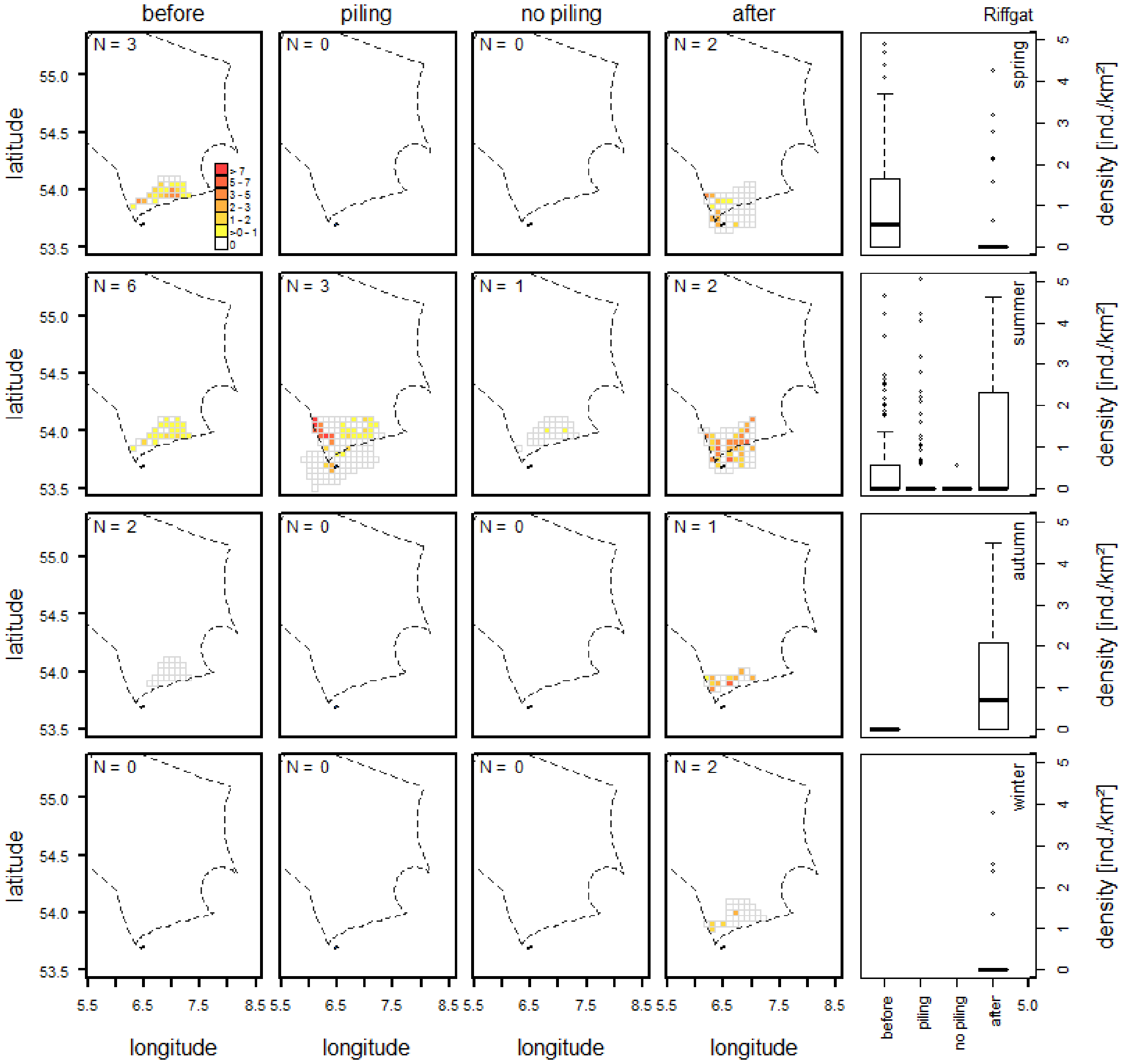


Figure A-30 Seasonal mean density estimates for harbour porpoises in the vicinity of the wind farm RG before, during and after the construction of the wind farm (before: flight was conducted before the first piling activity was realised and piling occurred at no nearby place; the period during the construction phase was divided in flights with piling and no piling activity; piling: piling ended within a 60 km radius and at least 60 hours before the flight; no piling: flight occurred at least 60 hours after the last piling and any other piling event at another wind farm occurred at a distance of 60 km; after: flight occurred after the end of any piling activity in the wind farm and piling occurred at no nearby place; N: number flights used per map).

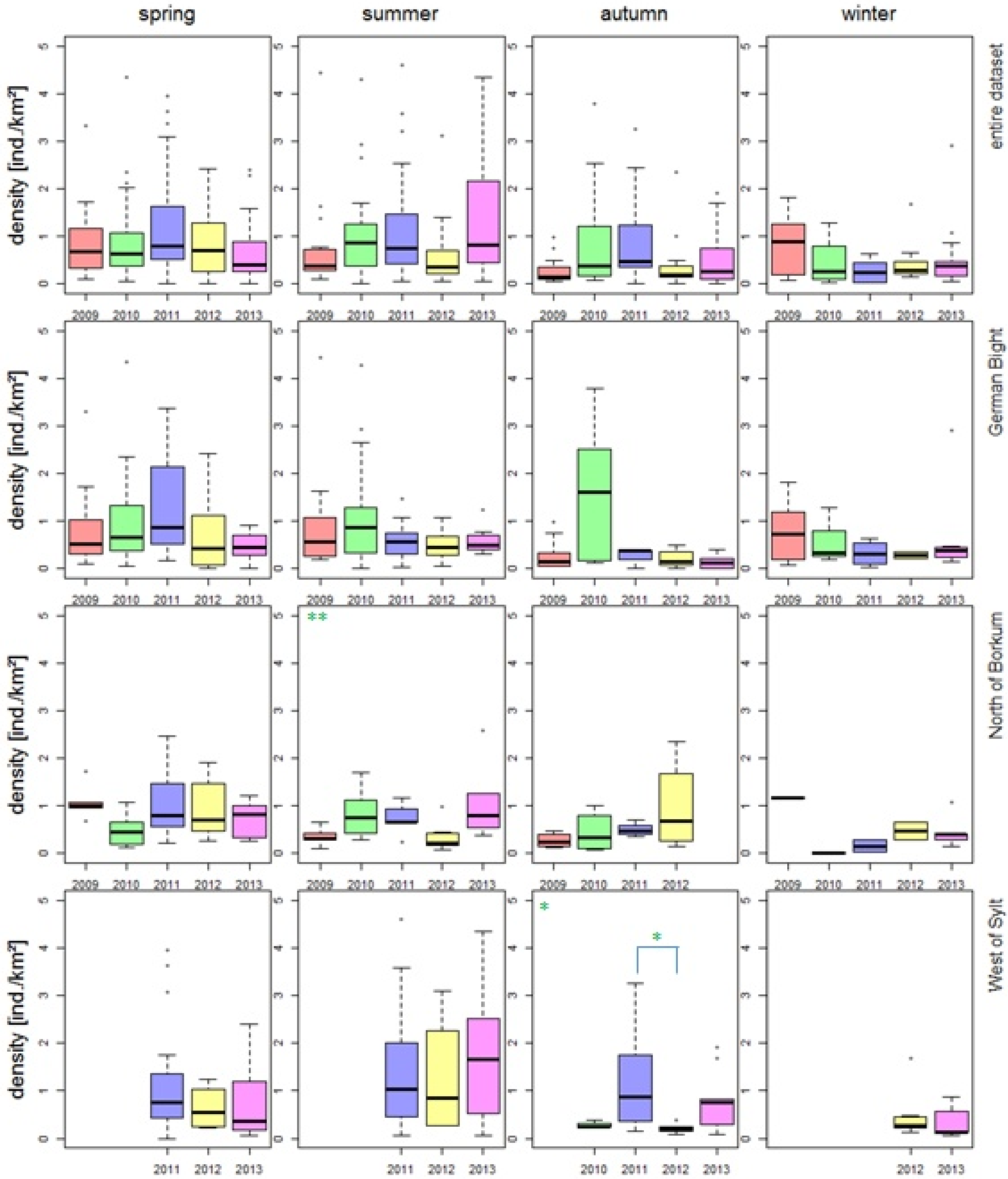


*Figure A-31 Density estimate of harbour porpoises using the entire dataset of spring and summer per sub-area (significance values for Kruskal and Nemenyi tests; **: p < 0.01;*: p < 0.05).*

A.6.2 GAM-plots not listed in the main text

Results from general models

Summaries of the general model outputs are presented in Table A-13 for the entire dataset, Table A-14 for the unaffected dataset and Table A-15 for the affected dataset. Table A-16, Table A-17,

Table A-18 and Table A-19 present the model outputs for specific subareas.

