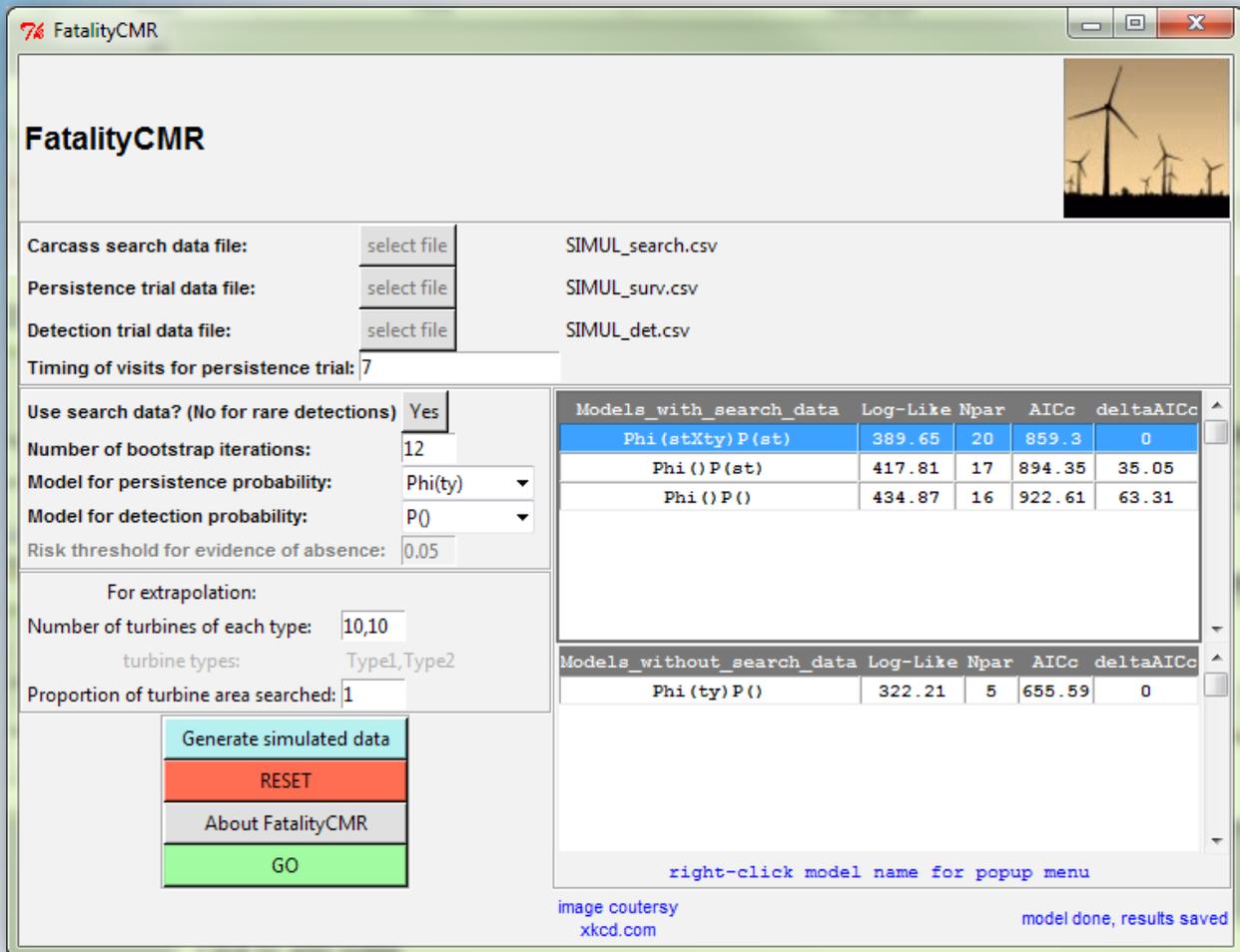




Prepared in cooperation with the U.S. Fish and Wildlife Service

# fatalityCMR—Capture-Recapture Software to Correct Raw Counts of Wildlife Fatalities Using Trial Experiments for Carcass Detection Probability and Persistence Time



**FatalityCMR**

Carcass search data file:  SIMUL\_search.csv

Persistence trial data file:  SIMUL\_surv.csv

Detection trial data file:  SIMUL\_det.csv

Timing of visits for persistence trial: 7

Use search data? (No for rare detections)  Yes

Number of bootstrap iterations: 12

Model for persistence probability: Phi(ty) ▼

Model for detection probability: P() ▼

Risk threshold for evidence of absence: 0.05

For extrapolation:

Number of turbines of each type: 10,10

turbine types: Type1,Type2

Proportion of turbine area searched: 1

Models_with_search_data	Log-Like	Npar	AICc	deltaAICc
Phi (stXty) P (st)	389.65	20	859.3	0
Phi () P (st)	417.81	17	894.35	35.05
Phi () P ()	434.87	16	922.61	63.31

Models_without_search_data	Log-Like	Npar	AICc	deltaAICc
Phi (ty) P ()	322.21	5	655.59	0

right-click model name for popup menu

image courtesy xkcd.com model done, results saved

Techniques and Methods 7–C11

**Cover.** Control window which appears after running fatalityCMR software.

# **fatalityCMR—Capture-Recapture Software to Correct Raw Counts of Wildlife Fatalities Using Trial Experiments for Carcass Detection Probability and Persistence Time**

By Guillaume Péron and James E. Hines

Prepared in cooperation with the U.S. Fish and Wildlife Service

Techniques and Methods 7–C11

**U.S. Department of the Interior  
U.S. Geological Survey**

**U.S. Department of the Interior**  
SALLY JEWELL, Secretary

**U.S. Geological Survey**  
Suzette M. Kimball, Acting Director

U.S. Geological Survey, Reston, Virginia: 2014

For more information on the USGS—the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment, visit <http://www.usgs.gov> or call 1–888–ASK–USGS.

For an overview of USGS information products, including maps, imagery, and publications, visit <http://www.usgs.gov/pubprod>

To order this and other USGS information products, visit <http://store.usgs.gov>

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Although this information product, for the most part, is in the public domain, it also may contain copyrighted materials as noted in the text. Permission to reproduce copyrighted items must be secured from the copyright owner.

Suggested citation:

Péron, Guillaume, and Hines, J.E., 2014, fatalityCMR—Capture-recapture software to correct raw counts of wildlife fatalities using trial experiments for carcass detection probability and persistence time: U.S. Geological Survey Techniques and Methods 7–C11, 14 p., <http://dx.doi.org/10.3133/tm7C11>.

ISSN 2328-7055 (online)

## Preface

This report documents a user interface for the analysis of data resulting from carcass search and experimental trials in the context of the environmental assessment of wind farms. The interface was developed in the computer language R. The interface has been tested for accuracy by using multiple datasets, and the software offers the possibility to generate simulated data. If users find or suspect errors, please contact James E. Hines of the U.S. Geological Survey (USGS) at [jhines@usgs.gov](mailto:jhines@usgs.gov) to report any errors that may be found or suspected.

The authors have made every effort to ensure that the spreadsheet interface is error free; however, errors may exist in the spreadsheet interface. The distribution of the spreadsheets does not constitute any warranty by the USGS, and the USGS assumes no responsibility for any errors that may exist, nor for any consequences thereof.

## Acknowledgments

The authors thank the U.S. Fish and Wildlife Service for funding the work leading to this software, B.A. Millsap and E. Bjerre for guidance and useful discussions of carcass surveys, M. Huso and D. Dalthorp for important insight into the evidence of absence routine, M.J. Clement for useful suggestions about an earlier version of this manual, and J.D. Nichols and W.L. Kendall for instigating the project and general support.

