

OFFSHORE RENEWABLES JOINT INDUSTRY
PROGRAMME (ORJIP) FOR OFFSHORE WIND



AppSaS – Apportioning seabirds seen-at-sea

WP4 – Tool development

February 2024



Contents

ORJIP Offshore Wind	3
Acknowledgements	3
Who we are.....	3
1. Introduction	5
2. Apportioning of guillemot and razorbill from GLS data	5
Overview	5
Methodological approach	6
Worked examples and comparisons with existing tools	8
Inter-relationship with other tools	10
Tool development process.....	11
Testing scenarios	11
Future work	11
Potential improvements to the ANBS tool (after v1.0 release).....	11
Integration with the CEF	12
Future data updates	13
Future research and development work	15
3. Apportioning of lesser black-backed gull from GPS data.....	16
Background	16
Methodological approach	16
Apportioning calculations	16
Adjustment for proportion of time on land	17
Classification of colonies as urban or non-urban	18
Other differences from GPS-based apportioning for other species.....	18
Delivery of a user-friendly tool via integration into the CEF	19
Extended set of user options	19
New data files within the CEF Data Store	21
Changes in internal functionality	22
Future work	22
4. NatureScot Apportioning Tool (inverse distance decay) extensions	24
Background	24
Methodological approach	24
Extension of the model	24
Relating the decay parameter to mean and maximum foraging ranges	25
Estimation of the decay parameter β using published foraging ranges	27
Results of estimation.....	27

Integration into the CEF	31
Future work	32
5. Comparison of breeding season apportioning methods	34
Introduction.....	34
Methodology	34
Selection of methods and species	34
Choice of footprints	36
Choice of SPAs.....	36
Results.....	36
6. Conclusions.....	39
7. References	40
8. Appendices.....	42
Appendix A. Extended distance decay model.....	42
Appendix B. Decay parameter estimation: SPA-level results.....	43
Appendix C. Full results of comparison of breeding season apportioning methods.....	45

ORJIP Offshore Wind

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

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

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

The current stage is a collaboration between The Carbon Trust, EDF Energy Renewables Limited, Ocean Winds UK Limited, Equinor ASA, Ørsted Power (UK) Limited, RWE Offshore Wind GmbH, Shell Global Solutions International B.V., SSE Renewables Services (UK) Limited, TotalEnergies OneTech, Crown Estate Scotland, Scottish Government (acting through the Offshore Wind Directorate and the Marine Directorate), and The Crown Estate Commissioners.

For further information regarding the ORJIP Offshore Wind programme, please refer to the [Carbon Trust website](#), or contact Ivan Savitsky (ivan.savitsky@carbontrust.com) and Žilvinas Valantiejus (zilvinas.valantiejus@carbontrust.com).

Acknowledgements

This document was produced on behalf of ORJIP Offshore Wind by Biomathematics and Statistics Scotland (BioSS), UK Centre for Ecology & Hydrology (UKCEH) and British Trust for Ornithology (BTO). The report was authored by Esther Jones, Katherine Whyte, Deena Mobbs, Kate Searle, Adam Butler, Ana Couto, Aonghais Cook, Chris Thaxter, Lila Buckingham, Maria Bogdanova and Francis Daunt.

The project has been advised by the ORJIP Offshore Wind Steering Group, and AppSaS Project Expert Panel. We would like to thank the following organisations for their advice and support of the project via participation on the Project Expert Panel:

- Natural England
- Natural Resources Wales
- NatureScot

This report was sponsored by the ORJIP Offshore Wind programme. For the avoidance of doubt, this report expresses independent views of the authors.

Who we are

Our mission is to accelerate the move to a decarbonised future.

We have been climate pioneers for more than 20 years, partnering with leading businesses, governments and financial institutions globally. From strategic planning and target setting to activation and communication - we are your expert guide to turn your climate ambition into impact.

We are one global network of 400 experts with offices in the UK, the Netherlands, South Africa, China, Singapore and Mexico. To date, we have helped set 200+ science-based targets and guided 3,000+ organisations in 70 countries on their route to Net Zero.

1. Introduction

Work Package 4 involves translating three aspects of functionality into user-friendly tools: (1) novel GLS-based apportioning estimates for the non-breeding season for razorbill and guillemot, (2) novel GPS-based apportioning estimates for the breeding season for lesser black-backed gulls (LBBG) and (3) extensions to the NatureScot apportioning tool. The first of these has involved substantial work within WP4, in part because the functionality included within this tool goes beyond that in the current apportioning tool (BDMPS), and in part because of technical challenges in using the GLS-based maps. As a result, and because integration into the Marine Scotland Cumulative Effects Framework (CEF) involves addressing some challenges, this part of the work has been delivered through a self-contained tool (albeit one that is futureproofed for future inclusion in the CEF). A live version of this tool has been produced, and this is summarised here and in the [User Guide](#).

The other two elements of WP4 involved smaller amounts of work within WP4 (even though the LBBG work involved substantial modelling work in [WP3](#)), and are readily incorporated into the CEF, so are being delivered via the CEF.

2. Apportioning of guillemot and razorbill from GLS data

Overview

Monthly estimates of the spatial distribution of guillemot and razorbill from tracked colonies through the non-breeding season have been derived from geolocator data, together with associated quantification of uncertainty ([WP3](#)), and the results compared against BDMPS at the level of BDMPS regions.

A stand-alone Apportioning in the Non-Breeding Season (ANBS) tool has been developed, and is available to download on [Github](#). The tool consists of the following elements:

[ANBS User Guide v1.0](#) (.docx) – user guide describing how to use the apportioning in the non-breeding season tool including set-up, user inputs, outputs, methodology and constraints. Case studies are examples showing different species, hypothetical offshore renewable footprints and modes of use and are provided to users as appendices.

ANBS R code (.rmd) - R code to run the ANBS tool.

ANBS data (various formats) – data required to run the ANBS tool.

The User Guide is an accessible user-facing document for the tool, and provides a practical, non-technical, description of the steps involved in running the tool using the open-source software R. The tool produces outputs in the form of an automated PDF report, and examples of this report are included to illustrate the content and format. This report does not seek to duplicate the User Guide or worked examples, or the [WP3](#) report, but instead provides additional context by (1) outlining the justification for the methodological approach taken within the tool, (2) describing the inter-relationship between this tool and other products

and tools (primarily the CEF), and (3) by outlining future work, including the potential for future integration of the tool into the CEF.

Methodological approach

The structure of the tool is similar to that used for current breeding season apportioning tools: users have the option to either select *map mode* (colonies of interest are selected and mean estimates of apportioning with uncertainty are mapped), or *footprint mode* (mean estimates of apportioning and uncertainty are estimated from relevant colonies for the area of an uploaded footprint). Unlike in breeding season tools, however, users can select the month or season that they wish to consider, since the GLS-based maps are monthly. Given the variation in spatial distribution between months and the challenges in producing biological definitions of seasons, this flexibility is designed to allow users to be able to use the tool to either apportion for seasons, whose start and end months they can specify, or for individual months. This will futureproof the tool against any possible future changes in the way in which assessments define seasons, and against any potential future switch from apportioning at a seasonal level to apportioning at a monthly level. Partial months (e.g. 2 weeks) cannot be selected by users because the underlying GLS-based utilisation distributions for tracked colonies are calculated at a month level. Therefore, the smallest unit of disaggregation in the ANBS tool is a month. In our view, if UD's were calculated on timescales smaller than a month (e.g. weekly or daily), the sparsity of data would introduce estimation bias and increase uncertainty, and would prohibitively increase computational processing time.

Apportioning relies upon knowing the spatial distributions associated with every colony, not only those colonies with tracking data. For the geolocator-based models to be used for apportioning it was necessary to either be able to use these models to produce spatial distributions for colonies without GLS tracking data, or else to use an alternative approach for colonies without GLS tracking data. The only obvious alternative approach to use in this context was BDMPS (Furness et al., 2015). Since there are limitations to BDMPS, we extrapolate the geolocator-based models out to colonies without GLS tracking data where these colonies are within a certain distance threshold GLS-tracked colony, and use BDMPS where the distance to the nearest colony with GLS tracking data exceeded this threshold. This provides a hierarchical approach, in which the best available method is used for each colony. Users can specify the distance threshold, which also allows sensitivity of the results to the choice of distance threshold to be evaluated. The tool outputs a red-amber-green (RAG) map that indicates whether, for the selected distance threshold, each UK SPA for the species has GLS tracking data (green), does not have tracking data but is within the distance threshold of a colony with GLS tracking data (amber), or is beyond the distance threshold and therefore use BDMPS rather than GLS (red).

Within each simulation each untracked colony is assumed to be associated with a particular tracked colony, where the probability of being linked to each tracked colony is equal to:

$$(distance\ threshold - distance\ by\ sea\ from\ untracked\ colony\ to\ tracked\ colony) / (distance\ threshold)$$

[Equation 1]

Tracked colonies that are closer to the untracked colony are therefore more likely to be selected than those that are further away. This process is similar to the calculation of a spatial distribution based on a

weighted average of distributions for colonies with tracking data but differs in explicitly accounting for the uncertainty associated with the linkage of untracked to tracked colonies.

The predicted number of birds from each colony in each grid square is calculated to be the utilisation distribution (UD) for that colony multiplied by the colony size. This calculation is performed separately for each simulated UD, allowing uncertainty to be propagated through the calculations. These predicted numbers can be converted into densities per km² by dividing by the grid cell area.

Map mode visualizes these distributions for the selected colonies, if these colonies are ones for which the distance threshold allows GLS-based maps to be produced. Footprint mode aggregates abundance within the footprint (uploaded by the user) and presents estimates of total abundance and density within the footprint, along with corresponding confidence intervals that represent the uncertainty associated with this. The proportion of birds within the footprint that are estimated to come from each population is then also shown, together with uncertainty.

Populations are taken to be exactly as in BDMPS (Furness et al., 2015) for each of the two species under consideration. This approach was taken to: (a) allow direct comparisons with the existing apportioning method (BDMPS) and to ensure that all functionality is also possible within the tool, (b) circumvent the difficulties in aligning Seabird 2000 subsites with SPA boundaries, (c) make abundance estimates as easy to update as possible (by allowing users to directly specify the size of each colony, with the BDMPS population sizes being used as defaults), and (d) allow non-SPA and non-UK populations to be accounted for (these are always assumed to use BDMPS: the GLS-based approach is only considered for UK SPAs).

For each population at which BDMPS is assumed to be the best method (UK SPAs that lie further than the distance threshold from a colony with GLS tracking data, as well as all UK non-SPA and non-UK populations, see [WP3](#) report) the abundance of birds within the footprint is assumed be:

*Breeding pairs within this population * 2 * Proportion of birds from this population that lie within the BDMPS region that contains the footprint, according to BDMPS * Proportion of the total area of the BDMPS region that lies within the footprint*

[Equation 2]

This is essentially the same as the calculation involving adult birds within BDMPS itself (immature birds are not currently considered within the tool), except that it is rescaled to relate to the footprint rather than to the whole BDMPS region. If BDMPS is used everywhere, then the proportion of birds arising from each population will not depend upon this proportion, so it can be ignored, but the multiplication by this proportion is necessary in order to translate the values obtained using BDMPS into the same units as those derived from the GLS-based approach, and it is this translation that enables a hierarchical approach to be used.

The tool allows uncertainty to be incorporated into BDMPS in a simple way by assuming that the uncertainty in the proportion of time spent in the footprint within BDMPS is a user-specified scaling factor multiplied by the level of uncertainty in the GLS-based maps (as derived from calculating the standard deviation of this proportion across colonies and simulations). The default scaling-factor is 1.5 so that BDMPS uncertainty is assumed to be one and a half times as large as the uncertainty derived in the GLS-based maps. This approach, although simplistic, enables the uncertainty to be treated in a consistent way across all populations, avoiding the paradoxical situation in which the uncertainty appears to be lowest in the situations where evidence is most sparse.

The tool allows the option for users to use BDMPS at all colonies, to set the uncertainty in BDMPS to zero, and to use BDMPS population sizes (these are the default population sizes), and thereby provides a straightforward way for users to compare the results obtained using GLS-based methods against those obtained using BDMPS.

Worked examples and comparisons with existing tools

The [WP3](#) report compared the results obtained using GLS-based methods and BDMPS at the level of entire BDMPS regions. A comprehensive comparison at the level of individual footprints is difficult, because the number of logically possible footprints is virtually unlimited, and there are sensitivities in running tools using footprints associated with actual projects. We therefore exploit a set of three artificial footprints generated within the ORJIP Seabird Sensitivity Mapping Tool, whose aim was to reproduce realistic characteristics of projects of recent interest (e.g. footprint size, distance from coast, distance from key SPAs) without using actual footprints. The footprints used are shown in Figure 1. These footprints are designed to be plausible but are not intended to be representative of all projects that could be of interest.

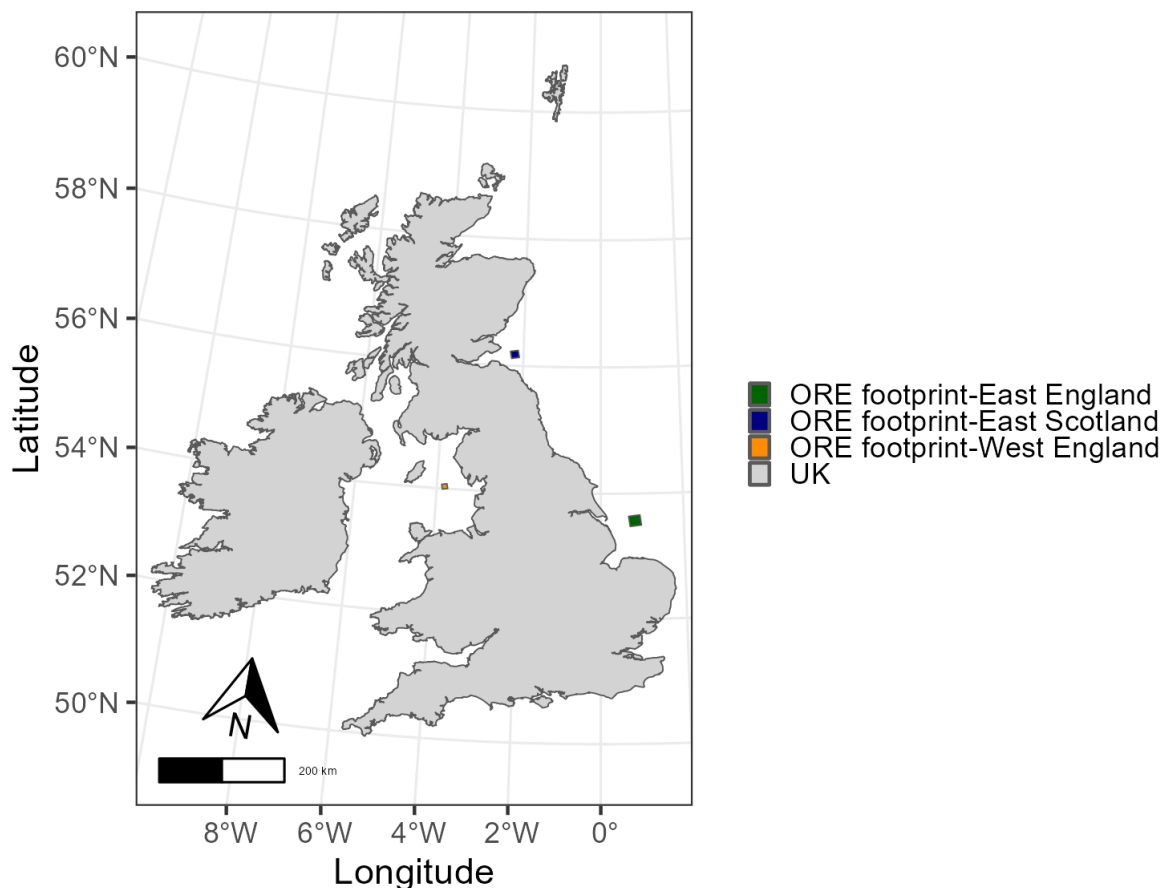


Figure 1. Three hypothetical offshore renewables energy (ORE) footprints used in the worked examples and comparisons.

We apply the tool to a set of five case studies, which are designed to show the different options of input parameters, including using the footprints (**Error! Reference source not found.**). However, the case studies are not exhaustive in terms of input parameters, they are rather designed to show examples of

the different options that can be used within the tool. In the [appendix of the User Guide](#), we show the output reports generated by the five worked examples, in order to illustrate the format of the output reports.

Table 1. Parameters of the five case studies used to illustrate running the ANBS tool.

Case study	Species	Mode	Usage	Distance threshold	Season	Months	BDMPS region
1	Common guillemot	Footprint – East Scotland	GLS & BDMPS	270 km	non-breeding season (Aug-Feb)	Aug-Feb	UK North Sea and Channel waters
2	Razorbill	Footprint – East England	GLS & BDMPS	150 km	migration seasons (Aug-Oct; Jan-Mar)	Aug-Oct; Jan-Mar	UK North Sea and Channel waters
3	Common guillemot	Footprint – West England	BDMPS-only	0 km	non-breeding season (Aug-Feb)		Western waters
4	Razorbill	Map	GLS & BDMPS	270 km	winter (Nov-Dec)	Nov-Dec	Western waters
5	Razorbill	Map	BDMPS-only	0 km	migration seasons (Aug-Oct; Jan-Mar)		UK North Sea and Channel waters

The reports are designed to be in a format that is comparable to that currently used in BDMPS, and to exploit BDMPS information for populations at which GLS-based information are unavailable or (based on the distance threshold selection) inappropriate. Table 2 highlights the tables in the BDMPS report that provide comparable results to the ANBS Tool calculations for each species, season and BDMPS region, and highlights the name of the column within each table that contains an estimate of the total number of adults from each population in the region. These totals can be converted into estimates of density per km², thereby enabling direct comparison against the GLS-based estimates, by dividing by the area of the BDMPS region (389636km² for UK Western waters, 341464km² for UK North Sea waters and Channel). This division by BDMPS region size allows the ANBS tool to align BDMPS and GLS-based outputs, enabling the hierarchical approach.

Table 2. Tables in the BDMPS report (Furness et al., 2015) that contain data relevant to the ANBS tool for each species, season and BDMPS region, and the name of the column in each BDMPS table that allows, via division by the BDMPS region size, an estimate of density per km² that is comparable to that produced from GLS data.