Table A-13 Results model A1 - GAM analysing annual trends and spatial distribution patterns of harbour porpoises using the entire dataset (n = 32,054, AIC = 66307.12, dev. Explained = 12.6 %).

variable	regression technique	category	edf	Chi²	p-value
Position ID	mrf smooth	factor	328.32	884.65	<0.001
day of year 2009	cyclic smooth	numeric (1-365)	7.478	181.04	<0.001
day of year 2010	cyclic smooth	numeric (1-365)	4.976	85.15	<0.001
day of year 2011	cyclic smooth	numeric (1-365)	6.175	189.78	<0.001
day of year 2012	cyclic smooth	numeric (1-365)	7.604	115.87	<0.001
day of year 2013	cyclic smooth	numeric (1-365)	7.419	204.12	<0.001
year	factor	numeric (2009-2013)	-	66.75	<0.001
moon illumination	thin plate smooth	numeric (0.0-1.0)	3.528	16.94	<0.001
flight time	thin plate smooth	numeric (0.0-1.0)	4.183	69.74	<0.001
latitude : longitude : spring	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	6.81	16.01	0.0316
latitude : longitude : summer	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	14.569	51.05	<0.001
latitude : longitude : autumn	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	9.832	37.59	<0.001
latitude : longitude : winter	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	15.256	81.07	<0.001
HH	thin plate smooth	numeric (4-18 hrs)	4.494	30.87	<0.001
water depth	thin plate smooth	numeric (0-50 m)	3.59	12.26	0.0164
SSTa	thin plate smooth	numeric (-3.0 - 3.0)	6.9	29.57	<0.001
cum piling / 15 days 60 km	thin plate smooth	numeric (0-21)	5.616	33.59	<0.001

Table A-14 Results model A2 - GAM analysing annual trends and spatial distribution patterns of harbour porpoises using unaffected dataset (n = 24,324, AIC = 51,445.72, dev. Explained = 9.0 %).

variable	regression technique	category	edf	Chi²	p-value
Position ID	mrf smooth	factor	297.62	664.70	<0.001
day of year 2009	cyclic smooth	numeric (1-365)	7.365	149.67	<0.001

variable	regression technique	category	edf	Chi²	p-value
day of year 2010	cyclic smooth	numeric (1-365)	5.781	86.68	<0.001
day of year 2011	cyclic smooth	numeric (1-365)	5.763	15.50	<0.001
day of year 2012	cyclic smooth	numeric (1-365)	6.537	79.11	<0.001
day of year 2013	cyclic smooth	numeric (1-365)	4.819	155.12	<0.001
year	factor	numeric (2009-2013)	-	20.4	<0.001
moon illumination	thin plate smooth	numeric (0.0-1.0)	7.736	39.83	<0.001
flight time	thin plate smooth	numeric (0.0-1.0)	4.475	101.33	<0.001
latitude : longitude : spring	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	10.014	23.92	0.016
latitude : longitude : summer	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	13.159	34.09	<0.001
latitude : longitude : autumn	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	10.390	48.72	<0.001
latitude : longitude : winter	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	14.786	90.14	<0.001
HH	thin plate smooth	numeric (4-18 hrs)	6.449	35.43	<0.001
water depth	thin plate smooth	numeric (0-50 m)	3.675	10.52	0.037
SSTa	thin plate smooth	numeric (-3.0 - 3.0)	4.60	18.95	0.003

Table A-15 Results model A3 - GAM analysing annual trends and spatial distribution patterns of harbour porpoises using affected dataset (n = 7,730, AIC = 14,368.06, dev. Explained = 25.3 %).

variable	regression technique	category	edf	Chi²	p-value
Position ID	mrf smooth	factor	138.394	283.15	<0.001
day of year 2009	cyclic smooth	numeric (1-365)	2.057	10.91	<0.001
day of year 2010	cyclic smooth	numeric (1-365)	5.750	62.46	<0.001
day of year 2011	cyclic smooth	numeric (1-365)	5.384	68.92	<0.001
day of year 2012	cyclic smooth	numeric (1-365)	6.912	78.47	<0.001
day of year 2013	cyclic smooth	numeric (1-365)	7.329	76.00	<0.001
moon illumination	thin plate smooth	numeric (0.0-1.0)	7.475	56.34	<0.001
latitude : longitude : spring	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	9.452	34.50	<0.001
latitude : longitude : summer	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	7.351	15.23	0.073
latitude : longitude : autumn	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	7.660	10.12	0.468
latitude : longitude : winter	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	12.146	56.12	<0.001
water depth	thin plate smooth	numeric (0-50 m)	4.195	12.545	0.026
SSTa	thin plate smooth	numeric (-3.0 - 3.0)	6.429	35.540	<0.001
year	factor	numeric (2009-2013)	-	15.16	0.004
season	factor	categorical	-	9.46	0.024

Table A-16 Results model A4 - GAM analysing annual trends and spatial distribution patterns of harbour porpoises in spring and summer using all grid cells in northern part of subarea German Bight NW (n = 5,928, AIC =13,400.66, dev. Explained = 13.6 %).

variable	regression technique	category	edf	Chi²	p-value
Position ID	mrf smooth	factor	11.305	12.177	0.160
day of year 2009	cyclic smooth	numeric (1-365)	7.,273	134.114	<0.001
day of year 2010	cyclic smooth	numeric (1-365)	7.845	119.632	<0.001
day of year 2011	cyclic smooth	numeric (1-365)	4.487	26.765	<0.001
day of year 2012	cyclic smooth	numeric (1-365)	4.186	41.902	<0.001
flight time	thin plate smooth	numeric (0.0-1.0)	3.661	77.325	<0.001
HH	thin plate smooth	numeric (4-18 hrs)	2.739	3.688	0.3676
Cum 0.5 month 60 km	thin plate smooth	numeric (0-20)	4.343	10.328	0.0762
moon illumination	thin plate smooth	numeric (0.0-1.0)	6.290	40.136	<0.001
latitude : longitude : 2009	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	10.720	22.539	0.050
latitude : longitude : 2010	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	8.215	22.070	0.018
latitude : longitude : 2011	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	3.012	2.811	0.4254
latitude : longitude : 2012	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	4.492	2.081	0.8810
“day and distance“	factor	10 categories	-	19.830	<0.001
year	factor	numeric (2009-2013)	-	1.213	0.750

Table A-17 Results model A5 - GAM analysing annual trends and spatial distribution patterns of harbour porpoises in spring and summer using southern part of subarea German Bight NW (n = 4,985, AIC =10,825.9, dev. Explained = 19.3 %).

