

Fatality Estimator User's Guide

Data Series 729
Version 1.2, December 2018

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

Fatality Estimator User's Guide

By Manuela Huso, U.S. Geological Survey; and Nicholas Som and Lew Ladd,
EcoStats, LLC

Data Series 729

Version 1.2, December 2018

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

U.S. Department of the Interior
KEN SALAZAR, Secretary

U.S. Geological Survey
Marcia K. McNutt, Director

U.S. Geological Survey, Reston, Virginia
First release: 2012
Revised: 2015
Revised: 2018

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 <https://www.usgs.gov> or call 1–888–ASK–USGS.

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

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

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

Although this report is in the public domain, permission must be secured from the individual copyright owners to reproduce any copyrighted materials contained within this report.

Suggested citation:

Huso, Manuela, Som, Nicholas, and Ladd, Lew, 2018, Fatality estimator user's guide (ver. 1.2, December 2018): U.S. Geological Survey Data Series 729, 22 p., <https://doi.org/10.3133/ds729>.

Contents

Abstract.....	1
Introduction.....	1
Description of Product.....	1
Installation Instructions.....	1
Getting Started	2
Main Page	3
Input Files	3
Fatality	3
Searcher Efficiency (SE).....	3
Carcass Persistence (CP).....	3
Path to R	3
Output Files	3
Results Observed and Bootstrap	3
Results SE and CP Estimates	3
Results Fatality Estimates.....	3
Parameters.....	4
Total Number of Turbines at Site	4
Total Number of Turbines in Study.....	4
Number of Bootstrap Resamples	4
Fatality Data	4
Summary List	4
Required Variables	4
Searcher-Efficiency Model.....	4
Required Variables	4
Select Variables for Model	4
Carcass-Persistence Model	5
Required Variables	5
Select Variables for Model	5
Alpha	5
Tutorial	5
Input Files	6
Fatality	6
Searcher Efficiency (SE).....	7
Carcass Persistence (CP).....	7
Path to R	8
Output Files	8
Results Observed and Bootstrap	8
Results SE and CP	8
Results Fatalities Estimates	8

Contents—Continued

Tutorial—Continued	
Parameters.....	9
Fatality Data	10
Searcher-Efficiency Model	10
Carcass-Persistence Model	11
Results	11
SE Model Selection	11
CP Model Selection	12
Final Estimates.....	13
Details	15
Data Formats—Case Sensitivity, Blanks within Columns, and Preceding Numbers	15
Missing Values	15
Input File Formats.....	15
Example Fatality Input Data	15
Example SE Trial Input Data	16
Example CP Trial Input Data	16
Output Files Contents	17
Results Observed and Bootstrap	17
Results SE and CP Estimates	17
Results Fatality Estimates.....	18
Density-Weighted Proportion Area Searched	19
Model Construction and Assumptions	19
SE Models	19
CP Models	20
Estimating Turbine-Specific Fatality	20
References Cited.....	20
Appendix A. Flow Chart of Estimation Process and Bootstrap Variance	21

Figures

1. Fatality estimator interface	2
2. Completed Input Files window	8
3. Completed Output Files window	8
4. Completed Parameters window	9
5. Completed Fatality Data window	10
6. Completed Searcher-Efficiency Model window	10
8. AICc for the SE Model based on Visibility Class	11
7. Completed Carcass-Persistence Model window	11
9. AICc results for the CP Model based on Visibility Class and a Weibull distribution	12
10. AICc results for the CP Model based on Visibility Class and a lognormal distribution	12
11. Kaplan-Meier curve (stair-step) and fitted models for four carcass-persistence distributions	13
12. Original fatality data and bootstrap sample estimates of SE, CP, r, and adjusted fatality	13
13. Estimated SE and bootstrap confidence intervals based on 2000 bootstrap iterations	14
14. Estimated CP, r, and bootstrap confidence intervals based on 2000 bootstrap iterations	14
15. Estimated fatality for all bats and for each species, and bootstrap confidence intervals based on 2000 bootstrap iterations	14
16. Example of fatality data as they appear in a spreadsheet	15
17. Example of searcher-efficiency trial data as they appear in a spreadsheet	16
18. Example of carcass-persistence trial data as they appear in a spreadsheet	16
19. Example bootstrap results output	17
20. Example SE results output	17
22. Example Fatality results output	18
21. Example CP results output	18

Tables

1. First few rows of fictitious example of Fatality data in Site1.Fatality.csv	6
2. First few rows of fictitious example of SE Trial data in Site1.Trial SE.csv	7
3. First few rows of fictitious example of CP Trial Data in Site1.Trial CP.csv	7
4. AICc results for SE Models with and without Visibility Class	11
5. AICc results for CP Models using lognormal distribution with and without Visibility Class	12

Fatality Estimator User's Guide

By Manuela Huso, U.S. Geological Survey; and Nicholas Som and Lew Ladd, EcoStats, LLC

Abstract

This publication is the User's Guide for software developed to estimate wildlife fatalities at wind-power facilities, although the software is applicable to a variety of circumstances in which the objective is to estimate the size of a superpopulation and the probability of detection of the individuals is less than one. Simple counts of carcasses do not accurately reflect fatality and do not allow comparison among locations because carcasses may be detected at different rates. This software uses data collected during carcass searches and knowledge of detection rates to accurately estimate the number of fatalities and to provide a measure of precision associated with the estimate. These estimates are fundamental to understanding acute and cumulative effects of wind power on wildlife populations.

Only carcasses judged to have been killed after the previous search should be included in the fatality data set submitted to this estimator software. This estimator already corrects for carcasses missed in previous searches, so carcasses judged to have been missed at least once should be considered “incidental” and not included in the fatality data set used to estimate fatality. **Note:** When observed carcass count is <5 (including 0 for species known to be at risk, but not observed), USGS Data Series 881 (<http://pubs.usgs.gov/ds/0881/>) is recommended for fatality estimation.

Introduction

Description of Product

This software is developed for use by anyone interested in estimating the size of a superpopulation of a group of animals based on observations of individuals whose probability of detection is less than one. A superpopulation is defined as the total number of animals occurring in a study area during a fixed period of time. The population is not closed, and the probability of detection can vary among individuals based either on the physical characteristics of the individual, such as size, or on the environmental conditions in which the individual exists, such as vegetation. The software was developed to estimate the total abundance of carcasses at wind-power facilities from the number of observed carcasses. Typically at these sites, an area beneath

each turbine is searched by humans, sometimes with the aid of dogs, to discover carcasses of animals that may have been killed by the rotating turbine blades. The simple count of observed carcasses does not accurately represent the actual population of animals killed by turbines because it is practical to search only a fraction of any given site, because scavengers sometimes remove carcasses before the search is conducted, and because observers simply overlook carcasses due to obstructing vegetation, topography, and simple bad luck. More importantly, the simple count is not linearly related to the population, so rarely can it be used as an index of mortality. The Fatality Estimator software uses data provided by the user to estimate the probability that a searcher will miss a carcass and the probability that a carcass will be removed before a searcher has the opportunity to observe it. These estimates are combined with the observed casualties and an estimate of the fraction of the population of killed animals expected to be in the search area to form an estimate of total abundance of carcasses over a specified period of time. Much of the conceptual framework for this approach can be found in Huso (2011).

Installation Instructions

- Make sure R software is installed on your machine and you know its location.
- R is open-source statistical software available at <http://cran.r-project.org/>. The Fatality Estimator software is Windows based, so select R for Windows.
- Make sure the survival package in R is also installed on your machine.
- Download the archive file “ds729_EstimatorSetup.zip” to your computer.
- Unzip the contents to a folder on your system. In this zip file, you will find the following:
 - Estimator Project User's Guide.pdf
 - EstimatorSetup.msi
 - Site1.Fatality.csv
 - Site1.Trial.CP.csv
 - Site1.Trial.SE.csv
 - Site1.Search Schedule.csv

2 Fatality Estimator User's Guide

The comma-delimited (.csv) files are example files from a fictitious site. They show expected file formats and present the data used in the tutorial included in this user's guide.

- Double click on EstimatorSetup.msi and follow instructions for installing the software. The InputPlaceholder.txt file that will be installed with the Fatality Estimator contains the R code for the program, and most updates will occur through updates to this file.
- You can now safely delete the zip file.

Getting Started

Start the Fatality Estimator by double-clicking on the shortcut created during installation. The first time you run the program all input windows of the main screen will be empty ([fig. 1](#)). After you run the program once, the software remembers the most recent input. On the main page are three blocks: **Input Files**, **Output Files**, and **Parameters**. Each block is described below. A schematic of the general process used by this software is given in [appendix A](#).