Species	Season	BDMPS region	Table, and name of key column within table, in BDMPS report (Furness et al., 2015) that is comparable and relevant to ANBS calculations
Guillemot	Non-breeding season (August to February)	UK North Sea & Channel waters	Table 62: UK N Sea & Channel Number adults
		UK western waters	Table 63: UK western waters number of adults
Razorbill	Migration seasons (August to October, and January to March)	UK North Sea & Channel waters	Table 64: UK N Sea & Channel Number adults
		UK western waters	Table 65: UK western waters number of adults
	Winter (November and December)	UK North Sea & Channel waters	Table 66: UK N Sea & Channel Number adults
		UK western waters	Table 67: UK western waters number of adults

Inter-relationship with other tools

The tool is designed to use a similar structure as BDMPS, and so should be readily compatible with the way that non-breeding season apportioning is currently carried out. In particular, the format of outputs is similar and the set of populations used is the same for each species as in BDMPS. There are two main differences in terms of presentation from BDMPS, both of which are necessary to exploit the advantages of using geolocator-based maps:

- uncertainty in apportioning is presented, in addition to the mean estimates (where GLS colonies are included);
- the tool presents abundance at the level of the footprint or grid cell (in footprint mode), and at the level of the entire (selected) BDMPS region (in map mode), so the apportioning results for a footprint will depend on the location of that footprint, not just on the BDMPS region that it belongs to.

The ANBS tool has been developed as a stand-alone tool, and is not currently integrated into the CEF, but the outputs from this tool can be used as inputs within the CEF. Within the structure of the CEF,

apportioning constitutes a user-defined input, rather than one of the main set of linked tools, so the CEF allows functionality for users to specify apportioning proportions directly. Default apportioning values for the non-breeding season in the CEF are derived from BDMPS, but users could opt not to use these, and instead to input their own values based on the outputs from this tool.

The CEF needs an apportioning proportion for each project in each non-breeding season. These values could be obtained by running the ANBS tool once, using the relevant project and season. The CEF and the ANBS tool both provide a high level of flexibility in defining seasons. When using outputs from the ANBS tool in the CEF only those SPAs that are being considered within the CEF run need to be inputted (but note that other SPAs, and non-SPA and non-UK populations, have already been taken account of in deriving the apportioning values for the SPAs of interest).

Note that the CEF currently only accepts apportioning estimates – it does not account for uncertainty in apportioning, but future work (see below) could allow the CEF to be extended to consider this.

Tool development process

A live version 1.0 of the tool has now been produced. Changes were made to the draft tool based on reviewer feedback, increasing functionality of the tool to incorporate many more scenarios, and extensive testing. There is currently one outstanding medium priority change that will be addressed in a future release of the tool (version 1.1), beyond the lifetime of this project. A failsafe stop mechanism will be implemented into the tool, which stops the tool run if there is a fatal error. An example of this is when a user has uploaded a footprint that is outside the UK EEZ. The code has been written to catch this error but the failsafe mechanism has not been implemented yet (in v1.0 of the ANBS tool) because it requires an additional round of testing.

Testing scenarios

The tool has undergone several stages of testing. One of the final stages was to create a set of scenarios for a 'user' (i.e. not the tool developer) to test using their own operating system and environment. There are thousands of possible combinations of user input parameters. From these, a selection of 20 testing scenarios were chosen randomly, but also ensuring that a wide a range of input parameters were included. All 20 of these tests were passed.

Future work

In this section we outline future work, structured into four distinct areas:

1. Improvements that could be made to the ANBS tool following release
2. Integration of the tool into the CEF
3. Future data updates that could be integrated into the tool
4. Potential for wider research and development work related to the tool

Potential improvements to the ANBS tool (after v1.0 release)

There are several improvements to the tool that would be possible beyond the end of the current project:

1. It would be desirable for the tool to be developed into an R package, to make it as widely accessible as possible, and to allow technical documentation for individual functions within the tool to be provided in a standardized way. The tool code has been written with the aim of making this as straightforward as possible, but it will nonetheless require sufficient additional work that this is unlikely to be feasible within the timelines of the project.
2. The datasets used by the tool are large, so computational speed is a potential issue. There has already been substantial work within the draft tool to make the code as efficient as possible, but there would be value in further optimising the code for speed so that processing time is reduced.
3. There would be value in improving the frequency and clarity of warning and error messages, to allow users to identify the sources of any issues as easily as possible.
4. The packages associated with spatial functionality in R are currently undergoing substantial changes – the newest versions of the key spatial R packages were not sufficiently stable to allow them to be used within version 1.0 of the tool, but it will be necessary in future to update the tool to use the newer spatial R packages (*terra* and *sf*) to prevent issues when the existing packages that underpin the tool are deprecated.
5. There is potential to include immature as well as adult birds within the tool – however, as the GLS-based maps relate to adult birds, immature birds could currently only be included in simple ways (e.g. by adopting the same approach as in BDMPS, or else by assuming that the GLS-based maps can also be applied to immature birds).

Integration with the CEF

The ANBS tool has been developed as a stand-alone tool but designed and structured to be integrated into the CEF in a straightforward way as possible. The datasets that are used both in the ANBS tool and the CEF use the same structure as in the CEF (for folders, naming and versioning), making it straightforward to align these with the CEF in future. The tool code is also structured in a similar way to the CEF, using similar data import routines and coding functions. The structure of the outputs from the tool is designed to link closely to the way that non-breeding season apportioning is currently included within the CEF.

Advantages of future integration into the CEF would be:

1. Ability to exploit the web-based IT infrastructure that underpins the CEF.
2. Full alignment of the input data required for the tool with the CEF Data Store.
3. Ability to link the tool directly into the rest of the CEF, without the need for users to manually input outputs from this tool into the CEF.
4. Ability to check for inconsistencies, for example, checking that the species, months and seasons used within this tool are equivalent to those specified for the CEF.
5. Alignment with the reporting structure of the CEF.

A further benefit would be the ability to propagate uncertainty from this tool into the CEF, although this only requires the CEF itself to be extended to account for uncertainty in apportioning, and so could also potentially be achieved without, or prior to, integration of the ANBS tool into the CEF.

The practical work involved in integrating the tool into the CEF would be:

- Addition of the GLS-based maps, along with some other datasets used within the tool (e.g. BDMPS region shapefiles), to the CEF Data Store and creation of the metadata required to do this.

- Modification of the CEF user interface to allow users to specify inputs that are required specifically for this tool – some of the user-defined inputs required for the tool (e.g. footprint, selection of SPAs, choice of months), are already inputs to the CEF, but others (e.g. distance threshold, uncertainty scaling for BDMPS) would need to be added to the interface.
- Alignment of the tool with the internal CEF functionality, which involves aligning inputs and outputs from the tool with relevant objects within the CEF code.
- Extending the CEF functionality to allow uncertainty in apportioning values.
- Testing of the CEF to ensure that the integration has not introduced unforeseen errors or issues.

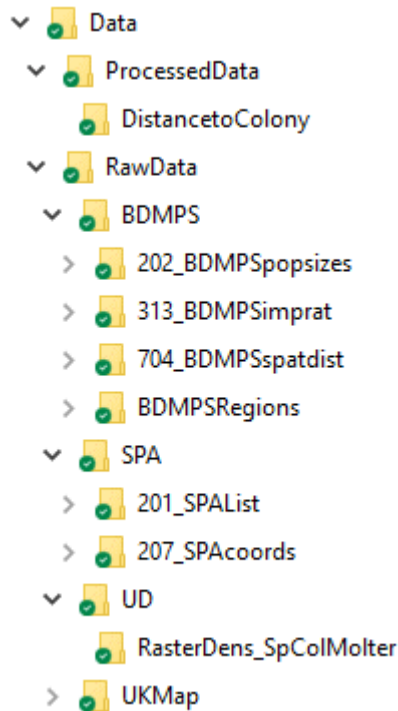
Future data updates

The tool is designed to be as straightforward as possible to update in future as future datasets and model outputs become available. Once new or revised data become available, the process that is involved in updating the tool to use these data will depend upon the nature of the changes/additions that have been made to the data. One way in which the tool has been designed to make updates to the data as straightforward as possible is by allowing key inputs to be directly user-specified. In particular, the population sizes of each SPA can be directly specified by the user, so changes to the underlying data are not required to use updated population size data within the tool. Similarly, the distance threshold is user-specified, so changes to this threshold do not require any changes to the underlying data.

The simplest other forms of data update involve revisions to the values within the current input data but leave the structure of the input data unchanged. In general, such updates should be straightforward to make and should immediately propagate through into the operation of the model. In particular:

1. Changes to the GLS maps could, if the set of colonies, the grid resolution and the grid projection remained identical, be achieved by substituting the existing files by a new set of raster files, whilst keeping the folder structure as it is at present.
2. Changes to the numeric values used for BDMPS population sizes or BDMPS spatial distribution values can be made by simply updating the values in the corresponding CSV files (note that population sizes are also user-modifiable within each run of the tool, so even without changing the underlying data it is possible to run the tool using different population sizes)
3. Changes to the BDMPS region polygons could, if the number and names of the regions remains unchanged, be incorporated by simply replacing the existing shapefiles with revised shapefiles

More information on the input data used are given in Section 4 of the [User Guide](#). If updates to datasets are made the structure of the input data folders needs to be preserved:



g

Although such changes are straightforward, note that a system of version control is needed to keep track of modifications that have been made, and to avoid ambiguity about the version used. Integration into the CEF would achieve this because a standard process of version control and data updates has been established within the CEF. The BDMPS and SPA data in the tool already use the CEF structures (e.g. for folders, versioning etc). This was designed so that integration within the CEF is as straightforward as possible, and if the data are updated in the CEF before the tool is integrated, the tool can be updated similarly (with additional cost). The tool is coded in a similar way to the CEF, using similar data import routines and coding functions similarly - this means that future updates to the tool can easily be aligned with changes to the CEF.

More complex updates are those that change the structure of the data, as well as the values themselves. Such updates may involve the addition of files for new species, modifications to the set of populations being considered, or revisions to the projection and resolution of the GLS-based maps. In general, such updates will require not only the dataset themselves to be updated, but also the tool code to be modified to account for the revised data structure. When the tool code has been modified it will also be necessary to re-run tests.

It is difficult to provide general guidance on the process for updating the data, or to futureproof against all such updates, in situations where the data structure changes, because this will depend on the exact nature of the changes that have been made, and the exact format of the revised/additional data. The code underpinning the tool will, however, be clearly documented, and it is structured in a way that makes it as straightforward as possible to see where each dataset is used within the calculations.

Changes that go beyond simple updates to the data files, and that involve changes to the structure of these files, are likely to require the involvement of someone with experience in R programming, in order to evaluate whether changes to the code are necessary, and to test the impact of such changes.

Note that the key input data for the tool are the GLS-based maps: the production of new or revised GLS-based maps would be a substantial piece of modelling work, and the appropriate approach to take would depend upon the design of the study and the empirical characteristics of the data. We therefore assume here that such models have already been fitted and used to produce estimated utilisation distributions (UDs) for each tracked colony, along with a quantification of uncertainty. The [WP3](#) report outlines the approach that was taken within this project to the generation of such spatial distributions from GLS data, but a general description of the appropriate approach to take for analysing such data is beyond the scope of this project.

Future research and development work

Improved quantification of uncertainty

The uncertainty quantification within the tool is designed to capture a key source of uncertainty within GLS-based maps – the relatively substantial locational uncertainty associated with GLS tags – and to incorporate some uncertainty into BDMPS estimates. The approach taken is a fairly straightforward one, but this could potentially be improved upon in future work. In particular, the approach:

1. Currently only considers one source of uncertainty in GLS-based uncertainty, albeit a key one - locational uncertainty. Other sources of uncertainty (e.g. in colony sizes) would ideally also be incorporated.
2. Focuses on the uncertainty associated with tracked birds at tracked colonies, and only deals in a partial way with the extrapolation from this to the uncertainty associated with considering all birds at all colonies.
3. Taken to quantify uncertainty in BDMPS does not directly consider uncertainty in BDMPS, but rather makes assumptions around the relationship between GLS-based uncertainty and BDMPS uncertainty.

Additional work on GLS tracking and modelling

Although the GLS data set on guillemot and razorbill constituted a step change in winter distribution data availability from multiple UK colonies, there were still important gaps in coverage, which could be filled with new field campaigns. However, it would be important to account for inter-annual variation when adding additional colonies to the data set. Improvements in modelling of GLS data is an ongoing focus in academic studies, and any new developments could be incorporated into future analyses. Estimating location from geolocation technology will remain challenging, however, and any technological developments that allow year-round deployments of GPS loggers are eagerly anticipated.

3. Apportioning of lesser black-backed gull from GPS data

Background

In [WP3](#), GPS tracking data from lesser black-backed gulls (LBBG) were used to model the spatial distribution of LBBG for colonies around Britain and Ireland. Using the approach developed in Wakefield et al. (2017), colony-specific habitat use was modelled by examining the relationship between bird habitat use and several covariates for colonies where GPS tracking data were available, and using this model to predict the distribution of birds from untracked colonies. We also adapted the models to allow for movement of birds across areas of land when assessing habitat accessibility, an adjustment which was not necessary for the seabird species considered previously using this approach (guillemot, razorbill, kittiwake, shag). LBBG distribution was found to be related to distance from the colony, habitat availability, competition, and the at-sea environment (depth). The resulting predicted colony-specific distributions were saved, alongside colony counts from the Seabird 2000 survey, to allow these results to be incorporated into an apportioning tool.

During analysis of the LBBG tracking data, it was observed that the time spent on land differed substantially between tagged individuals from different colonies. As these differences would likely alter the predicted density of birds attributed to at-sea locations during apportioning calculations (see Equation 1), it was necessary to explore and incorporate these differences further during WP4. Here, we outline how the difference in time spent on land is included within apportioning calculations, based on dividing both tracked and untracked colonies into two types (urban and non-urban). We then outline the approach taken to incorporate this into the CEF.

Methodological approach

The underlying methodology for estimating colony-specific spatial distribution of LBBG was outlined in the [WP3](#) report: in this section we focus upon the methodology used for translating these distribution maps into an apportioning tool.

Apportioning calculations

An estimate of the density of breeding birds within each grid cell from each colony (i.e. the number of bird per km²) is calculated via:

*size of the colony * (1 – proportion of time that birds from this type of colony spend on land) * Value of the estimated colony-specific UD within the grid cell / (Size of grid cell)*

[Equation 3]

The proportion of birds within grid cell *i* that can be apportioned to colony *j* is then equal to the density of breeding birds arising from colony *j* divided by the summed density of breeding birds arising from all colonies.

Colonies are assumed to be Seabird 2000 subsites, and colony sizes to relate to numbers of breeding pairs.

This is essentially the same approach as that used in the GPS-based apportioning tool developed for Marine Scotland (Butler et al., 2020), which uses Wakefield et al. (2017) to provide GPS-based colony-specific estimates of utilisation distributions and uses Seabird 2000 data to provide colony size information, but with the important difference that the LBBG version includes an adjustment for the proportion of time spent on land. This is necessary because LBBG spend a large proportion of their time on land, and, crucially, because the proportion of time spent foraging on land varies substantially between different kinds of colony (Langley et al., 2023). Note that, as in Butler et al. (2020) we focus here on time spent in all behaviours, not just foraging.

Adjustment for proportion of time on land

We assume here that colonies can be divided into two types, urban and non-urban, and that the proportion of time spent on land is different for these two types, but constant within each type. GPS tracking data are available for colonies within each of these two types, so we calculate the mean proportion of time spent on land by averaging across the tracked colonies within each type.

The mean percentage of time spent on land, based on data from the tracked colonies is:

Urban: 98.7039% (n = 54 bird-years, 2 colonies Belfast and Barrow)

Non-urban: 89.9318% (n = 352 bird-years, all other colonies)

Note, however, that there is considerable variation, with the standard deviations being 3.0629% and 13.6040% respectively. The number and location of colonies are detailed in [WP3](#) (Table 1), with a total of nine colonies across the UK used, of which seven were non-urban and two urban.

The division into two types is clearly an over-simplification of reality, since the proportion of time spent on land is likely to vary between colonies within these groups (particularly the “non-urban” group, which contains a wide range of colony characteristics) and between individuals within colonies. However, we did not think that a division into more than two groups was useful/meaningful in this context because:

- the vast majority of colonies are untracked, with a fairly small number of colonies having been tracked, so we need to have an approach that allows us to estimate the proportion of time spent on land for each untracked colony using the information from the limited number of tracked colonies;
- the level of uncertainty in the mean for each group is likely to become more pronounced as the number of groups considered increased.

The grouping of colonies into a dichotomous urban/not-urban split, although broad scale, represents a relevant ecological separation by breeding habitat type. Typically, LBBG have bred at coasts around north-west Europe (Ross-Smith et al. 2014), including habitats such as estuarine lagoons, spits, islands and some inland areas, such as moorlands. However, in recent years, LBBG have increasingly utilised urban areas for both breeding and foraging (Balmer et al. 2013, Langley et al. 2023). Further, there are known differences in the behaviour of LBBG breeding in natural and urban locations (Spelt et al. 2019, Booth Jones et al. 2022, Langley et al. 2023). Recent studies in the UK have shown that urban breeders spent most time in urban settings (Bristol, Spelt et al. 2019), and have core foraging areas within urban locations spending less time in marine areas than natural nesters (Urban Belfast vs natural Copeland). Further,

Langley al. (2023) found that urban breeders at Barrow-in-Furness positively selected urban habitats whereas natural nesting coastal birds at nearby Walney favoured coastal habitats and travelled further from the colony than urban birds (Thaxter et al. submitted).

Thus, the expected time spent in offshore and onshore environments is a priori expected to vary across colonies across a simple urban/non-urban gradient. Further grouping of sites could be possible in sub-habitat types, but as explained above, this was not feasible with the limited sample size of observed colonies.

Classification of colonies as urban or non-urban

Colonies were classified as either urban or not urban following recent methods used in developing urban survey design protocols for large gulls (see Thaxter et al. 2017). We used a grid mesh of 1 km squares covering the British Isles and extracted proportional cover of different habitat types in those squares using a combination of the Land Cover Map (2007) dataset¹ (Morton et al. 2011) for England, Scotland and Wales, and Corine Land Cover data² for Northern Ireland and the Republic of Ireland. Urban categories were identified and summed from those datasets per square. Squares were then divided into urban and non-urban categories using a simple 2% rule and 3.8% rule for LCM and Corine data, respectively, used previously as a robust delineation (Thaxter et al. 2017). Colony locations for LBBG were then extracted from the Seabird 2000 Census (Mitchell et al. 2004) and overlain onto spatial distribution of urban/non-urban squares. LBBG colonies were further filtered following the same methods as described in AppSaS [WP3](#), which thus allowed classification of all colonies to one of the two urban/non-urban strata.

