JNCC Report
No. 677

JNCC Report on the Correct treatment of uncertainty in ornithological assessments

Searle, K.R., Jones, E.L., Trinder, M., McGregor, R., Donovan, C., Cook, A., Daunt, F., Humphries, L., Masden, E., McCluskie, A. & Butler, A.

May 2021 (Revised June 2021)

© JNCC, Peterborough 2021

ISSN 0963 8091

For further information please contact:
Joint Nature Conservation Committee
Monkstone House
City Road
Peterborough PE1 1JY
https://jncc.gov.uk/

This report should be cited as:
Searle, K.R., Jones, E.L., Trinder, M., McGregor, R., Donovan, C., Cook, A., Daunt, F., Humphries, L., Masden, E., McCluskie, A. & Butler, A. 2021. JNCC Report on the Correct treatment of uncertainty in ornithological assessments. *JNCC Report No. 677*, JNCC, Peterborough, ISSN 0963-8091.

Acknowledgments:
We especially with to thank Nancy McLean (EDF-EN) for her input into the section on communicating uncertainty to decision makers, along with all members of the project steering group for their comments on drafts of this report (Sue O'Brien, Orea Anderson, Tom Evans (MSS), Richard Berridge and Rebecca Jones (NE)). We also wish to especially thank all those who attended the workshop for their time and insight: Julie Black (JNCC), Janelle Braithwaite, Tom Evans, Julie Miller (Marine Scotland), Matty Murphy, Alex Scorey, Mike Bailey (NRW), Glen Tyler (NatureScot) and Clare McNamara (Daera).

JNCC EQA Statement:
This report is compliant with JNCC's Evidence Quality Assurance Policy
https://jncc.gov.uk/about-jncc/corporate-information/evidence-quality-assurance/

Revision (June 2021):
This report was revised to correct a number of minor typographical errors.

Summary

Assessing impacts of offshore renewable developments (ORDs) on marine birds is challenging, as it involves a number of key interlinked elements, from observations of birds' behaviour in response to ORDs to predicted population trajectories. The synthesis of multiple sources of data and a range of modelling approaches must be used to understand a set of complex behavioural, energetic and demographic processes, operating in a marine environment that is both spatially and temporally highly dynamic and logistically challenging to work in. Consequently, there is considerable uncertainty associated with such assessments.

Currently, the assessment process does not quantify the overall uncertainty associated with the impacts of ORDs in a scientifically robust, evidence-based manner. The degree and defensibility of uncertainty quantification varies between different stages of the assessment process, and there is a lack of consensus on how the uncertainties associated with different stages of the process should be combined together in order to provide an "end-to-end" quantification of scientific uncertainty.

A key requirement for quantification of uncertainty within the assessment process is that uncertainty and environmental or natural variation (hereafter 'environmental variation' or 'variation') must both be accounted for, and that the assessment process must correctly distinguish between variation and uncertainty. The key general distinction between environmental variation and uncertainty is that environmental variation is an inherent feature of the system (e.g. arising from seabird biology), and so cannot be reduced through additional data collection, whereas uncertainty is a feature of the state of knowledge, and so can, at least in principle, be reduced through additional data collection and improved understanding, thereby enhancing validity of models.

It is important to stress that full quantification of uncertainty is as important as the reduction of uncertainty in supporting the decision-making process. This is because apparent reductions in uncertainty that arise in the context of an inadequate quantification of uncertainty are liable to create a false sense of certainty, and so increase the risk of unanticipated outcomes. The reduction of uncertainty can only meaningfully be prioritised and evaluated within the context of a comprehensive quantification of uncertainty, hence why much of the focus of the recommendations developed in this project is on the quantification of uncertainty, as well as the reduction of this uncertainty (Table 4).

In this project we have, through a workshop-based process of consultation with relevant experts and stakeholders produced a framework for how scientific uncertainty can be quantified and reduced throughout the assessment process, to facilitate the development of more precise ORD impact estimates. The project focuses upon offshore wind farm developments, rather than on other ORD technologies such as tidal or wave developments. We highlight key areas in which new empirical data and research are required in order to reduce uncertainties, and outline, in broad terms, the resources and activities required for their delivery. This work has involved bringing together a wide range of relevant experts and stakeholders to:

1. review the methods that are currently used to quantify uncertainty within the assessment process, and evaluate the ways in which these uncertainty estimates are currently used within the assessment process
2. highlight key areas in which the quantification and interpretation of uncertainty could be improved, either through statistical modelling, additional data collection or adaptation of the assessment process

3. provide a framework for the end-to-end quantification of uncertainty, which brings together estimates of uncertainty associated with individual stages of the assessment process
4. develop recommendations for the research required to both better quantify uncertainty, and to reduce it, to better underpin and inform potential reductions in consenting risk for future offshore wind development through more certainty about likely impacts of planned developments

Glossary

Apportioning: The process, for seabird species, of estimating the percentage of individuals within a specified area of sea (e.g. footprint) over a particular period of time (e.g. season or month) that can be attributed to each breeding colony.

Calibration: the process of comparing the outputs of a model against observed data and revising the values of input parameters in order to reduce the discrepancy between observations and model outputs.

Covariate: a measured quantity that can be used to describe variation, through a statistical model, in the quantity of interest.

Estimation: the process of inferring the values of the parameters of a model using observed data on the model outputs. Similar in many ways to "calibration", but the term "estimation" is typically used in the context of statistical models and the term "calibration" in the context of mechanistic models.

Parameter: a quantity within a model whose value is unknown, but for which the value may be informed by data or expert judgement.

Parameter uncertainty: the uncertainty in model outputs that arises from lack of knowledge regarding the values of parameters within the model (the term "estimation uncertainty" is also often used to capture this, and in Masden *et al.* (2015) the term "sampling uncertainty" is used).

Population Viability Analysis (PVA): A framework for translating effects on annual demography into impacts on longer-term abundance.

Sensitivity Analysis: a process for evaluating the extent to which the outputs of a model are sensitive to uncertainty/variation in each of the input parameters.

Stochastic model: a model that allows for variation in one or more of the inputs.

Uncertainty: limitations to our knowledge and understanding that arise from a lack of data, or limitations to the interpretation of the available data. Uncertainty is a feature of human knowledge, and so can, in principle, be reduced through data collection and/or further analyses and interpretation of existing data.

Structural uncertainty: uncertainty that arises from fundamental differences between the assumptions of the model and reality.

Variation: natural differences that occur - e.g. over time or space, or between individuals. Variation, unlike uncertainty, cannot be reduced through data collection or analysis, as it is an inherent feature of the system being studied.

Table of Contents

Figures

Tables

1 Background

UK Government has set targets to generate 50% of overall energy consumption from renewable sources by 2030 and to have decarbonised the energy system almost completely by 2050. However, the Government has a duty to ensure that Offshore Renewable Developments (ORDs) are delivered in a sustainable manner, in accordance with the requirements for Habitats Regulation Assessment and Environmental Impact Assessment. Offshore renewable developments have the potential to affect seabirds that are protected under the Habitats Regulations, notably from collisions with turbine blades and through displacement from important habitat (Drewitt & Langston 2006; Busch *et al*. 2013; Thaxter *et al*. 2015; Dierschke *et al.* 2016; Welcker & Nehls 2016).

Assessing impacts of ORDs on marine birds is challenging, as it involves a number of key interlinked elements, from observations of birds' behaviour in response to ORDs to predicted population trajectories. The process of assessing the potential impact of a planned development on a protected marine bird population involves predicting the potential demographic consequences of any mortality caused by the development. These predicted demographic consequences are then used to evaluate whether there will be an Adverse Effect on the Integrity of SPA interest features and the SPA network, which is an assessment of whether mortality will be sufficient to cause a decline in the ecological coherence of an SPA interest feature, such that the feature no longer meets the Conservation Objectives for that feature at that site. These predictions are underpinned by a series of interconnected processes, designed to estimate the behavioural and subsequent demographic consequences and population level response to the ORD. This can be depicted using the following schematic:

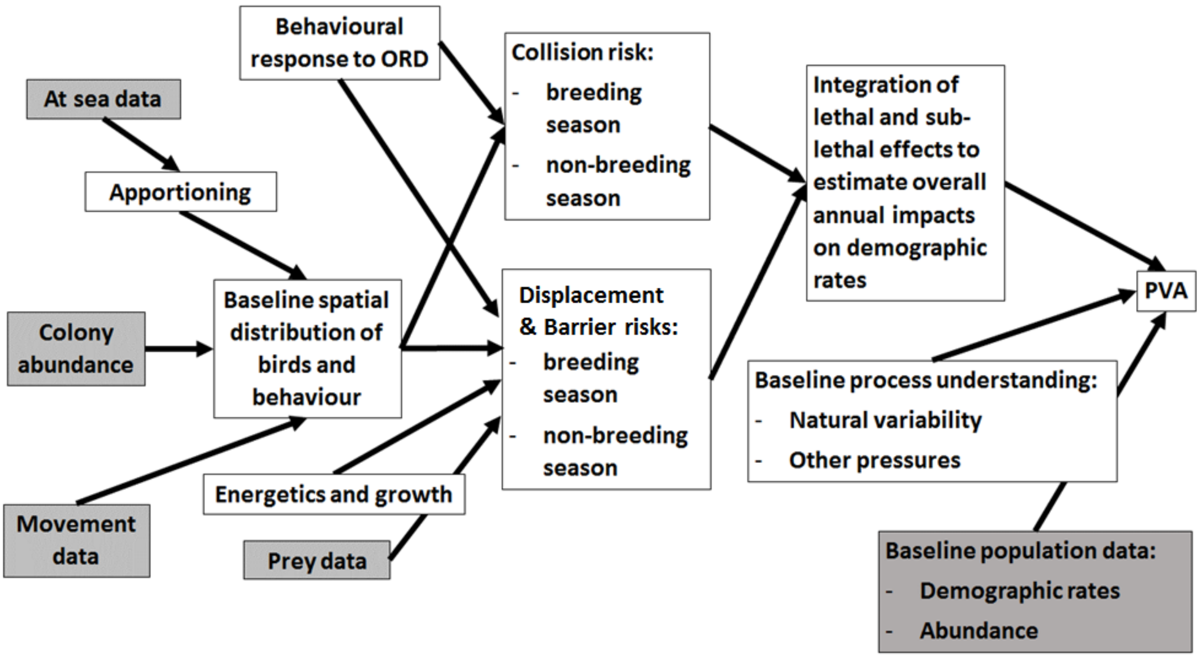


Figure 1: A schematic diagram illustrating the models and steps involved with the ornithology and offshore wind impact assessment process.

The synthesis of multiple sources of data and a range of modelling approaches must be used to understand a set of complex behavioural, energetic and demographic processes, operating in a marine environment that is both spatially and temporally highly dynamic and logistically challenging to work in. Consequently, there is considerable uncertainty associated with such assessments.

Currently, the assessment process does not quantify the overall uncertainty associated with the impacts of ORDs in a scientifically robust, evidence-based manner. The degree and defensibility of uncertainty quantification varies between different stages of the assessment process, and there is a lack of consensus on how the uncertainties associated with different stages of the process should be combined together in order to provide an “end-to-end” quantification of scientific uncertainty.

The consequence of this is that uncertainty around the magnitude of impacts of OW development on marine bird populations is not reliably known. Furthermore, this uncertainty is generally viewed as a feature of the process that can only be managed through additional empirical data collection. However, better statistical treatment of uncertainty and a holistic approach to managing uncertainty from beginning to end of the assessment process are likely to yield greater confidence in predicted impacts and quantitative estimates of uncertainty associated with them, reducing the need for precautionary approaches.

The precautionary principle exists for situations where scientific data does not exist or is incomplete and therefore it is not possible to complete a full evaluation of the possible risks a plan, project or activity may cause to the environment, including possible danger to humans, animal or plant health, or to the environment in general (RSPB 2019). The European Commission’s Precautionary Principle guidance[1] states that it should apply when a phenomenon, product or process may have a dangerous effect, identified by a scientific and objective evaluation, if this evaluation does not allow the risk to be determined with sufficient certainty. As such the degree of precaution applied to an evaluation, or assessment, can be seen to be directly proportional to the extent of scientific uncertainty inherent in that assessment. As the guidance goes on to recommend, “The implementation of an approach based on the precautionary principle should start with a scientific evaluation, as complete as possible, and where possible, identifying at each stage the degree of scientific uncertainty.” (RSPB 2019).

In this project we have, through a workshop-based process of consultation with relevant experts and stakeholders produced a clear framework for how scientific uncertainty can be quantified and reduced throughout the assessment process, to facilitate the development of more precise ORD impact estimates. The project focuses upon offshore wind farm developments, rather than on other ORD technologies such as tidal or wave developments. We highlight key areas in which new empirical data and research are required in order to reduce uncertainties, and outline, in broad terms, the resources and activities required for their delivery. This work has involved bringing together a wide range of relevant experts and stakeholders to:

1. review the methods that are currently used to quantify uncertainty within the assessment process, and evaluate the ways in which these uncertainty estimates are currently used within the assessment process
2. highlight key areas in which the quantification and interpretation of uncertainty could be improved, either through statistical modelling, additional data collection or adaptation of the assessment process
3. provide a framework for the end-to-end quantification of uncertainty, which brings together estimates of uncertainty associated with individual stages of the assessment process

---

[1] https://ec.europa.eu/environment/integration/research/newsalert/pdf/precautionary_principle_decision_making_under_uncertainty_FB18_en.pdf; https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/714379/180511_EUWB_Environmental_Protections_factsheet_10_May_18.pdf

4. develop recommendations for the research required to both better quantify uncertainty, and to reduce it, to better underpin and inform potential reductions in consenting risk for future offshore wind development through more certainty about likely impacts of planned developments

2 Review of sources of uncertainty in ornithological offshore wind assessments

2.1 Overview

A first step in understanding uncertainty in any process is to identify the various types and sources of uncertainty and variability that are involved. In the assessment of environmental impacts, the level and form of uncertainty not only depends upon the availability of relevant empirical data, but also upon data collection and sampling methodologies, analysis and modelling methods, linguistics used by different stakeholders, and policy frameworks.

When understanding and quantifying ecological processes, it is particularly important to recognise the different contributions and impacts of uncertainty versus natural variability (Figure 2). Natural variability is a property of natural systems, which may have many causes such as variation between individuals within a colony related to physiology, age or sex (often termed ‘individual variation’); or variation between colonies due to differing habitat characteristics, and variation across years due to variation in weather or other aspects of the ecosystem (often termed ‘environmental variation’; Figure 2). Importantly, because natural variability is a property of the ecological system, it cannot be reduced. It can, however, be characterised and quantified through measurement. This quantification of natural variability may then be used within models or analyses of ecological processes. If we could perfectly measure the natural variability in all the processes affecting ornithological interactions with offshore windfarms, we could include this variability within assessments, clearly separating its impacts from those arising from uncertainty.

Uncertainty itself is a function of how well we understand, measure and represent an ecological process. It is introduced due to the limitation of our knowledge and understanding of a system, whereby we often only have an imperfect ability to describe the ecological process of interest. This has been termed ‘knowledge uncertainty’ (Masden *et al*. 2015; Figure 2). Within ornithological offshore wind assessments, this knowledge uncertainty captures uncertainty that arises from our ability to understand and represent all of the ecological processes through which seabirds interact with offshore wind developments (hereafter referred to as ‘ORDs’). For instance, the assessment process typically considers three main types of seabird interactions: displacement from habitat, barrier effects, and collision impacts. However, these broad categories capture a myriad of underlying behavioural mechanisms, some of which may be explicitly represented within the assessment process, but many of which are currently not included, such as habituation, impacts on other trophic levels affecting predator-prey interactions, and foraging site fidelity. For almost all of these interactions between seabirds and ORDs, we are only able to partially describe and measure the underlying behavioural mechanisms, resulting in knowledge uncertainty that affects ornithological assessment outcomes. Further, all behaviours have energetic and fitness consequences for individuals. As with behavioural interactions, we are only able to partially describe and measure the energetic consequences of ORD impacts on the behaviour of seabirds, or the translation of these energetic consequences into fitness consequences on demographic rates of individuals and populations.

Our ability to understand, quantify and reduce this knowledge uncertainty is linked to how we describe these interactions between seabirds and ORDs. Knowledge uncertainty is driven by two key elements: 1) our descriptions or models of the relevant ecological processes, termed ‘structural uncertainty’ or sometimes ‘process uncertainty’, and 2) our ability to obtain data that adequately captures the states and processes underpinning interactions, termed ‘sampling uncertainty’, ‘estimation uncertainty’ or (in the context of specific models for interactions) ‘parameter uncertainty’ (Figure 2). Therefore, increasing or improving data collection, and using these data to improve understanding and analytical or model descriptions of behaviour and processes will all lead to reduced uncertainty. It is this uncertainty, ‘structural’ and ‘sampling’ on which we focus in this report. Our particular focus is on assessing how environmental variation and structural and sampling uncertainty are recognised, quantified, and used, and on how they are propagated through the assessment process. We also focus on how we can reduce knowledge uncertainty by increasing our knowledge of key ecological processes through data collection and modelling or statistical analysis, with the ultimate goal being to better quantify and reduce uncertainty to facilitate more robust decision making. However, we note that the complexities of natural systems, coupled with their inherent natural variability, mean that it will likely never be possible to perfectly quantify uncertainty, or minimise it to such an extent as to entirely remove risk in consenting decisions.

Finally, within ornithological impact assessments, uncertainty also arises through linguistic and decision-making processes. Linguistic uncertainty arises because language is vague and/or the precise meaning of words changes over time or between disciplines (Masden *et al.* 2015). For instance, the use of the word ‘precautionary’ within assessments is designed to have a precise meaning and interpretation, and yet it means many different things to different stakeholders. Decision-making uncertainty relates to how knowledge and predictions are interpreted, communicated and used in the management and policy arena (Masden *et al.* 2015). Whilst important, these two additional sources of uncertainty fall outside the direct scope of this project, and hence we do not propose recommendations for solutions to reducing uncertainty arising from these sources. However, decision-making uncertainty is discussed in brief below, because this provides important context to guide the prioritisation of work to quantify and reduce knowledge uncertainty.

