Environmental Impact
Assessment Report

NISA

North Irish Sea Array

Volume 9: Appendices (Offshore)

Appendix 14.3
Marine Mammal Uncertainties and Limitations

Statkraft
CIP Copenhagen Infrastructure Partners
GoBe APEM Group
ARUP

Data Limitations

Authors: Booth, C; Chudzinska, M; Sinclair, RR; Wilder, F; Klementisova, K

Appendix 14.3: Data limitations

There are uncertainties relating to the underwater noise modelling and impact assessment for the proposed development. Broadly, these relate to predicting exposure of animals to underwater noise, predicting the response of animals to underwater noise, and predicting potential population consequences of disturbance from underwater noise. Further detail of such uncertainty is set out below.

PTS assumptions and limitations

There are no empirical data on the threshold for auditory injury in the form of PTS onset for marine mammals, as to test this would be inhumane. Therefore, PTS onset thresholds are estimated based on extrapolating from TTS onset thresholds. For pulsed noise, such as piling, NOAA have set the onset of TTS at the lowest level that exceeds natural recorded variation in hearing sensitivity (6dB), and assume that PTS occurs from exposures resulting in 40dB or more of TTS measured approximately four minutes after exposure (NMFS 2018).

Proportion impacted

It is important to note that it is expected that only 18-19% of animals are predicted to actually experience PTS at the PTS-onset threshold level.

This was the approach adopted by Donovan et al. (2017) to develop their dose-response function implemented into the SAFESIMM (Statistical Algorithms For Estimating the Sonar Influence on Marine Megafauna) model, based on the data presented in Finneran et al. (2005). Therefore, where PTS-onset ranges are provided, it is not expected that all individuals within that range will experience PTS. Therefore, the number of animals predicted to be within PTS-onset ranges are precautionary, since they assume that all animals are impacted.

Exposure to noise

There are uncertainties relating to the ability to predict the exposure of animals to underwater noise, as well as in predicting the response to that exposure. These uncertainties relate to a number of factors: the ability to predict the level of noise that animals are exposed to, particularly over long periods of time; the ability to predict the numbers of animals affected, and the ability to predict the individual and ultimately population consequences of exposure to noise. These are explored in further detail in the paragraphs below.

The propagation of underwater noise is relatively well understood and modelled using standard methods. However, there are uncertainties regarding the amount of noise actually produced by each pulse at source and how the pulse characteristics change with range from the source. There are also uncertainties regarding the position of receptors in relation to received levels of noise, particularly over time, and understanding how the position of receptors in the water column may affect received level. Noise monitoring is not always carried out at distances relevant to the ranges predicted for effects on marine mammals, so effects at greater distances remain un-validated in terms of actual received levels. The extent to which ambient noise and other anthropogenic sources of noise may mask signals from the offshore wind farm construction are not specifically addressed. The dose-response functions for porpoise include behavioural responses at noise levels down to 120dB SELss which may be indistinguishable from ambient noise at the ranges these levels are predicted.

Data Limitations

Cumulative PTS

The cumulative sound exposure level (SEL_{cum}) is energy-based and is a measure of the accumulated sound energy an animal is exposed to over an exposure period. An animal is considered to be at risk of experiencing "cumulative PTS" if the SEL_{cum} exceeds the energy-based threshold. The calculation of SEL_{cum} is undertaken with frequency-weighted sound levels, using species group-specific weighing functions to reflect the hearing sensitivity of each functional hearing group. To assess the risk of cumulative PTS, it is necessary to make assumptions on how animals may respond to noise exposure, since any displacement of the animal relative to the noise source will affect the sound levels received. For this assessment, it was assumed that animals would flee from the pile foundation at the onset of piling. A fleeing animal model was therefore used to determine the cumulative PTS impact ranges, to determine the minimum distance to the pile site at which an animal can start to flee, without the risk of experiencing cumulative PTS.

There is much more uncertainty associated with the prediction of the cumulative PTS impact ranges than with those for the instantaneous PTS. One reason is that the sound levels an animal receives, and which are cumulated over a whole piling sequence are difficult to predict over such long periods of time, as a result of uncertainties about the animal's (responsive) movement in terms of its changing distance to the sound source and the related speed, and its position in the water column.

Another reason is that the prediction of the onset of PTS (which is assumed to be at the SEL_{cum} threshold values provided by Southall et al. (2019) is determined with the assumptions that:

- the amount of sound energy an animal is exposed to within 24 hours will have the same effect on its auditory system, regardless of whether it is received all at once (i.e., with a single bout of sound) or in several smaller doses spread over a longer period (called the equal-energy hypothesis); and
- the sound keeps its impulsive character, regardless of the distance to the sound source.

However, in practice:

- there is a recovery of a threshold shift caused by the sound energy if the dose is applied in several smaller doses (e.g., between pulses during pile driving or in piling breaks) leading to an onset of PTS at a higher energy level than assumed with the given SELcum threshold; and
- pulsed sound loses its impulsive characteristics while propagating away from the sound source, resulting in a slower shift of an animal's hearing threshold than would be predicted for an impulsive sound.

