

# Evidence of Absence Software User Guide

Data Series 881



# **Evidence of Absence Software User Guide**

By Dan Dalthorp, Manuela Huso, David Dail, and Jessica Kenyon

Data Series 881

**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:

Dalthorp, Dan, Huso, Manuela, Dail, David, and Kenyon, Jessica, 2014, Evidence of absence software user guide: U.S. Geological Survey Data Series 881, 34 p., <http://dx.doi.org/10.3133/ds881>.

ISSN 2327-638X (online)

# Contents

Abstract.....	1
Introduction.....	1
Installation Instructions.....	1
Requirements.....	1
Installation.....	2
RExcel .....	2
Evidence of Absence Software .....	2
Getting Started .....	3
Statistical Approach and Terminology .....	4
General Idea .....	4
Terminology and Parameters.....	4
Overall Detection Probability .....	4
Prior and Posterior Distributions.....	4
Management Parameters .....	5
Ecological Parameters.....	5
Data Management.....	6
Entering Data for Estimation .....	6
One Site, One Year .....	6
Entering Parameters .....	6
Informed Prior Distributions.....	7
Multi-Year Total .....	9
Multiple Sites .....	9
Restore Defaults .....	10
Entering Data for Design .....	10
Background Parameters .....	10
Management Parameters .....	10
Graphing Options .....	10
R graphs.....	10
Analysis and Interpretation: A Tutorial .....	11
Scenario 1: Analyzing Search Data (Eagles).....	11
Management Variables.....	12
Background Parameters .....	12
Scenario 2: Optimizing the Search Protocol (Eagles).....	15
Scenario 3: Indiana bat.....	18
Post-Search Estimation: Data Entry and Calculations .....	18
Post-Search Estimation: Interpretation .....	20
Informed Prior.....	20
Optimizing the Search Protocol.....	22
Scenario 4: Multiple sites .....	25
Acknowledgments.....	26
References Cited.....	26
Appendix A. Accounting for Uncertainty in Searcher Efficiency and Carcass Persistence .....	27

## Contents—Continued

Appendix B. Choice of Representative Interval Length for Calculating $r$ .....	29
Appendix C. Parameterizations of Distributions Used in the Software .....	30
Appendix D. Using Previous Year’s Search Data to Inform Current Year’s Prior Distribution .....	31
Appendix E. Using Several Previous Years’ Data to Inform Current Year’s Prior Distribution .....	32
Appendix F. Estimating Total Number of Fatalities Over Several Years (or Several Sites) .....	33

## Figures

1. Screen capture of the Evidence of Absence software Parameters sheet.....	3
2. Screen capture of the Evidence of Absence software Edit Parameters form for entering search parameters .....	11
3. Screen capture of the Evidence of Absence software Weibull Parameters form for entering persistence distribution parameters .....	12
4. Graph generated by Evidence of Absence software showing Posterior Distribution of the eagles example for scenario 1, analyzing search data. ....	13
5. Screen capture of Design Tradeoffs form for eagles example for scenario 2, optimizing the search protocol.....	15
6. Graph generated by Evidence of Absence software showing design tradeoffs for search interval and search coverage for long carcass persistence time.....	16
7. Graph generated by Evidence of Absence software showing design tradeoffs for search interval and search coverage for short carcass persistence time .....	17
8. Screen capture of the Edit Parameters form showing initial search parameters for the Indiana bat example for scenario 3 .....	18
9. Graph generated by Evidence of Absence software showing the posterior distribution of $M$ for the Indiana bat example for scenario 3 using uniform prior distribution .....	19
10. Graph generated by Evidence of Absence software showing posterior distribution of $M$ for the Indiana bat example for scenario 3 using informed prior distribution.....	21
11. Screen capture showing the Design Tradeoffs form for the Indiana bat example .....	22
12. Graphs generated by Evidence of Absence software showing design tradeoffs between search coverage, searcher efficiency, and search interval for the Indiana bat example for scenario 3.....	23
13. Graph generated by Evidence of Absence software showing posterior distribution for the Indiana bat example with informed prior distribution and intensified search schedule .....	24
14. Graph generated by Evidence of Absence software showing posterior distribution for multiple sites for scenario 4.....	25

## Tables

1. Example format of comma-separated values (.csv) file for input of search parameters.....	8
2. Example format of .csv input file for several years of previous data to inform current year’s prior distribution.....	9
3. Example format of .csv input file for calculating total mortality at several sites.....	10
4. Full posterior distribution results from Parameters sheet for the eagles example in scenario 1, analyzing search data .....	14

# Evidence of Absence Software User Guide

By Dan Dalthorp<sup>1</sup>, Manuela Huso<sup>1</sup>, David Dail<sup>2</sup>, and Jessica Kenyon<sup>1</sup>

## Abstract

Evidence of Absence software (EoA) is a user-friendly application used for estimating bird and bat fatalities at wind farms and designing search protocols. The software is particularly useful in addressing whether the number of fatalities has exceeded a given threshold and what search parameters are needed to give assurance that thresholds were not exceeded. The software is applicable even when zero carcasses have been found in searches. Depending on the effectiveness of the searches, such an absence of evidence of mortality may or may not be strong evidence that few fatalities occurred. Under a search protocol in which carcasses are detected with nearly 100 percent certainty, finding zero carcasses would be convincing evidence that overall mortality rate was near zero. By contrast, with a less effective search protocol with low probability of detecting a carcass, finding zero carcasses does not rule out the possibility that large numbers of animals were killed but not detected in the searches. EoA uses information about the search process and scavenging rates to estimate detection probabilities to determine a maximum credible number of fatalities, even when zero or few carcasses are observed.

## Introduction

The Evidence of Absence software addresses two broad questions:

1. Planning (pre-sampling): What monitoring protocol (search interval, searcher efficiency, search coverage, carcass persistence pattern, etc) will allow us to assert with  $100(1 - \alpha)\%$  credibility that some fatality threshold ( $\tau$ ) was not exceeded given that  $x$  carcasses are observed?
2. Estimation (post-sampling): Given the number of carcasses observed (potentially zero), search effort, and estimated detection probability parameter values, what fatality levels can be ruled out? Users

will be able to make statements like, “Given the observed carcass count and search parameters, the probability that actual number of fatalities of RareSpeciesX exceeded a set limit,  $\tau$ , is  $\leq \alpha$ .”

Other statistical tools we currently have for estimating actual fatality from observed carcasses (for example, Shoenfeld, 2004; Huso, 2011; Korner-Nievergelt and others, 2011; Warren-Hicks and others, 2013) are fairly robust when observed carcass counts are high, but these tools cannot estimate a non-zero fatality rate if no carcasses are observed. EoA uses Bayes’ formula to calculate credible intervals for the actual mortality. The conceptual framework for this approach can be found in Huso and others (in press).

## Installation Instructions

The freeware statistics software, R, is the main calculation and graphics engine for the Evidence of Absence software. Error checking, data management, and user-interface are executed through Excel<sup>®</sup>. Communication between R and Excel is accomplished through an Excel add-in called RExcel and its companion statconnDCOM.