The screenshot shows the 'EstimatorWindow' interface. It is divided into three main sections: 'Input Files', 'Output Files', and 'Parameters'.
- **Input Files:** Contains four text input fields with 'Select...' buttons: 'Fatality:', 'Searcher Efficiency (SE):', 'Carcass Persistence (CP):', and 'Path to R:'.
- **Output Files:** Contains three text input fields with 'Select...' buttons: 'Results Observed and Bootstraps:', 'Results SE and CP Estimates:', and 'Results Fatality Estimates:'.
- **Parameters:** Contains several input fields and buttons: 'Total Number of Turbines at Site', 'Total Number of Turbines in Study', 'Number of Bootstrap Resamples', 'Fatality Data' (with button 'Specify Fatality Data Input and Summary Groups'), 'Searcher Efficiency Model' (with button 'Specify Searcher Efficiency Model'), 'Carcass Persistence Model' (with button 'Specify Carcass Persistence Model'), and 'Alpha' (with value '0.05').
At the bottom, there are three buttons: 'GO', 'View R Input', and 'Review Last Output'.

Figure 1. Fatality estimator interface.

Main Page

All input data should be in comma-delimited file format (.csv). The main screen is divided into three sections: **Input Files, Output Files, and Parameters**. Three separate input data files in comma-delimited format (.csv) are required. Three separate output files in comma-delimited format (.csv) are produced. The user will have the opportunity to provide information about which columns in the trial data represent responses and explanatory variables, if applicable.

Input Files

Fatality

The user must browse to and select the appropriate file in her/his system containing the records of carcasses found during searches and any covariates that might affect detection probability. The file must contain at least four columns: a turbine identifier, the date the carcass was discovered, the date of the previous search at that turbine, and the proportion of carcasses expected to fall within the searched area of the turbine (see sections “[Parameters](#)” and “[Details](#)” for more information). Only carcasses judged to have been killed after the previous search should be included in the fatality data set submitted to this estimator software. This estimator already corrects for carcasses missed in previous searches, so carcasses judged to have been missed at least once should be considered “incidental” and not included in the fatality data set used to estimate fatality.

Searcher Efficiency (SE)

The user must browse to and select the appropriate file in her/his system containing the records of carcasses used during searcher-efficiency (SE) trials. The file must contain at least one column—an indicator (numeric or character) indicating whether the trial carcass was found by searchers or not (see sections “[Parameters](#)” and “[Details](#)” for more information).

Carcass Persistence (CP)

The user must browse to and select the appropriate file in her/his system containing the records of carcasses used during carcass-persistence (CP) trials. The file must contain at least three columns—one indicating the left censor time, one indicating the right censor time, and one representing the censoring type (see sections “[Parameters](#)” and “[Details](#)” for more information).

Path to R

Because all systems are uniquely configured, the user is asked to navigate to the location of the local R installation on the system, that is, the location of R.exe, which typically is in C:\Program Files\R.

Output Files

The Output Files screen is used to navigate to the locations and assign names to the three output files where (1) bootstrap samples, (2) SE and CP estimates, and (3) fatality estimates will be written. All three output files will be comma-delimited (.csv) files. The files cannot be written if they are open in other applications.

Results Observed and Bootstrap

The user must browse to a location in her/his system and name a file in which s/he wants to save the fatality data plus all bootstrap samples of fatality data and bootstrap estimates of SE (\hat{p}), average CP time (\hat{t}), and estimated proportion of carcasses persisting through the interval (\hat{r}) for each carcass.

Results SE and CP Estimates

The user must browse to a location in her/his system and name a file in which she/he wants to save the SE and CP estimates. This file will contain a summary of the original SE trial data and estimated SE (\hat{p}) with $(1-\alpha)$ bootstrap confidence intervals, based on a user-defined model of \hat{SE} ; a summary of original CP trial data, estimated CP time (\hat{t}) and proportion of carcasses persisting through the interval (\hat{r}) with $(1-\alpha)$ bootstrap confidence intervals based on the user-defined distribution and model of CP time.

Results Fatality Estimates

The user must browse to a location in her/his system and name a file in which she/he wants to save the per-turbine and total-site-fatality estimates with $(1-\alpha)$ bootstrap confidence intervals, for each group selected by the user.

Parameters

The Parameters screen is used to input site and study information and to allow the user to define models of SE and CP time.

Total Number of Turbines at Site

The user enters the total number of wind turbines at the location of the study—that is, the total number of turbines out of which the sample turbines were selected and to which the user wants to draw inference. The ratio of this number to the number of turbines sampled in the study will be used to expand estimated per-turbine fatality to all turbines at the study site.

Total Number of Turbines in Study

The user enters the number of turbines at the site where searches were conducted for carcasses. This number must be less than or equal to the total number of turbines at the site. If this number is smaller than the total number of turbines located at the study site, these turbines will be assumed to be a representative sample of all turbines. It is possible that no carcasses were found at some of the searched turbines. The software checks to assure that the number of unique turbine codes in the fatality data does not exceed the number of turbines searched in the study because such a situation would indicate that fatalities are included at turbines that are not part of the searched set.

Number of Bootstrap Resamples

The user enters the number of bootstrap iterations desired. Reported results should be computed using no fewer than 1,000 bootstrap resamples. We recommend using 5,000 resamples. The amount of time the program takes to complete is greatly affected by this number. It is common to choose fewer than 1,000 resamples, for example 1, to conduct a practice-run of the Fatality Estimator to ensure that all data are formatted correctly. This trial run also allows the user to determine if the output contains requested information. An additional use of a trial run is at the model-selection phase when the user only wants information regarding AICc for a particular model. If a number smaller than 1,000 is placed in this field for the number of bootstrap iterations, the output will contain a warning that the results should not be considered valid.

Fatality Data

Summary List

A list of variable names in the Input Fatality Data is available from the pull-down menu in this field. The user selects variable(s) from this menu that contain subgroup identifiers for each carcass and for which summary estimates of fatality are desired. The selected variables must be categorical, not continuous, with no missing values. To remove variables to this list, the user highlights the variable in the list on the right, and then selects the “Remove Selection” button.

Required Variables

Turbine ID: The user selects the column representing a unique turbine identifier.

Density-Weighted Proportion Area Searched: The user selects the column representing the proportion of carcasses (greater than 0 and less than or equal to 1) expected to fall within the searched area of the turbine (see section “[Details](#)” for more information).

Date Carcass Found: The user selects the column representing the date on which the carcass was found, in mm/dd/yyyy format.

Date of Previous Search: The user selects the column representing the date on which the previous search of that turbine occurred, in mm/dd/yyyy format. An error message is generated if the date of the previous search is greater than or equal to the date the carcass was found.

Searcher-Efficiency Model

Required Variables

Response Column: The user selects the column (numeric or character) representing an indicator of whether the trial carcass was found by searchers or not. The larger value will be assumed to indicate that the carcass was found.

Select Variables for Model

A list of variable names in the Input Searcher-Efficiency Trial Data is presented. The user must formulate the model, following syntax used by R. All models are additive models, that is, additive in the coefficients. All effects are separated by “+”, with interactions of two variables indicated by “:”.

The first main effect is included by double-clicking on a variable name. If further covariates are to be included, the “+” operator must be selected by double-clicking. Then further covariates can be selected, separated by “+”. Interactions between variables that have previously appeared as main effects are indicated by “.”. The user selects the variable(s) she/he wants to include in the model, separated by the selected operator. Consistent with R syntax, interactions between variables are represented by the semicolon operator. The selected variables must be categorical, not continuous. The Model-Building Help button provides brief instructions for building models.

Carcass-Persistence Model

Required Variables

Censoring Type: The user selects the (numeric) column representing the censoring type of each observation. Possible values are 0, 1, 2, or 3 (see sections “[Parameters](#)” and “[Details](#)” for more information).

Left: The user selects the (numeric) column representing the left censor time in days or fractions of days.

Right: The user selects the (numeric) column representing the right censor time in days or fractions of days.

Failure Time Distribution: The user selects the desired failure time distribution from the pull-down menu. Options are exponential, Weibull, loglogistic, and lognormal (Bispo and others, 2012).