## Contents

Preface .....	iii
Acknowledgments .....	iii
Abstract .....	1
Introduction.....	1
Data Structure .....	2
Carcass Search Data File .....	2
Persistence Trial Data File.....	3
Detection Trial Data File.....	3
Data Simulation .....	5
Data Analysis: Software Options and Interface .....	6
Timing of Visits to Planted Carcasses for the Persistence Trial (4).....	6
Option of Using Search Data or Not (5).....	6
Bootstrap Variance Computation (6) .....	8
Biological Effects Acting on Detection and Persistence Probabilities (7) and (8).....	8
Extrapolation to a Larger Area (10).....	8
Model Output.....	8
If the Search Data are Used .....	8
If the Search Data are Not Used.....	9
About fatalityCMR.....	11
Utility .....	11
References Cited.....	11
Appendix 1. Simulation study to illustrate the expected precision of fatality estimates.....	13

## Figures

1. Screen capture showing commands needed to open fatalityCMR from an R session once the necessary packages have been installed .....	2
2. Screen capture showing example of the formatting of the carcass search data file .....	3
3. Screen capture showing example of the formatting of the persistence trial data file .....	4
4. Screen capture showing example of the formatting of the detection trial data file .....	4
5. Screen capture showing the data generation interface with default parameter values.....	5
6. Screen capture showing the fatalityCMR main interface window.....	7
7. Screen capture showing output file for models that use the carcass search data.....	9
8. Screen capture showing output file for models that do not use carcass search data .....	10
9. Screen capture showing pop-up window that appears after clicking on “About FatalityCMR” .....	11

# fatalityCMR—Capture-Recapture Software to Correct Raw Counts of Wildlife Fatalities Using Trial Experiments for Carcass Detection Probability and Persistence Time

By Guillaume Péron and James E. Hines

## Abstract

Many industrial and agricultural activities involve wildlife fatalities caused by collision, poisoning, or other involuntary harvest, including wind turbines, highway networks, utility networks, tall structures, and pesticides. Affected wildlife may have official protection, including a monitoring requirement. Carcass counts are typically conducted to quantify the number of fatalities, but they need to be corrected for carcass persistence time (as influenced by removal by scavengers and decay) and detection probability (searcher efficiency). This report introduces new software developed by the U.S. Geological Survey (USGS) in cooperation with the U.S. Fish and Wildlife Service that fits a superpopulation capture-recapture model to raw count data. It uses trial data to estimate detection and daily persistence probabilities and combine these estimates with the carcass count data in an integrated model. A recurrent issue is that fatalities of rare, protected species are infrequent, in which case the software offers the option to switch to an “evidence of absence” mode based on binomial laws and estimate the probability of not finding a carcass. The software allows users to distinguish between different location types (for example, different vegetation cover or different technical properties of the devices causing fatalities), as well between two carcass age-classes or states with transition between those classes (for example, fresh and old). A data simulation feature can be used at the planning stage to optimize sampling design. Resulting mortality estimates can be used to (1) quantify the required amount of compensation, (2) inform mortality projections for proposed development sites, (3) inform decisions about management of existing sites, and (4) improve the design of carcass search protocols and trial experiments.

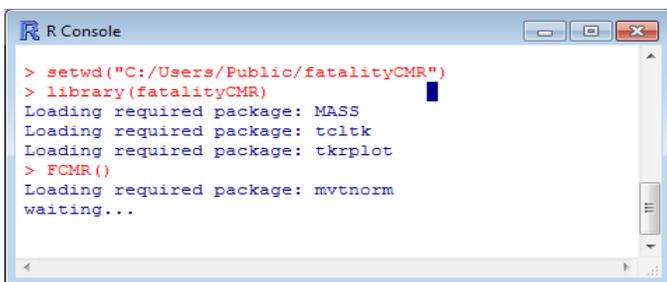
## Introduction

The environmental impact of human activities can be quite direct, such as when human-made structures or

compounds molecules kill wildlife. Collisions with vehicles, for example, have been deemed a major limitation of turtle population growth in the United States (Gibbs and Shriver, 2002), and mortality by pesticide ingestion is claimed to reduce the abundance of several bird species (Mineau and Whiteside, 2013). Recently, collisions with wind turbines have been identified as an important threat to bird and bat species (Drewitt and Langston, 2006; Bell and Smallwood, 2010). A major concern for stakeholders is that some affected species benefit from official protection. For example, in the United States, each and every death of a migratory bird caused by human activities is illegal in the absence of a permit (Migratory Bird Treaty Act of 1918). To assess the extent of the issue, it is often possible to directly document the “take” of protected species by searching for carcasses near the mortality source. The raw carcass counts, however, represent minimum estimates of mortality and need to be corrected for (1) persistence probability (loss of carcasses to scavengers, decay, weather, and tide) and (2) detection probability (observers may miss some of the carcasses that are present during surveys) (Arnett and others, 2008; Smallwood and others, 2010; Huso, 2011). In most cases, these probabilities are estimated by using trial experiments—that is, planting carcasses and monitoring their subsequent detection and fate. Trial experiments can also be based on naturally occurring carcasses.

Péron and others (2013) recently described a superpopulation capture-recapture approach (Crosbie and Manly, 1985; Schwarz and Arnason, 1996) that corrects for both persistence and detection. The origin of the method was developed with wind farms in mind, and this origin is reflected in the wording below, but the method can be applied to any type of wildlife fatality problem. This report introduces software developed by the U.S. Geological Survey (USGS) in cooperation with the U.S. Fish and Wildlife Service that allows implementation of the approach without extensive familiarity with capture-recapture models or any programming language. The key feature that distinguishes the statistical modeling in this software from other fatality estimators (Bernardino and others, 2013) is the use of the capture-recapture modeling framework, which does not force strong assumptions about “bleed through” (carcasses

that persist undetected over multiple sampling occasions; Warren-Hicks and others, 2013). The software also accommodates differences among turbine “types” and carcass “states.” Turbine types can account for different technologies or different vegetative cover, both factors known to affect persistence and (or) detection probabilities (Smallwood, 2013; Warren-Hicks and others, 2013). Turbines of different types can also pose different risk levels for wildlife (for example, turbines located on a ridge top as opposed to turbines located on a gentle slope) (de Lucas and others, 2012). Carcass states account for the effect of decay or desiccation on detection and persistence probability and, therefore, for departures from the exponential distribution of time-to-removal and time-to-detection of the carcasses. A maximum of two states is allowed (for example, fresh and dry, or intact and partially scavenged). There is currently little evidence about the effect of carcass “ageing” on fatality estimation. It is hoped that by making age effects readily available in the estimation framework, this software can help determine whether this variable (carcass state) should be routinely recorded and promote the adoption of common protocols for carcass age determination. The software works from within the software package R (R-Development-Core-Team, 2010) as function `FCMR()` in the package `fatalityCMR` (fig. 1). `fatalityCMR` is freely available at <http://www.mbr-pwrc.usgs.gov/software/fatalityCMR.html>. `fatalityCMR` relies on packages `tcltk`, `tkplot`, `mvtnorm`, and `MASS`, all of which are readily available from the R-CRAN Web site <http://cran.r-project.org/> and will be automatically installed if not already present (the first session should be run with administrator rights in this case). The statistical modeling on which this software is based is detailed in Péron and others (2013).



```

R Console
> setwd("C:/Users/Public/fatalityCMR")
> library(fatalityCMR)
Loading required package: MASS
Loading required package: tcltk
Loading required package: tkrplot
> FCMR()
Loading required package: mvtnorm
waiting...

```

**Figure 1.** Commands needed to open `fatalityCMR` from an R session once the necessary packages have been installed. (Note that the parenthesis must be used when `FCMR()` is prompted.)

