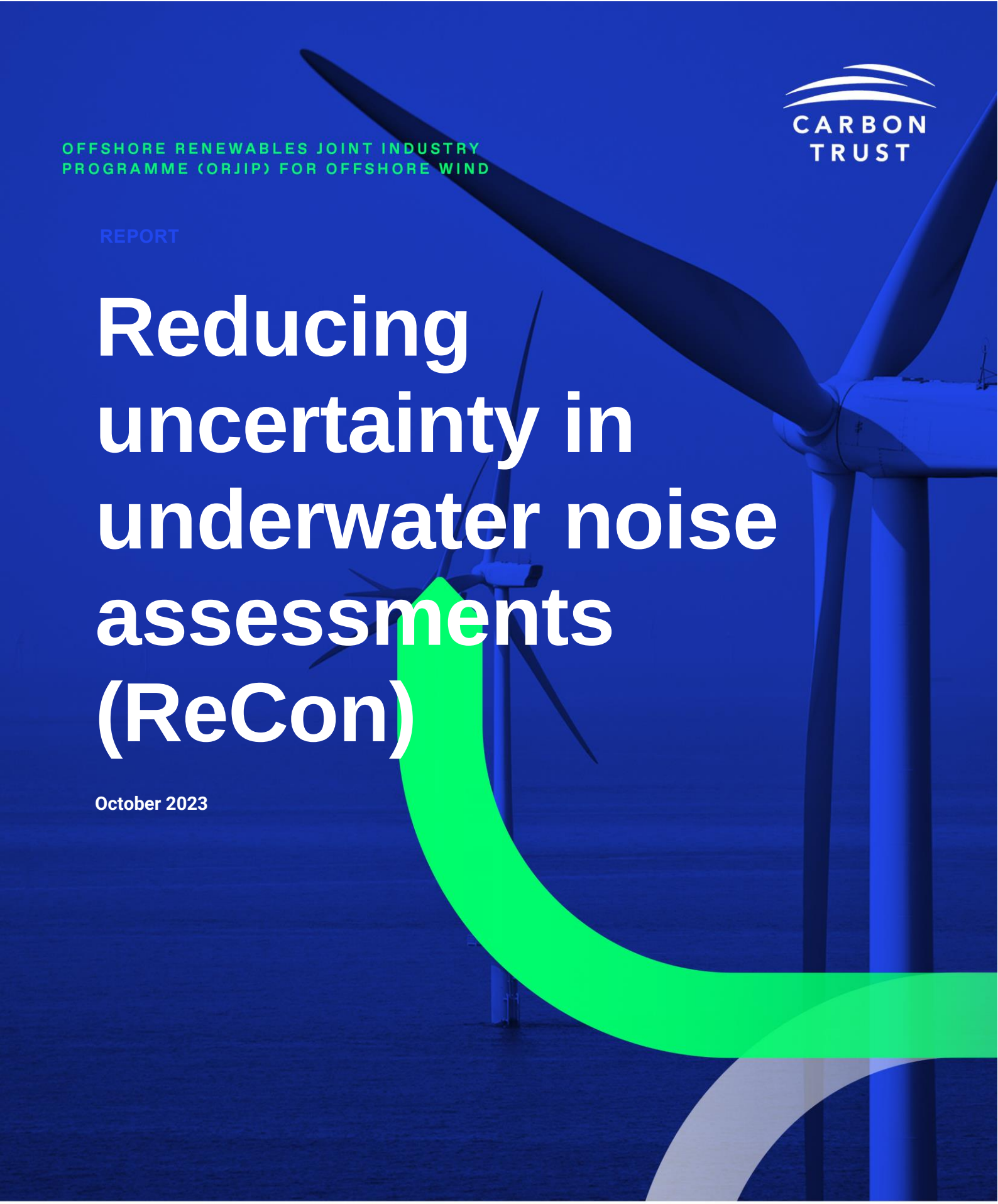

OFFSHORE RENEWABLES JOINT INDUSTRY PROGRAMME (ORJIP) FOR OFFSHORE WIND

CARBON TRUST

REPORT

Reducing uncertainty in underwater noise assessments (ReCon)

October 2023

orjip
working to accelerate offshore consenting

CARBON TRUST | OFFSHORE WIND

Crown Estate Scotland
Oighreachd a' Chrùin Alba

Offshore Wind Evidence + Change Programme | THE CROWN ESTATE

EDF renouvelables

equinor

Ørsted

OW

RWE

Scottish Government
Riaghaltas na h-Alba

SSE Renewables

TotalEnergies

INTRODUCTION

ORJIP Offshore Wind

The Offshore Renewables Joint Industry Programme (ORJIP) for Offshore Wind is a collaborative initiative that aims to:

- Fund research to improve our understanding of the effects of offshore wind on the marine environment
- Reduce the risk of not getting or delaying consent for offshore wind developments
- Reduce the risk of getting consent with conditions that reduce viability of the project.

The programme pools resources from the private sector and public sector bodies to fund projects that provide empirical data to support consenting authorities in evaluating the environmental risk of offshore wind. Projects are prioritised and informed by the ORJIP Advisory Network that includes key stakeholders such as statutory nature conservation bodies, academics, non-governmental organisations, and others.

ORJIP for Offshore Wind Stage 2 is a collaboration between The Carbon Trust, EDF Energy Renewables Limited, Ocean Wind UK Limited, Equinor ASA, Ørsted Wind Power A/S, RWE Renewables GmbH, Shell Global Solutions International B.V, SSE Renewables Developments UK Limited, TotalEnergies E&P UK Limited, Crown Estate Scotland, The Scottish Government, and The Crown Estate Commissioners, with this study additionally supported by Red Rock Power Limited.

Acknowledgments

This document was produced on behalf of ORJIP Offshore Wind by SMRU Consulting and itap, with review by the ORJIP Offshore Wind Steering Group and ReCon Project Expert Panel. Contributing authors were as follows: Verfuss, UK, Remmers, P, Ryder, M, Palmer, K, Wood, JD & Bellmann, MA.

SMRU Consulting
understand ♦ assess ♦ mitigate

itap
INSTITUT FÜR TECHNISCHE UND ANGEWANDTE PHYSIK GMBH

Who we are

We are a trusted, expert guide to Net Zero, bringing purpose led, vital expertise from the climate change frontline. We have been pioneering decarbonisation for more than 20 years for businesses, governments, and organisations around the world.

CARBON TRUST

The Carbon Trust's mission is to accelerate the move to a decarbonised future.

Contents

Tables

Figures

Abbreviations

ADD	Acoustic Deterrent Device
AIC	Akaike's Information Criterion
BE	Best Estimate
DCOs	Development Consent Orders
DMLs	Deemed Marine Licences
EIA	Environmental Impact Assessment
ES	Environmental Statement
GLMM	Generalised Minear Mixed-effect Models
LF	Low Frequency

ADD	Acoustic Deterrent Device
MMMP	Marine Mammal Mitigation Protocol
MMO	Marine Management Organisation
MS-LOT	Marine Scotland Licensing Operations Team
OLS	Ordinary Least Squared
OWF	Offshore Wind Farm
PS	Piling Strategy
PTS	Permanent Threshold Shift
TTS	Temporary Threshold Shift
UB	Upper Bound
VHF	Very High Hrequency
VIF	Variance Inflation Factors

Units

Unit	Description
µPa	Micro Pascal
dB	Decibel
s	Second

Metrics

Unit	Description
be	Blow Energy
E	Sound Exposure

E_{cum}	Cumulative Sound Exposure
$L_{p,pk}$	Zero-to-peak Sound Pressure Level
m	Meter
p(t)	Time Variant Sound Pressure
p_0	Reference Sound Pressure
pd	Pile Diameter
SEL_{cum}	Cumulative Sound Exposure Level
SEL_{ss}	Single Strike Sound Exposure Level
TT	Averaging Time
wd	Water Depth
X_{90}	90% Confidence Interval
p_{pk}	Maximum Sound Pressure
τ_{90}	90% Energy Signal Duration

1. Introduction

A key issue for offshore windfarm consent is the emission of underwater noise during the installation of turbine and offshore substation foundations. Percussive piling of pin piles or monopiles into the seabed using large hydraulic hammers, introduces high-amplitude impulsive underwater noise into the marine environment. Such noise can induce auditory injury and behavioural responses in marine mammals and fish (Lucke et al. 2009, Kastelein et al. 2014, Popper et al. 2014, Southall et al. 2019). The rapid development of the offshore wind industry and the drive for higher generation capacity has resulted in larger, higher-capacity turbines. Larger turbines require larger foundation structures, larger pile diameters and/or longer embedded piles. Consequently, higher blow energies are required for piling, resulting in greater concern over noise emissions (Bellmann et al. 2020). All cetaceans are listed under Annex IV of the EU Habitats Directive as European Protected Species (EPS) of Community Interest and require strict protection. The Habitats Regulations and the Offshore Marine Regulations make it an offence to injure or disturb any EPS. Therefore, Environmental Impact Assessments (EIAs) for offshore wind farms (OWF) in UK waters require quantitative assessments of the potential for auditory injury and disturbance impacts. It is important that impact assessments neither overestimate nor underestimate impact ranges, as this could result in unnecessary mitigation requirements or a delay in construction, respectively.

In the UK, the consenting process involves the use of models to predict noise levels and noise impact ranges on marine taxa. Underwater noise propagation models (and their predictions) rely on a set of assumptions which may have varied influence on the predicted outcome. These assumptions are made with a level of precaution. In the following sections, we summarise the current state of the assessment process to understand which factors may influence the magnitude of the predicted impact.

Current noise impact assessments for EIAs in the UK use a combination of noise propagation modelling and exposure criteria to quantitatively predict the area over which animals may be impacted. The following steps are conducted in each assessment:

1.1. Defining piling scenarios

UK OWF EIAs typically provide both a Maximum Design Scenario (mostly call Upper Bound – UB), and a Most Likely Scenario (Best Estimate – BE) in their underwater noise impact assessments. This is to ensure that the consent application covers the “noisiest condition” (worst case) within the project envelope for consenting purposes and provides context for a more realistic piling scenario. Piling scenarios are defined by the pile type (i.e., pin pile or monopile), pile diameter, maximum blow energy used, and the piling sequence. The piling sequence is characterised by a soft start (a series of pile strikes of low blow energy), a ramp up (with pile strikes of increasing blow energy up to maximum blow energy) and continued piling (pile strikes of high blow energies). The piling sequence is then defined by the duration of the soft start, ramp up and continued piling duration, the blow rate and blow energy used for each of the piling phases. Scenarios also consider if piles are installed as a single event (one pile per 24 hours), concurrent (two piles at the same time) or sequential (more than one pile per 24 hours).

1.2. Defining thresholds for auditory injury and behavioural responses

Different sets of noise thresholds have been used over time to determine impact ranges as piling noise assessment methodologies have changed over time, adapting to the most recent information available

(Verfuss et al. 2016, Faulkner et al. 2018). For current EIAs, the most recent auditory injury guidelines are used (Popper et al. 2014, National Marine Fisheries Service 2018, Southall et al. 2019). These require zero-to-peak Sound Pressure Level (Lp,pk) and the cumulative Sound Exposure Level (SELcum) to be predicted to obtain instantaneous PTS , TTS-onset impact ranges (based on Lp,pk) or cumulative PTS or TTS-onset impact ranges (based on SELcum). For fish, the SELcum is based on unweighted single strike SEL (SELss) and, for marine mammals, group-specific frequency weighting curves are applied. For the cumulative PTS-onset impact range, the sound energy an animal is exposed to is accumulated over the entire piling sequence, considering a scenario(s) of predicted animal movement. The cumulative PTS-onset impact range is then defined as the 'safe distance' at which an animal must be at before the start of piling, so as not to be at risk of PTS. To determine behavioural response impact ranges, mostly Lp,pk or SELss are used, either as a fixed threshold value or in combination with a dose-response curve, but no standard guidelines exist for this.

1.3. Conducting underwater noise modelling

With project and site-specific information as inputs, models are developed to predict noise source characteristics and its propagation into the marine environment. Underwater noise propagation models are strongly dependent on the assumed sound characteristics (sound level and frequency spectrum) emitted and the transmission loss. Predictions of cumulative impact ranges are sensitive to assumptions on animal movement (e.g., stationary vs fleeing in a single direction vs more complex simulations of movement) and the piling sequence (i.e., the soft start, ramp up and the variation in blow rate and blow energy).

When comparing impact assessments, there are no national or international standards for underwater noise modelling (beyond some national guidelines or best practice reports) and no standardised modelling software or standardised benchmarks (Müller and Zerbs 2013, de Jong et al. 2021). For modelling the propagation of impulsive pile-driving noise, two different model approaches are available: (i) a numerical approach, which is based on physical fundamentals (Wang et al. 2014) and (ii) an empirical model approach based on measured data.

1.3.1. Numerical approach

Numerical solutions can be used to calculate the sound pressure at any point in the water. The approaches are usually very computationally intensive. The advantages of numerical models are that any radiating noise (e.g., produced by the interaction between hammer and pile, pile and soil, and pile and water etc.) can be modelled and that the sound propagation (including interactions between water and soil, and water and sea surface) can be determined. The limitations are that these models require many detailed input parameters to precisely model the sound source itself as well as the propagation path. In most cases, the precision required of the input parameters is not available prior to construction. Therefore, such input parameters must be simplified (e.g., detailed soil profile at the pile site, soil layer conditions inside and outside the OWF area etc. are not available). Each simplification or assumption will lead to an uncertainty; therefore, an intensive error propagation should be performed to characterise the overall uncertainty. However, this is often not available for the current models in use. Ultimately, numerical models must be calibrated, evaluated, and validated with a huge set of measured data. Like-to-like comparison tests, where the same set of input parameters were used independently with different numerical models, show that predictions from close-range models can vary by a few dB, even in a very simplified pile-driving scenario (with a very simplified soil layer structure and flat bathymetry), and up to 15 fdB at 50 km distance in long-range models (Lippert et al.

2016). With increasing complexity of the scenario setting for any modelling task the uncertainty between different model approaches increased as well.

1.3.2. Empirical approach

Geometric dispersion functions are used as an alternative to the numerical solutions. These dispersion functions are often based on an empirical data set, i.e., measurements obtained in the field. The advantages of empirical models are that they require only a very limited number of data inputs, and the uncertainty of the model will significantly decrease with an increasing empirical dataset to help inform the model. A limitation of the empirical approach is that the model should be used within the parameter spread measured. An extrapolation is possible, but the uncertainty significantly depends on the quality of the underlying empirical data. However, the most influential parameters on piling noise must be identified to characterise the required input for the modelling.

Both approaches need to be adjusted to a source level or near field model, modelled by a finite-elements model or based on an empirical model. The first step in all approaches, whether numerical or empirical, is therefore to determine the sound amplitude level of the source (source level of the pile). The source level is typically defined as the sound level at 1 m, but can, depending on the model used, be between 1 m and a few 100 m from the pile. This source level is then coupled to a propagation model. Both the source level and the propagation model introduce uncertainties that add up. The total uncertainty increases with increasing distance from the source as the influence of the sound propagation model increases (Lippert et al. 2016). Most commercial suppliers have developed their own proprietary underwater noise modelling tools making direct comparison of the same piling scenario challenging. Moreover, it is often not known which kind of data input a certain proprietary software requires for modelling.

1.4. Estimating potential impacts and mitigation requirements

Outputs of noise modelling include impact ranges where noise levels, either as a single pulse or cumulative dose, are predicted to result in auditory injury and behavioural responses. These are typically combined with species-specific density information to estimate the number of animals potentially impacted and relate this to the relevant population unit. To reduce the risk of impacts to acceptable levels, including a reduction in risk of permanent auditory injury to negligible, mitigation measures are designed and implemented through a Marine Mammal Mitigation Protocol (MMMP) or Piling Strategy (PS). Consequently, reducing conservatism, where appropriate, in the above elements is critical to ensure appropriate mitigation measures are developed.

1.5. Measured (real-world) pile-driving noise data

The overall aim of the underwater noise monitoring during pile-driving activities is to validate the compliance with the MMMP/PS. Non-compliance due to underestimated impact ranges may lead to the requirement of adapting mitigation measures and might delay the construction process. In case of overestimated impact ranges, mitigation measures might be reduced with the approval of the relevant authorities (e.g., reduce the application period of Acoustic Deterrent Devices (ADDs)) to minimise their impact on the animals (Thompson et al. 2020). Key objectives of the validation are to characterise the soft start, the sound amplitude levels as a function of range from the source, the frequency spectra of the pile strikes, and to compare these to the modelled predictions and estimated marine mammal impact ranges. Validation of the

noise impact assessment model predictions is a standard licence condition imposed by the Marine Management Organisation (MMO) in Development Consent Orders (DCOs) / Deemed Marine Licences (DMLs) for OWF projects in English waters (e.g., MMO 2014, Statutory Instruments 2016). Typically, monitoring of the noise emitted during the installation of the first four driven or part-driven piles of each discrete foundation type is required. Noise monitoring is also conducted during the installation of driven pile foundations in Scottish OWFs (e.g., Beatrice and Moray East) in the context of their approved PS and Project Environmental Monitoring Plans – documents required under conditions of their Section 36 Consents and Marine Licences granted by the Marine Scotland Licensing Operations Team (MS-LOT) (e.g., Bellew 2017, Moray Offshore Windfarm (East) Limited 2018, Moray Offshore Windfarm (East) Limited 2019).

These licensing requirements have led to a large body of data that could help better inform pre-consenting acoustic modelling. However, crucially for this project, many noise measurements conducted during OWF construction were done using different measurement approaches and will have differed significantly with respect to measurement distances to the sound source, noise metrics and minimum requirement of documentation.

1.6. Study aim

The understanding of key factors influencing the noise levels emitted during percussive piling, and thereby influencing the predicted impact ranges, will allow for a better estimate of the noise impact on marine animals. This will allow the identification of key assumptions which could be justifiably amended to reduce uncertainties in noise modelling used to inform EIA, which in turn will support a better decision making on the mitigation measures required in the frame of MMMPs and PSs.

In the frame of this study, the following steps were undertaken:

- Modelling reports were reviewed detailing noise impact assessments for OWF piling noise to understand the modelling approaches used in the UK EIA process;
- Field measurements (real world data) acquired across a variety of different OWFs during the installation of pin piles and monopiles by using percussive pile-driving were analysed in a standardised manner according to ISO 18406 to compare these measured data from different projects and to investigate the influence of different factors on the emitted piling noise by using a cross-project analysis;
- Where data and model predictions were available, the noise levels observed in the field were compared with those predicted in the impact assessment;
- To understand the correlation between each identified key factor and the predicted noise levels a sensitivity analysis was conducted, changing one key factor at a time on a given piling scenario.

2. Methods

2.1. Data availability

2.1.1. Modelling reports

Piling noise modelling reports and the associated marine mammal impact assessments from nine OWFs constructed in UK waters were reviewed[1], including the six UK OWF for which piling noise measurement data were available (see section 2.1.2). These modelling reports have been published between 2005 and 2017 for the initial Environmental Statement (ES) or as a post-consent document to inform the MMMP or PS. Information gathering has been complemented by reviewing the physical processes chapter of the ES, publications, and OWF websites.

2.1.2. Field measurements

Measurements of piling noise during foundation construction between 2012 and 2020 using percussive pile driving activities were identified and available for analysis from 13 OWFs located in the UK, German, or Dutch North Sea (Figure 1, Table 1). These measurements were done under a broad range of site-specific factors like water depth and project specific factors like pile type and blow energy. For 11 OWFs, noise measurements at 750 m from the pile site were available. These were included in a cross-project analysis as described below. For the six UK OWFs, noise measurements were compared with modelling results.

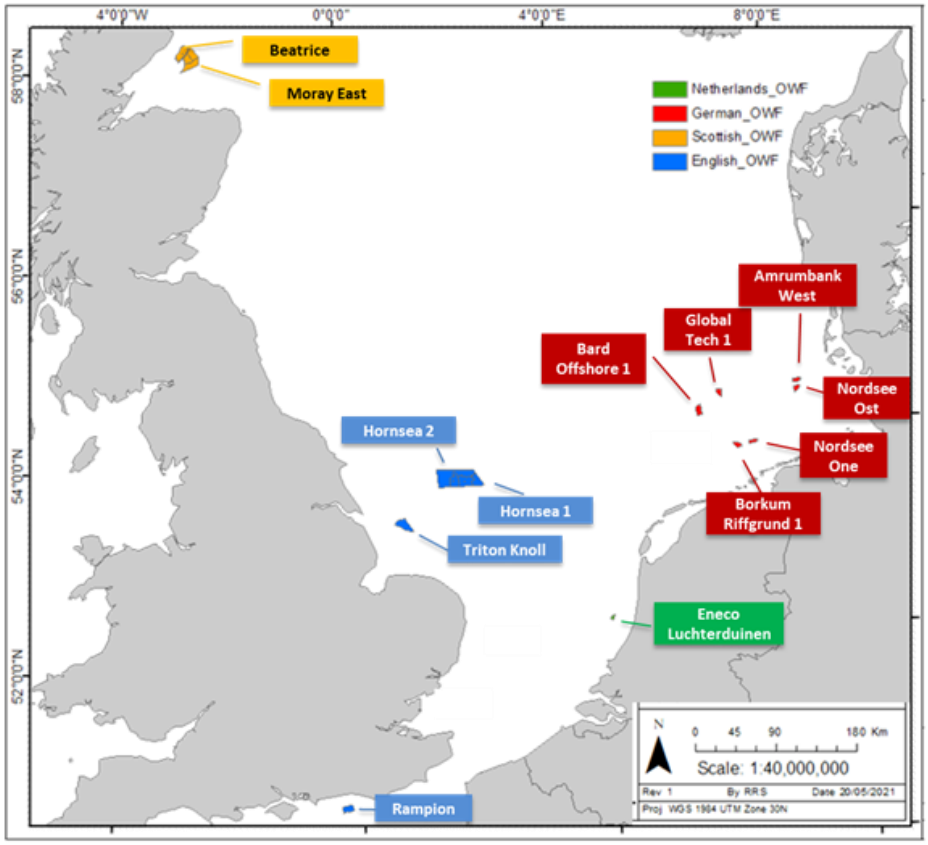

Figure 1: OWS included in this study. Further details please see Table 1.

Table 1: OWFs for which piling noise measurements were suitable for the cross-project analysis and modelling comparison. The table gives information on the data that were available to the project, i.e., information on pile type, water depth at the OWF site, the number of foundations monitored and the related noise monitoring positions per foundation and their relative distance to the pile site. Noise monitoring positions were either at a fixed position throughout the installation of all piles from an OWF (fixed) or were moved in-between piling of different OWF foundations (mobile).