Select Variables for Model

A list of variable names in the Input Carcass-Persistence Trial Data is presented. The user selects the variable(s) she/he wants to include in the model, separated by the selected operator. Consistent with R syntax, interactions between variables are represented by the semicolon operator. The selected variables will be treated as categorical not continuous. The Model-Building Help button provides brief instructions for building models.

Alpha

The user enters a number greater than 0 and less than 1, which is used to define the width of the confidence intervals computed by the estimator. For example, if 90-percent confidence intervals are desired, then $\alpha = 1 - 0.90 = 0.10$.

Tutorial

In this tutorial, we walk you through an analysis of the fictitious example data provided for Site 1. At this site, 15 of 35 turbines at the site were randomly selected to be searched weekly for bats and birds from July 28 through September 26, 2008. A scheduled search of Turbine 11 on August 18 was not completed due to turbine maintenance by the owner (see Site1. Search Schedule.csv). Clean-out searches were conducted the week of July 21–25 to remove any existing carcasses. Three turbines were searched per day by the same searcher over a 5-day work week. The data provided reflect only the bat data. At least one bat carcass was found at each turbine, except at Turbine 22 where none was found. The proportion of bat carcasses expected to fall within the searchable area beneath each turbine (Hull and Muir, 2010) as estimated and coded in a column named DWP (Density-Weighted Proportion). In this study, the habitat beneath each turbine was mapped into three visibility classes: Easy, Moderate and Difficult. It was hypothesized that both average searcher-efficiency and carcass-persistence time would differ significantly among these three classes. The objectives of the analysis are:

1. Determine whether Visibility Class significantly affects searcher efficiency, that is, average SE differs by more than 10 percent between classes. We will compare models incorporating Visibility Class to those with no covariate using AICc, and also compare estimated SE.
2. Determine the best carcass-persistence distribution and whether visibility class significantly affects average carcass-persistence time, and consequently the proportion of carcasses expected to persist through a search interval (one week). Visibility class will be considered significant if expected proportions differ by more than 10 percent between classes. We will compare models incorporating Visibility Class to those with no covariates using AICc, and also compare estimated proportion expected to persist through a search interval (r).
3. Estimate bat fatality (and 95-percent confidence interval) at this site from July 28 through September 26, 2008.

Input Files

Fatality

Browse to and select the file containing Fatality Data *Site1.Fatality.csv*. Each row of this dataset corresponds to a single observed bat carcass. As this was a study to estimate bat fatality and all records are bats, there is no column distinguishing birds and bats ([table 1](#)). Only carcasses judged to have been killed after the previous search were included in this fatality data set. This estimator already corrects for carcasses missed in previous searches, so carcasses judged to have been missed at least once were considered “incidental” and not included in these data.

Table 1. First few rows of fictitious example of Fatality data in Site1.Fatality.csv.

Turbine	DWP	Date	DateLast	Visibility	Species	Freq	Age	Sex	Condition
2	0.61	08-11-2008	08-04-2008	Easy	LACI	Low	U	U	LateDecomp
2	0.61	08-18-2008	08-11-2008	Difficult	LACI	Low	U	U	Desiccated
2	0.61	08-25-2008	08-18-2008	Easy	LANO	Low	A	M	Fresh
2	0.61	09-15-2008	09-08-2008	Easy	LABO	High	A	M	Fresh
4	0.44	08-12-2008	08-05-2008	Easy	UNKN	UNKN	A	F	Fresh
4	0.44	08-26-2008	08-19-2008	Difficult	LANO	Low	A	F	EarlyDecomp

Column	Description	Possible values
Turbine	Unique identifier for each turbine.	2, 4, 7, 8, 10, 11, 12, 15, 17, 19, 22, 23, 25, 28, 31
DWP	The distance-weighted proportion of bat carcasses expected to fall within the searchable area beneath the turbine.	0.01–1.00
Date	Date the carcass was found.	07-28-2008–09-26-2008
DateLast	Date of the previous search of this turbine.	07-21-2008–09-19-2008
Visibility	Visibility class representing the vegetation density in which the carcass was found.	Easy, Moderate, Difficult
Species	4 character code for the bat species. UNKN represents unknown.	EPFU, LABO, LACI, LANO, MYLU, PESU, UNKN
Freq	Call frequency group for this species of bat. UNKN represents unknown.	High, Low, UNKN
Age	Single letter code representing reproductive status of the bat. Possible values are Adult, Juvenile, Unknown.	A, J, U
Sex	Single letter code representing sex of the bat. Possible values are Female, Male, Unknown.	F, M, U
Condition	Condition of the carcass when found. Possible values are Fresh, EarlyDecomp[osition], LateDecomp[osition], Desiccated.	Fresh, EarlyDecomp, LateDecomp, Desiccated

Searcher Efficiency (SE)

Browse to and select file containing Searcher-efficiency (SE) Trial Data *Site1.Trial.SE.csv*. Each row of this dataset corresponds to a trial carcass used to estimate searcher efficiency ([table 2](#)).

Table 2. First few rows of fictitious example of SE Trial data in Site1.Trial SE.csv.

Turbine	Placed	Visibility	Species	Found
22	07-28-2008	Difficult	LABO	0
12	07-29-2008	Easy	LACI	0
15	07-30-2008	Difficult	LACI	0
25	07-30-2008	Easy	LABO	1
8	07-31-2008	Moderate	LABO	1
10	08-01-2008	Moderate	LABO	1

Column	Description	Possible values
Turbine	Unique identifier for each turbine.	2, 4, 7, 8, 10, 11, 12, 15, 17, 19, 22, 23, 25, 28, 31
Placed	Date the carcass was placed in the trial.	7-28-2008–9-26-2008
Visibility	Visibility class representing the vegetation density in which the carcass was found.	Easy, Moderate, Difficult
Species	4 character code for the bat species. UNKN represents unknown.	EPFU, LABO, LACI, LANO, MYLU, PESU, UNKN
Found	Indicator of whether the searcher found the trial carcass (1) or not (0).	0, 1

Carcass Persistence (CP)

Browse to and select file containing Carcass-persistence (CP) Trial Data *Site1.Trial.CP.csv*. Each row of this dataset corresponds to trial carcass used to estimate carcass-persistence models ([table 3](#)).

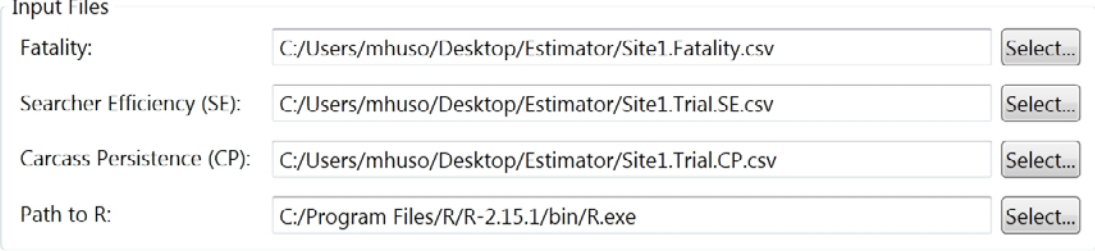
Table 3. First few rows of fictitious example of CP Trial Data in Site1.Trial CP.csv.

Turbine	Placed	Visibility	Species	Left	Right	Event
23	07-29-2008	Moderate	LACI	0.5	0.5	2
10	08-01-2008	Moderate	LABO	19.5	19.5	0
22	07-28-2008	Difficult	LABO	0.4	1.5	3
12	07-29-2008	Easy	LACI	0.6	1.6	3
15	07-30-2008	Difficult	LACI	0.5	1.6	3
11	08-04-2008	Moderate	LABO	18.6	18.6	0
25	07-30-2008	Easy	LABO	0.6	1.4	3

Column	Description	Possible values
Species	Four-character code for the bat species. UNKN represents unknown.	
Turbine	Turbine at which the carcass was placed.	
Visibility	Visibility class representing the vegetation density in which the carcass was placed.	Easy, Moderate, and Difficult
Placed	Date the carcass was placed in the trial.	
Left	Time (in days) when carcass was last known to be present.	
Right	Time (in days) when carcass was first known to be absent or undetectable.	
Event	Censoring code. This is the response in the survival models used to estimate carcass-persistence time (see section “ Details ” for code definitions).	

Path to R

Browse to the location of the local R.exe installation on the system. The Input Files window should now be filled in ([fig. 2](#)).