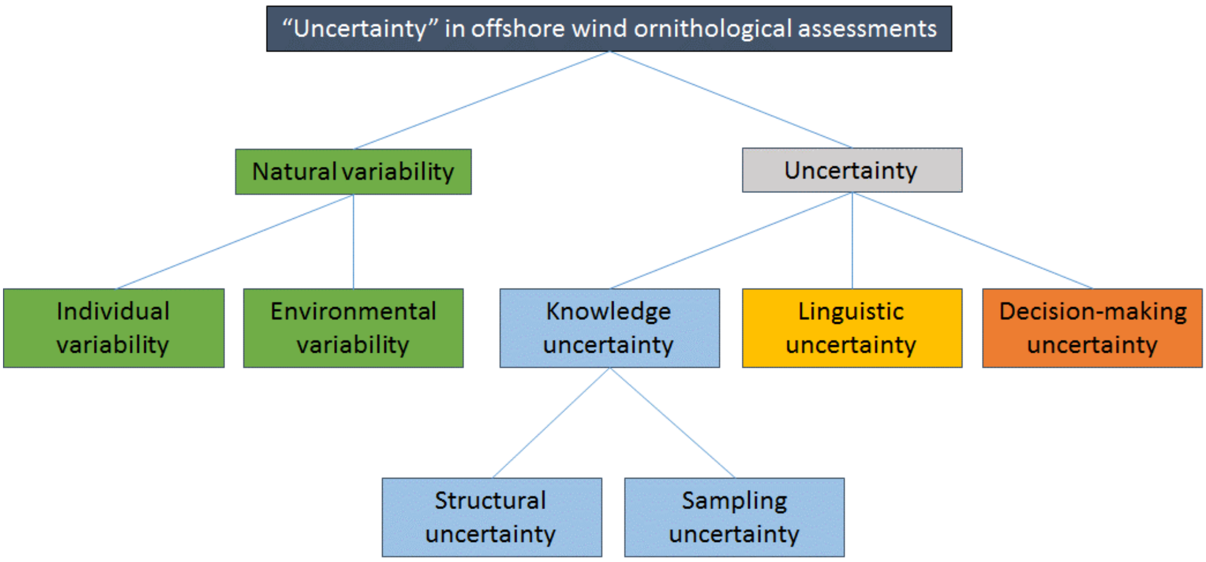


Figure 2: Summary of the sources of uncertainty affecting ornithological offshore windfarm assessments. Adapted from Masden *et al.* 2015.

2.2 Uncertainty and natural variability

A key requirement for quantification of uncertainty within the assessment process is that uncertainty and environmental or natural variation (hereafter ‘environmental variation’ or ‘variation’) must both be accounted for, and that the assessment process must correctly distinguish between variation and uncertainty. The key general distinction between environmental variation and uncertainty is that environmental variation is an inherent feature of the system (*e.g.* arising from seabird biology), and so cannot be reduced through additional data collection, whereas uncertainty is a feature of the state of knowledge, and so can, at least in principle, be reduced through additional data collection and improved understanding, thereby enhancing validity of models (Figure 2).

In practice, however, the typical mathematical and statistical ways of representing environmental variability and uncertainty are essentially the same (via probability distributions), and in complex systems, such as interactions between seabirds and ORDs, the two processes are often difficult to disentangle.

To illustrate the similarities and differences between uncertainty and environmental variability, and the challenges in disentangling them, we consider a simple hypothetical example, in which interest lies in estimating the adult mass at the end of the chick-rearing period (a key proxy for over-winter adult survival), for one particular colony of interest. Such an example could be of high relevance to assessing impacts of OWFs, because one of key ways in which displacement and barrier effects are suspected to reduce over-winter adult survival is through a reduction in adult mass at the end of the chick-rearing period: this is the mechanism by which barrier and displacement effects alter over-winter adult survival within the individual based model, SeabORD (Searle *et al*. 2014). This loss of mass may result from displacement forcing individuals to forage in less suitable areas of habitat, or from barrier effects forcing individuals to expend more energy by flying further to reach their foraging areas. We assume for this hypothetical example that data on mass can be collected for a random sample of n birds, in each of m years, and further assume that mass follows, at least approximately, a normal distribution. We also assume that no other data (*e.g*. on relevant covariates are available). In this situation, the obvious way to model these data would be through a linear mixed model (LMM) of mass, which contains a random effect for “year”. The LMM would contain three parameters, each of which would be estimated from the data:

a) the overall mean mass (across all years and all individuals);
b) the inter-annual variance, derived from the estimated random effect; and
c) the residual variance, which corresponds to variability between individual birds within a particular year.

The statistical model (the LMM) contains parameters that explicitly represent and quantify variability (the inter-annual variance, the residual variance), but no parameters that explicitly represent uncertainty. In this example, the magnitudes of the two forms of variability (inter-annual variability and inter-individual variability within each year) can be explicitly estimated from the available data through the model.

So, where is the uncertainty within this model? Software for fitting LMMs will always report the uncertainty (standard error) associated with the overall mean mass. Some mixed model software will also report the uncertainties associated with the inter-annual and residual variances; other software (*e.g*. the widely used lme4 package in R) does not. How the uncertainty is quantified will depend, however, upon the focus of the analysis. If interest in the model ultimately lies in estimating the overall mean mass (across all years and all individuals) then the model provides a direct estimate of this – the standard error for the associated parameter within the LMM. Alternatively, if interest is focused on predicting the mean mass (across individuals) for a randomly selected future year, then the calculation of uncertainty in this would involve both the standard error for the overall mean mass and the estimate of inter-year variability. Finally, if interest lay in predicting the mass of a particular individual in a particular future year, then the calculation of uncertainty would involve the standard error for the overall mean mass, the estimate of inter-year variability, and the estimate of residual variability (variability between individuals in a particular year). This leads us to a second key point: uncertainties in model outputs may or may not need to account for variability, and may depend on some, or all, sources of variability – this depends on which output(s) are of interest.

What happens if $m=1$, so that data are only available for a single year? In this case, the inter-annual variance cannot be estimated from the data, and the overall mean mass cannot be distinguished from the year-specific mean mass. In this situation, only some sources of variability can be quantified, and uncertainty can only be partially quantified. This leads us to a third important point: the available data do not always allow uncertainty and variability to be neatly separated, nor completely quantified.

The separation of uncertainty and variability is, consequently, challenging. But is it necessary? In general terms, we firmly believe that it is – because incorrectly treating uncertainty as variability, or vice versa, can lead to substantial under-estimation or over-estimation of uncertainty. In this simple example, if we are interested in predicting the mass for a particular future year, but we base the uncertainty in this solely upon the standard error for the overall mean mass (across all years and individuals), then we are liable to strongly under-estimate uncertainty, by ignoring the effect of inter-annual variability in generating uncertainty in predictions of mass for any particular year.

In the actual ORD assessment process, the separation of uncertainty and variability is far more complex than in this simple example. We return to the question of how this separation is currently achieved, and how this could be improved, within the PVA section (Section 3).

2.3 End-to-end propagation throughout ORD assessment process

The standard assessment process for estimating ORD impacts upon seabirds essentially involves running a linked set of tools, and then using the outputs of these tools to inform the decision-making process. In most cases (i.e. for species where either collision or displacement are relevant, but not both) the assessment process essentially involves running a series of tools, such as a collision risk model, then an apportioning tool, and then a population viability analysis (PVA). For some species (*e.g*. black-legged kittiwake) both collision and displacement can be relevant, and in this case the outputs of collision risk and displacement risk tools need to be combined (usually added) together before being inputted into a PVA. In either case, it is crucial to ensure that uncertainty is propagated through the series of tools.

The process of linking the tools used within assessments involves subjective judgement (*e.g*. regarding which sources of input data to use), and there is also uncertainty associated with the consenting decision-making process itself. In addition, the framework that is used to link the different tools (*e.g*. CRM models, displacement tools, apportioning tools, PVA) is likely to contain structural uncertainties. Such structural uncertainty might arise if there are impacts of offshore renewables upon seabirds other than those currently considered within the assessment process, but they could also arise if components of the process that are currently encapsulated by distinct tools interact. For example, for species that are potentially impacted by both collision and displacement assessments typically assume that these processes are independent, and that their effects can simply be added together, thereby ignoring the potential for biological interactions between the movement and behavioural processes that underpin displacement and collision effects. Some of these structural uncertainties are potentially resolvable through improved modelling. For example, recent extensions of individual-based models (SeabORD; Searle *et al.* 2020) allow interactions between collision risk and displacement risk to be accounted for within assessments, albeit in a relatively simplistic way – whilst others are much harder to resolve.

Aside from these broader uncertainties (structural uncertainty, decision-making uncertainty, uncertainty arising from subjective judgements in the selection of input data), there are also uncertainties that arise from the process of linking the different tools used within the assessment process, each of which itself contains uncertainty. This "end-to-end" quantification of uncertainty is important, because it allows the uncertainties associated with the final, decision-relevant, outputs (*e.g*. PVA metrics) to reflect all of the quantifiable uncertainties that arise from the linked tools that have been used to derive these outputs. Each tool within the assessment process is a model of one element of interactions between seabirds and ORDs, and the entire assessment process can also be regarded as a larger model, or meta-model, that is formed by linking the individual tools together. End-to-end quantification of uncertainty therefore involves quantifying the uncertainty associated with this larger, "meta-model". The simplest approach for linking uncertainties between tools, both conceptually and practically, is via simulation. The Marine Scotland 'SEANSE' project (Searle *et al.* 2020) provided an initial attempt to do this within the context of a case study on the impacts of offshore renewables on seabirds, and the Marine Scotland CEF project is currently systematising and automating this approach.

The simulation-based approach essentially involves running multiple simulations, each of which involves running all of the linked tools. Each simulation randomly generates the values of any inputs that contain uncertainty, and/or any internal tool components that involve stochasticity. This simulation-based approach is therefore able to account for both uncertainty and variability within a common framework. The distribution of the assessment process outputs (*e.g*. PVA metrics) across simulations then quantifies the "end-to-end" uncertainty associated with the assessment process.

A simple hypothetical example, to illustrate how this approach works in practice is given in Appendix B. The tools involved in the assessment process are much more complicated than in this simple example, and the number of tools being linked is larger, but the principle for linking together uncertainties through a simulation approach remains the same.

There are, however, three main limitations of the simulation-based approach:

a) the resulting estimates of "end-to-end" uncertainty will only be meaningful if the uncertainty quantification within each of the individual tools, and inputs, is comprehensive and statistically defensible;
b) the process only yields defensible/stable estimates of uncertainty if the number of simulations is large;
c) the simulation-based approach, at least in its simplest form, assumes that the various tools operate independently of each other.

All of these represent potentially substantive issues in the context of ORD assessments, but the most natural solution to all three issues arises through improved quantification of uncertainty within the individual tools, or within the input data to the tools. In particular, the uncertainty quantification within individual tools and inputs is currently inconsistent, and in some cases very limited (see next section). Within this report we provide a series of suggestions/recommendations for how these issues can be addressed.

The issue that a large of number of simulations is required to give a defensible representation of uncertainty is a substantive issue for those tools that are computationally intensive to run. Current collision risk models (sCRM) and PVA models (NE/JNCC PVA tool; Butler *et al.* 2020a; Searle *et al.* 2019) require some degree of computational effort, and so impose some constraints on the number of simulations that can realistically be run. However, the use of individual-based models (IBMs) such as SeabORD are by far the most computationally intensive methods, and hence, in contexts where IBMs is used, it is this that imposes the main practical limits upon the number of simulations that are possible. This issue is currently being addressed within the MSS CEF project – where the IBM SeabORD is being redesigned to be less computationally intensive, which will make the running of larger numbers of simulations feasible in future.

Finally, relaxing the assumption that the individual component tools of the assessment process operate independently depends upon the availability of data and biological knowledge to enable such relaxation. The Marine Scotland Collision and Displacement Integration project (Searle *et al*., in press) has made a first attempt at relaxing the assumption that collision and displacement processes operate independently, and highlighted ways that this assumption could be further relaxed through greater integration of IBMs like SeabORD with sCRM models.

2.4 Current estimation and use of uncertainty in assessments

The current tools used in assessments vary in the extent to which uncertainty is considered. The more complex tools such as the sCRM, SeabORD and NE/JNCC PVA tool, all consider variability, and some elements of uncertainty. All three tools use a probabilistic, simulation-based, approach to quantify and represent variability and uncertainty. The similar approaches to representation of uncertainty have allowed the sCRM and SeabORD to be linked together, and to be linked to the NE/JNCC PVA tool. There remain, however, key elements of uncertainty that are not currently considered within these tools, which we discuss in later sections.

Of the more simple tools used within assessments, the Displacement Matrix provides a visual representation of the uncertainty associated with displacement risk, but does not attempt to quantify this in a probabilistic way, and so cannot readily be incorporated into an end-to-end assessment of uncertainty. Similarly, the SNH and MSS apportioning tools do not currently represent uncertainty or variability at all, although some progress has been made to include variability in foraging ranges between breeding colonies within the context of the SNH apportioning tool (Searle *et al*. 2020).

2.5 Summarising the need for improved uncertainty quantification

Delivering the underpinning science to enable accurate, robust and defensible ornithological ORD impact assessments requires developing and advancing a credible line of inference from our conceptual understanding of the ecological and behavioural processes involved through to quantitative estimates with uncertainty (Hobbs & Hooten 2015). This involves representing our knowledge and understanding of the interactions between seabirds and ORDs with models and observations of the key processes shaping these responses, such as seabird spatial habitat use, displacement and barrier effects, and collision impacts. All models, whether conceptual, theoretical or statistical, are simplified abstractions of reality, and we rely on the proper quantification of uncertainty to bridge the gap between reality and our representations of it to help us draw inference from our models and to understand their validity and usefulness for shaping decision-making and policy. Similarly, the data that we collect to inform a model will often only partially capture the true, underlying state of the process we are trying to observe. A failure to recognise or quantify these uncertainties in models and data necessarily pushes the application of understanding towards subjective decision-making where the rationale is unclear, rather than towards transparent, objective evidence-based decision-making. It is therefore imperative that we underpin sound decision-making with a full attempt at quantifying uncertainty within ornithological ORD impact assessments. It is only through identifying and quantifying all sources of uncertainty, that we may then see clearly how to develop our science to reduce this uncertainty, thereby obtaining a better understanding of the potential impacts of offshore wind development on the environment (European Union Law[2]).

3 Population Viability Analysis

3.1 Context

Within the assessment process, offshore wind farms are assumed to alter annual demographic rates (primarily survival and productivity) via collision, displacement and barrier effects. These effects on annual rates translate into longer-term impacts on abundance. Population Viability Analysis (PVA) provides an established statistical framework for translating effects on annual demography into impacts on longer-term abundance (Soulé 1986; Beissinger & McCullough 2002). Not all impact assessments require PVA (e.g. some assessments use percent habitat lost or percent of SPA area lost), however the vast majority of offshore wind assessments utilise PVA, which plays an important role in the assessment process for most wildlife species.

Currently, the NE/JNCC PVA tool (Butler *et al.* 2020a; Searle *et al*. 2019) provides the primary tool for implementing PVA methods for ornithological ORD assessments. The NE/JNCC PVA tool is based upon Leslie matrix models (Leslie 1945), which provide a flexible, unifying statistical framework for linking demography and abundance. In practice, PVA involves running the Leslie matrix models forward in time for both impacted and unimpacted populations, and comparing these against each other, using a range of different metrics (Jitlal *et al.* 2017). Whilst it is possible to calculate metrics that relate to the absolute state of the impacted population (*e.g*. quasi-extinction probability), the metrics that are advised for use in assessments provide relative comparisons of impacted and baseline simulations. A widely used example of the latter is the impact on annual growth rate (*AGR*), which is calculated to be:

$$\text{AGR} = (\text{Final population size in impacted population} / \text{Final population size in unimpacted population})^{1/n}$$

where n is the number of years of impact.

The key rationale, that relative rather than absolute metrics are used, has been suggested because ratio metrics have been found to be less sensitive to misspecification of baseline demographic rates (Cook & Robinson 2016). Key inputs to PVAs are the initial population size, the estimated combined annual impacts of the ORDs on demographic rates, and the baseline demographic rates (age-specific survival, productivity, and age at first breeding). PVAs used for impact assessments assume closed populations, i.e. no immigration or emigration. Some models incorporate forms of density dependence whilst others are density independent. In general, relative metrics of impact are likely to be less sensitive to the values of the baseline demographic rates and initial population size than absolute metrics of impact (Cook & Robinson 2016).

[2] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=LEGISSUM%3Al32042

3.2 Current quantification of uncertainty and variability

Uncertainty in PVA metrics will arise from both uncertainty and variability in the inputs to the PVA, from structural uncertainty within models (i.e., a failure of the population model to capture all of the biological processes that operate), and from stochasticity within the PVA model itself. The NE/JNCC PVA tool uses a "stochastic" Leslie matrix model that includes both demographic stochasticity and environmental stochasticity. Annual demographic rates tell us the expected number of births and deaths, but there is still variability in the actual number of births and deaths that occur in each year – this is demographic stochasticity. Even if demographic rates were fixed across time, demographic stochasticity still leads to variability in the growth rate of abundance, particularly when population sizes are small. Environmental stochasticity accounts for natural temporal variation in annual demographic rates, for instance as a result of variability in weather conditions.

The NE/JNCC PVA tool allows uncertainty in annual ORD impacts on demographic rates to be incorporated into the PVA, but does not explicitly allow for uncertainty in any of the other PVA inputs, such as the mean baseline demographic rates, level of environmental stochasticity, age at first breeding, maximum brood size, initial population size, and (where considered) level of density dependence. It also does not account for uncertainties associated with the underlying model assumptions – i.e. the assumption that the population is closed.

The PVA approach is simulation-based, and the key distinction between variability and uncertainty within this context is whether the values of inputs are simulated once for each simulation run (and the same value applied to all years), or simulated for each year within that simulation run. The NE/JNCC PVA tool accounts for variability, but not uncertainty, in productivity and survival because it simulates productivity and survival rates to use for each year within each simulation run but generates these in the same way within all runs. In contrast, it accounts for uncertainty, but not variability, within ORD effects because it simulates a single annual ORD effect for each simulation run, which is applied to all years within that run.

3.3 Potential for improvements to uncertainty and variability

We have identified, via the workshop, four broad areas in which improvements can be made to the current representation of uncertainty and variability in the application of PVA in ornithological ORD assessments.

3.3.1 Sensitivity analysis

The first of these is the use of sensitivity analysis, to investigate the extent to which the PVA outputs of interest are sensitive to uncertainty in each of the input parameters. This is important in helping to prioritise areas for further work. A key point here is that any such sensitivity analysis needs to focus on the outputs that are used in assessments, most often ratio-based metrics. It may be expected, and previous work suggests (Jitlal *et al*. 2017), that these relative metrics are primarily sensitive to inputs that relate to annual ORD impacts on demographic rates, and comparatively insensitive to the values of inputs relating to baseline conditions such as baseline demographic rates and initial population size. Sensitivity analysis can also be valuable in determining whether potential extensions to the current PVA model (see below) are likely to lead to substantive changes in key PVA outputs, and therefore can be useful in prioritising which of these extensions will lead to substantial improvements in the application of PVA within ornithological assessments.