OWF	Country	Pile-type	Cross-project analysis	Modelling comparison	Water depth (m)	Year of construction	# Datasets included from pilings with noise monitoring	# Noise monitoring positions per foundation (distance to pile site, km)
Hornsea Project Two	UK	Monopile, pin pile	✓	✓	35-41.4	2020	5 pin piles from one foundation, 4 monopiles	3 mobile (0.75, 1.5, 12)
Triton Knoll	UK	Monopile	✓	✓	20	2020	3	2 mobile (0.5, 0.75 m), 1 fixed (between 0.58 and 1.4)
Moray East	UK	Pin pile	×	✓	35-55	2019	5	1 fixed (between 3.2 and 6.7)
Hornsea Project One	UK	Monopile	✓	✓	20	2018	5	4 mobile (0.75, 1.5, 3, 10)
Beatrice	UK	Pin pile	×	✓	38-60	2017	4	2 fixed (between 0.8 and 11)
Rampion	UK	Monopile	✓	✓	20-37	2016	7	1 mobile (0.75), for one foundation 3 mobile (0.25, 0.75, 1.5)
Nordsee One	GER	Monopile	✓	×	28	2015	2	3 mobile (0.75)
Borkum Riffgrund 1	GER	Monopile	✓	×	26.8	2014	1	13 fixed (between 0.25 and 5)
Amrumbank West	GER	Monopile	✓	×	20	2014	1	2 mobile (0.75)
Eneco Luchterduine	NL	Monopile	✓	×	21.8	2014	1	1 mobile (0.75)
Nordsee Ost	GER	Pin pile	✓	×	25	2013	4	4 mobile (0.75 in different directions)
London Array	UK	Monopile	✓	×	11, 23	2012	2	1 mobile (0.75)
Global Tech I	GER	Pin pile	✓	×	40	2012	3	4 mobile (0.75 in different directions)

2.2. Review of modelling reports

Documents have been reviewed to obtain the following information:

- Project specific construction details, such as water depth, pile type, pile diameter, hammer type, maximum consented blow energy and number of foundations accompanied with noise monitoring during installation,
- Noise modelling approaches used for impact predictions, including impact parameters, and executing company,
- Environmental parameters such as seabed properties at the OWF project area,
- Impact assessment methodologies and noise criteria used to model marine mammal impact ranges.

2.3. Cross-project analysis

2.3.1. Data processing

For the cross-project analysis, the following noise metrics according to ISO (2017a) and ISO (2017b) were obtained for each single pile strike with a signal-to-noise ratio of greater than 6 dB (see appendix for definitions):

- Zero-to-peak SPL – $L_{p,pk}$,
- 90 percent energy signal duration – τ_{90}
- Single strike SEL - SEL_{ss},
- One-third octave band frequency spectrum

For measured data sets of OWFs, for which SEL_{cum} was modelled in the impact assessment. The SEL_{cum} as defined by Energistyrelsen (2022) was analysed for comparison.

The blow energy used at the time of recording a pile strike was allocated to each set of single strike noise metrics. Information on the blow energy was obtained from the corresponding hammer logs. In most cases, such logs do not save information for each pile strike, but rather present the number of piles strikes per 25 cm with a corresponding blow energy. Linear 1dimensional interpolation and extrapolation of the hammer logs was therefore used to estimate blow energy for each pile driving stroke.

Next to blow energy, the current penetration depth of the pile into the soil as well as the related number of strikes used per 25 cm penetration (blow count) was allocated to each pile strike. This information was only available for seven OWFs (six monopile, one pin pile installation).

Furthermore, pile type (monopile or pin pile), pile diameter and water depth at the pile site or, where missing, of the OWF project area, were saved with each pile strike data set, along with the distance of the measuring position to the pile site, and a unique ID for each OWF, pile site, pile and measuring position, respectively. For the cross-project analysis, only measurements from 750 m distance to the pile site were included in the analysis to investigate the effect of different variables on the SEL_{SS}. Furthermore, pile strikes were divided into the soft start or installation phase of the piling sequence. Soft start was defined as the first 100 strikes

of a piling sequence, or until 40% of the maximum blow energy was reached, whatever came first. Thereafter, pile strikes belonged to the installation part of the piling sequence. SEL_{SS} values < 160 re 1µPa^{2}s (regression analysis) or < 150 dB re 1µPa^{2}s (mixed models) and blow energies < 200 kJ (regression analysis) or 150 kJ (mixed models) were excluded as likely measurement errors. The rationale for this is that SEL_{SS} values below 150 dB might have a poor signal-to-noise ratio and might be affected by background noise, whereas single strikes of blow energies less than 200 kJ might still be part of the soft start, though not within the first 100 strikes.

One of two pile installations at the London Array offshore wind farm was installed in 11 m water depth. In such shallow waters, unimpeded sound propagation of piling noise is impossible for frequencies below 76 Hz (Urick 1983). This leads to significantly lower sound inputs levels, compared to the ranges measured in deeper waters, justifying the exclusion of this installation from the cross-project analysis.

Table 2 details the range of water depth, pile diameters and maximum blow energies from the OWFs included in the cross-project analysis. Sound recordings of a total of 36 pile installations (24 monopiles and 12 pin piles) were successfully processed, resulting in data from 170,799 single strikes.

Table 2: Number of OWF projects and pile types for which noise monitoring data were processed and included in the cross-project analysis. Ranges in water depth, pile diameter and maximum blow energies are provided, as well as the number of foundations and piles. For the subset of data that contain penetration depth information these numbers are given in brackets, as this subset was used in a Mixed Model analysis.

Pile type	# OWF	Water depth at foundations (m)		Pile diameter (m)		Max blow energy (kJ)		# Noise recordings from	
		from	to	from	to	from	to	foundations	piles
Monopile	9 (6)	20 (20)	41 (27)	5 (5)	9.5 (6.5)	478 (830)	2,893 (1,905)	24 (13)	24 (13)
Pin pile	3 (1)	25 (25)	40 (25)	2.4 (2.4)	2.5 (2.4)	597 (842)	2,545 (842)	4(1)	12 (4)

2.3.2. Statistical analysis

Regression analysis

In order to identify the most influential factors on piling noise, a pairwise comparison of the SEL_{SS} and $L_{p,pk}$ has been conducted with:

- τ_{90}.
- Blow energy.
- Pile diameter.
- Water depth.
- Penetration depth.
- Phase of piling sequence.

In this comparison a linear and a logarithmic regression model was estimated for each parameter via ordinary least squared method (OLS). Subsequently, using the regression function thus obtained, the SEL_{SS} or $L_{p,pk}$ was normalised to a fixed value of the respective parameter. The parameters that reduced the 90% confidence interval of the SEL_{SS} or $L_{p,pk}$ were then included in a combined regression model, resulting in Equation 2 (section 3.2.2.5), which was then used for re-modelling the given scenarios for comparison with the measurements (see section 2.4.1).

Spectrum analysis

Frequency spectra of each pile strike pulse were normalised to a broadband level of 0 dB and compared separately for monopiles and pin piles. Due to a high degree of similarity of the spectra, a more in-depth investigation was deemed unnecessary.

Mixed Models

Generalised linear mixed-effect models (GLMMs) were used to characterize SEL_{SS} across the included windfarm locations. This process necessitated further data cleaning to remove rank deficient variables and outliers. Like in the regression analysis, pile strikes from the soft start phase were excluded from the GLMMs. The number of blows required to submerge a pile 25 cm was also included as a proxy for sediment density. The SEL_{SS} of a few piles were measured in different directions from the pile, i.e., at multiple measurement locations. To differentiate between these in the statistical model, a new covariate that combined both pile ID and measurement location ID was created (termed 'PileMeasure'). For the pin pile analysis, only one windfarm was surveyed and as such the 'PileMeasure' consisted only of the different pin piles within that windfarm.

Given the methodological differences between the installation of pin piles and monopiles, an independent model was built for each pile type. The most parsimonious GLMM models for each pile type (monopile and pin pile) was selected using a combination of Akaike's information criterion (AIC) and visualization of the model covariates. Final models were assessed for multi-collinearity using generalized VIF. Any parameter with VIF scores greater than 4 were excluded. Linear mixed effect models were created using the LMER4 package in R (Bates et al. 2014, R Core Team 2021).

For pin piles, candidate GLMM models consisted of various combinations of known predictive parameters of SEL_{SS} including blow energy, blow count, the individual piles, and penetration depth. While these variables are all important in understanding the SEL_{ss} at a given range, the order of importance and the shape of the relationship between covariates and SEL is less well understood. For this reason, we investigated both linear and polynomial versions of these parameters. We included the pile name and measurement as a random effect (PileMeasure) to provide a more generalized model for each location. This accounted for variation in physical parameters between the piling locations that affect sound propagation but for which no data were available (e.g., sediment type or other site-specific activities). AIC selection indicated that Model 3 containing polynomial effects for all parameters was the best fit (Table 3). There was only limited indication (variance inflation factors VIF =3.4) of collinearity and as such all model terms were retained.

Table 3: Linear mixed models tested on the pin pile data. Number of parameters (Npar), AIC, Chisq (Chi squared) and Chi squared probabilities (Pr(>Chisq)) are provided. In all models the windfarm name and measurement were included as random effects. 'Poly' terms indicate nth order polynomials fit to the data where n is in parentheses. Random effect terms are listed as 're'.

Model Number	Formula	Npar	AIC	Chisq	Pr(>Chisq)
1	penetration_depth + blow_energy blow_count+pilename + (re: PileMeasure)	9			
2	poly(penetration_depth, 2) + poly(blow_energy, 2) + poly(blow_count, 2) + PileMeasure + (re: PileMeasure)	12	105,399	1,491	< 2.2e-16
3	poly(penetration_depth, 3) + poly(blow_energy, 2) + poly(blow_count, 2) + PileMeasure + (re: PileMeasure)	13	102,802	2,599	< 2.2e-16
4	penetration_depth+ poly(blow_energy, 2) + poly(blow_count, 2) + PileMeasure + (re: PileMeasure)	11	106,888	1,151	< 2.2e-16

As with the pin pile data, for the monopiles, several models were initially compared using a Chi squared test and AIC scores (Table 4). For the modelling approach, we again selected various combinations of the known predictive factors of SEL_{SS}; including blow energy, blow count, the individual piles, the penetration depth, and pile diameter. Random effects of pile measure were nested within the windfarm.

The ability of the selected models to generalize was investigated using k-fold cross validation. In this process, data were divided into a train and a test portion. The selected models are re-fit using the training section and evaluated against the data that were held out for the evaluation section. This process is repeated *k* times, and the average performance of the selected model is reported. For the monopiles, the train/test split was based on windfarms wherein one windfarm was held out for the test and data from the remaining windfarms formed the test set. As data for pin piling were collected from a single site, the train/test split was based on the 'PileMeasure' variable. Model error is defined as the difference between the model predicted value and the observed value for the given covariates.

Table 4: Linear mixed models tested on the mono pile data. Number of parameters (Npar), AIC, Chisq (Chi squared) and Chi squared probabilities (Pr(>Chisq)) are provided. In all models the windfarm name and measurement were included as random effects. 'Poly' terms indicate nth order polynomials fit to the data where n is in parentheses. Random effect terms are listed as 're'. In this application pile measure was a random effect within each windfarm.

Model Number	Formula	Npar	AIC	Chisq	Pr(>Chisq)
1	penetration_depth + blow_energy blow_count + Windfarm + diameter_m + (re: Windfarm:PileMeasure)	12	113,185.02		

Model Number	Formula	Npar	AIC	Chisq	Pr(>Chisq)
2	poly(penetration_depth, 2) + poly(blow_energy, 2) + poly(blow_count, 2) + Windfarm + diameter_m + (re: Windfarm:PileMeasure)	15	102,806.15	10,385	< 2.2e-16
3	poly(penetration_depth, 3) + poly(blow_energy, 3) + poly(blow_count, 3) + Windfarm + diameter_m + (re:Windfarm:PileMeasure)	18	98,606.72	4,205	< 2.2e-16
4	poly(penetration_depth, 3) + poly(blow_energy, 4) + poly(blow_count, 4) + Windfarm + diameter_m + (re:Windfarm:PileMeasure)	21	97,939.37	673	< 2.2e-16

2.4. Comparison of measured data with modelling results

2.4.1. Single strike sound exposure level

Comparison of pre-construction model results with measurement results was complicated by model predictions and field measurements not occurring at matching distances from pile driving activity. For the comparison of the field data with the modelling results, we have, therefore:

- Compiled modelled SEL_{SS} versus distance data pairs by using the impact ranges, estimated for the various noise thresholds used in the noise impact assessment, and as detailed in the reviewed modelling reports;
- Used SEL_{SS} versus distance data pairs from the noise measurements analysed for this study;
- Estimated the transmission loss for the measured data based on a frequency dependent empirical model from Thiele and Schellstede (1980);
- Estimated (remodelled) source level by using Equation 2 (see section 3.2.2.5) and transmission loss with Thiele and Schellstede (1980).

To retrieve modelled SEL_{SS} versus distance data pairs, impact ranges for noise thresholds for auditory injury and behavioural responses (e.g., Southall et al. 2007, Lucke et al. 2009, National Marine Fisheries Service 2018) as estimated in the modelling reports were used. The use of the estimated impact ranges reveals which sound level values were achieved at different distances to the pile site. One uncertainty, however, is that some of the impact ranges given were for frequency weighted SEL_{SS} (e.g., Southall et al. 2007, National Marine Fisheries Service 2018), and these had to be back calculated to unweighted SEL_{SS} without knowledge of the assumed spectrum. In these cases, the median third octave spectrum from the spectrum analysis was considered as obtained from the measured data of this study (see Figure 2).

SEL_{SS} versus distance data pairs were taken from all measurement positions available for each OWF. These data pairs were then filtered for a blow energy range that fit the blow energy modelled. At some measurement positions, piling noise was not recorded continuously, and filtering for a certain blow energy

range would not result in enough data to compare with. In such instances, all data pairs available were used for the comparison regardless of the blow energy.

Based on the measured (and filtered) data pairs, where possible, the transmission loss was estimated based on a logarithmic function curve using the ordinary least squared method:

$$SEL = k \log_{10}(R) + \alpha R + b \qquad \text{Equation 1}$$

with

k = constant

R = distance to sound source (m)

α = absorption coefficient

b = offset

For comparison, an estimate of the SEL_{SS} versus distance was re-modelled for the blow energy as used in the pre-construction modelling using the curve fitting formula resulting from the measured data of this study (Equation 2, section 3.2.2.5), considering a transmission loss function according to Thiele and Schellstede (1980) for coastal regions in the North Sea under fair weather conditions.

All data were then plotted for comparison.

2.4.2. Cumulative sound exposure levels

SEL_{cum} is now regularly used in modelling as a threshold criterion for auditory injury. One of the reviewed modelling reports included SEL_{cum} impact ranges and sufficient information to allow for a comparison with the measured data. The calculated SEL_{cum} impact ranges usually refer to a moving receiver. For a comparison of modelled estimates with those based on measurements, details of the piling sequence were sourced in the modelling (blow rate, course of the blow energy and duration of the piling sequence). Furthermore, the modelled SEL_{SS} versus distance relationship was needed.

From the measured data, the same information was required to calculate the SEL_{cum}.

2.5. Sensitivity analysis

2.5.1. Single strike sound exposure level

To understand how much a change in the influencing parameters blow energy, pile diameter and water depth would have, a standard piling scenario was defined and the SEL_{SS} was modelled based on Equation 2 and one parameter at a time was changed, while the other two factors were kept constant.

The standard piling scenario was defined as a pile diameter of 5 m, blow energy of 2,500 kJ at a water depth of 30 m. Pile diameter changes ranged from 1 m to 10 m, blow energy from 500 to 5,500 kJ and water depth from 10 m to 60 m based on the OWF projects of this study and currently planned OWF projects for the near future.

2.5.2. Cumulative sound exposure level

The resulting impact ranges based on SEL_{cum} are determined by the piling sequence and the frequency spectrum of the single strikes. For modelling, a certain piling sequence is assumed with a defined number of single strikes, sequence of the blow energy and blow rate, soft start and ramp up duration. Furthermore, the pile diameter is assumed as well as the frequency spectrum of the blow, which is likely based on published spectra or archived recordings. To understand the influence of changes in such assumptions, these parameters were included in the sensitivity analysis. The influence of the SEL_{SS} during soft start was also investigated, as, in contrast to the installation phase, measured data showed a poor correlation between blow energy and the SEL_{SS} during the soft start. To understand the effect of this phenomenon, the effect of varying SEL_{SS} for the soft start on the resulting impact ranges was investigated.

For the calculation of the SEL_{cum}, the SEL received by an animal is considered. Marine mammals have been shown to flee from a piling site (e.g., Tougaard et al. 2009, Dähne et al. 2013), and modelling of SEL_{cum} impact ranges is usually based on a fleeing animal model. As the fleeing speed of an animal influences the sound levels it receives, and thereby the resulting SEL_{cum}, the effect of fleeing speed was also included into the investigations. For modelling, a simplified fleeing animal model was applied, assuming a straight swim trajectory directly away from the pile site.

We defined standard scenarios to investigate the influence of the:

a) Frequency spectrum.

b) Total number of single strikes in a piling sequence.

c) Blow energy.

d) Pile diameter.

e) Soft start duration.

f) Blow rate during soft start, and

g) SEL_{SS} during the soft start.

h) Animal fleeing speed on the resulting impact ranges.

Impact ranges were calculated based on the most recent noise thresholds for auditory injury (permanent threshold shift (PTS) by Southall et al. (2019).

Standard scenario for investigating the influence of factors a) to d) was as follows: an 8 m monopile at 40 m water depth with a standard frequency spectrum based on the measured data (Figure 2), 5,000 single strikes of a constant blow energy of 4,000 kJ, a blow rate of 30 strikes per minute and a fleeing speed of the animal of 1.5 m per second.

Factors a) to d) were changed as follows:

a) 1/3rd octave spectrum, which was shifted along the frequency axis from -1 octave to + 1 octave (see Figure 2);

b) Number of pile strikes from 1 to 10,000;

c) Blow energy from 200 kJ to 5,000 kJ;

d) Pile diameter from 3 m to 15 m.

The standard scenario for investigating the influence of factors e) to h) included a soft start of 20 minutes with 1 pile strike per minute at 400 kJ blow energy, a 20-minute ramp up with a stepwise increase in blow energy from 400 kJ to 4,000 kJ in 400 kJ steps every 2 minutes, followed by 5,000 pile strikes of 4,000 kJ and a blow rate of 30 strikes per minute starting with ramp up. For comparison, the same scenarios were re-modelled with a blow rate of 28 strikes per minute during soft start.

Factors were changed as follows:

a) Soft start duration from 0 to 60 minutes;

b) Changes in SEL_{SS} within the soft start from 0 dB up to 20 dB increase in standard SEL_{SS};

c) Blow rate from 1 pile strike/min to 29 pile strikes/min;

d) Fleeing speed from 0 m/s up to 5 m/s.

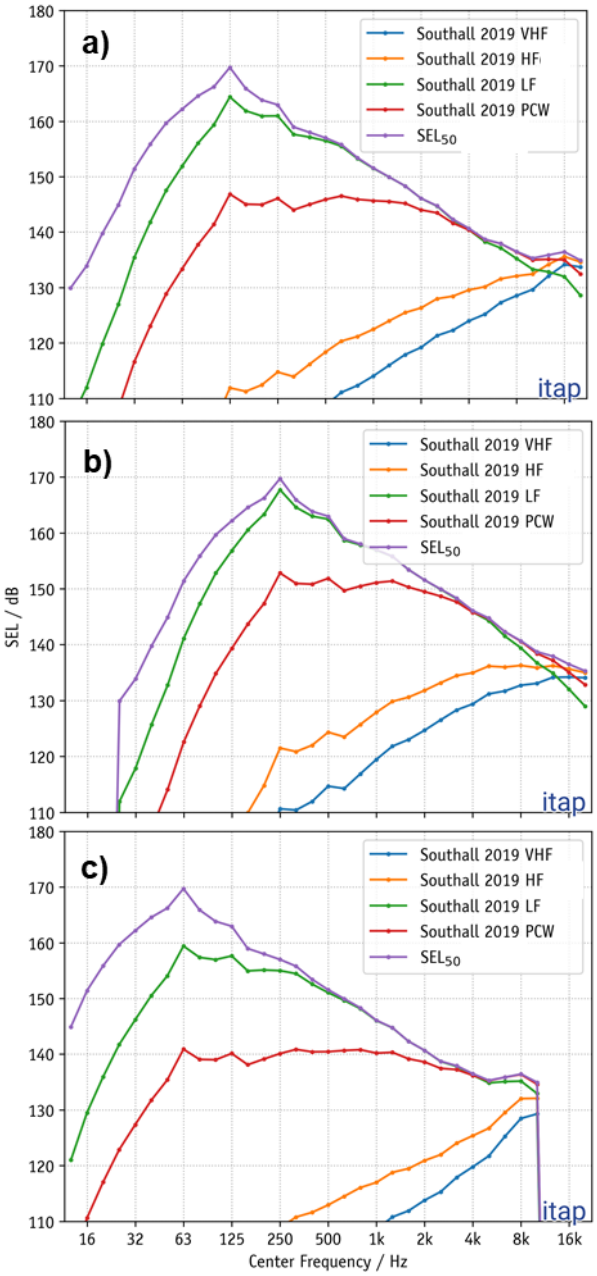

Figure 2: a) Median frequency spectrum of the SELSS measured at 750 m distance during monopile installation (a), unweighted (purple) and weighted with Southall et al. (2019) weighting functions for very high cetacean (VHF, blue), high frequency cetacean (VHF, blue) high frequency cetacean (HF, orange), low frequency cetacean (LF, green) and phocid seals in water (PCW, red, shifted by +1 octave (b) and -1octave (c)

3. Results and Discussion

3.1. Review of noise impact assessment reports

3.1.1. Modelling approaches

The underwater noise modelling for the impact assessments was conducted by four different noise modelling companies. While each company used a different type of model, all models have a two-step approach: the use of an empirical model to determine the sound amplitude level at the source (source level), and numerical or semi-empirical model approaches to calculate transmission loss (Figure 5).