Both assumptions, therefore, lead to a conservative determination of the impact ranges and are discussed in further detail in the sections below.

Modelling the SEL_{cum} impact ranges of PTS with a 'fleeing animal' model, as is typical in noise impact assessments, are subject to both above-mentioned uncertainties and the result is a highly precautionary prediction of impact ranges. As a result of these and the uncertainties on animal movement, model parameters, such as swim speed, are generally highly conservative and, when considered across multiple parameters, this precaution is compounded therefore the resulting predictions are very precautionary and very unlikely to be realised.

Equal energy hypothesis

The equal-energy hypothesis assumes that exposures of equal energy produce equal amounts of noise-induced threshold shift, regardless of how the energy is distributed over time. However, a continuous and an intermittent noise exposure of the same SEL will produce different levels of TTS (Ward 1997). Ward (1997) highlights that the same is true for impulsive noise, giving the example of

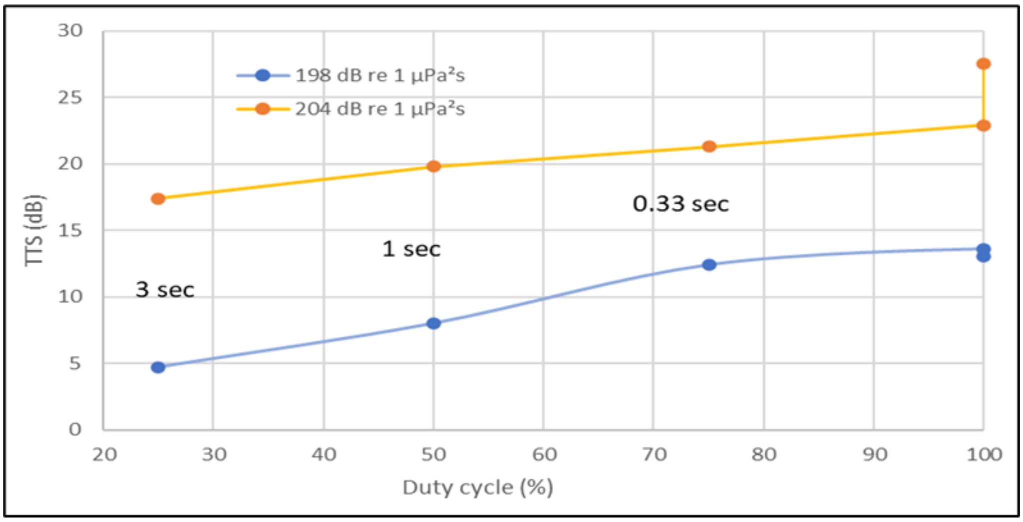
Data Limitations

simulated gunfires of the same SEL_{cum} exposed to human, where 30 impulses with an SPLpeak of 150dB re 1m Pa result in a TTS of 20dB, while 300 impulses of a respectively lower SPLpeak did not result in any TTS.

Finneran (2015) showed that several marine mammal studies have demonstrated that the temporal pattern of the exposure does in fact affect the resulting threshold shift (e.g., Kastak et al. 2005, Mooney et al. 2009, Finneran et al. 2010, Kastelein et al. 2013a). Intermittent noise allows for some recovery of the threshold shift in between exposures, and therefore recovery can occur in the gaps between individual pile strikes and in the breaks in piling activity, resulting in a lower overall threshold shift, compared to continuous exposure at the same SEL.

Kastelein et al. (2013a) showed that, for seals, the threshold shifts observed did not follow the assumptions made in the guidance regarding the equal-energy hypothesis. The threshold shifts observed were more similar to the hypothesis presented in Henderson et al. (1991), whereby hearing loss induced due to noise does not solely depend upon the total amount of energy, but on the interaction of several factors such as the level and duration of the exposure, the rate of repetition, and the susceptibility of the animal. Therefore, the equal energy hypothesis assumption behind the SEL_{cum} threshold is not valid, and as such, models will overestimate the level of threshold shift experienced from intermittent noise exposures.

Another detailed example to give is the study of (Kastelein et al. 2014), where a harbour porpoise was exposed to a series of 1-2kHz sonar down-sweep pulses of 1-second duration of various combinations, with regard to received sound pressure level, exposure duration and duty cycle (% of time with sound during a broadcast) to quantify the related threshold shift. The porpoise experienced a 6 to 8dB lower TTS when exposed to sound with a duty cycle of 25% compared to a continuous sound (Graph 1). A one second silent period in between pulses resulted in a 3 to 5dB lower TTS compared to a continuous sound (Graph 1).