## Requirements

- **Microsoft Windows<sup>®</sup>** XP, 7, 8 or later; 64-bit or 32-bit.
- **Microsoft Excel<sup>®</sup>** 2003 or later (32-bit only). Although 64-bit Excel<sup>®</sup> is available in recent Microsoft Office<sup>®</sup> releases, the 32-bit version is the default installation even on 64-bit machines because of compatibility issues with other programs.
- **R** software, version 2.15.1 or higher.
- **RExcel** and associated components.
- **Evidence of Absence, v.x.x.xlsm**, Excel<sup>®</sup> workbook and associated R source files.

---

<sup>1</sup>U.S. Geological Survey.

<sup>2</sup>JPMorgan Chase, Columbus, Ohio.

### Installation

There are several steps to installing the software. Do not skip any steps.

#### RExcel

1. Log into Windows with administrator privileges
2. Use your internet browser to navigate to the latest installation instructions for RExcel at: [http://homepage.univie.ac.at/erich.neuwirth/php/rcomwiki/doku.php?id=wiki:how\\_to\\_install](http://homepage.univie.ac.at/erich.neuwirth/php/rcomwiki/doku.php?id=wiki:how_to_install)
3. Select scenario that best describes your situation and follow instructions for that scenario.
  - **Scenario 1: You do not use R, and it is not installed on your computer.**
  - **Scenario 2: R is already installed on your computer, but you have to (or wish to) upgrade to a newer version.**
  - **Scenario 3: You already have a compatible version of R installed and do not wish to upgrade to a newer version.**

**Scenario 1:** You do not use R, and it is not installed on your computer.

1. Use your internet browser to navigate to <http://cran.r-project.org/mirrors.html> and click the location nearest to you.
2. Click **Download R for Windows > install R for the first time > Download R x.x.x for Windows to access installation .exe file.**
3. **Save** the installation file and run it manually after it has finished downloading.
4. After successful installation of R, proceed to scenario 3.

**Scenario 2: R is already installed on your computer, but you have to (or wish to) upgrade to a newer version.**

1. Uninstall your previous version of R then follow the instructions in scenario 1 for installing the latest version of R
2. Associate any existing R packages with the newly installed R. Copy the packages from your old library (in the folder for the old installation) to the library folder in the new installation. Then, run the command `update.packages(checkBuilt=TRUE, ask=FALSE)` in the new R.
3. After successful installation of R, proceed to scenario 3.

**Scenario 3: You already have a compatible version of R installed and do not wish to upgrade to a newer version.**

1. Use your internet browser to navigate to [http://homepage.univie.ac.at/erich.neuwirth/php/rcomwiki/doku.php?id=wiki:how\\_to\\_install](http://homepage.univie.ac.at/erich.neuwirth/php/rcomwiki/doku.php?id=wiki:how_to_install) and follow the instructions in the section, “How to install RExcel when R is already installed.” The procedure is not complicated, but there is little room for error. When RExcel is installed, a number of informational message boxes may pop up. These provide more information than most users will need. However, one critical piece of information is that after the installation is complete, it may be necessary to navigate to `statconn | RExcel` among Program Files in Windows Explorer and click **Activate RExcel Add-in**. (It is not always necessary to do this, but it is easier to just do it than to risk the frustration of the program not working and having no idea why not.)

#### Evidence of Absence Software

1. Use your Internet browser to navigate to <http://pubs.usgs.gov/ds/0881/>.
2. Click on **evidence\_of\_absence\_v1.00.zip**.
3. Download the folder to a convenient location on your computer.
4. Ensure:
  - all files are extracted from the compressed zip file
  - the `r_functions` folder are in the same folder where the `.xslm` workbook is stored
  - Excel® security setting allows macros to be run

The Excel® workbook has many macros for error-checking data and sending commands back and forth between Excel and R, so permissions may need to be managed on some machines.

NOTE: If you needed to click on a yellow bar in the workbook to enable permission to run macros, click the **Reset R** button at the bottom of the “Actions” list before running other commands. Otherwise, you may get a long line of indecipherable error messages from R (R functions that are normally defined at startup were not defined because macros were disabled) and maybe even one or two from Excel® as well.

The software automatically checks if the required R packages have been installed (`actuar`, `Matrix`, `tensorA`, `MASS`, and `VGAM`). If packages have not been previously installed, a window may open, prompting you to select a `CRAN Mirror`

(a location to from which to download the R package(s)). Select the nearest site. Downloading and installing the packages is done automatically, but the process can take several minutes.

## Getting Started

Start the Evidence of Absence software by opening the Excel® workbook that was downloaded with the package. This workbook is richly equipped with macros for doing internal calculations and for communicating with R. Users have the option of entering data through forms in the workbook or through properly formatted comma-separated values (.csv) files. The most recent results are automatically saved in the workbook, but users also may want to manually save numerical results to .csv files or save graphics to .jpg's or other graphical formats, or both. Options for entering data and

saving results are straightforward. Some users will prefer to experiment with the buttons in the workbook right away, while others will prefer to first read the details in the this user guide.

The software opens to the **Parameters** worksheet ([fig. 1](#)), which shows the most recent parameter set that was saved. The first time the workbook is opened, an arbitrary (but valid) parameter set is loaded as a default. Internal calculations strictly require the parameter set on this worksheet to be valid (for example, non-negative integer carcass count, or prior distribution selected from a list of admissible distributions and spelled correctly). Users may edit the sheet only by using Excel® forms or by uploading .csv files. Before writing data to the **Parameters** worksheet from forms or .csv files, the workbook runs a series of error checks and loads the data only if there are no errors. If the error-checking routine fails and an invalid parameter set is written to the sheet, the program may stop working. The user can click **Restore Defaults** to reload the valid, default parameter set to restore full functionality.

Parameter	Current Value	Prior distribution	Sampling Dates
carcass count (X)	0	m P(M = m)	0
sampling coverage (A)	0.75	0 0.005	7
searcher efficiency (p)	0.5	1 0.005	14
min p	0.25	2 0.005	21
max p	0.75	3 0.005	28
k	0.8	4 0.005	35
Sampling Dates	Formula	5 0.005	42
interval (I)	7	6 0.005	49
span	182	7 0.005	56
Persistence Distribution	Exponential	8 0.005	63
α	0.1	9 0.005	70
β	10.0	10 0.005	77
min b	4.0	11 0.005	84
max b	20.0	12 0.005	91
mean persistence (CP)	10.0	13 0.005	98
P(persist until 1st search)	0.719	14 0.005	105
interval (I <sub>r</sub> )	7	15 0.005	112
Prior Distribution	Uniform	16 0.005	119
max	200	17 0.005	126
Credibility Level (1 - α)	0.9	18 0.005	133
Arrival Function	Uniform	19 0.005	140
---	NA	20 0.005	147
---	NA	21 0.005	154
		22 0.005	161
		23 0.005	168
		24 0.005	175
		25 0.005	182
		26 0.005	
		27 0.005	
		28 0.005	
		29 0.005	
		30 0.005	
		31 0.005	
		32 0.005	
		33 0.005	

**Figure 1.** Screen capture of the Evidence of Absence software **Parameters** sheet.

After an analysis, results of calculations are (typically) shown on the **Parameters** worksheet to the right of the parameter set from which the calculations were based. Graphical results, results from **Design Tradeoffs**, and results from multiple year or multi-site analyses are exceptions which have a different format.