A short review of the different types of used models is summarised as follow:

- Empirical models to calculate source levels are often proprietary and based on real world measurement data conducted by the respective company for comparable pile-driving projects. These models might be further informed by published literature. A best estimate fit is then performed based on the collated information.
- The numerical models to calculate transmission loss are the energy flux model from Weston (1976) and the parabolic equation model from Farcas et al. (2016). These separate an environment into discrete spatial points. The input at each spatial point depends on the output of the previous spatial point. The parabolic equation method solves the Helmholtz equation, i.e., the time-independent wave equation. The acoustic field is separated into incoming and outgoing waves. In the Parabolic Equation model, only the outgoing waves are used, the incoming waves (back-scattered energy) are therefore neglected. The energy flux method is a hybrid method based on rays and modes. This method was developed by Weston (Weston 1959, 1968) and divides the acoustic field into four regions. Spherical spreading near the source followed by cylindrical spreading further away. The third region is called the mode stripping region which eventually leads to single mode propagation (fourth region) far from the source. This energy flux method is very fast and can handle range-dependent environments.
- The semi-empirical models to estimate transmission loss are based on geometric propagation functions. The semi-empirical model ‘INSPIRE’ considers the geometrical spreading of sound from the source, the absorption of the sound by the seawater and seabed and the bathymetry between the source and receiver positions. It is informed by an increasing set of sound propagation transect data sampled by the company using this model at various piling locations in UK waters. This method models broadband noise and is therefore frequency independent. Thiele and Schellstede (1980) is a frequency-dependent approach, estimating transfer functions for different regions in the North Sea based on far range underwater noise measurements during detonations.

Table 5: Models used for underwater noise modelling during the consenting process of selected UK OWFs. The table also summarises the model input parameters, mainly as found in the corresponding documents. Documents reviewed were noise impact assessment reports for the Environmental Statement (ES), the Piling Strategy (PS), the Marine Mammal Mitigation Protocol (MMMP) or peer-reviewed publications post construction.

OWF	Purpose	References	Near Field/Source	Far Field	Input parameter
Beatrice	ES	BOWL (2012a), Subacoustech Environmental Ltd (2012a)	Empirical model (SPEAR)	Semi-empirical (INSPIRE)	Pile diameter, water depth
	PS	BOWL (2015)	Empirical model (proprietary)	Semi-empirical (INSPIRE)	No information available
			Empirical model (De Jong and Ainslie 2008)	Spherical spreading loss equation	
	Post-construction publication	Thompson et al. (2020)	Empirical model (De Jong and Ainslie 2008)	Parabolic equation (Farcas et al. 2016)	Blow energy, pile diameter, energy efficiency, sound speed based on salinity & temperature, density, water depth
East Anglia One	ES	Scottish Power Renewables (2012c), Scottish Power Renewables (2012b)	Empirical model (De Jong and Ainslie 2008)	Energy flux (Weston 1976)	Pile diameter, blow energy, water depth, seabed properties, surface scattering
Hornsea Project 1	ES	SMart wind Limited (2013b), Verfuss et al. (2018), Subacoustech Environmental Ltd (2017)	Empirical model (De Jong and Ainslie 2008)	Energy flux (Weston 1976)	Pile diameter, blow energy, water depth, seabed properties, surface scattering
	MMMP	SMRU Consulting (2017)	Empirical model (proprietary)	Semi-empirical (INSPIRE)	No information available
Hornsea Project 2	ES	SMart Wind Limited (2015b), itap (2020)	Empirical model (De Jong and Ainslie 2008)	Energy flux (Weston 1976)	Pile diameter, blow energy, water depth, seabed properties, surface scattering

OWF	Purpose	References	Near Field/Source	Far Field	Input parameter
	MMMP	GoBe (2020a), GoBe (2020b)	Empirical model (proprietary)	Semi-empirical (Thiele and Schellstede 1980)	Pile diameter, bathymetry, blow energy (piling profile)
London Array	ES	RPS (2005)	No information available		
Moray East	ES	Natural Power (2012)	Empirical model (proprietary)	Semi-empirical (INSPIRE)	Pile diameter, blow energy, water depth
	PS	Royal Haskoning DHV (2019)	Empirical model (De Jong and Ainslie 2008)	Parabolic equation (Farcas et al. 2016)	Blow energy, pile diameter, energy efficiency, sound speed depending on salinity & temperature, seabed density, water depth
Rampion	ES	RSK Environmental Ltd (2012b), Simpson (2016)	Empirical model (proprietary)	Semi-empirical (INSPIRE)	No information available
Triton Knoll	ES	RWE npower renewables (2012b), Subacoustech Environmental Ltd (2012b)	Empirical model (proprietary)	Semi-empirical (INSPIRE)	Pile diameter, blow energy, water depth

3.1.2. Environmental parameters

A variety of environmental factors influence the transmission loss while sound propagates through the water column. Water depth is one key factor that determines the geometrical spreading loss but may also limit the forming of plane sound waves: low frequencies have long wavelengths and, therefore, need sufficient space to propagate. But even if the space is available, the sound pressure waves cannot propagate unhindered. Every medium, in this case sea water, offers the sound a certain resistance (acoustic impedance). This depends on the density of the medium, which in turn depends on its composition and temperature. However, since a good mixing of the water can be assumed in the entire North Sea, it is expected that these parameters will not significantly influence the sound levels. Sound waves are additionally attenuated at the boundaries (water surface and sediment), and the attenuation depends strongly on the seabed properties (soil conditions and layers) and sea surface conditions (e.g., rough, or calm sea).

From the documents reviewed, seabed properties could be collated from the OWF project areas (Table 6).

Table 6: Seabed properties of the OWF project areas as detailed in the corresponding ES or PS documents.

OWF	Seabed properties	From source	Reference
Beatrice	Smith Bank is a bathymetric high in the Outer Moray Firth. The main body of the bank is relict and stable, comprising bedrock overlain by poorly sorted stiff clay till sediments, with a variably thick veneer of (occasionally shelly) marine sands and gravels. Smith Bank is therefore not a true sand bank and its overall shape and minimum water depth, etc, will therefore have negligible sensitivity to changes in sediment transport pathways. Side-scan sonar data indicate a predominance of granular surface sediments across the Wind Farm site, except in the shallowest parts near the crest of Smith Bank, where the underlying till is largely exposed with little sediment veneer. PSA data indicate that surface sediments are typically medium sands (250 to 500 µm diameter) with little (i.e., less than 5%) or no measurable content of fines (less than 63 µm). Typically, less than 3% of sediment volume is classed as gravel (greater than 2 mm). However, in 10% of locations, 10 to 20% of the sediment volume, and in a further 10% of locations, 20 to 30% of the sediment volume, may comprise gravels.	ES: Physical Processes and Geomorphology	BOWL (2012b)
East Anglia One	Seabed sediments across the East Anglia ONE site generally consist of Holocene slightly gravelly sand, with pockets of gravelly sand, sand and sandy gravel. Fine (silt and clay sized) particles are largely absent. Almost every sample contains a modal peak at approximately 375 µm (medium sand), indicating that this is the most common sediment type in this area.	ES: Chapter 6 Coastal processes	Scottish Power Renewables (2012a)
Hornsea Project 1	Within much of the southern North Sea, the seabed sediments generally form a thin veneer over Quaternary or older formations. The exceptions to this are areas of tidal sandbanks and large sand waves. Seabed sediment maps produced by the BGS suggest that the regional seabed sediment ranges between sand and gravel (BGS, 1987a; 1990). Generally, gravel-rich sediment is more common towards the coast, whereas sandy sediments are more prevalent further offshore.	ES: Section 7 Marine Processes	SMart wind Limited (2013a)
Hornsea Project 2		ES: Chapter 1 Marine Processes	SMart Wind Limited (2015a)
London Array	The seabed sediments across the survey area generally comprise sands and gravels, with the coarsest materials found within Knock Deep. Current induced sediment bedforms such as sand ripples, mega ripples and sand waves occur throughout the survey area. The sand waves are up to 5.0 m high, are generally orientated northwest to southeast and occur primarily on the slopes and crests of the sand banks. Long Sand stretches some 33 km from the North Edinburgh Channel in the south to Long Sand Head in the east. It is composed of fine to very fine sand that varies in thickness above uneven bedrock from 5 m to 40 m. Apart from its northern end, Long Sand is	ES: Chapter 6 Description of the Environment	RPS (2005)

OWF	Seabed properties	From source	Reference
	relatively stable, in particular, on its eastern side between Fisherman's Gat and 51° 44' N - the area selected for potential development. From Long Sand, the selected area extends eastwards into the Knock Deep, where the water deepens to up to 25 m and the seabed remains relatively stable. Seabed sediments comprise medium to coarse sands and gravels north of 51° 40' N, and predominantly fine sands in the middle and southern sections. Sediment thickness over bedrock varies once more with the greatest thickness being found in the Thames buried channel at the southern end.		
Moray East	The results show that within the Moray East site, the thickness of sandy marine sediments is highly variable. In the Telford Wind Farm the marine sediment veneer was found to be typically 1 to 3 m thick, increasing to 10 to 30 m in the central southern part of the site, and very thin or absent over the shallowest area in the western part of the site. In the Stevenson Wind Farm the marine sediment veneer is typically 1 to 3 m thick, but very thin or absent over the bathymetric highs in central and eastern parts of the site. In the MacColl Wind Farm the marine sediment veneer is typically 1 to 3 m thick, increasing to 5 m in the western part and 10 to 30 m at the northeastern edge of the site. The marine deposits in the three sites overlay glacial tills (compacted poorly sorted mixtures of fine and coarse material). Where the surface veneer is sufficiently thin, glacial till is exposed at the seabed surface.	Piling Strategy Geotechnical and Geophysical Survey Results	Royal Haskoning DHV (2019)
Rampion	The seabed of the east English Channel is generally interpreted as comprising a thin layer (veneer) of mobile Holocene sand and gravel overlying non-mobile lag deposits. These mixed grained deposits overlay older, more consolidated deposits of Tertiary and Cretaceous age (90-30 million years old).	ES: Section 6 Physical environment	RSK Environmental Ltd (2012a)
Triton Knoll	Sand and gravel sized material dominate the site and the wider area; 58% of samples are, according to the Folk classification, sandy Gravel (sG). Fines are only present in 0.6% of samples, although BGS sediment mapping suggests that muddy sandy Gravels (mSG) dominate within, approximately, half of the western sector.	ES: Chapter 2 Physical Processes	RWE npower renewables (2012a)

3.1.3. Impact assessment methodology

Over the years in which ESs were published, the marine mammal impact assessment criteria, and therefore the methodology on how the impact was assessed, has changed with new information becoming available over time (Verfuss et al. 2016). The absolute values of the thresholds used define the magnitude of the resulting impact estimated as does the metric that is used (e.g., peak-to-peak/zero-to-peak/root-mean-square SPL or SEL_{ss} / SEL_{cum}) and whether the sound is frequency-unweighted or weighted according to the auditory sensitivity (hearing ability) of the receptor species or species group, and which weighting function is used. The impact criteria can be used as thresholds of a fixed sound pressure or exposure level value or are

based on dose-response curves. For fixed thresholds, the area within the isopleth of a threshold is defined as the impact area, and all animals within this area are considered as potentially being impacted. Dose-response functions consider that the probability of behavioural response decreases with decreasing sound level from the pile site, and, correspondingly, the portion of animals impacted decreases accordingly. The sound metrics for the impact thresholds are usually based on single strike events, except when it comes to the SEL_{cum}. Here, the sound energy received by an animal is accumulated over the whole piling sequence and is dependent on the hearing sensitivity of the species and distance of the animal to the pile site. The piling sequence with the timing of each successive pile strike as well as the course of blow energy used, therefore, has a fundamental influence on when an impact threshold is reached in case a moving animal is considered. Most impact criteria used in the assessments reviewed were based on single strike metrics. However, the SEL_{cum} has become more prevalent in recent assessments.

All impact assessments reviewed have been published from 2012 onwards, except for the London Array (Table 7). The London Array ES was published in 2005, and the impact assessment methodology is not clearly laid out nor the impact criteria sufficiently described. While some of the 2012 ES considered and assessed the impact ranges for death and physical injury, this was not included other assessments. The assessment of auditory injury and behavioural response was included in all ES impact assessments. Marine mammal impact criteria were mostly based on fixed thresholds. A dose-response function was only developed for the Beatrice OWF.

For determining the area of potential death or physical injury, criteria proposed by Parvin et al. (2007) were used, based on fixed unweighted peak-to-peak SPL. Criteria proposed by Nedwell et al. (2007) and Southall et al. (2007) were used in earlier ES for the assessment of auditory injury and behavioural reaction in marine mammals. The approach used by Nedwell et al. (2007) is basically a generalisation of the human A-weighting approach that is used to consider the hearing sensitivity of humans. Nedwell's $dB_{(ht)}$ thresholds are based on frequency-weighted sound pressure levels, with the sound levels weighted for the hearing curves of the respective animal. The $dB_{(ht)}$ weighting is the only weighting that has not been quantified in any report or publication, is therefore not reproducible and could not be used for any comparison with measured data (section 2.4) Southall et al. (2007) propose a dual criterion for assessing the risk of auditory injury from impulsive noise (such as pile-driving noise) in marine mammals. Marine mammal species are grouped in hearing groups, for which thresholds are given in unweighted peak sound pressure level as well as weighted sound exposure levels. Whichever threshold is exceeded first is to be considered in the assessment. The Southall et al. (2007) M-weighting is species group-specific and based on the auditory sensitivity of the hearing group, with a flatter weighting function than would be obtained by an inverse audiogram function (the hearing curve). Behavioural thresholds based on Southall et al. (2007) are often derived from one of their extensive tables rating behavioural reactions of marine mammal species to anthropogenic sound with a so-called severity score, a measure of the magnitude of a reaction to the exposed sound levels, given in rms SPL.

The study of Lucke et al. (2009) provided evidence that harbour porpoises are more sensitive than estimated by Southall et al. (2007), and, therefore, the Southall thresholds were often combined with thresholds derived from the Lucke et al. (2009) paper, with the Lucke thresholds replacing those for harbour porpoise from Southall. In 2016, the National Marine Fisheries Service (NMFS) introduced an update to the Southall et al. (2007) thresholds for auditory injury (National Marine Fisheries Service 2018) with a focus on species relevant for the US jurisdiction. The M-weighting curves were replaced by curves resembling averaged hearing-group-specific inverse audiograms, and thresholds were adapted based on the most recent literature. These data were part of a comprehensive update of Southall et al. (2007) published by Southall et

al. (2019). These recent criteria were so far only used in the updated modelling post-ES and are applied in current environmental statements of OWFs in the consenting process (and therefore not part of this review).

Table 7: Overview of impact criteria as used for the impact assessment of piling noise on marine mammals.

OWF	Purpose	Year	Death/Physical injury	Auditory injury	Behavioural Reaction
Beatrice	ES	2012	Parvin et al. (2007)	Southall et al. (2007)	Nedwell et al. (2007)
	PS	2015	-	Southall et al. (2007)	Nedwell et al. (2007)
East Anglia One	ES	2012	-	Lucke et al. (2009), Southall et al. (2007)	Lucke et al. (2009), Southall et al. (2007)
Hornsea Project 1	ES	2013	-	Lucke et al. (2009) Southall et al. (2007)	Lucke et al. (2009), Southall et al. (2007)
	MMMP	2017	-	as in ES plus NMFS (2016)	-
Hornsea Project 2	ES	2015	-	Lucke et al. (2009), Southall et al. (2007)	Lucke et al. (2009), Southall et al. (2007)
	MMMP	2020	-	as in ES plus Southall et al. (2019)	-
London Array	ES	2005	-	Not clear	
Moray East	ES	2012	-	Southall et al. (2007)	Nedwell et al. (2007)
	MMMP	2016		Southall et al. (2007)	-
Rampion	ES	2012	Parvin et al. (2007)	Nedwell et al. (2007), Southall et al. (2007)	Nedwell et al. (2007)
Triton Knoll	ES	2012	Parvin et al. (2007)	Southall et al. (2007)	Nedwell et al. (2007)

3.2. Cross-project analysis

3.2.1. Soft start anomaly

The data set analysed within this study, including pin pile and monopile data, showed a good correlation between various parameters and the SEL_{SS} (discussed in the next sections), except for pile strikes at the start of a piling sequence, i.e., within the soft start, which led to the exclusion of those for further analysis. Figure 3 gives two examples of the variability in the soft start during the installation of monopiles.

In the example shown in Figure 3a, the soft start of the piling sequence begins with five single strikes that precede a continuous hammering. The SEL_{SS} of those single strikes are about 2 to 3 dB higher than the SEL_{SS} of the following pile strikes with the same minimum blow energy of 400 kJ (10% capacity of the used

hammer). In the example shown in Figure 3b, a difference of 8 dB can be seen for the first single strikes of similar blow energy at around 300 kJ.

Figure 3 demonstrates, that during soft start with approximately 10% hammer capacity, a considerable difference in the SEL_{SS} values can occur. There are possible explanations:

- The hammer technology is not sufficiently precise to determine the exact blow energy at low capacities;
- Depending on site- and project-specific factors (such as soil conditions at the sea floor surface, pile, and hammer interaction at the start of piling, etc.), the penetration of the pile into the sediment is either more or less successful, defining the ratio of blow energy being converted into kinetic energy and underwater sound energy;
- The stability of the pile during the penetration process may also influence how much of the hammer energy is converted into underwater sound energy. At small penetration depths during soft start, the pile may be looser than during advanced piling, leading to a varying amount of hammer energy being converted to underwater sound energy.

These examples show that the uncertainty of the SEL_{ss} during soft start is much higher than during installation piling with higher hammer capacities. Which parameters influence these anomalies must be investigated in further detail and require a more detailed monitoring of the piling procedure during soft start.

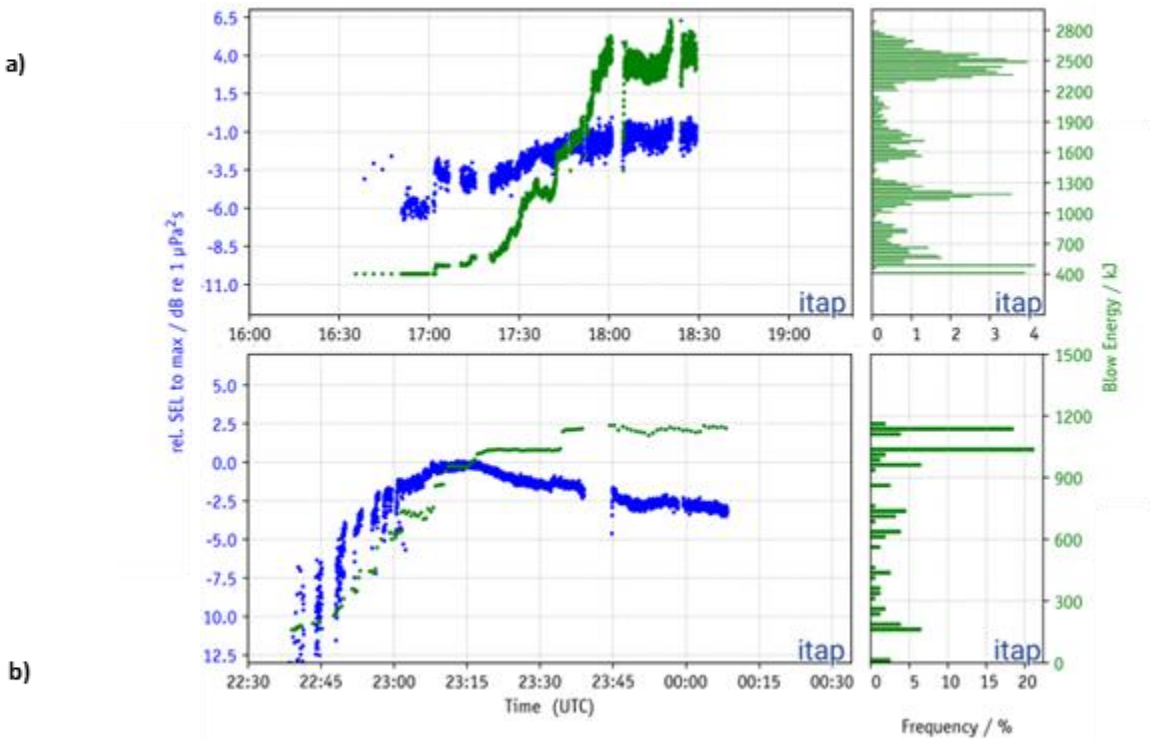

Figure 3: Temporal course of the SELSS (blue) and the blow energy (green) during two monopile installations (a, top and b, bottom) measured at approx. 750 m distance, showing unexpected high SELSS for the first five single strikes of the soft start (a) and a high SELSEE and a high variance in SEL_{SS} during soft start (b). The distribution of the blow energy is shown in a histogram (graphs on the right). Deviations of the times between blow energy (derived from hammer logs) and underwater noise measurement of a few minutes are possible due to the missing temporal synchronization

3.2.2. Regression analysis

Considering all single strikes (excluding soft start) of this study measured at ~750 m distance, sound level metrics $L_{p,pk}$ and SEL_{SS} showed a high linear positive correlation (r^2= 0.905, p≤ 0.0001). The $L_{p,pk}$ is 21.5 dB (+/- 3.5 dB) higher than the SEL_{SS}. This is in accordance with findings from Bellmann et al. (2020). Due to the high correlation between $L_{p,pk}$ and SEL_{SS}, further investigations focus on the SEL_{SS} only.

A regression analysis was conducted with the remaining parameters and the SEL_{SS}. For water depth and pile diameter only, a correlation with SEL_{SS} was not conducted for pin piles because three data sets for pin piles resulted in insufficient range of water depths and pile diameters. From the parameters considered, blow energy and penetration depth had the greatest influence on the SEL_{SS} (Table 8). However, the correlation of each parameter differs a lot from pile to pile. For the curve fitting, each parameter listed in Table 8 was included one by one, and only if the subsequent normalisation resulted in a reduction of the variability, this parameter was considered in the curve fitting. The influence of these parameters and the curve fitting are discussed below.

Although penetration depth shows a strong negative correlation with the SEL_{SS} (Table 8) for some of the pile installations, we only had information on penetration depth for around half of the data sets (Table 2). For pin piles, penetration depth information was only available for one OWF. Visual inspection of the pin pile data set shows a considerable influence of penetration depth on the SEL_{SS} (see below). Therefore, the influence of penetration depth was further investigated in the Mixed Model statistics (section 2.3.2.3). While information on the ground conditions were sourced in the ES documents (Table 6), no appropriate information was available to include ground condition as parameter. We found indications that the number of strikes per meter penetration depths might be a good proxy for ground conditions, seemingly having an influence on the SEL_{SS} (see below), which we therefore also included into the Mixed Model statistics (section 2.3.2.3).