Graph 1 Temporary threshold shift (TTS) elicited in a harbour porpoise by a series of 1-2kHz sonar down-sweeps of 1 second duration with varying duty cycle and a constant SEL_{cum} of 198 and 204dB re1 µPa²s, respectively. Also labelled is the corresponding 'silent period' in-between pulses. Data from Kastelein et al. (2014)

Kastelein et al. (2015b) showed that the 40dB hearing threshold shift (the PTS-onset threshold) for harbour porpoise, is expected to be reached at different SEL_{cum} levels depending on the duty cycle: for a 100% duty cycle, the 40dB hearing threshold shift is predicted to be reached at a SEL_{cum} of 196dB re 1µPa2s, but for a 10% duty cycle, the 40dB hearing threshold shift is predicted to be reached at a SEL_{cum} of 206dB re 1µPa2s (thus resulting in a 10dB re 1µPa2s difference in the threshold).

Data Limitations

Pile strikes are relatively short signals; the signal duration of monopile pile strikes may range between 0.1 seconds (De Jong and Ainslie 2008) and approximately 0.3 seconds (Dähne et al. 2017) measured at a distance of 3.3km to 3.6km. Duration will however increase with increasing distance from the pile site.

For the pile driving for the proposed development, the soft start is 10 blows/min and the ramp-up is 30 blows/min. Assuming a signal duration of around 0.5 seconds for a pile strike, the soft start has been an 8.3% duty cycle (0.5 seconds pulse followed by 5.5 seconds silence) and the ramp-up has been a 25% duty cycle (0.5 second pulse followed by 1.5 second silence). In the study of Kastelein et al. (2014), a silent period of 3 seconds corresponds to a duty cycle of 25%. The reduction in TTS at a duty cycle of 25% is 5.5–8.3dB. Assuming similar effects to the hearing system of marine mammals in the array area, the PTS onset threshold would be expected to be around 2.4dB higher than that proposed by Southall et al. (2019) and used in the current assessment, as reasoned in the following section.

Southall et al. (2009) calculates the PTS-onset thresholds based on the assumption that a TTS of 40dB will lead to PTS, and that an animal's hearing threshold will shift by 2.3dB per dB SEL received from an impulsive sound.

This means, if the same SEL elicits a ≥5.5dB lower TTS at 25% duty cycle compared to 100% duty cycle, to elicit the same TTS as a sound of 100% duty cycle, a ≥2.4dB higher SEL is needed with a 25% duty cycle than with a 100% duty cycle. The threshold at which PTS-onset is likely is therefore, expected to be a minimum of 2.4dB higher than the PTS-onset threshold proposed by Southall et al. (2019).

If a 2 or 3dB increase in the PTS-threshold is assumed, then this can make a significant difference to the maximum predicted impact range for cumulative PTS. Table 1 summarises the difference in the predicted PTS impact ranges using the current and adjusted thresholds. In summary, if the threshold accounts for recovery in hearing between pulses, the PTS impact ranges for the NE location decreases from 18.55km for harbour porpoise to 13.43km (+2dB) or 11.18km (+3dB). For minke whale the PTS impact ranges for the NE location decreases from 35.25km to 26.25km (+2dB) or 22.23km (+3dB).

Therefore, accounting for recovery in hearing between pulses by increasing the PTS onset threshold by 2 or 3dB significantly decreases the predicted PTS-onset impact ranges (Table 1). This approach to modelling cumulative PTS is in development and has not yet been fully assessed or peer reviewed. Therefore, the impact assessment will present the cumulative PTS impact ranges using the current Southall et al. (2019) PTS-onset impact threshold.

While more research needs to be conducted to understand the exact magnitude of this effect in relation to pile driving sound, this study proves a significant reduction in the risk of PTS even through short silent periods for TTS recovery as found in pile driving.

Table 1 Difference in predicted cumulative PTS impact ranges if recovery between pulses is accounted for and the PTS-onset threshold is increased by 2 or 3dB

Maximum disturbance distance		Max impact range (km)	Reduction in impact range (km)
Minke Whale			
PTS	183SELcum	35.250	-
PTS + 2dB	185SELcum	26.250	9.000
PTS + 3dB	186SELcum	22.225	13.025
Harbour porpoise			
PTS	155SELcum	18.550	-

Data Limitations

PTS + 2dB	157SELcum	13.425	5.125
PTS + 3dB	158SELcum	11.175	7.375

Impulsive characteristics

Southall et al. (2019) calculated the PTS onset thresholds based on the assumption that an animal's hearing threshold will shift by 2.3dB per dB SEL received from an impulsive sound, but only 1.6dB per dB SEL when the sound received is non-impulsive. The PTS onset threshold for non-impulsive sound is, therefore, higher than for impulsive sound, as more energy is needed to cause PTS with non-impulsive sound compared to impulsive sound. Consequently, an animal subject to both types of sound has been at risk of PTS at an SEL_{cum} that lies somewhere between the PTS onset thresholds of impulsive and non-impulsive sound.

Southall et al. (2019) acknowledges that, as a result of propagation effects, the sound signal of certain sound sources (e.g., impact piling) loses its impulsive characteristics and could potentially be characterised as non-impulsive beyond a certain distance. The changes in noise characteristics with distance generally result in exposures becoming less physiologically damaging with increasing distance as sharp transient peaks become less prominent (Southall et al. 2007). The Southall et al. (2019) updated criteria proposed that, while keeping the same source categories, the exposure criteria for impulsive and non-impulsive sound should be applied based on the signal features likely to be perceived by the animal rather than those emitted by the source. Methods to estimate the distance at which the transition from impulsive to non-impulsive noise are currently being developed (Southall et al. 2019).