variable	regression technique	category	edf	Chi²	p-value
Position ID	mrf smooth	factor	64.689	195.786	<0.001
day of year 2009	cyclic smooth	numeric (1-365)	5.655	75.987	<0.001
day of year 2010	cyclic smooth	numeric (1-365)	4.375	18.299	<0.001
day of year 2011	cyclic smooth	numeric (1-365)	3.062	37.074	<0.001
day of year 2012	cyclic smooth	numeric (1-365)	4.253	84.158	<0.001
day of year 2013	cyclic smooth	numeric (1-365)	4.,895	106.637	<0.001
moon illumination	thin plate smooth	numeric (0.0-1.0)	6.453	40.495	<0.001
water depth	thin plate smooth	numeric (0-50 m)	4.387	24.261	<0.001
latitude : longitude : 2009	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	4.669	4.721	<0.001
latitude : longitude : 2010	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	6.225	20.548	<0.001
latitude : longitude : 2011	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	3.003	5.753	<0.001
latitude : longitude : 2012	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	6.756	25.211	0.001
latitude : longitude : 2013	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	3.003	8.656	0.001
Cum 0.5 month 60 km	thin plate smooth	numeric (0-20)	3.929	14.014	0.034
"day and distance"	factor	10 categories	-	32.17	0.013
year	factor	numeric (2009-2013)	-	<0.0014	<0.001

Table A-18 Results model A6 - GAM analysing annual trends and spatial distribution patterns of harbour porpoises in spring and summer using all grid cells of subarea North of Borkum (n = 4,985, AIC =10,825.9, dev. Explained = 19.3 %).

variable	regression technique	category	edf	Chi²	p-value
Position ID	mrf smooth	factor	64.689	195.786	<0.001
day of year 2009	cyclic smooth	numeric (1-365)	5.655	75.987	<0.001
day of year 2010	cyclic smooth	numeric (1-365)	4.375	18.299	<0.001
day of year 2011	cyclic smooth	numeric (1-365)	3.062	37.074	<0.001
day of year 2012	cyclic smooth	numeric (1-365)	4.253	84.158	<0.001
day of year 2013	cyclic smooth	numeric (1-365)	4.,895	106.637	<0.001
moon illumination	thin plate smooth	numeric (0.0-1.0)	6.453	40.495	<0.001
water depth	thin plate smooth	numeric (0-50 m)	4.387	24.261	<0.001
latitude : longitude : 2009	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	4.669	4.721	<0.001
latitude : longitude : 2010	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	6.225	20.548	<0.001
latitude : longitude : 2011	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	3.003	5.753	<0.001
latitude : longitude : 2012	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	6.756	25.211	0.001
latitude : longitude : 2013	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	3.003	8.656	0.001
Cum 0.5 month 60 km	thin plate smooth	numeric (0-20)	3.929	14.014	0.034
“day and distance“	factor	10 categories	-	32.17	0.013
year	factor	numeric (2009-2013)	-	<0.0014	<0.001

Table A-19 Results model A7 - GAM analysing annual trends and spatial distribution patterns of harbour porpoises in spring and summer using all grid cells of West of Sylt (n = 6,835, AIC =17,934.70 dev. Explained = 20.5 %).

variable	regression technique	category	edf	Chi²	p-value
Position ID	mrf smooth	factor	122.573	299.57	<0.001
day of year 2011	cyclic smooth	numeric (1-365)	7.664	76.52	<0.001
day of year 2012	cyclic smooth	numeric (1-365)	3985	52.74	<0.001
day of year 2013	cyclic smooth	numeric (1-365)	7.072	112.39	<0.001
latitude : longitude : 2011	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	3.083	14.45	0.003
latitude : longitude : 2012	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	5.561	17.10	0.183
latitude : longitude : 2013	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	11.642	62.56	<0.001
HH	thin plate smooth	numeric (4-18 hrs)	4.523	55.80	<0.001
Cum 0.5 month 60 km	thin plate smooth	numeric (0-20)	3.397	12.52	0.016
SSTa	thin plate smooth	numeric (-3.0 - 3.0)	1.003	46.14	<0.001
“day and distance“	factor	10 categories	-	3.322	0.190
season	factor	4 categories	-	8.605	0.003
year	factor	numeric (2009-2013)	-	53.894	<0.001

Table A-20 Results model A8 - GAM analysing annual trends and spatial distribution patterns of harbour porpoises in spring and summer using all grid cells of German Bight NW (n = 15,819, AIC = 29,397.66, dev. Explained = 11.7 %).

variable	regression technique	category	edf	Chi²	p-value
Position ID	mrf smooth	factor	69.510	117.598	<0.001
day of year 2009	cyclic smooth	numeric (1-365)	6.981	150.683	<0.001
day of year 2010	cyclic smooth	numeric (1-365)	3.935	29.864	<0.001
day of year 2011	cyclic smooth	numeric (1-365)	3.640	69.226	<0.001
day of year 2012	cyclic smooth	numeric (1-365)	7.589	73.107	<0.001
day of year 2013	cyclic smooth	numeric (1-365)	7.068	78.359	<0.001
flight time	thin plate smooth	numeric (0-1)	4.481	89.707	<0.001
HH	cyclic smooth	numeric (0-24)	1.004	12.731	<0.001
Cum 0.5 month 60 km	thin plate smooth	numeric (0-20)	6.065	53.336	0.034
moon illumination	thin plate smooth	numeric (0.0-1.0)	6.319	65.246	<0.001
latitude : longitude : 2009	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	4.900	9.158	0.143
latitude : longitude : 2010	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	3.007	8.829	0.032
latitude : longitude : 2011	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	5.874	26.788	<0.001
latitude : longitude : 2012	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	7.033	11.370	0.2435
latitude : longitude : 2013	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	4.731	8.244	0.1941
year	factor	numeric (2009-2013)	4	76.99	<0.001
“day and distance“	factor	10 categories	9	37.49	<0.001

Preliminary results from distribution models per cluster

The results of the general distribution models presented above (chap. 6.3.2) showed mostly differences that were described best by seasonal and spatial variables. As the study, area was not covered spatially over the entire study period the focus was placed on the three subareas ("West of Sylt", "North of Borkum" and "German Bight NW"). To differentiate effects of pile driving on the distribution patterns of porpoise’s three datasets were compared. The entire dataset, the unaffected and affected data per subarea. The affected data were split per cluster in two or three models, because wind farms were either not constructed in the same year (BWII, and AV), not in the same season and/or were too distant from another (DT and MSO; GTI and BARD). In light of the limited amount of flights available per project and season (Figure 6-5), the GAMs were kept simple. First, models with differing complexity and amount of data are difficult to compare, second, models with the same variables were over parametrized and had to be reduced to the relevant variables (variable were tested if p-value < 0.1 and model selection was based on most parsimonious model; delta AIC > 2). The variables considered in these models were "area surveyed" as offset, the "neighbourhood matrix" for considering spatial autocorrelation and the "latitudinal and longitudinal coordinates" as tensor spline, which were grouped per season to give a seasonal flexibility to the model.

Focusing on the subarea "German Bight NW" (n=15,819), where BARD and GTI were constructed over several years and most of the flights conducted during or shortly after piling events were conducted next to them (Figure A-32).

The illustration of the smooth in summer indicated that when using the full data model, both BARD and GTI were located within an area of lower densities, while the use of unaffected data model showed very different distributions, which indicated that especially at BARD densities were significantly lower in summer. For GTI, the summer analysis indicated rather high densities near to the wind farm, while in the southwest part, beyond BARD, the densities were significantly higher. The variable describing the distribution in autumn indicated lower densities around BARD and GTI in both models. Investigating the flights conduct during construction indicated that the results associated to BARD showed a similar pattern, while at GTI the confidence interval indicated mean values that were higher in the northeast part, in the direction of the SAC Sylt Outer Reef, with respect to the values captured in the southern area. For the full dataset, the surveys deployed during the winter showed a band of higher densities stretching from north to south. In the three other models this pattern was, conversely, very different. Interestingly, BARD was within an area of significantly lower densities in the model with unaffected flights only and it was displaying higher densities than the model with affected surveys. The same pattern could be seen for the GTI dataset, that showed higher densities around the wind farm and lower values to the southwest at the German-Dutch border of the EEZ.

