# GenEst (v1.0.0) Software User Guide

By Juniper Simonis, Daniel Dalthorp, Manuela Huso, Lisa Madsen, Paul Rabie, and Jared Studyvin

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

Open File Report XXX

U.S. Department of the Interior

U.S. Geological Survey

U.S. Department of the Interior

RYAN K. ZINKE, Secretary

**U.S. Geological Survey**

James Reilly, Director

U.S. Geological Survey, Reston, Virginia: 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 (1–888–275–8747).

For an overview of USGS information products, including maps, imagery, and publications,  
visit [https:/store.usgs.gov](http://www.usgs.gov/pubprod).

Disclaimer: The findings and conclusions in this article are those of the author(s) and do not necessarily represent the views of the U.S. Fish and Wildlife Service or the U.S. Bureau of Land Management.

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: Simonis, Juniper, Dalthorp, Daniel, Huso, Manuela, Madsen, Lisa, and Rabie, Paul, 2018, GenEst (v1.0.0) software user guide: U.S. Geological Open File Report XXX, xx p., [https://doi.org/10.3133/](https://doi.org/10.3133/x)XXX.

ISSN 2327-638X (online)

Contents

GenEst (v1.0.0) Software User Guide 1

Abstract 6

1.0 Introduction 6

1.1 Description of Product 6

1.1.1 Software Modules 7

1.1.2 Comparison to Other Methods 7

1.2 Installation Instructions 7

1.3 Getting Started 8

1.4 Layout 8

2.0 Data Input 9

2.1 Data Files 9

3.0 Analyses 11

3.1 General Inputs 11

3.1 Searcher Efficiency 12

3.1.1 Model Inputs 12

3.1.2 Outputs 13

3.1.2.1 Data 13

3.1.2.1 Model Table 13

3.1.2.1 Figure 13

3.1.2.1 Model Selection 13

3.2 Carcass Persistence 14

3.3 Fatality Esitmation 14

3.4 Detection Probability 14

5.1 Subsection 14

5.1.1 Example subsubsection 14

4.0 Worked Examples 15

4.1 Example 1 15

4.1.1 Example subsubsection 15

4.2 Example 2 15

4.2.1 Example subsubsection 15

5.0 Conclusions 15

Option: An unnumbered but linked header 16

Acknowledgments 17

About the Authors and Collaborators 17

References Cited 17

Appendix A. Input File Formats 19

A.1 Search Efficiency 19

A.2 Carcass Persistence 19

A.3 Search Schedule 19

A.4 Carcass Observations 19

Appendix B. Analytical Details 19

B.1 Search Efficiency 19

B.2 Carcass Persistence 19

B.3 Search Schedule 19

B.4 Carcass Observations 19

Figures

**Figure 1.** Welcome screen disclaimer for GenEst, as opened in Mozilla Firefox (v58.0.2). 5

**Figure 2.** Browsing to the example data sets within the GenEst R package folder on a Windows (10.0.16299) machine. 6

**Figure 3.** Visible uploaded data in the GenEst GUI. 6

**Figure 4.** Required general inputs. 7

**Figure 2.** Screen captures of the EoA control panel and the minimized R console. 11

Tables

**Table 1.** Files needed for each analysis 10

**Table 1.** Example format of carcass persistence data. 16

Abbreviations

|  |  |
| --- | --- |
| AICc | Akaike information criterion corrected for finite sample size |
| ΔAIC | difference between AICc value for a given model and the lowest AICc value among models tested |
| .csv | comma-separated values (format for some input files) |
| CDF | cumulative distribution function |
| CCDF | complement of the cumulative distribution function or 1 – CDF |
| CI | confidence interval or credible interval (depending on context) |
| CP | carcass persistence |
| CPmax | the earliest time a persistence trial carcass is noted to be missing |
| CPmin | the minimum amount of time that a persistence trial carcass is known to have persisted |
| CRAN | Comprehensive R Archive Network (for downloading R statistical software and supplemental packages) |
| EoA | Evidence of Absence software |
| IQR | interquartile range; bounds of the central 50% of a distribution; the 0.25 and 0.75 quantiles of a distribution |
| LRT | likelihood ratio test |
| PDF | probability density function |
| PI | prediction interval |
| .rds | serialized R data object (format for some input files) |
| SE | searcher efficiency |