Table 8: Results of the regression analysis correlating the SELSS with different parameters. The table provides minimum and maximum correlation coefficients r2 and significance values p for the correlation analysis of each parameter with SELss measured at 750 m distance from the pile site (excluding soft start), separated by pin pile and monopile installation.

	Pin pile		Monopile	
Parameter	**r^2**	**p-value**	**r^2**	**p-value**
$L_{p,pk}$	0.849 – 0.874	<0.001	0.205 - 0.946	<0.001
τ_{90}	0.199 - 0.477	<0.001	0.001 - 0.454	<0.001 - 0.440
Blow Energy	0.094 - 0.151	<0.001	0.000 - 0.955	<0.001 - 0.975
Pile diameter	-	-	0.219	<0.001
Penetration depth	0.842	<0.001	0.400 - 0.977	<0.001-1.000
Water depth	-	-	0.842	<0.001

Blow energy

For monopile installation, blow energy has a logarithmic relationship with the SEL_{SS} (Figure 4). The correlation coefficient including all monopile installations is r^2=0.283, with p<0.001. Considering each

monopile installation separately, r^2 ranges from 0.000 to 0.955, with p-values from <0.001 to 0.975 (Table 8) indicating a large difference between monopiles.

For pin piles, penetration depth has a larger influence than blow energy (Table 8, see also below and section 3.2.2.4), which may be the reason for a larger variation of the SEL_{SS} compared to the variation observed in relation to the maximum blow energy used during installation (Figure 5). The larger variation in SEL_{SS} for pin piles compared to monopiles may also be caused by the vibration of the jacket or tripod foundation during piling due to coupling effects.

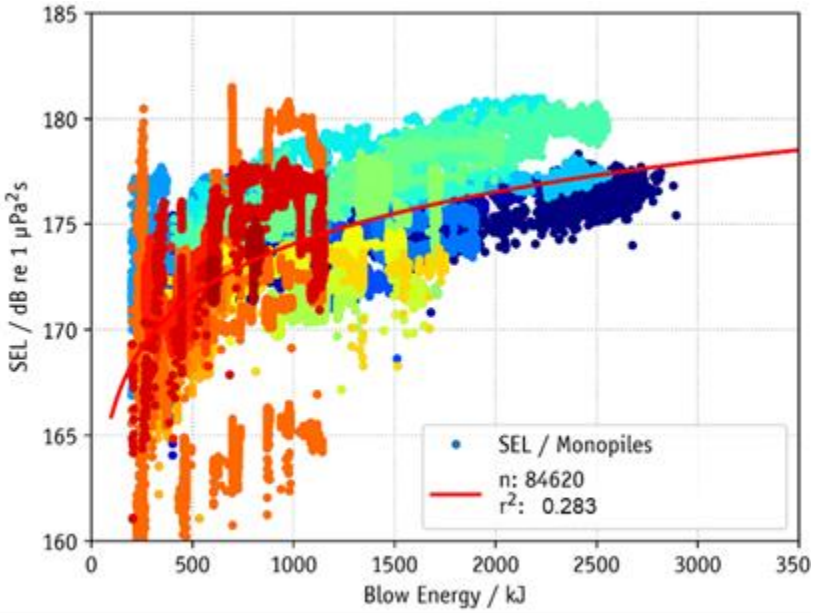

Figure 4: Comparison of the SELSS measured at 750 m distance from the pile against the Blow Energy. The dots show the Sound Exposure Level of 84,620 different single strikes from 25 monopile installations (indicated by different colours). The red line shows a logarithmic regression curve based on equation 2 (section 3.2.2) according to the ordinary least squared method (OLS)

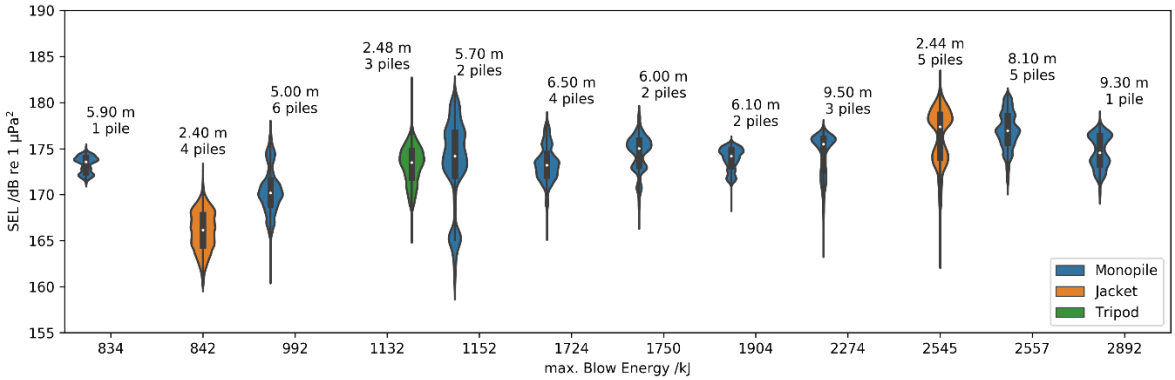

Figure 5: Violin plot of the distribution of the SELSS including all data of this study measured at 750 m. Data were pooled by pile diameter and plotted against the maximum blow energy of the pile installations. The width of the shape presents the univariate distribution of the SEL_{SS}, the black rectangles inside the violin shapes mark the interquartile range (75% to 25% exceedance) and the white dot the median Sound Exposure Level (SEL50). The different colours represent different foundation types. Pin piles are used for Jacket and Tripod installations.

Pile diameter

Bellmann et al. (2020) showed that the 5% exceedance level (95th percentile, i.e., level exceeded by 5% of all measured levels during a pile strike sequence) of the SEL_{SS} and the maximum $L_{p,pk}$ increase with increasing pile diameter (Figure 6). The data sets analysed in this project fit well within the regression ±5 dB tolerance range estimated by Bellmann et al. (2020), except for the piles with 2.44 m and 2.48 m diameter. One possible reason could be that the data from Bellmann et al. (2020) were all measured in German waters, where a limitation of the maximum blow energy is prescribed by the authorities; Since 2015, a so-called noise optimized piling procedure is requested by the BSH, meaning that the piling activity shall start with a soft-start (10% hammer capacity) and operators are only allowed to increase the blow energy if pile refusal is expected. Furthermore, piling permits issued by the BSH set a limitation to the blow energy that can be used for the first pile installations. If the noise emission of the first installation does not exceed the German mandatory noise limit, the blow energy does not need to be limited during further installations as long as the noise limit is not exceeded. The limitation of the maximum blow energy will be informed by the underwater noise prognosis and the latest pile-driving analysis of the respective OWF. In principle, the bigger the pile diameter the bigger the maximum blow energy allowed. For Pin piles the maximum blow energy ranges mostly between 1,500 and 2,000 kJ. For bigger Monopiles the max blow energy can be slightly higher but was always below 2,500 kJ.

However, the blow energy for the jacket installations included in the cross-project analysis from the UK was higher than this limit. The monopile installations in the UK and Germany used comparable blow energies.

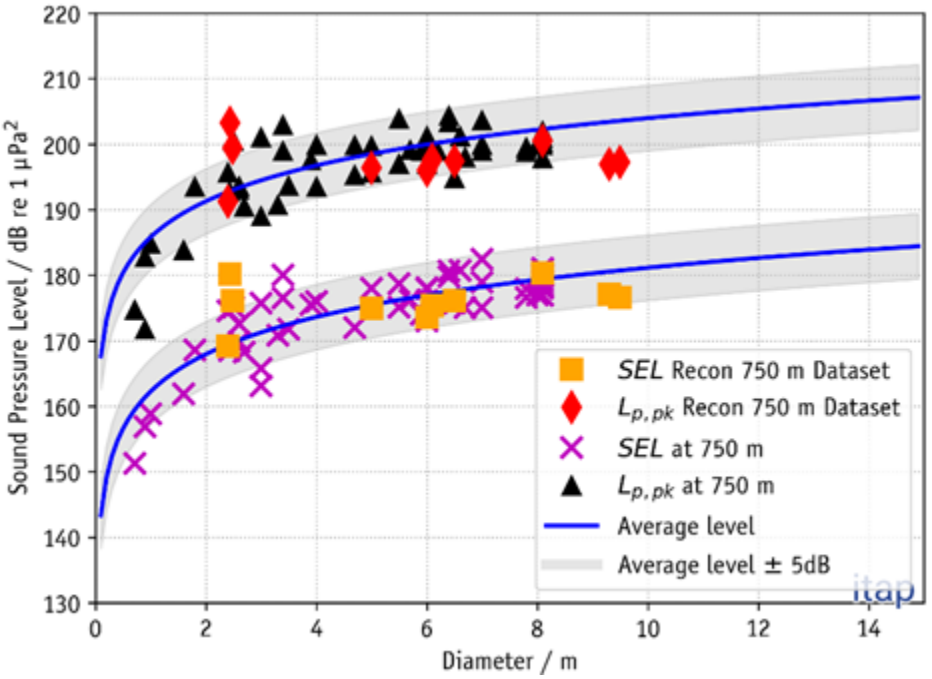

Figure 6: Measured 5% exceedance levels of SELSS and maximum level of Lp,pk across pile diameter from the current dataset compared with data from Figure 5 in Bellmann et al. (2020). Units for SELSS and for Lp,pk are dB re 1µPa2s.

Water depth

According to Jensen et al. (2011), water depth influences sound propagation in water. Below a certain cut-off frequency (f_g), unimpeded sound propagation is impossible. The shallower the water, the higher this cut-

off frequency (Figure 7). Sound around the cut-off frequency is reduced or damped to a large extent. An example of two measured 1/3 octave spectra in different water depths is given in Figure 8a. The cut-off frequency in 20 m water depth is approximately 40 Hz, meaning that below this frequency the spectrum drops off slightly steeper at such water depths compared to the frequency spectrum measured for >40 m water depth. The difference becomes prominent at shallower water depths (Figure 8b) because the spectral energy of the SEL_{SS} below 40 Hz is already low. For frequencies above the cut-off frequency, both spectra show a typical curve with a maximum between 63 Hz and 200 Hz and a decay to high frequencies with approx. 6 dB/octave. For the empirical dataset in this study, the impact of the cut-off frequency is negligible. However, the curve fitting (below) shows a negative influence of water depth and SEL_{ss}. This may be caused by the hammer energy being distributed over a larger pile surface and thereby reducing the sound energy per m^2 emitted by the pile.

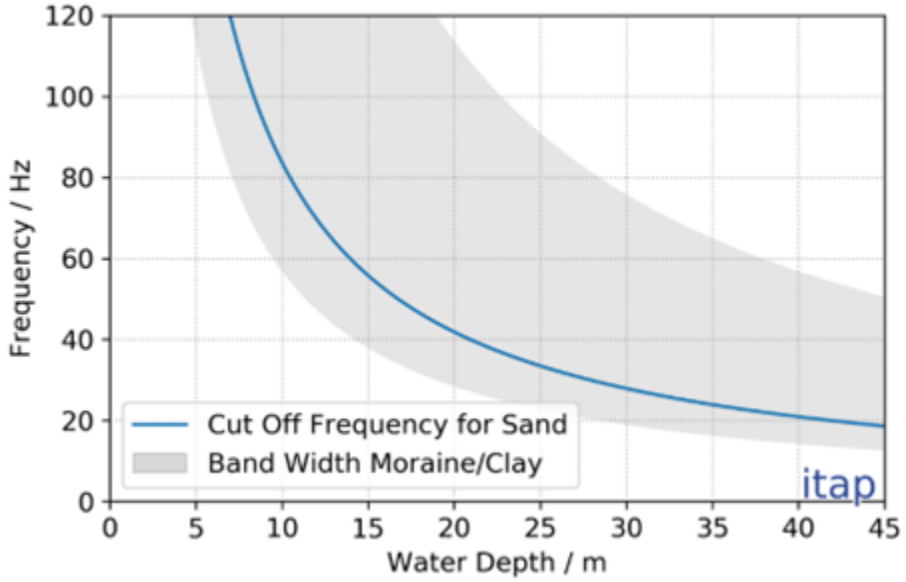

Figure 7: Theoretical lower (limit) frequency (fg) for undisturbed sound propagation in water as a function of the water depth for different soil types (example adapted from Urick (1983); Jensen et al. (2011); the example shows the possible range associated with different layers.

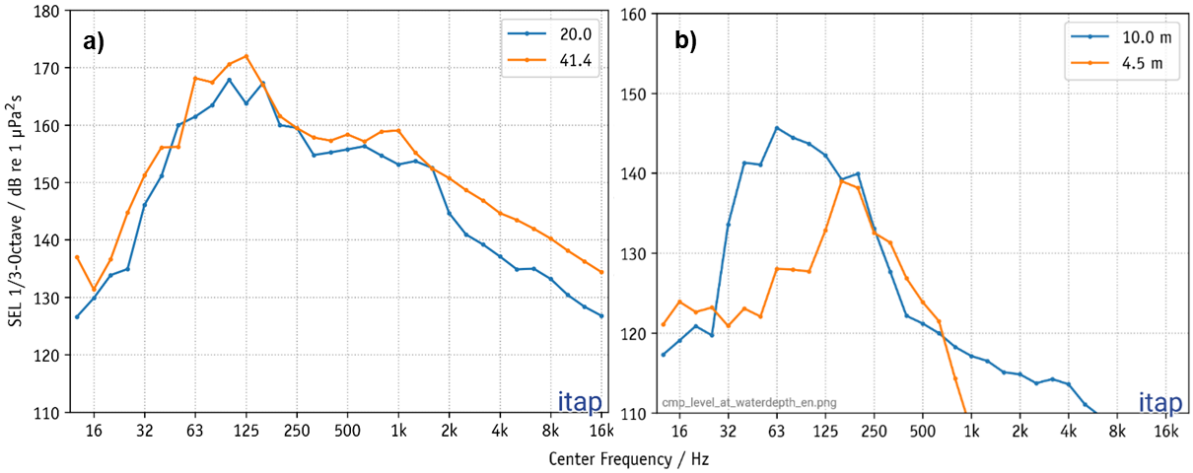

Figure 8: Comparison of the maximum 1/3 octave spectra of the SELSS for two Monopile installations at two different water depths measured at 750 m distance from the pile site. a) For the spectrum in 20 m water depth (blue line), a cut-off for frequencies below 40Hz can be seen. For the installation in 41.4 m water depth (orange line) the sound can propagate at full spectrum (data from this study). b) Two installations performed with comparable Noise Abatement Systems at 4.5 and 10 m water depth, sandy subsoil (Figure from Bellmann et al. 2020).

Penetration depth

For the one pin pile dataset included in the cross-project analysis, penetration depth has a significant influence on the SEL_{SS} during pin pile installation (Figure 9). The SEL_{SS} decreases with increasing penetration depth. The same phenomenon has been observed at the Beatrice OWF and described by Thompson et al. (2020). The amount of hammer energy that is converted to underwater sound reduces linearly with increasing penetration depth. The reason might be the continuously reducing surface area of the pile (stick-up length), which results in a reduction of the sound-radiating surface within the water column. Another reason might be the increasing pile stiffness with increasing penetration depth. The SEL_{SS} decreases continuously as soon as the hammer descends below the water surface (submerged hammering). Whereas with a monopile installation, the impact hammer is always above the sea surface and the radiating area in the water column remains almost constant.

Figure 9 shows an example of SEL_{SS} versus penetration depth from one pin pile installation, starting shortly before the hammer submerges below the water surface. The areas where the hammer reaches below the water surface and half the water depth are marked with a vertical line. The blow energy during the presented time window was nearly constant. The reduction in SEL_{SS} can therefore be attributed to the increase in penetration depth. The figure depicts a clear linear correlation between the sound-emitting surface in the water and the SEL_{SS}. In this example, a decrease in SEL_{SS} by 5 dB was observed with a reduction of the sound-emitting surface by 50%. In Lippert et al. (2017) this effect was modelled with 2.5 dB by halving the sound-emitting surface.

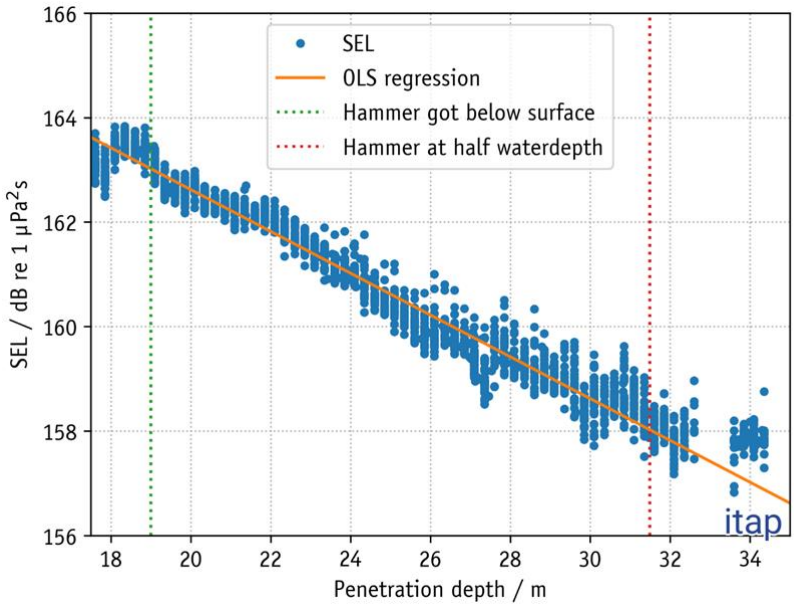

Figure 9: SELSS over the penetration depth from one selected pin pile installation. The green vertical line represents the penetration depth, where the hammer got below the water surface (submerged hammering) and the red vertical line presents the penetration depth, where the hammer reached half the water depth. The orange line shows the linear OLS regression which decreases with a slope of -0.4 over the penetration depth.

While there was no obvious influence of penetration depth on the SEL_{SS} for monopile installation (please see section 3.2.4.2 for in-depth investigations), qualitative inspection showed some obvious correlation between the change in numbers of single strikes needed per meter penetration depth (four times the blow count) and SEL_{SS}. In the example shown in the left graph of Figure 10, the blow count varies over penetration depth, while the blow count is relatively constant for most of the installation time in the example on the right of Figure 10. While the SEL_{SS} in the right example is relatively constant when normalised to a certain blow energy, it visibly fluctuates in the left example. The fluctuation shows a correlation with changing blow count: At low penetration depth, the SEL_{SS} increases with increasing blow count (box I) and decreases with decreasing blow count (box II). At deeper penetration depth, this correlation is less obvious (box III).

Blow counts change in instances where the soil conditions (soil resistance or cone penetration test) change during penetration. The harder the soil, the higher the number of blows with the same blow energy that are required to penetrate the pile by one meter. Based on the pile refusal criteria defined prior to piling and depending on the hammer used, it might be required to increase the applied blow energy in case the max blow count is reached. In hard soil conditions, less of the blow energy is converted into kinetic energy, and more will be converted to sound energy, compared to soft soil conditions. Therefore, we consider blow count as a good proxy to represent the soil resistivity. Penetration depth seems to have some influence on the relation between blow count and SEL_{SS}. With increasing penetration depth, the correlation gets weaker, likely due to the pile getting more stable and resonating less.

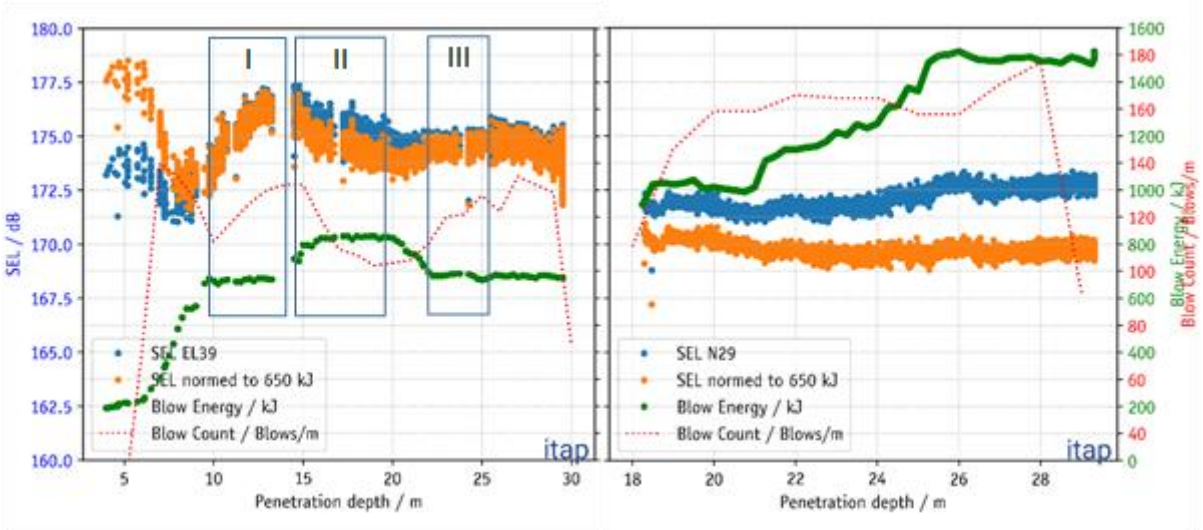

Figure 10: Broadband SELSS as recorded (blue dots), normalised to 650 kJ (orange dots), blow energy (green dots) and blow counts (dashed red line) for two monopile installations. Boxes I-III in the left graph are explained in the text.

Curve fitting

Blow energy, pile diameter and water depth were included into the curve fitting analysis via ordinary least squares to quantify the influencing factors for a forecast of the sound levels. Penetration depth was excluded from this analysis as this information was missing in hammer logs for more than half of the data sets (Table 2) but investigated further with the mixed model statistics (see section 3.2.2).

Blow energy, pile diameter and water depth cannot be considered independent of each other as usually larger pile diameters are used in deeper waters, and higher hammer energies are used for larger pile diameters. Therefore, in the following we do not attempt to determine dependency functions for each individual parameter, but to develop a model that adds the parameters sequentially. To eliminate the influence of the sound radiating area relative to the water depth (as penetration depth is not considered), only the SEL_{SS} of monopile installations were included in the analysis.

For all three parameters, a logarithmic dependency of the SELSS on the respective parameter is assumed, resulting in

$$SEL = 6.21 * \log(be) + 12.82 * \log(pd) - 8.19 * \log(wd) + 156.42 \qquad \textbf{Equation 2}$$

With

be - blow energy in kJ,

pd - pile diameter in m,

wd - water depth in m.