The screenshot shows the 'Input Files' window with four rows of input fields and 'Select...' buttons. The paths entered are as follows:

Input Field	Path
Fatality:	C:/Users/mhuso/Desktop/Estimator/Site1.Fatality.csv
Searcher Efficiency (SE):	C:/Users/mhuso/Desktop/Estimator/Site1.Trial.SE.csv
Carcass Persistence (CP):	C:/Users/mhuso/Desktop/Estimator/Site1.Trial.CP.csv
Path to R:	C:/Program Files/R/R-2.15.1/bin/R.exe

Figure 2. Completed Input Files window.

Output Files

Results Observed and Bootstrap

Browse to the folder in which you want to save the original data and the bootstrap iteration results and name it S1.Boot.

Results SE and CP

Browse to the folder in which you want to save the estimated values for SE and average carcass-persistence time and confidence intervals and name it S1.SECP.

Results Fatalities Estimates

Browse to the folder in which you want to save the estimated fatality and confidence intervals for the groups selected in the Parameters window (below) and name it S1.Fat. The Output Files window should now be filled in ([fig. 3](#)).



The screenshot shows the 'Output Files' window with three rows of input fields and 'Select...' buttons. The paths entered are as follows:

Input Field	Path
Results Observed and Bootstraps:	C:/Users/mhuso/Desktop/Estimator/Site1/S1.Boot.csv
Results SE and CP Estimates:	C:/Users/mhuso/Desktop/Estimator/Site1/S1.SECP.csv
Results Fatality Estimates:	C:/Users/mhuso/Desktop/Estimator/Site1/S1.Fat.csv

Figure 3. Completed Output Files window.

Parameters

- Total Number of Turbines at Site
Enter 35, the number of turbines at Site 1.
- Total Number of Turbines in the Study
Enter 15, the number of turbines randomly selected to be searched. This is critical to being able to accurately calculate per-turbine fatality when no carcasses were found at some turbines, and consequently, those turbines do not occur in the fatality data.
- Number of Bootstrap Resamples
Enter 1 to execute a test run to calculate AICc for model comparisons before running a full 1000 or more iterations.
- Alpha
Enter a number between 0 and 1, to request (1-alpha) percent confidence limits. The Parameters window should now be filled in ([fig. 4](#)).

Parameters	
Total Number of Turbines at Site	35
Total Number of Turbines in Study	15
Number of Bootstrap Resamples	1
Fatality Data	Specify Fatality Data Input and Summary Groups
Searcher Efficiency Model	Specify Searcher Efficiency Model
Carcass Persistence Model	Specify Carcass Persistence Model
Alpha	0.05

Figure 4. Completed Parameters window.

Fatality Data

From the Main Menu Parameter screen, select:

Specify Fatality Data Input and Summary Groups

and select relevant columns.

- **Summary Variables**
From the pull-down menu select Species as the variable indicating the groups for which we would like fatality to be estimated.
- **Required Variables**
 - Turbine ID: Select Turbine from the pull-down menu.
 - Density Weighted Proportion Searched: Select DWP from the pull-down menu.
 - Date Carcass Found: Select Date from the pull-down menu.
 - Date of Previous Search: Select DateLast from the pull-down menu.

The Fatality Data window should now be filled in ([fig. 5](#)).

Searcher-Efficiency Model

From the Main Menu Parameter screen, select

Specify Searcher Efficiency Model

and select relevant columns.

- **Required Variables**
 - Response Column: Select Found from the pull-down menu.
- **Select Variable(s) for Model:** Double-click on Visibility. Visibility will appear in the list of Variables in Model.

Click on Close. The SE Model window should now be filled in ([fig. 6](#)).

Figure 5. Completed Fatality Data window.

Figure 6. Completed Searcher-Efficiency Model window.

Carcass-Persistence Model

From the Main Menu Parameter screen, select:

Specify Carcass Persistence Model

and select relevant columns.

- Required Variables
 - Censoring Type: Select Event from the pull-down menu.
 - Left: Select Left from the pull-down menu.
 - Right: Select Right from the pull-down menu.
 - Failure Time Distribution: Select Weibull from the pull-down menu.
 - Select Variable(s) for Model: Double-click on Visibility. Visibility will appear in the list of Variables in Model.

Click on Close. The Carcass-Persistence Model window should now be filled in ([fig. 7](#)).

Figure 7. Completed Carcass-Persistence Model window.

Results

All input data are complete, and the program is executed by clicking on the GO button. A window with the estimates of SE, CP, and Fatality is automatically opened once all bootstrap iterations are complete. Because we are initially using the software to select the best models of SE and CP, we calculate only a single bootstrap iteration, and consequently, the reported confidence intervals in this exercise should be ignored.

SE Model Selection

Look at the results for the SE Model selection ([fig. 8](#)). AICc for the SE Model based on Visibility Class is 65.83. Make note of this, then edit the SE Model to remove Visibility Class from the model by clicking on *Clear*. Close the SE Model window and rerun by clicking on *GO*.

The AICc for the SE Model with no explanatory variables is 76.27 ([table 4](#)). The difference in AICc is greater than 10 units, and estimated SE differs among Visibility Classes by more than 10 percent; therefore, we will decide to model SE as a function of Visibility Class.

Table 4. AICc results for SE Models with and without Visibility Class.

AICc results for SE Models	
Explanatory variables	
Visibility Class	None
65.83	76.27

Results: SE

"User-defined alpha level of 0.05; all reported confidence intervals are 95% confidence intervals."
"

"Searcher Efficiency Estimates; AICc for model = Visibility: 65.83"

SF_Cov_Ind	Visibility	Found	Placed	SE	Lwr	Upr
Difficult	Difficult	4	18	0.22	0.28	0.28
Easy	Easy	15	18	0.83	0.83	0.83
Moderate	Moderate	11	18	0.61	0.61	0.61

Figure 8. AICc for the SE Model based on Visibility Class.

CP Model Selection

Look at the results for the CP Model selection ([fig. 9](#)). AICc for the CP Model based on Visibility Class and assuming a Weibull distribution is 224.16 ([fig. 9](#)). Make note of this, and then edit the CP Model to remove Visibility Class from the model by clicking on *Clear*. Close the CP Model window and rerun by clicking on *GO*. AICc for the CP Model with no explanatory variables and assuming a Weibull distribution is 220.59.

Results: CP

"Carcass Persistence Estimates; AICc for model = Visibility and distribution = weibull: 224.16"

"Estimates and Confidence Intervals for r based on an interval of 7 days."

CP_Cov_Ind	Visibility	Placed	CP	Lwr	Upr	r	r_Lwr	r_Upr
Difficult	Difficult	20	13.2	11.79	11.79	0.59	0.57	0.57
Easy	Easy	18	17.66	44.46	44.46	0.62	0.72	0.72
Moderate	Moderate	20	35.15	11.18	11.18	0.7	0.57	0.57

Figure 9. AICc results for the CP Model based on Visibility Class and a Weibull distribution.

Repeat the process selecting each of the three other distributions and models with and without Visibility Class. The results ([table 5](#)) indicate that the best model is the lognormal distribution with no covariates, although the loglogistic model is a close second and differs from the lognormal model by less than two units.

The model with minimum AICc was the lognormal with no covariates. Estimated proportion of carcasses persisting through a search interval of 7 days (r) did not differ by more than 10 percent among the Visibility classes ([fig. 10](#)). Therefore, we will decide to model CP as lognormal with no explanatory variables. The difference in estimated proportion of carcasses persisting among three of the distributions is minimal ([fig. 11](#)). Only the exponential distribution results in a significantly different persistence model that fits the empirical distribution poorly, and this outcome is reflected in its high AICc value.

Table 5. AICc results for CP Models using lognormal distribution with and without Visibility Class.

AICc results for carcass-persistence models		
Distribution	Explanatory variables	
	Visibility Class	None
Weibull	224.16	220.59
exponential	262.07	259.42
loglogistic	220.41	216.52
lognormal	219.24	215.26

Results: CP

"Carcass Persistence Estimates; AICc for model = Visibility and distribution = lognormal: 219.24"

"Estimates and Confidence Intervals for r based on an interval of 7 days."

CP_Cov_Ind	Visibility	Placed	CP	Lwr	Upr	r	r_Lwr	r_Upr
Difficult	Difficult	20	4.92	5.01	5.01	0.58	0.58	0.58
Easy	Easy	18	7.44	3.23	3.23	0.63	0.53	0.53
Moderate	Moderate	20	10.25	4.07	4.07	0.67	0.56	0.56

Figure 10. AICc results for the CP Model based on Visibility Class and a lognormal distribution.