Using the criteria of signal duration[1], rise time[2], crest factor[3] and peak pressure[4] divided by signal duration[5], Hastie et al. (2019) estimated the transition from impulsive to non-impulsive characteristics of impact piling noise during the installation of offshore wind turbine foundations at the Wash and in the Moray Firth. Hastie et al. (2019) showed that the noise signal experienced a high degree of change in its impulsive characteristics with increasing distance. Southall et al. (2019) state that mammalian hearing is most readily damaged by transient sounds with rapid rise-time, high peak pressures, and sustained duration relative to rise time. Therefore, of the four criteria used by Hastie et al. (2019), the rise-time and peak pressure may be the most appropriate indicators to determine the impulsive/non-impulsive transition.

Based on this data it is expected that the probability of a signal being defined as "impulsive" (using the criteria of rise time being less than 25 milliseconds) reduces to only 20% between ~2 and 5km from the source.

Predicted PTS impact ranges based on the impulsive noise thresholds may therefore be overestimates in cases where the impact ranges lie beyond this. Any animal present beyond that distance when piling starts will only be exposed to non-impulsive noise, and therefore impact ranges should be based on the non-impulsive thresholds.

[1] Time interval between the arrival of 5% and 95% of total energy in the signal.

[2] Measured time between the onset (defined as the 5th percentile of the cumulative pulse energy) and the peak pressure in the signal.

[3] The decibel difference between the peak sound pressure level (i.e., the peak pressure expressed in units of dB re 1 µPa) of the pulse and the root-mean-square sound pressure level calculated over the signal duration.

[4] The greatest absolute instantaneous sound pressure within a specified time interval.

[5] Time interval between the arrival of 5% and 95% of total energy in the signal.

Data Limitations

It is acknowledged that the Hastie et al. (2019) study is an initial investigation into this topic, and that further data are required in order to set limits to the range at which impulsive criteria for PTS are applied.

Since the Hastie et al. (2019) study, Martin et al. (2020) investigated the sound emission of different sound sources to test techniques for distinguishing between the sound being impulsive or non-impulsive. For impulsive sound sources, they included impact pile driving of four 4-legged jacket foundations installed at around 20m water depth (at the Block Island Wind Farm in the USA). For the impact piling sound, they recorded sound at four distances between ~500m and 9km, recording the sound of 24 piling events. To investigate the impulsiveness of the sound, they used three different parameters and suggested the use of kurtosis[6] to further investigate the impulsiveness of sound. Hamernik et al. (2007) showed a positive correlation between the magnitude of PTS and the kurtosis value in chinchillas, with an increase in PTS for a kurtosis value from 3 up to 40 (which in reverse also means that PTS decreases for the same SEL with decreasing kurtosis below 40). Therefore, Martin et al. (2020) argued that:

- Kurtosis of 0-3 = continuous sinusoidal signal (non-impulsive);
- Kurtosis of 3-40 = transition from non-impulsive to impulsive sound; and
- Kurtosis of 40 = fully impulsive.

For the evaluation of their data, Martin et al. (2020) used unweighted as well as LF-Cetacean (C) and VHFC weighted sound, based on the species-specific weighting curves in Southall et al. (2019) to investigate the impulsiveness of sound. Their results for pile driving are shown in Graph 14.27. For the unweighted and LFC weighted sound, the kurtosis value was >40 within 2km from the piling site. Beyond 2km, the kurtosis value decreased with increasing distance. For the VHFC weighted sound, kurtosis factor is more inconclusive with the median value >40 for the 500m and 9km measuring stations, and at 40 for the stations in between. However, the variability of the kurtosis value for the VHF-C weighted sound increased with distance.

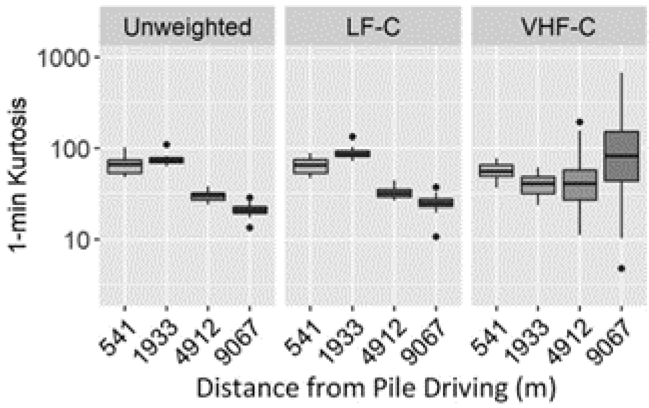


Figure 0.1 The range of kurtosis weighted by LF-C and VHF-C Southall et al. (2019) auditory frequency weighting functions for 30 min of impact pile driving data measured in 25m of water at the Block Island Wind Farm. Boxplots show the median value (horizontal lines, interquartile range (boxes) and outlier values (dots)

From these data, Martin et al. (2020) conclude that the change to non-impulsiveness *"is not relevant for assessing hearing injury because sounds retain impulsive character when SPLs are above EQT [effective quiet threshold]"* (i.e., the sounds they recorded retain their impulsive character while being at sound levels that can contribute to auditory injury).

