User documentation for the Stochastic Collision Risk Assessment for Movement (SCRAM)

Developed by:

Biodiversity Research Institute

The University of Rhode Island

U. S. Fish and Wildlife Service

With funding from:

The Bureau of Ocean Energy Management

10 June 2022

Disclaimers

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government. The findings and conclusions in this article are those of the author(s) and do not necessarily represent the views of the U.S. Fish and Wildlife Service.

Acknowledgements

This study was funded in part by the U.S. Department of the Interior, Bureau of Ocean Energy Management, Office of Renewable Energy Program, Washington, DC, through Inter-Agency Agreement Number M19PG00023 with the Fish and Wildlife Service (USFWS).

Add anyone who provides beta feedback in 2022

For More Information

The SCRAM Tool for which this User Manual was written is available at: https://briloon.shinyapps.io/SCRAM/. For more information on the tool or provide comments, contact Andrew Gilbert at the Biodiversity Research Institute (Andrew.gilbert@briwildlife.org). The R code for SCRAM is provided at the SCRAM GitHub repository: https://github.com/Biodiversity-Research-Institute/SCRAM. Update request and bugs can be posted post at https://github.com/Biodiversity-Research-Institute/SCRAM/issues.

Citation

Gilbert, A. T., Adams, E. M., Loring, P., William, K. 2022. User documentation for the Stochastic Collision Risk Assessment for Movement (SCRAM). Available at https://briloon.shinyapps.io/SCRAM/. XXX pp.

Contents

Overview

This is a user guide to an online web tool that provides access to a model that simulates collision risk to birds from existing or planned offshore wind energy development in the eastern United States. The underlying model is adapted from the widely used framework developed by Band (2012), which is often referred to as a collision risk model (CRM). This CRM for the eastern U.S. is run in the open source computing software R (R Core Team 2021) using code adapted from Masden (2015) and Trinder (2017). There is also a user interface for running the CRM in a web browser, similar to the application developed by Marine Scotland (McGregor et al. 2018), but the results are specific to the eastern United States. Collectively, we are calling this adaptation of the CRM and user interface the Stochastic Collision Risk Assessment for Movement (SCRAM).

This user manual for SCRAM has been developed to communicate the basics of the model and to guide users in its execution via the user interface.

What is the goal of SCRAM?

The model and web application have been initially implemented for three birds that are often the focus of species impact assessments in the western Atlantic: Roseate Tern (*Sterna dougallii*), Red Knot (*Calidris canatus*), and Piping Plover (*Charadrius melodus*). While SCRAM can be run for any species for which a user has data, here we provide habitat-use and flight height data for these key species derived from Motus telemetry data (Loring et al. 2018, Loring et al. 2019) and past research on the three focal species.

The model and web application have been initially implemented using Motus telemetry data (Loring et al. 2018, Loring et al. 2019) for three birds that are often the focus of species impact assessments in the western Atlantic: Roseate Tern (*Sterna dougallii*), Red Knot (*Calidris canatus*), and Piping Plover (*Charadrius melodus*). Future updates should allow the ability to provide data for other species, but that is not yet implemented in this version of SCRAM.

SCRAM can provide decision support for both environmental assessments and research related to collision risk from offshore wind. In SCRAM we are facilitating the use of CRMs by providing all necessary inputs to generate projections for any location in the Atlantic Outer Continental Shelf for the target species. The underlying statistical models are aiming to capture the large-scale processes that drive variation across this planning area, and therefore might not fully capture micro-scale variation. SCRAM does, however, allow the user to supersede the built-in data if higher resolution data exist for a particular site.

For more specialized applications, the underlying code for our adaptation of the Band, Masden, and Trinder CRM is available on GitHub. For a more detailed discussion of the strengths and limitations of SCRAM and other CRMs, see Appendix I.

Who is the intended audience of SCRAM

The intended audience for SCRAM, an adaptation of Band (2012), is anyone with an interest in understanding collision risk from wind for decision-making, planning, policy, or environmental assessments. This audience includes conservation practitioners, state and federal agencies, non-governmental organizations, and industry. The web application was developed for users who do not have previous experience with statistical or computational modeling, but as with previous iterations of the Band model (Masden 2015 and Trinder 2017), the code for both the CRM and web application are open source.

What is a Collision Risk Model?

At its core, a Collision Risk Model (CRM) estimates the number of collisions between a given bird or bat species and an array of turbines. The key pieces of information that are: (1) how many individuals of a given species are in the area that will be developed, (2) how many of those animals could pass through the rotor swept zone of the turbines, (3) the flight behavior of the animals, (4) and the probability that the animal will avoid the turbine blades through meso- or micro-avoidance. The rules of the simulation are determined by first principles of physical phenomena – e.g., blade rotation frequency is used to determine how often they would strike objects passing through them– as well as basic ecological models that estimate the likelihood of birds being in the vicinity of turbines. The simulation relies on array-specific data for physical turbine characteristics (e.g., number of turbines and blades, rotor speed, altitude of the rotor-swept zone) and the characteristics of target species (e.g., typical flight speed, passage rates through the area of interest, bird size, and avoidance behavior). This type of collision risk model includes all of the major components that are likely to influence risk for a proposed or existing array, but it does not integrate information from other arrays in the region. It is, therefore, best suited for array-specific assessments of risk. Further description of the goals, strengths, and limitations of CRMs is included in a CRM starter guide in Appendix I.

How does SCRAM differ from previous models?

One major difference in our adaptation of this CRM framework is in how bird passage rates are parameterized. Previous models typically have used observational survey data, typically collected from vessels, to estimate passage rates and flight heights. However, observational line-transect surveys at sea are primarily intended to obtain data on marine birds (as opposed to shorebirds like Piping Plovers and Red Knots), and may not be an optimal study method for obtaining information on at-sea behaviors or potential for interactions with anthropogenic structures (Camphuysen et al. 2012, Ronconi et al. 2015). Thus, SCRAM parameterizes bird passage rates using publicly accessible data from the Motus Wildlife Tracking System, an automated radio telemetry network (Taylor et al. 2017, Loring et al. 2018, Loring et al. 2019). See Appendix II for more information on the differences between SCRAM and previous implementations of Band's (2012) collision risk framework. For additional technical details of the modeling framework used in SCRAM see Band (2012), Masden (2015), and Trinder (2017), as well as Appendices II, IV, V, VI.

Data limitations specific to the current version of SCRAM

SCRAM is an evolving tool that will be updated as additional data and methods become available. SCRAM's CRM currently uses static flight height distributions, as opposed to distributions that vary over space and time, which are more realistic (Péron 2020). While we feel that treating flight heights as a statistical distribution, as we have done, implicitly accounts for some of this potential variation, more research is needed to determine whether unexplained variation is likely to influence collision risks (e.g., if flight heights increase substantially with increasing distance from land). One particular challenge to estimating passage rates from movement data is determining the overall number of individuals that could potentially encounter arrays. While we are using the latest regional population estimates for the target species, there are currently limitations in our knowledge of how representative movement data are of these populations.

What software and/or hardware is required for SCRAM?

Masden (2015) adapted the Band (2012) model for the programming language of the computing software R (R Core Team 2021). Trinder (2017), McGregor et al. (2018), Chris Field at URI, and these authors have further adapted this R code for use with Motus data as well as numerous other enhancements. McGregor et al. developed online user interfaces using the R package 'shiny' (Chang et al. 2021), which allows users to run computational tasks in R on a remote server. SCRAM built on the code from McGregor et al. and was further modified to run on the remote Shinyapps.io remote server by the team at BRI. No software is needed other than a up to date web browser running on any PC, Mac, or Linux device.

Updates to the Tool

Users experiencing problems with the operation of the tool should contact Andrew Gilbert at Andrew.gilbert@briwildlife.org or post a bug request at the SCRAM GitHub repository (https://github.com/Biodiversity-Research-Institute/SCRAM/issues). Updates to the tool and/or this user manual will be published at the following location:
https://briloon.shinyapps.io/SCRAM/

How to use the web application (SCRAM)

Overview

SCRAM requires two types of data: 1) “Wind farm data”, which are provided via a single spreadsheet of turbine and array characteristics, and 2) “Species data”, which are incorporated into the tool for the three target species. The application is built as a dashboard-type layout with input being added on the left-hand side of the screen (the sideboard) and outputs available on the tabs to the right of the sideboard in the main part of the dashboard. There are currently four tabs: “Start Here”, “Species Data”, “Wind Farm Data”, and “CRM Results”. A description of each tab is below, with further details given as we proceed in describing it’s use.