Variables and Parameters

|  |  |
| --- | --- |
|  | significance level (distinguished from shape parameter by context) |
|  | credibility level, confidence level (depending on context) |
|  | shape and scale parameters for persistence distributions |
|  | spatial coverage, fraction of carcasses falling into searched area |
|  | density weighted proportion; fraction of the total carcasses that arrive in a searched area |
|  | sampling fraction; proportion of potential sampling units that were searched |
|  | overall detection probability within the searched area, within the period of study |
|  | search interval; number of days between searches |
|  | assumed search interval for reporting an estimate of |
|  | factor by which searcher efficiency decreases with each successive search |
|  | mortality rate over a set time period |
|  | mortality or number of fatalities |
|  | estimated mortality or number of fatalities |
|  | searcher efficiency, probability of observing a carcass that is present in the searched area at the time of the search |
|  | probability that a carcass persists until the first search after arrival |
| span | the length of the monitoring period |
|  | temporal coverage, fraction of total carcasses that arrive during the monitored period |
|  | number of carcasses observed in searches |

GenEst (v1.0.0) Software User Guide

by Juniper Simonis[[1]](#footnote-1), Daniel Dalthorp[[2]](#footnote-2), Manuela Huso2, Lisa Madsen[[3]](#footnote-3), Paul Rabie[[4]](#footnote-4), and Jared Studyvin4,

# Abstract

GenEst (Generalized Estimator) is a software tool for estimating the total number of individuals arriving in an area during a specific time period when their detection probability is unknown but estimable. In particular, its development was motivated by the need to accurately estimate the number of bird and bat fatalities occurring at renewable-energy facilities. The software package is available with a user-friendly graphic interface as well as a flexible and powerful command-line implementation. GenEst includes tools for estimating searcher efficiency, carcass persistence, and other detection probability parameters from experimental field trials. Included in the software are example datasets for analyses, standard R package help files, this user guide, and vignettes detailing use at the command-line.

# 1.0 Introduction

## 1.1 Description of Product

GenEst is a software package designed to addresses the general problem of estimating numbers of fatalities over an extended period of time using systematic counts of carcasses and adjustments of the carcass counts to account for imperfect detection. Imperfect detection may be due to any of several possible detection biases, for example: (1) search teams fail to find carcasses that are present in the searched area at the time of the search, (2) scavengers remove carcasses before searches are conducted, (3) carcasses fall outside the searched area, or (4) fatalities occur outside the monitored period. The detection rate () is estimated primarily from results of field trials in which carcasses are placed at known locations within the searched areas at the site and monitored for persistence times and for evaluating the efficiency of search teams in detecting carcasses that are not scavenged. Combining the number of carcasses () found in the systematic carcass searches with information about the detection rate, GenEst estimates the total mortality () and quantifies the uncertainty associated with the estimation.

GenEst (Generalized Estimator) is a software package designed for use by anyone analyzing data associated with estimating bird or bat fatalities at renewable-energy facilities, such as wind and solar farms, but has applicability in other situations, as well. GenEst is written in R (R Core Team 2017) and available as a fully documented R package (<package submission information>) including a command-line interface and as a user-friendly Graphic User Interface (GUI) coded using shiny (Chang et al. 2017). This User Guide focuses on the GUI version of GenEst. Vignettes within the R package detail the command-line interface.

In tracking fatalities, carcasses are usually collected during distinct searches, and searching is repeated through time. Because some carcasses are removed (via, e.g., scavengers) before searches and some carcasses are not seen by searchers even when they are present, not all carcasses that fall within a search area are found. Investigators therefore typically conduct trial experiments with placed carcasses to estimate the impact of efficiency and persistence.

-density weighted proportion

-fraction of units searched

Modules allow for estimation of searcher efficiency (Searcher Efficiency), carcass persistence (Carcass Persistence), and detection probability (Detection Probability) parameters as well as for the overall fatality estimation (Fatality Estimation).