Focusing on the subarea "North of Borkum" (Figure A-33), the wind farms constructed here were AV, BWII and RG. AV was constructed in spring and summer, BWII was constructed in spring and autumn while RG was constructed only in summer, 2015.

The spring smooth related to the full dataset (n = 6712) and the unaffected dataset (n = 4962) exhibited significantly lower densities around the wind farm AV. In the affected data, however, there was no indication of lower densities around this wind farm. At BWII, the wind farm did not exhibit indications of higher or lower densities in any model. In summer, densities around AV were higher, both in the affected and unaffected models. In contrast, the model with the affected data showed significantly lower densities around the wind farm and significantly higher estimates directly south of it. Regarding RG, densities were significantly lower in the model with all data and in the unaffected one, but the trajectories indicated that densities were significantly higher there during piling. In autumn, during the construction of BWII, all three models consistently exhibited that the density west of the wind farm in the SAC "Borkum Reef Ground" was higher, while around the wind farm itself, the density estimate did not deviate from the mean.

In the subarea "West of Sylt" (Figure A-34), which was surveyed primarily from autumn 2010 to winter 2013, the wind farms MSO, NSO and DT were constructed. The three wind farms were constructed almost simultaneously, but the coverage of the flights was better at MSO and NSO than at DT. MSO and NSO were analysed together, because these wind farms were next to each other and the effect of the construction was continuous at both.

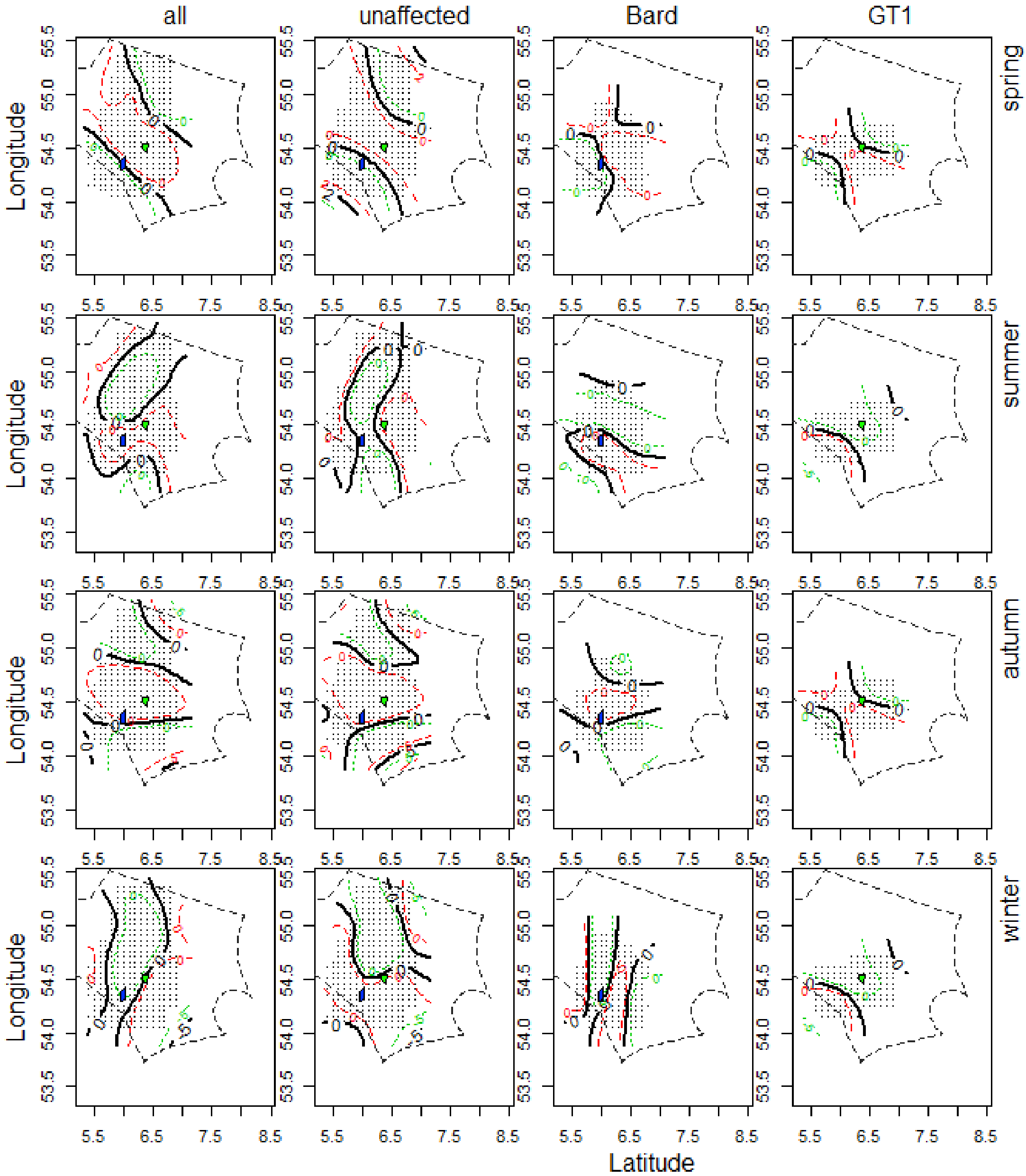


Figure A-32 Illustrations of different GAMs of the subarea "German Bight NW" showing tensor spline smooths of latitude and longitude per season for all data, unaffected data of this subarea and affected data of BARD and GTI.

The distribution during spring indicated higher estimates in the SAC "Sylt Outer Reef". Lower densities during the construction of MSO and NSO were expected. However, significantly lower values were detected around DT. In summer, the distribution in this subarea indicated higher densities in the SAC in both the affected and unaffected dataset. Again, the densities during the construction phase indicated significantly lower densities around DT, with a reduced spatial extension of the effect compared to spring. At MSO and NSO, the affected dataset showed no indication of lower or higher densities than the mean. In autumn, flights were conducted only during the construction of MSO and NSO. In all three models, densities were lower around the construction sites. Finally, in winter, piling was conducted only during flights next to MSO and NSO. Here, there were no indications of lower densities.