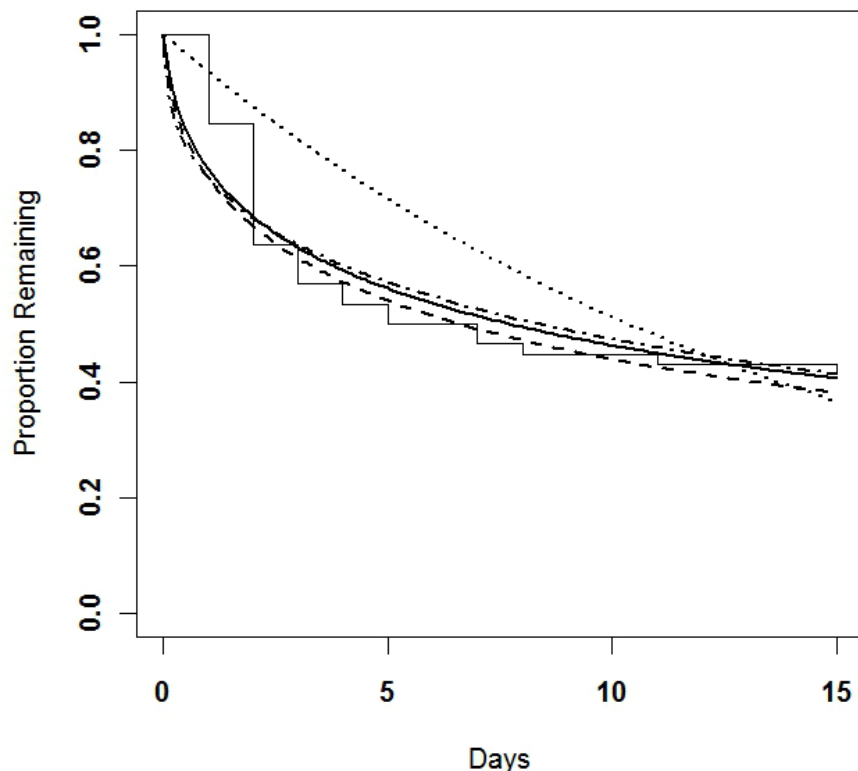


Figure 11. Kaplan-Meier curve (stair-step) and fitted models for four carcass-persistence distributions: lognormal (—), exponential (····), Weibull (—•—•—), and loglogistic (— — —).

Final Estimates

Return to the main menu, change the bootstrap iterations to 2000 (or even 5000), and click GO to estimate parameters and bootstrap confidence intervals. It will take several minutes to run through a large number of bootstrap iterations. A window will automatically open showing the three output files designated earlier: the original fatality data with the full set of bootstrap iterations appended ([fig. 12](#)); the estimates of average SE ([fig. 13](#)), average CP time and r , the proportion of carcasses expected to persist through the interval based on selected models ([fig. 14](#)); estimated Fatality for each requested group ([fig. 15](#)). The presented confidence intervals represent $(1-\alpha)$ percent empirical confidence intervals, but because the confidence intervals are bootstrapped, the ones you calculate may not be exactly the same as those presented here.

Results: Observed and BootStraps

Turbine	DWP	Date	DateLast	Visibility	Species	Freq	Age	Sex	Condition	SE_Cov_Ind	CP_Cov_Ind	SE_pred	CP_pred	rhat	AdjFat	boot_ID
2	0.61	8/18/2008	8/11/2008	Difficult	LACI	Low	U	U	Desiccated	Difficult	1	0.22	7.18	0.63	11.83	0
4	0.44	8/26/2008	8/19/2008	Difficult	LABO	Low	A	F	EarlyDecomp	Difficult	1	0.22	7.18	0.63	16.4	0
15	0.87	9/3/2008	8/27/2008	Difficult	LABO	High	A	F	Fresh	Difficult	1	0.22	7.18	0.63	8.29	0
23	0.37	8/5/2008	7/29/2008	Difficult	LACI	Low	J	M	Fresh	Difficult	1	0.22	7.18	0.63	19.5	0
19	0.89	9/26/2008	9/19/2008	Difficult	LABO	High	A	F	Fresh	Difficult	1	0.22	7.18	0.63	8.11	0
31	0.91	8/15/2008	8/8/2008	Difficult	LACI	Low	U	U	EarlyDecomp	Difficult	1	0.22	7.18	0.63	7.93	0
7	0.86	8/27/2008	8/20/2008	Difficult	LABO	High	U	U	LateDecomp	Difficult	1	0.22	7.18	0.63	8.39	0
28	0.73	9/11/2008	9/4/2008	Difficult	LABO	Low	A	M	EarlyDecomp	Difficult	1	0.22	7.18	0.63	9.88	0
28	0.73	8/28/2008	8/21/2008	Difficult	LABO	Low	A	F	EarlyDecomp	Difficult	1	0.22	7.18	0.63	9.88	0
7	0.86	9/3/2008	8/27/2008	Difficult	LABO	Low	A	F	Fresh	Difficult	1	0.22	7.18	0.63	8.39	0
7	0.86	7/30/2008	7/23/2008	Difficult	LACI	Low	A	M	Fresh	Difficult	1	0.22	7.18	0.63	8.39	0

Figure 12. Original fatality data and bootstrap sample estimates of SE, CP, r , and adjusted fatality.

14 Fatality Estimator User's Guide

Seven columns are appended to the original data. SE_Cov_Ind and CP_Cov_Ind are columns indicating the level of the combination of covariates selected in the model-selection process. Because our selected CP Model had no explanatory variables, this column for CP is simply filled with the number 1. For SE, the values in this column reflect the values in the Visibility variable selected as an explanatory variable for the SE Model. SE_pred and CP_pred are the estimated SE and average CP time for this bootstrap sample. rhat is the estimated proportion of carcasses persisting through the interval (calculated as the difference between DateLast and Date). AdjFat is the adjusted fatality estimate for the carcass, and boot_ID is an indicator of the bootstrap sample, with 0 representing the original data.

Results: SE

"User-defined alpha level of 0.05; all reported confidence intervals are 95% confidence intervals."

" "

"Searcher Efficiency Estimates; AICc for model = Visibility: 65.83"

SE_Cov_Ind	Visibility	Found	Placed	SE	Lwr	Upr	
Difficult	Difficult	4	18	0.22	0.06	0.44	
Easy	Easy	15	18	0.83	0.67	1	
Moderate	Moderate	11	18	0.61	0.39	0.83	

Figure 13. Estimated SE and bootstrap confidence intervals based on 2000 bootstrap iterations.

Results: CP

"Carcass Persistence Estimates; AICc for mean-only model and distribution = lognormal: 215.26"

"Estimates and Confidence Intervals for r based on an interval of 7 days."

CP_Cov_Ind	Placed	CP	Lwr	Upr	r	r_Lwr	r_Upr	
mean	58	7.18	3.22	20.67	0.63	0.53	0.75	

Figure 14. Estimated CP, r, and bootstrap confidence intervals based on 2000 bootstrap iterations.

Results: Fatalities Estimate

"This study included 15 turbines, and these estimates have been expanded to reference the 35 at the study site."

"User-defined alpha level of 0.05; all reported confidence intervals are 95% confidence intervals."

" "

"Requested Summaries"

Group	Level	Number_Found	PerTurb_Est	PerTurb_Lwr	PerTurb_Upr	SiteTot_Est	SiteTot_Lwr	SiteTot_Upr	
Overall		91	26.43	17.71	50.25	926	620	1759	
Species	LABO	46	13.61	8.86	26.66	477	311	934	
Species	LACI	45	12.82	8.13	25.04	449	285	877	

Figure 15. Estimated fatality for all bats and for each species, and bootstrap confidence intervals based on 2000 bootstrap iterations. Fatalities are estimated per turbine and extrapolated to the entire study area.

Details

Data Formats—Case Sensitivity, Blanks within Columns, and Preceding Numbers

- All input data must be in comma-delimited file format (.csv).
- The maximum number of columns is 16,384. The maximum number of rows is 1,048,573.
- All column names are case-sensitive and must begin with a letter, not a number.
- Do not name the first column in the dataset 'ID' because this may cause it to be recognized as a corrupted format.
- Values within any column should be either numeric or character, but no character value should begin with a number.
- Values within any column cannot include spaces or punctuation marks and are case-sensitive. That means values with different cases are considered different values, for example, Bat is not the same as bat.