[6] Kurtosis is a measure of the asymmetry of a probability distribution of a real-valued variable.

Data Limitations

However, SMRU Consulting interpret their results differently. GraphFigure 0.1 clearly shows (for unweighted and LF-C weighted sound) that piling sound loses its impulsiveness with increasing distance from the piling site; the kurtosis value decreases with increasing distance and therefore the sound loses its harmful impulsive characteristics. Based on this study and the study by Hastie et al. (2019), the argument is that the predicted PTS impact ranges based on the impulsive noise thresholds will over-estimate the risk of PTS-onset in cases and at ranges where the likelihood increases that an animal is exposed to sound with much reduced impulsive characteristics.

There are points that need consideration before adopting kurtosis as an impulsiveness measure with the recommended threshold value of 40. Firstly, this value was experimentally obtained for chinchillas that were exposed to noise for a five-day period under controlled conditions. Caution may need to be taken to directly adopt this threshold-value (and the related dose-response of increasing PTS with increasing kurtosis between 3 and 40) to marine mammals in the wild, especially given that the PTS guidance considers time periods of up to 24 hours. Secondly, kurtosis is recommended to be computed over at least 30 seconds, which means that it is not a specific measure that can be used for single blows of a piling sequence. Instead, kurtosis has been recommended to evaluate steady-state noise in order to include the risk from embedded impulsive noise (Goley et al. 2011). Metrics used by Hastie et al. (2019) computed for each pile strike (e.g. risetime) may be more suitable for inclusion in piling impact assessments, as the sound exposure levels received by an animal are considered for each pile strike. It is currently unknown which metric is the most useful and how they correlate with the magnitude of auditory injury in (marine) mammals.

Southall (2021) points out that *"at present there are no properly designed, comparative studies evaluating TTS for any marine mammal species with various noise types, using a range of impulsive metrics to determine either the best metric or to define an explicit threshold with which to delineate impulsiveness"*.

Southall (2021) proposes that the presence of high-frequency noise energy could be used as a proxy for impulsiveness, as all currently used metrics have in common that a high frequency spectral content result in high values for those metrics. This suggestion is an interim approach: *"the range at which noise from an impulsive source lacks discernible energy (relative to ambient noise at the same location) at frequencies ≥ 10 kHz could be used to distinguish when the relevant hearing effect criteria transitions from impulsive to non-impulsive"*.

Southall (2021), however, notes that: *"it should be recognized that the use of impulsive exposure criteria for receivers at greater ranges (tens of kilometers) is almost certainly an overly precautionary interpretation of existing criteria".*

Considering that an increasing proportion of the sound emitted during a piling sequence will become less impulsive (and thereby less harmful) while propagating away from the sound source, and this effect starts at ranges below 5km in all above mentioned examples, the cumulative PTS-onset threshold for animals starting to flee at 5km should be higher than the Southall (2021) threshold adopted for this assessment (i.e., the risk of experiencing PTS becomes lower), and any impact range estimated beyond this distance should be considered as an unrealistic over-estimate, especially when they result in very large distances.

For the purpose of presenting a precautionary assessment, the quantitative impact assessment for the proposed development is based on fully impulsive thresholds, but the potential for overestimation should be noted.

Cumulative PTS Conclusion

Given the above, SMRU Consulting considers that the calculated SEL_{cum} PTS-onset impact ranges are highly precautionary and that the true extent of effects (impact ranges and numbers of animals experiencing PTS) will likely be considerably less than that assessed here.

Data Limitations

Density

There are uncertainties relating to the ability to predict the responses of animals to underwater noise and the number of animals potentially exposed to levels of noise that may cause an impact is uncertain. Given the high spatial and temporal variation in marine mammal abundance and distribution in any area of the sea, it is difficult to predict how many animals may be present within the range of noise impacts. All methods for determining at-sea abundance and distribution suffer from a range of biases and uncertainties.

Predicting Response

In addition, there is limited empirical data available to inform predictions of the extent to which animals may experience auditory damage or display responses to noise. The current methods for prediction of behavioural responses are based on received sound levels, but it is likely that factors other than noise levels alone will also influence the probability of response and the strength of response (e.g., previous experience, behavioural and physiological context, proximity to activities, characteristics of the sound other than level, such as duty cycle and pulse characteristics). However, at present, it is impossible to adequately take these factors into account in a predictive sense. This assessment makes use of the monitoring work that has been carried out during the construction of the Beatrice Offshore Wind Farm and therefore uses the most recent and site-specific information on disturbance to harbour porpoise because of pile driving noise.