1) Start Here – this tab includes some basic instructions for use as well as buttons for downloading the manual, example species input data, and example wind farm data.
2) Species Data – this tab includes tables of species data that are included with SCRAM for Red Knot, Roseate Tern, and Piping Plover, or uploaded custom data. A plot of the flight height data are also available for uploaded species.
3) Wind Farm Data – A table showing the wind farm specs and operational data are provided for the uploaded wind farm as well as map of the wind farm location.
4) CRM Results – Here is where basic output is given as a histogram of the number of collisions per year for each iteration. Next steps such as performing a sensitivity analysis, downloading data, and downloading a report of the SCRAM run.

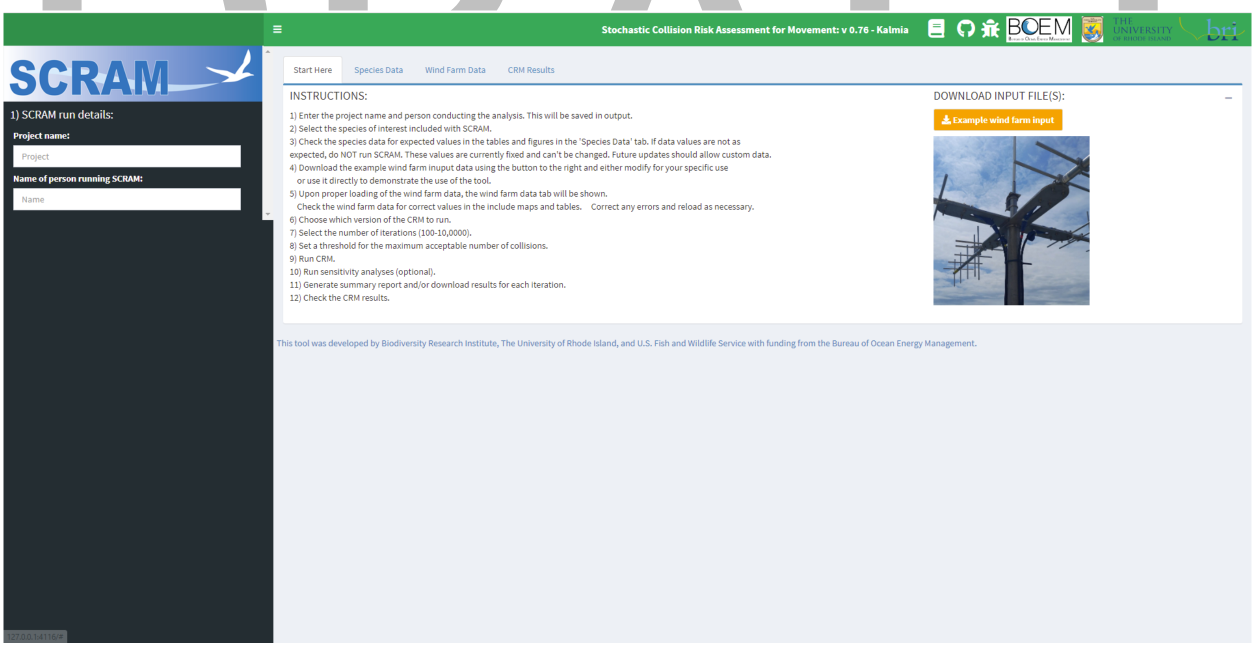


Figure 0-1. SCRAM initial start up screen.

SCRAM requires two types of data: 1) “Turbine data”, which are provided via a single spreadsheet of turbine and array characteristics, and 2) “Species data”, which are included in the model for three target species (Red Knot, Roseate Tern, and Piping Plover). Currently custom species data can **NOT** be uploaded and included species data can **NOT** be changed.

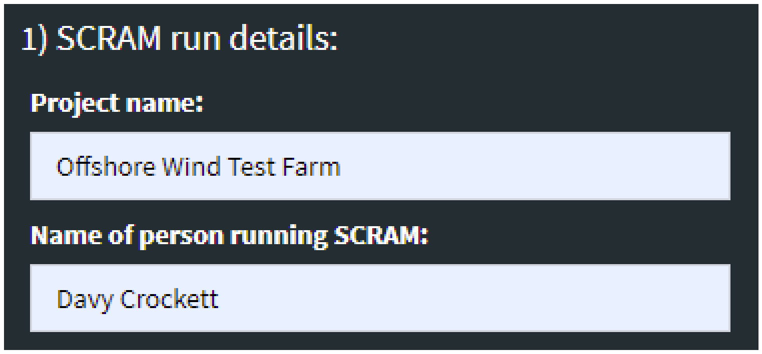
Examples of wind farm data input can be downloaded from the application interface using the “Example wind farm input” buttons shown on the “Start Here” tab, shown in Appendix III.

Once the data for turbine and array characteristics are compiled and formatted appropriately (see examples on interface and Appendix III), it should take 3-5 minutes to finish setting up SCRAM to run for a target species. The run time once the data are uploaded will depend on which version of the model specification is selected. The general steps for running SCRAM are discussed below. The interface was created to lead users through data input and model run with some inputs not available to the end-user until the prior input has been entered in SCRAM. A description of the basic steps is below:

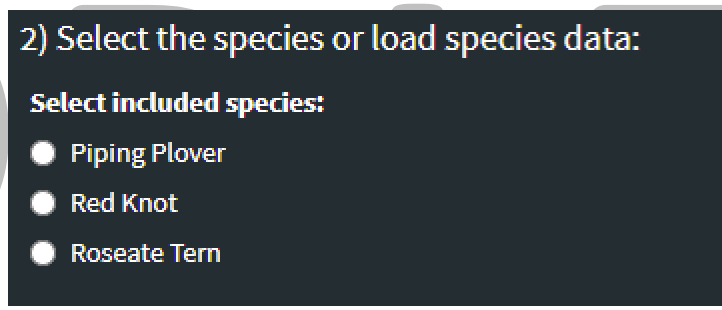
1) Enter the project name and person conducting the analysis.
2) Select the species of interest included with SCRAM.
3) Once the species is selected, the species data tab will be shown. Check that the species data and flight height data are as expected for this species. If not as intended, do NOT run SCRAM.
4) Download the example wind farm input data from the “start here” tab using the “example wind farm input” button to the right, and either use the example data directly or modify the file for your specific use.
5) Upon proper loading of the wind farm data, the wind farm data tab will be shown. Check the wind farm data are correct by examining the maps and tables in the wind farm data pane. Correct any errors and reload as necessary.
6) Choose which version of the CRM to run (the faster/approximate version or slower/more accurate version).
7) Select the number of iterations (100-10,0000).
8) Set a threshold for the maximum acceptable number of collisions. Estimates that are above the threshold number will be highlighted in model outputs.
9) Run CRM.
10) Run sensitivity analyses (optional).
11) Download model results (optional).
12) Generate output report (optional).
13) Check the CRM results.
14) Run SCRAM again as needed.