### 1.1.1 Software Modules

GenEst software includes two primary modules, each with submodules:

1. Data input—For pointing the software to the folder containing the data input files (Searcher Efficiency, Carcass Persistence, Search Schedule, Density Weighted Proportion, and Observed Fatalities), all in .csv format.

2. Analyses—For determining the best Searcher Efficiency and Carcass Persistence models and combining input data from several search classes and/or time periods into an estimate of total mortality.

### 1.1.2 Comparison to Other Methods

Text here describing when and why to use GenEst vs other methods. Reference to Paul’s table.

## 1.2 Installation Instructions

GenEst is available as an R package. Thus, you will need to download and install R (https://cran.r-project.org/) before using GenEst.

GenEst is available for download from the USGS GitHub repository. Within R, ensure that you have the devtools package installed and loaded into R by running

install.packages("devtools")

library(devtools)

Then download and install the GenEst package

devtools::install\_github("usgs/GenEst")

And load it into the R instance

library(GenEst)

GenEst will eventually be available from the R package repository (<http://cran.r-project.org>), at which point, the stable version of the software can be downloaded and installed locally by running

install.packages("GenEst")

And then loaded into the R instance via

library(GenEst)

Installing the GenEst package will also lead to all dependency R packages being installed and loaded into the R instance. In addition, GenEst comes with example data sets, vignettes, and help files, all located within the package folder and accessible via both the command-line and GUI routes.

## 1.3 Getting Started

Initializing the GenEst GUI is as simple as running

runGenEst()

which will launch the application in your default web browser (fig. 1). The welcome screen for the GUI has a disclaimer that must be acknowledged before proceeding. Note that opening the app ties up the active R session. As a result, if you would like to use an active R session alongside the GUI, you will need to initiate an additional R session.

## 1.4 Layout

The GUI is laid out using tabs, as shown in the top left corner of fig. 1: Home, Data Input, Analyses, and About. Once the user acknowledges the disclaimer, they are able to navigate between the tabs. The Home tab is a welcome page and the About tab includes additional information about the GUI and package. The Data Input and Analyses tabs allow active user interaction for analyses.