Missing Values

Missing values are not allowed for any of the required variables in any of the three input files nor for any of the potential covariates used in CP or SE models. While missing values could conceptually be acceptable in covariate columns for CP or SE data, their presence will result in slightly different datasets being compared in AICc model selection, thereby nullifying AICc results.

Input File Formats

Example Fatality Input Data

Each row of this dataset corresponds to a single observed fatality.

Turbine	DWP	Date	DateLast	Visibility	Species	Freq	Age	Sex	Condition
2	0.61	8/11/2008	8/4/2008	Easy	LACI	Low	U	U	LateDecomp
2	0.61	8/18/2008	8/11/2008	Difficult	LACI	Low	U	U	Desiccated
...									

Figure 16. Example of fatality data as they appear in a spreadsheet.

See [Data Formats](#) section for formatting details.

Example SE Trial Input Data

Each row of this dataset corresponds to a single trial carcass used to estimate SE.

Turbine	Placed	Visibility	Species	Found
22	7/28/2008	Difficult	LABO	0
12	7/29/2008	Easy	LACI	0
...				

Figure 17. Example of searcher-efficiency trial data as they appear in a spreadsheet.

See section “[Data Formats](#)” for formatting details. Any column used to model SE must also occur in the Fatality data, and all values must be consistently coded in both files.

Searcher-efficiency trials are generally conducted in this way. A study manager places carcasses in conditions [time of year, landscape characteristics, and carcass characteristics (size)] that will be used to model the probability that a searcher observes a carcass. Next, a searcher conducts a regularly scheduled search of the plot, unaware that a trial carcass has been placed there. If the carcass is found by the searcher, the trial is noted as a success (Found=1 in the example data). If the carcass is not found, the manager checks to see that the placed carcass was not scavenged and was potentially observable by the searcher. If so, the trial is noted as unsuccessful (Found=0 in the example data). If the carcass is missing, the trial should not be included in the SE trial data.

Example CP Trial Input Data

Each row of this dataset corresponds to a single trial carcass used to estimate CP.

Turbine	Placed	Visibility	Species	Left	Right	Event
23	7/29/2008	Moderate	LACI	0.5	0.5	2
10	8/1/2008	Moderate	LABO	19.5	19.5	0
...						

Figure 18. Example of carcass-persistence trial data as they appear in a spreadsheet.

See section “[Data Formats](#)” for formatting details. Any column used to model CP must also occur in the Fatality data, and all values must be consistently coded in both files. Carcass-persistence trials are generally conducted this way. Throughout the relevant study period, carcasses are placed in conditions [time of year, landscape characteristics, and carcass characteristics (size)] that will be used to model the average amount of time a carcass is expected to persist. Next, the carcass is visited at pre-determined times (not necessarily at equally spaced intervals) to ascertain its continued presence and observability. For instance, carcasses could be checked 1, 2, 3, 4, 5, 6, 7, 10, 14, 21, 28, and 35 days after placement. This dataset does not need to contain observations for each of the visits. It does need to contain columns for each of the explanatory variables, and three additional columns used for the statistical models. The columns in the example data are called Left, Right, and Event. Left and Right are measures of time that is always recorded as number of days or fractions of days since the carcass was placed in the study. Below are directions for constructing these columns:

If the carcass was already gone on the first visit: Both the Left and Right values are set to the time of the first check. The Event value would be the number 2, reflecting left censoring.

If the carcass was still present when the study ends: Both the Left and Right values are set to the time of the last check. The Event value would be the number 0 (zero) reflecting right censoring.

If the carcass disappears between checks, the most common case: The Left value is set to the time of the last check when the carcass was known to be present. The Right value is set to the time when the carcass was first known to be no longer observable. The Event value would be the number 3, reflecting interval censoring. For example, if the carcass was present during the check conducted on the 7th day after the carcass was placed but was gone when the carcass was rechecked on the 10th day, the Left value would be 7, the Right value would be 10, and the Event value would be 3. Similarly, if a particular carcass was still present on the 10th day, but that carcass was not monitored on the 14th and 21st days, and was then found to be gone on the 28th day, the Left value would be 10, the Right value would be 28, and the Event value would be 3.

If removal of the carcass was actually witnessed, so time of removal is known exactly: The Left and Right values are set to the time of removal, and the Event value would be the number 1 reflecting known Event time and no censoring.

Output Files Contents

Results Observed and Bootstrap

This output file will contain seven columns appended to the original data:

SE.Cov.Ind	CP.Cov.Ind	SE.pred	CP.pred	rhat	AdjFat	boot.ID
Difficult	1	0.22	7.18	0.63	11.83	0
Difficult	1	0.22	7.18	0.63	16.4	0
Difficult	1	0.22	7.18	0.63	8.29	0
...						

Figure 19. Example bootstrap results output.

SE.Cov.Ind and CP.Cov.Ind: Level(s) of the combinations of covariates identified in the SE and CP model selection processes, respectively. If there are no covariates in the SE or CP models, the column is simply a column filled with the number

SE.pred and CP.pred: Estimated SE and average CP time for the current bootstrap sample.

rhat: Estimated proportion of carcasses persisting through the interval, calculated from columns in the input Fatality data.

AdjFat: Adjusted fatality estimate based on Huso (2011) for each carcass.

boot.ID: Indicator of the bootstrap iteration, with 0 representing the original data.

Results SE and CP Estimates

This output file will contain separate sections for SE and CP.

The SE results will consist of a header describing the model and AICc results followed by six or more columns:

SE.Cov.Ind	Visibility	Found	Placed	SE	Lwr	Upr
Difficult	Difficult	4	18	0.22	0.06	0.44
Easy	Easy	15	18	0.83	0.61	1
Moderate	Moderate	11	18	0.61	0.39	0.83

Figure 20. Example SE results output.

SE.Cov.Ind: Level(s) of the combinations of covariates identified in the model selection process. The column value will be simply “mean” if there are no covariates in the SE Model.

Optional Column(s): These can exist with information indicating the level of each covariate in the SE Model individually. If there are no covariates in the SE Model, these columns will not occur.

Found: Number of trial carcasses in each covariate combination that was found.

Placed: Total number of trial carcasses in each covariate combination that were placed in the trial.

SE: Estimated SE for each covariate combination.

Lwr: Lower bootstrap (1-alpha) percent confidence interval for SE.

Upr: Upper bootstrap (1-alpha) percent confidence interval for SE.

The CP results will consist of a header describing the model and AICc results followed by eight or more columns:

CP.Cov.Ind	Placed	CP	Lwr	Upr	r	r.Lwr	r.Upr
mean	58	7.18	3.17	20.4	0.63	0.52	0.75

Figure 21. Example CP results output.

CP.Cov.Ind: Level(s) of the combinations of covariates identified in the model selection process. The column value will be simply “mean” if there are no covariates in the CP Model.

Optional column(s): Level of each covariate in the CP Model individually. If there are no covariates in the CP Model, the CP columns will not occur.

Placed: Total number of trial carcasses in each covariate combination that were placed in the trial.

CP: Estimated CP for each covariate combination.

Lwr: Lower bootstrap (1-alpha) percent confidence interval for CP.

Upr: Upper bootstrap (1-alpha) percent confidence interval for CP.

r: Estimated proportion of carcasses persisting through the interval for each covariate combination. The interval represented is the interval most commonly occurring in the Fatality data, that is, the mode.

r.Lwr: Lower bootstrap (1-alpha) percent confidence interval for CP.

r.Upr: Upper bootstrap (1-alpha) percent confidence interval for CP.

Results Fatality Estimates

This output file will consist of nine columns, and will have one section for each variable selected as a summary variable:

Group	Level	Number Found	PerTurb. Est	PerTurb. Lwr	PerTurb. Upr	SiteTot. Est	SiteTot. Lwr	SiteTot. Upr
Overall		91	26.43	17.82	51.04	926	624	1787
Species	LABO	46	13.61	8.97	26.98	477	314	945
Species	LACI	45	12.82	8	25.67	449	281	899

Figure 22. Example Fatality results output.

Group: Variable identified in the Fatality Data Summary List (see below) that contains subgroup identifiers for each carcass and for which summary estimates of fatality are desired.

Level: Subgroups of the Group variable.

Number Found: Summary of the total number of carcasses in each subgroup that were found during searches.