Other differences from GPS-based apportioning for other species

Two other differences from the apportioning tool developed in Butler et al. (2020) arise because the large number of colonies with data for LBBG in Seabird 2000, combined with the foraging range being larger than that for other species in the tool, mean that computational issues are more substantive than for other species in the tool. We deal with this by (a) using a coarser grid resolution (10x10km) and (b) excluding colonies with less than 10 pairs from the calculations and tool. A total of 382 colonies are considered within the tool, with a combined population of 119,680 pairs; 108 of these are urban colonies, containing a total of 10,176 pairs. Note, for urban gulls, the urban population value used here is considered an underestimate given the use of vantage point methodologies (Burnell et al. 2021a), however at the time of writing, the Seabird 2000 survey remains the most spatially comprehensive dataset available across the UK. Since Seabird 2000, the population of urban gulls has also increased (Ross-Smith et al. 2014), although exact magnitudes have not been quantified for the whole UK. However, within England, recent modelling work has showed large increases in urban populations for Lesser Black-backed Gull between Seabird 2000 and the latest Seabirds Count census, and declines in natural nesters (Burnell et al. 2021b), although with a high degree in uncertainty of estimates. Work is therefore ongoing to ascertain the potential changes that have occurred in urban populations across the UK relative to traditional natural sites.

¹ <http://www.ceh.ac.uk/services/land-cover-map-2007>

² <http://land.copernicus.eu/pan-european/corine-land-cover>

Delivery of a user-friendly tool via integration into the CEF

The GPS-based apportioning calculations for lesser- black gulls will be integrated into the Cumulative Effects Framework (CEF), a user-friendly web-based framework that incorporates a wide range of data relevant to UK assessments and allows assessment tools to be linked together, and to those data, in a transparent way. The integration of the LBBG apportioning calculations into the CEF is relatively straightforward, because it simply involves extending existing functionality within the CEF to be available for an additional species. As such, the changes to the CEF involved in integrating LBBG breeding season apportioning into the CEF are an extended set of user options if users are running the CEF for lesser black-backed gulls and the addition of new data files to datasets into the CEF Data Store to support this.

Extended set of user options

When users select the apportioning tool to use in the breeding season, the CEF was previously designed to offer two options – apportioning based on distance decay and foraging range (previously called the “SNH” or “NatureScot” Apportioning Tool) or apportioning based on GPS-based maps (previously called the “MSS” Apportioning Tool). Apportioning calculations for the former approach within the CEF utilize distance by sea for most species, but use distance by air for gull species that contain inland colonies, including lesser black-backed gulls (because distance by sea cannot be estimated for inland colonies).

The incorporation of the LBBG work from this project into the CEF enables this functionality to be extended, if lesser-black backed gulls have been selected, to allow users to select between three possible breeding season apportioning methods (at the same point in the user interface at which users currently select between breeding season apportioning methods for other species):

Method 1a: Distance decay and foraging range apportioning using "distance by air", with inland colonies included

Method 1b: Distance decay and foraging range apportioning using "distance by sea", with inland colonies excluded

Method 2: GPS-based maps, with inland colonies excluded

Method 1a was already part of the CEF, whereas the work in this project has allowed Methods 1b and 2 to also be included as options.

The later methods (Methods 1b and 2) are based on the set of colonies considered in this project (382 Seabird 2000 subsites, excluding inland colonies and those with very small counts), whereas Method 1a is based on a larger set of colonies (1134 Seabird 2000 subsites). “Distance by sea” calculations cannot meaningfully be used when inland colonies are excluded, hence the difference. Methods 1a and 1b involve the same calculations, but differ in the input data used, both in terms of the set of colonies considered and in terms of the distance values used.

The guidance associated with the CEF is being modified slightly to account for these changes.

The CEF contains four different “pathways” – one relating to marine mammals, and three relating to seabirds. The first main pathway relating to seabirds is called “risk assessment” and allows user to estimate collision and/or displacement risk, to combine these, and to look at long-term consequences via a Population Viability Analysis (PVA). The second main seabird pathway is called “spatial planning”, and provides an interactive visualization tool for exploring spatial variations in sensitivity – it arose from a

separate project (the ORJIP Seabird Sensitivity Mapping Tool) but has been fully integrated into the CEF. The third and fourth seabird pathways are specific, and related to running of a PVA, and to the use for calibration of the SeabORD model.

The “risk assessment” and “spatial planning” pathways for seabirds can both, depending on the options selected by users, involve breeding season apportioning. Within the “risk assessment” pathway, the apportioning calculations will be used to provide *default* values to use for apportioning: users can modify these for subsequent steps (e.g. running the sCRM, SeabORD, Displacement matrix and PVA tool) if they wish to do so. Within the “spatial planning” step the apportioning calculations are used in producing the maps or tables that are shown in the output.

An important distinction between the way that breeding season apportioning works within the two pathways is that the “risk assessment” pathway operates at an SPA level, whereas the “spatial planning” pathway shows colonies in relation to Seabird 2000 subsites.

The maps produced in this project were derived in relation to Seabird 2000 subsites (following Wakefield et al., 2017), so apportioning estimates based on these can be used directly in the “spatial planning” pathway in the CEF. Within the “risk” pathway of the CEF the apportioning estimates for each Seabird 2000 subsite are combined up to SPA level by using a dataset within the CEF Data Store that provides a simple geographical overlap of Seabird 2000 subsites with SPA polygons to estimate the proportion of each Seabird 2000 subsite that lies in each SPA (Data Store dataset #609)³. The total proportion of birds to apportion to each SPA is then calculated to be the sum of the product (e.g. multiplication) of the proportion of birds apportioned to each Seabird 2000 subsite and the proportion of that subsite that is within the SPA. In the special case that the proportion of subsite in SPA were equal to one for a single Seabird 2000 subsite, and to zero for all other subsites, then the proportion of birds apportioned to the SPA would therefore be equal to that for the Seabird 2000 subsite that had a proportion of one within the SPA. Only the 118 CEF SPAs (i.e. those relevant to any CEF species, as defined in consultation with JNCC) are considered within this approach, and default apportioning proportions for marine SPAs are always set to zero. Within this approach SPAs are considered whether they are designated for the species in question or not, since in the CEF users have the option to over-ride the default option of considering SPAs for which the species is designated. Note that the approach used to derive these overlaps (i.e. to produce dataset #609) is provisional, in the absence of an alternative generic way of aligning Seabird 2000 colony definitions with SPA boundaries. This dataset, and the approach used to generate it, has therefore been flagged as an element of the CEF that requires feedback from stakeholders and may be revised within subsequent iterations of the CEF.

³ If a particular Seabird 2000 subsite is defined by a start and end location, then a set of 100 linearly interpolated points are produced between these locations, and the proportion of the subsite lying within the SPA is assumed to be the proportion of these locations that lie within the SPA; if only a start location is available for the Seabird 2000 subsite then this calculation reduces down to just considering whether this location is within the SPA polygon (in which case we assume proportion = 1) or not (in which case we assume proportion = 0).

New data files within the CEF Data Store

The additional functionality described here has required the addition of new data files to datasets within the CEF Data Store. In each case, these are variants on existing datasets within the Data Store, so are classed as dataset versions (termed “appsas1.0” in each case, to distinguish from the version “1.0” which forms the main initial dataset version within the CEF). For each dataset these additional versions are only for lesser-black backed gull. The datasets that contain new versions are:

- a) GPS-based grid (dataset #205): this has a similar format to that used for the Wakefield et al. (2017) species, but the grid resolution is coarser than that used for other species (10 x 10km, rather than 2 x 2km or 0.5 x 0.5km in Wakefield et al.) to reflect the larger foraging range for LBBG and the computational implications of this.
- b) GPS-based maps (dataset #702): these contain raster files containing colony-specific maps of estimated spatial distributions for LBBG, as produced by the analyses in this project (within [WP3](#)), but adjusted to account for the proportion of time birds from each colony spend at sea (depending on whether that colony is classed as urban or not). The adjustment involves multiplying the raw GPS based maps for each colony by the proportion of time spent at sea for a colony of this type. These datasets closely match those already included in the CEF Data Store that were based on Wakefield et al. (2017) outputs, except for the different grid resolution.
- c) Seabird 2000 colony count data (dataset #204): this is a CSV file containing the Seabird 2000 colony count data used to produce the above maps. This differs from the main Seabird 2000 version used in the CEF because [i] the colony size/location dataset contains an additional column indicating whether each colony has been classed as urban or non-urban, and [ii] the set of colonies is slightly different to that in Seabird 2000 (e.g. because colonies with less than 10 pairs are excluded)
- d) Distance from the GPS-based grid to the colony (dataset #205): the new version of this has the same structure as the main CEF version, but using the modified set of colonies from the “appsas1.0” version of dataset #204
- e) Proportion of Seabird 2000 subsite in SPA (dataset #209): the new “appsas1.0” version of this has the same structure as the main CEF version, but using the modified set of colonies from the “appsas1.0” version of dataset #204
- f) Proportion of sea within foraging range of each Seabird 2000 subsite (dataset #210): the new “appsas1.0” version of this has the same structure as the main CEF version, but using the modified set of colonies from the “appsas1.0” version of dataset #204

Metadata are being produced for the new data versions.

Changes in internal functionality

When generating default values for breeding season apportioning, the CEF automatically determines whether to use the new “appsas1.0” data versions or the original “1.0” versions. The new data versions (“appsas1.0”) of the above datasets will be used if lesser black-backed gull has been selected and either Method 1b or Method 2 is being used for apportioning; in all other cases the main CEF versions (“1.0”) will be used.

Future work

The relatively coarse grid resolution used for modelling (10x10km) means that the apportioning tool may fail to produce results for smaller footprints, as the relatively crude approach that is currently used to link footprint polygons to grid cells (via the “fasterize” package in R) may not successfully identify the grid cells associated with footprints whose size is close to, or below, the size of a single grid cell. Further work is needed to improve the alignment of polygons and grid cells, and it may also be desirable in the longer term to re-run the modelling using a higher grid resolution.

The work carried out in this WP to divide tracked and untracked colonies into two groups (urban and non-urban) was a practical approach to incorporate the most apparent differences in the time spent on land between different colonies. Further work to validate and potentially divide these coarse groupings into further, biologically relevant, sets of colonies would be a valuable extension to this initial analysis. For example, it is possible that time spent on land may vary between urban colonies situated in different geographic locations (e.g. depends on distance to coast), and depend on the characteristics of the urban settlement (e.g. size of settlement, abundance of landfill sites, e.g. Langley et al. 2021). The behaviour of gulls from non-urban colonies may also differ according to geographic aspects of the surrounding environment (e.g. island vs. coastal colonies) and resource availability.

In applying these groupings to the analysis, there remains a challenge in how the differences observed in a few tracked colonies ($n=9$) can be applied with confidence to the large number of untracked colonies ($n\sim 400$). Additional tracking across colonies, particularly at urban sites (where sample sizes are currently low), would help to improve the precision of estimates for the mean proportion of time spent on land in urban and non-urban sites. Tracking of additional colonies may allow a more detailed classification (e.g. involving more than two categories) to be used. Tracking effort is in large part dependent on local resources, however.

The apportioning calculations currently only use the best estimate of mean proportion of time spent on land within each colony, but there is considerable uncertainty associated with these estimated means, and this would ideally be incorporated into the apportioning calculations. In addition to the (mean) proportion of time spent on land differing between colony types, it should be noted that there is variability between individuals, and between colonies within each colony type, and that the variability between individuals may also differ between different colony types. This would ideally also be incorporated explicitly into the apportioning calculations but is likely to require additional tracking data.

Modelling habitat association at sea, and then applying a land-sea correction, is a relatively simple way to deal with the fact that birds spend a substantial amount of time on land in the context of assessing impact of offshore wind (where a detailed understanding of habitat use on land is not required). However, it would be conceptually much neater to model habitat associations at sea and on land within a single model that

thereby also describes the proportion of time spent on land for each colony in relation to habitat covariates. Even though habitat use on land is not of direct interest, such a model may also have benefits for improving the apportioning of birds at sea, by providing a more mechanistic basis for extrapolating the proportion of time spent on land from tracked to untracked colonies.

Habitat use depends upon behaviour, and there would also be benefits in refining the modelling to account for differences in behaviour – not only to distinguish between foraging and commuting behaviour, but also to distinguish time spent on the colony from other behaviours. As with the classification of time into land and sea, the key challenge lies in the need to extrapolate from a relatively small number of tracked colonies to a much larger number of untracked colonies.

4. NatureScot Apportioning Tool (inverse distance decay) extensions

Background

The NatureScot (previously SNH) Apportioning Tool assumes that for a particular location or area at sea, the proportion of birds arising from each breeding colony is proportional to the population size multiplied by the inverse squared distance by sea from the location to the colony, divided by the proportion of the area within the foraging range that is sea. Proportions are assumed to be zero at distances beyond the foraging range.

The tool is motivated by an assumption that the density of birds originating from a particular colony shows an inverse square distance decay with distance from colony. The division by the proportion of the foraging range that is sea is used to account for the fact that the density of birds needs, all else being equal, to increase as this proportion decreases, because there is a smaller area of sea to fit within.

This is clearly a simple model that does not account for environmental heterogeneity or competition. However, distance from breeding colony does typically dominate spatial distributions within the breeding season, so it has potential to nonetheless provide a reasonable approximation to actual distributions, especially for species and populations for which the data needed to derive colony-specific spatial distributions (e.g. GPS tracking data) are sparse or unavailable. The CEF incorporates both NatureScot apportioning, and, for the species modelled by Wakefield et al. (2017), apportioning estimates derived from these GPS-based maps.

Methodological approach

Extension of the model

In the absence of GPS tracking data, however, it is still possible to extend the NatureScot Tool so that it estimates the rate of decay of bird density with distance empirically for each species, using published foraging ranges, rather than fixing densities to always decay in proportion to inverse distance squared. This is worthwhile because the effect of distance to colony is crucial in determining the spatial distribution of seabirds during the breeding season (e.g. Wakefield et al., 2017), and systematic biases in quantifying the relationship between bird density and distance to colony will therefore lead to substantial errors in the calculation of apportioning percentages.

The CEF includes both the NatureScot apportioning method and, for relevant species, the GPS-based apportioning approach. Even for species without sufficient GPS tracking data to allow alternative GPS-based approaches, however, there is still potential to extend the NatureScot apportioning tool by allowing the rate of decay with distance to be determined based on published foraging ranges, rather than fixed to always have the same value. Specifically, the assumption that the spatial distribution is proportional to d^{-2} , where d is distance by sea to colony, can be generalized to an assumption that the spatial distribution is proportional to d^β where β is a user-specified parameter that indicates the rate of decay. It is assumed that the distance is lower than the maximum foraging range, so that the spatial distribution of birds is effectively assumed to be zero beyond that range. Using a value of -2 for the decay parameter corresponds

to the current NatureScot apportioning tool, so this is effectively an extension of, rather than a distinct method from the current tool. The extended version also provides a natural way of validating one of the key assumptions of the existing NatureScot tool: if there is evidence that the rate of decay is not equal to two then this may indicate a failure in the assumptions underpinning the tool.

Relating the decay parameter to mean and maximum foraging ranges

The NatureScot tool is defined as an algorithmic approach, rather than as a probabilistic or statistical model, but it is possible (if the effects of land are ignored) to relate it, and the extension of it that we have outlined here, to an explicit probabilistic model in which the distance to colony has a bounded Pareto distribution (with shape parameter equal to $-(\beta + 2)$) and the angle from colony has a uniform distribution (Appendix A). We can use a more statistical characterisation to examine the properties of this model. In particular, standard properties of a bounded Pareto distribution allow us to calculate, for any particular value of the decay model parameter β , the value of the ratio R between the mean and maximum foraging ranges, via the formula

$$R = \left(\frac{\beta + 2}{\beta + 3} \right) \left(\frac{\varepsilon^{-(\beta+2)}}{1 - \varepsilon^{-(\beta+2)}} \right) \left(\frac{1}{\varepsilon^{-(\beta+3)}} - 1 \right)$$

[Equation 2]

Aside from the decay parameter, this formula depends on only one other quantity, ε , which represents the proportion of the maximum foraging range below which there are assumed to be no birds. This quantity is not explicitly part of the NatureScot tool, but is needed in order to translate both the existing tool and the extended version into an explicit probabilistic model of spatial distribution, because the density of birds would otherwise often become infinite when the distance to colony become zero (since, for example, the inverse of distance squared would be infinity when distance is zero). In practice, a lower cut-off of distance from colony is typically used when calculating foraging ranges from GPS tracking data, to avoid the inclusion of birds at the colony itself, so the quantity δ can be regarded as representing the fact that there is a distance below which the distance-decay model will not be plausible, with such distances excluded from calculations relating to the foraging range.

The relationship between the decay parameter β and the ratio of mean to maximum foraging range R is shown in Figure 2. The relationship is shown for four possible values for this additional parameter ε (0.025, 0.01, 0.001 and 0.00001), in order to illustrate the sensitivity of the relationship to this parameter.

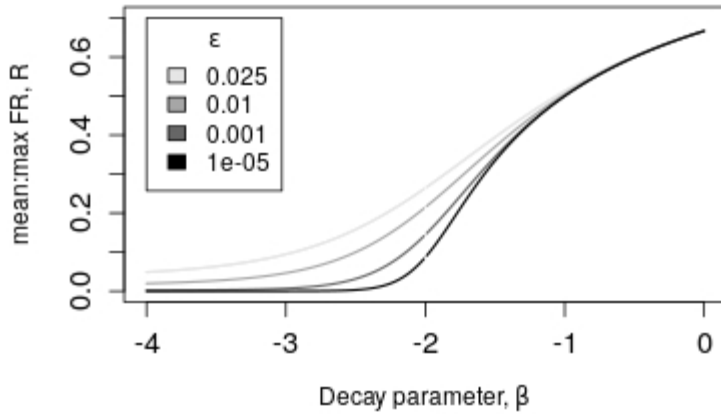


Figure 2. The ratio of mean to maximum foraging range (R) associated with different values of the distance decay parameter β . Results are shown for four different possible values of the lower threshold ϵ , which represents the proportion of the maximum foraging range below which the density of birds is assumed zero.

The reverse of this relationship can also be plotted: for each ratio of mean to maximum foraging ranges, R , we can calculate the value of the decay parameter β . There is no explicit formula that we can use to do this, but we can easily do this using a simple numerical optimization approach – i.e. trying a range of possible values of the decay parameter β , and, for each, using Equation 2 to calculate the ratio associated with this. We then select the value of the decay parameter that gives us the ratio that is equal, or (in practice) almost equal to the value of R that we are trying to achieve. We implement this in R using the optimize function, which uses a simple form of numerical optimisation to find the value of β that minimises the difference between the value of R associated with β and the value of R we want to achieve. This “reversed” relationship is shown in Figure 3. We can use this “reversed” relationship to help us identify plausible values of the decay parameter, since the ratio of mean to maximum foraging range is a quantity that can be derived from published foraging ranges.