There is also a lack of information on how observed effects (e.g. short-term displacement around impact piling activities) manifest themselves in terms of effects on individual fitness, and ultimately population dynamics (see Section 14.2.7 above on marine mammal sensitivity to disturbance and the recent expert elicitation conducted for harbour porpoise and both seal species) in order to attempt to quantify the amount of disturbance required before vital rates are impacted.

Duration of impact

The duration of disturbance is another uncertainty. Studies at Horns Rev 2 demonstrated that porpoises returned to the area between one and three days after piling (Brandt et al. 2011) and monitoring at the Dan Tysk Wind Farm as part of the Disturbance Effects on the Harbour Porpoise Population in the North Sea (DEPONS) project found return times of around 12 hours (van Beest et al. 2015).

Two studies at Alpha Ventus demonstrated, using aerial surveys, that the return of porpoises was about 18 hours after piling (Dähne et al. 2013). A recent study of porpoise response at the Gemini wind farm in the Netherlands, also part of the DEPONS project, found that local population densities recovered between two and six hours after piling (Nabe-Nielsen et al. 2018). An analysis of data collected at the first seven offshore wind farms in Germany has shown that harbour porpoise detections were reduced between one and two days after piling (Brandt et al. 2018).

Analysis of data from monitoring of marine mammal activity during piling of jacket pile foundations at Beatrice Offshore Wind Farm (Graham et al. 2017a, Graham et al. 2019) provides evidence that harbour porpoise are displaced during pile driving but return after cessation of piling, with a reduced extent of disturbance over the duration of the construction period. This suggests that the assumptions adopted in the current assessment are precautionary as animals are predicted to remain disturbed at the same level for the entire duration of the pile driving phase of construction.

TTS limitations

It is recognised that TTS is a temporary impairment of an animal's hearing ability with potential consequences for the animal's ability to escape predation, forage and/or communicate, supporting the statement of Kastelein et al. (2012c) that *"the magnitude of the consequence is likely to be related to the duration and magnitude of the TTS".* An assessment of the impact based on the TTS thresholds as currently given in Southall et al. (2019) or the former NMFS (2016) guidelines and

Data Limitations

Southall et al. (2007) guidance) would lead to a substantial overestimate of the potential impact of TTS. Furthermore, the prediction of TTS impact ranges, based on the sound exposure level (SEL) thresholds, are subject to the same inherent uncertainties as those for PTS, and in fact the uncertainties may be considered to have a proportionately larger effect on the prediction of TTS. These concepts are explained in detail below based on the thresholds detailed by Southall et al. (2019), as these represent the most up-to-date scientific knowledge.

It is SMRU Consulting's expert opinion that basing any impact assessment on the impact ranges for TTS using current TTS thresholds would overestimate the potential for an ecologically significant effect. This is because the species-specific TTS thresholds in Southall et al. (2019) describe those thresholds at which the onset of TTS is observed, which is, per their definition, a 6dB shift in the hearing threshold, usually measured four minutes after sound exposure, which is considered as *"the minimum threshold shift clearly larger than any day-to-day or session-to-session variation in a subject's normal hearing ability", and which "is typically the minimum amount of threshold shift that can be differentiated in most experimental conditions".*

The time hearing recovers back to normal (the recovery time) for such small threshold shifts is expected to be less than an hour, and, therefore, unlikely to cause any major consequences for an animal.

A large shift in the hearing threshold near to values that may cause PTS could however require multiple days to recover (Finneran 2015). For TTS induced by steady-state tones or narrowband noise, Finneran (2015) describes a logarithmic relationship between recovery rate and recovery time, expressed in dB/decade (with a decade corresponding to a ratio of 10 between two-time intervals, resulting in steps of 10, 100, 1000 minutes and so forth). For an initial shift of 5 to 15dB above hearing threshold, TTS reduced by 4 to 6dB per decade for dolphins, and 4 to 13dB per decade for harbour porpoise and harbour seals. Larger initial TTS tend to result in faster recovery rates, although the total time it takes to recover is usually longer for larger initial shifts (summarised in Finneran 2015). While the rather simple logarithmic function fits well for exposure to steady-state tones, the relationship between recovery rate and recovery time might be more complex for more complex broadband sound, such as that produced by pile driving noise.

For small threshold shifts of 4 to 5dB caused by pulsed noise, Kastelein et al. (2016) demonstrated that porpoises recovered within one hour from TTS. While the onset of TTS has been experimentally validated, the determination of a threshold shift that would cause a longer-term recovery time and is therefore potentially ecologically significant, is complex and associated with much uncertainty.

The degree of TTS and the duration of recovery time that may be considered severe enough to lead to any kind of energetic or fitness consequences for an individual, is currently undetermined, as is how many individuals of a population can suffer this level of TTS before it may lead to population consequences. There is currently no set threshold for the onset of a biologically meaningful TTS, and this threshold is likely to be well above the TTS-onset threshold, leading to smaller impact ranges (and consequently much smaller impact areas, considering a squared relationship between area and range) than those obtained for the TTS-onset threshold.