## Data Structure

`fatalityCMR` uses data on carcass searches, persistence trials, and detection trials to generate estimates of total fatalities. First, operators select a set of turbines and decide upon (1) a fixed area below each turbine, which they are going to search for carcasses; (2) a number  $T$  of sampling occasions; and (3) the time intervals between these occasions. The software does not allow the protocol to be varied at different turbines; each turbine in the set must be searched for carcasses at each occasion. Field technicians record the number of previously undetected carcasses, and, optionally, the type of turbine and the (dichotomous) age-class of the carcasses. No missing data (for example, turbines that are not visited on a given occasion) are allowed. Second, the software requires data from persistence trials and detection trials. Persistence trials consist of regular checks of planted carcasses for a fixed number of days or until they disappear; this information is hereafter called the “persistence trial data.” In detection trials, trial carcasses are placed throughout the survey area (or a portion of it), and investigators who are ignorant of the number and locations of carcasses search the area and record how many trial carcasses are detected. This information is hereafter called the “detection trial data.” The following three sections describe how these three sets of data are formatted and entered into the software.

## Carcass Search Data File

The carcass search data must be in a comma-separated-variable CSV file with specific formatting. The first line should contain specific headers, namely “State,” “ID,” “Type,” and “A1,” “A2,” ..., “AT,” where  $T$  is the number of sampling occasions (different searches).  $T = 12$  in figure 2. Headers are case sensitive, so that entering a file with “type” instead of “Type” will trigger an error. “State” is optional. The second line must contain interval durations in days in columns “A1,” “A2,” ..., “AT-1,” where the number entered under “A1” is the interval between sampling occasions 1 and 2, etc. The last cell under “AT” contains “NA” because there is no interval following the last occasion. Then, starting in the third line of the file, enter one line per turbine and per carcass state (“fresh” and “old” states). First enter as many lines as there are turbines in the study area. Under “ID,” enter the turbine identification numbers or any strings of characters to identify the turbines. Under “Type,” enter the turbine type; if there is no turbine type, enter the same string of characters throughout this column (for example, “NoType”). Note that the exact same turbine type denominations must also be used in the persistence and detection trial data files. Under “State,” enter “fresh” or any other appropriate name. Under “A1,” “A2,” ..., “AT,”

State	ID	Type	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
NA	intdur	NA	2	2	2	2	2	2	2	2	2	2	2	2 NA
fresh	1	Type1	0	0	0	1	0	1	0	1	0	2	0	2
fresh	2	Type1	0	3	0	1	2	1	2	1	1	0	1	0
fresh	3	Type1	2	1	0	0	1	1	0	0	0	0	0	0
fresh	4	Type1	0	2	0	3	1	0	0	1	0	0	0	3
fresh	5	Type1	0	1	0	0	1	1	0	0	0	0	0	1
fresh	6	Type1	0	1	0	0	0	0	0	0	1	2	1	1
fresh	7	Type1	0	1	1	0	2	1	1	1	0	0	0	0
fresh	8	Type1	3	0	0	0	0	0	0	2	0	1	0	0

**Figure 2.** Example of the formatting of the carcass search data file. Note: If different turbines are subjected to different search protocols (for example, different time intervals or a different number of time occasions), it is necessary to prepare one data file per type of search protocol and run separate analyses.

enter the number of fresh carcasses found under each turbine at each sampling occasion. All turbines must be searched at all sampling occasions; do not enter NA or 0 if the turbine was not searched. Then, enter a second series of lines (one per turbine). Under “State,” enter “dry” or the name of the second state. Under “A1,” “A2,” ..., “AT,” enter the number of previously undetected carcasses that were assigned to the second age-class upon detection. If carcasses are not assigned to age-classes, then remove the “State” column, and enter only one line per turbine. An example data file can be obtained by using the data simulation capacity.

## Persistence Trial Data File

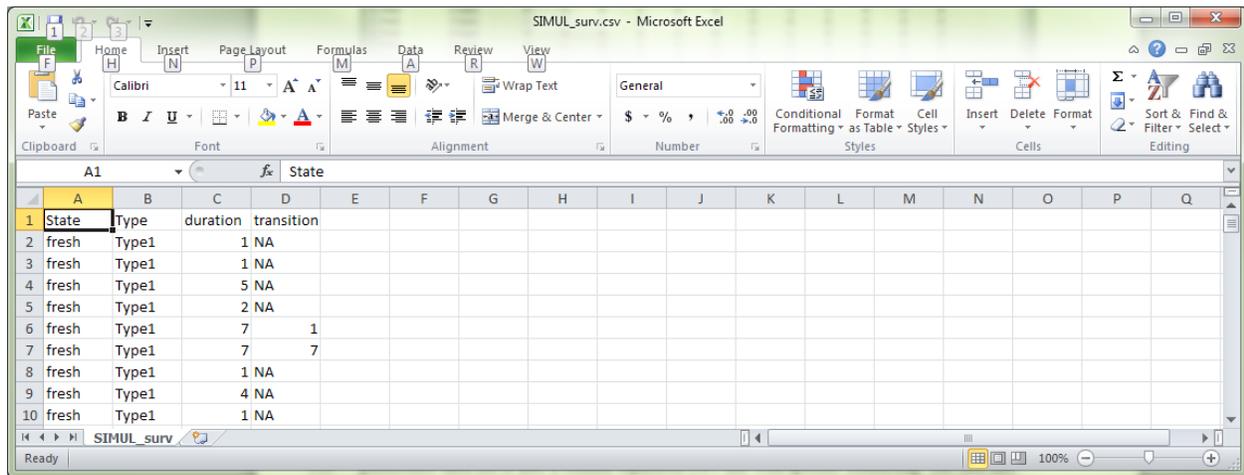
All trial carcasses used in the persistence trial must be subjected to the same protocol; they must be checked for continued persistence on the same days. The software does allow different carcasses to be checked on different days, but the duration of the interval can change throughout the study period (for example, daily checks for 2 weeks, or one check every 2 days for 10 days). The persistence data file (fig. 3) is also a comma-separated-variable (CSV) file with specific formatting. The first line contains the headers “State,” “Type,” “duration,” and “transition,” still case-specific. Below the header, there should be one line per trial carcass. Under “State,” enter the state of the carcass at the beginning of the trial. Under “Type,” enter the type of the turbine associated

with the trial carcass. Under “duration,” enter the persistence time in days. If the carcass is still present at the last visit, enter the duration of the persistence trial. Under “transition,” enter the time to transition, which is the last occasion at which the carcass was considered to be in the first state. For example, if the carcasses are checked on days 2, 3, 4, 6, 8, and 10, and on day 8 the focal carcass is considered to have transitioned to the second state, then enter 6 under “transition” for that carcass. If no transition is recorded, or if the carcass was already in the second state at the start of trial, enter “NA” under “transition.” Any state × type combination found in the carcass search data must also be represented in the persistence data file. An example persistence data file can be obtained by using the data simulation capacity.