Detailed description of SCRAM usage

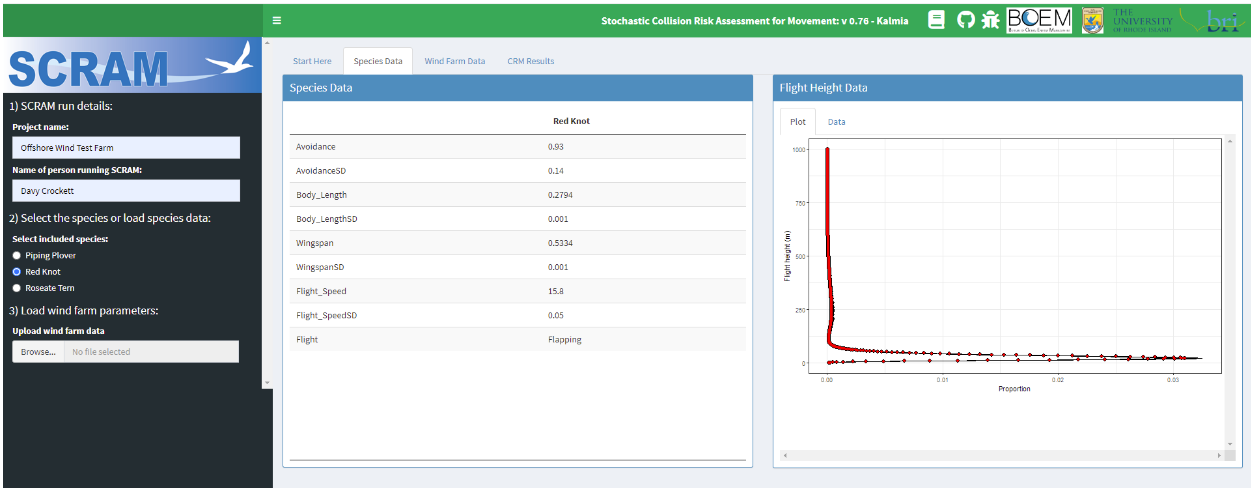
1) *Enter the project name and person conducting the analysis.* Choose whatever project name will be informative for you; this information will be saved in the output once SCRAM is run.



2) *Select the species of interest or select the option for providing your own species data.* Select the included target species (Piping Plover, Red Know, Roseate Tern)

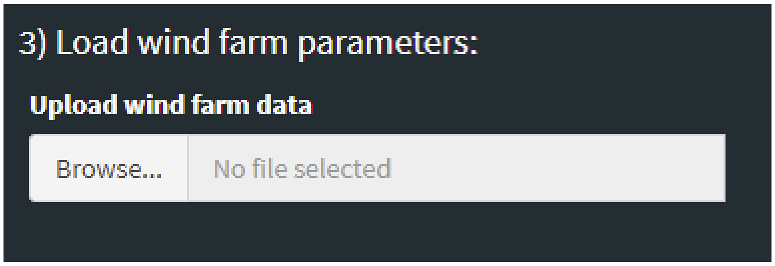


3) *Check the species data tables for correct values.* Make sure the values included in SCRAM are appropriate for these models by examining the tables and figures in the species data pane. Examine the flight height graph to see if the flight height data make sense and are appropriate for the model. Do NOT run SCRAM if data is not as expected. Future updates will include the ability to add custom species data.

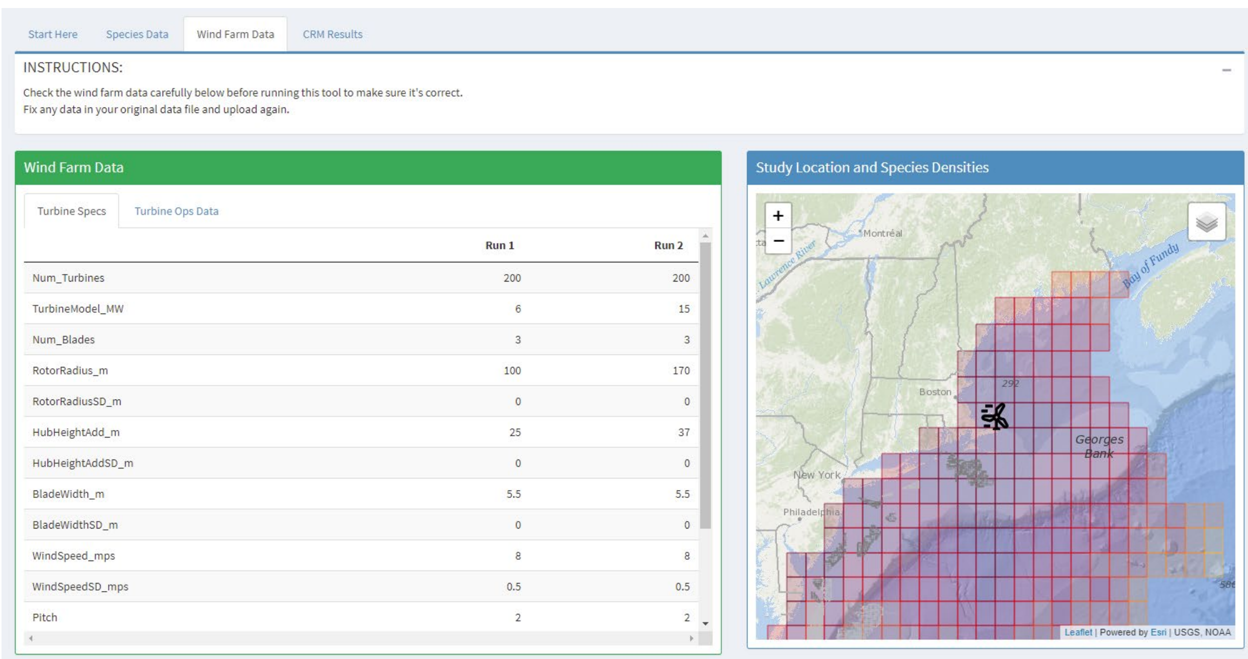


4) *Download the example wind farm input data from the "start here" tab to either use the example data directly or modify the file for your specific use.* Data for turbine and array characteristics are required to run any version of SCRAM. Turbine and array characteristics include the physical and geographic characteristics of the wind farm, including the dimensions of the turbine model, power targets, and the width and geographic coordinates of the turbine array (Table 1 in Appendix III). Note that the included example file (TurbineData_inputs_2run_example.csv) shows an example of wind farm options.

a. The application accepts these data as a single .csv file that has alternate options for arrays specified as rows. You can provide as many options as you desire or only one option by adding rows or removing the second option row if only a single option is desired.
b. SCRAM can run using only **ONE geographic location at a time**, but the specified array options can vary by other parameters including power, size, and turbine model specifications. If the rows for the turbine and array data include more than one set of geographic coordinates (latitude/longitude), only the coordinates from the first row will be used for model output. In order to run multiple locations for SCRAM, you must run SCRAM multiple times changing geographic coordinates for each run.
c. The naming convention of the file itself does not matter, as long as the appropriate fields are included and correctly named.
d. Turbine specifications, such as blade pitch and width, can be found on the manufactures web pages (e.g. models from General Electric) or the U.S. Geological Survey's Wind Turbine Database.
e. Upload the data file once you are satisfied with the values in the file.

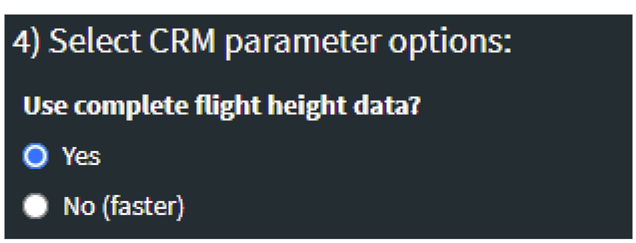


5) *Check the wind farm data tables for correct values*. Make sure the values you uploaded to SCRAM are appropriate for these models by examining the tables in the wind farm data pane. Correct any errors and reload as necessary. The map on the right of the “Wind Farm Data” tab shows the occurrence probability surface generated from the modeled Motus data for each species. The user can also show the 95% confidence interval range for those modeled data by clicking on the symbol at the upper right corner of the map. This is the layer tool which allows the user to turn layers on and off including BOEM lease and wind energy areas, the wind farm location, and the occurrence data.