3.3.2 Improved representation of uncertainty and variability within PVA models

A key component of uncertainty that is not currently accounted for within PVA models is uncertainty in initial population size. It would, in principle, be straightforward to incorporate uncertainty in this quantity into the NE/JNCC PVA tool. The key challenge lies in quantifying the form and magnitude of this uncertainty, which depends, in turn, upon constructing a plausible statistical model of observation error for seabird count data (*e.g*. from Seabird Monitoring Programme [SMP] count data, or other empirical observations of population abundance). A model of SMP count data would need, for example, to account for the differences in sampling methodology used at different colonies (such as the use of plot counts for some colonies and species, and whole colony counts for others), and for the different units used for counting (such as breeding individuals, breeding pairs, occupied burrows, etc.) and the uncertainties that arise from translating these into a common currency for use within PVA models.

Environmental stochasticity is accounted for within the NE/JNCC PVA tool, but currently under the assumption that stochastic variations in demographic rates are independent from year to year, and that variations in different demographic rates (productivity and survival) are not correlated with one another. These assumptions are unlikely to be biologically realistic. Inter-annual variation in demographic rates is unlikely to be independent because the underlying drivers, such as climate, exhibit patterns of temporal dependence. Moreover, correlations between demographic rates such as productivity and survival are likely to arise because (a) stochastic environmental effects act simultaneously on multiple demographic processes (*e.g*. poor weather conditions can impact on both productivity and survival), and (b) individuals may compensate from impacts upon one demographic process via other processes (*e.g*. adult birds may choose to prioritise their own survival over productivity). In practice, the inclusion of correlations (between demographic rates and over time) in PVAs is fairly straightforward – the R package associated with the NE/JNCC PVA tool already includes an option to specify correlation between rates, and it would be a relatively straightforward extension to also include correlations between years. The key challenge lies in empirically estimating these levels of correlation. The most promising avenue for doing so would be to focus on populations for which sufficient long-term data are available to be able to produce defensible annual estimates of both survival and productivity. In practice, the key challenge here is the lack of data that can be used to estimate survival in different populations, however new work is underway to assess ways of resolving this in the context of one species, black-legged kittiwakes (JNCC project: ‘Feasibility Study of Large-Scale Deployment of Colour-Ringing on Black-Legged Kittiwake Populations’), which will provide insight into addressing this issue more widely.

3.3.3 Validation and calibration of PVA models

The “baseline” PVA model essentially provides predictions for trends in population abundance. By running PVA models retrospectively (using the initial population size from a past year) the resulting predicted trends can be compared against the observed trends seen in the population abundance data. Discrepancies between predicted and observed trends indicate either errors in the values of PVA inputs, and/or structural errors in the model underpinning the PVA. Statistical models have been developed that use this discrepancy to estimate poorly known demographic rates, which may be only poorly constrained by inference from expert judgement or due to a lack of direct empirical data, as is common for juvenile survival in most seabird species. Although these models have been used in some contexts such as for seabird populations in the Forth-Tay (Freeman *et al.* 2014), they have not been used elsewhere. These models, which are effectively a form of data integration, merit further development, and have broader applicability than has currently been utilised within ornithological assessments.

3.3.4 Reduction of structural uncertainty in PVA models

The most substantive and wide-ranging area of improvement for PVA models relates to resolving structural errors in the current models by making their underlying assumptions more biologically realistic. There are a range of different ways in which the models could be refined to be more biologically plausible, and within this project we have identified the following as those with the most relevance to improving the use of PVAs within ornithological assessments by improving quantification and reduction of uncertainty:

- linking environmental stochasticity within PVA models to prey availability;
- linking environmental stochasticity within PVA models to climate change;
- Including and empirically parameterising density dependent processes;
- consideration of inter-specific interactions;
- inclusion of interactions between different ORD impacts (*e.g.* determining whether such impacts are synergistic or antagonistic);
- consideration of carry-over effects;
- Consideration of dispersal, immigration and emigration, potentially within the context of metapopulations

Of these extensions, only one (the inclusion of density dependence) is possible with the current NE/JNCC PVA tool. For all these extensions, however, the key challenge lies not in the relatively straightforward extension of the PVA models themselves, but rather in parameterising the additional processes in a defensible way. This depends, in turn, upon relevant analyses of empirical data – the outputs of these analyses can then be used to structurally improve current PVA models. In deciding which of these extensions to prioritise, it is important to consider the trade-off between model complexity, and the ability to defensibly parameterise this additional complexity, and the likely impact of the extension upon the PVA model outputs. The highest priority extensions are those where the additional processes can be prioritised using data (either existing or future), and for which the inclusion of additional processes is likely to have a substantive impact upon the key PVA outputs.

4 Displacement

4.1 Context

Two methods have been used to estimate the demographic impacts arising from displacement effects: the 'Displacement Matrix' approach (hereafter the 'matrix approach'), and the use of individual-based models, or IBMs. These two methods differ greatly in terms of the simplicity of the approach, the data used to underpin the approach, the structural complexity of the ecological processes and the biological realism involved, and in the treatment of uncertainty quantification.

4.2 Current quantification of uncertainty and variability

4.2.1 Matrix Approach

The matrix method is based on a simple calculation in which the density of birds within the wind farm footprint (plus a buffer) is estimated from local, at-sea survey data, and is subsequently multiplied by a user-specified displacement rate (% of birds within the footprint that are assumed to be displaced) and user-specified displacement mortality rate (% of displaced birds that are assumed to suffer mortality as a consequence of displacement). This results in an estimate of the number of birds killed by displacement impacts, which must then be apportioned back to relevant source populations (in the case of HRA) and converted to a change in survival rate for use in subsequent PVA. The Joint SNCB Interim Displacement Advice Note provides advice on how to present assessment information on the extent and potential consequences of seabird displacement from Offshore Wind Farm (OWF) developments (Joint-SNBC 2017). This advice requires assessments to use published indices of disturbance (*e.g*. Furness *et al.* 2013; Wade *et al*. 2016) to assign a range of displacement levels for each species individually, with consideration of modifications arising from emerging new evidence and discussions with SNCBs to agree appropriate levels of likely adult mortality associated with particular displacement levels, for each species individually. Assessments are then advised to use these two metrics (displacement rate and displacement mortality rate) to compile a 'Matrix Approach' table (*i.e*. representing proportions of birds potentially displaced/dying as a result of OWF development). The advice specifies that this table should be presented from 0-100%, in 10% increments for displacement levels. Percentage increments for mortality should also be presented between 0-100% but including smaller increments at lower values (*e.g*. 0%, 1%, 2%, 5%, 10%, 20%, etc).

There is no explicit consideration of uncertainty in the matrix approach. Instead, uncertainty in the two rates – displacement rate and displacement mortality rate – is visualised by use of the table, spanning a range of potential values for these two parameters. Uncertainty in the estimate of bird density, is also not considered. SNCBs recommend assessing impacts of displacement based on the overall mean seasonal peak numbers of birds (averaged over the years of survey) in the development footprint and appropriate buffer. SNCBs advise that at least two full years of monthly survey data should be collected pre-construction, which should be considered the bare minimum for assessment purposes. This provides a combined estimate of the number of birds on the water (corrected for survey coverage and distance analysis/diving species availability bias, if appropriate) and of the number of birds in flight (corrected for survey coverage). The methodology does, through use of multiple surveys per month over several years, attempt some consideration of environmental variation when performed at the scale of 'breeding' and 'non-breeding' seasons. However, a single mean seasonal peak value is used within the table calculations, so there is no formalised incorporation of environmental variation within the method.

In summary, the matrix approach provides a visual, qualitative consideration of uncertainty in displacement impacts. It depends upon three quantities; the input density data: the mean seasonal peak in the number of birds observed in the footprint and appropriate buffer area; and two parameters: the proportion of these birds that are displaced, and the proportion of these displaced birds that suffer mortality. The input density data includes a minimal consideration of environment variation, but no quantification of uncertainty. The two parameters include no formal quantification of environmental variation or uncertainty and instead use a qualitative, visualisation approach to presenting uncertainty in the output metric for the number of birds predicted to suffer mortality as a result of displacement impacts.

4.2.2 Individual-based Models

SeabORD is a stochastic, dynamic, individual-based model of seabird behaviour, energetics and demography during the chick-rearing period (Searle *et al.* 2014, 2018). The stochastic nature of SeabORD means that it is able to incorporate a range of different sources of variability. It allows for inter-individual variability in body mass, chick body mass and daily energy requirements at the start of the chick breeding season, and in displacement susceptibility (some individuals are simulated to be susceptible to displacement and/or barrier effects, whilst others are not, but this susceptibility is assumed to be constant within an individual over time). Temporal and inter-individual variability in the choice of foraging location chosen is also accounted for, as is variability in the choice of alternative foraging location if an individual is displaced by the ORD. Temporal and inter-individual variability in time budgets is incorporated indirectly, because time budgets are assumed, within SeabORD, to be linked to the choice of foraging location. These sources of variability all mean that there is variability in the final mass of individual birds at the end of the chick rearing period, and variation in whether or not their chicks survive to fledging. There is also assumed to be stochastic variation in the actual outcomes for each adult bird -- final mass is assumed to be related (via a logit-linear model) to the probability of over-winter survival, but there is still stochastic variation in whether any individual bird actually survives or not. If SeabORD is coupled with the sCRM to incorporate both collision and displacement and barrier effects, there is also variability in whether each individual dies from collision at each model time step, with the sCRM and simulated daily time budget determining the probability of collision for each bird at each time step.

Quantification of uncertainty within SeabORD is currently, in contrast, very limited. The values of most of the twenty or so input parameters, and, where relevant, the levels of variability in these parameters, are currently assigned based on published literature or expert judgement, and no uncertainty in these parameters is currently considered. Two parameters controlling intake rate are estimated, at a species level, by calibration against the mean number of foraging trips made per day, and the mean/range of time spent foraging per day, but the uncertainty associated with this calibration is also not currently quantified. A final, key, parameter, the total amount of prey, is calibrated for each new population and location, against empirical data on adult mass change and chick survival, and SeabORD does account for uncertainty in this parameter. The current advice to users is to run SeabORD multiple times (a relatively small number of runs, ten, being the standard choice, due to the model being computationally intensive to run), with a different level of total prey being used for each run. Current work within the Marine Scotland Cumulative Effects Framework project involves increasing the computational speed of SeabORD, which should ultimately enable larger numbers of simulations to be used in capturing uncertainty (a critical step if uncertainty in additional inputs is to accounted for), and the automation of the calibration process for the total prey parameter.

4.2.3 Estimating displacement and displacement mortality rates

To quantify the consequences of displacement by an offshore wind farm on a seabird population, we need an estimate of the proportion of birds displaced and, an estimate of the impact of that displacement on the population demographic rates. Rates of displacement are typically assessed used a comparison of pre- and post-construction monitoring data (e.g. Dierschke *et al.* 2016; Vanermen *et al.* 2015). Such studies highlight clear inter-specific differences in displacement rates. Some species, such as divers, gannets and auks, show a consistent negative response to wind farms. Others, such as cormorants show evidence for attraction, and several show no clear response (Dierschke *et al.* 2016). However, until very recently accurately quantifying the proportion of birds displaced has proven challenging (but see Heinänen *et al.* 2020; Peschko *et al.* 2020a, 2020b; Peschko *et al.* 2021).

Survey data are often over-dispersed and zero-inflated. Analysis of data collected from Thorntonbank Offshore Wind Farm in Belgium gives an indication of the scale of the problem (Vanermen *et al.* 2015). These data suggest that after 10 years, of 12 species present, it would only be possible to detect a decline of 50% for gannet and common guillemot. Whilst these analyses were based on data collected using boat surveys, similar results have been demonstrated for standard aerial surveys (Maclean *et al.* 2013). More recently, similar work has been undertaken in relation to digital aerial surveys (Donovan & Caneco 2020). In these studies, the inherent variability of the marine environment has contributed to the difficulty in quantifying displacement rates. However, in the past, these problems have been exacerbated by inconsistent approaches to, and, poor design of, post-construction monitoring studies (Marine Management Organisation 2014), and in the UK it only since the development of the MRSea model (Mackenzie *et al.* 2013) that more consistent methodologies have been applied to address these issues.

The expansion of GPS tracking studies in relation to offshore wind farms offers the potential to explore displacement in more detail. Consistent with the survey data, initial studies suggest that there may be little response by gulls to the presence of an offshore wind farm (Thaxter *et al.* 2018), both gannet and common guillemot may show a stronger response (Garthe *et al.* 2017; Peschko *et al.* 2020, 2021). Such data enable us to explore individual differences in response to offshore wind farms in more detail. For example, whilst the majority (89%) of gannets avoided the wind farm completely, a proportion (11%) made repeated trips through the wind farm when commuting between foraging areas and their breeding colony (Peschko *et al.* 2021). This indicates that the impact of displacement is unlikely to be distributed evenly across the population, with some individuals more likely to be affected than others.

After accounting for individual differences, analyses of GPS data suggest that observed displacement rates may be lower than those recorded using survey data (Peschko *et al.* 2021). A key explanation for this may be that GPS data are collected during the breeding season whilst data from surveys are collected year-round. During the breeding season, birds are constrained as central place foragers and may be willing to accept greater risks when foraging than in the winter when they are not constrained by the need to provision chicks. Such data would imply that the consideration of season-specific displacement rates may be appropriate.

Having quantified the proportion of birds displaced from a wind farm, it is necessary to consider what the impact of that displacement is likely to be. At a population level, displacement is likely to affect birds indirectly, through a reduction in survival as a consequence of the energetic costs for losing an area used for commuting or foraging and/or a reduction in productivity due to the increased energetic costs of provisioning young. At present, there is very little evidence with which to quantify the impacts of displacement on demographic rates. However, during the breeding season, evidence suggests that birds will attempt to buffer the impacts of increased energetic cost through reduced parental investment in chicks (Regular *et al.* 2014; Suryan *et al.* 2006) and that this may result in reduced productivity. Further analysis with an individual based model suggests that this may also have a more significant effect on adult survival during the breeding season than has been assumed in previous assessments (Searle *et al.* 2014, 2018, 2020). The increased use of GPS tracking technology offers the potential to investigate links between displacement from Offshore Wind Farms during the breeding season and demographic rates through the comparison of productivity and survival of birds that do, and do not, use offshore wind farms. Outside the breeding season, it is clear that there can be significant displacement of birds from offshore wind farms (Mendel *et al.* 2019). However, the demographic consequences of this displacement when birds are not constrained by the need to provision chicks and may therefore be more able to make use of alternative areas, is unclear.

4.3 Potential for improvements to uncertainty and variability

We first outline three broad ways in which the representation of uncertainty and variability in the individual based model, SeabORD, could be improved. The first of these is through inclusion of additional, direct, information of the levels of uncertainty/variability in some or all of the input parameters, and other inputs (*e.g*. prey maps, bird distribution maps). The second is through use of statistical methods to capture uncertainties associated with the process of calibrating SeabORD against empirical data. The third is through the improved representation of biological mechanisms, and the variability associated with these mechanisms, within the model. We discuss these further below.

4.3.1 Direct information on uncertainty and variability in inputs

SeabORD currently fixes the values of many of the behavioural and energetic parameters within the model. These include those relating to chick growth, chick daily energy requirements, intake rate parameters, energetic costs of different activities, the body mass below which adults and chicks are assumed to suffer mortality, the condition of the adult that stimulates abandonment of the breeding attempt, the probability of chick death due to adult unattendance at the nest, and the slope parameter determining the relationship between adult mass at the end of chick-rearing and subsequent survival. Most of these are based upon published studies, empirical data from the long-term study on the Isle of May, or in the case of intake rate parameters, calibration against time-activity budgets and foraging trip data from the Isle of May. However, in all cases, available information on both natural variability and uncertainty in these parameters is very limited or non-existent. Moreover, some parameters were derived from a study of a single species or distant location, which were taken to indicate a species-level value for all populations or applied to other species where data were lacking. This situation could be improved by undertaking a new, thorough literature review to identify further sources of information with which to add natural variability and uncertainty into model parameters. However, for many parameters, improvements will likely only be achieved by new empirical studies, dedicated to capturing natural variability (both environmental and between individual) and uncertainty in the processes of interest. Our assessment is that the most important parameters to prioritise for this effort are those relating to the mass-survival relationship (because this converts time-energy budgets into the currency of survival), and those relating to intake rate of foraging birds (which, when coupled with accurate prey availability data, will determine the day to day success of foraging adults, and hence their subsequent condition and provisioning rates to offspring). A recent Marine Scotland project estimated the form and strength of this relationship for four seabird species breeding on the Isle of May, including the estimation of uncertainty (Daunt *et al.* 2020), and outcomes from this work should now be incorporated into assessment tools relating end of breeding season mass to subsequent survival, such as SeabORD.
There can be situations in which it is difficult to derive knowledge of the values of input parameters from the literature or available data, but where experts nonetheless have useful knowledge regarding both the value of the parameter, and the level of uncertainty associated with this. Expert elicitation provides a mechanism for encapsulating this knowledge in a quantitative way, and typically involves assessing judgements on the level and form of uncertainty alongside judgement on the true value of the input parameters. Elicitation exercises typically involve multiple experts, in order that the judgements they incorporate relate to a community of experts, rather than to a single individual. Guidance on best practice have been developed that attempt to overcome known pitfalls and to minimise bias when eliciting information on uncertainty (EFSA 2014; Peel *et al.* 2018).

Accurate input data is key to driving accurate model outputs; for SeabORD this relates to seabird utilisation distributions around breeding colonies, and spatial prey availability across the foraging range of all colonies included within the simulations. We discuss seabird density/utilisation distributions in a later section. However, the development of spatially explicit, high quality, robust prey availability maps for breeding seabirds would greatly enhance the validity of SeabORD model outputs and allow for model structural enhancements to be made to develop more realistic foraging behaviours and patterns within the model.

4.3.2 Accounting for uncertainty when calibrating

The values of those input parameters that cannot be directly estimated from empirical data can instead be estimated by calibrating/estimating the model against observed data on model outputs, and the uncertainties associated with this process of calibration can be quantified. Standard calibration processes involve using numerical optimisation methods to identify the sets of input parameters that provide the best match, according to some metric (*e.g*. sum of squared differences, deviance), to observed data on one or more of the model outputs. However, many commonly used calibration methods do not account for the uncertainty associated with calibration. Likelihood-based methods do allow for quantification of uncertainty (Azzalini 1996) but rely on the likelihood of the model being calculable, and this is not feasible for complex simulation-based stochastic models such as SeabORD.