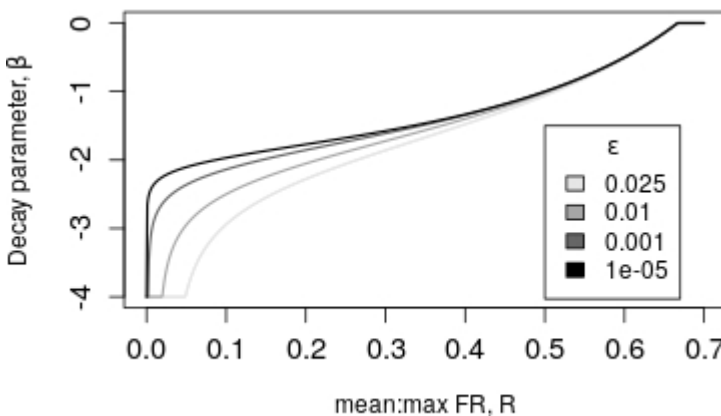


Figure 3. The distance decay parameter β associated with different values for the ratio of mean to maximum foraging range (R). Results are shown for four different possible values of the lower threshold ϵ , which represents the proportion of the maximum foraging range below which the density of birds is assumed zero.

Estimation of the decay parameter β using published foraging ranges

We estimate decay rates associated with published foraging ranges searching across a range of possible values for the decay rate, and, for each set of published foraging ranges, identifying the decay rate that provides the best match to the ratio of mean to maximum foraging ranges. We do this both using both SPA-specific foraging ranges, where available within the CEF Data Store, and using the overall species-level ranges derived in Woodward et al. (2019). When using the former we use the population-level maximum values; when using the latter we consider four possible metrics to represent the maximum: (a) mean-max (i.e. the population-level maximum, averaged across populations), (b) mean-max plus 1 standard deviation, (c) mean-max plus 2 standard deviations, and (d) max-max (the maximum range across all populations).

Results of estimation

In Table 3 we show the estimates of the distance-decay parameter for each species when using Woodward et al. (2019) foraging ranges, focusing on the mean-max plus 1 SD as the metric for the maximum.

Results suggest that the best estimates are sensitive, and sometimes very sensitive, to the value of the lower distance threshold ε , and hence to the assumptions that are made about locations in the immediate vicinity of the colony. For many species, the estimated parameters for at least some values of ε are relatively close to -2, indicating that the published ranges may be reasonably consistent with the existing NatureScot model. The species whose estimates differ most strongly from -2 is Black Guillemot, for which the values are always between -0.8 and -0.9. This is an unusual species in this context, however, in the sense that the ratio of mean to maximum foraging range is 0.54 (whereas it is below 0.4 for all other species).

Table 3. Mean and maximum foraging ranges, the ratio of the mean to the maximum, and associated estimates of the decay parameter β , based on published species-level foraging ranges from Woodward et al. (2019). “Max” foraging ranges are based on mean-max plus one standard deviation. Decay parameter value estimates are given for four different values of the lower threshold ε parameter; bird densities are assumed to be zero when the ratio of distance to max foraging range is less than this threshold.

Species	Max	Mean	Mean: max ratio	Decay parameter estimation			
				$\varepsilon =$ 0.025	$\varepsilon =$ 0.01	$\varepsilon =$ 0.001	$\varepsilon =$ 0.00001
Arctic Tern	40.5	6.1	0.15	-2.55	-2.24	-1.98	-1.87
Atlantic Puffin	265.4	62.4	0.24	-2.12	-1.93	-1.77	-1.71
Black Guillemot	9.1	4.9	0.54	-0.87	-0.85	-0.83	-0.83
Black-legged Kittiwake	300.6	54.7	0.18	-2.37	-2.12	-1.90	-1.81
Common Guillemot	153.7	33.1	0.22	-2.21	-2.00	-1.82	-1.74
Common Tern	26.9	6.4	0.24	-2.11	-1.92	-1.76	-1.70
Cormorant	33.9	7.1	0.21	-2.23	-2.02	-1.83	-1.75
European Shag	23.7	9.2	0.39	-1.52	-1.44	-1.38	-1.37
Great Skua	931.2	67	0.07	-3.32	-2.70	-2.25	-2.04
Herring Gull	85.6	14.9	0.17	-2.41	-2.15	-1.92	-1.82
Lesser Black-backed Gull	236	43.3	0.18	-2.36	-2.11	-1.90	-1.80
Manx Shearwater	2365.5	136.1	0.06	-3.66	-2.85	-2.32	-2.08
Northern Fulmar	1200.2	134.6	0.11	-2.83	-2.43	-2.10	-1.94
Northern Gannet	509.4	120.4	0.24	-2.12	-1.93	-1.77	-1.70
Razorbill	164.6	61.3	0.37	-1.58	-1.49	-1.42	-1.41
Roseate Tern	23.2	4.1	0.18	-2.40	-2.14	-1.91	-1.82
Sandwich Tern	57.5	9	0.16	-2.51	-2.22	-1.97	-1.85

Table 4 contains corresponding results for an alternative metric mean-max plus 2SD, and max-max. Results for mean-max foraging range plus 2 SD are similar to those for foraging range plus 1 SD, except that the estimate for black guillemot is now substantially closer to -2 (reflecting a substantially smaller mean:max ratio).

Table 4. As Table 3, but with “maximum” foraging ranges assumed to be mean-max plus 2 standard deviations.

Species	Max	Mean	Mean: max ratio	Decay parameter estimation			
				$\delta =$ 0.025	$\delta =$ 0.01	$\delta =$ 0.001	$\delta =$ 0.00001
Arctic Tern	55.3	6.1	0.11	-2.84	-2.44	-2.10	-1.95
Atlantic Puffin	393.7	62.4	0.16	-2.50	-2.21	-1.96	-1.85
Black Guillemot	13.4	4.9	0.37	-1.61	-1.51	-1.44	-1.42
Black-legged Kittiwake	445.1	54.7	0.12	-2.74	-2.37	-2.06	-1.92
Common Guillemot	234.2	33.1	0.14	-2.60	-2.29	-2.01	-1.88
Common Tern	35.8	6.4	0.18	-2.39	-2.13	-1.91	-1.81
Cormorant	42.2	7.1	0.17	-2.44	-2.17	-1.94	-1.83
European Shag	34.2	9.2	0.27	-1.98	-1.82	-1.69	-1.64
Great Skua	1419.1	67	0.05	-4.08	-2.99	-2.38	-2.11
Herring Gull	112.4	14.9	0.13	-2.66	-2.33	-2.03	-1.90
Lesser Black-backed Gull	345	43.3	0.13	-2.72	-2.36	-2.06	-1.91
Manx Shearwater	3384.2	136.1	0.04	-4.63	-3.11	-2.43	-2.14
Northern Fulmar	1858.1	134.6	0.07	-3.31	-2.70	-2.25	-2.04
Northern Gannet	703.6	120.4	0.17	-2.43	-2.16	-1.93	-1.83
Razorbill	240.5	61.3	0.25	-2.04	-1.87	-1.72	-1.67
Roseate Tern	33.8	4.1	0.12	-2.75	-2.38	-2.07	-1.92
Sandwich Tern	80.7	9	0.11	-2.83	-2.43	-2.10	-1.94

Results for max-max foraging range (Table 5) are also similar to those for foraging range plus 1 SD except that there are some species that can be included for this metric that were not included in other tables, and that have distance-decay parameter estimates very far from -2 due to the large mean:max ratios (Little tern, Mediterranean Gull, red-throated diver).

Table 5. As Table 3, but with “maximum” foraging ranges assumed to be max-max ranges

Species	Max	Mean	Mean: max ratio	Decay parameter estimation			
				$\delta =$ 0.025	$\delta =$ 0.01	$\delta =$ 0.001	$\delta =$ 0.00001
Arctic Tern	46	6.1	0.13	-2.66	-2.33	-2.03	-1.90
Atlantic Puffin	383	62.4	0.16	-2.47	-2.19	-1.95	-1.84
Black Guillemot	8	4.9	0.61	-0.43	-0.42	-0.42	-0.42
Black-headed Gull	18.5	7	0.38	-1.56	-1.47	-1.41	-1.39
Black-legged Kittiwake	770	54.7	0.07	-3.34	-2.71	-2.25	-2.04
Common Eider	22.5	3.2	0.14	-2.60	-2.28	-2.01	-1.88
Common Guillemot	338	33.1	0.10	-2.96	-2.51	-2.15	-1.97
Common Tern	30	6.4	0.21	-2.22	-2.01	-1.82	-1.75
Cormorant	35	7.1	0.20	-2.27	-2.04	-1.85	-1.77
European Shag	46	9.2	0.20	-2.28	-2.05	-1.86	-1.77
Great Black-backed Gull	73	16.7	0.23	-2.15	-1.95	-1.78	-1.72
Great Skua	1003	67	0.07	-3.42	-2.75	-2.27	-2.05
Herring Gull	92	14.9	0.16	-2.48	-2.20	-1.95	-1.84
Lesser Black-backed Gull	533	43.3	0.08	-3.17	-2.63	-2.21	-2.01
Little Tern	5	3.5	0.70	0.33	0.33	0.33	0.33
Manx Shearwater	2890	136.1	0.05	-4.09	-2.99	-2.38	-2.12
Mediterranean Gull	20	11.5	0.58	-0.67	-0.65	-0.65	-0.65
Northern Fulmar	2736	134.6	0.05	-3.98	-2.96	-2.37	-2.11
Northern Gannet	709	120.4	0.17	-2.43	-2.17	-1.93	-1.83
Razorbill	313	61.3	0.20	-2.30	-2.07	-1.87	-1.78
Red-throated Diver	9	4.5	0.50	-1.06	-1.02	-1.00	-1.00
Roseate Tern	24	4.1	0.17	-2.43	-2.16	-1.93	-1.83
Sandwich Tern	80	9	0.11	-2.82	-2.43	-2.10	-1.94

The results with mean-max foraging range (Table 6) differ fairly substantially from the other tables, in that a range of species in this case have high mean-max ratios, and therefore decay parameters that are substantially larger than -2. These species include Atlantic puffin, common guillemot, European shag and razorbill, in addition to the species already mentioned for other metrics.

Table 6. As Table 3, but with “maximum” foraging ranges assumed to be mean-max s. Cells in grey denote a situation in which the decay parameter cannot be estimated.

Species	Max	Mean	Mean: max ratio, R	Estimates of the decay parameter, β			
				$\varepsilon = 0.025$	$\varepsilon = 0.01$	$\varepsilon = 0.001$	$\varepsilon = 0.00001$
Arctic Tern	25.7	6.1	0.24	-2.11	-1.92	-1.76	-1.70
Atlantic Puffin	137.1	62.4	0.46	-1.26	-1.20	-1.17	-1.16
Black Guillemot	4.8	4.9	1.02				
Black-headed Gull	18.5	7	0.38	-1.56	-1.47	-1.41	-1.39
Black-legged Kittiwake	156.1	54.7	0.35	-1.67	-1.56	-1.48	-1.46
Common Eider	21.5	3.2	0.15	-2.56	-2.25	-1.99	-1.87
Common Guillemot	73.2	33.1	0.45	-1.27	-1.21	-1.18	-1.17
Common Tern	18	6.4	0.36	-1.65	-1.55	-1.47	-1.45
Cormorant	25.6	7.1	0.28	-1.95	-1.80	-1.67	-1.62
European Shag	13.2	9.2	0.70	0.30	0.30	0.30	0.30
Great Black-backed Gull	73	16.7	0.23	-2.15	-1.95	-1.78	-1.72
Great Skua	443.3	67	0.15	-2.54	-2.24	-1.98	-1.86
Herring Gull	58.8	14.9	0.25	-2.04	-1.87	-1.73	-1.67
Lesser Black-backed Gull	127	43.3	0.34	-1.70	-1.59	-1.51	-1.48
Little Tern	5	3.5	0.70	0.33	0.33	0.33	0.33
Manx Shearwater	1346.8	136.1	0.10	-2.93	-2.49	-2.13	-1.97
Mediterranean Gull	20	11.5	0.58	-0.67	-0.65	-0.65	-0.65
Northern Fulmar	542.3	134.6	0.25	-2.07	-1.89	-1.74	-1.68
Northern Gannet	315.2	120.4	0.38	-1.55	-1.46	-1.40	-1.38
Razorbill	88.7	61.3	0.69	0.24	0.24	0.24	0.24
Red-throated Diver	9	4.5	0.50	-1.06	-1.02	-1.00	-1.00
Roseate Tern	12.6	4.1	0.33	-1.76	-1.64	-1.55	-1.52
Sandwich Tern	34.3	9	0.26	-2.01	-1.84	-1.70	-1.65

Table B1, in Appendix B, shows the corresponding results for the SPA level foraging ranges. These results show a much higher level of variability than the results based on the Woodward et al. (2019) ranges, possibly in part reflecting noise in the estimation of SPA-level ranges.

Integration into the CEF

For all species for which the distance and foraging range (“NatureScot”) apportioning is available, the CEF functionality has been extended, by addition of a new input into the user interface, to allow an extended version of the apportioning tool to be used. This extension assumes that the proportion of birds arising from each colony is proportional, all else being equal, to $d^{-\beta}$, where d is equal to the distance by sea to the colony, and β is a decay parameter. The extension to the interface allows users to specify the value of the rate of decay parameter β , if they have opted to use breeding season apportioning to do this using the distance decay and foraging range (“NatureScot”) approach;. The value of this parameter was previously fixed to be equal to two within the CEF, so could not be specified by users. The “NatureScot”

tool is a special case of this extended version, in which β is fixed to be equal to two, and two is therefore the default value for the new parameter.

As in the existing “NatureScot” tool, the proportion of birds arising from each colony is assumed to be proportional to the colony size multiplied by the distance decay relationship, divided by the proportion of area within the foraging range of the colony that is sea. For gull species that have inland colonies, “distance by air” is used, and the proportion of area that is sea is not adjusted for; for all other species, “distance by sea” is used and an adjustment is applied.

Future work

The analyses undertaken here to estimate the decay parameter β are crude, because they estimate the values using only two pieces of information: an estimate of the maximum foraging range (either mean-max, mean-max plus 1SD, mean-max plus 2SD, or max-max) and an estimate of the mean foraging range. The resulting estimates are therefore likely to be very sensitive to any inaccuracies in these values, and the results obtained here indicate that they are also sensitive to the choice of metric used. Whilst we have looked at variability (e.g. between SPAs, and metrics) in estimates, the limited information also makes it impossible to use the published foraging ranges alone to quantify uncertainty in these estimates. A further empirical exploration of these issues using GPS tracking data would be valuable.

The distance-decay relationship considered here is of the form $d^{-\beta}$, but a different distance-decay relationship, $\exp(-d\beta)/d$, is considered for the distance-decay option within SeabORD, and other simple distance-decay models (e.g. $\exp(-d\beta)$) are also possible. It would be useful to compare the theoretical properties, practical robustness, and, in situations where GPS tracking data are available, empirical performance of these models. The rationale for using $\exp(-d\beta)/d$ rather than $d^{-\beta}$ in SeabORD is that the latter is particularly sensitive to the inclusion of very small distances, which is, in turn, likely to be very sensitive to the exact choice of grid. This is a problem for SeabORD, which requires the entire spatial distribution, hence the use of an alternative distance-decay model within SeabORD. Arguably, however, this issue is much less important for apportioning, since the values of the distribution are only needed for the location(s) at which apportioning is required (which are typically offshore, and so do not involve very small values for distance to colony). However, as the results in this section have shown, these problems do arise in the context of apportioning once the decay parameter is estimated empirically, rather than being fixed, because the estimation of the parameter in the power law model (d^β) is very sensitive to the way that distances close to zero are dealt with. Further exploration of the relative properties of different distance decay models would be valuable.

The process for estimating the decay parameter is also based on an idealised colony that is surrounded by sea, and so ignores the effects of local geography or of using distance by sea rather than distance by land, although these factors can still be accounted for (and within the CEF are) when using the estimated parameters to run the apportioning calculations. To avoid inconsistencies, however, it would be useful to extend the current approach to explicitly estimate the parameters in a way that accounts for the effects of land. The way that the “NatureScot” apportioning tool calculations adjust for land could also usefully be revisited and modified. This adjustment involves a division by the proportion of the area within the foraging range of the colony that is sea. This adjustment is motivated by the idea that the calculations need to be adjusted to account for the area available – i.e. that if only half of the foraging range area is sea then the density of birds in that area will, all else being equal, be double the density if the entire area

within the foraging range were sea. The principle is sound, but the way the adjustment is currently applied is crude, because it does not account for the fact that the density of birds, according to the distance-decay model, varies with distance (and, in the presence of land, angle). There is a risk that this may lead to internal inconsistencies, for example to the overall weight being assigned to each colony not being proportional to colony size when summed across all locations. Thus, further exploration of alternative approaches to adjust for the effects of land would be valuable.

Simple distance-decay relationships are unlikely to capture all of the key characteristics of colony-specific spatial distributions, so approaches that use a wider range of empirical data (e.g. GPS or GLS tracking data) and also account for the effects of competition and environmental variation are likely to provide more accurate estimates of apportioning. The simple extension of the NatureScot apportioning tool described here is therefore a straightforward interim solution to try to investigate the performance of the existing model, and, where appropriate, to extend it, but it would ideally ultimately be superseded by additional data collection and associated modelling.

5. Comparison of breeding season apportioning methods

Introduction

In the non-breeding season BDMPS is effectively the only method currently available for most species, so no comparison of methods is possible. For guillemot and razorbill the GLS-based method, which has been developed into the ANBS tool, can be compared against BDMPS, and this comparison is delivered via the case studies associated with the tool ([ANBS User Guide](#)).

In this section we therefore focus on the comparison of methods for the breeding season. In this case there are effectively two general methods – a distance-decay and foraging range approach, or an approach based on GPS-based maps. The latter approach is now available for five species (whereas it was previously, prior to this project, available for four). The distance-decay approach is available for a wider set of species (in the CEF for 13 species). The distance-decay approach depends upon the choice of foraging range, and this project has also allowed the decay parameter within it to be varied.

We compare results obtained using the distance decay and GPS-based apportioning methods, and using different implementations of the distance-decay approach.

Methodology

For each footprint, for each species, we use each method to estimate the proportion of birds in each footprint that are estimated to arise from each SPA. We then compare these proportions between methods.

Selection of methods and species

We consider twelve species (all of the species considered in the ORJIP Seabird Sensitivity Mapping Tool, with the exception of Red Throated Diver; this species is excluded because it lacks the data required to implement the apportioning methods). The species, and the available methods for each species, are shown in Table 7. This table also summarises those methods that have been made available, or have enhanced functionality, as a result of the work within this project.