## Statistical Approach and Terminology

### General Idea

When we are interested in estimating the total number of animals killed in an area, within a set period of time, based on the number of carcasses observed during our searches, there are two scenarios that would result in zero carcasses being observed: (1) there were zero fatalities, or (2) there were some fatalities but all the carcasses were missed in the searches. If there is perfect detection of carcasses, then “zero carcasses observed” means “zero animals killed.” In the case of imperfect detection, though, we cannot rule out the possibility that some animals were killed but the carcasses were missed because they landed in unsearched areas, were eaten or carried off by scavengers, or were missed for some other reason (for example, difficult terrain, bad lighting, or bad luck). These factors combine to give an overall probability of detecting a carcass. Although we cannot be certain of how many actual fatalities there were, we can use information about the overall probability of detection to rule out the possibility of large numbers of fatalities. For example, with a protocol that results in a relatively high probability of detecting a carcass, we might be able to credibly assert that there were no more than, say, three fatalities—if there had been more than that, we would have observed some carcasses. With a protocol that results in a relatively low probability of detecting a carcass, we might not be able to rule out 20, 30, or even more fatalities. That is because our detection probability is so low that even if there were many fatalities, there is still a reasonable chance we would miss all of them in our searches.

The EoA software evaluates the overall probability of detection and then uses Bayes’ formula to quantify the credibility of assertions such as, “the total mortality  $M$  did not exceed 3.” The software does the calculation not just for  $M = 3$  but for all whole numbers  $m$  and compiles the results into a list called the *posterior distribution* of  $M$ , which represents the current state of knowledge about the total mortality.

The posterior distribution depends on the overall probability of detection and the observed carcass count. In addition, prior knowledge (or lack thereof) about expected mortality levels can be used to inform the calculation of the posterior distribution. For example, using our prior knowledge

about the Hawaiian hoary bat (*Lasiurus cinereus semotus*), we could rule out  $M > 0$  at a wind farm in Antarctica without doing any searches. Our posterior distribution would then be  $P(M = 0) = 1$  and  $P(M = m) = 0$  for  $m > 0$ . By contrast, at a wind farm on Oahu, we cannot rule out anything ahead of time. In that case, we might assume a uniform *prior distribution* (that is,  $P(M = m)$  is the same for all  $m$  up to some large number) to reflect the view that anything is possible. If we then find zero (or very few) carcasses despite an effective monitoring protocol, we can rule out the possibility that large numbers were killed. We can then use what was learned about the system in one year to build a refined prior distribution for use in the following year. In this way, we can incrementally improve the estimates each year.

### Terminology and Parameters

#### Overall Detection Probability

The EoA software estimates the overall probability ( $g$ ) of detecting a carcass that arrives at a site (or group of sites) during the monitoring season(s). The estimate of  $g$  takes into account scavenging rate, searcher efficiency, search schedule, change in searcher efficiency for older carcasses, and changes in expected carcass arrival rates through time. A point estimate of the total number of fatalities is:

$$\hat{M} = X / \hat{g} \quad (1)$$

where

$X$  is the number of carcasses observed, and  
 $\hat{g}$  is the estimated overall detection probability.

#### Prior and Posterior Distributions

The *posterior distribution* is a quantitative representation of the current state of knowledge about the total number of fatalities  $M$  during the monitoring period. The posterior distribution is expressed as a probability distribution  $P(M = m)$  or  $P(M > m)$  for integer values of  $m$ . It is used to construct credible intervals (CIs) for  $M$  to rule out mortality levels outside the interval.

The *prior distribution* is a quantitative representation of the state of knowledge about what the total number of fatalities  $M$  is before taking into account the monitoring data and evaluation of the sampling protocol. In the absence of any such prior knowledge, a uniform, “anything is possible,” prior distribution is used. Knowledge from previous years’ monitoring can be used to inform and improve the prior distribution used in the later years.

Central to the calculation of the posterior distribution are a number of parameters that determine the overall probability of detection. The parameters fall into two broad categories, *management parameters* and *ecological* (or *background parameters*). The management parameters are largely a function of management decisions about sampling intensity (in both time and space). The ecological parameters are less obviously amenable to change.

## Management Parameters

*Sampling coverage* ( $a$ ) is the fraction of the total carcasses expected to arrive in the searched area. For example, if there are 10 turbines and 3 are intensively searched to a wide enough radius to encompass all the carcasses and there is no unsearched area within that search radius, then  $a = 0.3$ . Another example: if there is one turbine that is only partially searched,  $a$  is the fraction of carcasses that are expected to land in the searched area around the turbine.  $a$  must be in the interval  $(0, 1]$ . It must be emphasized that  $a$  is not the fraction of the area that is searched but the fraction of carcasses arriving in the searched area. The number of carcass falling at a given distance from a turbine tends to decrease with distance whereas the area increases, so area nearer to a turbine accounts for more carcasses than does an area of equal size at a greater distance from a turbine. Thus,  $a$  should be a density-weighted proportion (DWP) of the area sampled (Huso and Dalthorp, 2014).

*Searcher efficiency* ( $p$ ) is the probability of observing a carcass that is present in the searched area at the time of the search. Values must be in the interval  $(0, 1]$ . Searcher efficiency values are typically determined by searcher efficiency (SE) trials conducted on an ongoing basis during routine searches throughout the season (Huso and others, 2012). The software accounts for uncertainty in searcher efficiency and requires users to enter the bounds on a 95 percent confidence interval (95% CI) for searcher efficiency. The CI for search efficiency is based on field trials and is calculated using other software (for example, Huso and others, 2012; [appendix A](#)).

$k$  is the factor by which search efficiency changes with each successive search. Carcasses missed in one search tend to be more likely to be missed in subsequent searches because:

- older carcasses are harder to detect, and
- carcasses that are missed in a search may have been missed because they landed in a spot that was particularly difficult to see.

For example, the probability of detecting a carcass in the first search after it arrives is  $p$  (assuming that it persists unscavenged until the search). If it is not found in the first search but is still present at the time of the second search, the probability of detection drops from  $p$  to  $k \cdot p$ . If it is missed again on the second search but persists until the third search,

the probability of detection decreases by another factor of  $k$  to  $k^2 p$ , and similarly decreases by a factor of  $k$  with each search. If searcher efficiency is the same no matter how long a carcass has been in the field or how many times it has been missed, then  $k = 1$ . If a carcass missed in the first search after arriving has no chance of being found in later searches (or not included in the carcass counts), then  $k = 0$ .