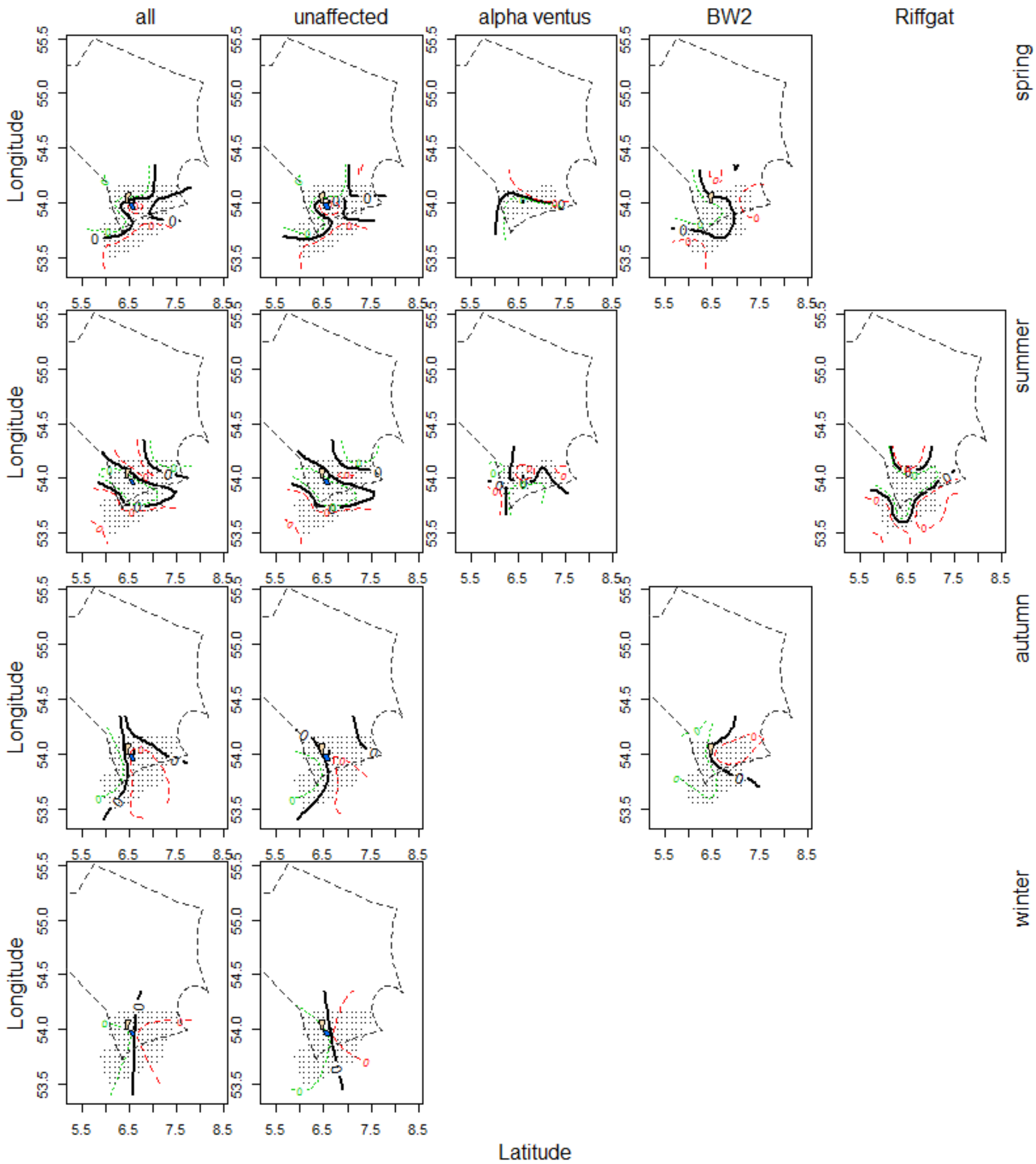


Figure A-33 Illustrations of different GAMs of the subarea "North of Borkum" showing tensor spline smooths of latitude and longitude per season for all data, unaffected data of this subarea and affected data of "AV", BWII and RG.

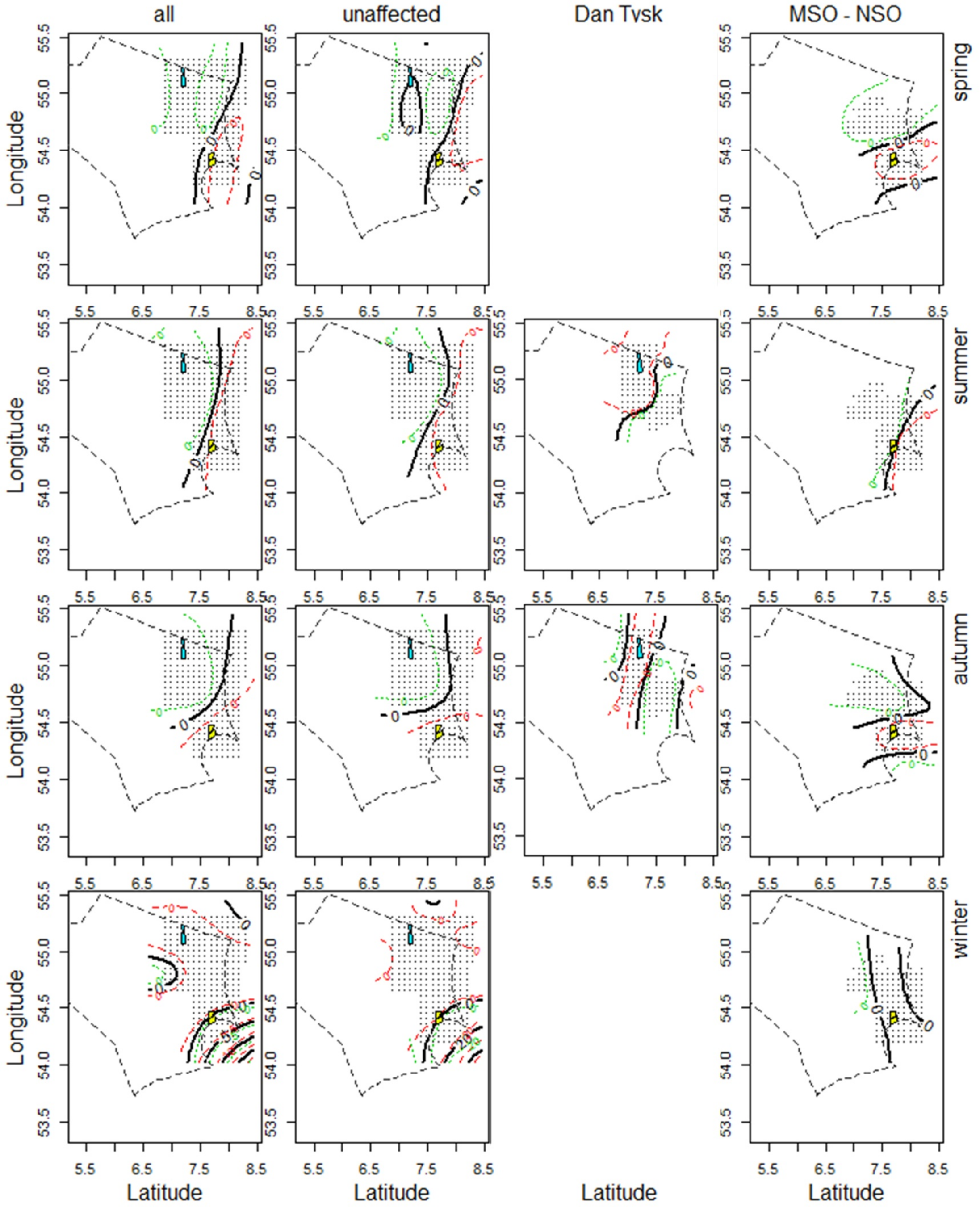


Figure A-34 Illustrations of different GAMs of the subarea "West of Sylt" showing tensor spline smooths of latitude and longitude per season for all data, unaffected data of this subarea and affected data of DT, MSO and NSO.

Results from effect model

Figure A-35 presents the seasonal porpoise densities according to distance to piling. Table A-21 presents the model outputs of one GAM describing the temporal and spatial effect of piling events on the distribution of harbour porpoises. Figure A-36 presents 5 sub-models testing the reliability of distance effects according to distance and time to piling. Figure A-37 presents the distribution of density estimates per distance class and time since piling ceased for each wind farm. Figure A-38 illustrates the two-dimensional smooth of distance from piling and minimal time since piling ceased. And finally, Table A-26 gives the model outcomes of effect range of piling gams grouped per wind farm.