Table 7. Species considered, and the apportioning methods that are available, and hence compared, for each species. Entries with a single asterisk (*) denote methods for which functionality has been extended within this project, entries with a double asterisk () denote methods that are newly available due to the analyses within this project.**

Species	Inland colonies considered?	Type of distance used for distance-decay approach	GPS-based apportioning approach available?
Atlantic Puffin	Not relevant	By sea*	No
Black-legged Kittiwake			Yes
Common Guillemot			Yes
European Shag			Yes
European Storm Petrel			No
Great Skua			No
Greater Black Backed Gull	Yes	By air*	No
Herring Gull			
Lesser Black Backed Gull	No**	By sea*	Yes**
Manx Shearwater	Not relevant	By sea*	No
Northern Fulmar	Not relevant	By sea*	No
Razorbill	Not relevant	By sea*	Yes

For the distance-decay and foraging range approach we consider four possible foraging metrics (mean-max, mean-max mean-max plus 1SD, mean-max plus 2SD, and max-max) in order to explore the sensitivity to the choice of metric. We also explore sensitivity to the value of the distance-decay rate parameter by implementing the approach with decay parameter values of 1.5, 2 and 2.5. We consider a fixed set of decay parameter values, rather than the estimated values outlined in the previous section, because of the

sensitivity of estimates to the assumptions that are made regarding the distribution very close to the colony (i.e. to the choice of value for the lower distance threshold).

Choice of footprints

We consider apportioning for each of the nine simulated footprints generated in the ORJIP Seabird Sensitivity Mapping Tool.

These footprints were created by selecting nine actual footprints that cover a range of geographical regions and sizes (Beatrice, Hornsea Project 3 (HOW03) Wind Farm, East Anglia Three Wind Farm, Dogger Bank A, Triton Knoll Wind Farm, Walney Extension (WOW03) Wind Farm, SeaGreen Alpha, Neart na Gaoithe Wind Farm and Islay Demo Zone). Each of these nine footprints was then transformed in such a way as to be sufficiently distinct from the actual footprint to ensure that the results relate to a hypothetical, rather than a real, footprint, by:

1. Shifting the entire footprint either north or south (selected at random) by 10km
2. Either increasing or reducing (selected at random) the size of the footprint by 50% (subject to a constraint that the hypothetical footprint cannot have an area of less than 10km²)
3. Changing the shape of the footprint to be square

Choice of SPAs

For each footprint we estimate the proportion of birds in each simulated footprint that are estimated to arise from each of the 118 CEF SPAs (the set of SPAs identified, in consultation with JNCC, as being relevant for the seabird species considered within the Marine Scotland CEF project) using each of the above methods. However, 13 of the 118 SPAs are Marine SPAs, and apportioning for these is assumed to be zero (as the methods considered here cannot currently be implemented for these SPAs, and a naïve implementation of apportioning for these SPAs would risk double counting). In practice, 105 SPAs are therefore considered.

Results

Table 8 shows the mean number of SPAs per project that have a non-zero proportion of birds apportioned to them, based on a foraging range approach (using each of the four different foraging range metrics) and, for species where it is possible, a GPS-based approach.

Note that the GPS-based approach could not be applied to all of the simulated footprints for all species. For guillemot and razorbill there were two simulated footprints in the southern North Sea that lay beyond the grid extent used for the GPS modelling, and for shag there were four simulated footprints (again in the southern North Sea): since the grid extends were based on the locations of colonies, and foraging ranges, these footprints would essentially have had zero overlap with the modelled distributions, so their exclusion from apportioning calculations is consistent with the underlying modelling and data. For lesser black-backed gull there were two simulated footprints, in the Irish Sea, that were excluded for a different reason - because the footprint size was sufficiently small, relative to the grid resolution, that no grid cells were classed as being within the footprint. This latter issue is an important limitation for LBBG, given the coarser grid resolution (10x10km) used for this species, but is only likely to be important for other species if very small footprints are considered (since the other species are mapped at 2x2km).

Table 8. Mean number of SPAs per project that have a non-zero proportion of birds apportioned to them, based on a foraging range approach (using each of the four different foraging range metrics) and, for species where it is possible, a GPS-based approach. For lesser black backed gull results are shown for all colonies included, and with inland colonies excluded. Dark grey cells indicate those for which apportioning was impossible (either because the GPS-based approach had not been applied for the species or Woodward et al., 2019, were not able to provide a standard deviation for the species). Light grey cells indicate those for which distance by air is used; for all other cells distance by sea was used.

Species	Distance decay and foraging range				GPS-based
	MeanMax	Max-Max	MeanMax+1SD	MeanMax+2SD	
All colonies included					
Atlantic Puffin	2.56	10.78	6.44	11.33	
Black Legged Kittiwake	4.56	34.78	9.67	17.67	11.67
Common Guillemot	1.11	10.22	3.89	5.78	11.89
European Shag	0.00	0.56	0.22	0.33	0.56
European Storm Petrel	2.56	2.56			
Greater Black Backed Gull	1.89	1.89			
Great Skua	8.56	29.78	27.67	31.00	
Herring Gull	1.22	3.56	3.22	4.78	
Lesser Black Backed Gull	4.22	38.00	10.11	22.67	
Manx Shearwater	9.89	10.00	10.00	10.00	
Northern Fulmar	28.44	66.00	63.22	66.00	
Northern Gannet	2.56	7.78	4.67	7.67	
Razorbill	2.22	9.56	4.78	6.33	10.78
Urban colonies excluded					
Lesser Black Backed Gull	2.11	12.22	4.78	8.22	15.67

It can be seen that the mean number of SPAs with non-zero apportioning estimates varies substantially between species, as might be expected, and is highest for the species with the highest foraging ranges. The mean number is also heavily dependent on the foraging range metric, and on whether a GPS-based approach is used (because this effectively again assumes a different foraging range metric). For European Shag there are no SPAs within the mean-max foraging range of any of the projects, so the mean is equal to zero.

Table 9 focuses on looking at the correlation between the non-zero apportioning proportions that are derived using different methods. The results suggest that the correlations are very high when comparing proportions obtained from different decay parameter values (always > 0.98), and relatively high when comparing the GPS-based approach against the distance decay approach with mean-max foraging range (always > 0.95). Results comparing the distance decay approach with mean-max foraging range with the same approach with other metrics are more variable. Correlations are high across different foraging metrics for many species, including kittiwake, guillemot and razorbill, but are much lower for other species, with the lowest levels of correlation being seen for northern gannet and northern fulmar. More detailed results are given in Appendix C.

Table 9. Correlation between apportioning proportions obtained using a distance-decay and foraging range approach with mean-max foraging range and a decay parameter of $\beta = -2$ and the equivalent apportioning proportions calculated using other apportioning methods. Correlations are calculated using all combinations of SPAs and simulated footprints for which the apportioning proportions are non-zero using both of the methods being correlated. Dark grey cells indicate those for which apportioning was impossible (either because the GPS-based approach had not been applied for the species, because Woodward et al., 2019, were not able to provide a standard deviation for the species, or because there were no SPAs within the mean-max foraging range for any project). Light grey cells indicate those for which distance by air is used; for all other cells distance by sea was used.

Species	Distance decay and foraging range					GPS-based
	Max-Max	MeanMax +1SD	MeanMax +2SD	Mean-Max		
	$\beta = -2$			$\beta = -1.5$	$\beta = -2.5$	
<i>All colonies excluded</i>						
Atlantic Puffin	0.806	0.932	0.807	0.996	0.996	
Black Legged Kittiwake	0.997	0.999	0.998	0.996	0.997	0.975
Common Guillemot	0.989	0.988	0.992	0.998	0.998	0.960
European Shag						
European Storm Petrel	1.000			0.992	0.994	
Greater Black Backed Gull	1.000			0.987	0.988	
Great Skua	0.835	0.836	0.835	0.995	0.997	
Herring Gull	1.000	1.000	0.999	0.999	1.000	
Lesser Black Backed Gull	0.843	0.896	0.857	0.995	0.997	
Manx Shearwater	1.000	1.000	1.000	0.994	0.996	
Northern Fulmar	0.765	0.770	0.765	0.964	0.985	
Northern Gannet	0.767	0.834	0.767	0.998	0.999	
Razorbill	0.998	0.997	0.998	0.994	0.996	0.958
<i>Urban colonies excluded</i>						
Lesser Black Backed Gull	0.887	0.936	0.909	0.994	0.996	0.942

6. Conclusions

Within WP4 we have developed a new user-friendly ANBS tool for apportioning of razorbill and guillemot in the non-breeding season, and have integrated new functionality on breeding season apportioning, particularly of lesser black backed gulls, into the CEF. This work package has therefore translated the results of the analyses of [WP3](#) into practical tools. We have highlighted key areas of future work, including the potential for future integration of the new ANBS non-breeding season tool into the CEF. We have also compared the results obtained using different apportioning methods within the context of hypothetical footprints, building on the comparisons undertaken in [WP3](#). The key feature of the breeding season comparisons is the importance of the foraging range, with the results showing substantial sensitivity to the choice of foraging range metric.

7. References

- Balmer, D.E., Gillings, S., Caffrey, B.J., Swann, R. L., Downie, I.S. & Fuller, R.J. (2013) Bird Atlas 2007–11: the breeding and wintering birds of Britain and Ireland. BTO Books, Thetford.
- Booth Jones, K., Thaxter, C., Clewley, G., Wolsey, S., Calbrade, N., Atkinson, P., Calladine, J. & Burton, N. (2022) Belfast's urban gulls: an assessment of breeding populations, breeding season movements and winter population. BTO Research Report 734, BTO, Thetford, UK.
- Burnell, D (2021a). Urban nesting Herring Gull *Larus argentatus* and Lesser Black-backed Gull *Larus fuscus* population estimates: devising species-specific correction models for ground-based survey data. Natural England publication ref: JNCC21_01
- Burnell, D. (2021b). Population estimates for urban and natural nesting Herring Gull *Larus argentatus* and Lesser Black-backed Gull *Larus fuscus* in England. Natural England publication ref: JNCC21_02
- Butler, A., Carroll, M., Searle, K.R., Bolton, M., Waggitt, J., Evans, P., Rehfish, M., Goddard, B., Brewer, M., Burthe, S. and Daunt, F. (2020). Attributing seabirds at sea to appropriate breeding colonies and populations. Scottish Marine and Freshwater Science Vol 11 No 8. DOI: 10.7489/2006-1
- 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 Report, 164.
- Langley, L.P., Bearhop, S., Burton, N.H.K., Banks, A.N., Frayling, T., Thaxter, C.B., Clewley, G.D., Scragg, E. & Votier, S.C. (2021) GPS tracking reveals landfill closures induce higher foraging effort and habitat switching in gulls. *Movement Ecology*, 9:56
- Langley, L.P., Bearhop, S., Burton, N.H.K., Banks, A.N., Frayling, T., Thaxter, C.B., Clewley, G.D., Scragg, E. & Votier, S.C. (2023) Urban and coastal breeding lesser black-backed gulls (*Larus fuscus*) segregate by foraging habitat. *Ibis*, 165, 1, 214-230.
- Mitchell, P.I., Newton, S.F., Ratcliffe, N. & Dunn, T.E. 2004. Seabird Populations of Britain and Ireland. T. & A.D. Poyser, London.
- Morton, D., Rowland, C., Wood, C. Meek, L., Marston, C., Smith, G., Wadsworth, R., Simpson, I.C. 2011. Final Report for LCM2007 - the new UK land cover map. Countryside Survey Technical Report No 11/07 NERC/Centre for Ecology & Hydrology (CEH Project Number: C03259).
- Ross-Smith, V.H., Robinson, R.A., Banks, A.N., Frayling, T.D., Gibson, C.C. & Clark, J.A. (2014) The Lesser Black-backed Gull *Larus fuscus* in England: how to resolve a conservation conundrum. *Seabird*, 27, 41-61.
- Spelt, A., Williamson, C., Shamoun-Baranes, J., Shepard, E., Rock, P. & Windsor, S. (2019) Habitat use of urban-nesting lesser black-backed gulls during the breeding season. *Nature Scientific Reports*, 9, 10527.
- Thaxter, C.B., Horswill, C., Ross, K.E., Austin, G.E., Balmer, D.E. & Burton, N.H.K. (2017) Urban Breeding Gull Surveys: A Survey Design Simulation, BTO Research Report No. 699
- Thaxter, C.B., Quinn, L., Atkinson, P., Booth Jones K.A., Burton, N.H.K., Clark, N.A., Clewley, G.D., Johnston, D.T., Masden, E.A., Ross-Smith, V.H., Sage, E., Scragg, E., Taylor, R. & Humphreys, E.M. (submitted)

Importance of breeding and foraging habitat in determining foraging ranges of sympatric generalist species.

Wakefield, E.D., Owen, E., Baer, J., Carroll, M.J., Daunt, F., Dodd, S.G., Green, J.A., Guilford, T., Mavor, R.A., Miller, P.I., Newell, M.A., Newton, S.F., Robertson, G.S., Shoji, A., Soanes, L.M., Votier, S.C., Wanless, S. & Bolton, M. (2017) Breeding density, fine-scale tracking, and large-scale spatial modeling reveal the regional distribution of four seabird species. *Ecological Applications*, 27, 7, 2074-2091.

Woodward, I., Thaxter, C.B., Owen, E. & Cook, A.S.C.P. 2019. Desk-based revision of seabird foraging ranges used for HRA screening, Report of work carried out by the British Trust for Ornithology on behalf of NIRAS and The Crown Estate, ISBN 978-1-912642-12-0.

8. Appendices

Appendix A. Extended distance decay model

The extended version of the NatureScot model is based on the assumption that the probability of a randomly selected bird from a particular colony being at any particular point in the sea (x, y) is, in the absence of land, proportional to d^β , where $d = \sqrt{x^2 + y^2}$ denotes the distance between this point and the colony (which we can assume without loss of generality has location $(0,0)$). This implies that there is a probability density function of the form $f_{XY}(x, y) \propto d^\beta$.

We assume that the density is zero for distances in excess of the maximum foraging range, D . We will also need to assume that it is zero for distance below a distance $D\varepsilon$, to avoid the distribution including distances of zero (which would, according to this model, typically have an infinite density, something that makes the model intractable, as well as being biologically implausible). This implies $f_{XY}(x, y) \propto d^\beta I(D\varepsilon < d < D)$, where $I(\cdot)$ is the indicator function.

It is more natural, and will assist with further calculations, in transforming the distribution from cartesian coordinates into polar coordinates, so that we consider the distribution in relation to distance from colony d and angle from colony θ rather than in terms of geographical location. The standard conversion to polar coordinates gives us

$$f_{d\theta}(d, \theta) = df_{XY}(x, y) \propto d^{\beta+1} I(D\varepsilon < d < D)$$

We convert the distance d into a relative distance $r = \frac{d}{D}$, that indicates the ratio of distance d to the maximum foraging range D beyond which the density is assumed to be zero. This implies

$$f_{r\theta}(r, \theta) = rDf_{XY}(x, y) \propto r^{\beta+1} I(\varepsilon < r < 1)$$

By recognition, we can see that the PDF on the right hand side corresponds to a PDF in which:

1. Angle and distance from colony are independent, so that

$$f_{r\theta}(r, \theta) = f_r(r)f_{d\theta}(d, \theta)$$
2. The angle from colony θ is uniformly distributed over the range $(0, 2\pi)$, so that $f_\theta(\theta) = 1/2\pi$, implying that the model does not contain any directional effects;
3. The ratio of distance to colony to maximum foraging range, r , has a bounded Pareto distribution over the range $(\varepsilon, 1)$, with shape parameter $\alpha = -(\beta + 2)$, so that

$$f_r(r) = \frac{\alpha \varepsilon^\alpha r^{-\alpha-1}}{1 - \varepsilon^\alpha} = \frac{-(\beta + 2) \varepsilon^{-(\beta+2)} r^{\beta+1}}{1 - \varepsilon^{(\beta+2)}}$$

This therefore provides a probabilistic model that is consistent (in the absence of land) with the extended version of the NatureScot model.

Appendix B. Decay parameter estimation: SPA-level results

Table B1. Mean and maximum foraging ranges, the ratio of the mean to the maximum, and associated estimates of the decay parameter β , based on SPA-level foraging range information from the CEF Data Store. “Max” and “mean” foraging ranges are both based on population-specific values. Decay parameter value estimates use a lower threshold ε parameter value of 0.00001; bird densities are assumed to be zero when the ratio of distance to max foraging range is less than this threshold. Grey cells denote situations in which estimation of the decay parameter was impossible (e.g. due to the mean and maximum being equal, a situation that is not consistent with the model assumptions).