A range of modern statistical approaches to calibration do allow uncertainty to be quantified, even in contexts where the likelihood cannot be evaluated. Methods of likelihood-free inference, such as Approximation Bayesian Computation (ABC; Beaumont *et al.* 2002; Marjoram *et al.* 2003; Sisson *et al.* 2007), could in principle be used for this purpose, but are unlikely to be feasible in practice because they rely upon the ability to generate a large number of simulations from the model, and this is not currently feasible for SeabORD given the computational effort required to run the model. Emulation provides an alternative branch of statistical methodology, which also allows calibration to be performed in a way that accounts for uncertainty, and which is explicitly designed for models that are computationally intensive to run. Emulation involves approximating the process-based model (*e.g*. SeabORD) using a statistical model (Kennedy & O'Hagan 2001): the central idea of emulation is to (a) run the mechanistic model (*e.g*. SeabORD) for a relatively small number of sets of input parameters, and (b) construct a statistical model that describes how the key outputs of the mechanistic model vary in relation to the values of the input parameters. This statistical model can then be used to quantify the uncertainty associated with calibration, whilst accounting for the uncertainty that arises from the relatively small number of runs of the process-based model. Emulation methods can also be used to quantify, and account for, the presence of structural error in the model – structural uncertainties arise if the model systematically deviates from reality even with an optimal choice of input parameters. Emulation methods were originally designed for use with complex deterministic models, but more recent variants of the methodology also allow it to be applied to stochastic models, such as SeabORD (*e.g.* in the context of individual-based models, Oyebamijia *et al.* 2017).

4.3.3 Improving realism of model assumptions and resolving structural errors

In order to fully quantify, and ultimately reduce, structural uncertainties, it is necessary to identify areas of SeabORD in which the biological assumptions are unrealistic, and to replace these with more biologically plausible alternatives. In practice, seabird behaviour is extremely complicated, and consequently it is likely that none of the biological assumptions that underpin a model like SeabORD are likely to be entirely true. This does not always matter - some model assumptions may be incorrect, but nonetheless lead to models that provide a very good approximation to reality. However, other assumptions may be sufficiently incorrect that they lead to substantial levels of structural uncertainty. It remains the case that the current version of SeabORD has had to make a series of simplistic assumptions about some of the key behavioural processes driving model outcomes, primarily due to a lack of data on which to build more realistic approximations.

There are many ways in which the structure of SeabORD could be refined to improve biological plausibility, but our judgement is that the key priorities for further development are those that have already been identified as priorities elsewhere (Nature Scot Marine Bird Impact Assessment Guidance Workshop Report, February 2020):

- improved representation of overall prey levels and availability, and of spatial heterogeneity in prey;
- improved representation of flight paths, of the estimated bird density maps that underpin these, and of the joint distribution between seabird and prey spatio-temporal dynamics;
- improved representation of displacement, barrier and collision effects;
- improved representation of behaviour, energetics and ORD interactions outside the chick-rearing period;
- improved representation of the relationship between adult mass at the end of the chick rearing period and subsequent over-winter survival.

We regard these as high priorities for improving the structure of SeabORD in part because they all represent key components of the model, but also because, in each case, we can propose specific actions that can be pursued in order to make the biological assumptions of the model more realistic, by refining the way that data are used to inform the model structure and parameters. There are other components of the model that are potentially influential, and currently contain biological assumptions that are likely to be over-simplistic, but where it is difficult to see how the model could usefully be improved, given current data or additional data that could currently be collected.

4.3.4 Displacement impacts

For displacement impacts, there is a clear need for the development of data and analytical methods to provide better estimates and uncertainty quantification for the displacement rate and the displacement mortality rate, particularly in relation to variation in the environment, seasonal differences, and the characteristics of ORDs. Although individual-based models like SeabORD are designed to estimate the displacement mortality rate as a model output, there is still a clear need for empirical data on this rate to both validate IBM outputs and improve model structure, thereby reducing structural uncertainty.

There is emerging evidence that displacement rates may vary spatially, such that the application of a single displacement rate to both the ORD footprint and buffer area may be unrealistic, and that different colonies may have different displacement rates when interacting with the same ORD site. There is a strong need for empirical data and analytical methods for developing understanding of how and why displacement rates and displacement mortality rates vary both spatially and temporally, such that better predictive models can be developed and validated to both improve the precision of displacement rates and displacement mortality rates within assessments, and to better quantify the uncertainty associated with these estimates so it may be propagated and properly reflected in assessment outputs. This will undoubtedly need to be underpinned by the use of individual-level tracking data that can be linked to condition and fitness (productivity and survival) measurements over multiple seasons and years. This is because of the inherent difficulty in trying to measure and reliably quantify these processes at the population level, where any displacement-driven changes in seabird distribution or population demography will be masked by the co-occurrence of many other pressures, such as changing environmental conditions and other anthropogenic activities such as fishing and climate change.

There is also a need for empirical data and analytical methods for quantifying barrier effects in seabirds, and the consequences of these effects on condition and fitness. Barrier effects are modelled and estimated within IBMs like SeabORD, but are currently represented very simplistically, and are reliant upon a number of untested assumptions about the nature of an individual's behavioural response to a perceived barrier. SeabORD does allow users to vary the barrier rate and displacement rate separately (both are user-specified inputs to the model), such that birds may be classified into different susceptibility categories – birds that are unaffected by wind farms, birds that are displaced but not barriered by wind farms, and birds that are barriered and displaced by wind farms. The model assesses these processes simultaneously during simulations, and model output can be used to compare resulting demographic changes for birds in each category. However, barrier effects are represented very simplistically within SeabORD, assuming no change in barrier rates over time or in relation to bird state or environmental conditions and using simplified assumptions about the flight paths of barriered birds that encounter ORDs. The collection of individual level tracking data would allow for better measurement of the occurrence and form of barrier effects, how they vary between individuals, in relation to environmental and ORD characteristics, and over time. The matrix approach is designed to primarily consider the impacts of displacement on individuals and does not explicitly consider barrier effects. However, the inclusion of birds in flight in the input data to the matrix approach means that the method assumes that the resultant mortality is a combination of displacement (foraging birds) and barrier effects (birds in flight). This method therefore has the underpinning assumption that the proportion of birds affected by displacement and barrier effects are the same (i.e., there is one 'displacement rate' that is applied to both birds in flight and birds assumed to be foraging/resting at sea), and that the resultant mortality is the same between the two processes. A better understanding and quantification of barrier effects, alongside that of displacement effects, would therefore greatly enhance the quantification and potential reduction of uncertainty within assessments. Importantly, the matrix approach also does not take account of turnover in individual birds observed within the footprint surveys, thereby contributing additional, unquantified uncertainty in outputs.

Finally, the two methods currently considered for use in estimating displacement impacts in ornithological ORD assessments vary considerably from a simpler, generalizable approach (the matrix method) to a more biologically realistic, but data-hungry and location specific approach (SeabORD). Both approaches have their advantages and disadvantages, however it may be useful to invest resources in developing more of a 'middle-ground' approach to estimating displacement effects and their impact on demography. For instance, using more of a habitat-based approach to estimate how ORDs alter or remove habitat from seabird colonies, and the impact of this upon productivity and survival.

5 Collision

5.1 Context

Collision risk estimates for UK wind farms are calculated using the deterministic methods developed by Band *et al.* (2007) and Band (2012), and subsequently built into a simulation tool (sCRM) to allow stochastic variations in parameters (Masden 2015; McGregor *et al.* 2018). Three sets of parameters are used in the model (McGregor *et al.* 2018); site specific seabird data (monthly densities of birds in flight, site-specific flight height distributions), generic seabird data (biometrics and flight characteristics) and turbine data (rotor size, hub height, RPM, etc.).

5.2 Current quantification of uncertainty and variability

Prior to the development of the sCRM the alternative means to represent uncertainty in collision estimates was to use upper and lower confidence estimates for key input parameters (*e.g*. seabird density, flight height, avoidance rate, etc.) to give an indication of the likely bounds of the collision predictions. However, this is undertaken for single parameters at a time and omits the context present in the underlying probability distributions, which in the case of density estimates can often be heavily skewed. This also results in impact assessments presenting several alternative versions of the collision predictions, introducing additional complexity to the process.

The sCRM, by simulating across multiple parameter distributions, is therefore a significant step forward for presenting uncertainty. Indeed, the assessment process is developing to reflect the changes introduced by the sCRM. The key challenge that remains is how to integrate the outputs from the survey data analysis into the sCRM in a robust and repeatable manner, and how to take the outputs from the sCRM and combine them in a consistent manner with those from other aspects of the assessment (*e.g*. displacement), and finally how to use these as inputs to PVA (these are tasks being tackled currently by the CEF project).

5.3 Potential for improvements to uncertainty and variability

5.3.1 Areas for development

Whilst the sCRM has addressed the need to properly consider parameter variability, the underlying model remains unchanged and lacks what are likely to be important features of seabird behaviour, how these are related to weather conditions and how seabirds will interact with the turbines themselves.

Weather conditions, and most pertinently wind speed, are likely to influence flight characteristics (height and speed) and obviously also influence rotor operation (blade angle and RPM). Collision risk predictions are positively related to RPM, and both positively and negatively related to flight speed (less risk of collision on rotor transit, but increased number of transits). If higher wind speeds also affect flight heights, this could also be an important consideration. Disentangling these will require a combination of observations at wind farms, and tag-based studies incorporating collection of local weather conditions. There are also likely to be seasonal considerations to these relationships, for example reflecting daily foraging requirements in the breeding season versus long distance seasonal migration and over-wintering distributions. However, it is also worth remembering that the temporal unit for collision estimation is month, reflecting the frequency of survey data collection. Consequently, unless a smaller temporal unit is adopted, the requirement for the collision estimates is that they are representative of conditions throughout the month in question. Thus, improving understanding of these relationships needs to keep this requirement in mind.

Collision predictions are most sensitive to the avoidance rate value, and this is a critical focus for impact assessment purposes. Current estimates are based on an amalgamation of data sources very little of which has been collected at operational offshore wind farms, due to the difficulty of undertaking long-term studies at these locations. As with other aspects of behaviour, there are also likely to be relationships with weather conditions. Currently avoidance is applied as a single (albeit stochastic in the case of sCRM) figure across all estimates (i.e. this lacks any seasonal variation). While improving confidence in avoidance rates should remain a priority for monitoring studies it must be acknowledged that considerable effort (both time and expense) is likely to be required to achieve improvements in our understanding.

An example of the kind of detailed study that may be required is tracking of flight paths in three dimensions through wind farms. These will likely require high resolution GPS tags, high resolution stereo cameras and tracking algorithms or combinations of both. While such technology is either available or in development, it will need to be deployed on a large scale to obtain sample sizes sufficient to begin addressing the behavioural and interaction questions of interest. Thus, it will be the move from current proof of concept stage to commercial deployment as a standard monitoring option for new wind farms that will be necessary for the results to feed back into impact assessments.

It is clear from the brief summary above that it would be helpful to determine a priority list of monitoring that balances gains (in reduced uncertainty) against time, expense and potentially likelihood of success. One option for this would be to conduct a fairly high-level sensitivity analysis that could be used as a guide for decision making.

As previously indicated with regards PVA, sensitivity analysis is a valuable tool for practically examining components of a mathematical/statistical abstraction of a system. This is particularly true in the case of simulations with Monte Carlo treatment of uncertainties, such as found within the sCRM, IBMs and PVAs.

Models can be complex, with a number of linked elements providing their outputs – such that the contribution of individual inputs to the outputs cannot be determined by simple inspection. This becomes more pronounced when several modelling tools themselves are linked, with their own concomitant uncertainties, *e.g*. an IBM feeding into a CRM then to a PVA. A sensitivity analysis evaluates the practical importance of the various inputs to a model, by perturbing these with resulting changes in outputs examined practically (*e.g*. Donovan *et al.* 2017). In the case of a CRM, the output is primarily animal mortalities – so inputs with greatest effect on these numbers require the greatest understanding, accuracy and precision. In practice this may be as simple as fixing all inputs but one, which is subject to carefully considered simulation of uncertainty, leading to a distribution of mortalities. This is iterated over all inputs to provide a ranking by sensitivity.

A sensitivity analysis is very informative about where research effort ought to be focussed. If the analysis suggests the model is sensitive to particular assumptions or parameters, then research to confirm the assumptions or increase the precision of the parameter estimates, is high priority. Conversely, non-influential assumptions or parameters warrant lesser consideration – the model is robust to these inputs. The analysis may also extend to alternative or additional modelling components. Balanced against budgetary constraints, the sensitivity analysis informs a return-on-investment for research priorities. In the context of uncertainty here, the sensitivity analysis would indicate which inputs contribute most to the precision of outputs, and thereby a priority list for reducing these uncertainties. This is not limited to collision, but all modelling components of impact assessment whose contribution to uncertainty can't be easily evaluated by inspection.

6 Density estimation and apportioning

6.1 Context

6.1.1 Apportioning

Apportioning is currently used within the assessment process to partition seabirds in the breeding and non-breeding seasons by colony. Apportioning during the breeding season can be defined as the expected number of birds N_{AC} from colony i of size C_i within an area A_i, which is $N_{AC} = C_i \sum_i A$. The total number of colonies is M. Therefore, the proportion of birds (defined as the apportioning proportion) within A_i that originate from i is:

$$\frac{N_{AC}}{\sum_{i=1}^{M} N_{AC}}$$

Therefore, apportioning requires an estimate of the utilisation distribution associated with each colony and of the colony size. Producing utilisation distributions is reliant on spatial data, the nature of which is dependent on how and why it is collected. For example, aerial or boat-based surveys are used to produce mean or median bird densities within ORD footprints. GPS tracking data are used to estimate utilisation distributions as an input to SeabORD (Searle *et al.* 2014, 2018). Apportioning tools currently used within the assessment process are the SNH Apportioning Tool (SNH 2018), MS Apportioning Tool (Butler *et al.* 2020b), and the Biologically Defined Minimum Population Scales (BDMPS; Furness *et al.* 2015).

6.1.2 Spatial data & modelling approaches for density estimation

Spatial data are utilised in relation to the assessment process at differing spatial scales: broad-scale, project-level, and colony-level.

Broad-scale data encompass the UK and the European continental shelf and are based on off-shore aerial and boat-based surveys that capture year-round spatial distributions of birds (and marine mammals). Generalised estimating equations – generalised linear models (GEE-GLMs) were used to synthesise different data sets and produce predicted distributions over time. They were presented at 10km resolution and provide insight into seasonal offshore space use (Waggitt *et al.* 2019).

Project-level data are collected through aerial surveys that collect strip transect (or grid) based data over the proposed development site plus a buffer (usually 4km). Data are collected along transects in each calendar month for a minimum period of two years (i.e. 24 surveys) thereby providing two estimates in each month. Survey data are generally analysed using either design-based methods (i.e. extrapolation of observed data to the unobserved areas) or model-based methods (*e.g*. spatial model such as MRSea) which incorporate spatial smoothers and covariates to permit estimation of abundance in unobserved areas within the survey area. The R package MRSea (Scott-Hayward *et al.* 2013, 2014) uses a Spatially Adaptive Smoothing Algorithm (SALSA; Walker *et al*. 2010) to account for missing data by making the assumption that the density of animals varies smoothly over space. The approach allows adjustment for the presence of missing data by exploiting empirical relationships between abundance and other variables. MRSea allows the quantification of uncertainty through bootstrapping the mean estimates. Whilst mean estimates from spatial models are robust, the models perform poorly for species present in low numbers and for these it is necessary to use design-based estimates. Bootstrapping of images along transects can be used to obtain confidence intervals for such estimates.

Colony-level utilisation distributions in the breeding season are derived using the distance decay relationship described above combined with colony-specific GPS tracking data. Wakefield *et al.* (2017) used multi-colony GPS tracking data for four seabird species (guillemot, razorbill, kittiwake and shag) to build statistical models that empirically describe the colony-specific spatial distributions of birds from these species in relation to both accessibility and environmental heterogeneity. The models are based on generalised linear mixed models (GLMMs). The full likelihood is unknown, and the residuals are relative rather than absolute and so uncertainty cannot be fully quantified.

6.2 Current quantification of uncertainty and variability

6.2.1 Apportioning

The SNH Apportioning Tool makes strong and biologically unrealistic assumptions and fails to quantify uncertainty. It assumes that the abundance of birds at a location of distance d from the colony is proportional to d^(-2), and the number of locations at this distance is proportional to 2πd. The total number of birds at distance d is therefore proportional to d^(-2) 2πd, and proportional to d^(-1), representing a strong central-place foraging effect (Butler *et* *al.* 2020b). As the utilisation distribution is proportional to d^(-2) until the foraging range, and zero beyond, this means that the foraging range is effectively the only unknown parameter. The underpinning assumptions that distance by sea to colony is the only important explanatory covariate, density decays in proportion to the inverse distance squared, and the foraging range is static for each species, are biologically unrealistic. These assumptions can introduce structural uncertainty into calculations as they fail to account for the heterogeneity in environmental (and geographical) space (Wakefield *et al.* 2017) and for competition effects (Bodey *et al.* 2013), and so errors can arise in the calculation of apportioning percentages.

The MSS Apportioning Tool is based on habitat use models derived from tracking data (Wakefield *et al.* 2017) for four species of seabird (European shag, black-legged kittiwake, common guillemot, razorbill). An advantage of this approach is that habitat use is species and colony specific, informed by the environmental availability of each colony. A disadvantage is that only partial estimates of uncertainty are available as the full residuals are not available. However, for species where tracking data are available, the MSS Apportioning Tool is a more defensible approach (than the SNH Apportioning Tool), and as more species are tracked, this approach can be used for estimating the foraging ranges for these species.

6.3 Potential for improvements to uncertainty and variability

6.3.1 SNH apportioning tool

Rather than assuming the foraging range is static over all colonies for a species, potential improvements to this tool could produce more viable foraging ranges disaggregated by colony and quantify uncertainty. For species where tracking data are not yet available, foraging ranges can be derived from published distributions (*e.g*. Woodward *et al.* 2019) by estimating the rate of decay of utilisation with distance. The inter-colony variation can be defined so that foraging ranges can be disaggregated by colony, region, or meta-population as appropriate. Using the inter-colony variability in foraging range, uncertainty in apportioning percentage can then be estimated using a simulation-based approach (MS SEANSE; Searle *et al.* 2020).

6.3.2 MSS apportioning tool

To quantify uncertainty within this tool requires more thorough statistical approaches to properly address the intrinsic complexities within tracking data such as spatial and temporal autocorrelation, such that appropriate uncertainties around estimated habitat utilisation distributions can be incorporated within the tool.

6.3.3 BDMPS approach

The BDMPS approach is currently used to define the reference populations for all seabird species in the non-breeding season. However, while BDMPS was developed from a comprehensive review of migration and movement literature (Furness *et al.* 2015), there remains scope for refinement through expansion of the type of large scale tagging studies conducted for kittiwakes (Frederiksen *et al.* 2012) and more recently auks (Buckingham *et al.,* in prep). Further improvements to apportioning in the non-breeding season could be made through updating population sizes and demographic rates (survival rates, age of first breeding, productivity) to more accurately assess the numbers of immature birds present in different regions. Including available data from the timing of breeding and migration from populations breeding inside and outside of the UK could address some of the issues within the BDMPS that arise from paucity of information.