One has to bear in mind that the TTS-onset thresholds as recommended first by Southall et al. (2007) and further revised by Southall et al. (2019) were determined as a means to be able to determine the PTS-onset thresholds and represents the smallest measurable degree of TTS above normal day to day variation.

A direct determination of PTS-onset thresholds would lead to an injury of the experimental animal and is therefore considered as unethical. Guidelines such as National Academies of Sciences Engineering and Medicine (2017) and Southall et al. (2007) therefore rely on available data from humans and other terrestrial mammals that indicate that a shift in the hearing threshold of 40dB may lead to the onset of PTS.

Data Limitations

For pile driving for offshore wind farm foundations, the TTS- and PTS-onset thresholds for impulsive sound are the appropriate thresholds to consider. These consist of a dual metric, a threshold for the peak sound pressure associated with each individual hammer strike, and one for the cumulative sound exposure level (SELcum), for which the sound energy over successive strokes is summated. The SEL_{cum} is based on the assumption that each unit of sound energy an animal is exposed to leads to a certain amount of threshold shift once the cumulated energy raises above the TTS-onset threshold. For impulsive sound, the threshold shift that is predicted to occur is 2.3dB per dB noise received; for non-impulsive sound this rate is smaller (1.6dB per dB noise) (Southall et al. 2007). Please see the section above for further details on the limitations of SEL_{cum} thresholds (the same limitations apply to TTS as PTS).

Modelling of the SEL_{cum} impact ranges of PTS with a 'fleeing animal' model (as is typical during in noise impact assessments) is subject to both of these precautions. Modelling the SEL_{cum} TTS impact ranges will inherit the same uncertainties, however, over a longer period of time, and over greater ranges as the TTS impact ranges are expected to be larger than those of PTS. Therefore, these uncertainties and conservativisms will have a relatively larger effect on predictions of TTS ranges.

It is also important to bear in mind that the quantification of any impact ranges in the environmental assessment process, is done to inform an assessment of the potential magnitude and significance of an impact. Because the TTS thresholds are not universally used to indicate a level of biologically meaningful impact of concern per se but are used to enable the prediction of where PTS might occur, it would be very challenging to use them as the basis of any assessment of impact significance.

All the data that exists on auditory injury in marine mammals are from studies of TTS and not PTS. SMRU Consulting agrees with the studies' conclusion expressing confidence in predicting the range at which any TTS may occur.

However, this is not necessarily very useful for the impact assessment process. SMRU Consulting accept that scientific understanding of the degree of exposure required to elicit TTS may be more empirically based than our ability to predict the degree of sound required to elicit PTS, it does not automatically follow that our ability to determine the consequences of a stated level of TTS for individuals is any more certain than our ability to determine the consequences of a stated level of PTS for individuals. It could even be argued that predicting the consequences of a permanent effect is more reliable than predicting the consequences of a temporary effect of variable severity and uncertain duration.

It is important to consider that predictions of PTS and TTS are linked to potential changes in hearing sensitivity at particular hearing frequencies, which for piling noise are generally thought to occur in the 2-10kHz range and are not considered to occur across the whole frequency spectrum. Studies have shown that exposure to impulsive pile driving noise induces TTS in a relatively narrow frequency band in harbour porpoise and harbour seals (reviewed in Finneran 2015), with statistically significant TTS occurring at 4 and 8kHz (Kastelein et al. 2016) and centred at 4kHz (Kastelein et al. 2012a, Kastelein et al. 2012b, Kastelein et al. 2013b, Kastelein et al. 2017). Our understanding of the consequences of PTS within this frequency range to an individual's survival and fecundity is limited, and therefore our ability to predict and assess the consequences of TTS of variable severity and duration is even more difficult to do.

TTS Conclusion

Predicted TTS impact ranges and the number of animals within those ranges are presented in this impact assessment. However, no assessment of the sensitivity of marine mammals to TTS is provided, nor is the magnitude of TTS assessed.

Data Limitations

Population Modelling

There is a lack of empirical data on the way in which changes in behaviour and hearing sensitivity may affect the ability of individual marine mammals to survive and reproduce.

Therefore, in the absence of empirical data, the iPCoD framework uses the results of an expert elicitation process conducted according to the protocol described in Donovan et al. (2016) to predict the effects of disturbance and PTS on survival and reproductive rate. The process generates a set of statistical distributions for these effects and then simulations are conducted using values randomly selected from these distributions that represent the opinions of a "virtual" expert. This process is repeated many 100s of times to capture the uncertainty among experts.

There are several precautions built into the iPCoD model and this specific scenario that mean that the results are considered to be highly precautionary and likely over-estimate the true population level effects. These include:

- The fact that the model assumes bottlenose dolphins will not forage for 24 hours after being disturbed;
- The lack of density dependence in the model (meaning the population will not respond to any reduction in population size); and
- The level of environmental and demographic stochasticity in the model.