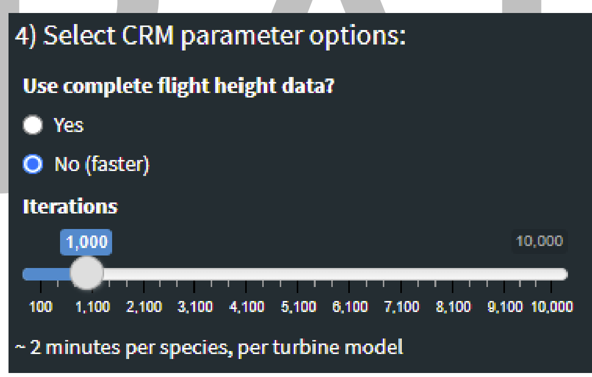
	Run 1	Run 2
Num_Turbines	200	200
TurbineModel_MW	6	15
Num_Blades	3	3
RotorRadius_m	100	170
RotorRadiusSD_m	0	0
HubHeightAdd_m	25	37
HubHeightAddSD_m	0	0
BladeWidth_m	5.5	5.5
BladeWidthSD_m	0	0
WindSpeed_mps	8	8
WindSpeedSD_mps	0.5	0.5
Pitch	2	2

6) *Choose which version of the CRM to run (the faster/approximate version or slower/more accurate version).*

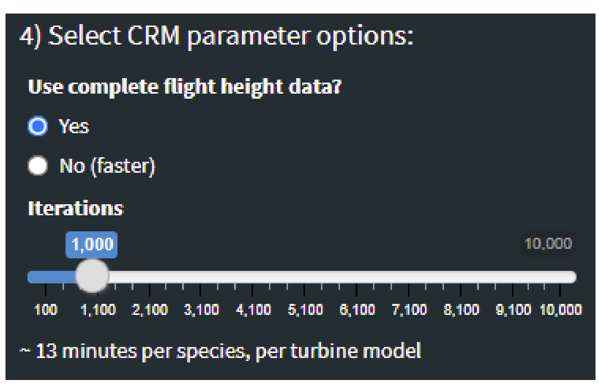


The principles that the CRM uses to simulate collision risk are simple, but there are two flavors for how these principles are executed that differ in how input data are used in the underlying calculations. Band (2012), Masden (2015), and Trinder (2017) provided several options, which we synthesized to provide two options – one that we have shown performs best and one that gives approximate estimates in much less time. In SCRAM, both options use the same input data and any necessary data manipulations are done automatically. In option 1 (Yes – use complete flight height data); SCRAM takes into account risk along the rotor blades using complete flight height data and thus provides a more accurate accounting of collision risk (Trinder 2017), but is slower to run. Option 2 (No (faster) – do not use complete flight height data); risk is not modeled along the rotor but is presumed to be constant risk.

Because these options estimate collision risk using the same data input, **we recommend selecting "Yes" for "Use complete flight height data" whenever the user is not severely limited by computation time**. Estimates of time are given below the iterations slider and are dependent on both the option used (yes or no) and the number of model iterations desired (100-10,000). For example, an estimate of two minutes is given for option 2 (No) with 1,000 iterations.

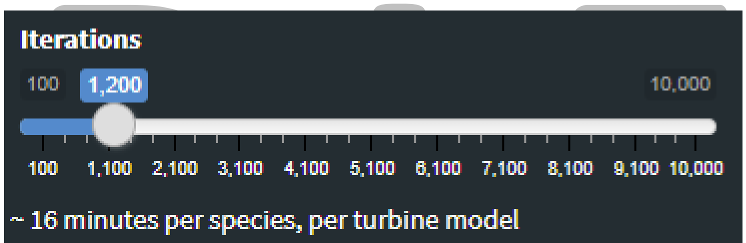


The same number of iterations takes ~13 minutes using full flight height data and accounting for differences in risk along the blade.



We do not recommend running both options for comparison because, in most cases, the option that uses complete flight height data will provide the more accurate estimate (Appendix V). Selecting “No” reverts to the “Standard” option in Band (2012), which in SCRAM summarizes the full flight height distribution and provides NO estimate of risk along the blade. For SCRAM’s target species, we recommend selecting “Yes” to use the full flight height distributions to account for the variation in risk along the rotor blades (Trinder 2017). For more detail, Appendix VI shows the underlying mathematical routines that differ between options and a test of their relative speed.

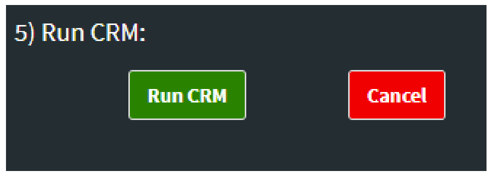
7) *Select the number of iterations*. Use the slider to specify how many iterations of the model will be run in order to propagate the influence of parameter uncertainty on the simulation results. A dialog box at the bottom of the application interface will give an estimated run time for the currently selected option. In this CRM framework, uncertainty – i.e. variation in the results among iterations – is a result of the variance estimates provided for the input parameters. Increasing the number of iterations will give more accurate estimates for the model outputs, until the error associated with estimating outputs via stochastic simulation is arbitrarily small, which is around 10,000 iterations for this model.



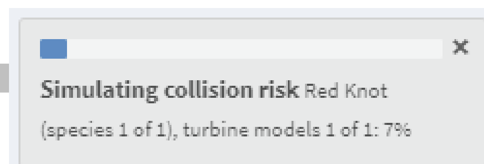
We recommend running **at least 1,000 iterations**, but SCRAM defaults to a minimum of 100. For final estimates for collision risk a user and to get the best estimate from SCRAM **10,000 iterations should be run** but note that this can take ~2 hours per turbine model/wind farm specification and the user must watch SCRAM to make. The user must pay attention to SCRAM as the model is running because once finished, the application will timeout after 30 minutes of inactivity. Results are not automatically saved so the user must export results within that time period the model finishes or interact with SCRAM to make sure the application is closed.

8) *Set a threshold for the maximum acceptable number of collisions.* If desired, specify a threshold for the maximum acceptable number of collisions (this number can be zero). When provided, SCRAM will calculate the proportion of iterations that produce a collision estimate larger than the specified threshold. The application will show the threshold value alongside the results for reference, and will include the probability of exceeding this value in a dialog box.

9) *Run CRM.* The button to run the CRM appears when the minimum conditions for running the model are met.



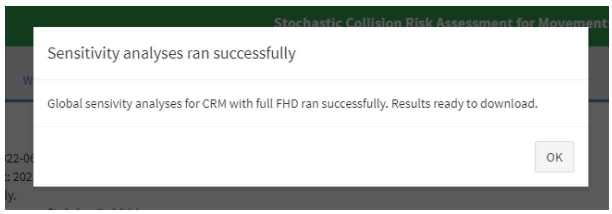
A status bar will update progress through the specified species and turbine models. If this button is clicked in error, clicking “cancel” will stop the current run.



Once the model has completed, the CRM results tabs displays basic results including details of the model run times, the model that was run, probability of exceeding the threshold, and histograms (one for each wind farm option) of the number of collisions per year for each iteration. This histogram is also provided in the output report.

10) *Run sensitivity analysis (optional)*. Once the model has been run successfully, SCRAM provides the option to run a simple sensitivity analysis that determines the relative contribution of the input parameters to the uncertainty bounds of the results. We have provided this option as a general guide for determining where (e.g., for which parameters) more precise data are likely to lead to the biggest gains in our understanding of collision risk for the species and arrays of interest. We do not recommend running sensitivity analyses when the number of iterations is less than 1000. The results are saved to a .csv file that can be exported from the application using the “Download model runs” button.

Successful sensitivity analysis will show the following dialog box:



For the 10 most influential parameters, the analyses provide estimates of the proportion of the variation in the results that is contributed by each parameter. For example, if the value for turbine avoidance rate is 0.12, approximately 12% of the width of the final uncertainty bounds is a result of uncertainty in our understanding of avoidance behavior. The analyses provided in SCRAM are approximate (see Borcard 2002 for details on the methods) due to computational constraints, so we recommend using this option as a rough guide for understanding parameter sensitivity or potentially planning additional sensitivity analyses.

11) *Download model results (optional).* Once the model has run successfully, which will be indicated in a dialog box at the bottom of the interface, the option to download the full results will appear. When the user clicks the button a file save dialog box will appear with the ability to browse to a location for saving as well as change the compressed file name. A compressed file will be downloaded containing the following files

   i. Model output Rdata file that can be directly loaded in R
   ii. Collision estimates for each month in each iteration as a csv file
   iii. The stochastic draws of all input parameters for each iteration as a csv file
   iv. Sensitivity analysis results as a csv file

12) *Generate output report (optional).* The user can also download a custom report for the model runs by clicking the “Generate output report” button. The report provides details about the model run details (SCRAM version, run times, project, user, and probability of exceedance), model input parameters including both species and wind farm parameters, histogram of the number of collisions per year for each iteration, figure showing the predicted mean and 95% confidence intervals for the number of collisions per month, a figure showing the number of collisions per month for each species and turbine model

combination, and histograms comparing the difference in the number of collisions per year between models.