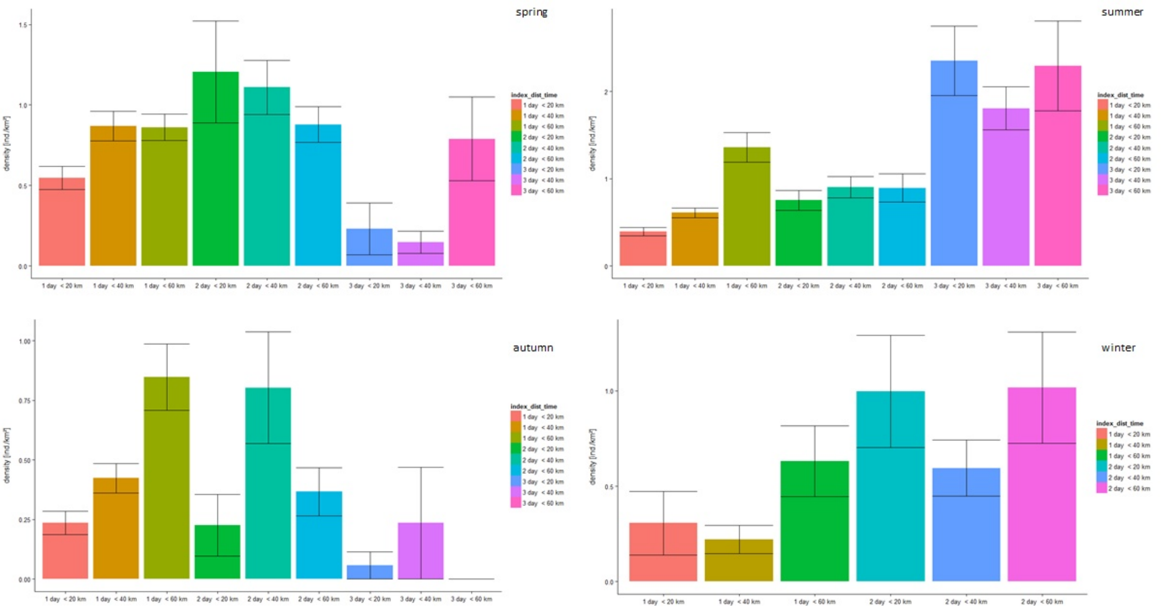


Figure A-35 Seasonal barplot of harbour porpoise's densities grouped by „day and distance" (illustration of mean and standard deviation).

Table A-21 Results model ST1 - GAM describing the temporal and spatial effect of piling events on the distribution of harbour porpoises (hours relative to piling 0 – 60; distance: 0 – 60 km; n = 6,682; AIC =12,368.08; dev. Explained = 25.9 %).

variable	regression technique	category	edf	Chi²	p-value
Position ID	mrf smooth	factor	114.80	185.233	<0.001
day of the year 2009	cyclic smooth	numeric (1-365)	0.005	0.001	0.6936
day of the year 2010	cyclic smooth	numeric (1-365)	1.913	12.504	<0.001
day of the year 2011	cyclic smooth	numeric (1-365)	5.256	35.758	<0.001
day of the year 2012	cyclic smooth	numeric (1-365)	6.065	38.133	<0.001
day of the year 2013	cyclic smooth	numeric (1-365)	6.735	43.468	<0.001
moon illumination	thin plate smooth	numeric (0.0-1.0)	5.560	27.151	<0.001
flight time	thin plate smooth	numeric (0.0-1.0)	4.265	25.983	<0.001
latitude : longitude : spring	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	7.064	26.333	<0.001
latitude : longitude : summer	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	7.356	28.904	<0.001
latitude : longitude : autumn	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	3.000	2.229	<0.001
latitude : longitude : winter	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	11.23	46.895	0.001
latitude : longitude : 2013	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	12.54	87.244	0.001
water depth	thin plate smooth	numeric (0-50 m)	3.949	10.777	<0.001
min_dist_deterence	thin plate smooth	numeric	4.563	29.396	<0.001
year	factor	10 categories	-	17.96	0.002
season	factor	numeric (2009-2013)	-	28.96	<0.001

Table A-22 Results model ST2 - GAM describing the temporal and spatial effect of piling events on the distribution of harbour porpoises (hour relative to piling 0 – 60; distance: 0 – 40 km; n = 4,675; AIC = 7892.69; dev. Explained = 24.7%).

variable	regression technique	category	edf	Chi²	p-value
Position ID	mrf smooth	factor	78.5092	114.321	<0.001
day of the year 2009	cyclic smooth	numeric (1-365)	0.0018	0.001	0.2812
day of the year 2010	cyclic smooth	numeric (1-365)	1.9757	20.555	<0.001
day of the year 2011	cyclic smooth	numeric (1-365)	1.8354	5.072	0.033
day of the year 2012	cyclic smooth	numeric (1-365)	1.1429	2.428	0.075
day of the year 2013	cyclic smooth	numeric (1-365)	0.0011	0.000	0.635
moon illumination	thin plate smooth	numeric (0.0-1.0)	5.0052	33.947	<0.001
flight time	thin plate smooth	numeric (0.0-1.0)	3.1157	28.826	<0.001
latitude : longitude : spring	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	6.2108	43.050	<0.001
latitude : longitude : summer	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	4.9297	12.508	0.042
latitude : longitude : autumn	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	5.2687	37.064	<0.001
latitude : longitude : winter	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	8.6591	26.746	0.003
latitude : longitude : 2013	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	8.1944	52.587	<0.001
water depth	thin plate smooth	numeric (0-50 m)	3.0880	7.441	0.094
min_dist_deterence	thin plate smooth	numeric	3.8545	11.019	0.042
year	factor	10 categories	4	27.151	<0.001
season	factor	numeric (2009-2013)	3	5.872	0.188

Table A-23 Results model ST3 - GAM describing the temporal and spatial effect of piling events on the distribution of harbour porpoises (hours relative to piling 0 – 24; distance: 0 – 60 km; n = 1,819; AIC = 2854.28; dev. Explained = 34.1%).

variable	regression technique	category	Redf	Chi²	p-value
Position ID	mrf smooth	factor	4.159	39.258	0.999
day of the year 2009	cyclic smooth	numeric (1-365)	0.004	0.000	0.618
day of the year 2010	cyclic smooth	numeric (1-365)	1.232	3.908	0.034
day of the year 2011	cyclic smooth	numeric (1-365)	2.433	9.762	0.004
day of the year 2012	cyclic smooth	numeric (1-365)	4.129	0.000	0.430
day of the year 2013	cyclic smooth	numeric (1-365)	0.002	0.002	0.927
moon illumination	thin plate smooth	numeric (0.0-1.0)	1.000	0.008	<0.001
Flight time	thin plate smooth	numeric (0.0-1.0)	2.160	16.405	0.017
latitude : longitude : spring	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	5.536	15.562	0.606
latitude : longitude : summer	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	3.004	1.846	0.019
latitude : longitude : autumn	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	3.000	9.892	0.025
latitude : longitude : winter	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	5.375	14.557	<0.001
latitude : longitude : 2013	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	7.499	28.428	0.001
water depth	thin plate smooth	numeric (0-50 m)	2.967	3.327	0.441
min_dist_deterence	thin plate smooth	numeric	3.329	8.635	0.073
year	factor	10 categories	4	23.840	<0.001
season	factor	(spring - winter)	3	4.265	0.234

Table A-24 Results model ST4 - GAM describing the temporal and spatial effect of piling events on the distribution of harbour porpoises (hours relative to piling 0 – 48; distance: 0 – 60 km; n = 6,325; AIC = 11,340.39; dev. Explained = 25.5 %).

variable	regression technique	category	edf	Chi²	p-value
Position ID	mrf smooth	factor	111.212	175.450	<0.001
day of the year 2009	cyclic smooth	numeric (1-365)	2.234	16.442	<0.001
day of the year 2010	cyclic smooth	numeric (1-365)	1.710	8.384	0.002
day of the year 2011	cyclic smooth	numeric (1-365)	6.154	49.963	<0.001
day of the year 2012	cyclic smooth	numeric (1-365)	6.475	46.035	<0.001
day of the year 2013	cyclic smooth	numeric (1-365)	7.311	64.815	<0.001
moon illumination	thin plate smooth	numeric (0.0-1.0)	6.067	38.233	<0.001
Flight time	thin plate smooth	numeric (0.0-1.0)	4.107	27.693	<0.001
latitude : longitude : spring	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	6.336	28.096	<0.001
latitude : longitude : summer	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	7.680	20.163	0.020
latitude : longitude : autumn	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	7.792	7.035	0.694
latitude : longitude : winter	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	11.454	39.273	<0.001
latitude : longitude : 2013	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	11.684	81.064	<0.001
water depth	thin plate smooth	numeric (0-50 m)	3.944	12.667	0.022
min_dist_deterence	thin plate smooth	numeric	4.149	22.004	<0.001
year	factor	10 categories	4	14.040	<0.001
season	factor	(spring - winter)	3	28.980	<0.001