Duration of disturbance: bottlenose dolphins

The iPCoD model for bottlenose dolphin disturbance was last updated following the expert elicitation in 2013 (Harwood et al. 2014). When this expert elicitation was conducted, the experts provided responses on the assumption that a disturbed individual would not forage for 24 hours. However, the most recent expert elicitation in 2018 highlighted that this was an unrealistic assumption for harbour porpoises (generally considered to be more responsive than minke whales and bottlenose dolphins), and was amended to assume that disturbance resulted in 6 hours of non-foraging time (Booth et al. 2019). Unfortunately, bottlenose dolphins were not included in the updated expert elicitation for disturbance, and thus the iPCoD model still assumes 24 hours of non-foraging time for bottlenose dolphins. This is unrealistic considering the current understanding of marine mammal behavioural responses to pile driving. A recent study estimated energetic costs associated with disturbance from sonar, where it was assumed that 1 hour of feeding cessation was classified as a mild response, 2 hours of feeding cessation was classified as a strong response and 8 hours of feeding cessation was classified as an extreme response (Czapanskiy et al. 2021).

Assuming 24 hours of feeding cessation for bottlenose dolphins in the iPCoD model is significantly beyond that which is considered to be an extreme response and is therefore considered to be unrealistic and will over-estimate the true disturbance levels expected from the proposed development.

Lack of density dependence

Density dependence is described as "the process whereby demographic rates change in response to changes in population density, resulting in an increase in the population growth rate when density decreases and a decrease in that growth rate when density increases" (Harwood et al. 2014). The iPCoD scenario run assumes no density dependence since there is insufficient data to parameterise this relationship. Essentially, what this means is that there is no ability for the modelled impacted population to increase in size back up to carrying capacity following disturbance. At a recent expert elicitation, conducted for the purpose of modelling population impacts of the Deepwater Horizon oil spill (Schwacke et al. 2021), experts agreed that there would likely be a concave density dependence on fertility, which means that in reality, it would be expected that the impacted population would recover to carrying capacity (assumed to be equal to the size of un-impacted population which is assumed to be at carrying capacity) rather than continuing at a stable trajectory that is smaller than that of the un-impacted population.

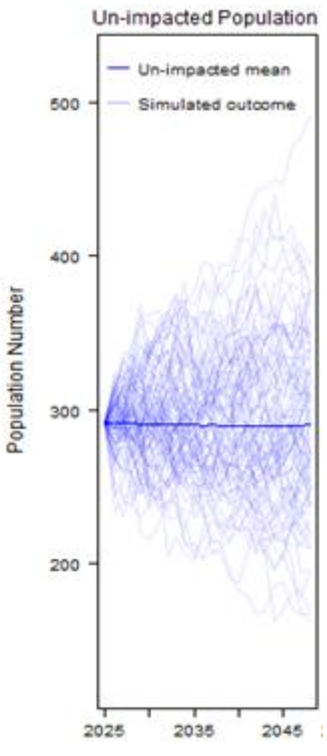
Environmental and demographic stochasticity

The iPCoD model attempts to model some of the sources of uncertainty inherent in the calculation of the potential effects of disturbance on marine mammal population. This includes demographic stochasticity and environmental variation. Environmental variation is defined as "the variation in demographic rates among years as a result of changes in environmental conditions" (Harwood et al. 2014). Demographic stochasticity is defined as "variation among individuals in their realised vital rates as a result of random processes" (Harwood et al. 2014).

The iPCoD protocol describes this in further detail:

"Demographic stochasticity is caused by the fact that, even if survival and fertility rates are constant, the number of animals in a population that die and give birth will vary from year to year because of chance events. Demographic stochasticity has its greatest effect on the dynamics of relatively small populations, and we have incorporated it in models for all situations where the estimated population within an MU is less than 3000 individuals. One consequence of demographic stochasticity is that two otherwise identical populations that experience exactly the same sequence of environmental conditions will follow slightly different trajectories over time. As a result, it is possible for a "lucky" population that experiences disturbance effects to increase, whereas an identical undisturbed but "unlucky" population may decrease" (Harwood et al. 2014).

This is clearly evidenced in the outputs of iPCoD where the un-impacted (baseline) population size varies greatly between iterations, not as a result of disturbance but simply as a result on environmental and demographic stochasticity. In the example provided in Graph 2, after 25 years of simulation, the un impacted population size varies between 176 (lower 2.5%) and 418 (upper 97.5%). Thus, the change in population size resulting from the impact of disturbance is significantly smaller than that driven by the environmental and demographic stochasticity in the model.



Data Limitations

Graph 2 Simulated un-impacted (baseline) population size over the 25 years modelled

Summary

All of the precautions built into the iPCoD model mean that the results are considered to be highly precautionary. Despite these limitations and uncertainties, this assessment has been carried out according to best practice and using the best available scientific information. The information provided is therefore considered to be sufficient to carry out an adequate assessment, though a level of precaution around the results should be taken into account when drawing conclusions.

In addition to this, it is noted that iPCoD is not available for common dolphins.