6.3.4 Density Estimation

Producing distributions using habitat association models for more species (*e.g*. lesser black-backed gull, fulmar and Northern gannet) is plausible as more tracking data becomes available. Outputs from approaches such as Wakefield *et al.* (2017) are currently used in seabird tools such as SeabORD and are being implemented into the MS Cumulative Effects Framework (CEF). Therefore, additional species distributions would be straightforward to implement into the existing framework, providing more defensibility than the current approach of using foraging ranges, and would in turn reduce structural uncertainty. However, there are two issues to producing plausible uncertainty estimates: one is the limitation of the GLMM approach and the other is how much uncertainty can be quantified and incorporated into a modelling framework. It should be recognised that some types of uncertainty cannot be included and/or propagated through every modelling or statistical approach. An example for the GLMM approach may be to account for within-individual variation but not propagate this uncertainty through to the final estimate (Patrick *et al*. 2014).

One potential solution is to use a spatially explicit analytical framework so that estimates of uncertainty can be produced. Integrated Nested Laplace Approximations (INLA; Rue *et al.* 2009) is a fast-fitting hierarchal Bayesian framework that can be used to model tracking data as a spatial point-process with a specified two-dimensional random field. This modelling framework is particularly useful for dealing with autocorrelation, which is prevalent in tracking data as locations are related in both time and space. If underlying autocorrelation in tracking data is not dealt with properly, covariates can be falsely identified as being important to a species' habitat selection because uncertainty in the parameter estimates can be underestimated.

Non-breeding season utilisation distributions can be derived using the BDMPS model combined with geolocator data (GLS) from tracked individuals. Geolocator data are light-level data loggers, which are lightweight and long-lasting. Because position is estimated using ambient light intensities and elapsed time, GLS locations have relatively large uncertainties up to hundreds of kilometres (Merkel *et al.* 2016). However, these data can offer insight into the movement and distribution of seabirds during the non-breeding season, and uncertainty can be somewhat reduced using various methodologies (Merkel *et al.* 2016). If colony-specific distributions can be estimated from GLS data then, as with GPS data in the breeding season, these distributions can be combined with counts of colony size and be used to apportion birds to colonies within the non-breeding season. Producing spatial distributions from GLS data has similarities with the modelling of GPS data, but there are some important differences:

a) GLS data are much lower frequency than GPS data – typically 1-2 records per day, which means detailed modelling of behaviour and local spatial movement are not possible.
b) Levels of observation error in GLS data are much higher than for GPS data, and are sufficiently large that models ignoring observation error are unlikely to be defensible.
c) The level of observation error in GLS data are likely to be heterogeneous, and because it is likely that they will vary according to known factors (*e.g*. time of year) this variability can be modelled.

These differences mean that the methods used to build models that can be used to apportion in the non-breeding season will necessarily differ from those used in the breeding season. Utilising available geolocator data (on guillemots and razorbills) on birds tagged in the non-breeding season could provide insight into broad-scale distributions of multiple tracked birds which could be scaled to population level with uncertainty quantified. The INLA approach can be combined with Stochastic Partial Differential Equations (SPDE; Lindgren *et al.* 2011) which allows modelling over non-regular grid cells of varying scales. A mesh can be used to cover the areas of interest using different spatial scales. As seabirds can be wide-ranging in the non-breeding season when the central-place foraging constraint is relaxed (Furness *et al.* 2015), spatial analysis needs to be able to deal with a broad spatial extent over varying spatial scales. INLA-SPDE models are fast-fitting and all colonies with data can be fitted as separate components of the same model. However, GLS data have large location errors and attempting to use environmental variables as explanatory covariates may not be feasible. Therefore, to predict distributions for colonies where count data are available, but no birds were tagged with geolocators, the spatial random field generated by the fitted model could be used as a prediction surface for these colonies.

6.3.5 Data integration

More data are now being collected in relation to ORD developments and spatial planning, with technological advances allowing more varied surveys to be undertaken and data types to become available: drones (Rush *et al.* 2018), aerial and boat-based surveys (Hammond *et al.* 2002, 2013, 2018), camera imaging on wind farms to assess collisions (Skov *et al.* 2018), and biologging devices that track location and collect in-situ environmental information (Cleasby *et al.* 2015; Isaksson *et al.* 2021), movement through accelerometers (tri-axial movement) (Warwick-Evans *et al.* 2017), and behaviour through time-depth records (Peschko *et al.* 2020).

The question arises of how can we best use these data? Varied data may require specific statistical analysis techniques to address intrinsic issues such as autocorrelation, but more knowledge will be gained, and hence uncertainty reduced, if at least some of these data can be integrated (Matthiopoulos *et al*. 2020). Obvious advantages are broad-scale coverage of spatial distributions of species when combining surveys that are collected using different survey techniques. An example is the SCANS surveys which are approximately decadal census of cetaceans around the UK and European continental shelf through combining design-based boat and aerial surveys of cetaceans (Hammond *et al.* 2002, 2013, 2018). Extended temporal coverage can be achieved when combining surveys from different years or season. This could be particularly useful for extracting additional value from appropriate assessments, which in isolation represent two years of project-level surveys but when combined could provide information on inter-annual variability. If multiple at-sea footprint surveys across time and space could be compiled, then an estimation of baseline inter-annual variability could be obtained.

For effective integration, two criteria need to be met: data need to overlap or align either spatially or temporally, and state-of-the-art statistical methods must be developed to deal with intrinsic data issues and propagate uncertainty through the model. It is obvious that not all data can be integrated, but here we give an example of where data integration could advance seabird assessments. The distribution of non-breeding birds during the breeding season is not normally considered and leads to differences between assessments based on GPS data and those based on at-sea surveys (Sansom *et al*. 2018; Searle *et al.* 2020). Integrated modelling of these two data sources would allow the distribution of non-breeding birds to be estimated, and the uncertainty associated with this component of the population to be quantified. When considering integration, differences in the nature of the data should be considered. In this example, GPS and at-sea data could represent differing spatial extent (coastal vs. more pelagic), spatial scale (fine-scale vs. broader-scale), and temporal (discrete time-steps vs. snapshots), and any statistical assumptions made within the models used for integration would need to account for these differences. We set out two modelling approaches that could be used to integrate these data.

6.3.6 Spatial point process approach

A potential approach for integrating these data uses INLA-SPDE to model the data as a spatial point process using joint response modelling (with separate random fields) to accommodate the different data types. There are several advantages to this approach. Varying scales of data can be handled using Stochastic Partial Differential Equations (SPDE), using a non-regular mesh. This allows areas of interest where there are fine-scale observations to be mapped at a high resolution, and remaining areas to be mapped at lower resolutions. Explanatory covariates used in the model can be shared between the joint responses, potentially conserving processing time. Uncertainty can be propagated through so a single estimate of uncertainty can be extracted from the model. One disadvantage of this approach is the difficulty in producing model diagnostics and validation (Yuan *et al*. 2017) as metrics such as the Bayesian Information Criterion (BIC) cannot be calculated (Schwarz 1978) and so comparative visualisation is typically used.

Caution should be taken when integrating data. Integrating partial datasets, where much of the data are missing, is likely to lead to a misspecification of the model, biased estimates, and incorrect uncertainty estimates. The same (if not higher) data integrity and quality thresholds should be used when integrating data as when data are used in isolation, as data integration cannot fix issues, only account for intrinsic properties of the data.

6.3.7 Movement modelling approach & model validation

Movement models are used to predict behaviours (*e.g*. foraging, resting at sea, diving) and estimate activity budgets of seabirds fitted with biologgers to investigate flight paths with respect to collision risk and displacement (Cleasby *et al.* 2015; Warwick-Evans *et al*. 2017; Peschko *et al.* 2020). A class of movement models that has become popular in ecology for analysing tracking data are Hidden Markov Models (HMMs), which are state-space time series models that assume the observed (state-dependent) time series is driven by an unobservable (‘hidden’) state process. They are used to sequence behaviours (states) and can account for serial dependence between observations (Patterson *et al.* 2008; Langrock *et al.* 2012). HMMs are straightforward to implement, aided by R packages such as moveHMM (Michelot *et al.* 2016) and momentuHMM (McClintock & Michelot 2018). HMMs are implemented by assuming equally spaced locations form a bivariate time series with step-length (lt, distance between two locations) and turning angle (φt, angle between two locations) defining the changes between consecutive locations. Depending on the complexity of the behavioural states required, combining locational data with ancillary information such as accelerometer, time-depth recorders (TDRs), or environmental covariates can produce more plausible models. For example, where at-sea behaviour is required to be disaggregated into behaviour states beyond foraging and flying (*e.g*. resting on water, flapping flight, gliding flight, foraging, and taking off), accelerometer data can provide additional information to delineate between these behaviours (Berlincourt *et al.* 2015). Within the context of quantifying and/or reducing uncertainty, there are two limitations to this movement modelling approach: model validation and propagating uncertainty.

Typically, model validation is difficult to achieve because ground-truth data are generally unavailable. However, where animals have been fitted with a device which records GPS and time-depth records, there is an opportunity to fit a movement model using only location data and use the depth information to validate model accuracy in determining diving and non-diving behavioural states (Browning *et al.* 2017). Validating a location-only movement model could be useful in circumstances where only some individuals had TDRs but all had GPS functionality in the tag and a general movement model was required, where further telemetry deployments could only use GPS tags (for example, due to prohibitive cost of TDRs), or where research had found that the weight of TDRs caused adverse effects to individuals.

Typically, HMMs do not consider observation error on location but treat the state as part of a stochastic process (Patterson *et al.* 2008). Continuous-time Markov chain Monte Carlo models use velocity and momentum (rather than step-length and turning angle) and allow for behavioural switching to occur continuously in time rather than at (discrete) observational times (Parton & Blackwell 2017). They can account for observation error and for irregular observations. Using continuous-time models can allow for uncertainty to be quantified and for more 'realistic' (foraging) trips to be simulated (Blackwell 2019). In this way, GPS and at-sea survey data could be integrated through sampling from a utilisation distribution generated by the movement model.

6.3.8 Environmental covariates – prey distribution and availability

Species distribution or habitat preference maps that form the inputs to displacement and collision risk models are produced using spatial data from seabirds such as GPS tracks or at-sea surveys (Wakefield *et al.* 2017; Waggitt *et al*. 2019). Habitat preference models associate animal space use with characteristics of their environment (Aarts *et al.* 2008). When these models are used to predict space use, choosing appropriate explanatory covariates is important. The marine environment is dynamic and mostly inaccessible so collecting and defining appropriate covariates can be challenging. Currently, covariates that represent proxies of prey fields are used due to paucity of information. Additionally, the marine ecosystem is complex, seasonal, and dynamic. Associating top predators with oceanographic covariates such as sea surface temperature when there are many complex biological and physical processes between them can make habitat association modelling difficult due to weak explanatory power in a model where covariates do not adequately capture heterogeneity in environmental space. However, assessments of the impacts of ORDs on seabird populations need to account for pre, during, and post development activities along with seasonal variation in seabird habitat use due to life history (pre-breeding, breeding, chick incubation, non-breeding, and migratory), as well as population response to environmental variability. Understanding the complexity of the impacts as well as proper quantification of uncertainty can only be achieved through the collection of good quality covariates of direct prey of seabirds that are required to produce habitat association models or individual-based models to explain seabird behaviour and activity budgets. Using prey data (instead of proxies) allows us to account not only for environmental variability but provides a direct link to causal mechanisms of key drivers in seabird behaviour. Understanding these drivers and producing accurate spatio-temporal species distribution maps for seabird species is essential for assessing how anthropogenic activities such as ORDs will impact seabird populations. Information on prey fields can then be combined with oceanographic covariates to identify and characterise different scales of seabird distribution and the underlying mechanisms that drive change over space and time.

7 Post-consent monitoring

The tools within the assessment process are essentially "predicting" the likely impacts of future ORDs. A key potential mechanism for reducing uncertainty, therefore, is through the incorporation of data that quantify the impacts of existing ORDs. These data include both post-consent monitoring data that developers are required to, or elect to, collect (*e.g*. at sea survey data, radar data to detect collisions and micro-avoidance), but may also include other monitoring data (*e.g*. foraging, provisioning and nest attendance behaviour, demographic rates, colony counts) that can be used to retrospectively assess the impacts of ORDs upon seabird populations.

The most obvious use of such post-consent monitoring data is to refine the estimates of key input parameters, such as displacement and avoidance rates. As the amounts of available data increase, the levels of uncertainty associated with these inputs to the assessment process should decrease, with the result that the incorporation of post-consent monitoring data has the potential to reduce uncertainty.

The other key role of post-consent monitoring data is in validation -- attempting to detect structural errors within the tools used for assessment. SeabORD, for example, makes specific assumptions about the paths that birds will take to avoid wind farms (barrier effects), and post-consent GPS monitoring data should allow us to evaluate whether these assumptions are plausible. By identifying structural errors, and providing the empirical basis to resolve these by making existing tools more realistic, the incorporation of post-consent monitoring data may therefore appear to increase uncertainty – in reality, however, the presence of unidentified structural errors in tools would mean that uncertainty is currently being underestimated, so that their resolution leads to a more accurate, and defensible, quantification of uncertainty.

Post-consent monitoring data can be used to inform specific components of the assessment process (Table 1). Broader-scale data (*e.g*. on population size and abundance) can also be used to detect whether the overall ORD impacts produced by assessments are consistent with the levels of change in demography and abundance that are seen after construction. However, these broader-scale data are not able to distinguish the cause of any discrepancies – which components of the assessment process are introducing error and are also likely to have low statistical power to detect differences (Cook *et al.* 2019). The key focus of post-consent monitoring data, therefore, should be on informing and validating specific inputs and component tools used within the assessment process.

Many of the more substantial knowledge gaps, or topics for which uncertainty could be reduced are ones that operate across large spatial and/or temporal scales and therefore need to be addressed through strategic studies, rather than as individual offshore wind project level post-construction monitoring studies. Administered by an advisory group with a core scientific remit with funding provided by the relevant stakeholders (wind farm developers/operators, regulators and statutory agencies), a strategic approach to monitoring would be able to undertake the large-scale studies needed.

8 Communication of uncertainty to decision makers

A key risk in the consenting of Round 3 and Scottish Territorial Waters (STW) offshore wind farms has been uncertainty in environmental data and the consequential conservative assumptions made within the modelling of potential impacts on environmental receptors. This has resulted in projects being refused consent (*e.g*. Docking Shoal), reduced in size from licence applications (*e.g*. Race Bank, Beatrice and Moray Firth Offshore Wind Farms), or being challenged in court (i.e. the Forth and Tay offshore wind farms). An existing framework on the consistent treatment of uncertainty using calibrated language was published by the Intergovernmental Panel on Climate Change (IPCC) (Mastrandrea *et al*. 2010). This has been used in the past as part of environmental impact assessments, *e.g*. the seal assessment framework (Thompson *et al*. 2013), the Greater Wash determination in 2012 and, Marine Scotland’s Acceptable Biological Change framework. The seal assessment framework (Thompson *et al.* 2013) made use of the guidance on the interaction of evidence and agreement on that evidence (Table 2), to determine confidence. The framework then applied precaution until expert and stakeholder confidence was sufficiently high that the assessment of impact from construction noise on seals addressed the uncertainty in the environmental data appropriately.

Table 1: The interaction of evidence and agreement statements, and their relationship to confidence. Confidence increases to the top right corner as suggested by the increase strength of shading. From Mastrandrea *et al.* (2010).

Agreement ↑				Confidence ↑
	High agreement Limited evidence	High agreement Medium evidence	High agreement Robust evidence	
	Medium agreement Limited evidence	Medium agreement Medium evidence	Medium agreement Robust evidence	
	Low agreement Limited evidence	Low agreement Medium evidence	Low agreement Robust evidence	
	Evidence (type, amount, quality, consistency →			

The IPCC guidance on calibrated language for describing quantified uncertainty was used in the North Norfolk SPA Sandwich tern HRA assessment by DECC, and was applied in the Marine Scotland appropriate assessments for the Forth and Tay wind farms using their Acceptable Biological Change ("ABC") approach to interpreting PVA outputs (Table 3).

Table 2: Likelihood scale providing calibrated language for describing quantified uncertainty. From Mastrandrea *et al*. (2010).

Term	Likelihood of the outcome
Virtually certain	99-100% probability
Very likely	90-100% probability
Likely	66-100% probability
About as likely as not	33-66% probability
Unlikely	0-33% probability
Very unlikely	0-10% probability
Exceptionally unlikely	0-1% probability

These approaches have applied specific parts of the IPCC guidance to parts of impact assessment. However, no attempt has been published which describes a structured approach to communicating data and model confidence, and uncertainty and precaution, using the integrated approach of the IPCC across all aspects of environmental impact assessments.

The IPCC guidance was reviewed to determine how the treatment of uncertainty and confidence could be used in the environmental impact assessment of offshore wind farms. The advice was assessed in relation to existing guidance from the European Commission on the use of precautionary principal (European Union 2000), in particular the "structured approach to the analysis of risk". By considering the IPCC guidance as a useful means to provide a "common approach and calibrated language", a framework for the communication of uncertainty and confidence in impact assessment results was produced to aid the "analysis of risk" using a "structured approach" by decision makers.

Here an integrated approach to transparently identify, record and communicate confidence and uncertainty in environmental impact assessments is described. By identifying and communicating where key uncertainties lie, where confidence is low and agreement between stakeholders is poor, a risk-based approach to decision making can be taken. Once applied, this framework can also be used to determine cost-effective application of post-construction monitoring resources and to aspects of risk assessment for due diligence studies.

Using the advice from the IPCC, we created a structured approach to the qualitative assessment and recording of data and model confidence across a range of important data criteria: type, amount, quality, consistency and agreement (Table 3). Criteria for each data value within each data dimension that the IPCC recommends are provided, though these could be adapted to differing circumstances.

Table 3: Criteria for assessing the value of different data dimensions (adapted from Mastrandrea *et al.* 2010).

Dimensions	Criteria	Value
Type of evidence	Qualitative data	Limited
	Semi-quantitative data	Medium
	Quantitative data	Robust
Amount of evidence	Small sample size	Limited
	Medium sample size	Medium
	Large sample size	Robust
Quality of evidence	Apply expert opinion and record reasoning	Limited
		Medium
		Robust
Consistency of evidence	Few studies agree	Limited
	Most studies agree	Medium
	All studies agree	Robust
Agreement	Few parties agree	Low
	Majority of parties agree	Medium
	Most, but not all, parties agree	High

Input data are assessed for their quality, consistency, amount and type using the criteria described in Table 3. The results of the values assigned to each data dimension is then used to determine an overall evidence summary term (limited, medium, or robust). It is important that authors provide a traceable account of this evaluation of summary evidence. Where this overall evidence value is determined to be “limited” we recommend that precaution is applied, until the agreement term is valued as “medium” or “high”.

While this assessment is underway, the relevant stakeholders are consulted on the data values and data sources. Agreement is sought on data use, and the agreement is recorded. Where overall agreement is “low”, precaution is added to the datum value in order to create an agreement level of “medium” or “high”. A worked example of this process is provided in Appendix C.