Figure 11 shows to what extent the variation in SEL_{SS} can be reduced by considering the respective parameter based on the monopile data set consisting of 77,648 single strikes. In this Figure, the broadband SEL_{SS} for each monopile single strike of our data set is shown in grey, then stepwise and additive, the influence of blow energy (orange), pile diameter (red) and water depth (blue) is eliminated using Equation 2,

leading to a reduction of the variance from 9.7 dB to 6.2 dB. Therefore, by considering these three parameters, 90% of the SEL_{SS} values could be predicted with an uncertainty of ± 3.1 dB.

These results confirm the results of Bellmann et al. (2020), which have shown that the project-specific parameters blow energy, pile diameter and foundation type have a significant influence on pile driving noise. The amount of hammer energy emitted into the sea as sound is very much dependent on the foundation design, the diameter of the pile and the blow energy used to install the foundation piles by impact pile-driving. Simply put, it can be said that the stronger the pile vibrates and the larger the radiating area is in the water, the higher is the transmitted sound energy. For this reason, the blow energy is an important impact parameter in all underwater noise models, as is the radiating surface of the pile in water, which is related to the pile diameter and the water depth (Bellmann et al. 2020). During the installation of jacket foundations, not only the pile but also the guiding frame or entire foundation (if it is already in the water) is also stimulated to vibrate, which adds to the emitted sound energy.

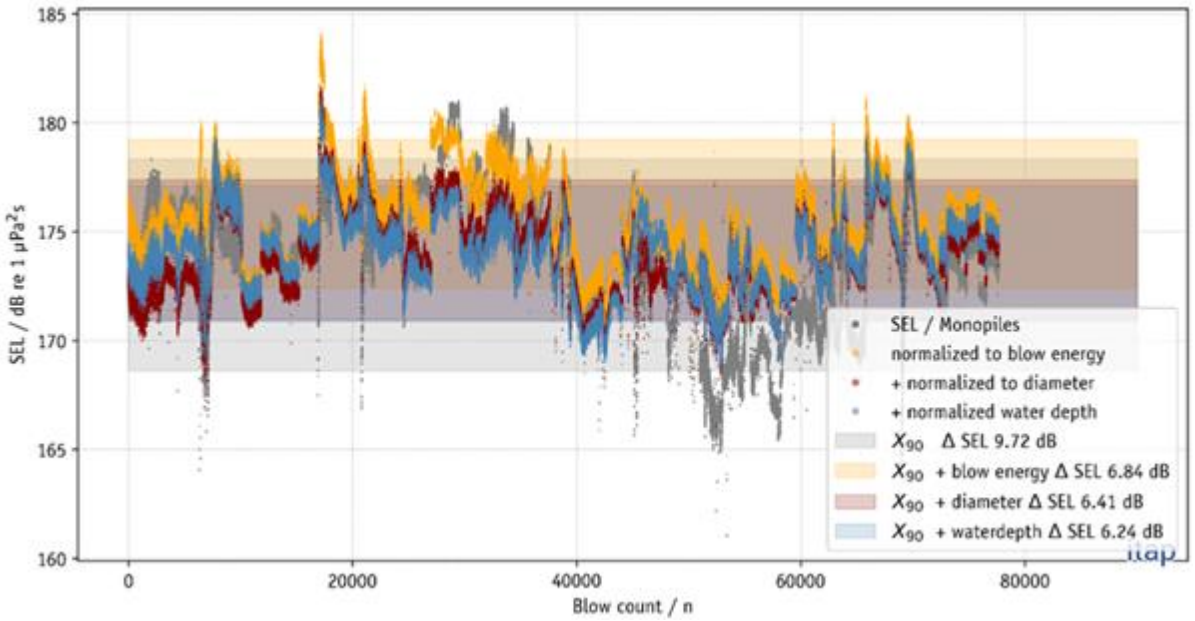

Figure 11: SELSS of each single strike of the monopile installations included in this study (excluding soft start), consisting of 77,648 blows (grey). The SELSS values were normalised over the blow energy, diameter and water depth; the 90% confidence interval X90 for each normalized SEL_{SS} value is depicted as well: normalised for blow energy (orange), for blow energy and diameter (red), and for blow energy, diameter and water depth (blue).

3.2.3. Frequency spectra

The frequency spectra of single strike piling noise are relatively broadband and low frequency, with a peak frequency of 125 Hz for monopiles and pin piles, and a -3dB bandwidth below 200 Hz and -10 dB bandwidth of below 1 kHz (Table 9, Figure 12). The frequency spectra of pin pile installations show more energy at the high frequency end of the spectrum compared to monopile installation. This is likely due to the smaller pile diameter of pin piles compared to monopiles leading to a higher resonance frequency. Figure 13 shows the differences between pin piles and monopiles from Bellmann et al. (2020) for a different dataset. The presented pin piles and monopiles in Figure 12 are similar compared to the range shown in Figure 13.

Table 9: Peak frequency and bandwidth of pile strike sound recorded at 750 m from the pile site for monopile and pin pile installation.

Foundation type	Peak frequency (Hz)	-3 dB bandwidth (Hz)	-10 dB bandwidth (Hz)
Pin pile	125	170	960
Monopile	125	80	370

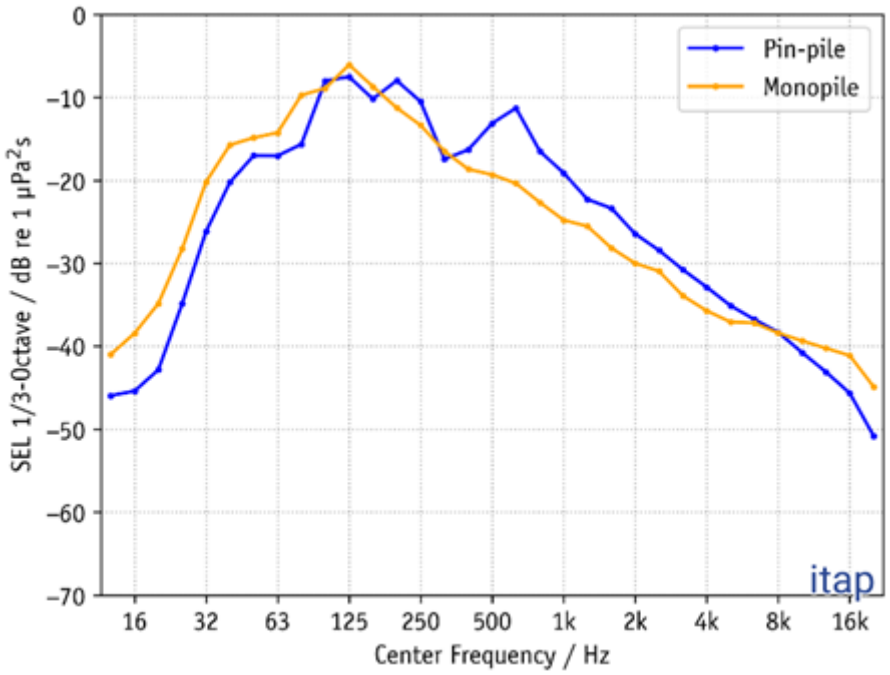

Figure 12: Normalised median 1/3 octave spectra for monopile and pin pile installations based on the data from this study.

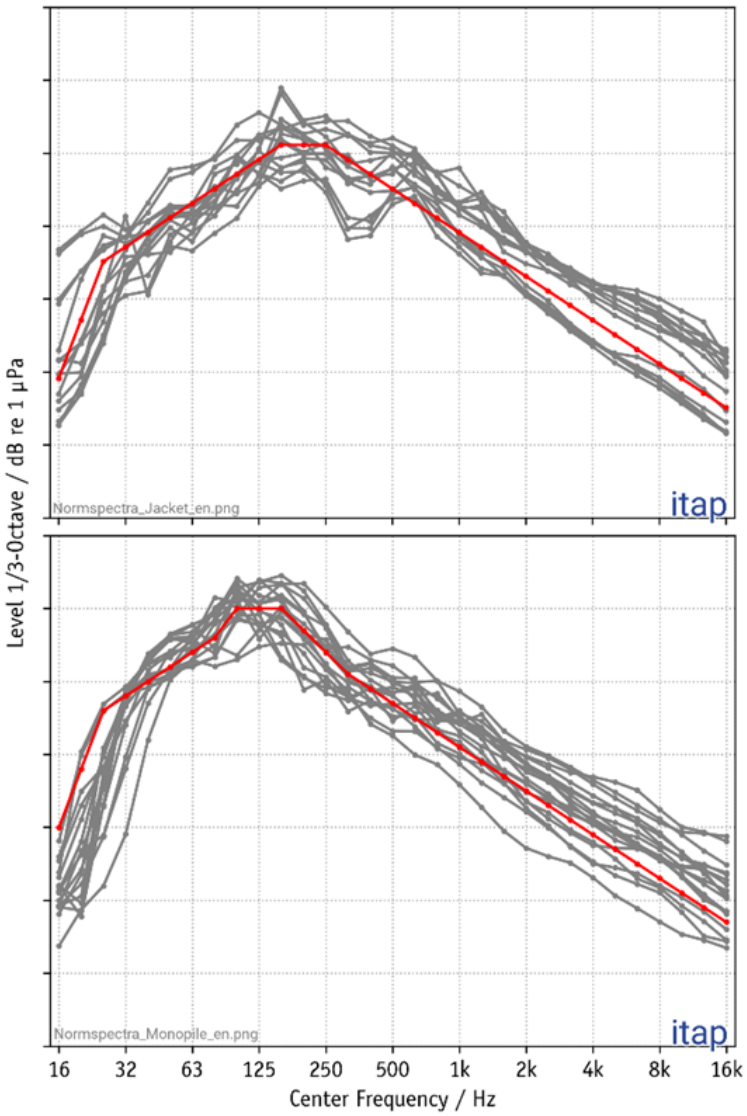

Figure 13: 1/3-octave-spectra of impulse pile-driving in different OWF construction projects, measured at 750 m distance from Bellmann et al. (2020) . The pile-driving activities were performed without the application of technical Noise Abatement Systems. Top: grey shaded lines mark the real measurement data of different pile diameters up to a maximum diameter of approx. 3.5 m (piles for Jackets); the red line characterizes an averaged, theoretical model spectrum (median). Bottom: grey shaded lines mark the real measurement data of different diameters (minimum 6 m, monopiles); the red line characterizes the averaged, theoretical model spectrum (median). The presented level grid is 10 dB.

3.2.4. Mixed models

To conduct an in-depth analysis on the influence of penetration depth and blow count (number of single strikes per 25 cm) on the SEL_{SS}, a GLMM approach was taken. This analysis has been conducted on a subset of the data involved in this study (see Table 2), resulting in only a small coverage of water depths (20 to 27 m) and pile diameter (5 to 6.5 m) for monopile installations, and only four pin pile data sets from the same foundation installation.

Pin piles

The GLMM analysis shows a positive relationship between SEL_{SS} and blow count and blow energy and a negative effect between SEL_{SS} and the pile penetration depth (Figure 14, Table 10). Penetration depth had the largest impact on SEL_{SS} values recorded at 750 m with a predicted spread of ~7 dB between the shallowest and deepest penetration values. Blow energy and blow count also had significant influence on SEL_{ss} but larger uncertainties surrounding the estimates and lower overall influence. While the SEL_{SS} increases by around 4 dB from hammer energies below 300 kJ to above 800 kJ, further supporting the results of the regression analysis (section 3.2.2), the blow count only has a minimal effect on the SEL_{SS}.

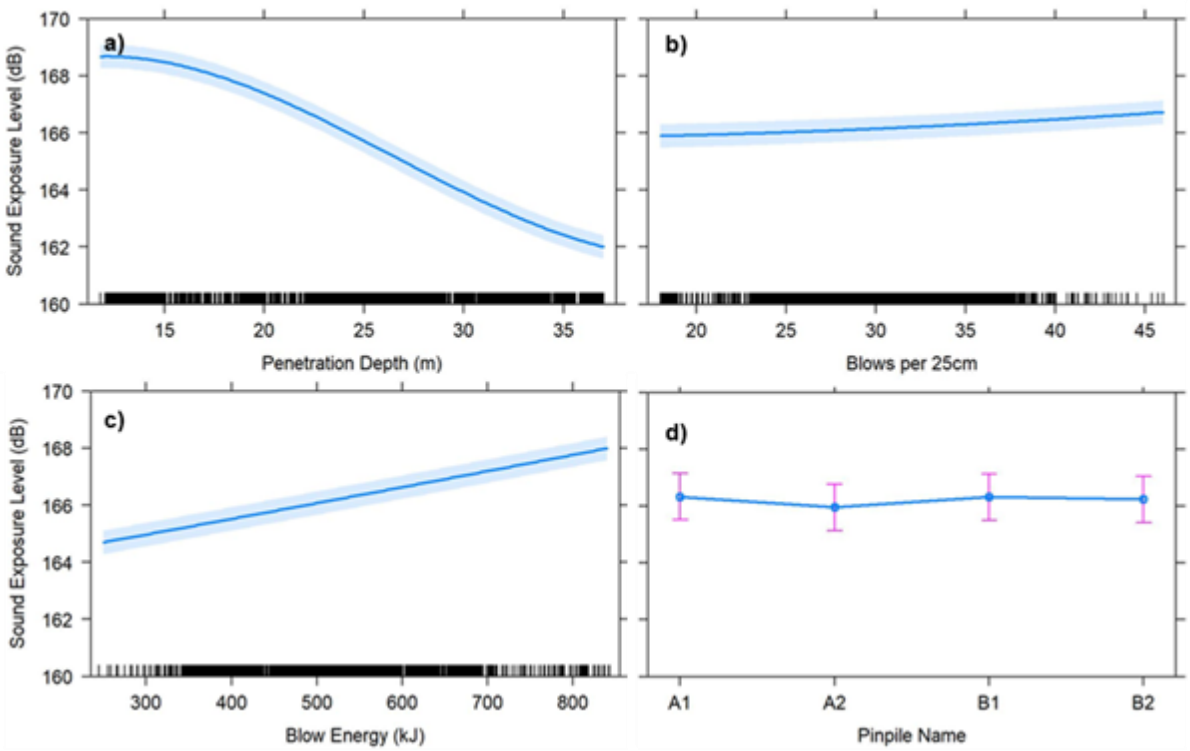

Figure 14: Partial effects on sound exposure levels during pin pile installation in the selected model for each factor: penetration depth (a), blow count (b), blow energy (c) and pin pine name (d). Blue lines (continuous variables) and dots (factor level variables) represent predicted value. Shaded areas (continuous variables) and whiskers (factor level variables) represent the 95% confidence intervals around the prediction.

Model parameters for the best fit model are shown in Table 10. High p-values indicate that there the parameter had little impact on the received level of the impulse. In this case, there was little difference between the piles (pilename) within the windfarm (P>0.5; Figure 14). The second term of the energy

relationship was also non-significant. However, as this was part of polynomial relationship, direct interpretation is not meaningful. Penetration depth, blow energy, and blow count (blows per 25 cm of penetration) were all meaningful (P<.001) in predicting the received level of the blows.

Table 10: Model summary of the best fit model for the pin pile data.

Model Parameters	Estimate	Std. Error	df	t-value	Pr(>|t|)
Intercept	166.10	0.42	12.00	398.31	< 2e-16 ***
poly(penetration_depth, 3)1	-478.10	0.99	4,9650.00	-485.31	< 2e-16 ***
poly(penetration_depth, 3)2	-44.26	0.87	49,650.00	-50.61	< 2e-16 ***
poly(penetration_depth, 3)3	39.05	0.76	49,650.00	51.65	< 2e-16 ***
poly(blow_energy, 2)1	122.70	1.13	49,650.00	109.07	< 2e-16 ***
poly(blow_energy, 2)2	0.76	0.92	49,650.00	0.82	0.41
poly(blow_count, 2)1	29.63	0.79	49,650.00	37.31	< 2e-16 ***
poly(blow_count, 2)2	3.72	0.74	49,650.00	5.05	.35e-07 ***
pilenameA2	-0.37	0.59	12.00	-0.62	0.545
pilenameB1	-0.01	0.59	12.00	-0.02	0.984
pilenameB2	-0.09	0.59	12.00	-0.14	0.888

K-fold cross validation of the selected mode indicated that the prediction error was generally ±3 dB and the accuracy varied between pile locations (Figure 15). The largest within-pile location variation in model error was observed in the A1 location where the model error was 6 dB between random effect locations. Model error was also skewed by 3 dB at one of the A2 locations.

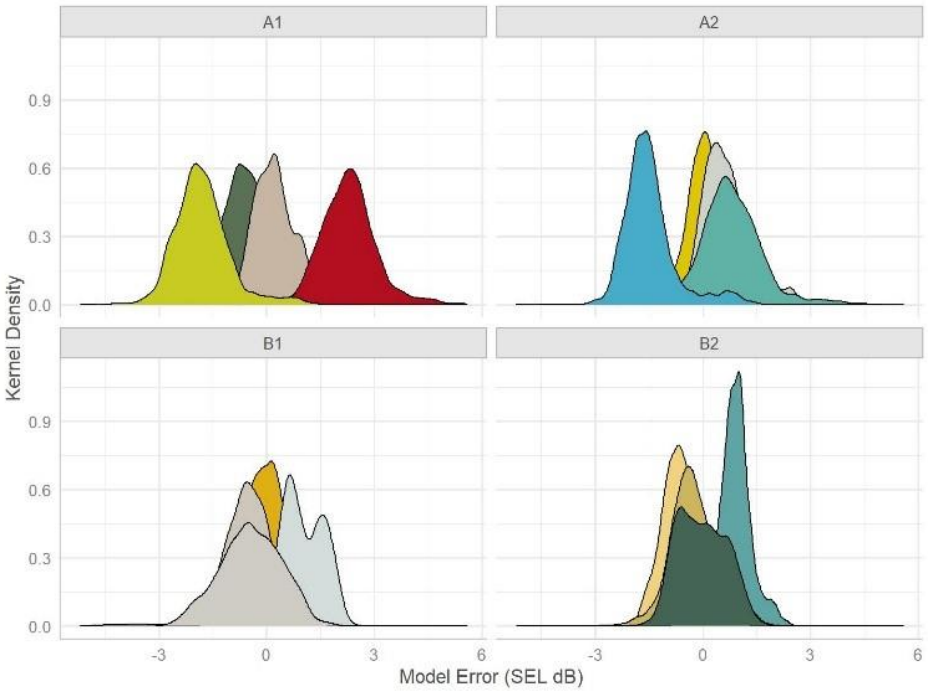

Figure 15: K-fold model error. Difference (dB) between observed and predicted model values for each point in the pin pile data sets (with pile names A1, A2, B1, B2) used for the GLMM.

Monopiles

AIC selection strongly favoured the fourth model (Table 4), but k-fold cross validation indicated lower model error using the third model. Similarly, while no VIF score exceeded 6 for any model, VIF scores for the fourth model were highest. Given that no model predicted the SEL_{SS} specifically well at other wind farm locations, we assume that the higher order polynomial selected by AIC likely represents over-parametrization compensating for latent variables. We therefore selected the third model which had the lowest K-fold error, and second lowest AIC score.

The relationships between the observed parameters and the SEL_{SS} was complex, and the models did not generalise well; models built with data from different windfarms did not accurately predict the SEL_{SS} values at the windfarm that was held out. Final model selection indicated a positive relationship between SEL_{SS} and blow energy (Figure 16, Table 11), further supporting the results of the regression analysis (section 3.2.2). Modelling also indicated a negative effect between SEL_{SS} and the pile penetration depth, which might be explained by the pile vibrating less in the water column as it becomes stabilized by the surrounding sediment. Blow count also showed a negative relationship with SEL_{SS} at low blow counts as well as high blow counts, while no correlation is obvious for the middle range of blow counts. This result does not confirm the hypothesis based on the data shown in Figure 10, and will need further investigations to fully understand the relationship. Very low blow counts, however, were partly connected to a pile reaching or being at its final penetration depth, and the inclusion of such data may have masked any effect potentially existing in the middle range of blow counts. The effect of pile diameter was, as expected, not significant, due to the small range of diameters included in the analysis.

K-fold cross validation of the selected model Indicated that the prediction error was generally ±5 dB and the accuracy varied considerably both within and between windfarms (Figure 17). The smallest within-windfarm variation in model error was at OFW 1 where the median model error was ~2 dB between pile locations. At

OFW 6 the median k-fold error was >7dB between pile locations. This result indicates that the modelling does not include all covariates necessary to enable a sufficient prediction of the SEL_{SS} at future locations.