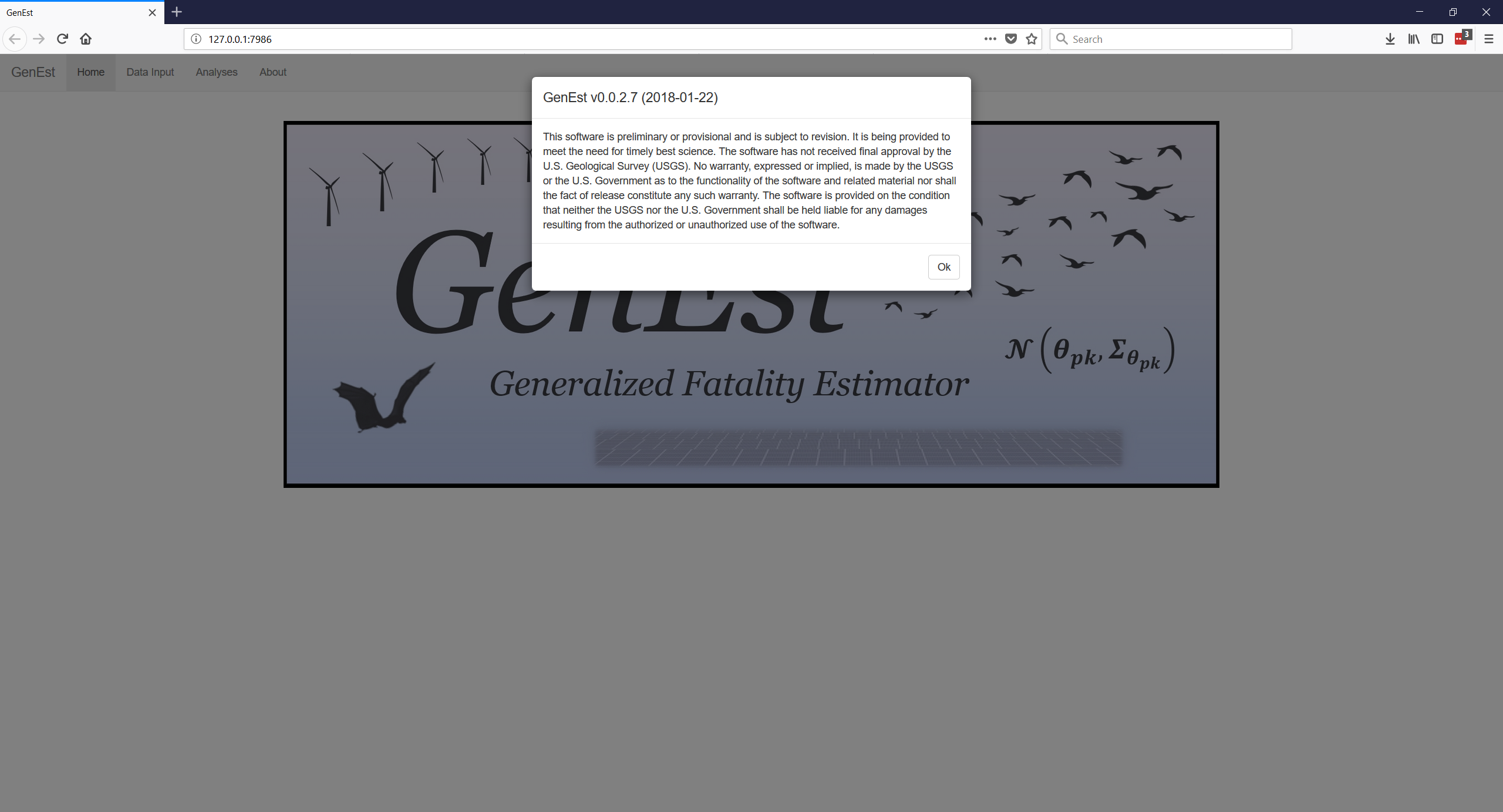


Figure 1. Welcome screen disclaimer for GenEst, as opened in Mozilla Firefox (v58.0.2).

# 2.0 Data Input

## 2.1 Data Files

Having acknowledged the disclaimer, the GUI is now able to accept data for use (fig. 2). Example data files are provided in the R package that can be used within the GUI (in the extdata subfolder in the GenEst package folder, which is located with other R package folders on the user’s computer). Clicking on the “Browse…” button allows the user to locate the files for upload.

Currently, the GUI can take one of each of four data files: Searcher Efficiency, Carcass Persistence, Search Schedule, and Carcass Observations. The files should be formatted with comma separation (either a comma separated value (.csv) or a comma delimited regular text (.txt) file). Appendix A details the necessary and available formats for each of the files. The files required to upload depend on the user’s specific desires (table 1). Upon upload, each data file becomes available for inspection (fig. 3) as well as use throughout the GUI.