13) *Check the CRM results.* Check the results of the model to see that they are sensible. Tips for interpreting the results are discussed below.

14) *Run again.* The model may be run again at this point by selecting another species or choosing a new options (you can also refresh the browser if you prefer to start with a clean slate).

Tips for interpreting the results

SCRAM provides several types of visualizations to aid with interpreting results in the model output report (and in a more limited way in the "CRM Results" tab), as well as the option of downloading spreadsheets and RData files of the raw model output (Step 10, above to conduct further analysis and/or generate other figures and tables. All figures include uncertainty, either summarized as bars or shown as variation in the model runs across stochastic iterations. In this CRM framework, uncertainty in the collision risk estimates is the result of variation in key parameters (e.g., annual variation in regional population sizes or flight heights) as well as uncertainty in our estimation of these parameters (e.g., uncertainty in flight height estimates due to estimation error or uncertainty in habitat use due to variation in Motus coverage in the region). Increasing the number of iterations will give more accurate estimates for the model outputs, until the error associated with estimating outputs via stochastic simulation is arbitrarily small, which is around 10,000 iterations for this model. Note, however, that we do not measure the bias or uncertainty of the collision risk model itself, rather we propagate uncertainty in model parameters through the deterministic CRM. If model parameters are biased or uncertainty estimates are in accurate, then CRM results may be inaccurate or biased regardless of their precision.

The plot on the application interface shows the number of simulation iterations associated with the values on the x-axis, with wider distributions being a result of greater uncertainty in the number of collisions per year. We have also provided the option to summarize this uncertainty probabilistically by specifying a threshold for the acceptable number of collisions per year. This threshold can be any integer between zero and an arbitrarily large number. When provided, SCRAM will calculate the proportion of iterations that produce a collision estimate larger than the specified threshold (which can be interpreted as an estimate of the probability that the number of collisions will exceed the specified threshold value). This value is a probability, rather than a known outcome, because there is uncertainty in the input data. For example, if all of the input data were known and provided as values with 100% certainty, as opposed to estimates with error bounds, a simple yes or no could be given for whether collisions exceed the threshold. Because input data will never be known with 100% certainty, however, SCRAM provides a probability of exceeding the specified threshold, the precision of which is scaled to account for the number of iterations – i.e. SCRAM will show more significant figures as the number of iterations is increased. This probability is also dependent on model assumptions. A plain language interpretation is “X is the probability of exceeding the specified threshold, taking into account the uncertainty of the input data and assuming that the model is a reasonable description for how collisions happen in reality.”

For more detailed visualizations, including variation by month, download the output report. We have included tables for the input parameters so that all results are associated with the input data, as the results of CRMs can be sensitive to certain aspects of the availability and quality of the underlying data (see step 9 above).

Appendix I. A starter guide for understanding the strengths and limitations of collision risk models

Christopher Field[1], Pam Loring[2], Brian Gerber[1]
University of Rhode Island[1]
U.S. Fish and Wildlife Service[2]

Overview

The aim of this guide is to clarify a few key points about collision risk models to help potential users interpret their results and better understand how they can use emerging tools, like the recently developed web application SCRAM, to inform their work.

The aim of collision risk modeling

Collision risk models (CRMs) generate projections for the likely number of bird collisions at wind farms, which can inform planning, mitigation, and assessments of the cumulative impacts of increased mortality for at-risk species. The primary benefit of CRMs is that they generate this evidence at the earliest stages of wind farm planning, when there is the greatest potential to influence the design of turbine arrays to minimize risk. During this critical time in the lifespan of a wind farm, it is currently not possible to collect data on collisions *per se*. Data collection is therefore limited to the factors that will likely influence collisions, including passage rates and flight heights. Projections models are one way to synthesize these data, which are likely to be essential components of pre- and post-construction monitoring, to provide estimates of the quantity we are often most interested in – collisions. Using CRMs to project the number of collisions also makes it possible to identify turbine and array characteristics that are most likely to influence collision risk. The greatest utility of CRMs is during the planning, leasing, and construction phases of wind farms. Over time, the projections from CRMs can be validated, and even perhaps eventually replaced, by direct monitoring of collisions during operation (Figure A 1).

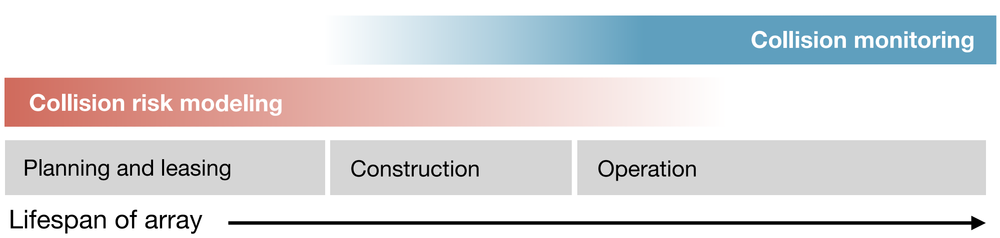


Figure A 1. A conceptual diagram showing relative utility of collision risk modeling vs. direct monitoring of collisions over the lifespan of an offshore wind farm. Collision risk modeling, which is necessary pre-construction, becomes less important as turbines are erected and empirical collision monitoring becomes more feasible.

What causes CRMs to be wrong?

The value of CRMs for planning depends on how successful they are at synthesizing the available data and generating reliable estimates of collision risk. When considering the performance of CRMs, we think it is helpful to distinguish between bias caused by misspecification of the CRM vs. bias in the underlying data. Bias and uncertainty in data for offshore areas are a general roadblock to understanding the impacts of wind, and CRMs are not more sensitive to these data limitations than other approaches for understanding potential impacts, such as verbal models. The reliability of a collision risk modeling framework is, instead, determined by 1) how it handles uncertain data and 2) whether the assumptions underlying the translation of these data sources to collision risk are appropriate.

Incorporating uncertain data is straightforward, as current implementations of the Band (2012) framework for collision risk modeling are stochastic – i.e. the bounds of the collision risk estimates accurately reflect the uncertainty of the input data. This framework even allows a user to, if appropriate, specify complete uncertainty for a parameter with no reliable data input (e.g., behavioral avoidance rates for some species). Most of the assumptions of CRMs – i.e. the processes that a CRM captures through its model specification – are straightforward or based on a strong body of evidence (e.g., the importance of the distribution of flight heights; Johnston et al. 2014). An important feature of the Band (2012) CRM framework is that the assumptions of the model are all transparent, straightforward, and based on first principles, making misspecification of the CRM unlikely. For example, at one step of the model the number of birds at risk is narrowed by the proportion of birds flying at risk height, which is simple (and uncontroversial) multiplication. The rotor risk component of the CRM, which models the process of birds flying through a spinning turbine blades, is based on simple, verifiable geometry. In fact, modeling these geometric relationships is the only means of making inferences about risk that incorporate the influence of blade length, shape, and speed. Other processes, such as correcting for behavioral changes around large arrays, are reasonable, but would benefit from validation either through additional research or direct validation of CRM projections when data are available. For some of these cases, such as large array corrections, the underlying assumptions are only likely to be consequential under a narrow set of circumstances (Band 2012).

When CRMs generate projections that do not accurately reflect reality, it is usually the result of bias in the underlying data, which can be caused by biases in collection or analysis. For example, several inputs to CRMs, both historically (flight height distributions) and with the SCRAM implementation (movement tracks), use estimates from statistical models. The aim of these statistical analyses is to find the best balance between uncertainty and bias. While uncertainty is fully accounted for in a stochastic CRM, any bias that remains falls under the category of “known unknowns”, highlighting the importance of identifying potential bias whenever possible. The potential issues associated with biased data are not limited to their use in CRMs, however. For example, if the estimated flight height distributions are skewed, the potentially skewed estimates of collision risk are a symptom of a general problem arising from our biased understanding of flight heights, and would affect any inference about risk including simpler or verbal models.