Where probabilities require communication, we recommend the calibrated language approach of the IPCC in assessing risks to the environment (Table 3). Thus, a probability datum would be described as having a defined likelihood and overall confidence in it.

8.1.1 Conclusions

Better recording and assessment of data representativeness, model confidence, uncertainty and stakeholder agreement applied throughout the impact assessment phase has benefits for decision making. Doing so provides a clear outline of where and why precaution was added to assessments. By combining consideration of confidence and uncertainty, a risk-based approach to decision making can be undertaken. In addition, the transparency of this approach can be used to target limited resources to post-construction monitoring identified as having low confidence or low stakeholder agreement. The approach here can be applied both during the assessment process itself, but also to cumulative and in-combination assessments.

9 Recommendations

Recommendations for proposed methodologies for better quantification and reduction of uncertainty in ornithological offshore wind farm assessments.

Here, we lay out a set of recommendations for the additional work required (both empirical data collection and the use of modern analytical methods to exploit information in existing data) required for achieving a full quantification of uncertainty in the ecological processes and behaviours determining outcomes of interactions between seabirds and ORDs, and in reducing this uncertainty (Table 4). We focus on a full assessment of uncertainty in all mechanisms underpinning these interactions, moving beyond the methods and tools that are in current use. For each proposed recommendation, we provide a qualitative assessment of its contribution to the full quantification of uncertainty, and to reducing uncertainty (high, medium or low), and an initial estimate of the resourcing (time and funding) required for delivery (Table 4). It is important to stress that full quantification of uncertainty is as important as the reduction of uncertainty in supporting the decision-making process. This is because apparent reductions in uncertainty that arise in the context of an inadequate quantification of uncertainty are liable to create a false sense of certainty, and so increase the risk of unanticipated outcomes. The reduction of uncertainty can only meaningfully be prioritised and evaluated within the context of a comprehensive quantification of uncertainty, hence why much of the focus of the recommendations is on the quantification of uncertainty as well as the reduction of this uncertainty.

Table 4: Summary of research priorities for better estimating and reducing uncertainty in ornithological offshore wind farm assessments, moving beyond current tools and methodologies. Priorities are ranked into low, medium and high contributions to a) full quantification of uncertainty, and b) reduction in uncertainty. Potential broad methodologies are proposed for addressing research priorities, with a specification of the form (desk-based and/or field study – field implies both field and analytical components are required) and likely size of resourcing required for delivery ('small': <£30k and <6 months (i.e. "quick wins"); 'medium': £30-100k, and 6-18 months; 'large': £100-250k, and 12 months plus; 'very large': over £250k and over multiple years).

Ecological Process & relevant stage of assessment	**Contribution to a full quantification of uncertainty**	**Contribution to reducing uncertainty**	**Methods**
Predator-prey interactions, relationship between prey density and prey availability, impacts of ORDs on prey distributions and availability ***Displacement & Collision, Density & Apportioning***	High	High	Collate existing data on distribution of key prey (sandeels, sprats) from MSS and other surveys, and conduct a spatio-temporal analysis to predict prey distribution in space and time in relation to environmental characteristics (Desk-based, medium)
			Collect new empirical data on the joint distribution of both prey and seabirds to develop spatio-temporal models for prey availability in relation to environmental characteristics (Field & desk-based, large)
			Collect new empirical data on prey distributions before, during and after ORD construction to estimate changes in prey distribution as a result of ORD (Field, very large)
			Collect GPS tracking data over multiple years during and post construction to assess permeability of barrier effects/avoidance and habituation over time (Field, very large)
Estimate link between displacement effects and changes in demographic rates (productivity and survival)	High	High	Design and conduct GPS tracking of individuals from different breeding colonies before, during and after ORD construction, and link to changes in body condition, breeding success and survival to estimate displacement rates and impacts of displacement effects upon productivity and survival, ideally in relation to environmental variation and ORD characteristics (needs a strategic approach) (Field & desk-based, very large)

Ecological Process & relevant stage of assessment	Contribution to a full quantification of uncertainty	Contribution to reducing uncertainty	Methods
Displacement, Density & Apportioning			Collect empirical data and apply statistical models to estimate and quantify uncertainty in the relationship between end of season condition and subsequent overwinter survival (Field & desk-based, very large)
			Conduct power analyses to estimate sample sizes needed to detect displacement effects through the use of at-sea distribution data and GPS tracking data of individuals (desk-based, small)
			Apply habitat and resource modelling methods to estimate loss/change in habitat during/after construction and its impact on demographic rates (desk-based, large)
Better understanding and quantification of the year-round impacts of displacement ***Displacement, Density & Apportioning***	Medium	High	Collect GPS tracking data for large gulls with year-round coverage, and a fine spatial and temporal resolution; analyse to estimate seasonal variation in behaviour and distribution to allow for assessment of year-round interactions with ORDs (field & desk-based, very large)
			Collect geolocator data for species across multiple breeding colonies to assess non-breeding season habitat use, behaviour and time-activity patterns (field & desk-based, large)
			Develop IBMs for periods outside of the chick-rearing period, including prey-laying, incubation, post-fledging and the non-breeding season (desk-based, medium)
Effects of displacement on different age classes, *e.g.* immatures and non-breeders ***Displacement***	Medium	Medium	Collect empirical data to estimate differential effects/sensitivity of different age classes and states (*e.g*. breeding/sabbatical/failed breeder) to displacement (field, very large)

Ecological Process & relevant stage of assessment	Contribution to a full quantification of uncertainty	Contribution to reducing uncertainty	Methods
Improve uncertainty quantification within IBMs ***Displacement & collision***	Medium	Medium	Use emulation methods to fully quantify the uncertainty involved in estimating key parameters by matching key outputs to empirical data on demographic rates, time-activity budgets, and mass change (desk-based, medium) Reduce model structural uncertainty by collecting empirical data to understand behavioural processes, energetics and consequences for fitness (field, very large)
Assess sensitivity of collision risk model outputs to variation in input and structural parameters; understand and quantify covariance between parameters used in collision risk models ***Collision***	Medium	Medium	Perform sensitivity analysis on collision risk models (desk-based, small)
Improve quantification of flight speed and flight height for species, quantify influence of environmental conditions, and quantify how variation in these parameters is related to behaviour (commuting versus foraging) ***Collision***	Medium	Medium	Collect GPS tracking data from multiple colonies, utilise behavioural classification models for GPS tracking data and link to environmental covariates (field, large) Develop and fit 3D models for GPS tracking data around individual turbines (field, large)
Improve quantification of avoidance rates, split into micro-, meso- and macro-avoidance, and quantify	High	High	GPS tracking of individuals from multiple colonies combined with turbine-mounted monitoring of meso and micro-avoidance (field, very large)

Ecological Process & relevant stage of assessment	**Contribution to a full quantification of uncertainty**	**Contribution to reducing uncertainty**	**Methods**
influence of environmental conditions upon avoidance rates ***Collision***			
Improve estimates for productivity, adult and immature survival, and inter-colony movements (including uncertainty in rates) ***PVA***	High	High	Empirical collection of survival data (adult and juvenile) from multiple colonies and years (field, large)
			Empirical collection of dispersal/immigration/emigration data across colony networks (field, large)
			New methods for statistical modelling to estimate survival rates using historical population abundance and productivity data (desk-based, medium)
			Application of statistical metapopulation models to colony networks (desk-based, large)
			Empirical data on permanently marked individuals to estimate net movements among sub-populations (field, very large)
Empirical estimation of correlation between environmental stochasticity in demographic rates & Improved models of observation error for abundance	Medium	High	Analyse historical time series of population or individual level survival and productivity data (desk-based, medium)
			Empirical data on repeated counts to quantify uncertainty in population size, and analysis of plot vs whole colony counts to estimate relationships and adjust monitoring if plots not representative (desk-based and field, medium)

Ecological Process & relevant stage of assessment	Contribution to a full quantification of uncertainty	Contribution to reducing uncertainty	Methods
PVA			Quantify observation error in historical survey methods with new methods (desk-based, medium)
Quantify relationship between demographic rates and prey availability, climate and other environmental variables to include in population forecasts ***PVA***	High	High	Improve measurement and use of covariates in models - critical for capturing and understanding variability leading to better predictive power and reduced uncertainty (desk-based, medium) Statistical analyses of historical data to estimate linkages between environmental variables and demographic rates (desk-based, medium)
Data integration and model fitting methods Sensitivity analyses for PVAs ***PVA***	Medium	Medium	Rationalising disagreement between "known" trajectory for a population and parameters borrowed from other populations – apply statistical methods to integrate understanding of population trajectory prior to impact within models (desk-based, medium) Sensitivity analysis to identify which parameters and inputs need to be known with the most precision in relation to output metrics used in ORD assessments to guide future research (desk-based, small)
Better understanding and quantification of density dependent processes in populations ***PVA***	Medium	Medium	Application of statistical models to estimate density dependence in demographic rates to historical data (desk-based, medium)

References

Aarts, G., MacKenzie, M., McConnell, B., Fedak, M. & Matthiopoulos, J. 2008. Estimating space-use and habitat preference from wildlife telemetry data. *Ecography*, **31(1)**, 140–160. https://doi.org/10.1111/j.2007.0906-7590.05236.x

Azzalini, A., 1996. *Statistical Inference Based on the likelihood*. Chapman & Hall, London.

Beaumont, M.A., Zhang, W. & Balding, D.J. 2002. Approximate Bayesian Computation in Population Genetics. *Genetics*, **162 (4)**: 2025–2035.

Berlincourt, M., Angel, L.P. & Arnould, J. P. Y. 2015. Combined use of GPS and accelerometery reveals fine scale three-dimensional foraging behaviour in the short-tailed shearwater. *PLoS ONE*, **10(10)**, 1–16. https://doi.org/10.1371/journal.pone.0139351

Beissinger, S.R. & McCullough, D.R. (eds). *Population Viability Analysis.* Chicago, IL: University of Chicago Press, 2002.

Blackwell, P.G. 2019. Integrated Continuous-time Hidden Markov Models. *ArXiv*, **1–18**. arXiv:1807.11907v3

Busch, M., Kannen, A., Garthe, S. & Jessopp, M. 2013. Consequences of a cumulative perspective on marine environmental impacts: Offshore wind farming and seabirds at North Sea scale in context of the EU Marine Strategy Framework Directive. Ocean & Coastal Management, Volume 71, Pages 213-224

Butler, A., Searle, K., Mobbs, D.C. & Daunt, F. 2020a. A Population Viability Analysis Modelling Tool for Seabird Species – Tool Testing: Report for methodology and results for testing of tool. JNCC Report No. 657, JNCC, Peterborough, ISSN 0963-8091. https://hub.jncc.gov.uk/assets/302a7a51-fe29-4633-95d1-b3ef458cb79a

Butler, A., Carroll, M., Searle, K., Bolton, M., Waggitt, J., Evans, P., Rehfisch, M., Goddard, B., Brewer, M., Burthe, S. & Daunt, F. 2020b. Attributing seabirds at sea to appropriate breeding colonies and populations (CR/2015/18). Scottish Marine and Freshwater Science Vol 11 No 8, 140pp.

Cleasby, I.R., Wakefield, E.D., Bearhop, S., Bodey, T.W., Votier, S.C. & Hamer, K.C. 2015. Three-dimensional tracking of a wide-ranging marine predator: Flight heights and vulnerability to offshore wind farms. *Journal of Applied Ecology*, **52(6)**, 1474–1482. https://doi.org/10.1111/1365-2664.12529

Cook, A.S.C.P. & Robinson, R.A. (2016) Testing sensitivity of metrics of seabird population response to offshore wind farm effects, JNCC Report No. 553, JNCC, Peterborough, ISSN 0963-8091.
Cook, A.S.C.P., Humphreys, E.M., Robinson, R.A. & Burton, N.H.K. 2019. Review of the potential of seabird colony monitoring to inform monitoring programmes for consented offshore wind farm projects. Research Report no. 712. British Trust for Ornithology, Thetford, Norfolk ISBN: 978-1-912642-07-6 163pp

Daunt, F., Fang, Z., Howells, R., Harris, M., Wanless, S., Searle K.R. & Elston, D. 2020. Improving estimates of seabird body-mass survival rates. *Scottish Marine and Freshwater Science* **Vol 11 No 13**. DOI: 10.7489/12329-1

Dierschke, V., Furness, R.W. & Garthe, S. (2016, October). Seabirds and offshore wind farms in European waters: Avoidance and attraction. Biological Conservation. Elsevier Ltd. https://doi.org/10.1016/j.biocon.2016.08.016

Donovan, C.R., Harris, C.M., Milazzo, L., Harwood, J., Marshall, M. & Williams, R. 2017. A simulation approach to assessing environmental risk of sound exposure to marine mammals. *Ecology and Evolution*, **7(7)**, 2101-2111.

Donovan, C.R. & Caneco, B.A.R. (2020) Seabird Survey Designs for the East Coast of Scotland. Scottish Marine and Freshwater Science Vol 11 No 19, 78pp.

Drewitt, A.L. & Langston, R.H.W. 2006. Assessing the impacts of wind farms on birds. IBIS 148, 29–42.

EFSA (European Food Safety Authority). 2014. Guidance on Expert Knowledge Elicitation in Food and Feed Safety Risk Assessment. *EFSA Journal 2014;* **12(6)**:3734. [278 pp.] doi:10.2903/j.efsa.2014.3734

Furness, R.W. 2015. Non-breeding season populations of seabirds in UK waters: Population sizes for Biologically Defined Minimum Population Scales (BDMPS). Natural England Commissioned Reports, Number 164.

European Union. 2000. Communication from The Commission on the precautionary principle. European Union, Brussels. [http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2000:0001:FIN:EN:PDF accessed 08 June 2017]

Freeman, S., Searle, K., Bogdanova, M., Wanless, S. & Daunt, F. 2014. Population dynamics of Forth & Tay breeding seabirds: review of available models and modelling of key breeding populations. Ref MSQ-0006. Final report to Marine Scotland Science

Furness, R.W. 2015. Non-breeding season populations of seabirds in UK waters. Population sizes for Biologically Defined Minimum Population Scales (BDMPS). Natural England Commissioned Reports, Number 164.

Garthe, S., Markones, N. & Corman, A.-M. (2017). Possible impacts of offshore wind farms on seabirds: a pilot study in Northern Gannets in the southern North Sea. Journal of Ornithology, 158(1), 345–349. https://doi.org/10.1007/s10336-016-1402-y

Hammond, P.S., Lacey, C., Gilles, A., Viquerat, S., Börjesson, P., Herr, H. & Øien, N. 2018. Estimates of cetacean abundance in European Atlantic waters in summer 2016 from the SCANS-III aerial and shipboard surveys. https://doi.org/10.1016/j.ijppaw.2016.02.001

Hammond, P.S., Macleod, K., Berggren, P., Borchers, D.L., Burt, L., Cañadas, A. & Vázquez, J.A. 2013. Cetacean abundance and distribution in European Atlantic shelf waters to inform conservation and management. Biological Conservation, 164, 107–122. https://doi.org/10.1016/j.biocon.2013.04.010

Hammond, P.S., Berggren, P., Benke, H., Borchers, D.L., Collet, A., Heide-Jørgensen, M.P. & Oien, N. 2002. Abundance of harbour porpoise and other cetaceans in the North Sea and adjacent waters. Journal of Applied Ecology, 39(2), 361–376. https://doi.org/10.1046/j.1365-2664.2002.00713.x

Heinänen, S., Žydelis, R., Kleinschmidt, B., Dorsch, M., Burger, C., Morkūnas, J., Quillfeldt, P. & G. Nehls. 2020. Satellite telemetry and digital aerial surveys show strong displacement of red-throated divers (*Gavia stellata*) from offshore wind farms. Marine Environmental Research, Volume 160, 104989.

Hobbs, N.T. & Hooten, M.B. 2015. Bayesian Models: A statistical primer for ecologists. Princeton University Press.

Isaksson, N., Cleasby, I.R., Owen, E., Williamson, B.J., Houghton, J.D.R., Wilson, J. & Masden, E.A. The Use of Animal-Borne Biologging and Telemetry Data to Quantify Spatial Overlap of Wildlife with Marine Renewables. J. Mar. Sci. Eng. 2021, 9, 263.

Jitlal, M., Burthe, S., Freeman, S. & Daunt, F. 2017. Testing and Validating Metrics of Change Produced by Population Viability Analysis (PVA). Scottish Marine and Freshwater Science Vol 8 No 23, 210pp. DOI: 10.7489/2018-1

Kennedy, M. & O'Hagan, A. 2001. Bayesian calibration of computer models (with discussion). J. R. Statist. Soc. Ser. B, 63, 425–464

Langrock, R., King, R., Matthiopoulos, J., Thomas, L., Fortin, D. & Morales, J. M. 2012. Flexible and practical modelling of animal telemetry data: Hidden Markov models and extensions. Ecology, 93, 2336–2342. https://doi.org/10.1890/11-2241.1

Leslie, P.H. 1945. The use of matrices in certain population mathematics. Biometrika, 33(3), 183–212.

Lindgren, F., Rue, H., & Lindström, J. 2011. An explicit link between gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach. Journal of the Royal Statistical Society. Series B: Statistical Methodology, **73(4)**, 423–498. https://doi.org/10.1111/j.1467-9868.2011.00777.x

Mackenzie, M.L, Scott-Hayward, L.A.S., Oedekoven, C.S., Skov, H., Humphreys, E. & Rexstad, E. (2013). Statistical Modelling of Seabird and Cetacean data: Guidance Document. University of St. Andrews contract for Marine Scotland; SB9 (CR/2012/05).

Maclean, I.M.D., Rehfisch, M.M., Skov, H. & Thaxter, C.B. (2013). Evaluating the statistical power of detecting changes in the abundance of seabirds at sea. Ibis, 155(1), 113–126. https://doi.org/10.1111/j.1474-919X.2012.01272.x

Marine Management Organisation. (2014). Review of post-consent offshore wind farm monitoring data associated with licence conditions.

Marjoram, P., Molitor, J., Plagnol, V. & Tavare, S. 2003. Markov chain Monte Carlo without likelihoods. Proc Natl Acad Sci U S A. 100 (26): 15324–15328. https://doi:10.1073/pnas.0306899100

Masden, E. (2015) Developing an avian collision risk model to incorporate variability and uncertainty. Scottish Marine and Freshwater Science Vol 6 No 14. Edinburgh: Scottish Government, 43pp. DOI: 10.7489/1659-1

Masden, E., McCluskie, A., Owen, E. & Langston, R. 2015. Renewable energy developments in an uncertain world: The case of offshore wind and birds in the UK. Marine Policy. 51. 169–172. 10.1016/j.marpol.2014.08.006.