PerTurb.Est: Estimated number of carcasses of each subgroup killed per turbine

PerTurb.Lwr: Lower bootstrap (1-alpha) percent confidence interval for PerTurb.Est

PerTurb.Upr: Upper bootstrap (1-alpha) percent confidence interval for PerTurb.Est

SiteTot.Est: Estimated total number of carcasses of each subgroup killed at the site

SiteTot.Lwr: Lower bootstrap (1-alpha) percent confidence interval for SiteTot.Est

SiteTot.Upr: Upper bootstrap (1-alpha) percent confidence interval for SiteTot.Est

Density-Weighted Proportion Area Searched

Even if we search every turbine at a site, our search areas vary in size from one site to another and may not include all areas in which a carcass might fall. In addition, a turbine's entire search area may not be searchable, for example, landscape characteristics might prevent a full-area search. Because carcasses are more likely to land close to a turbine than far from a turbine, the configuration and size of the searched area will affect the proportion of carcasses expected to land within it. The fatality data should contain a column indicating what proportion of the carcass density was estimated to be within searchable areas beneath each turbine (Hull and Muir, 2012). This is not the same as the fraction of the designated search plot that was searchable. The numbers in this column should all be greater than 0 and less than or equal to 1 (0.75 would be correct, but 75 percent would **NOT** be). This factor is used to expand estimates to the entire search area. If this factor is unknown, the default is 1, but selecting the default may lead to underestimated fatality.

Model Construction and Assumptions

All variables selected as model covariates will be interpreted as factors. Continuous covariates cannot yet be incorporated in SE or CP models. A minimum of 10 trial carcasses per defined factor combination is required to assure adequate samples to calculate bootstrap sampling distributions.

SE Models

The Searcher-Efficiency Trial input file must contain at least one column indicating whether each trial carcass was found or not. The variable selected by the user as the response is assumed to be distributed as a Bernoulli (binary) random variable, with found indicated by the highest numeric value (or highest alphabetic value if the response variable is character-valued).

NOTE: Not all data in any modeled combination of covariates for SE can be 0 or not found. In addition, the number of levels in any covariate used to model SE must be equal to or greater than the number of observed levels in the fatality data. For example, if size is a meaningful covariate of SE, then all levels of size occurring in the fatality data also must be represented in the SE trial data.

A generalized linear model of the binary response with a logit link is fit to the data. The modeled response is $\log_e(\text{odds of observing a carcass}) = \log_e(p/(1-p))$. If covariates are selected for the SE Model, they are converted to factors, and indicator variables are created by the software. An indicator variable is created for each level of a factor except the first, which is considered the reference level. For example, a covariate called season with four possible values ("Fall", "Spring", "Summer", "Winter") will be converted to three indicator variables, S1-S3 as follows:

$S_1 = 1$ if Season="Spring", 0 otherwise

$S_2 = 1$ if Season="Summer", 0 otherwise

$S_3 = 1$ if Season="Winter", 0 otherwise

"Fall" is the default reference level

A covariate called Size with two possible values ("L", "S") will be converted as follows:

$Z_1 = 1$ if Size="Small", 0 otherwise.

"Large" is the default reference level.

Examples of three different models involving season with and without size:

Model: SEASON

$$\log_e(\text{odds}) = \beta_0 + \beta_1 * S_1 + \beta_2 * S_2 + \beta_3 * S_3$$

Model: SEASON + SIZE

$$\log_e(\text{odds}) = \beta_0 + \beta_1 * S_1 + \beta_2 * S_2 + \beta_3 * S_3 + \beta_4 * Z_1$$

Model: SEASON + SIZE + SEASON : SIZE

$$\log_e(\text{odds}) = \beta_0 + \beta_1 * S_1 + \beta_2 * S_2 + \beta_3 * S_3 + \beta_4 * Z_1 + \beta_5 * S_1 * Z_1 + \beta_6 * S_2 * Z_1 + \beta_7 * S_3 * Z_1$$

In the models above, the estimate of β_0 represents the estimated $\log_e(\text{odds of observing a reference-level carcass})$, where the reference level is a carcass in the Fall in the first model and a Large carcass in the Fall in the second two models. β_1 represents how much the $\log_e(\text{odds of observing a reference-level carcass})$ are estimated to change if the search is conducted in the Spring.

NOTE: None of the parameter estimates is actually visible to the user. This information is provided to help the user understand how to formulate the model using correct syntax and how the difference between the second and third models is represented.

With some datasets, a bootstrap sample of SE trial carcasses may result in all observations being unsuccessful, that is, 0, which would lead to a runtime error when the estimator would attempt to divide by 0. In this case, SE is set to 0.05, representing a 1 in 20 chance of observing a carcass under the specified conditions in order to prevent division

by 0. This may result in slight underestimation of the upper confidence limits of fatality but will not affect the estimate or the lower confidence limit. Simply dropping samples that result in zero detection for any stratum would potentially affect both the upper and lower limits. The user is cautioned against using trial data for which SE is extremely low.

CP Models

The Carcass-Persistence Trial input file must contain at least three columns: one indicating the left censor time, that is, the time in days since being placed in the study when carcass was last known to be present; one indicating the right censor time, that is, the time in days since being placed in the study when carcass was first known to be absent or the study reached its end; and one representing the censoring type. These are used to estimate carcass-persistence time, and must be numeric, with only one of three possible values:

- 0=right censored, that is, the carcass was still present when the study ended. Both right and left censor times must have the same value and it is the number of days from when the carcass was first placed to when it was last observed at the end of the study
- 1=no censoring, that is, time of removal is known exactly (this will almost never be the case).
- 2=left censored, that is, the carcass was removed sometime between when it was placed and the first check. Again, both right and left censor times must have the same value and it is the number of days from when the carcass was first placed to the time of the first check when it was observed to have been removed.
- 3=interval censored, that is, the carcass was removed sometime between the previous check when it was last known to be present and the current check when it was first known to be absent. The value in the left censor time must be less than the value in the right censor time.

NOTE: Not all data in any modeled combination of covariates can be right censored.

In addition, the number of levels in any covariate used to model CP must be equal to or greater than the number of observed levels in the fatality data. For example, if size is a meaningful covariate of CP, then all levels of size occurring in the fatality data also must be represented in the CP trial data.

Carcass-persistence times are modeled via log-linear parametric accelerated failure time models. If we define T as a persistence-time random variable, then a general form of our model is $\log_e(T) = \beta_0 + \sigma\epsilon$, where β_0 is the mean, σ is a scale parameter, and ϵ is assumed to follow a probability distribution (Collett, 2003). When persistence time is not known exactly or is censored, the model is appropriately modified (see Collette, 2003 for details beyond the scope of this user's guide). Decisions about the distribution of ϵ dictate the distribution of T , and we offer four commonly used accelerated failure-time model options: exponential, Weibull, log-logistic, and lognormal. If covariates are selected for the CP Model, they are converted to factors and indicator variables are created by the software as described above in SE Models.

NOTE: None of the parameter estimates is actually visible to the user.

Estimating Turbine-Specific Fatality

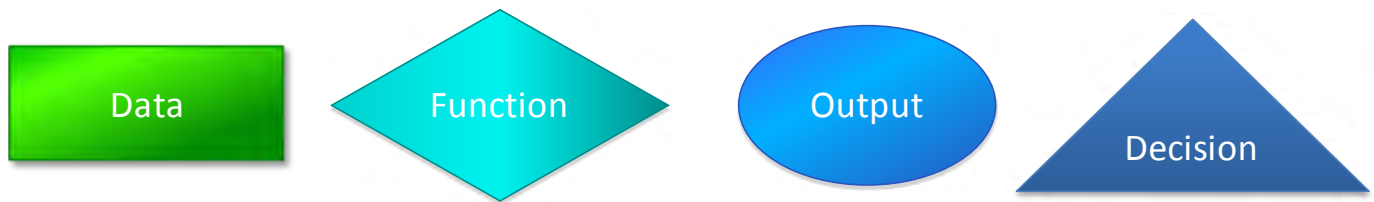
This software is designed to estimate per-turbine and total-site fatality, but is not designed to estimate fatality at individual turbines. However, the user can summarize adjusted fatality estimates provided in the Bootstrap Results file by turbine to generate these estimates if they are of interest and if there are adequate observations at each turbine to support such estimates.