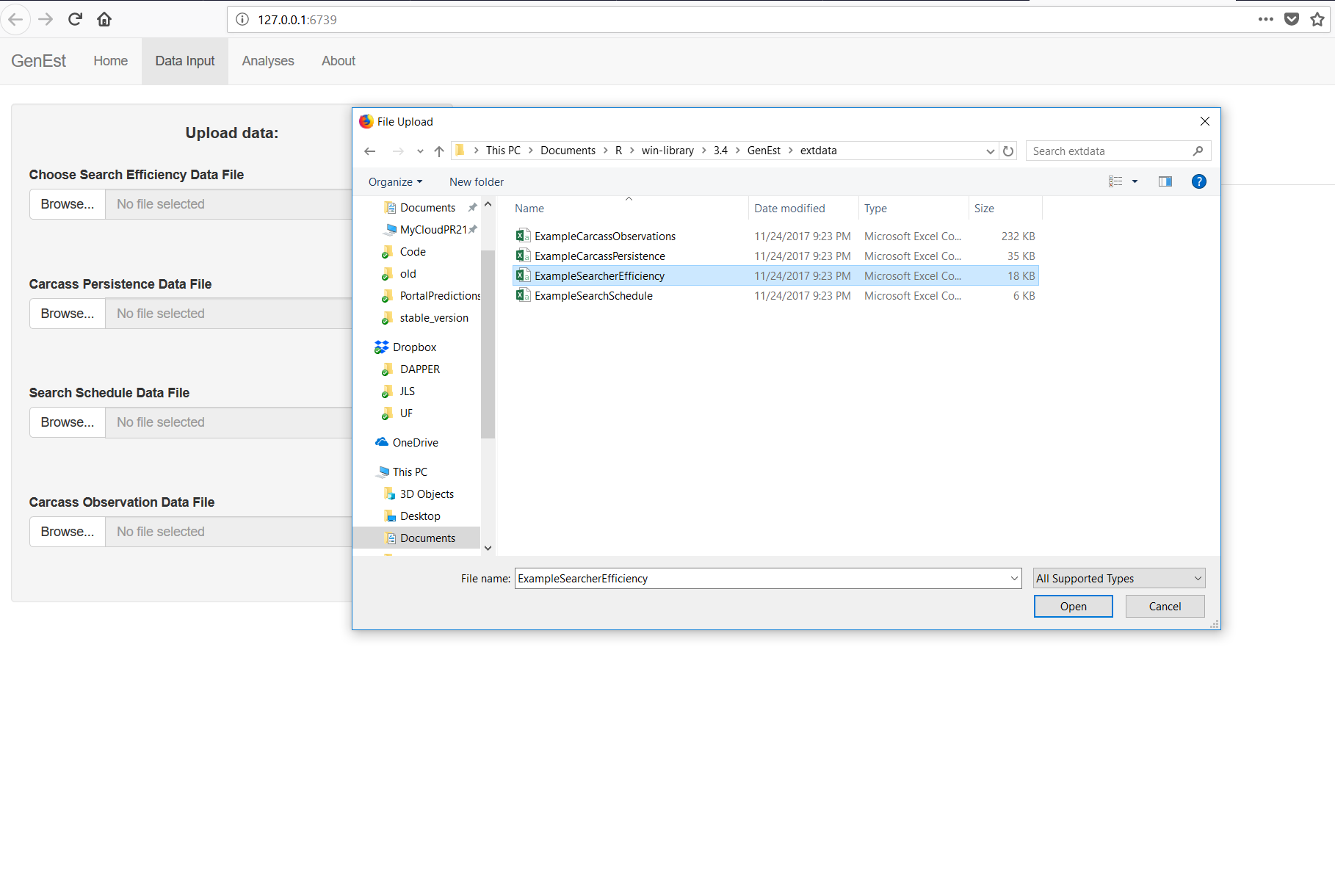


Figure 2. Browsing to the example data sets within the GenEst R package folder on a Windows (10.0.16299) machine.

Table 1. Files needed for each analysis

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Analysis | Search Efficiency | Carcass Persistence | Search Schedule | Carcass Observations |
| Search Efficiency | x |  |  |  |
| Carcass Persistence |  | x |  |  |
| Detection Probability | x | x | x |  |
| Fatality Estimation | x | x | x | x |