Mastrandrea, M.D., Field, C.B., Stocker, T.F., Edenhofer, O., Ebi, K.L., Frame, D.J., Held, H., Kriegler, E., Mach, K.J., Matschoss, P.R., Plattner, G.-K., Yohe, G.W. & Zwiers, F.W. 2010: Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent

Treatment of Uncertainties. Intergovernmental Panel on Climate Change (IPCC). Available at http://www.ipcc.ch

Matthiopoulos, J. & Trinder, M. 2020. Study to develop best practice recommendations for combining seabird study data collected from different platforms (ITQ-0247).

McClintock, B.T. & Michelot, T. 2018. momentuHMM: R package for generalized hidden Markov models of animal movement. Methods in Ecology and Evolution, 9(6), 1518–1530. https://doi.org/10.1111/2041-210X.12995

Mendel, B., Schwemmer, P., Peschko, V., Müller, S., Schwemmer, H., Mercker, M. & Garthe, S. (2019). Operational offshore wind farms and associated ship traffic cause profound changes in distribution patterns of Loons (*Gavia* spp.). Journal of Environmental Management, 231, 429–438. https://doi.org/10.1016/J.JENVMAN.2018.10.053

Merkel, B., Phillips, R.A., Descamps, S. *et al*. A probabilistic algorithm to process geolocation data. Mov Ecol 4, 26 2016. https://doi.org/10.1186/s40462-016-0091-8

Michelot, T., Langrock, R. & Patterson, T.A. 2016. moveHMM: An R package for the statistical modelling of animal movement data using hidden Markov models. Methods in Ecology and Evolution, 7, 1308–1315. https://doi.org/10.1111/2041-210X.12578

Oyebamiji, O.K., Wilkinson, D.J., Jayathilake, P.G., Curtis, T.P., Rushton, S.P., Li, B. & Guptad, P. 2017. Gaussian process emulation of an individual-based model simulation of microbial communities. Journal of Computational Science, 22, 69-84.

Parton, A. & Blackwell, P.G. 2017. Bayesian Inference for Multistate ‘Step and Turn’ Animal Movement in Continuous Time. Journal of Agricultural, Biological, and Environmental Statistics, 22(3), 373–392. https://doi.org/10.1007/s13253-017-0286-5

Patterson, T.A., Thomas, L., Wilcox, C., Ovaskainen, O. & Matthiopoulos, J. 2008. State–space models of individual animal movement. Trends in Ecology and Evolution, 23(2), 87–94. https://doi.org/10.1016/j.tree.2007.10.009

Patrick, S.C., Bearhop, S., Grémillet, D., Lescroël, A., Grecian, W.J., Bodey, T.W. & Votier, S.C. 2014. Individual differences in searching behaviour and spatial foraging consistency in a central place marine predator. Oikos, 123(1), 33–40. https://doi.org/10.1111/j.1600-0706.2013.00406.x

Peel, A., Jenks, M., Choudhury, M., Lovett, R., Rejon-Parrilla, J.C., Sims, A. & Craig, J. 2018. Use of Expert Judgement Across NICE Guidance-Making Programmes: A Review of Current Processes and Suitability of Existing Tools to Support the Use of Expert Elicitation. Applied Health Economics and Health Policy, 16(6), 819-836. https://doi.org/10.1007/s40258-018-0415-5

Peschko, V., Mercker, M. & Garthe, S. 2020a. Telemetry reveals strong effects of offshore wind farms on behaviour and habitat use of common guillemots (*Uria aalge*) during the breeding season. Marine Biology, 167(8), 1–13. https://doi.org/10.1007/s00227-020-03735-5

Peschko, V., Mendel, B., Müller, S., Markones, N., Mercker, M. & Garthe, S. 2020b. Effects of offshore windfarms on seabird abundance: Strong effects in spring and in the breeding season. Marine Environmental Research, Volume 162, 105157.

Peschko, V., Mendel, B., Mercker, M., Dierschke, J. & Garthe, S. 2021. Northern gannets (*Morus bassanus*) are strongly affected by operating offshore wind farms during the breeding season. Journal of Environmental Management, 279, 111509. https://doi.org/10.1016/j.jenvman.2020.111509

Regular, P.M., Hedd, A., Montevecchi, W.A., Robertson, G.J., Storey, A.E. & Walsh, C.J. (2014). Why timing is everything: Energetic costs and reproductive consequences of resource mismatch for a chick-rearing seabird. Ecosphere, 5(12), art155. https://doi.org/10.1890/ES14-00182.1

RSPB. (2019) Application by Norfolk Vanguard Limited for an Order Granting Development Consent for the Norfolk Vanguard Offshore Wind Farm'; RSPB response to the Applicant's Deadline 8 submission 'Precaution in ornithological assessment for offshore wind farms' (2019).

Rue, H., Martino, S. & Chopin, N. 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. Journal of the Royal Statistical Society. Series B: Statistical Methodology, 71(2), 319–392. https://doi.org/10.1111/j.1467-9868.2008.00700.x

Rush, G.P., Clarke, L.E., Stone, M. & Wood, M.J. 2018. Can drones count gulls? Minimal disturbance and semiautomated image processing with an unmanned aerial vehicle for colony-nesting seabirds. Ecology and Evolution, 8(24), 12322–12334. https://doi.org/10.1002/ece3.4495

Sansom, A., Wilson, L.J., Caldow, R.W.G. & Bolton, M. 2018. Comparing marine distribution maps for seabirds during the breeding season derived from different survey and analysis methods. PLoS ONE, 13(8), 1–17. https://doi.org/10.1371/journal.pone.0201797

Scott-Hayward, L., Oedekoven, C., Mackenzie, M. & Rexstad, E. 2013. Using the MRSea Package - Statistical modelling of bird and cetacean distributions in offshore renewables development areas. Contract for Marine Scotland (SB9:CR/2012/05), 9.

Scott-Hayward, L.A.S., Mackenzie, M.L., Donovan, C.R., Walker, C.G. & Ashe, E. 2014. Complex region spatial smoother (CReSS). Journal of Computational and Graphical Statistics, 23(2), 340–360. https://doi.org/10.1080/10618600.2012.762920

Scottish Natural Heritage. 2014; 2018. Interim Guidance of Apportioning Impacts from Marine Renewable Developments to Breeding Seabird Populations in Special Protection Areas.

Schwarz, G. 1978. Estimating the dimension of a model. The Annals of Statistics, 6(2), 461–464. https://doi.org/10.1214/aos/1176344136

Searle, K., Mobbs, D., Butler, A., Bogdanova, M., Freeman, S., Wanless, S. & Daunt, F. (2014) Population consequences of displacement from proposed offshore wind energy developments for seabirds breeding at Scottish SPAs (CR/2012/03). Report to Scottish Government.

Searle, K.R., Mobbs, D.C., Butler, A., Furness, R.W., Trinder, M.N. & Daunt, F. 2018. Finding out the Fate of Displaced Birds. Scottish Marine and Freshwater Science Vol 9 No 8. DOI: 10.7489/12118-1

Searle, K.R., Mobbs, D., Daunt, F. & Butler, A. 2019. A Population Viability Analysis Modelling Tool for Seabird Species. Natural England Commissioned Report NECR274. ISBN 978-1-78354-531-5

Searle, K.R.; Butler, A.; Mobbs, D.; Trinder, M.; Waggitt, J.; Evans, P.; Daunt, F. 2020. Scottish Waters East Region Regional Sectoral Marine Plan Strategic Ornithology Study: final report (Report No. NEC07184). Report by Natural Environment Research Council (NERC). Report for Marine Scotland Science. https://www.gov.scot/publications/scottish-waters-east-region-regional-sectoral-marine-plan-strategic-ornithology-study-final-report/

Searle, K.R., Butler, A., Mobbs, D., Trinder, M., McGregor, R., Cook, A., McCluskie, A., Caneco, B, & Daunt, F. (in press). Study to examine how seabird collision risk, displacement and barrier effects could be integrated for assessment of offshore wind developments. Report for Marine Science Scotland.

Sisson, S.A., Fan, Y. & Tanaka, M.M. 2007. Sequential Monte Carlo without likelihoods. Proc Natl Acad Sci U S A. 104 (6), 1760–1765. https://doi:10.1073/pnas.0607208104

Skov, H., Heinänen, S., Norman, T., Ward, R., Méndez-Roldán, S. & Ellis, I. 2018. ORJIP Bird Collision and Avoidance Study. Final report – April 2018. The Carbon Trust. United Kingdom.

Soulé, M.E. (ed.) 1986. Conservation biology. The science of scarcity and diversity. Sinauer Associates Inc., Sunderland (Mass.). 584 pages. ISBN 0-87893-795-1

Suryan, R.M., Irons, D.B., Brown, E.D., Jodice, P.G.R. & Roby, D.D. (2006). Site-specific effects on productivity of an upper trophic-level marine predator: Bottom-up, top-down, and mismatch effects on reproduction in a colonial seabird. Progress in Oceanography, 68(2–4), 303–328. https://doi.org/10.1016/j.pocean.2006.02.006

Thaxter, C.B., Ross-Smith, V.H., Bouten, W., Clark, N.A., Conway, G.J., Rehfisch, M.M. & Burton, N.H.K. 2015. Seabird–wind farm interactions during the breeding season vary within and between years: A case study of lesser black-backed gull *Larus fuscus* in the UK. Biological Conservation, Volume 186, Pages 347-358

Thaxter, C.B., Ross-Smith, V.H., Bouten, W., Masden, E.A., Clark, N.A., Conway, G. & Burton, N.H.K. (2018). Dodging the blades: new insights into three-dimensional area use of offshore wind farms by lesser black-backed gulls *Larus fuscus*. Marine Ecology Progress Series, 587, 247–253. https://doi.org/10.3354/meps12415

Thompson, P.M., Hastie, G.D., Nedwell, J., Barham, R., Brookes, K.L., Cordes, L.S., Bailey, H. & McLean, N., 2013. Framework for assessing impacts of pile-driving noise from offshore wind farm construction on a harbour seal population. Environmental Impact Assessment Review, 43, pp.73-85.

Vanermen, N., Onkelinx, T., Courtens, W., Van de walle, M., Verstraete, H. & Stienen, E.W.M. (2015). Seabird avoidance and attraction at an offshore wind farm in the Belgian part of the North Sea. Hydrobiologia, 756(1), 51–61. https://doi.org/10.1007/s10750-014-2088-x

Vanermen, N., Onkelinx, T., Verschelde, P., Courtens, W., Van de walle, M., Verstraete, H. & Stienen, E.W.M. (2015). Assessing seabird displacement at offshore wind farms: power ranges of a monitoring and data handling protocol. Hydrobiologia, 756(1), 155–167. https://doi.org/10.1007/s10750-014-2156-2

Wakefield, E.D., Owen, E., Baer, J., Carroll, M.J., Daunt, F., Dodd, S.G., … Bolton, M. 2017. Breeding density, fine- scale tracking, and large-scale modelling reveal the regional distribution of four seabird species: Ecological Applications, **27(7)**, 2074–2091.

Wakefield, E.D., Bodey, T.W., Bearhop, S., Blackburn, J., Colhoun, K., Davies, R., … Hamer, K.C. 2013. Space partitioning without territoriality in gannets. Science, 141–146. https://doi.org/10.1126/science.1236077

Walker, C.G., Mackenzie, M.L., Donovan, C.R. & O'Sullivan, M.J. 2011. SALSA - a spatially adaptive local smoothing algorithm. Journal of Statistical Computation and Simulation, 81(2), 179–191. https://doi.org/10.1080/00949650903229041

Warwick-Evans, V., Atkinson, P.W., Walkington, I. & Green, J.A. 2018. Predicting the impacts of wind farms on seabirds: An individual-based model. *Journal of Applied Ecology*, **55(2)**, 503–515. https://doi.org/10.1111/1365-2664.12996

Welcker, J. & Nehls, G. 2016. Displacement of seabirds by an offshore wind farm in the North Sea. Mar Ecol Prog Ser 554:173-182. https://doi.org/10.3354/meps11812

Woodward, I., Thaxter, C.B., Owen, E. & Cook, A.S.C.P. 2019. *Desk-based revision of seabird foraging ranges used for HRA screening.* BTO research report number 724

Yuan, Y., Bachl, F.E., Lindgren, F., Borchers, D.L., Illian, J.B., Buckland, S.T., … Gerrodette, T. 2017. Point process models for spatio-temporal distance sampling data from a large-scale survey of blue whales. Annals of Applied Statistics, **11(4)**, 2270–2297. https://doi.org/10.1214/17-AOAS1078

Appendix A: Workshop summary

December 16th 2020 – online workshop

This project is aimed at producing a set of clear recommendations for how scientific uncertainty can be quantified and reduced throughout the ornithological offshore wind assessment process, leading to more precise ORD impact estimates. It will highlight where new empirical data and research are required in order to reduce uncertainties. The specific aims of the workshop were to:

1. review the methods that are currently used to quantify uncertainty within the assessment process, and evaluate the ways in which these uncertainty estimates are currently used within the assessment process
2. highlight key areas in which the quantification and interpretation of uncertainty could be improved, either through statistical modelling, additional data collection or adaptation of the assessment process
3. provide a framework for the end-to-end quantification of uncertainty, which brings together estimates of uncertainty associated with individual stages of the assessment process
4. develop recommendations for the research required to both better quantify uncertainty, and to reduce it, in order to reduce consenting risk and increase headroom for future offshore wind development through more certainty about likely impacts of planned developments.

Therefore, each of the workshop discussion sessions selected a stage of the assessment process (see agenda below), and was structured as follows (Table 5):

i. Delivery of a brief presentation by a member of the project team on how uncertainty is currently incorporated into assessments – this was comprised of identifying which sources of uncertainty are currently quantified, and how they are currently used within the assessment process
ii. Presentation of all relevant sources of uncertainty (being addressed in the project) that should be included in the assessment stage, including those that are currently excluded
iii. Open discussions (using Jamboard) to identify where we can quantify uncertainty better, the priorities for quantifying uncertainty better, methods for reducing uncertainty, and prioritisation of methods for reducing uncertainty

Table 5: Summary of presentations and discussion for each of the workshop discussion sessions, centred on a single stage of the assessment process.

How is uncertainty currently quantified? Which sources of uncertainty are quantified? How are these sources of uncertainty currently used within assessments?	Short presentation (couple of slides) by project team, followed by short open session on any comments/queries
Can we quantify uncertainty better?	Jamboard – project team will propose methods, followed by open discussion
Priorities for quantifying uncertainty	Jamboard – project team will propose priorities, followed by open discussion
What can we do to reduce uncertainty?	Jamboard – project team will propose methods, followed by open discussion
Priorities for how to reduce uncertainty	Jamboard – project team will propose priorities, Jamboard by open discussion

There were four main discussion sessions around the main stages of assessment:

1. PVA, Cumulative impacts and uncertainty propagation, end to end quantification of uncertainty
2. Displacement modelling
3. Collision modelling
4. Density data and apportioning methods

Finally, there was one further discussion session on the more general topic of:

5. Integration of post-consent monitoring data (specifically, how post-consent monitoring data may be used to quantify and reduce uncertainty in the assessment process)

The final discussion session of the day proposed and refined a high-level set of recommendations for research to better quantify and reduce uncertainty throughout the assessment process.

Workshop attendees

Kate Searle, Francis Daunt (UKCEH)
Adam Butler, Esther Jones (BioSS)
Mark Trinder, Ross McGregor (MacArthur Green)
Aonghais Cook, Liz Humphries (BTO)
Aly McCluskie (RSPB)
Carl Donovan (DMP Statistics)
Elizabeth Masden (UHI)
Sue O'Brien, Julie Black (JNCC)
Janelle Braithwaite, Tom Evans, Julie Miller (Marine Scotland)
Matty Murphy, Alex Scorey, Mike Bailey (NRW)
Glen Tyler (NatureScot)
Clare McNamara (Daera)

General recommendations

- Properly communicating assumptions of models to decision makers

Population Viability Analysis recommendations

Key focus for quantifying and reducing uncertainty:

- Better understanding and quantification of density dependence in population processes
- Better understanding and quantification of correlation between demographic rates
- Better understanding and quantification of metapopulation dynamics and assessment of populations at ecologically relevant scales
- Application of methods that incorporate all available data, including population trend prior to impact, solid quantification of baseline population
- Inclusion of climate change impacts within population models
- Better quantification and understanding of demographic rates for immatures, and the processes affecting recruitment and age structure in populations
- Importance of data collection at appropriate scales - both spatially (*e.g.* distributions of birds at sea) and temporally (*e.g.* across years and across the annual cycle)

Table 6: Summary of main processes, data collection and analytical methods for quantifying and reducing uncertainty in Population Viability Analysis (PVA).

Process	Empirical Data	Analyses/tools/models
Improve estimates for productivity, adult and immature survival		Address lack of survival /dispersal data compared to other demographic parameters time series of population or individual level survival and productivity data Potential for different effects/sensitivity of different age classes and states (*e.g.* breeding/sabbatical/failed breeder) - methods to quantify and consider this
Understand and estimate carry-over effects on productivity and survival		time series of population or individual level survival and productivity data
Empirical estimation of correlation between environmental stochasticity in demographic rates		Including rate correlation/covariance could have large effect in reducing uncertainty and data available? time series of population or individual level survival and productivity data
Improved models of observation error for abundance	Repeat counts to quantify uncertainty in population size Plot vs whole colony counts; adjust monitoring if plots not representative	Quantify error in historical survey methods with new methods
Quantify and include uncertainty in model parameters		time series of population or individual level survival and productivity data
Quantify and include uncertainty in population sizes used in models		
Quantify relationship between prey availability and demographic rates so changes to prey can be included in population forecasts		Improve measurement and use of covariates in models - critical for capturing and understanding variability leading to better predictive power and reduced uncertainty
Proper accounting of uncertainty in future population projections		
Data integration and model fitting methods		Rationalising disagreement between "known" trajectory for a population and parameters borrowed from other populations

Process	Empirical Data	Analyses/tools/models
		Integrate understanding of population trajectory prior to impact within models
Sensitivity analyses for PVA – which parameters/inputs need to be known with most precision?		should be conducted at scale of impact assessment: where should effort be focussed? On PVA, impact estimation methods, or source seabird data?
Quantify and include inter-annual variability in ORD effects; quantify and model habituation over time		Incorporation of seasonal information into estimates of annual OR effects Studies or models looking at magnitude and direction of environmental impacts on age and stages of seabirds
Understand interactions of ORD impacts – are different impacts/processes synergistic or antagonistic?		
Include climate change effects within PVA models	Diet data and changes in prey availability	time series of population or individual level survival and productivity data able to be linked with appropriate scaled climate data
Better understanding and quantification of variability in population dynamics around the UK		
Are interspecific interactions important in determining demographic rates for some species, and impacts of ORDs?		
Better understanding and quantification of density dependent processes in populations		Quantification of scales (spatial and temporal) at which density dependence works
Consider trade-off in model complexity versus underpinning data for parameters		
Metapopulations and dispersal/immigration/emigration	permanently marked individuals Net movements among sub-pops	Julie Miller's analysis (black-legged kittiwakes)

Displacement recommendations

Key focus for quantifying and reducing uncertainty:

- Better quantification of displacement rate
- Better quantification of displacement mortality rate
- Development of 'middle-ground' methods that are more biologically realistic than the Displacement Matrix but not as data hungry as individual based models (IBMs)
- Improved use of individual-level data (tracking, condition and fitness) for estimating displacement rates and displacement mortality rates
- Better understanding and quantification of barrier effects

Table 7: Summary of main processes, data collection and analytical methods for quantifying and reducing uncertainty in displacement impact modelling.