## Detection Trial Data File

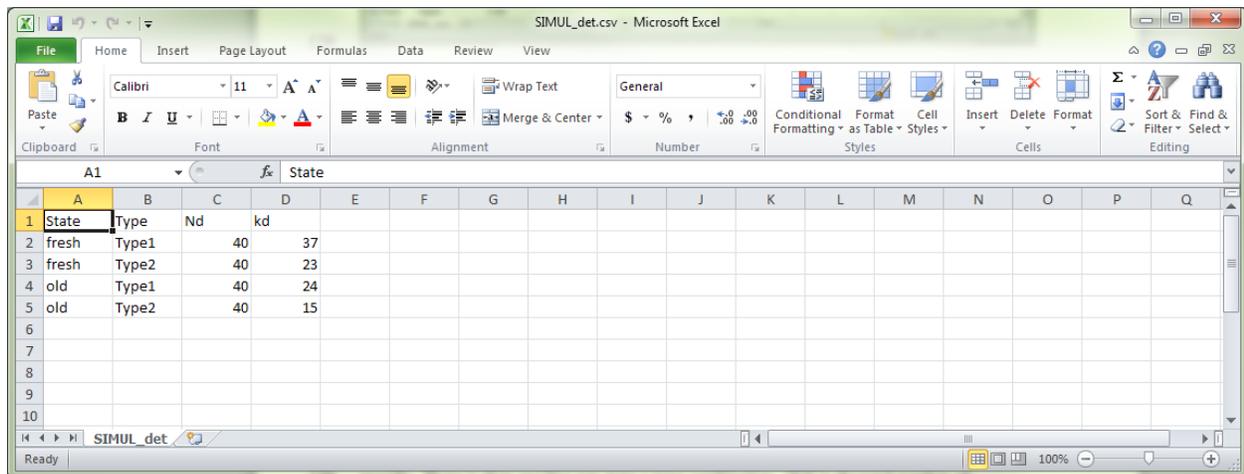
The detection data file (fig. 4) is also a comma-separated-variable (CSV) file with specific formatting. The first line contains the headers “State,” “Type,” “Nd,” and “kd,” also case-specific. There should be one line per state × type combination. “Nd” is the total number of carcasses that were planted in this combination. “kd” is the number of those carcasses that were detected. Any state × type combination found in the carcass search data must also be represented in this file. An example data file can be obtained by using the data simulation capacity.

## 4 fatalityCMR—Capture-Recapture Software to Correct Raw Counts of Wildlife Fatalities



State	Type	duration	transition
fresh	Type1	1	NA
fresh	Type1	1	NA
fresh	Type1	5	NA
fresh	Type1	2	NA
fresh	Type1	7	1
fresh	Type1	7	7
fresh	Type1	1	NA
fresh	Type1	4	NA
fresh	Type1	1	NA

Figure 3. Example of the formatting of the persistence trial data file.



State	Type	Nd	kd
fresh	Type1	40	37
fresh	Type2	40	23
old	Type1	40	24
old	Type2	40	15

Figure 4. Example of the formatting of the detection trial data file. Note: The persistence and detection trial data files can be used for projects other than the one from which they originate, with the assumption that persistence and (or) detection probabilities are the same.

## Data Simulation

The fatalityCMR software can generate datasets with two types of turbines and two carcass states (fig. 5). The different parameters as they are designated in the data simulation window are reviewed below; this review also serves as a review of model parameters (see Péron and others, 2013).

- Number of search sessions: The number of sampling occasions during which the area is searched for new carcasses.
- Interval duration(s): The duration in days between the successive sampling occasions. If fewer than  $T - 1$  numbers are entered (separated by commas), then the interval duration is assumed to be fixed, and only the first value is used.
- Number of turbines that are searched for carcasses.
- Total number of fatalities (superpopulation size): The total number of carcasses that are available for detection during at least one of the sampling occasions and that are not already dry at the first occasion.
- Number of dry carcasses already on the ground at  $T = 1$ : The number of carcasses that are already dry at the first occasion.
- Proportion of sampled turbines in type 1: One minus this number indicates the proportion of sampled turbines in type 2.
- Temporal distribution: Choose either “uniform” to fix the entry rate of new carcasses to a constant daily

The screenshot shows a window titled "76 FatalityCMR simulations" with a list of parameters and their values. At the bottom, there are two buttons: "Reset values to defaults" (red) and "Generate input files" (green).

Number of search sessions:	12	
Interval duration(s):	2	days
Number of turbines:	20	
Total number of fatalities (superpopulation size):	20	
Number of dry carcasses already on the ground at t=1:	150	
Proportion of sampled turbines in type 1:	0.5	0 < <1
Temporal distribution:	uniform	
Detection prob.state=Fresh,type=1:	0.8	0 < <1
Detection prob.state=Fresh,type=2:	0.5	0 < <1
Detection prob.state=Dry,type=1:	0.6	0 < <1
Detection prob.state=Dry,type=2:	0.4	0 < <1
Daily persistence prob.state=Fresh,type=1:	0.8	0 < <1
Daily persistence prob.state=Fresh,type=2:	0.8	0 < <1
Daily persistence prob.state=Dry,type=1:	0.9	0 < <1
Daily persistence prob.state=Dry,type=2:	0.9	0 < <1
Daily state transition probability:	0.01	0 < <1
Sample size of detection trial:	40	
Sample size of survival trial:	40	
Length of persistence trial:	7	days
Risk factor:	0.8	

**Figure 5.** The data generation interface with default parameter values. (This pop-up window appears after clicking on “Generate Simulated Data.”)

value, or “pulsed” to mimic a pulsed phenology centered in the middle of the study period.

- Detection prob.state = Fresh, type = 1: The detection probability of fresh carcasses under turbines of type 1.
- Detection prob.state = Fresh, type = 2: The detection probability of fresh carcasses under turbines of type 2.
- Detection prob.state = Dry, type = 1: The detection probability of dry carcasses under turbines of type 1.
- Detection prob.state = Dry, type = 2: The detection probability of dry carcasses under turbines of type 2.
- Daily persistence prob.state = Fresh, type = 1: The daily persistence probability of fresh carcasses under turbines of type 1.
- Daily persistence prob.state = Fresh, type = 2: The daily persistence probability of fresh carcasses under turbines of type 2.
- Daily persistence prob.state = Dry, type = 1: The daily persistence probability of dry carcasses under turbines of type 1.
- Daily persistence prob.state = Dry, type = 2: The daily persistence probability of dry carcasses under turbines of type 2.
- Daily state transition probability: The daily probability that a fresh carcass becomes dry.
- Sample size of detection trial: The number of carcasses of each type  $\times$  state combination that is used in the detection trial. If only one value is entered, it is considered to represent the number of carcasses used in each of the trial combinations, so that the total trial sample size is four times the entered value.
- Sample size of survival trial: The number of carcasses of each type  $\times$  state combination that is used (initially placed) in the survival trial. If only one value is entered, it is considered to represent the number of carcasses used in each of the trial combinations, so that the total trial sample size is four times the entered value.
- Length of persistence trial: The number of days each trial carcass is monitored for persistence.
- Risk factor: The multiplication factor used to assign carcasses to one or the other turbine type. Risk factors less than 1 indicate that type 1 is more dangerous (produces more dead birds), whereas risk factors greater than 1 indicate that type 2 is more dangerous.

## Data Analysis: Software Options and Interface

In the text below, the various options available to tailor the analysis are reviewed. In the following subheadings, numbers in parentheses refer to the numbers shown in figure 6.

### Timing of Visits to Planted Carcasses for the Persistence Trial (4)

The software gives the option to either provide a single number in this space, which is interpreted as the duration of the entire trial, in days, with locations visited every day, or to enter a list of increasing integers, which are the days on which the locations were visited (for example, 1, 2, 3, 6, 7 if trial carcasses were not checked on days 4 and 5 of a week-long trial). Beware that there is no data checking for this feature. If the persistence trial data file indicates that a carcass was last recorded on a day when the location was not visited as specified in the timing of visits, no error message is generated, and this carcass will be considered to have persisted until the final day of the trial.