Table A-25 Results model ST5 - GAM describing the temporal and spatial effect of piling events on the distribution of harbour porpoises (hours relative to piling 0 – 24; distance: 0 – 60 km; n = 4,535; AIC =7,859.726; dev. Explained = 34.7%).

variable	regression technique	category	edf	Chi²	p-value
Position ID	mrf smooth	factor	152.50	288.471	<0.001
day of the year 2009	cyclic smooth	numeric (1-365)	0.772	2.118	<0.001
day of the year 2010	cyclic smooth	numeric (1-365)	0.000	0.000	0.592
day of the year 2011	cyclic smooth	numeric (1-365)	5.781	29.507	<0.001
day of the year 2012	cyclic smooth	numeric (1-365)	5.285	23.863	<0.001
day of the year 2013	cyclic smooth	numeric (1-365)	3.072	19.488	<0.001
moon illumination	thin plate smooth	numeric (0.0-1.0)	4.804	24.540	<0.001
Flight time	thin plate smooth	numeric (0.0-1.0)	3.632	10.329	0.041
latitude : longitude : spring	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	4.641	4.103	0.569
latitude : longitude : summer	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	9.287	23.321	0.015
latitude : longitude : autumn	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	9.691	15.334	0.203
latitude : longitude : winter	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	4.153	6.992	0.195
latitude : longitude : 2013	2-D tensor spline	numeric (53-56 / 5.4 - 8.3)	1.324	65.394	<0.001
water depth	thin plate smooth	numeric (0-50 m)	1.243	0.466	0.622
min_dist_deterence	thin plate smooth	numeric	5.261	27.775	<0.001
year	factor	10 categories	4	4.889	0.299
season	factor	(spring - winter)	3	9.094	0.028

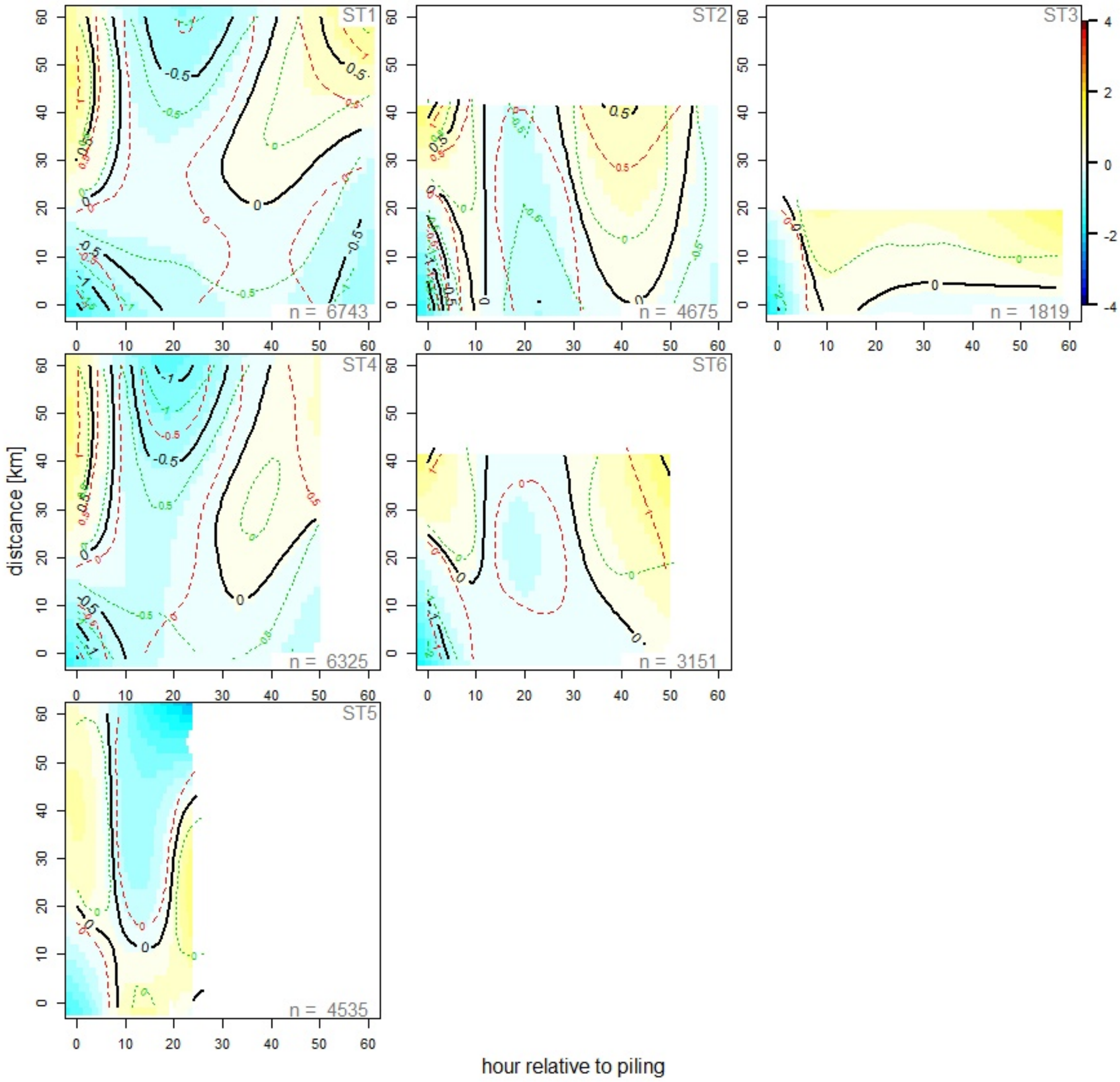


Figure A-36 GAM-plots of models ST2 – ST6 to test the reliability of distance effects by changing the variables "distance" and "hour relative to piling".

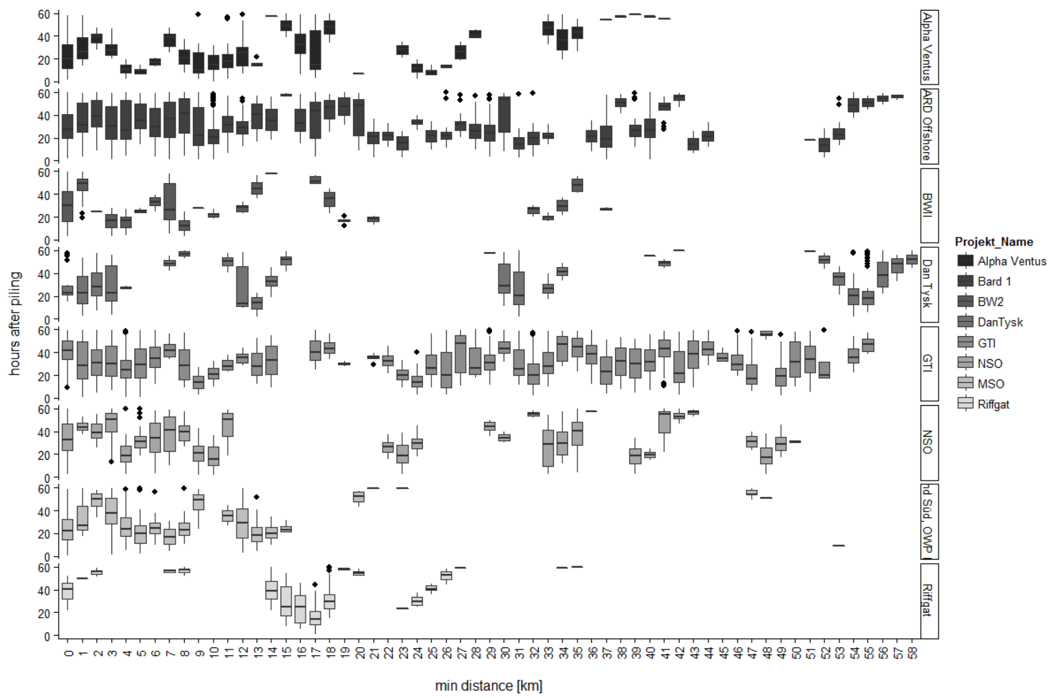


Figure A-37 Box-Whisker-plots of the distribution of density estimates per distance class and time since piling ceased for each wind farm (95 % quantile: lower bar; 25 % to 75 % quantiles mark borders of upper, black bar: median and lower box, 5% quantiles are marked by upper error bar).