$I$  and *span* are the time interval (days) between searches and the total time spanned by the searches, respectively. Users are given the option to enter either  $I$  and *span*, or a list of sampling dates.

## Ecological Parameters

*Carcass persistence*. —The amount of time a carcass persists without being lost to scavenging or decay is a random variable characterized by a probability distribution. The user can select the distribution family (exponential, Weibull, lognormal, or loglogistic) and parameters ( $\alpha$  = shape and  $\beta$  = scale; or CP = mean carcass persistence time and  $r$  = the probability that a carcass that arrives in an interval of length  $I_r$  persists through the interval ([appendix B](#)). Carcass persistence distributions are generally determined by carcass persistence trials conducted during the monitoring season.

*Distributions* ([appendix C](#)).—The user selects one of four distributions for carcass persistence times: exponential, Weibull, loglogistic, or lognormal. The exponential distribution makes a mathematically convenient assumption that the scavenging rate does not depend on age of carcass. It requires only one parameter, which the user may enter as either  $1/\lambda = \text{meanCP}$  or as  $r = \text{probability that a carcass persists until the first search after it arrives}$ . In practice, the rate of carcass removal tends to change with time, and the exponential distribution usually does not accurately reflect the true persistence distribution. Frequently, it underestimates the scavenging rate for fresh carcasses and, as a consequence, underestimates the number of fatalities. Instead, the Weibull, log-logistic, or lognormal persistence model is preferred (Bispo, and others, 2012; Warren-Hicks and others, 2013). All three are more flexible and often more realistic than the exponential.

With the lognormal distribution, the attrition rate starts at zero for fresh carcasses, rises sharply to a peak a short time later, and gradually declines. The Weibull has more flexibility for reflecting a variety of patterns for fresh carcasses (for example, attrition rate may either increase or decrease initially), but may not fare as well for longer persistence times. The log-logistic has a “heavy tail” and can accommodate a situation in which some carcasses persist for a very long time despite having a high initial rate of disappearance.

## 6 Evidence of Absence Software User Guide

*Uncertainty.*—For all distributions, the user is required to provide a 95% CI for the scale parameter ( $\beta$ ) or the probability that a carcass persists ( $r$ ) until the first search after it was killed (assuming a search interval of  $I_r$ , [appendix B](#)). Uncertainty parameters are based on analysis of field trials and are calculated using other software (for example, the Fatality Estimator by Huso and others, 2012). The Evidence of Absence software accounts for uncertainty about these parameters.

*Relative arrival rate.*—Is fatality rate uniform throughout the sampling period, or does it vary by season? The software will accommodate various seasonal carcass arrival patterns, for example, if more carcasses are present in the fall than in the spring or summer, the user may select an arrival rate function that reflects expected populations in the field. The fatality estimates are robust with respect to arrival rate assumptions, and estimates of total fatality usually will not vary noticeably with choice of arrival function.

## Data Management

Most of the data entry is done using Excel® forms, which implement extensive error-checking. Data-entry fields turn yellow when invalid parameters are entered, and invalid parameter sets cannot be analyzed or saved by the software. Users sometimes also have the option of entering data using .csv files or Excel® worksheets.

### Entering Data for Estimation

#### One Site, One Year

The estimation of number of fatalities requires the user to enter the observed carcass count with a parameter set characterizing the monitoring protocol. Parameters include sampling coverage, searcher efficiency,  $k$ , sampling schedule, carcass persistence distribution, carcass arrival function, and prior distribution. Parameters may be entered using either Excel® forms or .csv files.

Clicking the **Edit Parameters** button opens a form for entering the parameter values. By default the form is populated by values read directly from the **Parameters** sheet, which contains the last parameter set saved. The user may then modify any of the parameter values. The form performs extensive error-checking. Valid parameter sets may then be saved to the main **Parameters** worksheet (when the user clicks **Save**) or saved and sent directly to R for analysis (when the user clicks **Calculate Posterior and Save**). Invalid

parameter sets cannot be saved or sent to R for analysis. Errors in parameter values are highlighted in yellow, and the **Save** and **Calculate Posterior and Save** buttons remain disabled as long as there are any errors.

#### Entering Parameters

*Carcass Count ( $X$ )* represents the total number of carcasses observed in all searches during the season.  $X$  must be a non-negative integer.

*Sampling Coverage ( $a$ ), Searcher Efficiency ( $p$ ), and Proportional Change in Searcher Efficiency with Each Search ( $k$ )* must all be in the interval (0, 1], but  $k = 0$  also is permissible. Bounds on a 95% CI for  $p$  also must be entered.

*Credibility Level ( $1-\alpha$ )* is the desired level of credibility for making statements like, “Based on the search data, we can assert with  $100(1-\alpha)\%$  credibility that there were not more than  $\tau$  fatalities.” The credibility level in Bayesian statistics is analogous to confidence level in traditional statistics.

*Persistence Distribution:* Select the desired distribution family (exponential, Weibull, log-logistic, or lognormal). Default parameters are the last saved values for whichever distribution is selected. To edit parameters, double-click the distribution name or click the **Edit Parameter Values** button to bring up a form. Parameters for exponential distribution may be entered either as meanCP = average persistence time or as  $r$  = probability carcass persists until the first search after arrival (length of the search interval,  $I_r$ , that  $r$  is based on also must be given). In addition, bounds for a 95% CI for the meanCP or  $r$  also must be entered. Parameters for the other distributions may be entered either as  $\alpha$  and  $\beta$  (according to parameterizations given in [appendix A](#)) or as meanCP and  $r$ . The value of  $r$  depends on the length of the interval for which it is calculated. The user provides the interval ( $I_r$ ) that the  $r$  value is based on. Normally,  $r$  would be calculated for an interval that is in some sense typical for the given search protocol (for example, Huso and others, 2012) uses the mode of the search intervals), but the choice of interval length is not critical as long as the  $r$  value that is provided by the user is based on the  $I_r$  that also is provided. Bounds for the 95% CI for  $\beta$  or  $r$  also must be provided by the user. These values are based on analysis of field trial data and are calculated outside the EoA software (for example, Huso and others, 2012). Within the forms for editing the persistence distribution parameters, clicking **Cancel** closes the form and retains the previous values for the selected distribution. **View** calls on R to draw a graph of the distribution. The user may save the graph, leave the window open and continue working, or close the graph by clicking the red “X”.

*Sampling Dates (I and Span):* Sampling dates are entered as the number of days since the start of the monitoring period. For example, if there are weekly searches beginning after a June 15 “clean-out” search, the dates are entered as 0, 7, 14, and so on, corresponding to June 15, June 22, June 29, and so on. There are two ways to enter dates:

- Dates may be entered by **Formula**, wherein the user enters the interval ( $I$ ) between the searches and the total time spanned ( $span$ ) by the monitoring season. The span must be evenly divisible by  $I$ . For example, if there are weekly searches ( $I = 7$ ),  $span$  can be 77 but not 79.
- If search intervals vary, a **Custom** list of dates (for example, 0, 7, 12, 19, 22, 29) may be entered into an Excel® worksheet and loaded into the program. The user selects **Custom** dates in the **Edit Parameters** form, which will open a worksheet for reading dates. Dates may be entered by hand or by copy-and-paste from a different file. When a custom list of dates is entered, the caption in the **Sampling Dates** section of the form will show the span as the date of the last search and average interval = span/number of searches. This is strictly for convenience. All analyses use the actual list of dates as entered rather than the simple summary shown in the caption.