### Option of Using Search Data or Not (5)

The total number of detected carcasses,  $C$ , may be small. In this case, the general method (use search data=“yes” option) performs poorly. It is therefore recommended to choose the “No” option if  $C$  is less than or equal to five carcasses. The procedure then relies on the estimate for  $P_0$ , the probability that one carcass of unknown entry date is not detected. The estimate for  $P_0$  depends only on the trial data and on the timing of the sampling occasion; it does not make use of the result of the carcass searches. Under the assumption that each fatality event is independent of others, the software then computes two quantities. First, it provides the maximum number of carcasses that could have escaped detection, with a user-defined error risk  $\alpha$ —that is,  $N_{\max}$  such that

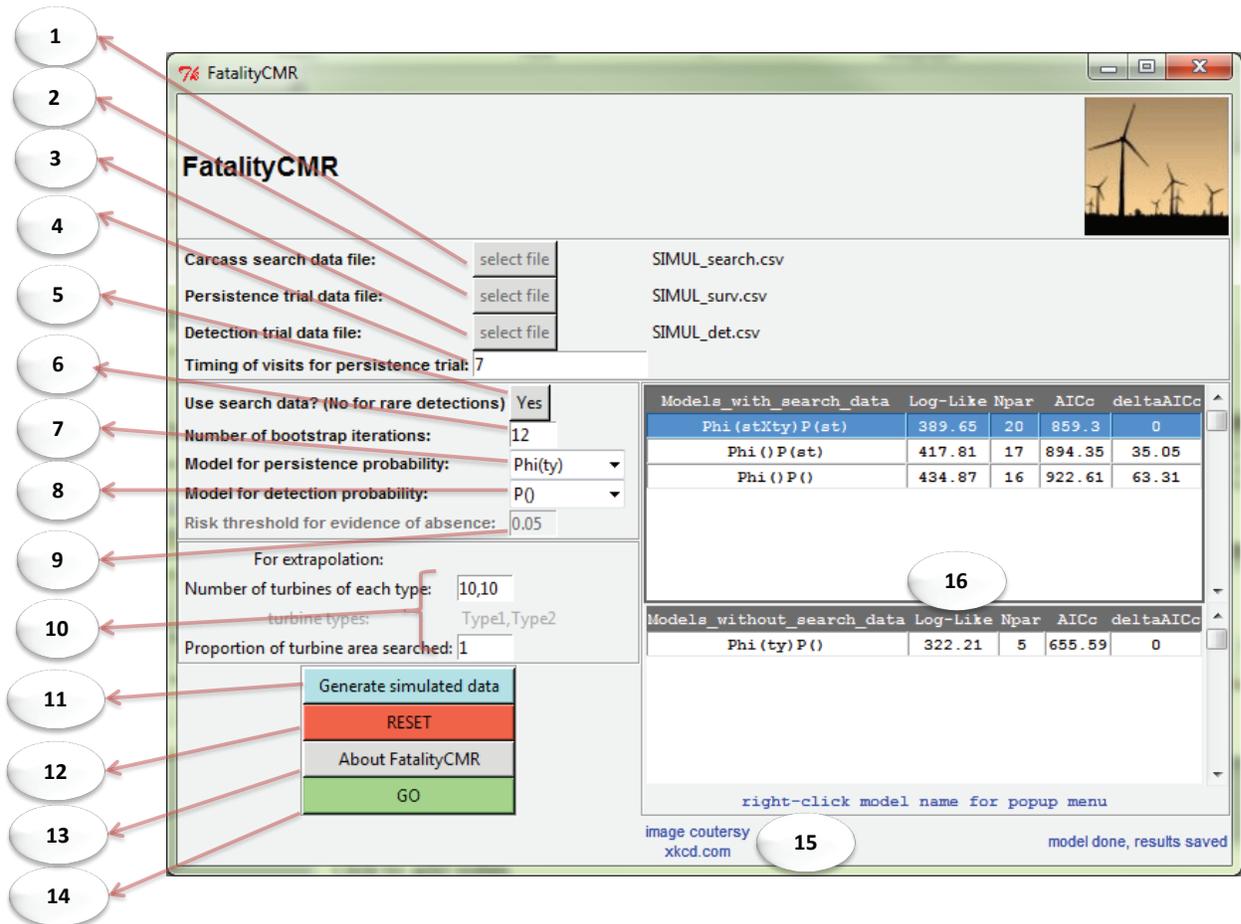
$$\binom{N_{\max}}{C} (1 - P_0)^C P_0^{N_{\max} - C} = \alpha.$$

By default,  $\alpha$  is set to 0.05. Second, it computes an ad hoc estimate of the number of fatalities that is,

$$\hat{N} = \sum_{n=C}^{\infty} [n \cdot \Pr(n|C)],$$

where in practice, the summation is stopped when the following condition was met:

$$\Pr(n|C) < 10^{-9}.$$



**Figure 6.** The fatalityCMR main interface window: 1, Data entry for the carcass search data; 2, Data entry for the persistence trial data; 3, Data entry for the detection trial data; 4, Data entry for the duration of the persistence trial (in days); 5, Option to use the carcass search data or not. If “No” is selected, then the software switches to “evidence of absence” mode (see text for details); 6, Number of replications in the parametric bootstrap for variance computation; 7 and 8, Model structure for persistence and detection probabilities, respectively; 9, Risk threshold for the “evidence of absence” routine (see text for details). This section becomes active only if the “No” option is chosen at (5); 10, Parameters for the extrapolation to a larger area (for example, whole wind farm) (see text for details); 11, Button to enter the data simulation interface; 12, Reset button to erase all previously entered data, as well as the results of previous analyses of these data; 13, “About” button for e-mail contact and reference; 14, “Go” button to launch the analysis once the data have been entered and the correct options chosen; 15, Status bar with information printed in blue during computation. This is also where information is displayed at the data entry stage; 16, Akaike Information Criterion table with elements to compare model fit and parsimony. Right-click on model names to access a menu and display either the text output file or the graph of carcass distribution between turbine types.

## Bootstrap Variance Computation (6)

To compute standard errors for some of the derived quantities, the software performs a parametric bootstrap procedure using the estimated variance-covariance matrix of the parameters. The user can change the number of replications or set it to zero. In the bootstrap procedure, the software generates parameter values from the multivariate normal distribution centered around the estimates; it computes derived quantities such as  $P_0$  and then estimates the standard deviation of the set of simulated values.

## Biological Effects Acting on Detection and Persistence Probabilities (7) and (8)

Detection and persistence probabilities may vary with turbine type or carcass state. The user can choose from four pre-defined model structures for both persistence and detection probability: constant probability, denoted by a blank; one parameter per turbine type, denoted by “ty”; state-specific probability, denoted by “st”; and two-way interaction of type and state, denoted by “tyXst.” The effects are modeled on the logit scale. A “biological model” is then combinations of model structure for persistence and for detection. Biological models vary in their fit to the data. This fit and model parsimony are assessed using the Akaike Information Criterion and associated metrics (Burnham and Anderson, 2002), which are displayed in the model selection table (16 in figure 6). Estimates from the best ranking model should be considered to be the most reliable.