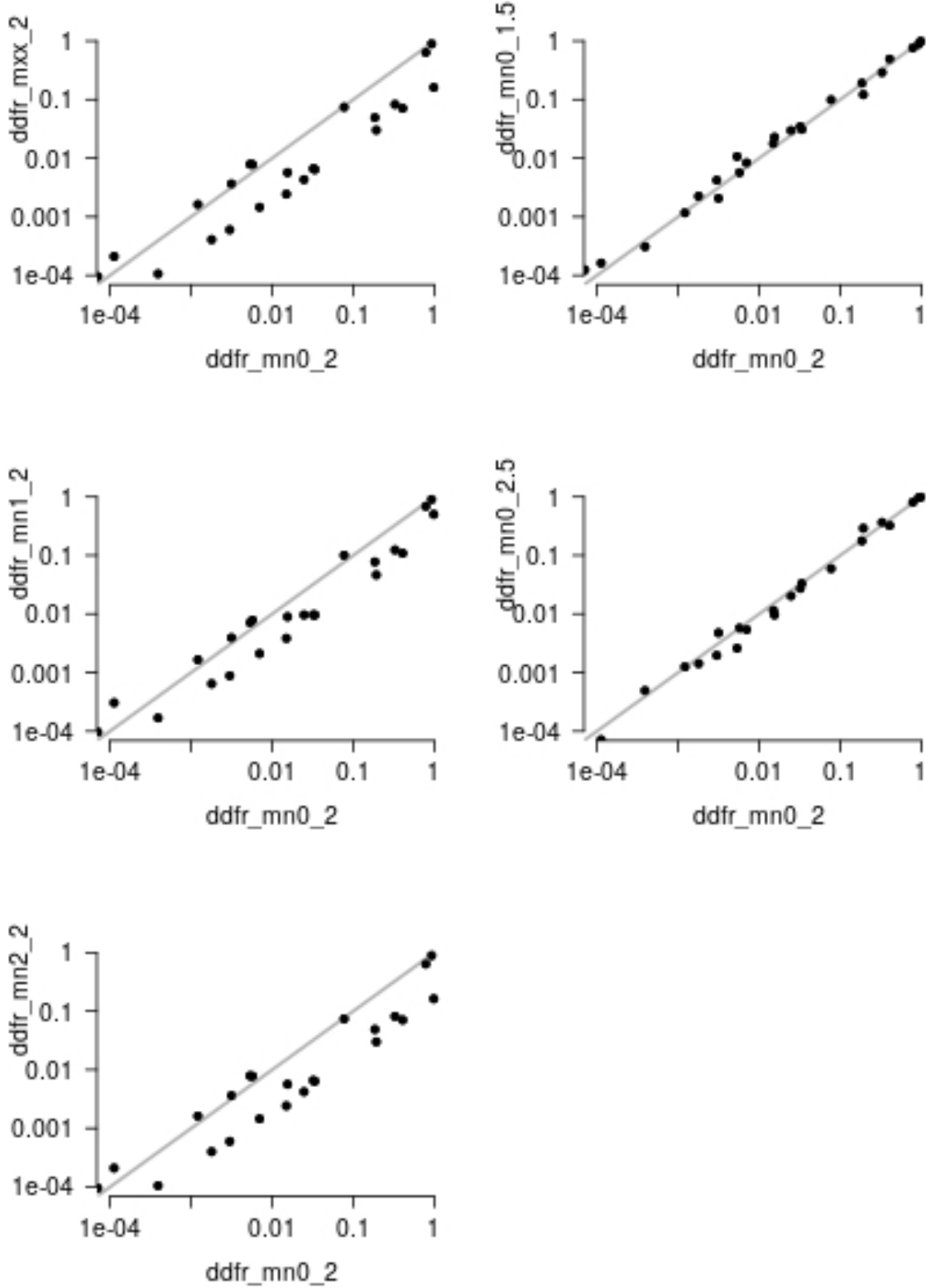
Receptor	SITE_CODE	Subsite name	Max	Mean	Mean: Max ratio	Estimate of β
Arctic Tern	UK9020291	Big Copeland; Co Down	39.6	5.8	0.15	-1.87
	UK9020271	Cockle Island; Groomsport; Co Down	46	15.7	0.34	-1.48
	UK9006031	Coquet Island	36	4.3	0.12	-1.93
	UK9013061	Skerries - Anglesey	29	8.1	0.28	-1.62
Atlantic Puffin	UK9002091	Fair Isle	241.7	106.5	0.44	-1.21
	UK9002061	Foula	95.4	69.5	0.73	0.68
	UK9002011	Hermaness; Shetland	383.3	89.5	0.23	-1.71
	UK9004171	Isle of May	65.5	42.6	0.65	-0.14
	UK9001041	The Shiant Isles	45.7	23.4	0.51	-0.95
Black-legged Kittiwake	UK9006101	Bempton Cliffs	226.9	86.2	0.38	-1.39
	UK9002491	Bullers of Buchan	81.9	59.4	0.73	0.64
	UK9001231	Cape Wrath	38	25.7	0.68	0.09
	UK9003171	Colonsay	144.5	49.8	0.34	-1.48
	UK9002151	Copinsay; Orkney	227.8	55.2	0.24	-1.69
	UK9002091	Fair Isle	183.4	44	0.24	-1.70
	UK9006021	Farne Islands	111.2	35.6	0.32	-1.53
	UK9006101	Filey	212.4	118.1	0.56	-0.75
	UK9006101	Flamborough Head	316.9	199.6	0.63	-0.30
	UK9002271	Fowlsheugh	172.3	97	0.56	-0.71
	UK9020011	Rathlin Island	76	37.1	0.49	-1.05
	UK9004271	St Abbs Head	108.7	82.4	0.76	1.13
	UK9002511	Sumburgh Head; Shetland	40	5	0.13	-1.92
	UK9002491	Whinnyfold	88.1	54.5	0.62	-0.38
Common Guillemot	UK9002151	Copinsay; Orkney	26.6	11.1	0.42	-1.28
	UK9002091	Fair Isle	338.4	145.4	0.43	-1.25
	UK9002271	Fowlsheugh	44.2	32.4	0.73	0.75
	UK9004171	Isle of May	65.1	11.5	0.18	-1.82
	UK9002511	Sumburgh Head; Shetland	9.4	2.9	0.31	-1.56
	UK9001041	The Shiant Isles	7	7	1.00	
Common Tern	UK9009031	Blakeney Point	9	2.3	0.26	-1.67

	UK9020042	Blue Circle Island; Larne Lough	30.4	4	0.13	-1.90
	UK9006031	Coquet Island	17.9	4.1	0.23	-1.72
	UK9003211	Glas Eileanan	13.7	3.2	0.23	-1.71
European Shag	UK9020288	Annet	9.6	5.9	0.61	-0.41
	UK9002091	Fair Isle	7.6	3.8	0.50	-1.00
	UK9006021	Farne Islands	14.1	3.2	0.23	-1.72
	UK9004171	Isle of May	17	9.7	0.57	-0.67
	UK9020288	Little Ganninick	3.9	1	0.26	-1.66
	UK9020288	Sansom	5.5	3.9	0.71	0.44
Great Skua	UK9002061	Foula	219	85.6	0.39	-1.36
Great Skua	UK9002141	Hoy	108	41	0.38	-1.39
Herring Gull	UK9020326	Walney Island	83.7	21.8	0.26	-1.66
Lesser Black-backed Gull	UK9009112	Havergate	22.5	17.1	0.76	1.17
	UK9009112	Orford Ness	124	49.9	0.40	-1.33
	UK9005103	Ribble	111.9	52.5	0.47	-1.12
	UK9014051	Skokholm	150.5	74	0.49	-1.03
	UK9020326	Walney Island	92.7	44.2	0.48	-1.09
Little Tern	UK9009271	Scroby Sands	5	3.5	0.70	0.33
Manx Shearwater	UK9013121	Bardsey Island	162.1	18.5	0.11	-1.94
	UK9014051	Skomer	1250	190.9	0.15	-1.86
Northern Fulmar	UK9002491	Bullers of Buchan	223.6	97.4	0.44	-1.23
	UK9002151	Copinsay; Orkney	479.5	154.3	0.32	-1.53
	UK9002371	Eynhallow; Orkney	2736	324.4	0.12	-1.93
	UK9002091	Fair Isle	246.8	109.7	0.44	-1.20
	UK9002061	Foula	120	35	0.29	-1.59
	UK9002491	Whinnyfold	108.4	108.4	1.00	
Northern Gannet	UK9003091	Ailsa Craig	296	159.5	0.54	-0.83
	UK9004171	Bass Rock	590	206.7	0.35	-1.46
	UK9006101	Bempton Cliffs	404.4	43.3	0.11	-1.95
	UK9014041	Grassholm	516.7	160.6	0.31	-1.55
Razorbill	UK9002091	Fair Isle	312.9	152.2	0.49	-1.05
	UK9001021	Flannans	92.2	50.8	0.55	-0.77
	UK9004171	Isle of May	52	18.4	0.35	-1.45
	UK9020011	Rathlin Island	74.4	74.4	1.00	
	UK9014051	Skomer	62	27.4	0.44	-1.21
	UK9001041	The Shiant Isles	36	29.1	0.81	2.22
Red-throated Diver	UK9002311	Orkneys	9	4.5	0.50	-1.00
Roseate Tern	UK9006031	Coquet Island	10.8	3.2	0.30	-1.58
Sandwich Tern	UK9020042	Blue Circle Island; Larne Lough	17.2	6	0.35	-1.47
	UK9006031	Coquet Island	27.6	5.3	0.19	-1.79
	UK9002221	Sands of Forvie (Ythan Estuary)	22.9	8.4	0.37	-1.42
	UK9009031	Scolt Head	54	11.1	0.21	-1.76

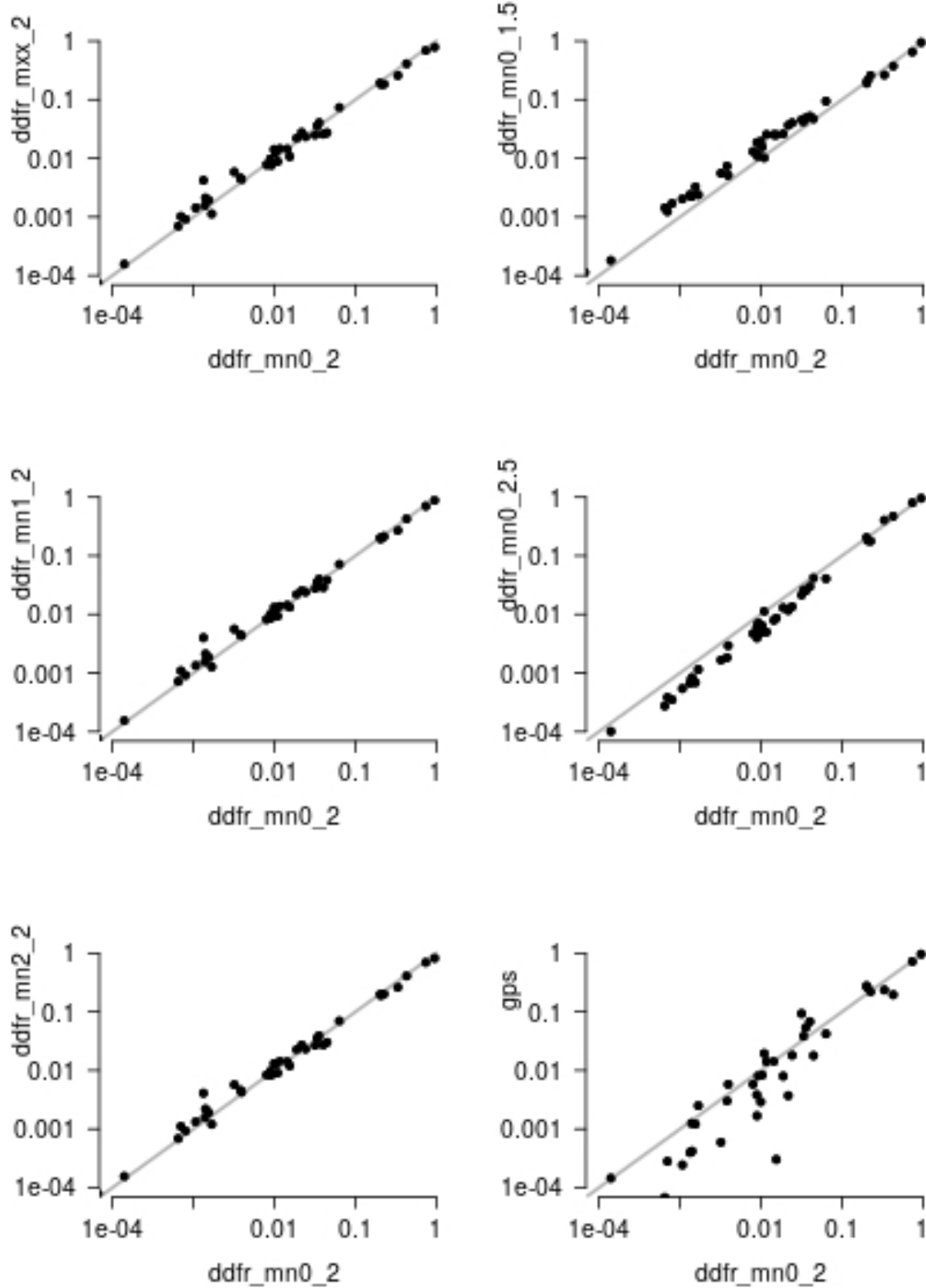
Appendix C. Full results of comparison of breeding season apportioning methods

Figure C1. Detailed comparison of apportioning results obtained using different methods for each species. Plots show the comparison of non-zero apportioning results obtained using the distance-decay approach with mean-max foraging range and a decay parameter of $\beta = 2$ against those obtained using: (a) the same approach with different foraging range metrics, (b) the same approach with different decay parameter values and (c) the GPS-based approach. The method use is shown on each case as either “ddfr” (distance-decay and foraging range approach) or “gps” (GPS-based approach). In the former case, this is followed by an underscore and the foraging range metric used, and then followed by a second underscore and the value of the decay parameter. Foraging range metrics are abbreviated as: “mmx” = max-max foraging range, “mn1” = mean-max plus 1 SD foraging range, “mn2” = mean-max plus 2 SD foraging range.

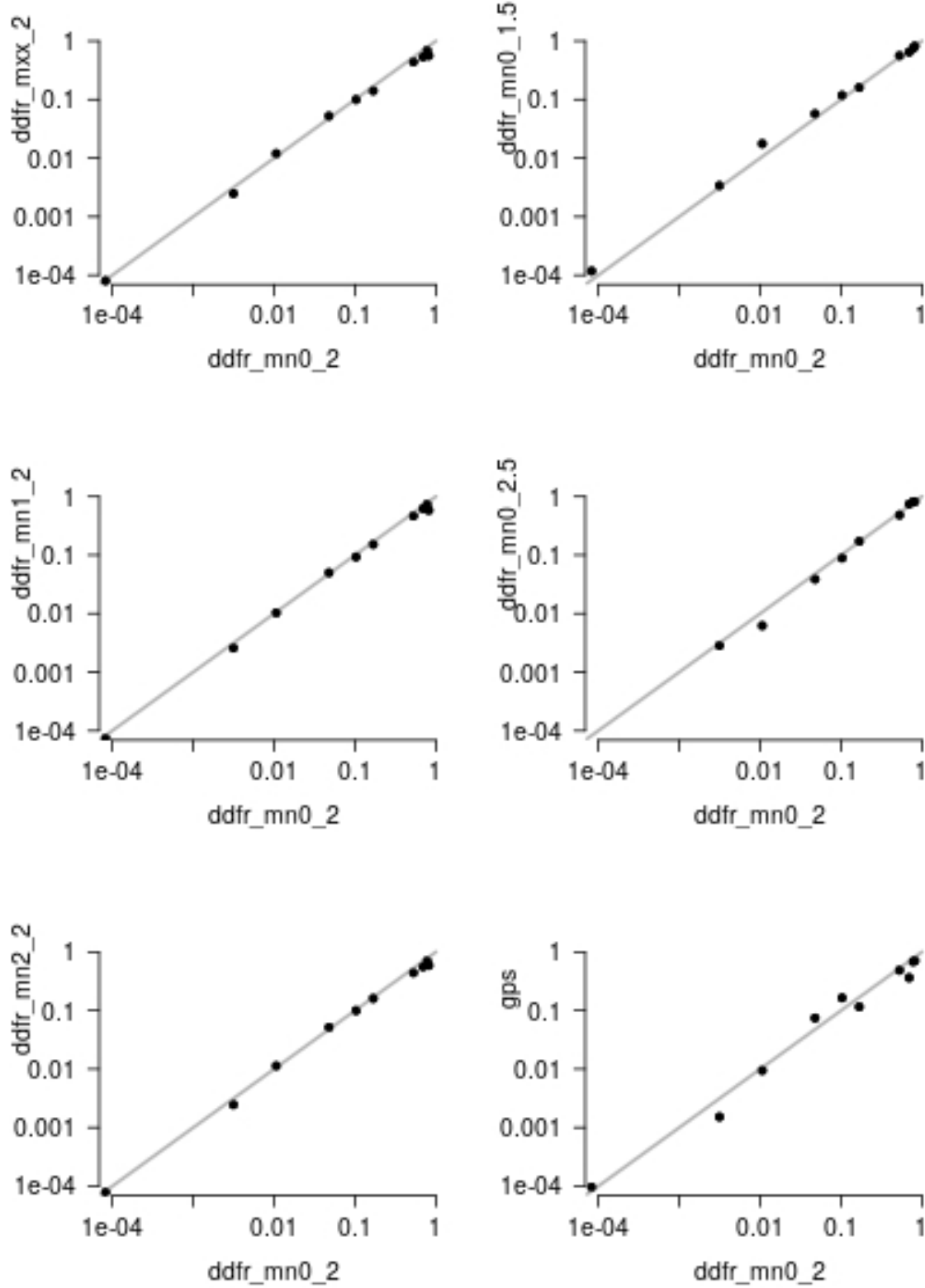
AtlanticPuffin



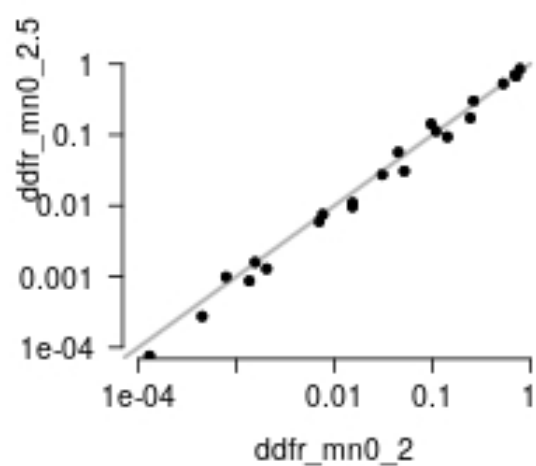
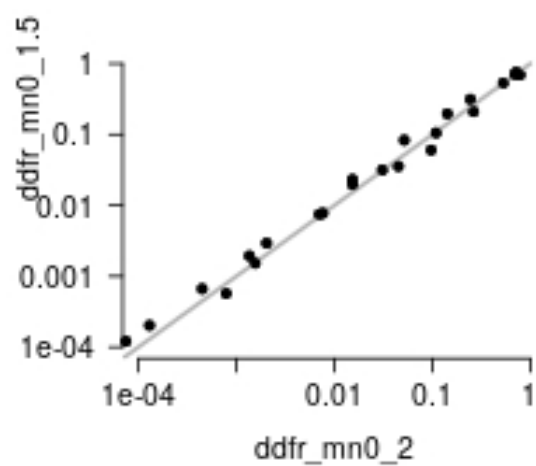
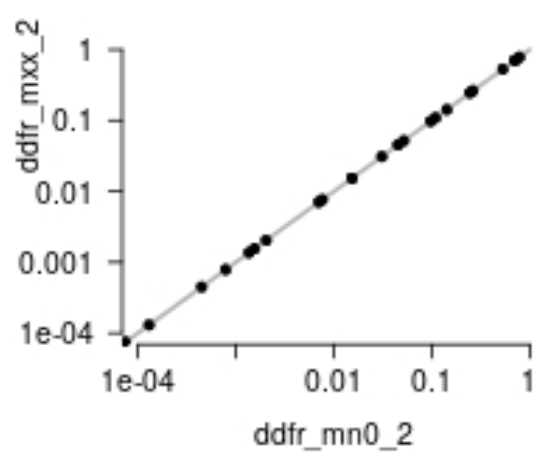
BlackLeggedKittiwake