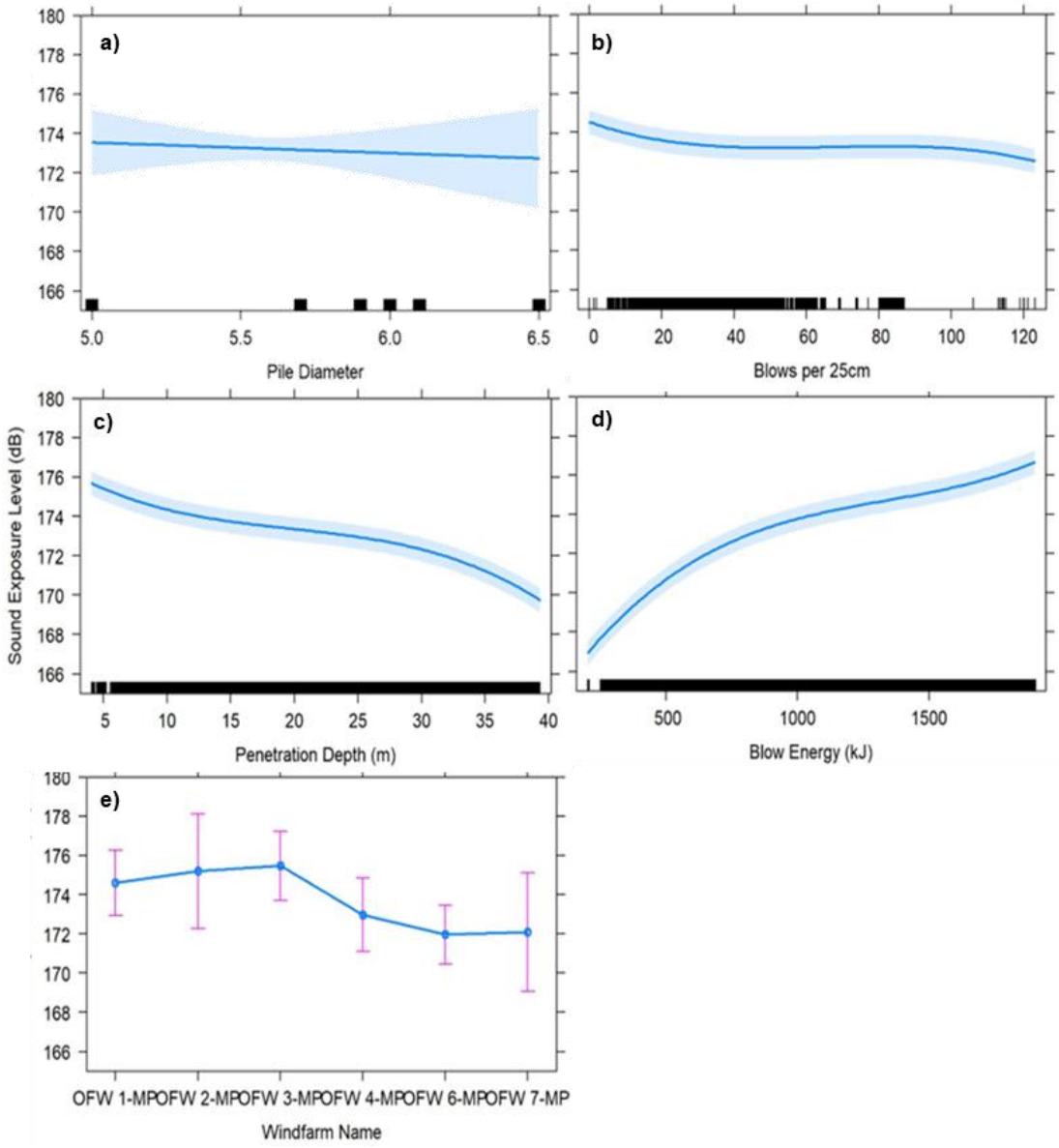

Figure 16: Partial effects on sound exposure levels during monopile installation in the selected model for each parameter: pile diameter (a), blow count (b), penetration depth (c), blow energy (d) and windfarm (e). Blue lines (continuous variables) and dots (factor level variables) represent predicted value. Shaded areas (continuous variables) and whiskers (factor level variables) represent the 95% confidence intervals around the prediction.

Table 11: Model summary of the best fit model for the monopile data.

Fixed effects:	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	176.8066	8.1476	10.9997	21.7	2.22E-10
poly(blow_energy, 3)1	458.2782	2.6655	42365.81	171.929	< 2e-16
poly(blow_energy, 3)2	-125.335	1.3287	42366.57	-94.33	< 2e-16
poly(blow_energy, 3)3	58.0544	1.0278	42363.87	56.482	< 2e-16
poly(penetration_depth, 3)1	-185.235	1.5539	42364.82	-119.203	< 2e-16
poly(penetration_depth, 3)2	-31.154	1.0012	42362.67	-31.116	< 2e-16
poly(penetration_depth, 3)3	-38.9475	0.8985	42362.97	-43.349	< 2e-16
poly(blow_count, 3)1	-30.6875	2.3915	42363.11	-12.832	< 2e-16
poly(blow_count, 3)2	7.054	1.0677	42362.39	6.607	3.98E-11
poly(blow_count, 3)3	-22.4965	1.1525	42362.26	-19.52	< 2e-16
diameter_m	-0.538	1.3755	10.9986	-0.391	0.7032
windfarmOFW 2-MP	0.5836	1.9067	11.0071	0.306	0.7653
windfarmOFW 3-MP	0.8537	1.1793	11.0234	0.724	0.4842
windfarmOFW 4-MP	-1.6481	0.9983	11.031	-1.651	0.1269
windfarmOFW 6-MP	-2.6505	1.3442	11.0142	-1.972	0.0743
windfarmOFW 7-MP	-2.5128	1.413	11.017	-1.778	0.1029

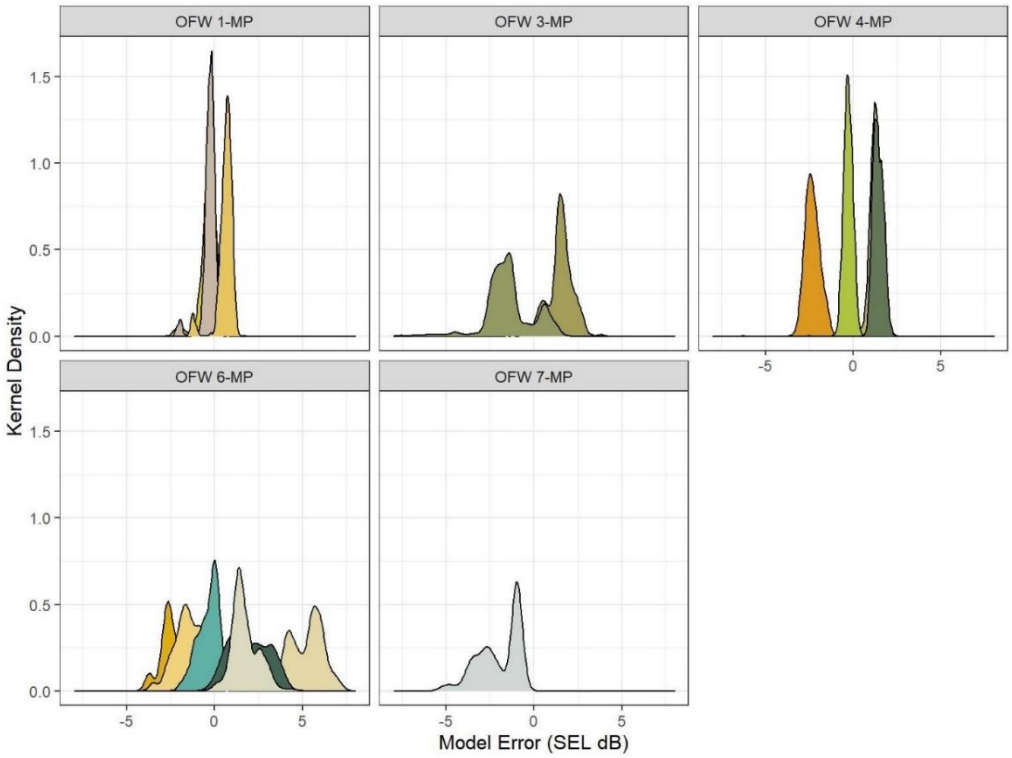

Figure 17: K-fold model error. Difference (dB) between observed and predicted model values for each point in the data set.

3.3. Comparison of measurements with modelled predictions

For the six UK OWFs there was a good agreement between the assumption of the pile diameter used for modelling and the actual pile diameter used. The maximum hammer energy assumed for modelling was, however, never reached during the installation of the piles that were subject to noise monitoring and included in the analysis of this study (Table 12). Further data from offshore installations are needed to understand if this is a common feature.

Table 12: Comparison of pile diameter and maximum hammer energy values used for modelling and used/measured at sea.

	Pile diameter		Maximum hammer energy (kJ)	
Pile type	**Modelled**	**Installed**	**Consented**	**Measured**
Monopile	8.1	8.1	3,000	2,031 to 2,557
Monopile	9.5	9.5	4,000	478 to 2,893
Monopiles	8.5	6.5	4,000	1,343 to 1,725
Pin pile	up to 2.5	2.5	2,250	1,012 to 1,748
Pin pile	2.2	2.2	2,500	826 to 1,735
Pin pile	2.5	2.44	3,000	1,952 to 2,545

3.3.1. Single strike sound levels

Comparisons of single strike sound levels that were modelled and measured are provided in Figure 18 by OWF. The most comprehensive data set are depicted in Figure 18a. This was the only situation where results of the unweighted SEL_{SS} were presented. The results obtained from the impact assessment modelling and the modelling for this report matched well with the noise levels from measurement data. Only at larger ranges (several 10s km) were the Upper Bound estimates overestimated. A comparison of the data from the OWF seen in Figure 18c also shows a high level of agreement, although the back calculation of weighted forecasts to unweighted SEL_{SS} introduces uncertainties. For Figure 18b, weighted forecasts also had to be back calculated to unweighted SEL_{SS} values. Here, some impact range predictions were way overestimated (~9 km instead of ~1 km). However, there is a good agreement between the transmission loss used in the impact assessment and the re-modelled transmission loss. Two fixed measurement locations were available for this OWF, resulting in two measurements at different distances for each foundation. However, due to the small number of measurement positions and the fact that they are often close to each other, the actual transmission loss cannot be sufficiently estimated.

The measurements shown in Figure 18d and Figure 18e were carried out at one measurement point that was fixed at one location throughout the installation of several foundations; meaning the distance to each foundation differs significantly from each other. The measurements obtained for different distances to the pile site at these windfarms, therefore, are from measurements taken during different pile installations. The scatter seen in these data are likely due to different source levels rather than due to different distances. The attempt to determine a realistic propagation loss from this proved to be unsuccessful.

The comparison of measured and modelled sound levels has proved to be difficult for most of the OWF involved in this comparison, as a) the modelling reports are focused on the noise impact assessment and thereby do not necessarily deliver sufficient or tailored information suitable for such comparison, and b) the measurement array deployed in the field is not sufficient to determine an accurate transmission loss. To determine a valid transmission loss function in the field a minimum four measurement position should be in place at distances of between 750 m and up to 8 to 12 km (doubling distances between two measurement locations) for each pile site. If this is not given a calculation of the site-specific transmission loss is not reliably possible.

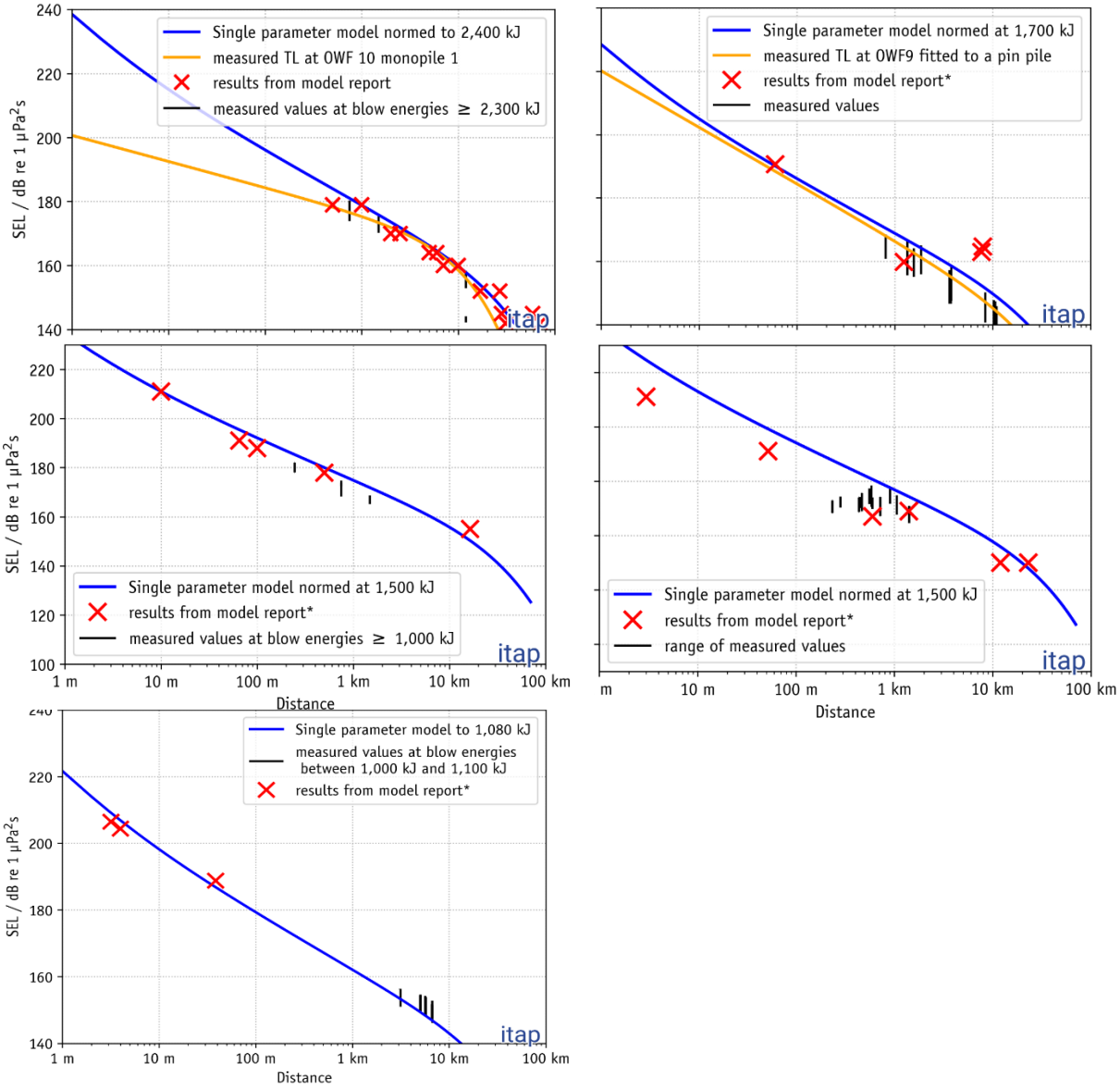

Figure 18: Comparison of the measured and modelled unweighted SELSS for five different OWFs. The transmission loss function (TL) based on the measurement results (orange) could only be determined for two OWFs due to insufficient measured data at different distances to piling. The blue line shows the modelled SEL_{SS} over the distance for a given blow energy by using the empirical model of itap within this study. The black vertical lines show the range of measured SEL_{SS} and the red crosses the results of the modelled impact ranges estimated based on certain noise thresholds.

3.3.2. Cumulative sound exposure level

The SEL_{cum} was only estimated in two OWF modelling reports. However, the measurement data which one of the noise monitoring is based upon were partly recorded as interval measurements (e. g., 10 minutes measurement time every 30 minutes). A comparison of modelled estimates with measured data was therefore only possible for the data of the other study, shown in Figure 18a. Although the SEL_{cum} was given in the model for a start distance of 750 m, for which measurement data are also available, the forecast considered a worst-case scenario with two installations in one day. Therefore, we have recalculated the

forecast with one installation per day. In Figure 13, the recalculated SEL_{cum} at 750 m is compared to the SEL_{cum} from the measurements of three monopile foundation installations. In addition, key pile driving data are given, such as the total number of pile strikes, the maximum blow energy, and the total piling duration without interruptions.

In the model, the unweighted SEL_{cum} was calculated for two monopile foundations (1 and 2). The same piling sequence and the usage of an ADD for 15 Minutes was considered for both foundations. The assumed piling sequence is based on a previous pile driving analysis. It is noticeable that the parameters blow energy, piling duration and total strikes differ significantly for the third foundation measured. However, foundation 1 and 2 are comparable and are more similar to the model predictions.

The maximum unweighted SEL_{SS} was overestimated by 3 dB in the model, based on a pile driving energy of 4,000 kJ. Based on the actual pile driving energies used, the levels would be 2 dB lower in both cases. The prediction results thus had a deviation of + 1 dB from the measurements and therefore is within a stated uncertainty of ± 2 dB.

For the modelling, 3,121 single strikes were considered; there were 156 fewer strikes than modelled at foundation 1 and 461 more strikes than modelled at foundation 2 (Figure 13). The actual pile driving duration was significantly less in both cases. The worst-case assumption of 4,000 kJ was not reached for any of the monopile installations, but somewhat compensated for at foundation 2 by the higher number of blows. The unweighted SEL_{cum} estimated for a stationary animal is 3 dB to 4 dB lower than the modelled SEL_{cum} for foundation 1 and 2, and 17 dB lower for foundation 3, which had the shortest installation time with the least strikes and lowest blow energy. The difference between modelled unweighted SEL_{cum} and measurements are similar or slightly larger when considering a moving animal (Figure 13). During the installation of foundation 2, these pauses were only 2 to 6 minutes (Figure 19). During the ramp up, 3 pauses of 8 minutes each were considered in the model (Figure 19). This does not result in any differences in the calculation of the stationary SEL_{cum}. For the moving SEL_{cum} this does make a difference. For example, if a break is 6 minutes longer, an additional 540 m more could be covered by the animal without receiving any sound energy.

All in all, the installations of the three monopiles were shorter in duration, with fewer strikes and a lower hammer energy than assumed for modelling. The unweighted SEL_{cum} based on the measured data were lower than the modelled SEL_{cum}. This could lead to the assumption that the SEL_{cum} based on the measured data is smaller than from the modelling results. This is true for foundation 3, as seen for the PTS- and TTS-impact ranges for VHF-cetacean (Table 13). For foundation 1 and 2, while the VHF PTS-impact range is small, the VHF TTS-impact range is more than 2 km further than the model assumption. This is likely due to the frequency spectrum of foundation 1 and 2 containing more energy in the high frequency part than the other foundations and the spectrum used for modelling (Figure 20), as well as higher SEL_{ss} throughout the soft start (Figure 19). The reasons for the discrepancies will have to be investigated in further detail.

Table 13: SELcum and SELSS modelled and measured at three foundation installations of one OWF at 750 m distance. In addition, relevant piling sequence parameter are listed.

OWF	Model	Measurement foundation 1	Measurement foundation 2	Measurement foundation 3
Unweighted max. SEL_{SS} at 750 m	183	180	180	174
VHF-weighted max. SEL_{SS} at 750 m	144	159	158	143
Unweighted stationary SEL_{cum} at 750 m	217	213	214	199
VHF-weighted stationary SEL_{cum} at 750 m	178	191	191	168
Unweighted moving SEL_{cum} starting at 750 m + 15 minutes ADD	198	193	193	177
VHF-weighted moving SEL_{cum} starting at 750 m + 15 minutes ADD	145	153	151	134
Total number of strikes	3121	2965	3582	769
Max. blow energy	4000 kJ	2869 kJ	2280 kJ	858 kJ
Duration (excl. breaks)	01:39:00	01:16:22	01:26:47	01:02:59
SEL_{cum} PTS impact range (VHF)	<0.1 km	<0.1 km	<0.1 km	<0.1 km
SEL_{cum} TTS impact range (VHF)	2.8 km	6.3 km	5.4 km	<0.1 km

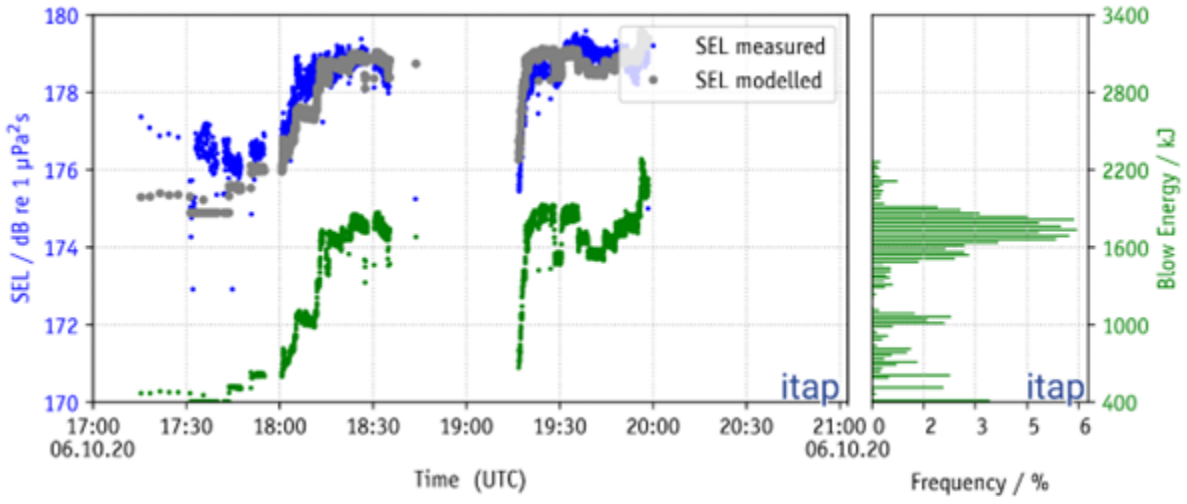

Figure 19: Time series of the measured SELSS (blue) at approx. 750 m and the blow energy (green) during monopile foundation 2 installation. Shown is also the modelled time series based on Equation 2 (grey).

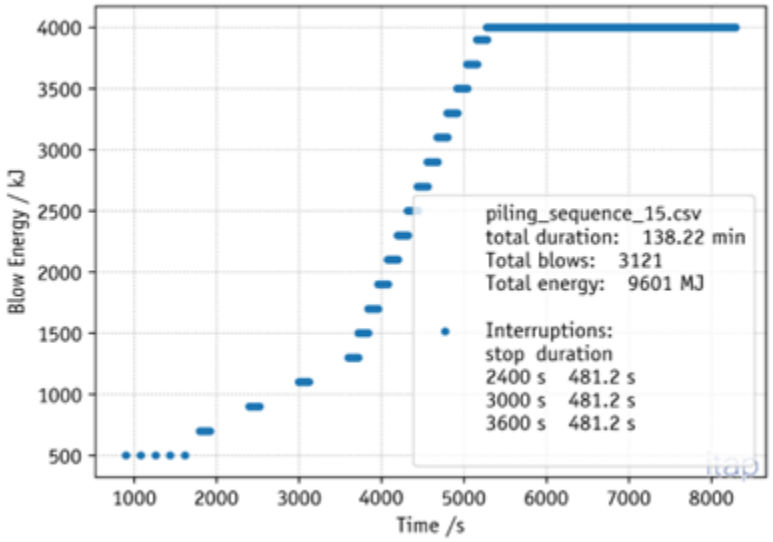

Figure 20: Time series of the blow energy as used for modelling the installation of monopile foundation 1 and described in the corresponding modelling report.