References Cited

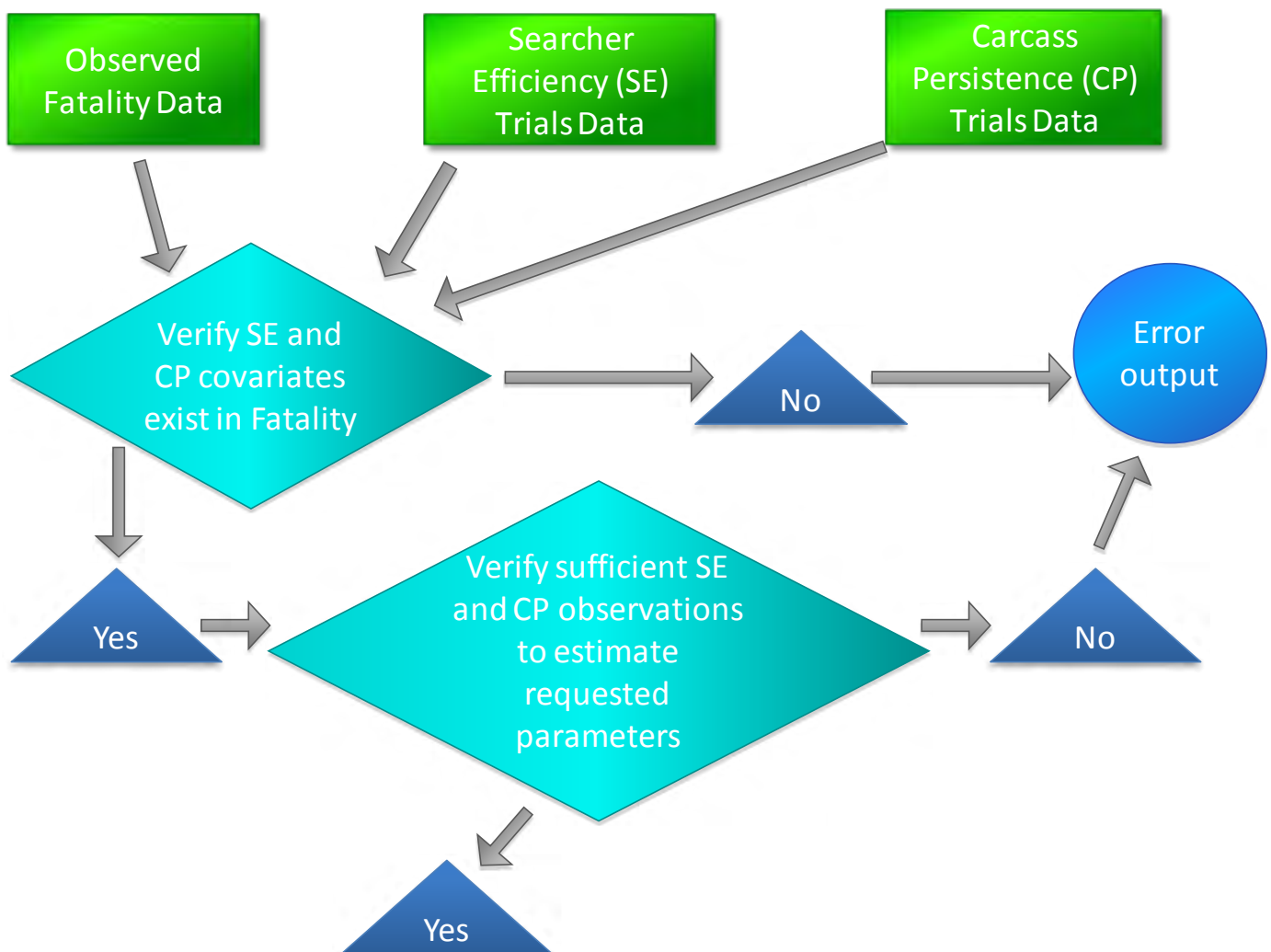
- Collett, D., 2003, *Modelling survival data in medical research*: Boca Raton, Florida, Chapman and Hall, 408 p.
- Bispo, R., Bernardino, J., Marques, T., and Pestana, D., 2012, *Modeling carcass removal time for avian mortality assessment in wind farms using survival analysis*: Springer, Environmental and Ecological Statistics, doi:10.1007/s10651-012-0212-5.
- Hull, C., Muir, S., 2010, Search areas for monitoring bird and bat carcasses at wind farms using a Monte-Carlo model: *Australian Journal of Environmental Management*, v. 17, p. 77–87.
- Huso, M.M.P., 2011, An estimator of wildlife fatality from observed carcasses: *Environmetrics*, v. 22, p. 318–329.

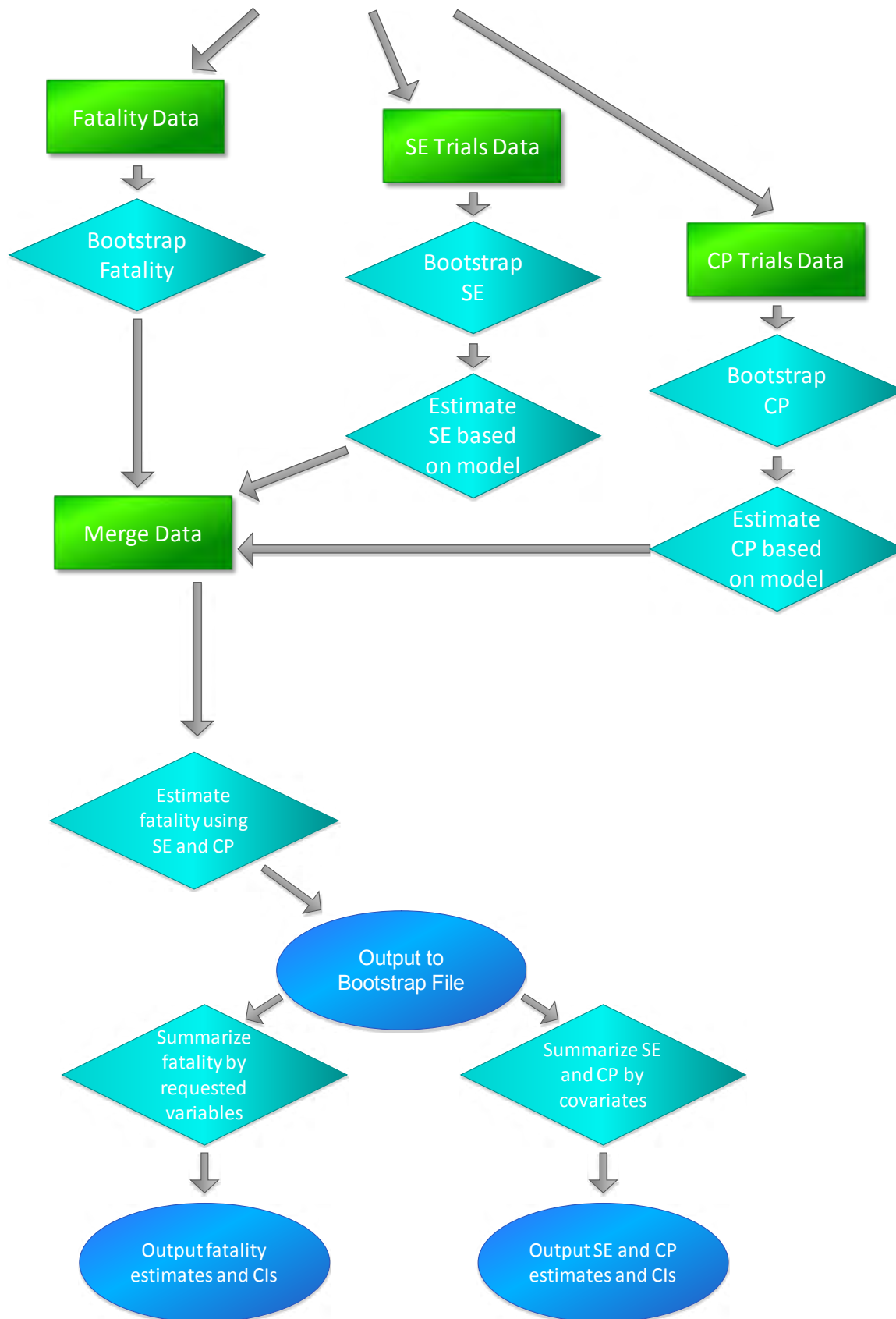
Appendix A. Flow Chart of Estimation Process and Bootstrap Variance

Key to flowchart symbols:



From three input files:



Appendix A.—Continued

Publishing support provided by the U.S. Geological Survey
Publishing Network, Tacoma Publishing Service Center

For more information concerning the research in this report, contact the
Director, Forest and Rangeland Ecosystem Science Center,
U.S. Geological Survey, 777 NW 9th Street
Corvallis, Oregon 97330
<http://fresc.usgs.gov>