## Extrapolation to a Larger Area (10)

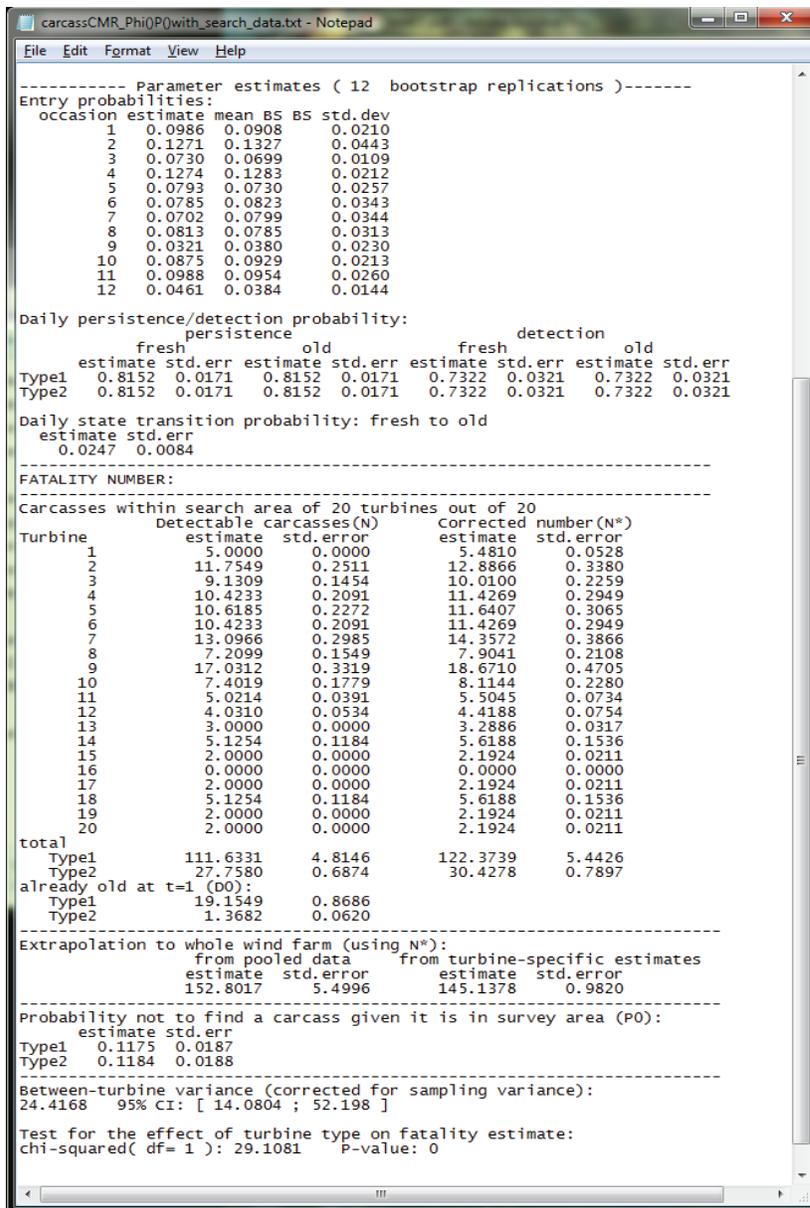
The software can perform simple extrapolations (cross multiplications) to extend the results from the search area to the entire wind farm. In the first space, the user can enter the total number of turbines of each type in the wind farm, separated by a comma. These numbers should include both the turbines selected for the surveys and those that were not monitored. If all turbines were searched, or if the total number is not known, leave this space blank. The parameter in the second space is a correction factor that accounts for unsearched areas within each turbine plot. This value can represent inaccessible areas (dense vegetative cover, crag, water) or areas beyond the search radius. For example, if it is known that  $1 - a = 20$  percent of the carcasses fall beyond the search radius, and that  $1 - b = 15$  percent of the area within the search radius could not be searched, then the correction factor is  $a * b = 0.68$ . Note, however, that inaccessible areas are assumed to be uniformly distributed with respect to distance from the turbine pole; the correction should be amended if unsearchable areas are disproportionately close to, or far from, the turbine pole. Smallwood (2013) reviews a substantial amount of information available to compute  $a$  and  $b$ . If only the estimate of the number of carcasses that fell *within the search area* is desired, enter 1 as the correction factor.

## Model Output

The fatalityCMR software provides (1) a model comparison table in which the Akaike Information Criterion, corrected for small sample sizes, is used to compare the fit and parsimony of the different models (Burnham and Anderson, 2002); (2) one text output file per model (fig. 7), with parameter estimates, derived quantities, data summary, and characteristics of the model; and (3) a graph displaying the observed distribution of carcasses between turbine types.

## If the Search Data are Used

The quantity most users will be interested in is the estimate of superpopulation size—that is, the number of carcasses that were available for detection at least once during the period between the initial and final sampling occasions. This estimate is located under “Detectable carcasses (N)” in the output file. Turbine-specific estimates are computed by fixing persistence and detection parameters to their values obtained when the pooled data are used and running the model using only those carcasses found under the focal turbine. A correction is then performed for carcasses that enter and are removed by scavengers during the same interval and therefore are never available for detection (see equation 8 in Péron and others, 2013). These corrected estimates are found under “Corrected number (N\*).” The extrapolated fatality numbers are located under “Extrapolation to whole wind farm.” They are computed by using corrected estimates (N\*). Two different extrapolations are performed. The first uses the numbers estimated from the pooled data and corrects for the potential nonrepresentative sampling of different turbines types. The second uses the turbine-specific estimates and does not correct for the potential nonrepresentative sampling of different turbines types. The standard error for the second extrapolation is computed by using equation 5 in Skalski (1994). The probability of not finding a carcass of unknown entry date ( $P_0$ ) is located under “Probability of not finding a carcass.” Note that this probability is conditional on the carcass being available for detection during at least one sampling occasion and within the search area. Last, the software provides two measures of between-turbine variation in the number of fatalities. The first measure is the variance in the turbine-specific fatality estimates, corrected for sampling error by using equation 4.2 in Burnham and others (1987). Note that if fatalities are rare, the algorithm commonly fails and displays “NA” for the variance estimate, indicating that either the variance is null or the sample is too small. The 95-percent confidence interval is computed by using the chi-squared distribution (Burnham and others, 1987). The second measure of between-turbine variation in mortality is a chi-squared test for the effect of turbine type on fatality number. It tests whether the distribution of estimated fatalities deviates from the expected distribution, given the sampling stratification. There is also an option to display a graph of the number of observed fatalities with respect to the number of expected fatalities for different turbine types.



- One entry probability value per sampling occasion
- One persistence and detection value for each turbine type × carcass state combination
- Daily transition probability between carcass states
- Turbine-specific estimated numbers of fatalities
- Estimated total number of fatalities per turbine type
- Estimated number of “old” carcasses at first sampling occasions
- Extrapolation to larger area
- Probability that a carcass is missed
- Variation between turbines

**Figure 7.** Output file for models that use the carcass search data (partial view lacking data summary and model specification).

### If the Search Data are Not Used

A reduced amount of information is included in the output (fig. 8) to avoid small sample biases and errors. The key quantity to estimate is the probability of not finding a carcass that was available for detection in the search area during at least one sampling occasion ( $P_0$ ). This quantity is located under “Probability of not finding a carcass (P0)” in the output file. It is computed for two different phenologies of carcass entry: uniform—that is, a fixed daily probability that a fatality