*Prior Distribution:* There are three choices for prior distribution of  $M$ : uniform, informed, and user-defined. In the absence of any prior information (especially search data from previous years), a **uniform** prior distribution is typically used because it gives full voice to the current data. User enters the maximum number of fatalities that may have occurred ( $M_{max}$ ). A value of  $M_{max} = 200$  is often reasonable. Larger values would not noticeably change the results but would slow the calculations, whereas very large values would crash the program. Small values could skew the results, especially if the overall probability of detection is small. Previous years’ search data can be incorporated into the model using an **informed** prior distribution ([appendix D](#)). Alternatively, knowledgeable users who have developed their own prior distribution based on other models and information, may enter their prior distribution as **User-defined**.

*Arrival Function:* The arrival function expresses the relative numbers of carcasses arriving at various times throughout the monitoring season. The analyses do not require input about the absolute numbers of animals being killed, but relative changes in mortality rates through time (for example, more fatalities in late summer than in early spring) are taken into account. However, if the sampling interval remains constant (or nearly constant) through the year, then the arrival function does not have a strong effect, and a Uniform relative arrival rate

is appropriate. Otherwise, user may select Beta distribution parameters to reflect expected variation in mortality rates through time.

Users may save input parameter sets to .csv files for later consideration. The **Save Parameters and Results to .csv** button on the **Parameters** sheet writes the current parameter set shown in the **Parameters** sheet to a .csv file after prompting the user for file name and path. The format of the saved .csv is compatible with the **Read Dates and Parameters from .csv** button, which loads previously saved data. Parameter sets are saved either with or without corresponding results, depending on whether results are shown on the **Parameters** sheet when saved. Alternatively, users may create their own .csv files for loading into the program. The format must match that of files saved using the **Save Parameters and Results to .csv** button, including the spelling and capitalization conventions of all names in the second column; valid parameter values entered into the proper cells; sampling dates beginning at  $t = 0$  with all other values positive and no duplicates; and prior distribution defined for consecutive non-negative integers beginning with 0 with probabilities all between 0 and 1 and summing to 1, etc. However, it is normally easier and safer to enter data using the buttons in the Evidence of Absence software rather than creating a .csv file by hand. An example of a parameter input file is shown in [table 1](#).

After the search parameters and prior distribution parameters have been entered, the posterior distribution can be calculated by: on the **Edit Parameters** form, click the **Calculate Posterior and Save** button; or, on the **Parameters** sheet, click the **Estimate Mortality for Given Parameter Set** or the **Plot Posterior Distribution** button. Results are displayed in the **Results** area to the right of the input parameter set. A graph of the posterior can be viewed by clicking **Plot Posterior Distribution**, and the complete set of inputs and associated posterior distribution can be saved to a .csv file for later review or subsequent analysis.

## Informed Prior Distributions

Search data from previous years can be used to inform the current year’s prior distribution and improve estimation ([appendix D](#)). This requires the user to enter search counts, overall detection probabilities, bounds on 95% CIs for the detection probabilities for each previous year’s data; and, if more than one previous year is considered, the relative weight among years as described in the following paragraphs. This is done from the **Edit Parameters** form by selecting **Informed** as the Prior Distribution and clicking **Edit Parameter Values** or by double-clicking “Informed.”

**Table 1.** Example format of comma-separated values (.csv) file for input of search parameters.

Parameter	Current value	Prior distribution		Sampling dates
		m	P(M = m)	
carcass count (X)	0	m	P(M = m)	0
sampling coverage (a)	0.75	0	0.0094	7
searcher efficiency (p)	0.5	1	0.0307	14
p_lwr	0.25	2	0.0583	21
p_upr	0.75	3	0.0842	28
k	0.8	4	0.1024	35
Sampling Dates	Formula	5	0.1106	42
interval (I)	7	6	0.1094	49
span	182	7	0.101	56
Persistence Distribution	Exponential	8	0.0884	63
a	0.1	9	0.074	70
b	10	10	0.0597	77
b_lwr	4	11	0.0466	84
b_upr	20	12	0.0355	91
mean persistence (CP)	10	13	0.0264	98
r = P(persist until search)	0.719164	14	0.0192	105
interval (I <sub>r</sub> )	7	15	0.0137	112
Prior Distribution	User-defined	16	0.0097	119
95th percentile	14	17	0.0067	126
Credibility Level (1 - $\alpha$ )	0.9	18	0.0046	133
Arrival Function	Uniform	19	0.0031	140
–	NA	20	0.002	147
–	NA	21	0.0013	154
		22	0.0009	161
		23	0.0005	168
		24	0.0003	175
		25	0.0002	182

To incorporate just one previous year's search data into the current year's analysis, the user may manually enter the previous year's carcass count ( $X$ ); the overall probability that a carcass that arrives in the area of interest is actually observed ( $g = P(\text{obs}|\text{arr})$ ); and the bounds on the 95% CI for  $g$ . These values may be calculated using the Evidence of Absence software with the previous year's search parameters or may be derived independently from a different program or model. Alternatively, if the user has saved the analysis (including results) from a previous year to a .csv file, the data may be read from that .csv. The model for creating a prior distribution from a single previous year's data assumes that site operational parameters relating to total kill rate (for example, number of turbines, cut-in speed, or deterrents) do not change from the prior year to the current year; however, search parameters (for example, searcher efficiency, carcass

persistence, or search schedule) may change. The new, informed prior distribution is calculated and saved by the software for use with the current year's carcass count and search parameters.

If more than one year of previous data are to be incorporated into the analysis, counts ( $X$ ) and overall detection probabilities ( $g$ ) and their 95% CIs for each year must be entered. Additionally, the user must provide an estimate of the relative arrival rate for each previous year compared to the current year. If site operational parameters change through time (for example, new turbines added or curtailment or deterrents used to decrease fatality rates), or if animal activity varies noticeably (for example, migration routes shift), the year-to-year changes should be accounted for in the values entered for relative weights. Data be may be entered using .csv

file(s) or manually. Data from separate years may be combined into a single file with separate columns for year; carcass count ( $X$ );  $g$ ;  $g_{lwr}$  = lower bound on 95% CI for  $g$ ;  $g_{upr}$  = upper bound on 95% CI for  $g$ ; and relative weight (rel\_wt) assigned to each year. The rel\_wt for the current year is assumed to be 1, and previous years' weights represent the relative expected number of carcasses for each year. For example, in [table 2](#), years 4 and 5 have rel\_wt = 1 because there is no reason to expect a systematic change in the number of fatalities from year 4 to year 5 to the current year—the number of turbines did not change, there were no fatality minimization measures taken, local population size did not change dramatically, etc. Relative weights for years 1, 2, and 3 reflect site expansion, as the number of turbines operating at the site increased from 41 to 83 to 100.