A related source of bias is micro-scale deviances from the input data caused by a spatial or temporal mismatch in the scale of the array of interest and the underlying data. For example, with SCRAM we are facilitating the use of CRMs by providing all necessary inputs to generate projections for any location in the Atlantic Outer Continental Shelf for the target species. The underlying statistical models are aiming to capture the large-scale processes that drive variation across this planning area, and therefore might not fully capture micro-scale variation.

The strengths of CRMs

Identifying the strengths and limitations of CRMs and applying them in the appropriate contexts is key to using them successfully. Inference in the Band (2012) framework is focused on individual arrays, which makes it well-suited for providing information on planning at this scale, including turbine models and characteristics, spacing, and operation time.

The Band model can be run as many times as needed to estimate risk for multiple arrays or different time periods. It does not, however, contain any methods for explicitly quantifying cumulative risk over space and time, making it less suited, in its current form, for questions about scenarios of total power capacity across the Atlantic OCS or questions about changing conditions such as bird population trends and range shifts.

Appendix II. Differences between SCRAM and previous implementations of the Band (2012) model

Our approach to developing SCRAM made full use of recent advancements in quantifying the potential impacts of offshore wind from Band (2012) and adaptions of the Band framework (Masden 2015, Trinder 2017, McGregor et al. 2018). We aimed to advance the implementation of this framework in the western Atlantic by 1) contributing updates to the primary model script and 2) developing an online interface that best addresses the specific needs of users and stakeholders in the eastern U.S. While there is significant overlap in the model description between our version and previous iterations, there are several important differences. The most consequential change to the underlying model is that we have re-worked the data inputs to work primarily with movement data – as opposed to surveys data – as these data are widely available through automated telemetry, such as the Motus Wildlife Tracking System (Figure A 2). In developing this new input, we also made it possible to use the Band model framework with missing values, allowing assessments to be conducted for areas or time periods with incomplete data availability.

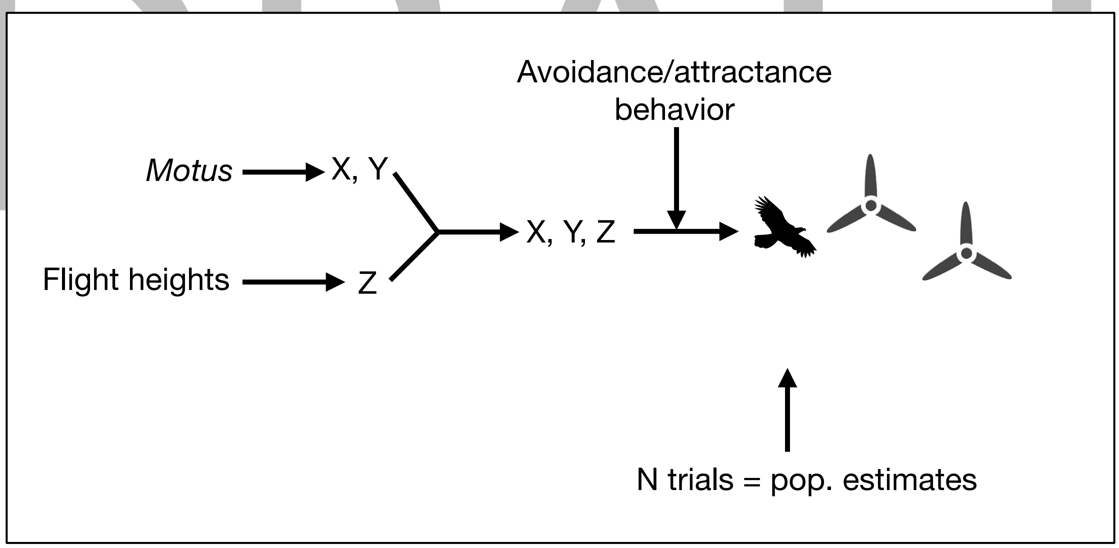


Figure A 2. Conceptual diagram of the primary model inputs for species data. Estimates of species movements from Motus data and species-specific flight height distributions determine the longitudinal (X), latitudinal (Y), and altitudinal (Z) locations of individuals over the course of a year, which in turn determines their risk of encountering rotor space. These estimated movements are combined with regional population estimates, corresponding to the sampling populations of the Motus projects, to estimate the total number of individuals likely to encounter rotor space for each month in which birds were tracked. Monthly collision risk is then estimated using the same assumptions as the other CRMs based on the Band framework (this component of the model is denoted by the square that contains turbines).

We have also revised several components of the primary model script, including an adjustment to how flight height distributions are integrated with risk along the rotor blade. We modified this component to treat flight heights as a statistical distribution, as opposed to point-wise sampling along the range of flight heights (see bullets below for more information). All of the changes to the underlying model code are tracked on GitHub.

The largest differences between our version and previous iterations relate to the delivery of the primary model script via the online interface. Our general philosophy was to make the interface as simple as possible, with most data input being accomplished using .csv files that the user can store locally, as opposed to data input on the interface itself. We also designed the tool to encourage linear advancement through the model specification process. We accomplished this by 1) having a defined, but not required, order of operations and 2) providing specific, evidence-based guidance to identify the most appropriate model option while discouraging the use of more than one option at a time. We have also created more points of dialog with the user and functionality that address the inconvenience of potentially long run times with this CRM framework. SCRAM provides the take-home results on the application interface, but most of model output is delivered via downloads of the raw results or a generated report that contains visualizations and input and output data tables. A more comprehensive list of differences between our version and previous iterations is given below.

Major differences in primary computational script

- SCRAM integrates the flight height distributions with risk along the rotor blade using cell-wise instead of point-wise probabilities. The consequence of this change is that the first probability of the flight height distribution (labeled as 1 m) corresponds to the band that is 0 – 1 m above sea level.
- SCRAM's primary species input is *Motus* data and therefore it uses flight speed and habitat-use estimates derived from tracking data rather than bird passage rates per unit time (referred to as "flux" in Band 2012).
- SCRAM uses spatially explicit occupancy probably derived movement models rather than density estimates derived through surveys. To appropriately scale occupancy to the entire population an estimate of population size (and uncertainty) is needed.
- The primary computational script was revised to include a preamble that conducts a set of checks on the input data sources to ensure they are uploaded correctly.
- SCRAM allows the user to conduct an approximated global sensitivity analysis to quantify the contribution of input data to the uncertainty bounds of the results.
- SCRAM allows for missing values (specified as NA) in the input data. This is useful, for example, when movement data are not available for every month. Missing values are automatically propagated through the model and displayed in the results accordingly.
- SCRAM calculates total operation time as wind availability*(1 – down time) to avoid the fact the negative values can theoretically happen with the original formulation (wind availability – down time).

- SCRAM estimates rotor speed using the relationship between tip speed ratio (TSR), wind speed (S), and rotor diameter (r): w = (TSR*S)/r * pi (in radians/s), which is converted to rpm.
- Fixed an error in the Riemann sum for rotor risk (used for the "extended" version of Band [2012]) that was causing a redundant loop
- The primary computational script is run asynchronously, using the framework of making "promises" with the 'promises' and 'future' R packages (Bengtsson 2020, Cheng 2020), to allow multiple users simultaneously and allow the ability to cancel computational tasks.
- SCRAM allows the user to download inputs and outputs from every iteration of the model run.
- Tidal offset and nocturnal activity are no longer user-specified parameters
- Simplified input so that a global avoidance is used instead of option-specific rates

Major differences in online interface

- SCRAM's interface was built from the ground up, focusing on simplicity and encouraging a linear path through the tool.
- Only the most appropriate options in SCRAM are available to the user, depending on the input data and model specifications, to minimize the chance of running the model in a way the user did not intend.
- The majority of SCRAM's results are given in a downloadable report rather than on the application interface.
- SCRAM includes additional warnings to caution the user about applying the application to areas with sparse data or low confidence.

Appendix III. Metadata for input datasets

SCRAM input datasets for turbine and array characteristics (Table A1) and species characteristics (Tables A2-A4) must match the specified input structure, including exact column names. Underlying movement data (Table A5), which specify the estimated probability that an individual from the target population will pass through the modeled area in each month, are "baked in" to the tool for the three focal species of Red Knot, Roseate Tern, and Piping Plover and do not need to be provided by the user.