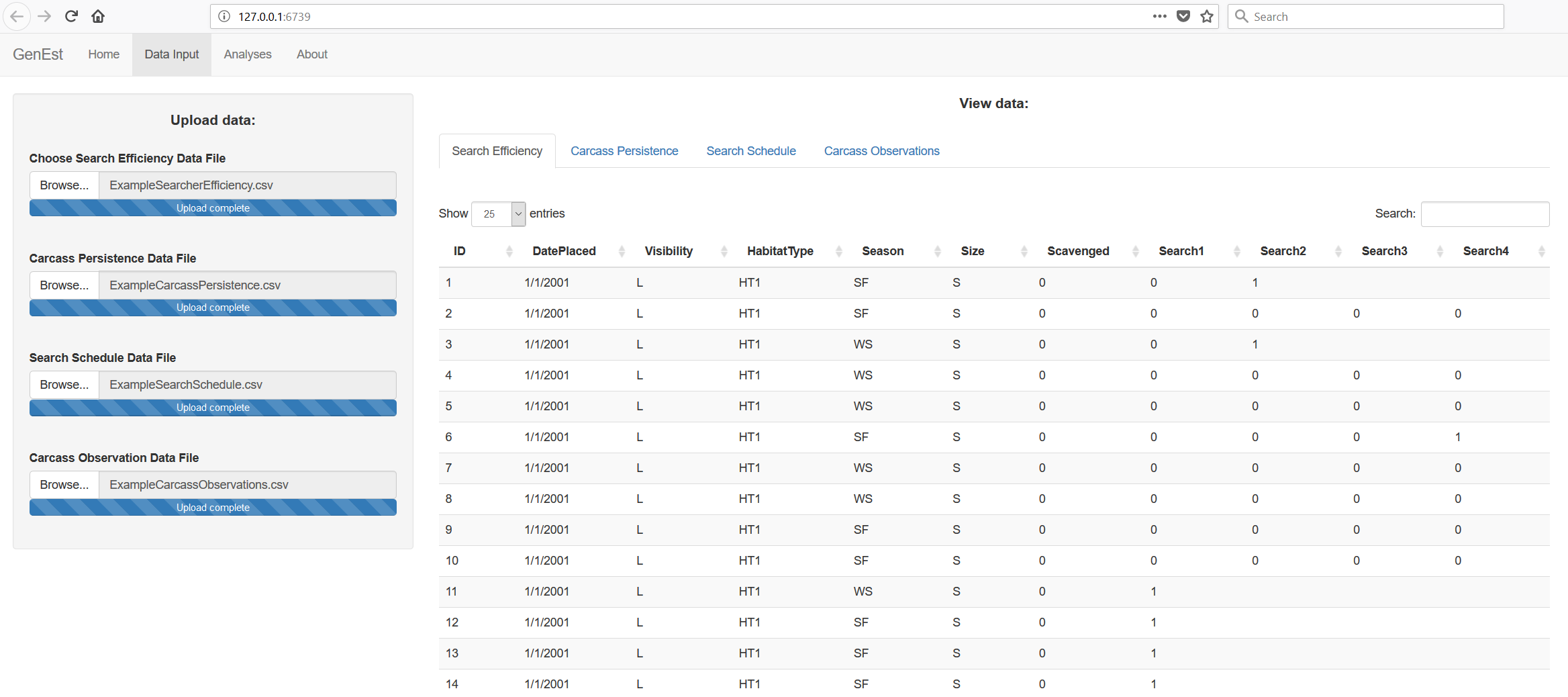


Figure 3. Visible uploaded data in the GenEst GUI.

# 3.0 Analyses

Following upload of the data files, the user should navigate to the “Analyses” tab.

## 3.1 General Inputs

Under the “General Inputs” subtab, the user provides two values: “Number of iterations” and “Confidence Level” (fig. 4) and is asked to supply a third (optional) value. The “Number of iterations” controls the number of replicates used to estimate parameters, defaults to 1,000, and must be an integer value 1 or larger, with current support up to 10,000. The “Confidence Level” defines the probability used to describe the spread of distributions (for parameters and state variables), defaults to 0.95 (a.k.a. 95%) and must be between 0 and 1. The user is asked to choose a “Size Class Column (optional).” The software offers as options those columns that occur in all three of Searcher Efficiency, Carcass Persistence and Carcass Observations input files.

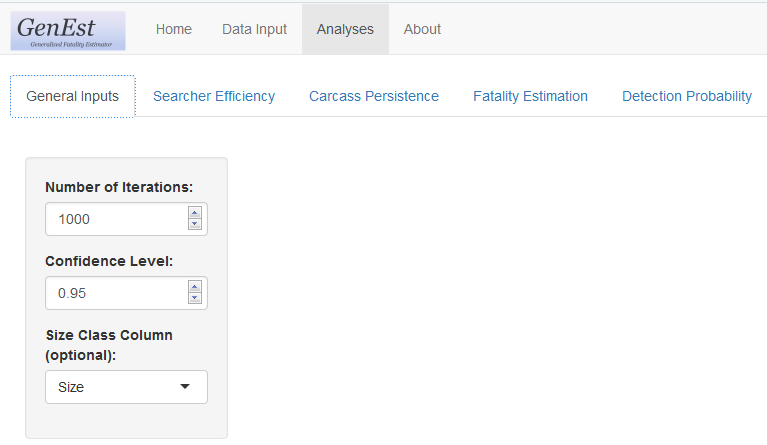


Figure 4. Required general inputs.