Process	Empirical Data	Analyses/tools/models
Predator-prey interactions, relationship between prey density and prey availability	Fish stocks – availability to seabirds and impact of ORDs on key fish species Collate all existing data on distribution of key prey (sandeels, sprats) from MSS and other surveys, and conduct an analysis to predict distribution in space and time	Importance of working with fisheries biologists and making use of their models Ecosystems models Predator-prey models
Link between displacement effects and changes in demographic rates	GPS tracking of individuals Relationship between end of season condition and subsequent overwinter survival Post-consent data: empirical estimation of displacement rates (GPS tracking and at-sea surveys) Post-consent data: empirical estimation of displacement mortality rates (GPS tracking linked to mass and survival of individuals)	Power analyses
Effect of environmental variation on displacement rate	Post-consent data Climate change	Habitat use models in relation to environmental variables Power analyses to examine statistical power of existing monitoring methods
Year-round impacts of displacement	GPS tracking for large gulls (year-round data with fine spatial and temporal resolution) Geolocator data for birds from multiple breeding colonies to	Develop IBM for non-breeding season (especially RTDs) Consider Dutch developed IBM model for non-breeding season displacement

Process	Empirical Data	Analyses/tools/models
	assess non-breeding season habitat use	
Habituation over time	Longer-term GPS tracking data of individuals	
Density-dependent effects of conspecifics on displacement mortality (competition and environmental quality)		
Better understanding of habitat loss and habitat quality, and how to represent this in models		Habitat and resource modelling methods Sundberg et al: "A mechanistic framework to inform the spatial management of conflicting fisheries and top predators." Journal of Applied Ecology (2020)
Interspecific resource competition and relationship with displacement effects		
Extent of site fidelity in foraging locations and its effect on displacement rates and mortality		
Condition/state dependence (*e.g*. breeding stage) in displacement rates/influence of personality or behavioural syndromes on displacement rates to understand individual variation		
Effects of displacement on different age classes, *e.g*. immatures and non-breeders		
Effects of displacement on different behaviours (commuting vs foraging)		
IBMs		emulation methods to fully quantify the uncertainty involved in estimating key parameters by matching key outputs to empirical data on demographic rates Reduce model structural uncertainty by collecting empirical data to understand behavioural processes, energetics and consequences for fitness

Collision recommendations

Key focus for quantifying and reducing uncertainty:

- Better quantification of correlation in parameters within the model (*e.g*. flight height and flight speed, wind speed and rotor speed)
- Better quantification of behaviour of birds in relation to weather conditions and state
- Better quantification of behaviours and variability in flight height/speed (*e.g*. commuting versus foraging flight)
- Better quantification of avoidance rate – better understanding and quantification of the processes and mechanisms that underpin avoidance rates
- Move towards models of individual flights of birds and interactions with individual turbines, and use of 3D models for flight paths
- Sensitivity analysis - what's most influential. Some obvious ones (avoidance, density, FHD) - beyond this?
    - And does this vary by species?

Table 8: Summary of main processes, data collection and analytical methods for quantifying and reducing uncertainty in collision risk models (CRM).

Process	Empirical Data	Analyses/tools/models
Flight speed/flight height – better measurement; variation in relation to environmental conditions; variation in relation to behaviour (commuting versus foraging)	GPS tracking data GPS tracking coupled with measurement of wind speed etc	Behavioural classification models for GPS tracking data Behavioural classification models linked to environmental covariates 3D models around individual turbines
Avoidance behaviour – does this change over time with experience	GPS tracking data of individuals for macro-avoidance Turbine-mounted monitoring of meso and micro-avoidance	3D flight paths from tracking data
Factors driving fine-scale variation in risk/avoidance rates	GPS tracking data of individuals for macro-avoidance Turbine-mounted monitoring of meso and micro-avoidance Finer resolution measurement of environmental variables (*e.g.* habitat quality and weather) Colony-specific variation – GPS tracking and comparison of birds from known provenance/breeding status	Models linking avoidance behaviour with environmental characteristics (*e.g.* wind speed/wind direction) or bird characteristics (*e.g*. age, breeding state, etc); placement of turbines Approach angle, should this be considered more? See Holmstrom *et al.* (2011) 3D flight paths from tracking data
Separation of micro-, meso- and macro-avoidance	GPS tracking data of individuals for macro-avoidance Turbine-mounted monitoring of meso and micro-avoidance	3D flight paths from tracking data

Process	Empirical Data	Analyses/tools/models
Understand and quantify covariance between parameters used in collision risk models		
better parameterisation of turbine rotation speed		
separation of environmental variation versus uncertainty within collision risk models		
Consider error of Band versus other models, *e.g*. Kleyheeg-Hartman and refine		
Use of post consent monitoring data: How to use to inform on CRM parameters, *e.g*. changes in flight height distribution post-construction not independent of avoidance rate		3D flight paths from tracking data
Understanding of collision during migration		
Converting densities into flux – better methods for this; relationship between flux rate and flight behaviour; Separate out how flight speed is used for pCol and flux calculations		
Validation of pColl against empirical data		
flapping rates and gliding proportion		

Density and apportioning recommendations

Key focus for quantifying and reducing uncertainty:

- Data integration methods for estimating spatial habitat use and at-sea densities
- Need for a strategic approach to data collection
- Better quantification of inter-annual variability and understanding of representativeness of survey data
- development of methods to leverage existing data to address this
- Need for more model validation, ground-truthing of model predictions
- Improved measurement and use of environmental covariates in models
- Improved measurements and use of prey data within model

Table 9: Summary of main processes, data collection and analytical methods for quantifying and reducing uncertainty in density data and apportioning methods.

Process	**Empirical Data**	**Analyses/tools/models**
Quantifying uncertainty in tracking models	GPS tracking, many individuals, many colonies	HMMs, continuous time models
Focus on appropriate spatial scale for data collection	Measure environmental variables at finer resolution to link better with GPS tracking or at-sea survey data Need to collect concurrent covariate data at the right scale for count and tracking data Collect covariates at spatial and temporal scales that match the bird data, *e.g.* (such as hydrographic features)	Strategic collection of at-sea survey data (*e.g.* government-led) Variable scales for data collection (‘fences’) and analysis (INLA SPDE, continuous time models)
Better estimation of baseline inter-annual variability	Compile multiple at-sea footprint surveys across time and space and investigate inter-annual variability in comparison to ‘snapshot’ surveys on, *e.g.* 1 day/month	Consider alternative survey designs to address temporal variation
Reducing structural uncertainty in tracking models		
Identifying behavioural states from tracking data		Ground truth HMMs using accelerometer/TDR data?? (anonymous animal X: are all states reliably identifiable from this? *e.g.* resting vs feeding) HMMs, continuous time models
Site turnover	Long-term tracking of individuals	

Process	Empirical Data	Analyses/tools/models
Direct measurements of prey, not use of proxies in models for density; consider prey availability not just abundance		
Understanding of habitat use outside of chick-rearing; habitat use by non-breeders	Need for tracking data from whole/more of breeding season, most tracking data being early chick-rearing currently	Data integration, *e.g*. GPS and at-sea
Improve upon use of distance-decay in apportioning methods	Detailed colony-level modelling with more covariates	For gannets use Wakefield et al 2013 to develop MS apportioning method for this species?
Better definition of biological seasons - dealing with overlapping months when migrants, breeders and non-breeders may be present		
Quantify statistical power of survey data to quantify abundance/density and changes		Need defined acceptable/ recommended alpha, beta and effect size levels Any power analysis needs to be tailored to the statistical test to be used, and considering the normality assumption, independence, etc.
How to defensibly quantify exposure (density essentially) for less common species, *e.g*. great black-backed gull		
Influence of climate change and potentially shifting distributions of prey and seabirds		

Post construction monitoring recommendations

Key focus for quantifying and reducing uncertainty:

- need for a strategic approach with targeted data collection:
    - post construction monitoring needs to be more than what the industry has to do due to their requirements
    - Need for strategic (non-project specific) licence conditions (Scotland already doing some of this) - would allow collection of more useful data for future assessments
    - Consider compatibility of data collected pre and post consent – methods for data collection for the site characterization versus post consent monitoring has meant it hasn't always been possible detect change
    - Reduce disconnect between planning condition requirements and useful monitoring
    - better alignment of data collection methods across sites, to facilitate analysing data in a combined way
    - Collaboration in post consent monitoring between developers; avoid duplication, bolster financial scope, bolster sample sizes, share data
    - Agreed guidelines on the format and storage of data in central accessible (open-) source
- test mitigation measures such as colouring turbine blades
- make data available for research both pre and post construction
    - reduce lag in time taken for data/reports coming out of post-consent monitoring to inform future ORD assessments
    - Need to make sure all PCM and other data becomes available in a repository - stipulation of consent process
- consider and measure 'source' population changes not related to construction
- need to convert understanding of where uncertainty lies in modelling/assessment and how this can be translated into monitoring requirements and how outcomes will reduce uncertainty
- Improved understanding and estimation of barrier effects
- Improved links between post-consent monitoring data and its use in validating models used in pre-consenting

Table 10: Summary of main processes, data collection and analytical methods for improving the collection and application of post-consent monitoring data.

Process	Empirical Data	Analyses/tools/models
Variation in displacement rates due to turbine density or wind farm layout		
Quantification of avoidance behaviours		a big sensitivity in CRM, so need to improve precision/accuracy
Understand and quantify below surface habitat changes post construction		
Quantify changes to prey post construction		

Process	**Empirical Data**	**Analyses/tools/models**
Habituation of seabirds to ORDs		
Validate collision risk models, at least individual components (*e.g.* pcoll & flux)		Use existing and emerging turbine monitoring to compare accuracy of pre-construction CRM predictions, *e.g*. different versions of Band, Kleyeeg Hartman etc. Identify improvements to model structure – behaviour/better characterising flux/etc
Consider longer-term shifts due to climate change	Demographic monitoring across multiple colonies	
Barrier effects	GPS tracking of individuals, longer term tracking to estimate habituation and links to environmental variation	

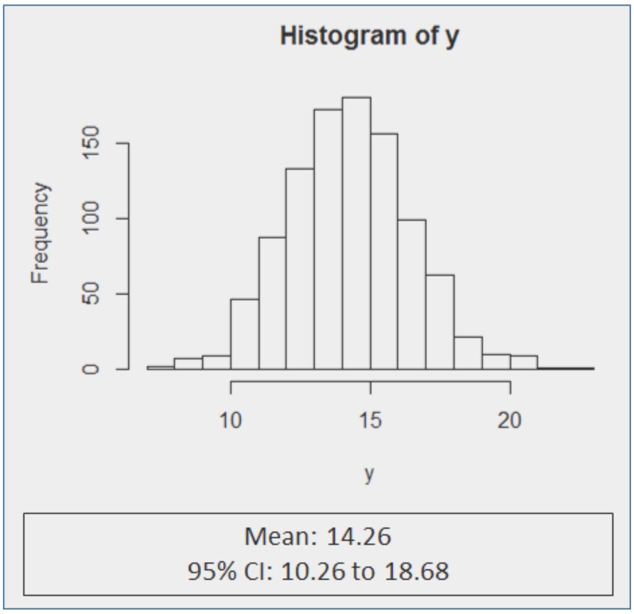
Appendix B: Simple hypothetical example of simulation-based approach

Consider the following hypothetical situation:

a) there are two “tools”, each of which depends on a single input, x;
b) the assessment process involves summing the outputs of the two tools together;
c) the value of the input x is uncertain; it can be assumed to have a normal distribution with mean 3 and standard error 0.2;
d) the first tool is stochastic, and involves simulating from a normal distribution with mean $7 * x$ and standard deviation 0.6;
e) the second tool is stochastic and involves simulating from a normal distribution with mean $x^2/20$ and standard deviation 0.2.

In this case, the following simple, block of R code allows the “tools” to be linked, and uncertainty in the resulting output to be quantified:

```
########################################################
nsim <- 1000
tool1 <- function(x, nsim){ rnorm(nsim, 7*x, 0.6) }
tool2 <- function(x, nsim){ rnorm(nsim, (x^2)/20, 0.2) }
x <- rnorm(nsim, 2, 0.3)
y <- tool1(x, nsim) + tool2(x, nsim)
hist(y) ## graphical representation of uncertainty
mean(y) ## best estimate
quantile(y, c(0.025, 0.975)) ## 95% confidence interval
########################################################
```



Appendix C: Hypothetical worked example for the communication of uncertainty in assessments

A worked example is provided here based on some of the key data used in collision risk modelling; here black-legged kittiwake (*Rissa tridactyla*) is used as an example. While many of the data values and sources are accurate, we have also included hypothetical data (*e.g.* bird aerial densities), which do not relate to an actual development (though these values do approximate those from some development sites).

The first step collates the bird data values and describes the source of these data (Table 11). There is then a simultaneous process of determining evidence quality and stakeholder agreement. The quality, consistency, amount, and type of evidence was assessed for each kittiwake datum.

Taking “nocturnal activity” as an example, the value proposed for use in the CRM is 0%. This is based on a single reference showing that both adult breeding birds of a pair return to their nest site at night (Coulson 2011). The quality of evidence was assessed as being of high quality, as it is based on a long-term study from a trusted source. Consistency of evidence was assessed as medium as, while there was only one data source, there was no contradictory data source found.

With respect to both the amount of evidence and the type of evidence, these evidences were assessed as “limited” as:

- This was only a report of a single study, from a single colony, thus the amount of evidence was limited; and
- The type of evidence was only from actively breeding birds at the colony, rather than from birds at sea, this was also “limited” to the situation being assessed.

The overall evidence value was rated as medium, due to a high-quality data source and medium consistency of evidence. It was thought that although assigned a ‘limited’ value, the amount and type of evidence was not sufficiently lacking to reduce the overall evidence value to limited.

Table 11: List of metric values and data sources, with assessment of the quality, consistency amount and type of evidence.

Metric	Value	Source	Quality of evidence	Consistency of evidence	Amount of evidence	Type of evidence	Overall evidence
Bird length	0.38 – 0.40 m (midpoint = 0.39 m)	Snow & Perrins (1998)	Medium	Robust	Robust	Robust	ROBUST
Bird wingspan	0.95 – 1.20 m (mid-point = 1.075 m)	Snow & Perrins (1998)	Medium	Robust	Robust	Robust	ROBUST
Flight speed	13.1 ± 0.4 ms^{-1}	Alerstam *et al.* (2007)	Medium	Robust	Robust	Robust	ROBUST

	(Mean ± S.D.)						
Nocturnal activity	0%	Coulson (2011)	High	Medium	Limited	Limited	MEDIUM
Flight type	Flapping	Snow & Perrins (1998)	Robust	Robust	Robust	Robust	ROBUST
Day time bird flight density	0.25 – 2.5 birds km^{-2}	Surveys of proposed wind farm	Robust	Limited	Medium	Robust	MEDIUM

The example of stakeholder agreement shown in Table 12 is entirely hypothetical. It was assumed that agreement would be sought from the regulator and their statutory advisor(s) for the use of the proposed values for the metrics presented in Table 11. In addition, independent experts could be consulted (*e.g*. academic scientists). Finally, key stakeholders would also be consulted (*e.g*. RSPB for birds). In this hypothetical example, the suggested nocturnal activity value of 0% was rejected by all of the consultees. In reality, we would hope that reasons would be given, and recommendations of alternative data values and source would be provided. In addition, if an alternative value was based wholly on precaution, this should also be recorded. Since all consultees disagreed with the value suggested for nocturnal activity, the agreement value was “low”.

The input from different stakeholder responses could either be given equal or unequal weighting. For instance, it may be that independent expert advice is given the highest weighting, or it may be that statutory advisors are given the highest weighting as their opinion carries some legal value. If weightings were applied to stakeholder responses, we recommend that these are transparent and recorded. Stakeholders should be made aware of any weightings before being asked for the consultation response.

Table 12: Record of stakeholder agreement on each datum and its source (hypothetical example).

Metric	Developer	Licensing authority	Independent expert	SNCB 1	SNCB 2	Key stakeholder(s)	Agreement
Bird length	Y	Y	Y	Y	N	N	MEDIUM
Bird wingspan	Y	Y	Y	Y	N	N	MEDIUM
Flight speed	Y	Y	Y	N	Y	N	MEDIUM
Nocturnal activity	Y	N	N	N	N	N	LOW
Flight type	Y	Y	Y	Y	Y	Y	HIGH
Day time bird density	Y	Y	Y	Y	N	N	MEDIUM

In this hypothetical example, we have assumed that the stakeholders all agreed that, after the introduction of a degree of precaution, a modified nocturnal activity value of 25% was acceptable (Table 13). We recommend that the value is modified until the stakeholder agreement value is high (most, but not all, stakeholders agree). In order to prevent inappropriate precaution being introduced, the nocturnal activity value in this example should not be modified until 100% agreement is reached. These values were then applied to the IPCC confidence matrix (Table 1).

Table 13: Modification of datum with low agreement and precaution value.

Metric	Value	Agreement	New value	Precaution
Nocturnal activity	0%	LOW	25%	High

The overall confidence value derived from the IPCC confidence matrix (Table 11) was recorded in Table 14. In the example here, the modified nocturnal activity value is highlighted in red, and shows a medium confidence value due to limited evidence but high agreement.

Table 14: Summary of overall evidence, added precaution and agreement values, and the final overall confidence value of each data.

Metric	Overall evidence	Precaution	Agreement	Confidence
Bird length	Robust	None	Medium	HIGH
Bird wingspan	Robust	None	Medium	HIGH
Flight speed	Robust	None	Medium	HIGH
Nocturnal activity	Limited	High	High	MEDIUM
Flight type	Robust	None	High	VERY HIGH
Day time bird density	Medium	None	Medium	MEDIUM

This example illustrates that only one value of our set of hypothetical data needed any modification. While some stakeholders disagreed with some values, their concerns were recorded and compared with other stakeholders. No modification of values, other than nocturnal activity, took place as agreement remained sufficiently high for the assessment to use the suggested value. In addition, the one value with low agreement was only modified sufficiently to reach high agreement, which we recommend is not 100% agreement, but a majority agreement. In order to achieve consistency with the planning process and legal authority of the Determining Authority and their statutory advisers, this majority agreement could be through a weighted agreement approach

References

Coulson, J.C. 2011. The kittiwake. A & C Black