Table A1. Turbine and array characteristics included in input datasets. Each turbine/array characteristic, and when appropriate its associated uncertainty, is specified in a column. The table should be 54 columns wide. Each row gives the specifications for a turbine array of interest, so the number of rows should be equal to the number of different wind farm arrays and will dictate how many times the model will run.

Turbine parameter name	Definition
Run	The model run value – wind farm array number
Num_Turbines	The number of turbines in the wind farm array
TurbineModel_MW	Megawatt rating of turbine model
Num_Blades	Number of blades for the turbine model
RotorRadius_m	Rotor radius (hub to blade tip; m)
RotorRadiusSD_m	Standard deviation of rotor radius
HubHeightAdd_m	Hub height (m)
HubHeightAddSD_m	Standard deviation of hub height
BladeWidth_m	Chord width of blade (m)
BladeWidthSD_m	Standard deviation of blade width
WindSpeed_mps	Wind speed rating of turbine model (m/s)
WindSpeedSD_mps	Standard deviation of wind speed rating
Pitch	Pitch angle of blades (degrees relative to rotor plane)
PitchSD	Standard deviation of pitch angle of blades
WFWidth_km	Wind farm width (km)
Latitude	Latitude (decimal degrees)
Longitude	Longitude (decimal degrees)
Prop_Upwind	Proportion (0 - 1) of birds flying in upwind direction
MonthOp *(x12)*	Wind availability (maximum amount of time turbines can be operational/month). One column for each month, e.g., JanOp, FebOp…
MonthOpMean *(x12)*	Mean time that turbines will not be operational ("down time"), assumed to be independent of "MonthOp" – i.e. total operation = MonthOp*(1 – MonthOpMean). One column for each month, e.g., JanOpMean, FebOpMean…
MonthOpSD *(x12)*	Standard deviation of mean operational time. One column for each month, e.g., JanOpSD, FebOpSD…

Table A2. Species characteristics included in input datasets. Each species characteristic, and when appropriate its associated uncertainty, is specified in a column. The table should be 10 columns wide.

Species parameters	Definitions
Species	Species name for associated data
Avoidance	Proportion of birds that avoid turbines
AvoidanceSD	Standard deviation of avoidance estimate
Body_Length	Body length of target species (m)
Body_LengthSD	Standard deviation of body length
Wingspan	Wingspan of target species (m)
WingspanSD	Standard deviation of target species
Flight_Speed	Flight speed of target species (m/sec)
Flight	Flight mode (“flapping” or “gliding”)

Table A3. Flight height data example with appropriate units. This dataset specifies the estimated flight height distribution for the species of interest. The flight height distribution gives the relative probabilities of an individual flying at each height across the range of possible heights, at 1 m intervals from 1 - 1000 m. Columns are samples from the uncertainty distributions of the relative probabilities, which can be bootstrap samples or draws from a posterior distribution. The table should contain 1000 rows (not including the header). The number of columns is flexible, but we recommend at least 100 to adequately represent the sampling distributions.

Species	Height_m	BootId_1	BootId_2	BootId_3	BootId_4	…	BootId_100
Roseate_Tern	1	0.10	0.08	0.10	0.08		0.09
	2	0.09	0.08	0.09	0.07		0.08
	.						
	1000	0.08	0.07	0.08	0.07		0.04

Table A4. Count data example. This dataset specifies regional population sizes associated with the movement dataset for each species of interest. Columns specify the mean and standard deviation for the estimated population size, which can vary by month. The table should be 25 columns wide. The number of rows should be equal to the number of species.

Species	Jan	JanSD	Feb	FebSD	Mar	MarSD	…	Dec	DecSD
Roseate_Tern	80	15	70	12	65	14		86	19

Table A5. Format of underlying movement data, which are "baked in" to the tool for Piping Plover, Red Knot, and Roseate Tern and do not need to be provided by the user. This dataset specifies the estimated probability that an individual from the target population will pass through the modeled area in each month. Each row is a sample from the uncertainty distributions of these estimates, is a draw from a posterior distribution. The table should be 13 columns wide.

Species	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Roseate_Tern	0.09	0.11	0.12	0.07	0.05	0.10	0.12	0.06	0.08	0.11	0.12	0.05
	0.08	0.12	0.09	0.05	0.03	0.14	0.13	0.04	0.10	0.11	0.12	0.04
	.	.	.	.	.	.	.	.	.	.	.	.
	0.07	0.13	0.11	0.06	0.03	0.12	0.12	0.07	0.09	0.10	0.11	0.04

Appendix IV. Options for Estimating Flight Height Distributions

We ran the model using Red Knot as a test case with identical inputs using 1) the option that integrates the full flight height distribution (FHD) and 2) the option that summarizes the FHD to obtain a single value for the proportion of birds flying at risk height. We then calculated the results from this latter option (2) as a proportion of the former (1). The resulting variable provides an estimate of the difference between options in terms of the percentage by which the simpler option is off in either direction. For example, a value of one means the options gave the same answer, while 1.8 means the simplified FHD gave an answer that was 80% larger than the more realistic option. Both options used the same input data and identical flight height data so that the only difference was that the simpler option summarized the FHD instead of using the full distribution.

The two options are related to the “extended” and “standard” options of Band (2012) (“Yes” and “No” for “Use complete flight height data?”, respectively, in SCRAM. They are compared in Figure A 3 using the above ratio in order to quantify the influence of several variables on the differences between the two flight height model options, including hub height, rotor size, and underlying shape of the flight height distribution. We recommend using the full FHD when running SCRAM for the tool’s three species (“Yes” for “Use complete flight height data?”).

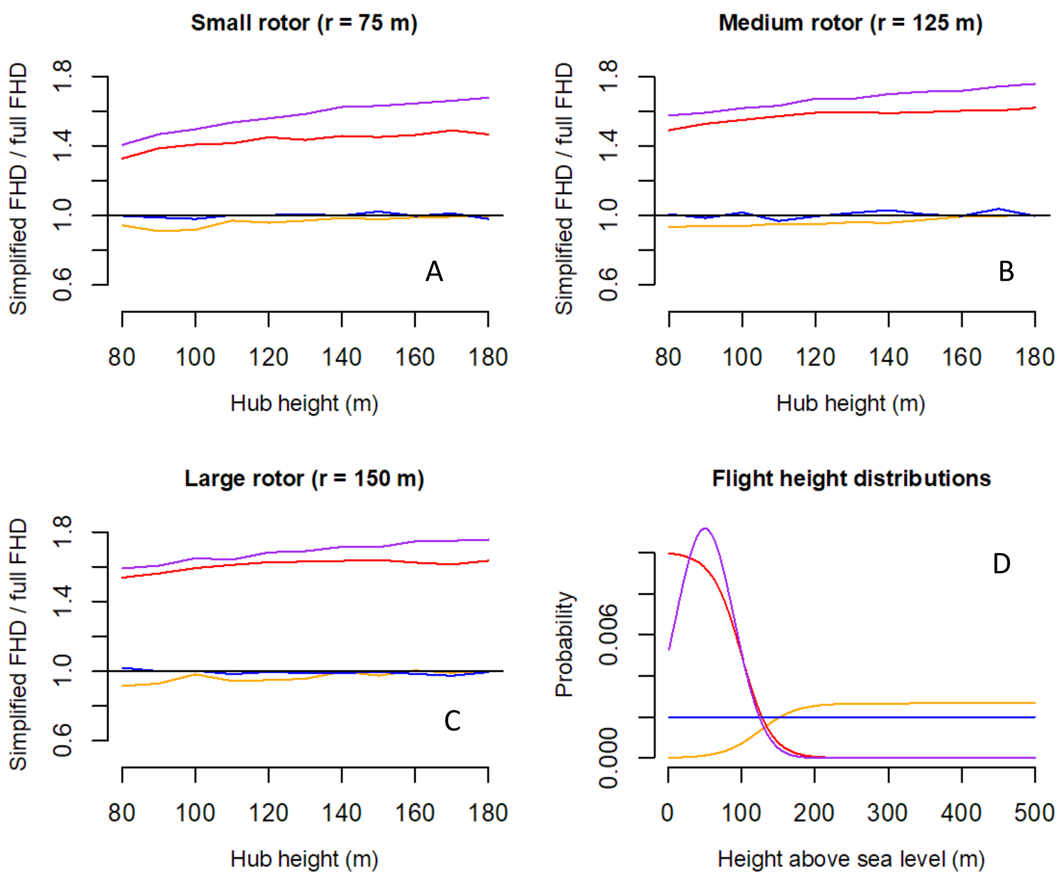
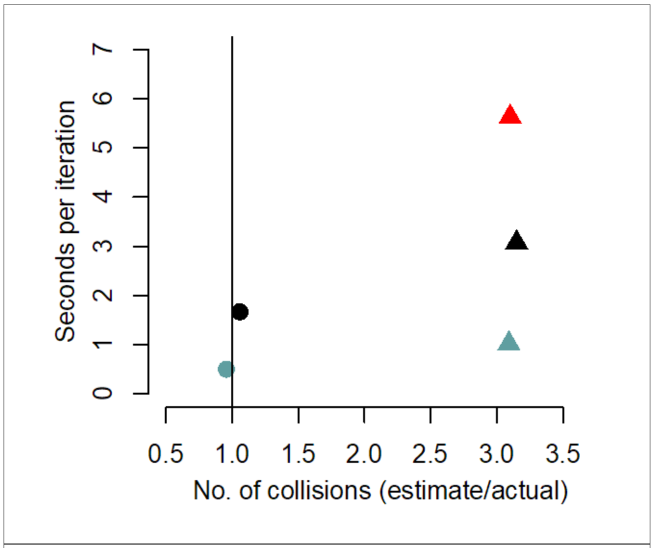