## 3.1 Searcher Efficiency

Searcher efficiency is the probability that a searcher will observe a carcass that is present in the search area when the search is conducted, represented by the parameter . It is typically measured by in trials conducted simultaneously with the search process, by individuals not conducting the searches. Trial carcasses, representing as closely as possible the species composition of fauna killed at the facility, are placed at regular intervals within the searched area, at locations known only to the trial conductor, not the searchers. After each search, if a trial carcass is not reported as found by the searchers, the trial conductor will visit the site to ascertain that indeed the carcass was present, but truly missed or no longer present and hence not eligible for inclusion in the searcher efficiency trial data. Trial conductors are encouraged to allow missed carcasses to remain in the field for potential discovery on a later search. When searches are conducted for several species simultaneously, searcher efficiency often varies with characteristics of the individual carcasses or conditions of its search, e.g., size of the carcass, carcass coloration, vegetation density, individual searcher, etc.

GenEst uses single or repeat carcass detection surveys to estimate searcher efficiency. Searcher efficiency is modeled with two parameters: , the probability that a carcass that is present is found during the first search after it landed, and , the proportional reduction in the probability that a present carcass is found on each subsequent search; means that carcasses are only ever found on the first search and means that the detection probability does not decrease over subsequent searches. Users also have the option to fix at any acceptable level (0-1), rather than estimate it from the data. If only one survey was completed, cannot be estimated and must be set by the user. Carcasses can be grouped into size classes and any combination of up to two discrete-valued predictors (e.g., season and vegetation class). The searcher efficiency parameters are estimated from the survey data in each class x size combination using maximum likelihood methods, as implemented with custom-written functions (<insert\_final\_function\_names>) for the GenEst package. For full details, see appendix B.1.

### 3.1.1 Model Inputs

Search efficiency model inputs are selected on the left-hand side panel of the Search Efficiency tab, Selected Data sub-tab (fig. 5).

* “Observations:”
  + The user is required to select the columns in the input data set that correspond to the search efficiency trial results (found = 1, missed = 0, NA = no longer available for each search occasion.
  + The columns must be selected in order of search occasion.
  + At least one column must be selected, and currently there is no maximum number of searches allowed.
* “Predictor Variables:”
  + The user can select up to two discrete-valued predictors they believe might affect searcher efficiency.
  + The predictor columns selected for searcher efficiency models must also be in the carcass observations data in order to be useable for fatality estimation.
  + The predictors selected in this analysis can be, but do not need to be, used in the carcass persistence analysis.
  + This input is entirely optional.
* “Fix k?”
  + The user has the option of estimating *k* if the data allow, i.e., repeated attempts to find the same carcasses, or fixing *k* at a particular value between 0 and 1 for all predictor variable and size combinations.

### 3.1.2 Outputs

Once all input data are selected, the user clicks on “Run Model.” A section titled “Table & Figure Selection:” will appear below “Model Inputs:” on the left and the “Model Comparison” sub-tab is automatically opened, but several output sub-tabs are available: Selected Data, Figures, Model Tables, Model Comparison, and Model Selection (**fig. 5**).

### 3.1.2.1 Data

Text.

### 3.1.2.1 Model Table

Text.

### 3.1.2.1 Figure

Text.

### 3.1.2.1 Model Selection

## 3.2 Carcass Persistence

Text.

## 3.3 Fatality Esitmation

## 3.4 Detection Probability

## 

# 4.0 Worked Examples

## 4.1 Example 1 – Wind facility, all animals, sampled turbines

Text.

### 4.1.1 Example subsubsection

Text.

## 4.2 Example 2 – Wind facility, bats, road and pad surveys of all turbines

Text.

### 4.2.1 Example subsubsection

Text.