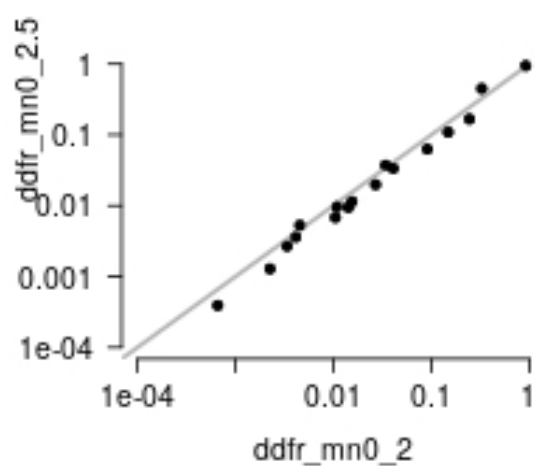
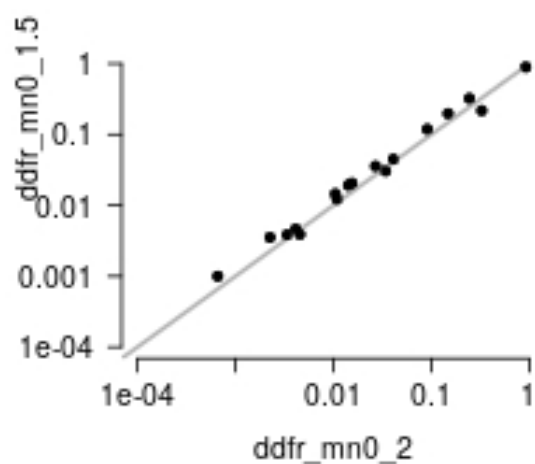
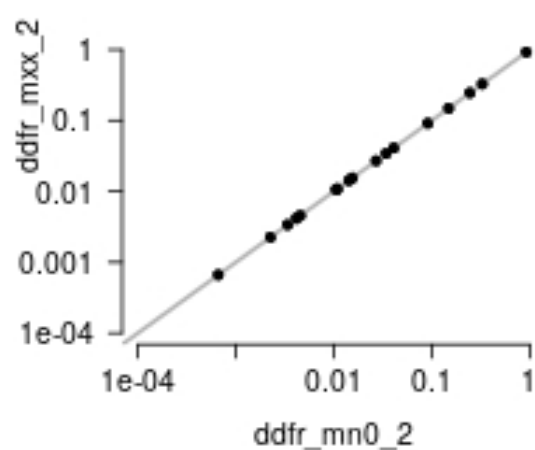
CommonGuillemot



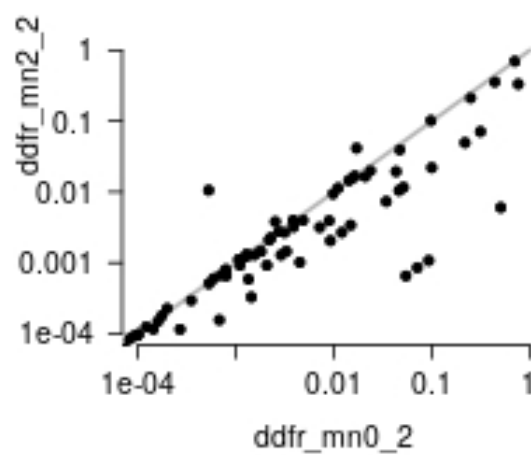
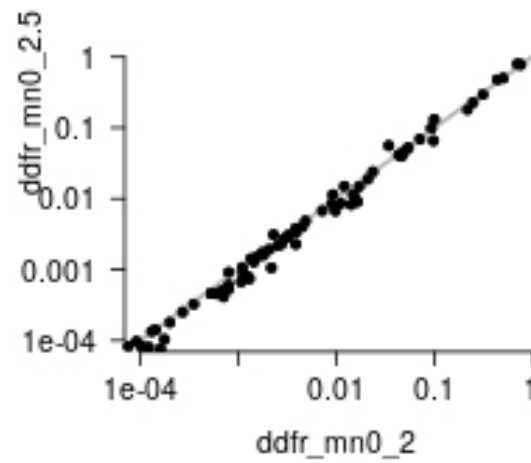
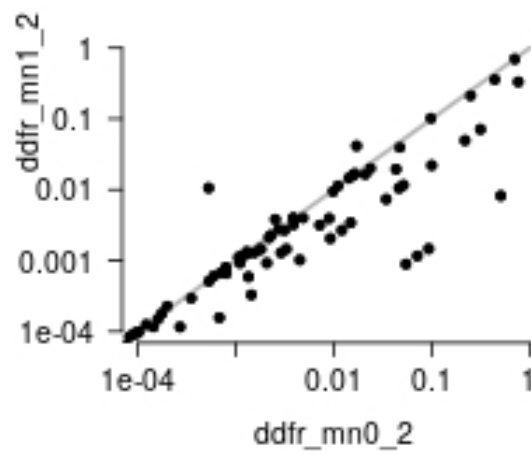
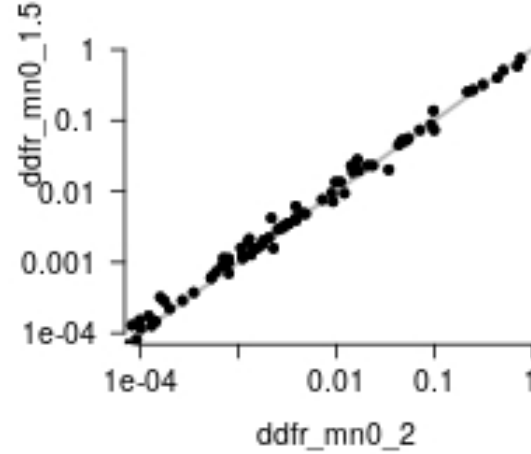
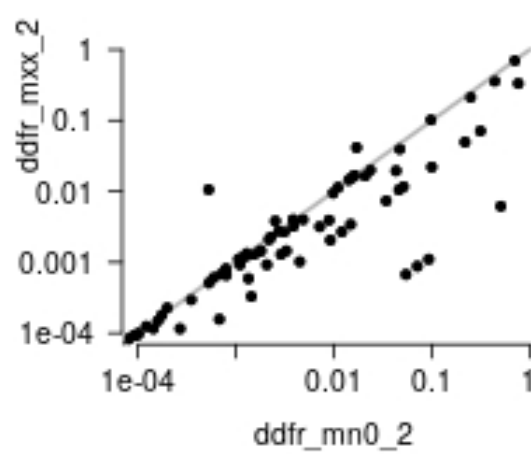
EuropeanStormPetrel



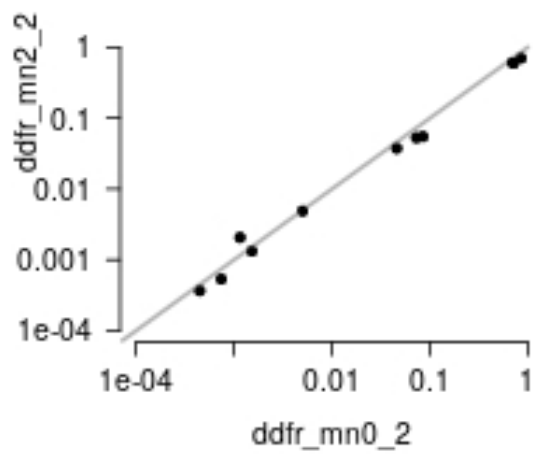
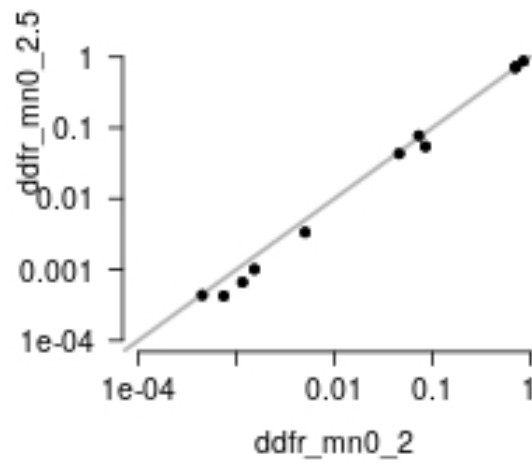
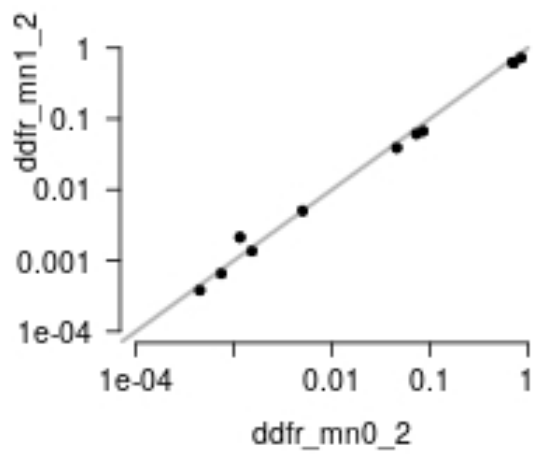
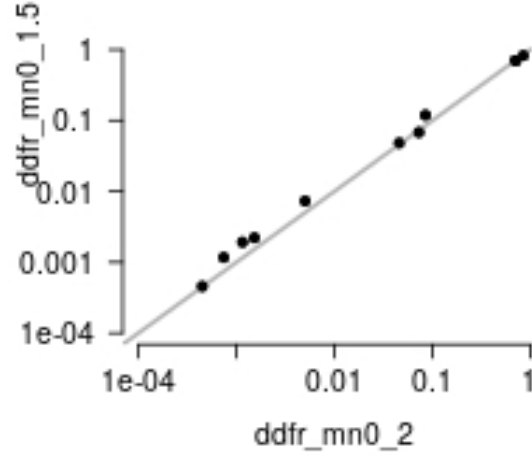
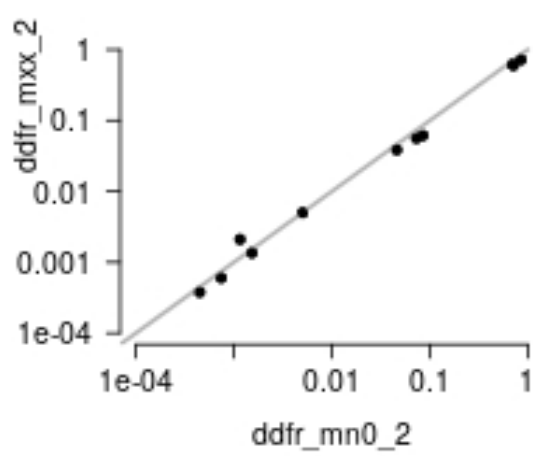
GreaterBlackBackedGull



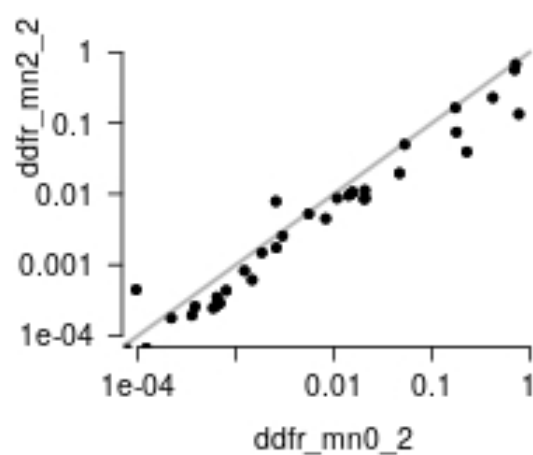
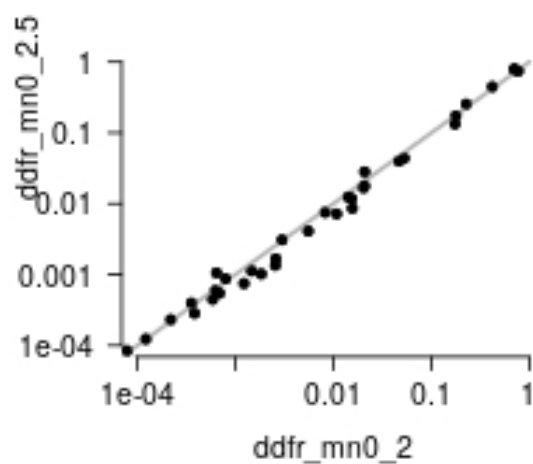
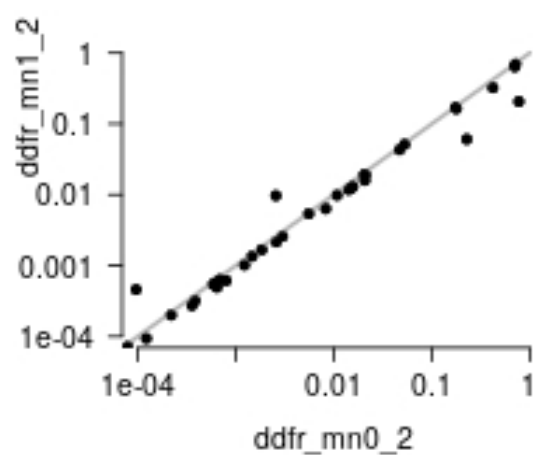
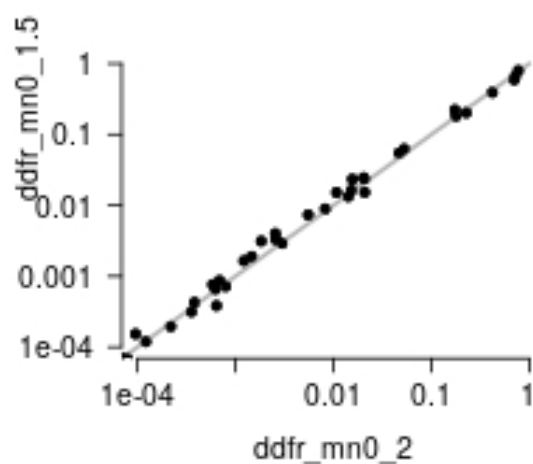
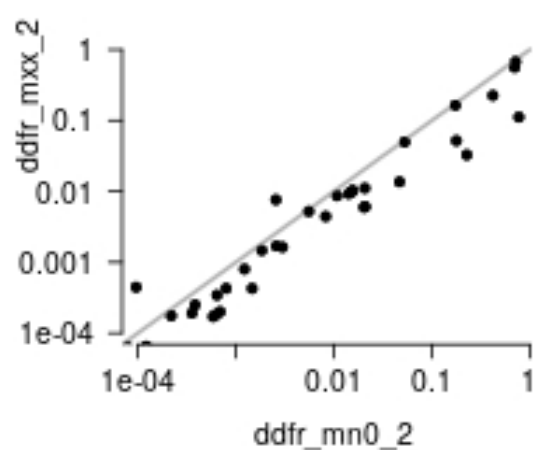
GreatSkua



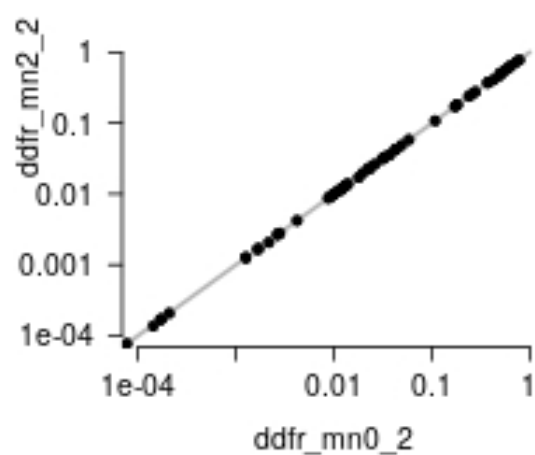
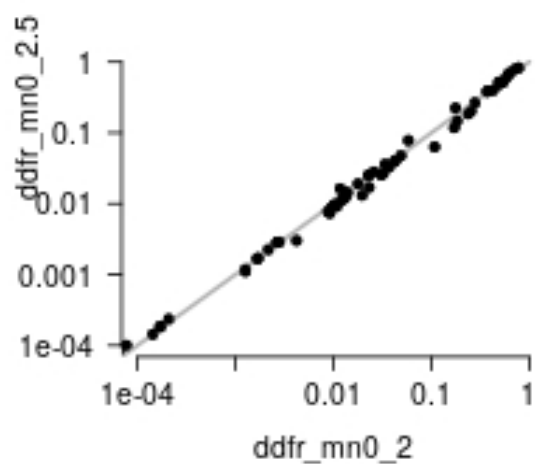
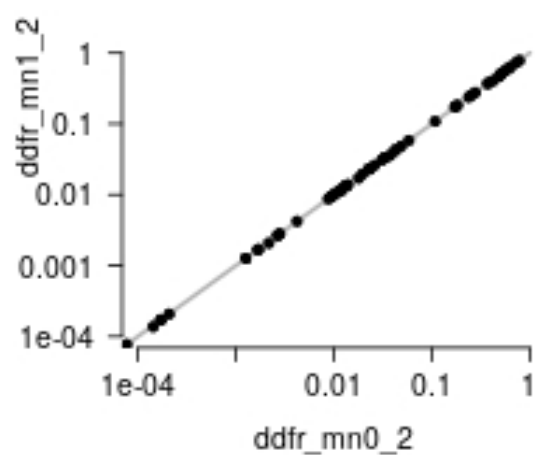
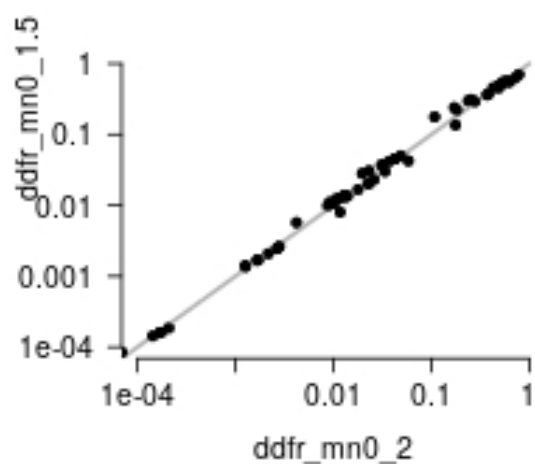
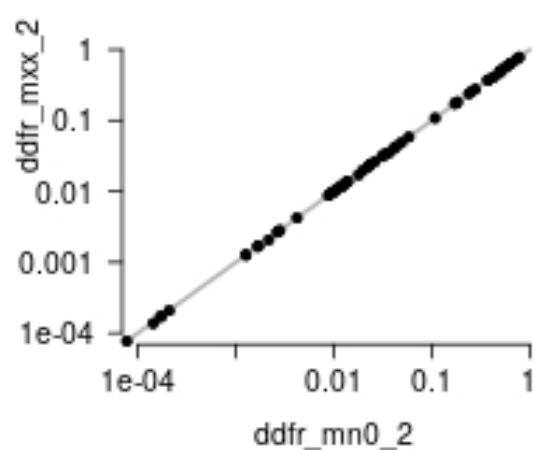
HerringGull