Figure A 3. Demonstration for a range of flight height distributions of the potential influence of model option on the estimates of the number collisions/year. The colors of the lines relate to the four FHD shapes shown in the bottom-right plot (Part D). The proportional difference between the two options (y-axes in Parts A-C) was assessed for a range of hub heights (x-axes in Parts A-C), four shapes for the FHD (Part D), and three rotor sizes (75 m, 125 m, and 150 m; Parts A-C) to quantify the influence of these variables on the differences between model options.

Appendix V. Speed vs. accuracy for five methods for estimating collision risk



Full integrated version

$$N * 2/pi \int_{-1}^{1} \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} d(y)\, p(xy)\, dx\, dy$$

Integration (analytical)

▲ $$N * (\int_{-1}^{1} d(y)\, dy)/2 * 2/pi \int_{-1}^{1} \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} p(xy)\, dx\, dy$$

Integration (approximated)

● $$N * 2/pi(\sum_{y=-1}^{1} \sum_{x=-\sqrt{1-y^2}}^{\sqrt{1-y^2}} d(y)\, p(xy))/xinc$$

▲ $$N * (\sum_{y=-1}^{1} d(y))/2 * 2/pi(\sum_{y=-1}^{1} \sum_{x=-\sqrt{1-y^2}}^{\sqrt{1-y^2}} p(xy))/xinc$$

Individual-based

● $$\sum_{i=1}^{N} rmultinomial(d) * rBernoulli(p(x_i, y_i))$$

▲ $$\sum_{y=\min(rotor)}^{\max(rotor)} d(y) * \sum_{i=1}^{N} rBernoulli(p(x_i, y_i))$$

The five options' estimates of the number of collisions over the course of one year, assuming no avoidance, are compared to the full double integral of Band (2012) (eqs. 9 and 10), which does not use approximations or simplifying assumptions. The options are categorized by their approach to estimating the full integral: blue is an approximation of the integral ("Yes" for "Use complete flight height data?" in SCRAM); red uses an analytical solution for the integral for rotor risk (not available in SCRAM); and black is an individual-based population simulation that uses a probabilistic framework instead of integration (not available in SCRAM). Versions of the options that estimate the proportion of birds flying at risk height and the risk over the rotor disk are shown as triangles ("No" for "Use complete flight height data?"). Versions that estimate these probabilities in a combined fashion, either by double integration or in a probabilistic framework are shown as circles ("Yes" for "Use complete flight height data?"). For each option, a general form of the equation that estimates the number of collisions is shown for reference.

References

Band, B. (2012). Using a collision risk model to assess bird collision risks for offshore windfarms. The Crown Estate as part of the Strategic Ornithological Support Services Programme, Project SOSS- 02.

Bengtsson, H. (2020). R package “future”: Unified parallel and distributed processing in R for everyone (1.21.0). https://CRAN.R-project.org/package=future.

Borcard, D. (2002). Partial r2, contribution and fraction [a]. Multiple and Partial Regression and Correlation. http://biol09.biol.umontreal.ca/borcardd/partialr2.pdf.

Camphuysen, C. J., Shamoun-Baranes, J., Bouten, W., & Garthe, S. (2012). Identifying ecologically important marine areas for seabirds using behavioural information in combination with distribution patterns. *Biological Conservation,* 156, 22–29.

Chang, W., Cheng, J., Allaire, J. J. et al. (2021). R package “shiny”: web application framework for R (1.6.0). https://CRAN.R-project.org/package=shiny.

Cheng, J. (2020). R package “promises”: Abstractions for promise-based asynchronous programming (1.1.1). https://CRAN.R-project.org/package=promises.

Conroy, M. J., & Peterson, J. T. (2013). *Decision Making in Natural Resource Management: A Structured, Adaptive Approach: A Structured, Adaptive Approach*. John Wiley & Sons, Ltd. DOI: 10.1002/9781118506196

Johnston, A., Cook, A. S. C. P., Wright, L. J., Humphreys, E. M., & Burton, N. H. K. (2014). Modelling flight heights of marine birds to more accurately assess collision risk with offshore wind turbines. *Journal of Applied Ecology*, 51, 31–41.

Judson, O. P. (1994). The rise of the individual-based model in ecology. *Trends in Ecology & Evolution*, *9*(1), 9–14.

Loring, P., McLaren, J., Smith, P., Niles, L., Koch, S., Goyert, H., & Bai, H. (2018). Tracking Movements of Threatened Migratory *rufa* Red Knots in U.S. Atlantic Outer Continental Shelf Waters. OCS Study BOEM 2018-046. US Department of the Interior, Bureau of Ocean Energy Management, Sterling, Virginia. 145 p.

Loring, P., Paton, P., McLaren, J., Bai, H., Janaswamy, R., Goyert, H., Griffin, C., & Sievert, P. (2019). Tracking Offshore Occurrence of Common Terns, Endangered Roseate Terns, and Threatened Piping Plovers with VHF Arrays. OCS Study BOEM 2019-017. US Department of the Interior, Bureau of Ocean Energy Management, Sterling, Virginia. 140 p.

Masden, E. (2015). Developing an avian collision risk model to incorporate variability and uncertainty. *Scottish Marine and Freshwater Science,* 6(14).

McGregor, R., King, S., Donovan, C., Caneco, B., & Webb, A. (2018). A stochastic collision risk model for seabirds in flight. Marine Scotland, Issue 1, Document number: HC0010-400-001.

Nicholson, E., & Possingham, H. P. (2007). Making conservation decisions under uncertainty for the persistence of multiple species. *Ecological Applications*, 17(1), 251-265.

Péron, G., Calabrese, J. M., Duriez, O., Fleming, C. H., García-Jiménez, R., Johnston, A., Lambertucci, S. A., Safi, K., & Shepard, E. L. C. (2020). The challenges of estimating the distribution of flight heights from telemetry or altimetry data. *Animal Biotelemetry*, 8(1), 5.

R Development Core Team (2021). *R: A language and environment for statistical computing* (4.0.3). https://www.r-project.org.

Ronconi, R. A., Allard, K. A., & Taylor, P. D. (2015). Bird interactions with offshore oil and gas platforms: Review of impacts and monitoring techniques. *Journal of Environmental Management*, 147(1), 34-45.

Taylor, P. D., Crewe, T. L., Mackenzie, S. A., Lepage, D., Aubry, Y., Crysler, Z., Finney, G., & Charles, M. 2017. The Motus Wildlife Tracking System : a collaborative research network. *Avian Conservation and Ecology*, 12(1).

Trinder, M. (2017). Offshore wind farms and birds: Incorporating uncertainty in collision risk models: A test of Masden (2015). Natural England Commissioned Reports, 237.