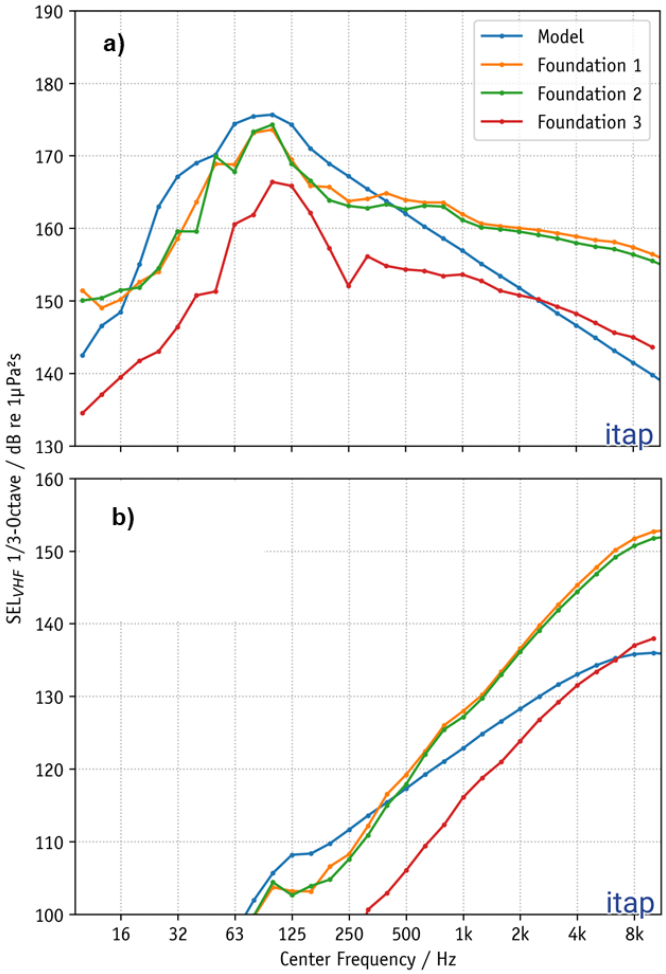

Figure 21: Comparison of unweighted (a) and VHF-weighted (b) frequency spectra used for modelling (blue) and as measured during foundation installation 1 (yellow), 2 (green) and 3 (red).

3.4. Sensitivity analysis

3.4.1. Single strike sound exposure level

All three parameters blow energy, pile diameter and water depth, have a logarithmic relationship with the SEL_{SS} (Figure 21). Therefore, variations towards lower values of a factor lead to larger changes of the SEL_{SS} than variations made towards higher values of a factor. SEL_{SS} increases with 1.8 dB per doubling blow energy, 3.9 dB per doubling pile diameter and decreases with 2.4 dB per doubling water depth.

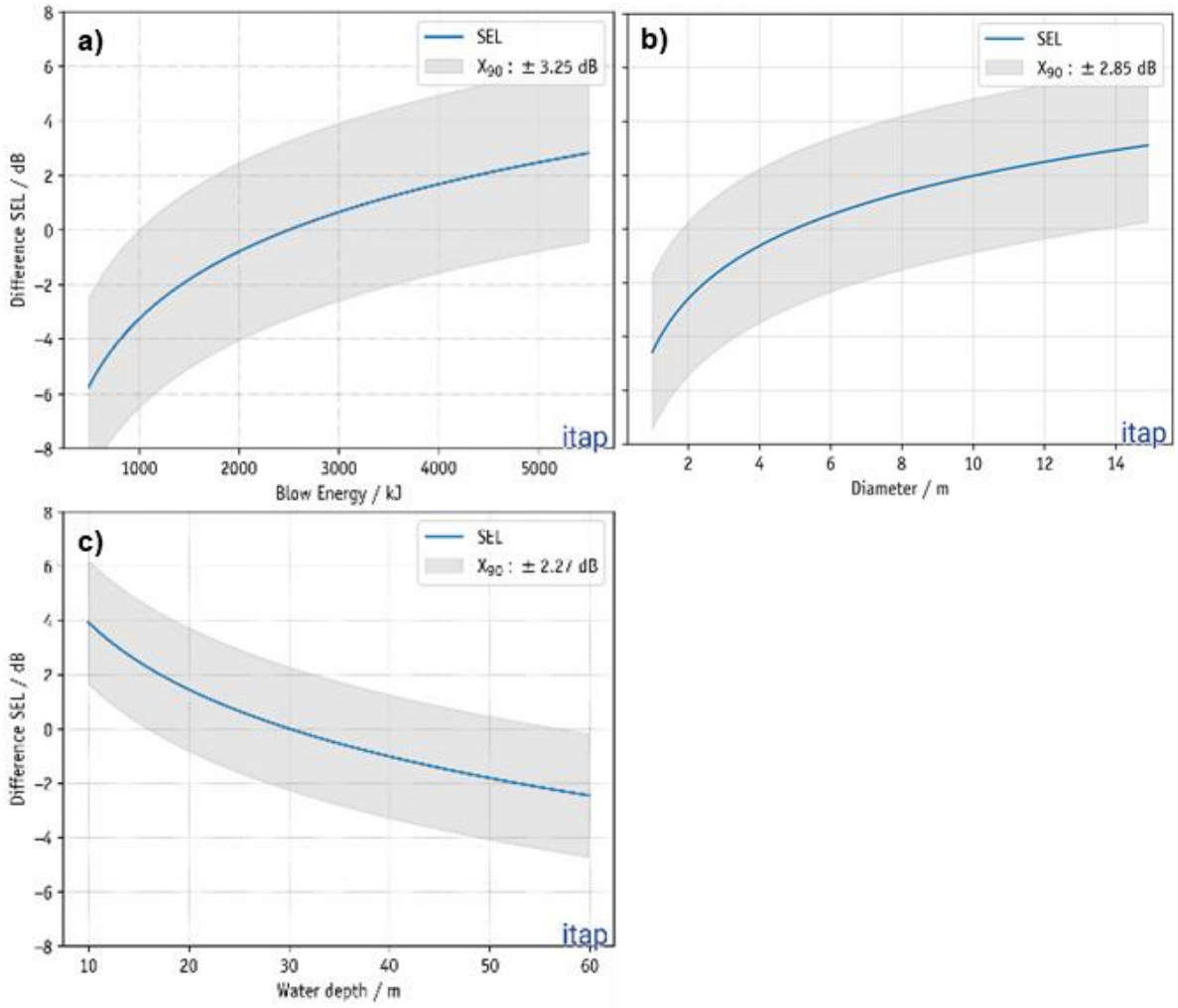

Figure 22: Modelled influence of blow energy (a), pile diameter (b) and water depth (c) on the single strike SELSS. This figure shows the difference in dB to the SELSS of a pile strike on a monopile with a standard scenario of 5 m diameter in 30 m water depth with 2,500 kJ blow energy. The grey shaded area shows the 90% percentile range.

3.4.2. Cumulative sound exposure level

Figure 22 shows the influence of changing either the frequency spectrum, number of blows, blow energy or pile diameter when modelling the SEL_{cum} PTS-impact ranges. It is obvious that the impact ranges for HF-cetacean such as dolphins are not influenced by changes in the factors considered here. Seals (PCW species group) are also hardly affected. The impact ranges for VHF cetacean such as the harbour porpoise and LF cetacean such as the minke whale are noticeably affected by changes to the considered factors.

For the frequency spectrum it is obvious that with increasing energy towards the higher frequency range, the impact ranges for VHF and LF increase noticeably. In practice this means that 1) frequency spectra of pin piles result in higher impact ranges than those of monopiles (given the same overall energy), and 2) it is important that a realistic frequency spectrum is assumed for modelling in order not to under- or overestimate the impact ranges.

Having a fair estimate of the number of pile strikes used for modelling the installation of a pile is important, especially when using a small number of pile strikes. With increasing number of pile strikes the error in impact range estimates caused by a mismatch between modelling and real data become smaller and, at some point, insignificant.

Changes in the blow energy has a larger effect on LF-cetacean than on VHF-cetacean but has generally a noticeable effect regardless of the magnitude of the blow energy (though for LF-cetacean a little less at higher values). This is similar to the relationship observed with changes in pile diameter.

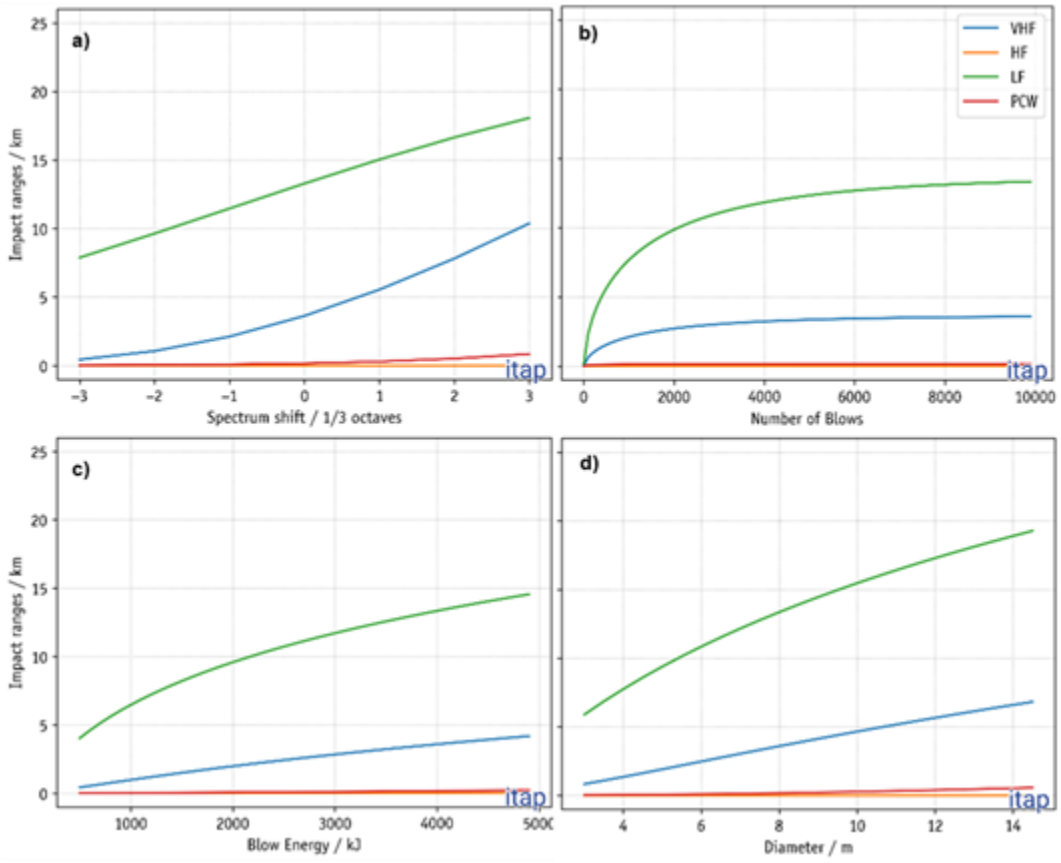

Figure 23: Modelled PTS-impact ranges based on the SELcum-thresholds for PTS defined by Southall et al. 2019 for different species groups in relation to changes in the frequency spectrum (a), total number of pile strikes (b), blow energy (c) and pile diameter (d) by using Equation 2 and the empirical transmission Loss function IIg according to Thiele and Schellstede (1980). Base Scenario: 8 m Monopile at 40 m water depth with 5,000 blows of 4,000 kJ blow energy and a blow rate of 30 blows/ minute.

Figure 23 shows the influence of changes in either fleeing speed, the soft start duration, SEL_{SS} levels during soft start and blow rate during soft start when modelling the SEL_{cum} PTS-impact ranges. As for the factors discussed above, changes to soft start duration, blow rate or the SEL_{SS} during soft start hardly affect the

impact range estimates for dolphins and seals, as those are already literally zero for the baseline scenario. Swim speed only affects the impact range for these species at very slow speeds. There is a noticeable effect of changes in fleeing speed, length of soft starts and sound level at soft start on VHF and LF cetacean. A change of the blow rate hardly has any effect on the impact ranges when considering that all other parameters are kept to baseline values. However, using a higher blow rate as a baseline value (28 strikes per minute instead of 1 strike per minute) shows that the influence of the other factors on the impact ranges becomes more prominent (Figure 24).

Increasing the fleeing speed of an animal results in smaller impact ranges, as the received levels decrease in a logarithmic manner with increasing distance to the pile site, and a faster animal therefore accumulates less sound energy than a slower animal. It is, therefore, important when modelling the impact ranges for marine mammals e.g., harbour porpoise and minke whales, to use a realistic swimming speed. Often, as in our example, a direct swim path away from the pile site is assumed. Estimated swim speeds are, however, often based on surfacing of animals, which may be faster than the speed of an animal when measured directly away from the sound source. McGarry et al. (2017), for example, observed swim speeds of minke whales moving away from a sound source, and analysed the swim speed of the animals in different ways. While the swim speed obtained from surfacing was around 15.1 km/h (± 4.7 sd), the speed the animals moved directly away from the sound source was 3.4 km/h (± 1.8 sd) lower.

Prolongation of soft start during modelling helps to reduce the impact ranges for harbour porpoise and minke whale to some extent. However, the sound levels of the soft start cannot yet be reliably modelled as data measured in the field show a higher, yet unexplainable variety in sound levels compared to those during the installation process. During the soft start, sound levels can be higher than expected through modelling, which increases the impact ranges for the species discussed here. However, the influence of higher SEL_{SS} during the soft start seems to be minor when the SEL_{SS} is only a few dB higher than expected, and the blow rate is kept at 1 strike per minute.

Unexpectedly, changes in the blow rate alone have hardly any effect on the impact ranges. This might be due to the low blow energy assumed in the standard scenario. Using a different blow rate in the standard scenario has an effect on the influence of the other parameters on the impact ranges. Consequently, when interpreting the results presented above, one must consider that only one parameter at a time was changed while all other parameters were kept constant. There may be interactions between parameters, e.g., that one parameter may have a larger influence on the impact range when other parameters are different to the current standard scenario chosen.

The estimation of SEL_{cum} impact ranges inherits further uncertainties that were not investigated in this study. These are the change of impulsiveness with increasing distance to the pile site (Hastie et al. 2016) and recovery of the animal's hearing threshold shift in-between pile strikes (Kastelein et al. 2015) which lead to conservatisms in the modelled impact ranges. Further research is needed to quantify those uncertainties.

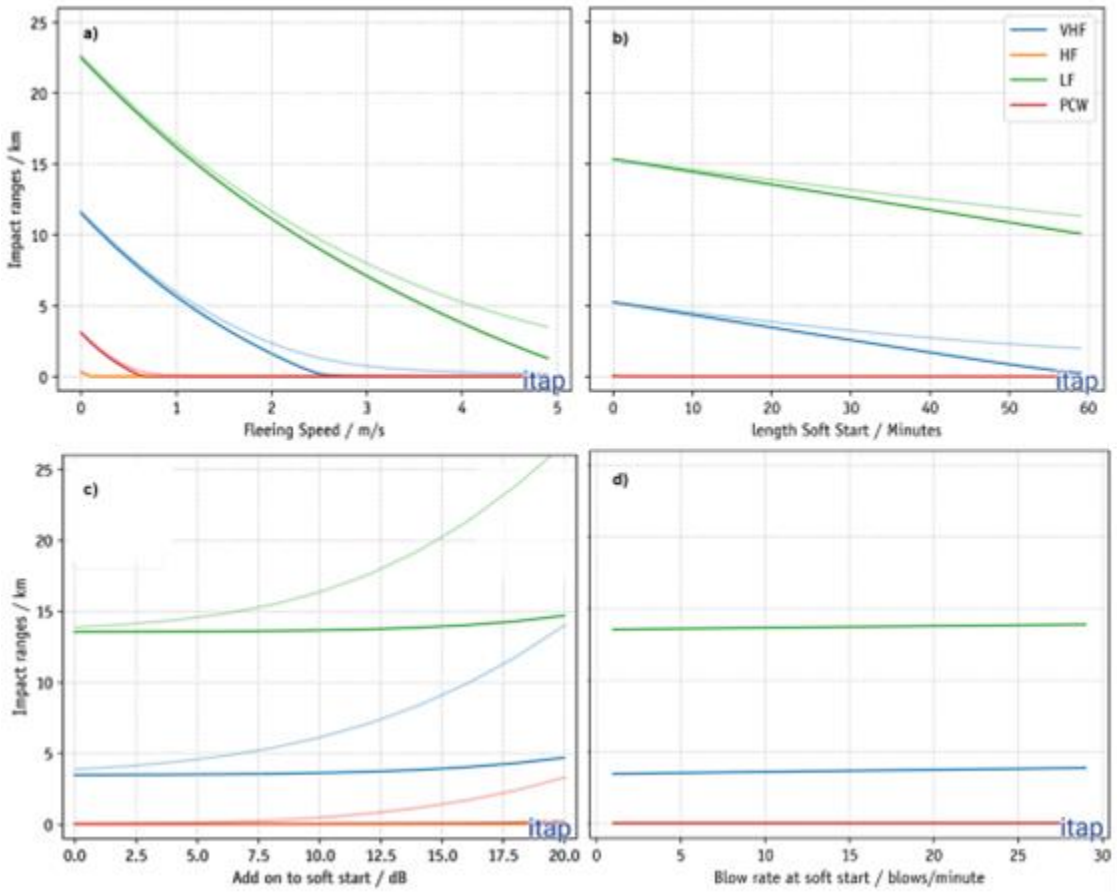

Figure 24: Modelled PTS-impact ranges based on the SELcum-thresholds defined by Southall et al. 2019 for different species groups in relation to changes in the animal's fleeing speed (a), the soft start duration (b), SELSS of soft start pile strikes (c) and blow rate during soft start (d). Base Scenario: 8 m Monopile at 40 m water depth with 5,000 blows of 4,000 kJ blow energy and a blow rate of 30 blows/minute, preceded by a 20 min soft start of 400 kJ strikes with a blow rate of 1 strike per minute, and a 20 min ramp up from 400 kJ to 4,000 kJ at a 30 strike per minute blow rate. For comparison, the same scenarios were repeated with a blow rate of 28 blows per minute during the soft start (lighter colour curves).

4. Conclusion

- Blow energy, pile diameter and foundation type have a significant influence on pile driving noise, confirming results of Bellmann et al. (2020). The amount of hammer energy emitted into the sea as sound is therefore very much dependent on the design envelope;
- The parameters mentioned above do not explain all the variation within the noise levels emitted during impact piling. Variation could be caused by soil influences, for example, which could not be investigated in detail this study due to a lack of information. Also, the quality of the installation execution (adaptation of the anvil and hammer to the pile, use of pile followers, decoupling of the pile gripper, etc.) both in the run-up to piling and afterwards cannot be recorded. Further possible influences on the sound emission could be the different types of installation vessels such as jack-up barges or floating vessels, different types of grippers and impact hammers. Thus, there are other possible influencing parameters here, the impact of which cannot be determined;
- Uncertainties, and thereby variation, in the measurement results can also be caused by methodological constraints, e.g., the assignment of hammer energies to the measured pile strike sound via the hammer logs, or the precision of the measuring equipment;
- For pin piles, penetration depth influences the piling noise significantly as soon as the pile submerges under water. This is likely mainly influenced by a decrease in the sound emitting pile surface in the water column with increasing penetration of the pile into the soil;
- The negative correlation between penetration depth and piling noise might also be due to the pile getting firmer with increasing water depth, leading to less vibrations of the pile;
- Water depth is generally negatively correlated with the SEL since the radiating surface of the pile under water increases with increasing water depth (unless the pile submerges below the water surface), and thereby the energy per m² pile surface decreases, given a constant hammer energy. However, by taking very shallow water (< 15 m) into account the SEL will be reduced significantly with decreasing water depth due to the increasing cut-off frequency of sound propagation;
- Soil conditions are likely an influencing factor on the pile strike sound levels. When penetrating harder soil layers, less of the hammer energy will be converted to kinetic energy while more of the hammer energy will be converted to underwater sound energy. The number of strikes needed per 25 cm penetration depth (blow count) might be a good proxy for the soil condition, as more strikes per penetration depth are needed for harder soil layers;
- The correlation between penetration depth, blow count (as a proxy for soil layer properties) and sound levels might need further investigations to quantify the relationship for inclusion into noise modelling;
- Drivability tests conducted post consent can give information on the blow count, which could inform the noise modelling. However, further investigations are needed to quantify the correlation between blow count and sound levels;
- The comparison of modelling results of the noise impact assessments of impact pile driving with measurements in the field has been difficult since modelling reports have not been tailored to inform such comparison, and at some OWFs, noise monitoring was not sufficient to inform transmission loss calculations;

- Minimum standard requirements are needed for modelling reports and noise monitoring to allow sufficient comparison;
- The use of the PTS-criterion based on an SEL_{cum} for impact assessment has scarcely been implemented in the impact assessments relating to the measured data sets used in this study, but is used more frequently in more recent noise impact assessments in currently consented OWFs;
- The comparison of measured versus modelled SEL_{cum} for the one OWF where this comparison was possible, showed an underestimation of modelled impact ranges vs the measurements for VHF cetacean, although an overestimate would have been expected due to the use of lower hammer energies and less pile strikes for the installation compared to the modelling. The underestimation is likely due to differences in the frequency content of the modelled vs measured spectrum of the pile strike; The measured spectrum has more energy in the higher frequency range than the spectrum used for modelling;
- The estimation of SEL_{cum} impact ranges with noise modelling is influenced by the assumed frequency spectrum of the pile strike, the number of pile strikes, the blow energy and diameter of the pile, as seen in the sensitivity analysis. Differences in these assumptions to the values occurring in the field during construction will mainly result in an over- or underestimation of impact ranges for the low frequency cetacean and very high frequency cetacean species group. Estimated impact ranges for pinnipeds and high frequency cetacean (dolphins) have been essentially zero for the baseline scenarios and were therefore hardly affected by changing the values of one of the parameters mentioned above;
- Differences in the soft start assumed in the noise modelling compared to the soft start occurring in the field during construction will again mainly affect impact ranges estimated for low frequency cetacean and very high frequency cetacean species group, as impact ranges for pinnipeds and dolphins are expected to be small based on the current Southall et al. (2019) PTS thresholds. Modelling pinniped impact ranges is, however, sensitive to fleeing speed when slow speeds are considered;
- While uncertainties in some parameters do have an influence on the impact ranges (such as the frequency spectrum, blow energy, pile diameter etc), others seem not to influence the impact ranges too much (such as a moderate increase in SEL during soft start or the blow rate during soft start). However, the later must be taken with caution as these results are only true for the given scenario, and only if the parameters are changed in isolation. Changing blow rate while also changing the SEL_{ss}, will certainly lead to an increase in impact ranges at some point.

5. Recommendations

Based on the results of this study, the following recommendations are provided to improve noise risk assessments, noise monitoring and reporting, as well as the validation of modelled impact predictions from technical and ecological points of view.

5.1. Noise risk assessments

The influencing factors hammer energy, pile diameter and foundation type should be considered in the noise modelling for risk assessments. Penetration depth has a major influence on the sound levels for pin pile installation (submerged pile driving activities), the mixed model also shows a negative correlation between penetration depth and SEL_{SS} for monopile installation. Penetration depth should therefore also be considered in modelling of piling noise. Information on and influence of soil condition on sound levels may be included in the modelling approach by considering the number of strikes per meter or 25 cm penetration depth. This factor could be considered in the post-consent modelling when drivability tests reveal such information. However, further investigations on the influence of sediment layer properties on SEL_{ss} are needed. The statistical approach here aimed to capture, through replication, the physics of the interaction between the hammer strike, sediment property, and depth of the pile at each windfarm. However, the poor ability of the models to generalise (k-fold cross validation) to windfarms out with the data they were built on indicated that additional covariates and physical properties are needed.