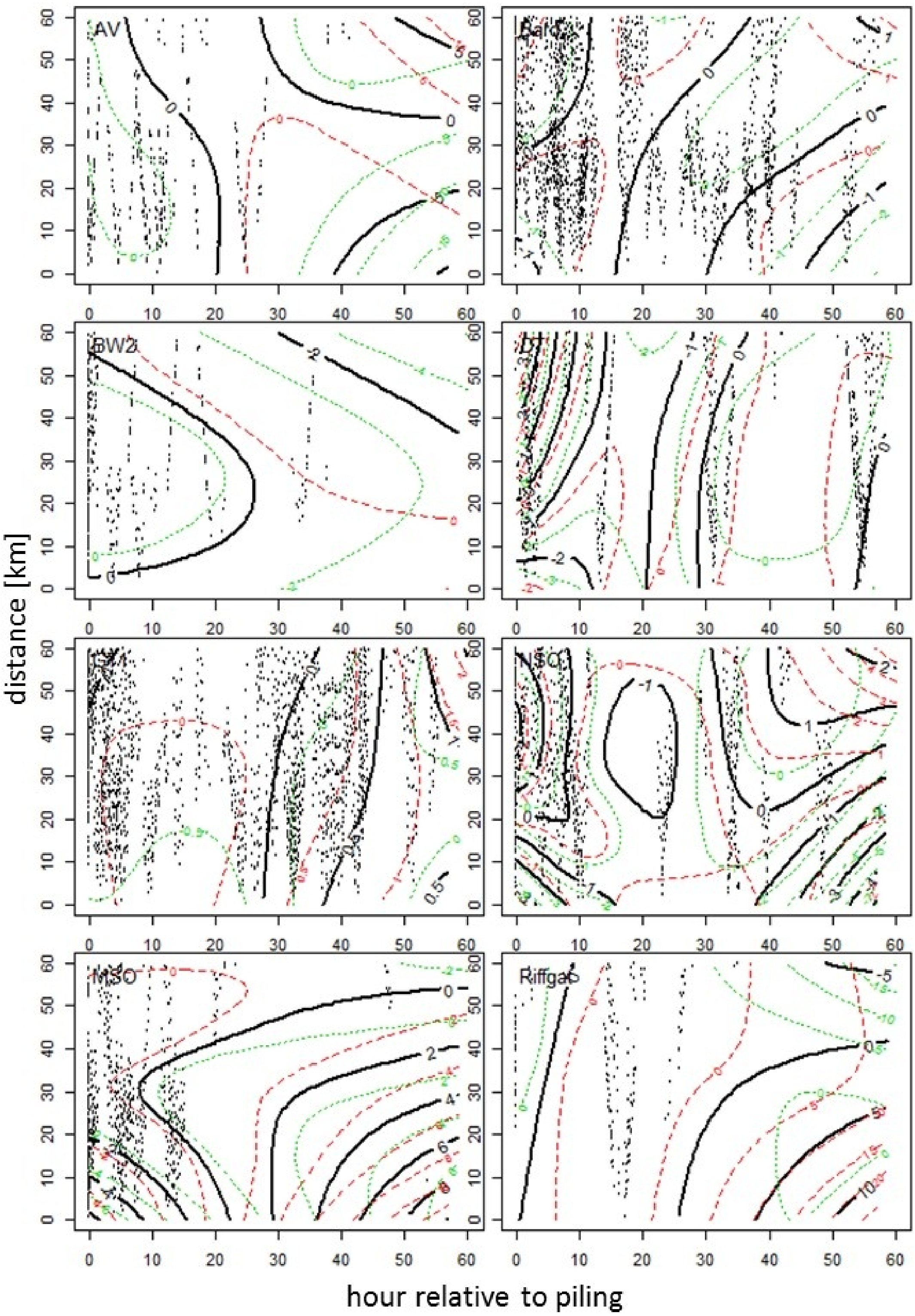


Figure A-38 GAM-plot illustrating the two-dimensional smooth of distance from piling and minimal time since piling ceased (left: isoclines with confidence bands and right coloured representation of the same graph; red values are negative values, yellow values are positive

Table A-26 Results model ST6 - GAM describing the temporal and spatial effect of piling events on the distribution of harbour porpoises per wind farm (N = 6.871, AIC = 12,328.2 dev. Explained = 28.7 %).

variable	regression technique	category	edf	Chi²	p-value
Position ID	mrf smooth	factor	107.7	166.075	<0.001
day of the year 2009	cyclic smooth	numeric (1-365)	0.000	0.000	0.587
day of the year 2010	cyclic smooth	numeric (1-365)	5.482	48.684	<0.001
day of the year 2011	cyclic smooth	numeric (1-365)	4.148	20.666	<0.001
day of the year 2012	cyclic smooth	numeric (1-365)	5.679	28.273	<0.001
day of the year 2013	cyclic smooth	numeric (1-365)	5.294	30.482	<0.001
moon illumination	thin plate smooth	numeric (2009-2013)	6.011	28.362	<0.001
flight time	thin plate smooth	numeric (0.0-1.0)	3.949	12.284	0.022
latitude : longitude : spring	2-D tensor spline	numeric	6.885	11.864	0.1741
latitude : longitude : summer	2-D tensor spline	numeric	5.741	9.693	0.1714
latitude : longitude : autumn	2-D tensor spline	numeric	6.389	11.359	0.138
latitude : longitude : winter	2-D tensor spline	numeric	11.57	50.110	<0.001
Hour relative to piling: AV	2-D tensor spline	numeric (0-60 hrs/km)	7.896	22.974	<0.001
Hour relative to piling: BARD	2-D tensor spline	numeric (0-60 hrs/km)	9.267	30.725	<0.001
Hour relative to piling: BWII	2-D tensor spline	numeric (0-60 hrs/km)	4.555	12.055	<0.001
Hour relative to piling: DT	2-D tensor spline	numeric (0-60 hrs/km)	7.080	43.804	<0.001
Hour relative to piling: GTI	2-D tensor spline	numeric (0-60 hrs/km)	3.350	1.582	0.764
Hour relative to piling: NSO	2-D tensor spline	numeric (0-60 hrs/km)	10.18	40.314	<0.001
Hour relative to piling: MSO	2-D tensor spline	numeric (0-60 hrs/km)	4.986	29.098	<0.001
Hour relative to piling: RG	2-D tensor spline	numeric (0-60 hrs/km)	3.002	13.663	<0.001
water depth	thin plate smooth	numeric (0-50 m)	4.367	14.124	<0.001
netto deterrence time	thin plate smooth	minutes	1.000	1.272	0.259
netto piling time	thin plate smooth	minutes	5.114	27.613	<0.001
year	factor		4	71.33	<0.001
season	factor	(spring, summer, autumn or winter)	3	20.68	<0.001
Project	factor		6	200.93	<0.001