The model is robust with respect to the relative weights, and moderate misspecification will rarely have a significant impact on the final estimates. However, it is crucial that weights are entered on a scale with the current year (yr = 6 and not represented in [table 2](#)) assumed to be 1.

Alternatively, if data from previous years have been individually analyzed and saved as .csv files using the Evidence of Absence software, these data may be serially loaded from the .csv file by entering each file name separately. Relative weights are then entered by hand.

The new, informed prior distribution is calculated and saved by the software for use with the current year's carcass count and search parameters.

Year to year changes in operational parameters (for example, cut-in speed, acoustic deterrents, or site expansion) or population dynamics (for example, population size, migration patterns, nesting sites, etc.) may be accounted for by the relative weights, whereas changes in search protocol are accounted for in the estimates of  $g$  for the individual years.

## Multi-Year Total

The total number of fatalities at a site over the course of several years also can be estimated by the software ([appendix E](#)). Counts ( $X$ ) and overall detection probabilities ( $g$  and 95% CI) for each year are entered in a format similar to that described in the previous section. The user enters the observed count and the search parameters that were in effect in a given year and calculates the overall detection probability (by clicking **Estimate P(obs carcass|arrive)**, **Estimate Mortality for Given Parameter Set**, or **Plot Posterior distribution**). Then the user saves each year's

**Table 2.** Example format of .csv input file for several years of previous data to inform current year's prior distribution.

Credibility level (1 - $\alpha$ )					
0.85					
Yr	X	g	min(g)	max(g)	rel_wt
1	2	0.75	0.72	0.92	0.41
2	0	0.8	0.7	0.83	0.83
3	1	0.8	0.7	0.83	0.83
4	0	0.65	0.52	0.79	1
5	0	0.46	0.29	0.75	1

analysis to a separate .csv file or enters the observed count and the calculated 95% CI for each year's overall detection probability ( $P(\text{observe}|\text{arrive})$ ) into a single .csv file ([table 2](#)), with  $X$  = carcass count,  $g = P(\text{observe}|\text{arrive})$ ,  $g_{lwr}$  = lower bound on 95% CI for  $P(\text{obs}|\text{arr})$ ,  $g_{upr}$  = upper bound on 95% CI for  $P(\text{obs}|\text{arr})$ , and rel\_wt = relative number of fatalities for each year (must be positive). Then the user clicks **Estimate Total over Multiple Years** button on the **Parameters** sheet. From there, the user is offered a number of options to proceed, including editing the data manually, estimating the total number of fatalities, and returning to the main analysis page.

## Multiple Sites

The total number of fatalities summed over a number of sites may be estimated ([appendix E](#)) if the user provides the carcass count, the overall detection probability (with 95% CI), and an estimate of the relative weight for each site. Relative weights are proportional to the total number of carcasses expected at each site. If animal population sizes and activity patterns are the same at each site and if the sites all have the same number of turbines and have the same operational characteristics, then all the relative weights should be identical. If animal activity is similar among sites, then the relative weights may be taken to be a measure of site size (for example, number of turbines).

On the **Parameters** sheet, the user clicks on **Estimate Total at Multiple Sites**. The multiple sites window will initially be populated by the most recent run, but the user can enter new data by clicking on **New Data from File(s)**. There are two options for submitting input data from the individual sites:

- *Combined*.—Carcass counts, overall detection probabilities (with 95% CIs), and relative weights from the sites are copied into a single .csv file ([table 3](#)). Relative weights must sum to 1.

## 10 Evidence of Absence Software User Guide

- *Serial .csv files.*—Data are stored in a series of .csv files that have been saved using the **Save Parameters and Results to .csv** button (one file for each site) after overall detection probabilities or posterior distributions have been calculated. After clicking **Go** (on the **Estimate Total at Multiple Sites** form), the user is prompted to enter the names of the files that contain data of the individual sites. The software reads only the carcass count and detection probability (with 95% CI) from the .csv file. Other data in the files are ignored. Relative weights for the sites are entered manually after the individual sites have been uploaded.

### Restore Defaults

The parameter values shown on the **Parameters** sheet are assumed to be valid and compatible with one another. The program may not function properly if faulty parameters are somehow written to this sheet or to one of the several hidden sheets used for storing data and doing intermediate calculations. If this occurs, **Restore Defaults** enters valid parameters sets to critical locations in the workbook to restore functionality.

**Restore Defaults** is intended solely as an emergency measure to restore functionality after an unforeseen problem. In addition to replacing the parameter set on the **Parameters** sheet to workable numbers, **Restore Defaults** also erases memory of recently used parameter sets. Data entry forms will open with new default values rather than values retrieved from previous sessions.

### Entering Data for Design

The Evidence of Absence software can be used to design a monitoring protocol to determine what combinations of management parameter values (search coverage [ $a$ ], search interval [ $I$ ], and search efficiency [ $p$ ]) are required to rule out the possibility that mortality  $M$  exceeds a given threshold  $\tau$  if  $X$  carcasses are found.

The **Design Tradeoffs** button in the **Parameters** worksheet opens a form that allows users to (1) specify a threshold  $\tau$  and either single values or ranges of  $a$ ,  $I$ , and  $p$ , and (2) graph the probability that  $M > \tau$  as a function of the management parameters for the given observed carcass count  $X$  and background parameters (persistence distribution, arrival function,  $k$ , and prior distribution).

The *Threshold* ( $\tau$ ) is the maximum acceptable mortality level. The software calculates the probability that  $M > \tau$  as a function of various combinations of management parameters  $a$ ,  $I$ , and  $p$ . The calculations are specific to a given set of background parameters (persistence, arrivals,  $k$ , and prior distribution) and observed carcass count  $X$ .  $\tau$  is entered directly into the **Design Tradeoffs** form and must be a non-negative number.

**Table 3.** Example format of .csv input file for calculating total mortality at several sites.

Credibility level ( $1 - \alpha$ )					
0.85					
Site	X	g	g_lwr	g_upr	rel_wt
Wind Rocks	1	0.31	0.17	0.52	0.14
Buffalo Springfield	0	0.53	0.4	0.68	0.2
Power Central	0	0.62	0.41	0.83	0.6
Enormity Ridge	0	0.52	0.15	0.8	0.06

### Background Parameters

Whereas  $a$ ,  $I$ , and  $p$  values are transparently amenable to management decisions, persistence distribution, arrival function,  $k$ , and prior distribution are more determined by the ecology. The last saved values of these parameters are loaded into the form by default, but values can be edited using the tools on the right side of the **Design Tradeoffs** form.

### Management Parameters

The user has the option to enter fixed values or ranges of the three management parameters  $a$ ,  $I$ , and  $p$ . Parameters  $a$  and  $p$  must be positive and less than or equal to 1, and  $I$  must be positive and (approximately) divide the *span*.