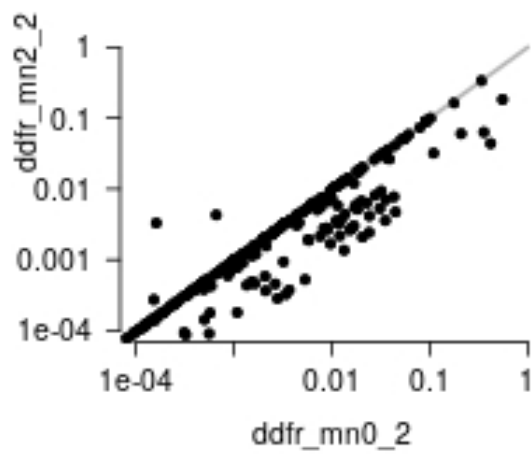
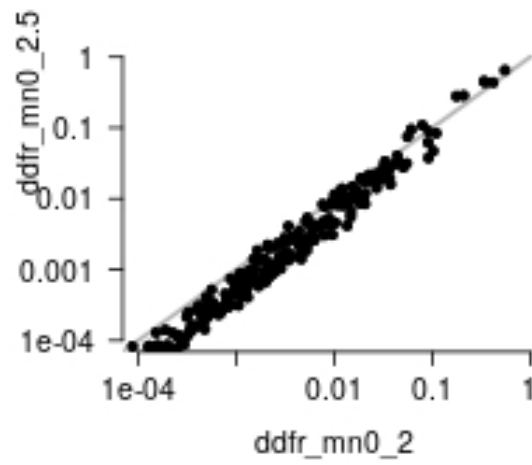
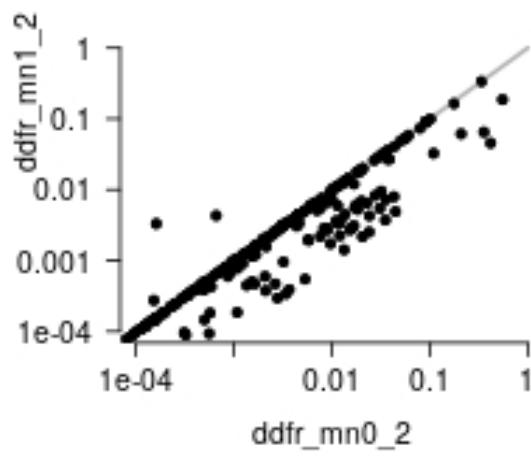
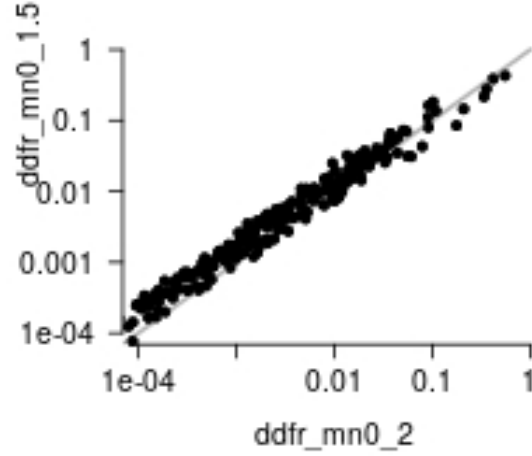
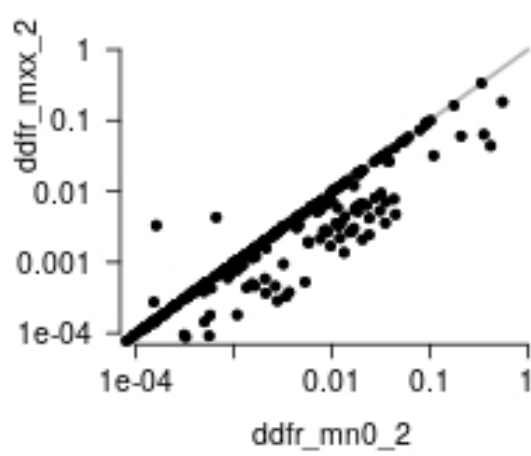
LesserBlackBackedGull



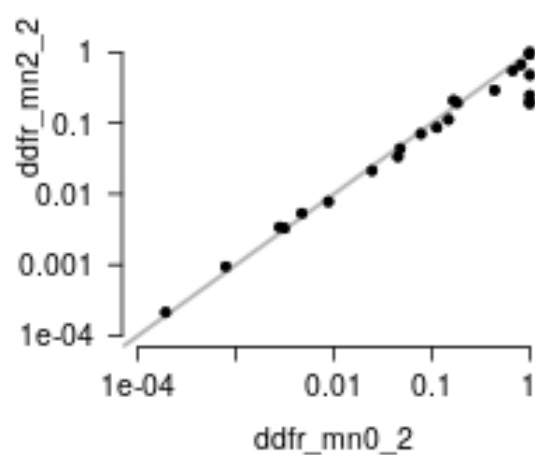
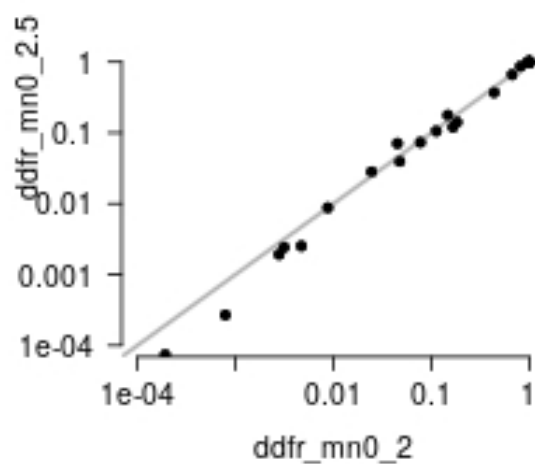
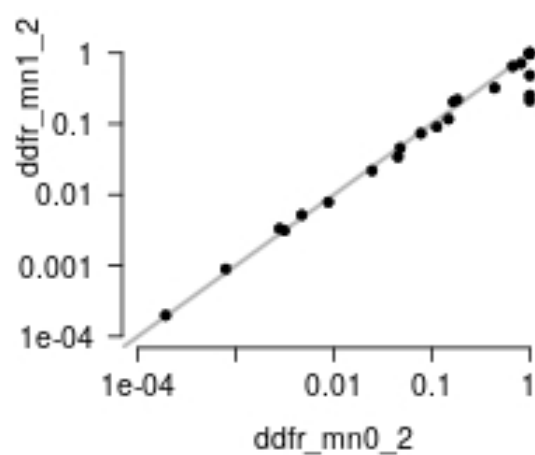
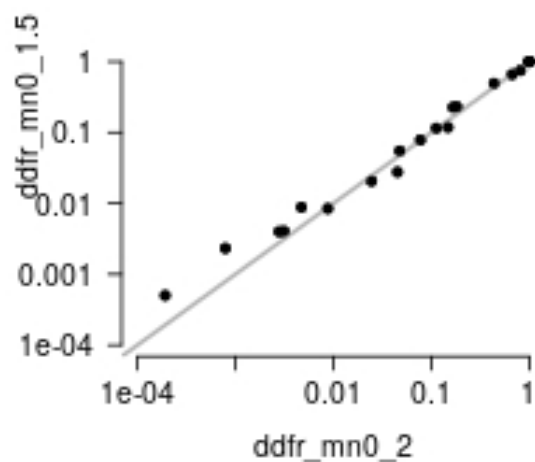
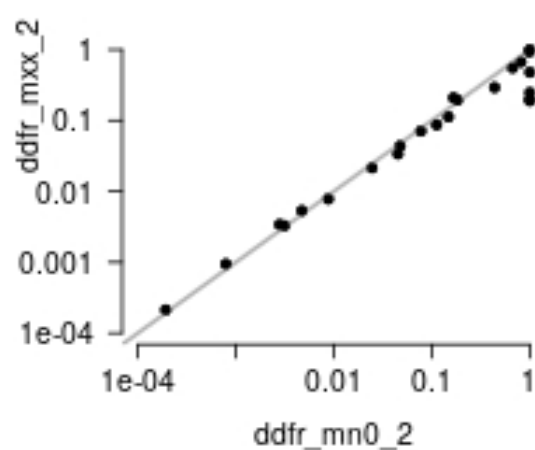
ManxShearwater



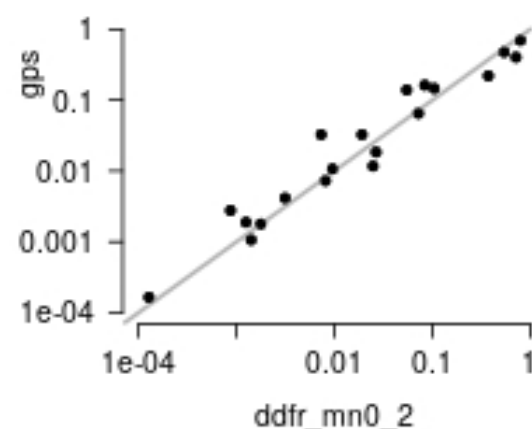
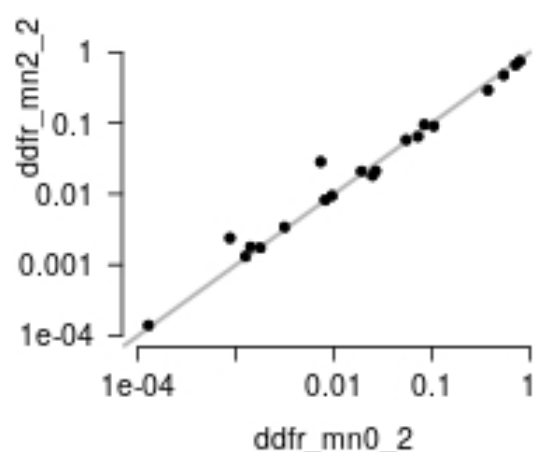
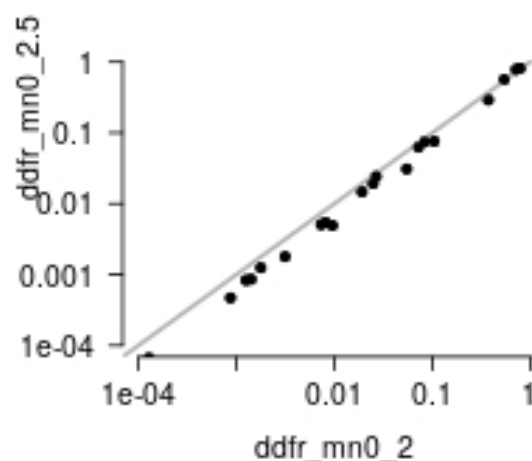
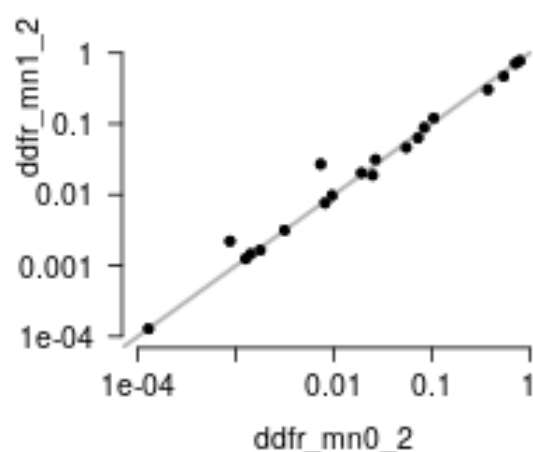
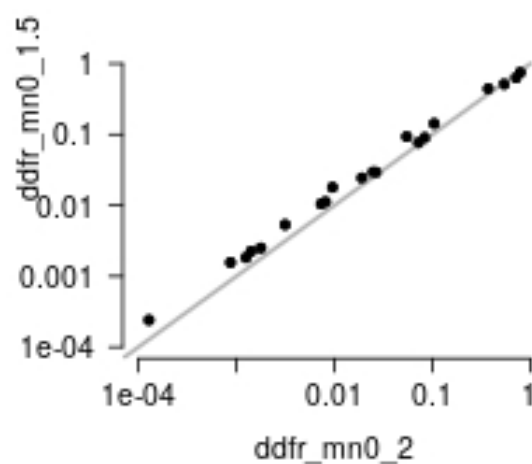
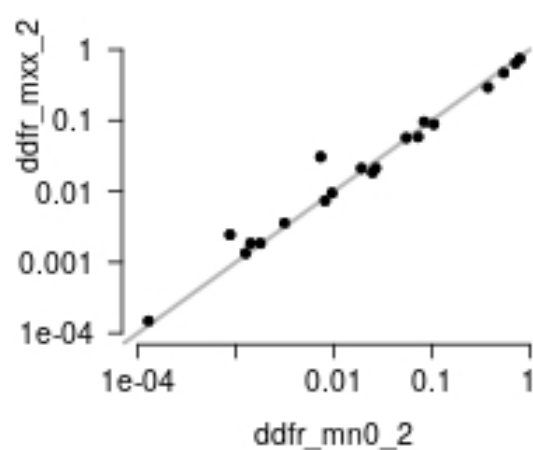
NorthernFulmar



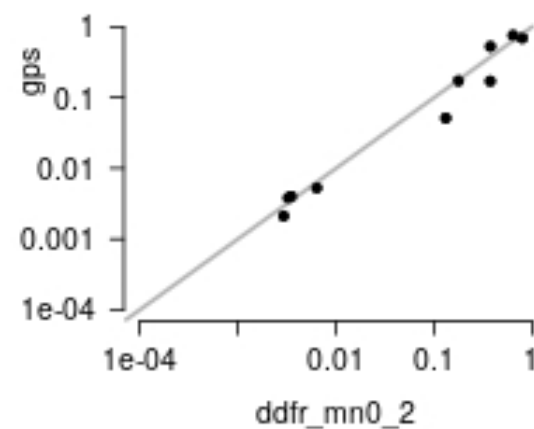
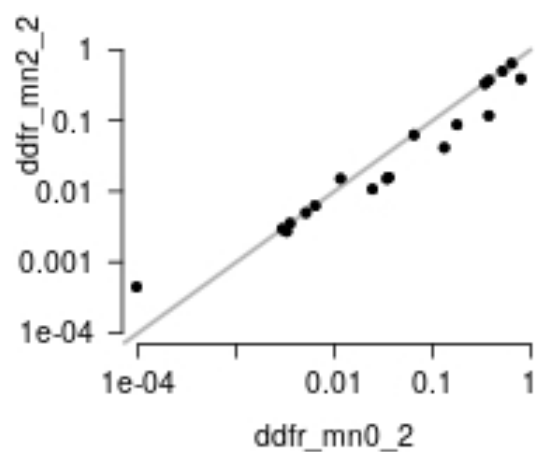
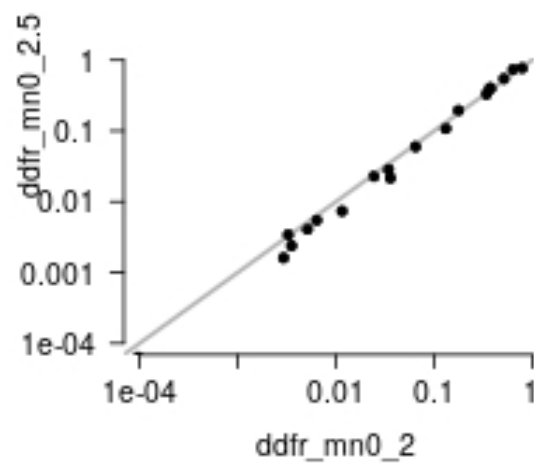
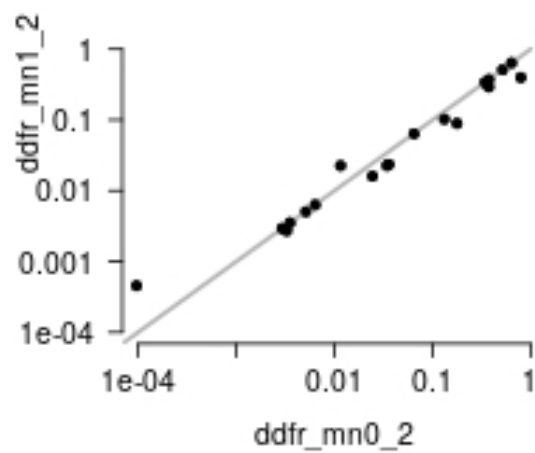
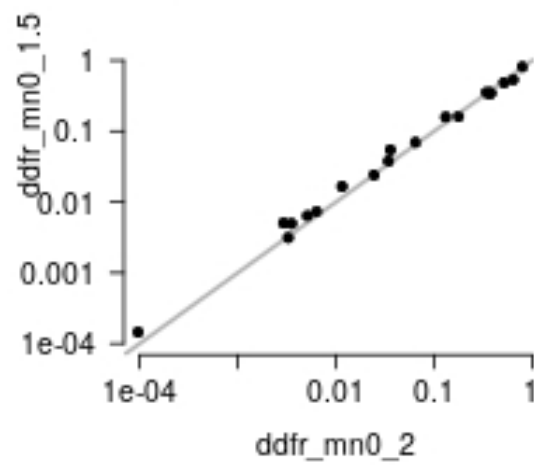
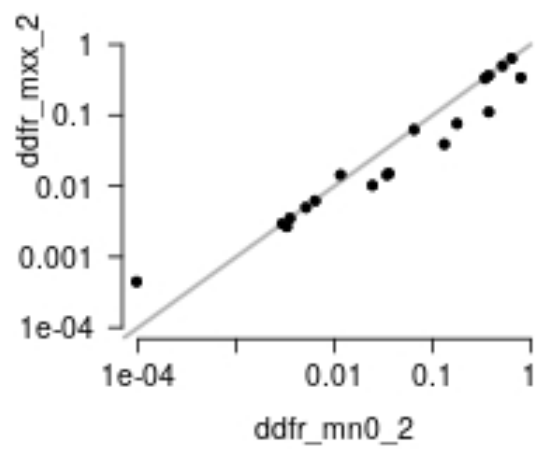
NorthernGannet



Razorbill



LBBG(InlandExcluded)



carbontrust.com

+44 (0) 20 7170 7000

Whilst reasonable steps have been taken to ensure that the information contained within this publication is correct, the authors, the Carbon Trust, its agents, contractors and sub-contractors give no warranty and make no representation as to its accuracy and accept no liability for any errors or omissions. Any trademarks, service marks or logos used in this publication, and copyright in it, are the property of the Carbon Trust. Nothing in this publication shall be construed as granting any licence or right to use or reproduce any of the trademarks, service marks, logos, copyright or any proprietary information in any way without the Carbon Trust's prior written permission. The Carbon Trust enforces infringements of its intellectual property rights to the full extent permitted by law.

The Carbon Trust is a company limited by guarantee and registered in England and Wales under Company number 4190230 with its Registered Office at: Level 5, Arbor, 255 Blackfriars road, London SE1 9AX.

© The Carbon Trust 2024. All rights reserved.

Published in the UK: 2024