## 4.3 Example 3 – Solar photovoltaic facility, all birds

Text.

### 4.3.1 Example subsubsection

Text.

## 4.4 Example 4 – Solar trough facility, all birds

Text.

### 4.4.1 Example subsubsection

Text.

## 4.5 Example 5 – Solar power tower facility, all animals

Text.

### 4.5.1 Example subsubsection

Text.

# 5.0 Conclusions

Text.

(1)



# Acknowledgments

# We would like to thank the GenEst steering committee for their tireless support of this project: Cris Hein with Bat Conservation International, Taber Allison with American Wind Wildlife Institute, Jerry Roppe with Avangrid Renewables and Dave Cowan with…. Special thanks go to Fränzi Korner Nievergelt of the Swiss Ornithological Institute and Robert Wolpert of Duke University for their contribution to the conceptual development of estimators and consultations on implementation of GenEst throughout our project. We thank [stats beta reviewers] for rigorous evaluation of our statistical approach and [user beta reviewers] for critical feedback that served to make this software and its user guide more user-friendly. [We are grateful to the many USFWS managers and industry consultants who tested and provided valuable feedback on earlier version of the software.] Funding for this project was provided by the U.S. Bureau of Land Management, the Ecosystems Mission Area Wildlife Program of the U.S. Geological Survey, the U.S. Department of Energy, WEST Inc. LLC and Avangrid Renewables. About the Authors and Collaborators

Numerous fatality estimators have been developed to estimate fatalities of birds and bats at wind energy facilities, but failure to meet their inherent assumptions can lead to radically different estimates of fatality, resulting in confusion and poor inferential capacity. The statisticians who developed and/or collaborated on GenEst are the authors of several of the estimators in current use [cite estimators]. Recognizing the commonalities and differences among our approaches [cite Wolpert], we realized they could (and should) all be incorporated under a single umbrella approach. The result is GenEst, or the Generalized Estimator (GenEst). It is “generalized” statistically and also in its application – it is generalized to apply to monitoring fatalities at solar- and wind-energy facilities, but can apply to many other situations, e.g., fisheries by-catch, power lines, oil spills, etc. The general approach we’ve taken doesn’t change with any of these. GenEst allows the user to evaluate assumptions regarding input parameters and select the approach that best reflects the situation and data.

Juniper Simonis is …

Dan Dalthorp is a statistician with the U.S. Geological Survey. He is author of the Evidence of Absence software and the Triggering OFR as well as several journal articles on estimation of mortality at renewable-energy facilities.

Manuela Huso is a research statistician with the U.S. Geological Survey. She is coauthor of the Evidence of Absence software and the Triggering OFR as well as several journal articles on estimation of mortality at renewable-energy facilities.

Lisa Madsen is…

Paul Rabie is…

Jared Studyvin is…

Robert Wolpert is…

Fränzi Korner-Nievergelt is…

# References Cited

Chang, W., Cheng, J., Allaire, J.J., Xie, Y., and McPherson, J., 2017, shiny: web application framework for R, version 1.0.5, <https://CRAN.R-project.org/package=shiny>.

R Core Team, 2017, R: A language and environment for statistical computing, v3.4.3: R Foundation for Statistical Computing, Vienna, Austria, [https://www.R-project.org/](%20https:/www.R-project.org/).

# Appendix A. Input File Formats

This appendix details the formats of the input files for the GenEst GUI.

## A.1 Search Efficiency

With text!

## A.2 Carcass Persistence

With text!

## A.3 Search Schedule

With text!

## A.4 Carcass Observations

With text!

# Appendix B. Analytical Details

This appendix provides additional information regarding the analyses underlying the GenEst software.

## B.1 Search Efficiency

With text!

## B.2 Carcass Persistence

With text!

## B.3 Search Schedule

With text!

## B.4 Carcass Observations

With text!

1. DAPPER Stats. [↑](#footnote-ref-1)
2. U.S. Geological Survey. [↑](#footnote-ref-2)
3. Oregon State University. [↑](#footnote-ref-3)
4. Western EcoSystems Technology, Inc.. [↑](#footnote-ref-4)