### Graphing Options

Most of the graphs depend on fairly intensive calculations. The finer resolution graphs look prettier but take more time to draw than the coarser resolution graphs. The user also has the option of showing either the full parameter set or a partial set on each graph. Note that each time a graph is created, the previous graph is overwritten. To save a particular graph, the user must save before creating a new graph.

### R graphs

At every stage of data entry and analysis, there are options for generating graphs in R. These graphs can be saved in a number of different ways:

- on the **File** menu, click **Save as...**, then in the graphing window, select a format, and type a file name,
- on the **File** menu, click **Copy to clipboard ► as a bitmap or as a metafile** (Metafiles are efficient, clear, scalable, and can be pasted seamlessly into Microsoft Word or Powerpoint, but metafiles may not be as easily compatible with other programs.), or
- right-click the graph and copy to clipboard or save to a file.

The R graphing windows remain open until they are closed by the user or the workbook is closed. If a new graph is created when a graphing window is already open, the first graph is overwritten by the new graph.

## Analysis and Interpretation: A Tutorial

**Problem #1:** Searcher efficiency and carcass persistence trials have already been done. Searches have been conducted and no carcasses have been found. Can we rule out the possibility that the take exceeded some predefined threshold? What is the smallest maximum number of fatalities that we can credibly rule out?

The questions are addressed by evaluating probabilities from the posterior distribution derived from the observed carcass count ( $X$ ), the search parameters ( $a$ ,  $I$ , and  $p$ ), and the background parameters (persistence distribution, arrival function,  $k$ , and prior distribution). The process is: (1) enter parameters using the **Edit Parameters** form or the **Read Dates and Parameters from .csv** button, (2) run the analysis using the **Plot Posterior Distribution** button or the **Estimate Mortality for Given Parameter Set** button, and (3) interpret the results.

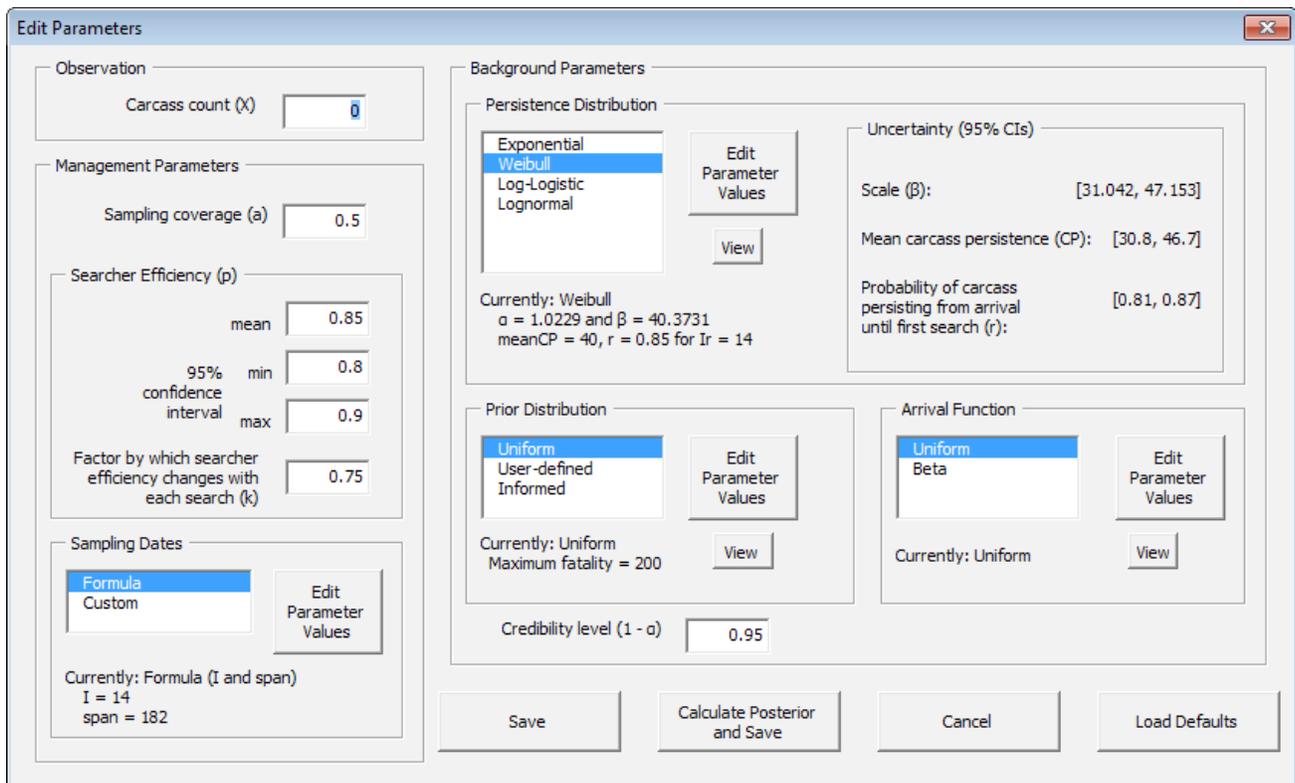
**Problem #2:** Searches have not been conducted yet, but we'd like to design a search protocol so that we can conclude with sufficient credibility that the number of fatalities did not exceed a given threshold if the total observed carcass count at the end of the season was zero (or few). What search interval, coverage, and searcher efficiency will we need?

The tutorial discusses some approaches to using the software to address these questions for several hypothetical scenarios.

### Scenario 1: Analyzing Search Data (Eagles)

We have completed a season of searches for eagles at a wind farm in eagle range and habitat, and we found zero carcasses. What is the maximum number of fatalities that could realistically have occurred given the strength of our search protocol? Or, what is the smallest number of fatalities that we can credibly rule out as having occurred?

**Step 1:** Click the **Edit Parameters** button and enter carcass count ( $X = 0$ ), management variables, prior distribution, and background parameters (fig. 2).



**Figure 2.** Screen capture of the Evidence of Absence software **Edit Parameters** form for entering search parameters.

## Management Variables

*Coverage* ( $\alpha$ ) = 0.5.—Half the turbines are thoroughly searched, that is, there is no unsearched ground around a searched turbine, and the search radius is large enough so there is little chance a carcass falls outside the searched area.

*Searcher Efficiency* ( $p$ ) = 0.85.—Large carcasses, thin vegetation cover, high probability of detection. Although not known precisely, field trials indicate that  $p$  is somewhere between 0.8 and 0.9 (95% CI).

$k = 0.75$ —This assumes that if a carcass is missed on a given search, it is only 75 percent as likely to be seen in the next search. (Aging carcasses are more difficult to spot, and the carcass might have been missed the first time because it was somehow hidden.)

*Search Interval* ( $I$ ) and *span*.—Select **Formula** and enter  $I = 14$ ,  $span = 182$  for searches once every two weeks for one-half year.