Modelling reports need to detail general description of the noise model used, its source level approach as well as the transmission loss approach. In general, the terminology of the ISO 18405 (2017) should be used for any metric.

The relevant input parameters must be described and defined, such as:

- A project and foundation specific 1/3 octave spectrum, which is unmitigated and realistic based on the pile design and site-specific conditions.
- Source levels and sound levels at 750 m (SEL_{SS} as well as zero-to-peak SPL $L_{p,pk}$).
- A quantification and visualisation of the used transmission loss over distance.
- A detailed description of the piling sequence incl. soft start, blow rate, hammer energies and number of pile strikes.
- Quantification of any weighting function used.
- In case of application of any noise mitigation or noise abatement system a realistic frequency dependant insertion loss (minimum 1/3 octave spectrum) shall be used.

For the assessment of impact ranges on marine mammals, BE and UB scenarios should be modelled, based on the drivability analysis if feasible. The likelihood of these scenarios should be considered in the overall assessment of the impact, e.g., how many piles of a windfarm would more likely need to be driven with upper bound conditions vs best estimate conditions. Furthermore, a realistic swim speed of the species must be assumed when modelling SEL_{cum}. It must be considered that a swim speed gained from animal borne tags or surfacing is faster than the speed an animal moves away perpendicular from the sound source in cases where perpendicular fleeing is assumed.

An assessment of SEL_{cum} impact ranges using a stationary animal is helpful solely for comparison reasons with the field data.

To enable a comparison of the performance of the different proprietary models used by the various noise model companies, a standardised piling scenario with project and site-specific information as inputs should be defined to enable a comparison of defined output (e.g., sound levels at various distances to the pile site and resulting impact ranges) (similar to tests described in Lippert et al., 2016). A qualitative statement of which model predicts the most accurate results is only possible by validation with field measurements.

5.2. Noise monitoring and reporting

Noise measurements during the installation of foundations need to be sufficient to enable a good estimate of the transmission loss. The field study and the reporting should be conducted according to recognised standards. In the following, we list the minimum standards required for noise measurements and reporting.

5.2.1. Field measurements

- Measurement devices, mooring and calibration should be selected in accordance with ISO 18406:2017.
- Ideally, a minimum of four measurement positions at increasing distances in one cardinal direction should be applied to determine project specific transmission loss in case of a flat bathymetry. The distance categories are:
    - 750 m.
    - 1,500 to 3,000 m.
    - 5,000 to 8,000 m.
    - > 10,000 m.
- In case of a non-flat bathymetry, it might be required to measure the project-specific transmission loss in different cardinal directions.

5.2.2. Analysis and Reporting

- Analysis and reporting should follow the ISO 18406:2017 guidance regarding noise metrics and site-specific conditions.
- In the case of measurements for the evaluation of an applied noise abatement system the measurements, analysis and reporting should be in accordance with the DIN SPEK 45653 (2017).
- For the comparison of noise measurements with modelling, the report should, for all measuring positions and piles, include as a minimum:
    - 1/3 octave spectrum of the piling noise.
    - A plot presenting SEL_{SS} versus blow energy.
    - Plots presenting SEL_{SS}, $L_{p,pk}$ and SPL versus time. The SPL gives an idea of the prevailing background conditions.

- Plot of the measured project specific transmission loss per pile.
- Table of the relevant impact ranges of the EIA based on the measured data and project-specific transmission loss.
- The hammer log should be included as an annex to the report including the relevant parameters like piling duration, number of single strikes, blow rate, blow count, penetration depth etc.

5.3. Validation of modelled impact predictions

A comparison of the unweighted SEL_{SS} versus distance from the measured and the modelled data will give a good idea if the noise model predictions are suitably representing the noise emitted during pile driving at sea. A comparison of the piling profile used for modelling and applied in the field will furthermore give an estimate on potential under- or overestimation of the SEL_{cum} impact ranges, and if the predefined piling profile was followed in the field.

For allowing an easy and suitable procedure to conduct the comparison, and any further studies involving the measured data, the following recommendations are made for the provision of the measurement data by the consultants acquiring the data / the developer to the regulators for potential further analysis by secondary consultants:

- Requirements for raw data submission for a quality assurance:
  - Sufficient calibration information in accordance with ISO 18406:2017 and ISO 17025:2018 (sensor sensitivity, calibration procedure, serial number, calibration certificates etc.).
  - Standardised format of raw data (e.g., PCM WAV).
  - Standardised time information (e.g., UTC; Not local time).
  - Sufficient accompanying documents e.g., digital readable hammer logs, mooring concept.
- Post processed data and metadata in digital readable tables/csv files according to the standardized data format of HELCOM and OSPAR, for example:
  - Sound levels versus time (broad band levels as well as 1/3 octave spectra).
  - Weighting functions (if not publicly available).
  - Hammer logs including time of pile strikes and associated hammer energy, penetration depth/blow count per 25 cm penetration depth.
  - Standardised time information (e.g., UTC; Not local time).
  - Sufficient accompanying documents e.g., calibration information, hammer log, mooring concept, measurement coordinates.

Acknowledgements

We are grateful to The Carbon Trust and the Offshore Renewable Joint Industry Project Steering Group for the support throughout the project: Oliver Patrick and Liam Leahy (The Carbon Trust), Ed Salter (The Crown Estate), Kate Bellow (Crown Estate Scotland), Scottish Government Officials (Marine Scotland and Marine Scotland Science) , Tom Anderson (RWE), Nancy Mclean (EDF), Catarina Rei (Ocean Winds), Magnus Eriksen (Equinor), Ben King (Red Rock Power), Gareth Edwards Johnson (Ørsted), Koen Broker (Shell), Emma Ahart (SSE Renewables) and Lars Oberbeck (TOTAL). Thank you to the Project Expert Panel: Claire Ludgate (Natural England), Sonia Mendes (JNCC), Caroline Carter (NatureScot), Tom Stringell (Natural Resource Wales), Jürgen Weissenberger (Equinor), Koen Broker (Shell), Ben King (Red Rock Power) and Matej Simurda (Ørsted).

We would like to thank the following companies and organisations for granting permissions to use and provide their data: Ørsted, Scottish Power, SSE Renewables, Ocean Winds, RWE, Ocean Breeze Energy, Global Tech I, Northland Power, ENECO and BSH.

The project was funded by Total Energies, Marine Scotland, Equinor, SSE Renewables, EDF Renouvelables, The Crown Estate, RWE, Shell, Crown Estate Scotland, SDIC RedRock Power Limited, Ocean Wind and Ørsted under Stage 2 of the Offshore Renewable Joint Industry Project for Offshore Wind.

Literature Cited

Bates, D., M. Mächler, B. Bolker, and S. Walker. 2014. Fitting linear mixed-effects models using lme4. arXiv preprint arXiv:1406.5823.

Bellew, S. 2017. Piling Strategy.

Bellmann, M., A. May, T. Wendt, S. Gerlach, P. Remmers, and J. Brinkmann. 2020. Underwater noise during percussive pile driving: Influencing factors on pile-driving noise and technical possibilities to comply with noise mitigation values. itap GmbH, Oldenburg.

BOWL. 2012a. Beatrice Offshore Wind Farm Environmental Statement. Section 12: Wind Farm Marine Mammals.

BOWL. 2012b. Beatrice Offshore Wind Farm Environmental Statement. Section 12: Wind Farm Physical Processes and Geomorphology.

BOWL. 2015. Beatrice Offshore Wind Farm Piling Strategy. LF000005.

Dähne, M., A. Gilles, K. Lucke, V. Peschko, S. Adler, K. Krugel, J. Sundermeyer, and U. Siebert. 2013. Effects of pile-driving on harbour porpoises (Phocoena phocoena) at the first offshore wind farm in Germany. Environmental Research Letters 8:025002.

de Jong, C., B. Binnerts, S. Robinson, and L. Wang. 2021. Guidelines for modelling ocean ambient noise. Report of the EU INTERREG Joint Monitoring Programme for Ambient Noise North Sea (JOMOPANS).

De Jong, C. A. f., and M. A. Ainslie. 2008. Underwater radiated noise due to the piling for the Q7 Offshore Wind Park. Journal of the Acoustical Society of America 123:2987.

Energistyrelsen. 2022. Guidelines for underwater noise, Prognosis for EIA and SEA assessments.

Farcas, A., P. M. Thompson, and N. D. Merchant. 2016. Underwater noise modelling for environmental impact assessment. Environmental Impact Assessment Review 57:114-122.

Faulkner, R. C., A. Farcas, and N. D. Merchant. 2018. Guiding principles for assessing the impact of underwater noise. Journal of Applied Ecology.

GoBe. 2020a. Hornsea Project Two Offshore Wind Farm. Generation Assets Marine Mammal Mitigation Protocol (MMMP).

GoBe. 2020b. Hornsea Project Two Offshore Wind Farm. Marine Mammal Mitigation Protocol for Substations.

Hastie, G., T. Gotz, D. Russell, V. Janik, P. Thompson, and N. D. Merchant. 2016. Range dependent characteristics of impulsive sounds: implications for marine mammal behavioural responses and auditory damage. Draftmanuscriptto DECC 4th May2016.

ISO. 1998. ISO 13261-1:1998 Sound power rating of air-conditioning and air-source heat pump equipment — Part 1: Non-ducted outdoor equipment.

ISO. 2017a. ISO 18405 Underwater Acoustics—Terminology. International Organization for Standardization Geneva.

ISO. 2017b. ISO 18406:2017 Underwater Acoustics—Measurement of radiated underwater sound from percussive pile driving. International Organization for Standardization Geneva.

itap. 2020. Hornsea 2 Offshore Wind Farm Technical report: Modelling of underwater noise emission during pile-driving construction work. Version 4., itap, Oldenburg.

Jensen, F. B., W. A. Kuperman, M. B. Porter, H. Schmidt, and A. Tolstoy. 2011. Computational ocean acoustics. Springer.

Kastelein, R. A., R. Gransier, M. Marijt, L. Hoek, and H. Winter. 2014. Hearing frequencies of a harbour porpoise (Phocoena phocoena) temporarily affected by played back offshore pile driving sounds. SEAMARCO.

Kastelein, R. A., R. Gransier, J. Schop, and L. Hoek. 2015. Effects of exposure to intermittent and continuous 6–7 kHz sonar sweeps on harbor porpoise (Phocoena phocoena) hearing. The Journal of the Acoustical Society of America 137:1623-1633.

Lippert, S., M. Huisman, M. Ruhnau, O. Estorff, and K. van Zandwijk. 2017. Prognosis of underwater pile driving noise for submerged skirt piles of jacket structures. Pages 2-8 in Proceedings of the UACE 2017 4th Underwater Acoustics Conference and Exhibition, Skiathos, Greece.

Lippert, S., M. Nijhof, T. Lippert, D. Wilkes, A. Gavrilov, K. Heitmann, M. Ruhnau, O. von Estorff, A. Schäfke, and I. Schäfer. 2016. COMPILE—A generic benchmark case for predictions of marine pile-driving noise. Ieee Journal of Oceanic Engineering 41:1061-1071.

Lucke, K., U. Siebert, P. A. Lepper, and M. Blanchet. 2009. Temporary shift in masked hearing thresholds in a harbor porpoise (Phocoena phocoena) after exposure to siesmic airgun stimuli. Journal of the Acoustical Society of America 125:4060-4070.

McGarry, T., O. Boisseau, S. Stephenson, and R. Compton. 2017. Understanding the Effectiveness of Acoustic Deterrent Devices (ADDs) on Minke Whale (Balaenoptera acutorostrata), a Low Frequency

Cetacean. Report for the Offshore Renewables Joint Industry Programme (ORJIP) Project 4, Phase 2. Prepared on behalf of the Carbon Trust.

MMO. 2014. Review of Post-Consent Offshore Wind Farm Monitoring Data Associated with Marine Licence Conditions. A report produced for the Marine Management Organisation, pp 194. MMO Project No: 1031. ISBN: 978-1-909452-24-4.

Moray Offshore Windfarm (East) Limited. 2018. Project Environmental Monitoring Programme: Telford, Stevenson and MacColl Offshore Wind Farms and associated Offshore transmission infrastructure.

Moray Offshore Windfarm (East) Limited. 2019. Wind Farm Piling Strategy. Royal Haskoning DHV.

Müller, A., and C. Zerbs. 2013. Offshore wind farms prediction of underwater sound: Minimum requirements on documentation. Federal Maritime and Hydrographic Agency (BSH), Hamburg, Germany.

National Marine Fisheries Service. 2018. Revisions to: Technical Guidance for Assessing the Effects of Anthropogenic Sound on Marine Mammal Hearing (Version 2.0): Underwater Thresholds for Onset of Permanent and Temporary Threshold Shifts. Page 167. U.S. Department of Commerce, NOAA, Silver Spring.

Natural Power. 2012. Moray Offshore Renewables Ltd Environmental Statement Technical Appendix 7.3 A - Marine Mammals EnvironmentalImpact Assessment.

Nedwell, J., A. Turnpenny, J. Lovell, S. Parvin, R. Workman, J. Spinks, and D. Howell. 2007. A validation of the dBht as a measure of the behavioural and auditory effects of underwater noise. Subacoustech Report No. 534R1231.

NMFS. 2016. Technical Guidance for Assessing the Effects of Anthropogenic Sound on Marine Mammal Hearing: Underwater Acoustic Thresholds for Onset of Permanent and Temporary Threshold Shifts. Page 189. U.S. Department of Commerce, Silver Spring.

Parvin, S., J. Nedwell, and E. Harland. 2007. Lethal and physical injury of marine mammals, and requirements for Passive Acoustic Monitoring. Subacoustech Report Reference: 565R0212, February.

Popper, A., A. Hawkins, R. Fay, D. Mann, S. Bartol, T. Carlson, S. Coombs, W. Ellison, R. Gentry, and M. Halvorsen. 2014. Sound exposure guidelines for fishes and sea turtles. Springer Briefs in Oceanography. DOI 10:978-973.

R Core Team. 2021. R: A language to environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Royal Haskoning DHV. 2019. Moray East Offshore Windfarm Wind Farm Piling Strategy.

RPS. 2005. London Array Offshore Wind Farm Environmental Statement. Volume 1: Offshore Works. London Array Ltd.

RSK Environmental Ltd. 2012a. Rampion Offshore Wind Farm ES Section 6 – Physical Environment.

RSK Environmental Ltd. 2012b. Rampion Offshore Wind Farm ES Section 10–Marine Mammals.

RWE npower renewables. 2012a. Triton Knoll Offshore Wind Farm Environmental Statement. Volume 2: Chapter 2 – Physical Processes.

RWE npower renewables. 2012b. Triton Knoll Offshore Wind Farm Environmental Statement. Volume 2: Chapter 5 - Marine Mammals.

Scottish Power Renewables. 2012a. East Anglia ONE Offshore Windfarm Environmental Statement Volume 2 Offshore: Chapter 6 - Marine Geology, Oceanography and Physical Processes.

Scottish Power Renewables. 2012b. East Anglia ONE Offshore Windfarm Environmental Statement Volume 2 Offshore: Chapter 8 - Underwater Noise and Vibration and Electromagnetic Fields.

Scottish Power Renewables. 2012c. East Anglia ONE Offshore Windfarm Environmental Statement Volume 2 Offshore: Chapter 11 - Marine Mammals.

Simpson, M. 2016. Noise monitoring during foundation installation (G04, J02, I04, F04, J04) Rampion Offshore Wind Farms. Baker Consultants.

SMart wind Limited. 2013a. Hornsea Offshore Wind Farm Project One. Environmental Statement Volume 2 - Offshore. Chapter 1 Marine Processes.

SMart wind Limited. 2013b. Hornsea Offshore Wind Farm Project One. Environmental Statement Volume 2 - Offshore. Chapter 4 Marine Mammals.

SMart Wind Limited. 2015a. Hornsea Offshore Wind Farm Project Two – Environmental Statement: Volume 2 – Offshore, Chapter 1 – Marine Processes.

SMart Wind Limited. 2015b. Hornsea Offshore Wind Farm. Project Two. Environmental Statement. Volume 2 - Offshore. Chapter 4: Marine Mammals.

SMRU Consulting. 2017. Hornsea Project One Offshore Wind Farm Offshore Environmental Monitoring Strategy – Marine Mammal Mitigation Protocol.

Southall, B., J. J. Finneran, C. Reichmuth, P. E. Nachtigall, D. R. Ketten, A. E. Bowles, W. T. Ellison, D. Nowacek, and P. Tyack. 2019. Marine Mammal Noise Exposure Criteria: Updated Scientific Recommendations for Residual Hearing Effects. Aquatic Mammals 45:125-232.

Southall, B. L., A. E. Bowles, W. T. Ellison, J. J. Finneran, R. L. Gentry, C. R. J. Greene, D. Kastak, D. R. Ketten, J. H. Miller, P. E. Nachtigall, W. J. Richardson, J. A. Thomas, and P. L. Tyack. 2007. Marine mammal noise exposure criteria: initial scientific recommendations. Aquatic Mammals 33:411-414.

Statutory Instruments. 2016. Infrastructure planning - The Hornsea Two Offshore Wind Farm Order 2016.in E. a. I. S. Department for Business, editor.

Subacoustech Environmental Ltd. 2012a. Beatrice Offshore Wind Farm Environmental Statement. Annex A: Underwater Noise Modelling Technical Report.

Subacoustech Environmental Ltd. 2012b. Underwater Noise Impact Modelling in Support of the Triton Knoll Offshore Wind Farm Development

Subacoustech Environmental Ltd. 2017. Additional underwater noise modelling for HOW01.

Thiele, R., and G. Schellstede. 1980. Standardwerte zur Ausbreitungsdämpfung in der Nordsee.

Thompson, P. M., I. M. Graham, B. Cheney, T. R. Barton, A. Farcas, and N. D. Merchant. 2020. Balancing risks of injury and disturbance to marine mammals when pile driving at offshore windfarms. Ecological Solutions and Evidence 1.

Tougaard, J., J. Carstensen, J. Teilmann, S. Henrik, and P. Rasmussen. 2009. Pile driving zone of responsiveness extends beyond 20 km for harbor porpoises (Phocoena phocoena (L.)) (L). Journal of the Acoustical Society of America 126:11-14.

Urick, R. J. 1983. Principles of underwater sound, 3rd ed. Peninsula Publishing, Los Altos.

Verfuss, U., M. Bellman, R. Kühler, P. Remmers, R. Plunkett, and C. Sparling. 2018. Underwater noise monitoring during monopile installation at the Hornsea Project One Offshore Windfarm – cumulative SEL report. Report number SMRUC-GOB-2018-014 provided to GoBe Consultants & Ørsted, October 2018 (unpublished).

Verfuss, U. K., R. Plunkett, C. G. Booth, and J. Harwood. 2016. Assessing the benefit of noise reduction measures during offshore wind farm construction on harbour porpoises. WWF-UK.

Wang, L., K. Heaney, T. Pangerc, P. Theobald, S. Robinson, and M. Ainslie. 2014. Review of underwater acoustic propagation models.

Weston, D. 1959. Guided propagation in a slowly varying medium. Proceedings of the Physical Society (1958-1967) 73:365.

Weston, D. 1968. Sound Focusing and Beaming in the Interference Field Due to Several Shallow-Water Modes. The Journal of the Acoustical Society of America 44:1706-1712.

Weston, D. 1976. Propagation in water with uniform sound velocity but variable-depth lossy bottom. Journal of Sound and Vibration 47:473-483.

Appendix 1:

Pulse length, percentage energy signal duration ($\tau 90$)

Time during which 90% of unweighted sound exposure occurs in seconds.

Zero-to-peak Sound Pressure Level ($L_{p,pk}$)

This parameter is a measure for sound pressure peaks. Unlike SPL and SEL_{SS}, Zero-to-peak Sound Pressure Level does not involve any averaging (ISO 2017a):

$$L_{p,pk} = 20\ \log_{10}\left(\frac{|p_{pk}|}{p_0}\right) \text{ [dB re 1µPa]}$$

Equation 3

where
$|p_{pk}|$ - maximum absolute determined Sound Pressure,

p_0 - reference sound pressure (1 µPa).

(Energy-) equivalent continuous Sound Pressure Level (SPL)

SPL is the most common amplitude measurement in acoustics and is defined according to ISO (2017a) as follows:

$$SPL = 10\ \log_{10}\left(\frac{1}{T}\int_0^T \frac{p(t)^2}{p_0^2}\ \mathrm{d}t\right) \text{ [dB]}$$

Equation 4

where
p(t) - time-variant sound pressure,
p_0 - reference sound pressure (1 µPa),
T - averaging time.

Sound Exposure Level (SEL)

The sound exposure – E and the resulting SEL are defined as follows (ISO 2017a):

$$E = \frac{1}{T_0}\int_{T_1}^{T_2} \frac{p(t)^2}{p_0^2}\ \mathrm{d}t$$

Equation 5

$$SEL = 10\ \log_{10}\left(\frac{1}{T_0}\int_{T_1}^{T_2}\frac{p(t)^2}{p_0^2}\,\mathrm{d}t\right)\ [\mathrm{dB}]$$

Equation 6

where

T_1 and T_2 - start and end time of the averaging (to be determined so that the sound event (pile strike) is between T_1 and T_2),

T_0 - reference 1 second.

Cumulative Sound Exposure Level (SEL_{cum})

A value for the dose of sound energy an animal may receive is the cumulative Sound Exposure Level (SEL_{cum}) and is defined as follows (Energistyrelsen 2022):

$$SEL_{cum} = 10\ \log_{10}\left(\frac{E_{cum}}{E_{ref}}\right)\ [\mathrm{dB}]$$

Equation 7

With the cumulative sound exposure E_{cum} for *N* sound events with the frequency sound exposure E_n

$$E_{cum} = \sum_{n=1}^{N} E_n$$

Equation 8

and the reference exposure $E_{ref} = p_{ref}^2 \cdot T_{ref}$, in which p_{ref} is the reference sound pressure 1 µPa and T_{ref} the reference duration 1 s.

One-third octave band frequency spectrum

The frequency is separated in frequency bands of one third of an octave. The frequency ratio between two successive frequency bands is $2^{1/3}$. One third octave bands are defined in ISO (1998).