occurs, or pulsed—that is, the carcass entry rate peaks in the middle of the study period.  $N_{max}$  (see “Option of using search data or not” above for definition of this notation) is located under “Maximum number of fatalities,” and  $\hat{N}$  is located under “Ad hoc estimate of the number of fatalities”. Note that values must be summed over the different turbine types to obtain the total estimated number of fatalities. The first set of values corresponds to carcasses within the search area; the second set corresponds to the values extrapolated by using the correction factors (see “Extrapolation to a larger area,” above).

```

carcassCMR_Phi(OP)without_search_data.txt - Notepad
File Edit Format View Help
carcassCMR ( version:2.0.2(2013-06-12) )      Wed Sep 11 13:33:23 2013

----- Data summary -----
data filename: SIMUL_search.csv
surv filename: SIMUL_surv.csv
det filename: SIMUL_det.csv

Number of turbines of each type: 10 10
Duration of the persistence trial: 7 days

12 sampling occasions
20 sampled turbines out of a total of 20 turbines in the wind farm
Type1 : 10 sampled turbines
Type2 : 10 sampled turbines

Numbers of carcasses found:
  by occasion: 1 0 1 2 3 0 0 0 0 2 5
  by turbine : 1 1 1 0 1 2 1 1 1 2 0 0 0 0 1 0 1 0 1 0
  Total      : 14

Persistence trial (duration: 7 days):
  160 carcasses which persisted 4.75 days on average
Detection trial: 160 carcasses of which 98 were found

Model: Phi(OP)without search data

Model Deviance= 305.3133
Number of estimated parameters: 4
Effective sample size: 40

----- Parameter estimates ( 12 bootstrap replications )-----
Daily persistence/detection probability:
      persistence                detection
      fresh      old      fresh      old
estimate std.err estimate std.err estimate std.err estimate std.err
Type1  0.8677  0.0194  0.8677  0.0194  0.6106  0.0341  0.6106  0.0341
Type2  0.8677  0.0194  0.8677  0.0194  0.6106  0.0341  0.6106  0.0341

Daily state transition probability: fresh to old
      estimate std.err
      0.0173  0.0136

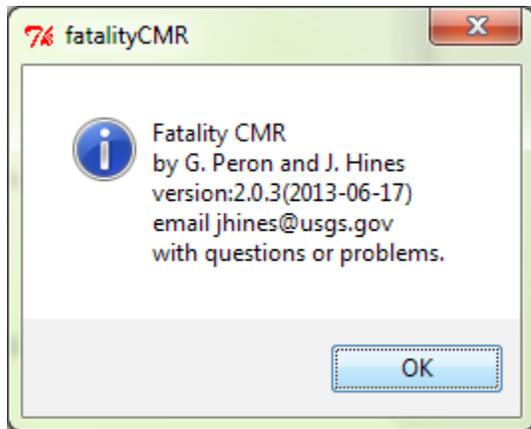
-----
Probability of not finding a carcass given it is in surveyed area (P0):
-----
Uniform entry rate Pulsed entry rate SE(unif) SE(pulse)
Type1      0.1653      0.1386  0.0241  0.0231
Type2      0.1653      0.1386  0.0241  0.0231
-----
Maximum number of fatalities (with risk threshold 0.05 )
-----
Uniform entry rate Pulsed entry rate SE(unif) SE(pulse)
Type1     16.5329     15.8343  0.6448  0.5998
Type2      6.0420      5.7131  0.3018  0.2835
-----
Ad-hoc estimate of the number of fatalities
-----
Uniform entry rate Pulsed entry rate SE(unif) SE(pulse)
Type1     15.3298     14.2970  0.9903  0.8753
Type2      4.2762      3.9656  0.3034  0.2671
-----
-----%-----
Extrapolations to whole wind farm -----
-----
Maximum number of fatalities (with risk threshold 0.05 )
-----
Uniform entry rate Pulsed entry rate SE(unif) SE(pulse)
Type1     16.5329     15.8343  0.6448  0.5998

```

**Figure 8.** Output file for models that do not use carcass search data. (The end of the output with the extrapolations to whole wind farm is omitted from this view.)

## About fatalityCMR

Figure 9 displays the version number (2.0.3) of the software and contact details for James E. Hines, who maintains it.



**Figure 9.** Pop-up window that appears after clicking on “About FatalityCMR.”

## Utility

It is envisioned that fatalityCMR will be used for four primary purposes. First, fatality estimates can be used to quantify the level of required mitigation. For example, wind projects that kill eagles are commonly required to retrofit utility poles. Older utility poles cause eagle deaths by electrocution, and retrofitting them compensates for eagle deaths caused by the wind farm. If a direct relation is assumed between the number of retrofitted poles and the number of eagle deaths avoided, one can estimate the number of utility poles that need to be retrofitted. Second, proposed wind projects are commonly required to assess their potential effect on protected species as part of a permitting process. Predicting eagle take requires linking eagle use (measured in eagle-minutes in the U.S. Fish and Wildlife Service protocol) to eagle fatalities. In a Bayesian framework, this means updating the predictive model with new fatality estimates when those become available. Third, fatality estimates determined by using the software can be used adaptively to modify the operation of existing wind farms. For example, wind farms may have to alter or cease their operation once a threshold of take has been reached. Fatality estimates can be used to compare the risk to eagles of different operation modes—for example, varying cut-off wind speeds, or times of day or year during which turbines are shut down. Fourth, the software, and specifically its data simulation capacity, can be used to design carcass search protocols and trial experiments (see appendix A). There

is a tradeoff between field effort, cost, and expected precision of the fatality estimates (Péron and others, 2013). This tradeoff depends on the specifics of the wind project under consideration, such as whether or not trial carcasses are easy to obtain, the remoteness or ease of access of the site, and the expected number of fatalities.

## References Cited

- Arnett, E.B., Brown, W.K., Erickson, W.P., Fiedler, J.K., Hamilton, B.L., Henry, T.H., Jain, A., Johnson, G.D., Kerns, J., Koford, R.R., Nicholson, C.P., O’Connell, T.J., Piorkowski, M.D., and Tankersley, R.D., 2008, Patterns of bat fatalities at wind energy facilities in North America: *The Journal of Wildlife Management*, v. 72, p. 61–78.
- Bell, D.A., and Smallwood, K.S., 2010, Birds of prey remain at risk: *Science*, v. 330, p. 913–913.
- Bernardino, J., Bispo, R., Costa, H., and Mascarenhas, M., 2013, Estimating bird and bat fatality at wind farms: a practical overview of estimators, their assumptions and limitations: *New Zealand Journal of Zoology*, v. 40, p. 63–74.
- Burnham, K.P., and Anderson, D.R., 2002, Model selection and multimodel inference: a practical information-theoretic approach (2d ed.): New York, Springer-Verlag, 488 p.
- Burnham, K.P., Anderson, D.R., White, G.C., Brownie, C., and Pollock, K.H., 1987, Design and analysis method for fish survival experiments based on release-recapture: *American Fisheries Society Monographs* 5 (1987): American Fisheries Society, Maryland, U.S.A. 437 p.
- Crosbie, S.F., and Manly, B.F.J., 1985, Parsimonious modeling of capture-mark-recapture studies: *Biometrics*, v. 41, p. 385–398.
- de Lucas, M., Ferrer, M., Bechard, M.J., and Muñoz, A.R., 2012, Griffon vulture mortality at wind farms in southern Spain: distribution of fatalities and active mitigation measures: *Biological Conservation*, v. 147, p. 184–189.
- Drewitt, A.L., and Langston, R.H.W., 2006, Assessing the impacts of wind farms on birds: *Ibis*, v. 148, p. 29–42.
- Gibbs, J.P., and Shriver, W.G., 2002, Estimating the effects of road mortality on turtle populations: *Conservation Biology*, v. 16, p. 1,647–1,652.
- Huso, M.M.P., 2011, An estimator of wildlife fatality from observed carcasses: *Environmetrics*, v. 22, p. 318–329.
- Mineau, P., and Whiteside, M., 2013, Pesticide acute toxicity is a better correlate of US grassland bird declines than agricultural intensification: *Plos One*, 8(2): e57457. doi:10.1371/journal.pone.0057457.

- Péron, G., Hines, J.E., Nichols, J.D., Kendall, W.L., Peters, K.A., and Mizrahi, D.S., 2013, Estimation of bird and bat mortality at wind-power farms with superpopulation models: *Journal of Applied Ecology*, 50(4), p. 902–911.
- R-Development-Core-Team, 2010, R: a language and environment for statistical computing: R Foundation for Statistical Computing, Vienna, Austria, accessed March 14, 2014, at <http://www.R-project.org>.
- Schwarz, C.J., and Arnason, A.N., 1996, A general methodology for the analysis of capture-recapture experiments in open populations: *Biometrics*, v. 52, p. 860–873.
- Skalski, J.R., 1994, Estimating wildlife populations based on incomplete area surveys: *Wildlife Society Bulletin*, v. 22, p. 192–203.
- Smallwood, K.S., 2013, Comparing bird and bat fatality-rate estimates among North American wind-energy projects: *Wildlife Society Bulletin*, v. 37, p. 19–33.
- Smallwood, K.S., Bell, D.A., Snyder, S.A., and Didonato, J.E., 2010, Novel scavenger removal trials increase wind turbine-caused avian fatality estimates: *Journal of Wildlife Management*, v. 74, p. 1,089–1,097.
- Warren-Hicks, W., Newman, J., Wolpert, R., Karas, B., and Tran, L. (California Wind Energy Association), 2013, Improving methods for estimating fatality of birds and bats at wind energy facilities: California Energy Commission, Publication Number: CEC-500-2012-086.

# Appendix 1. Example simulation study to illustrate the expected precision of fatality estimates.

## Control variables:

- Number of carcass searches: from 3 to 33
- Interval length: from 2 to 60 days
- Number of fatalities: from 5 to 100
- Detection probability: 0.5, 0.8
- Persistence probability: 0.5, 0.8
- Sample size of the trial: from 4 to 100
- Duration of the persistence trial: from 7 to 28 days

Not represented is the risk of impossible inference when the trial experiment sample size (number of trial carcasses that are planted) is small. Indeed, finding zero carcasses during the detection trial leads to an estimated detection probability of zero, leading to an infinite estimate of superpopulation size. Because the probability of finding zero trial carcasses is  $(1-p)^{N_D}$ , where  $p$  is the detection probability and  $N_D$  is the trial sample size, this problematic situation may arise more than half the time for small sample sizes. Similar problems occur with persistence probability (estimated to be 1 if all carcasses persist until the end of the trial). In this study, all simulations where the estimated coefficient of variation was greater than 1,000 or the point estimate of superpopulation size was greater than 1,000 were discarded.

## For small number of fatalities: $N = 5$

On average across simulation scenarios, the maximum value of superpopulation size under an error risk of 0.05 was 6.38 (median 6.18). This value indicates that in order to use this maximum estimated value as a valid estimate of superpopulation size, one may want to choose a higher error risk threshold—for example, 0.09 instead of 0.05. In this “evidence of absence” routine, the superpopulation size is estimated as the maximum value above which the probability of missing the carcasses would have exceeded the error risk threshold ( $N_{\max}$  is computed such that  $P_0^{N_{\max}} = 0.05$ ).  $N_{\max}$  is, therefore, not strictly speaking an estimate of superpopulation size but a conservative maximum value.  $\hat{N}$ , however, provides a more reliable estimate, not affected by the choice of risk threshold value.

**Table 1-1.** Coefficient of variation for the estimate of maximum superpopulation size ( $N_{\max} = 5$  carcasses) as a function of  $N_D$ , the sample size of the trial experiments and  $p$ , the detection probability.

$N_D$	$p$	
	0.5	0.8
4	1.56	0.46
10	0.64	0.19
20	0.32	0.12
50	0.18	0.06
100	0.12	0.04

**Table 1-2.** Linear model explaining the coefficient of variation for the estimate of maximum superpopulation size ( $N_{\max} = 5$  carcasses) as a function of  $N_D$ , the sample size of the trial experiments;  $p$ , the detection probability;  $\varphi$ , the daily persistence probability;  $I$ , the interval length in days;  $L$ , the duration of the persistence trial in days; and  $T$ , the number of sampling occasions.

[SE, standard error; P-value is from an analysis of variance (ANOVA); --, not applicable; <, less than]

Control variable	Estimate	SE	P-value
(Intercept)	1.22	0.058	--
$\log(N_D)$	-0.25	0.0070	< $10^{-6}$
$p$	-0.35	0.016	< $10^{-6}$
$\varphi$	-0.030	0.016	0.06
$\log(I)$	0.036	0.0070	< $10^{-6}$
$\log(L)$	0.0060	0.015	0.69
$\log(T)$	-0.012	0.010	0.25

**Table 1-3.** Linear model with best explanatory power explaining the coefficient of variation for the estimate of superpopulation size (5 carcasses) as a function of  $N_D$ , the sample size of the trial experiments;  $p$ , the detection probability; their interaction; and  $I$ , the interval length in days.

[SE, standard error; P-value is from an analysis of variance (ANOVA); --, not applicable; <, less than]

Control variable	Estimate	SE	P-value
(Intercept)	1.64	0.035	--
$\log(N_D)$	-0.39	0.0095	< $10^{-6}$
$p$	-1.19	0.043	< $10^{-6}$
$\log(I)$	0.036	0.0065	< $10^{-6}$
$\log(N_D) \times p$	0.27	0.013	< $10^{-6}$

A few notes and interpretation:

- Coefficient of variation (CV) decreases with higher detection probability, larger trial sample size, and shorter interval length.
- Persistence probability: It is probable that very low daily persistence would have had an effect on CV. Because the model is interval based, however, very long intervals with high persistence probability are equivalent to short intervals with low persistence probability.
- The total number of fatalities was constant and distributed over a time period proportional to the number of sampling occasions  $T$ . Therefore, it is not surprising that  $T$  had no effect. The effect of  $T$  is observed when the number of fatalities increases with  $T$ , as is expected if fatalities occur on a daily basis (that is, if  $T$  influences the sample size).
- Only trial lengths greater than 7 days were considered. For trial lengths less than 7 days, there may be an increase in CV or bias, but it would not be appropriate to plant trial carcasses and monitor them for only a very short period. Note that even with trial durations that are shorter than the expected persistence time of carcasses, CV does not increase noticeably.

### For large number of fatalities: $N \geq 10$

The average bias across simulation scenarios was -0.3 percent (as expected for an unbiased estimator).

**Table 1-4.** Coefficient of variation for the estimate of superpopulation size (10 to 100 carcasses) as a function of  $N_D$ , the sample size of the trial experiments and  $p$ , the detection probability.

$N_D$	$p$	
	0.5	0.8
4	0.86	0.54
10	0.40	0.28
20	0.20	0.16
50	0.13	0.10
100	0.11	0.08

**Table 1-5.** Linear model explaining the coefficient of variation for the estimate of superpopulation size (10 to 100 carcasses) as a function of  $N$ , the actual superpopulation size to be estimated;  $N_D$ , the sample size of the trial experiments;  $p$ , the detection probability;  $\varphi$ , the daily persistence probability;  $l$ , the interval length in days;  $L$ , the duration of the persistence trial in days; and  $T$ , the number of sampling occasions.

[SE, standard error; P-value is from an analysis of variance (ANOVA); --, not applicable; <, less than]

Control variable	Estimate	SE	P-value
(Intercept)	1.29	0.088	--
$\log(N)$	-0.14	0.012	< $10^{-6}$
$\log(N_D)$	-0.17	0.0096	< $10^{-6}$
$p$	-0.11	0.022	< $10^{-6}$
$\varphi$	0.0034	0.022	0.88
$\log(l)$	0.0028	0.0094	0.77
$\log(L)$	0.011	0.021	0.60
$\log(T)$	0.0058	0.019	0.76

**Table 1-6.** Linear model with better explanatory power explaining the coefficient of variation for the estimate of superpopulation size (10 to 100 carcasses) as a function of  $N$ , the actual superpopulation size to be estimated;  $N_D$ , the sample size of the trial experiments;  $p$ , the detection probability; their interaction;  $l$ , the interval length in days; and its interaction with  $p$ .

[SE, standard error; P-value is from an analysis of variance (ANOVA); <, less than]

Control variable	Estimate	SE	P-value
(Intercept)	1.42	0.065	< $10^{-6}$
$\log(N)$	-0.14	0.011	< $10^{-6}$
$\log(N_D)$	-0.21	0.013	< $10^{-6}$
$p$	-0.28	0.074	$1.84 \times 10^{-6}$
$\log(l)$	0.025	0.013	0.043
$\log(N_D) \times p$	0.090	0.019	$2.34 \times 10^{-6}$
$\log(l) \times p$	-0.047	0.018	0.0086

The coefficient of variation decreases with increasing superpopulation size to be estimated, increasing detection probability, increasing trial sample size, and decreasing interval length.

Prepared by the West Trenton Publishing Service Center

For more information, contact:  
Patuxent Wildlife Research Center  
U.S. Geological Survey  
Suite 4039  
12100 Beech Forest Rd.  
Laurel, MD 20708-4039

<http://www.pwrc.usgs.gov/>