*Credibility Level* ( $1 - \alpha$ ).—Desired credibility level for assertions that the number of fatalities did not exceed a given threshold. A higher credibility level makes it more difficult to rule out high numbers of fatalities. For this example, we use a high value of  $1 - \alpha = 0.95$ .

## Background Parameters

*Persistence Distribution*: Based on carcass persistence trials performed throughout the monitoring season, we assume a Weibull persistence distribution with  $meanCP = 40$  days and average probability that a carcass persists until a search is  $r = 0.85$  for a search interval of 14 days. The 95% CI interval on  $r$  is [0.81, 0.87]. To enter these parameters, click on **Weibull** persistence distribution and then click the **Edit Parameters** button. You will be given the option of entering either shape and scale parameters or meanCP and  $r$ . Click on **Mean CP and r** to bring up a form for entering the interval associated with  $r$ . Enter 14 and click **OK** to bring up the form for entering the parameters (fig. 3).

After entering the persistence distribution parameters, enter the remaining background parameters.

*Prior Distribution for Number of Fatalities*: Assume uniform—that is, we do not make assumptions ahead of time about the number of fatalities; there could have been 0 or 200 or anything in between.

*Arrival Function*: Assume uniform—that is, Golden Eagles are resident year-round so there is not great seasonal variation in the rate of eagle kills during the sampling period. Note: If the

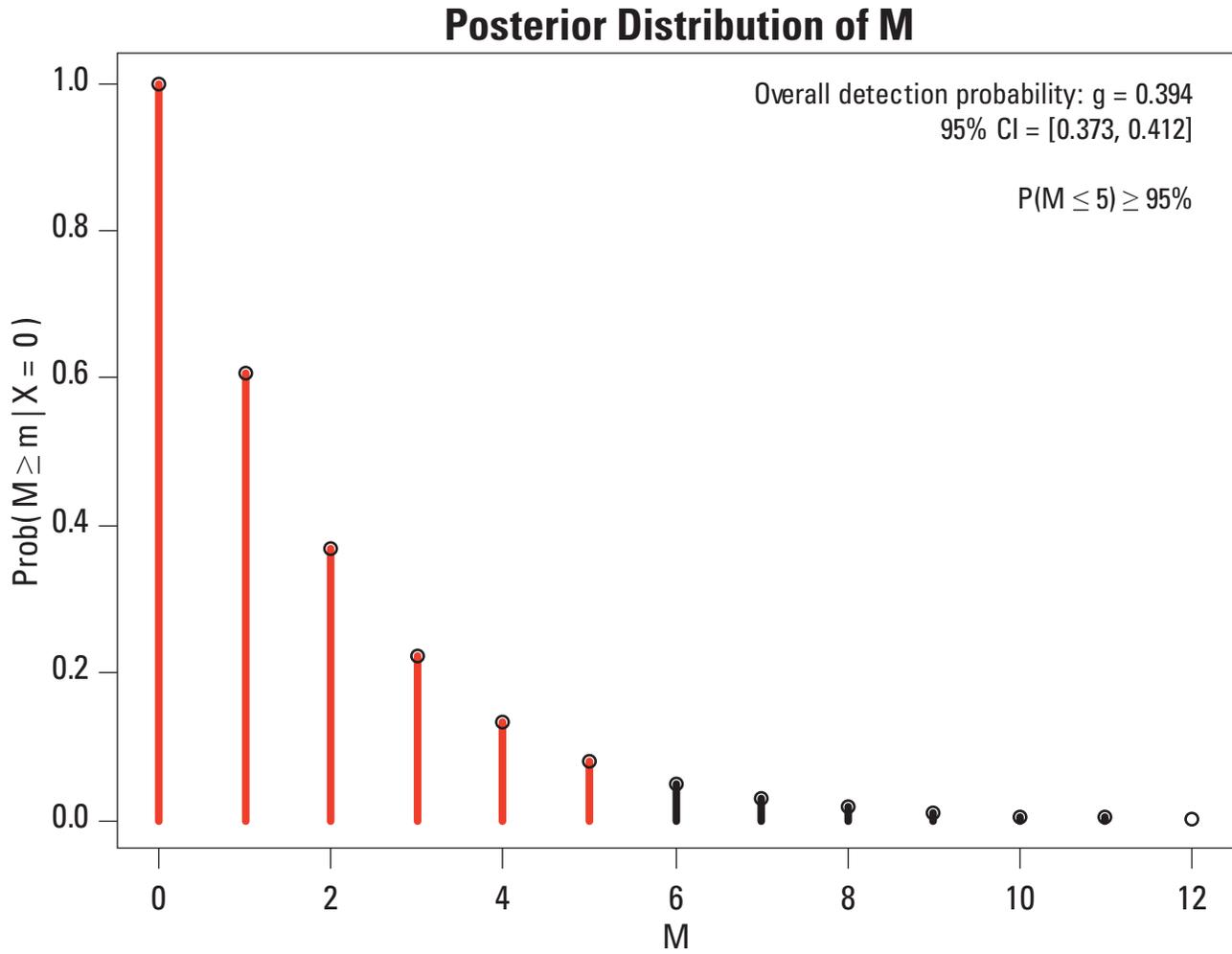
**Figure 3.** Screen capture of the Evidence of Absence software **Weibull Parameters** form for entering persistence distribution parameters.

search interval is regular, the relative arrival rate assumption has little effect. Choice of relative arrival function only has an effect if the search schedule is adjusted to match seasonal variation in the at-risk population or if most carcasses arrive in a pulse near the end of the monitoring season.

Click **Save** on the **Edit Parameters** form.

**Step 2:** Click **Plot Posterior** on the **Parameters** sheet to produce the graph of the Posterior Distribution of  $M$  (fig. 4). Even though 0 carcasses were observed, we cannot conclude that zero eagles were killed—we could have missed some carcasses in the searches. However, according to the information we have about the monitoring protocol and the ecology, we can conclude that it is highly unlikely that more than five were killed. If more than five had been killed, there is little chance (less than 1 in 20) we would have missed them all in our searches.

The bars in figure 4 represent the probability that the number of fatalities exceeded the given  $m$  value on the x-axis. The red bars give the 95 percent credible interval for  $M$ , that is, we can assert with 95 percent credibility that the number of kills did not exceed five given the effectiveness of our search protocol, our prior knowledge (or lack thereof) about how many might be killed (uniform prior), and the fact that we found zero carcasses.



Carcass count:  $X = 0$   
 Coverage:  $a = 0.5$   
 Sampling:  $l = 14$ , span = 182  
 Search efficiency:  $p = 0.85$ , 95% CI = [0.8, 0.9]  
 $k = 0.75$

Persistence: Weibull,  $\alpha = 1.023$ ,  $\beta = 40.373$   
 CP: mean = 40, 95% CI = [30.8, 46.7]  
 r: mean = 0.85, 95% CI = [0.81, 0.87]  
 Arrivals: Uniform  
 Prior: Uniform, mmax = 200

**Figure 4.** Graph generated by Evidence of Absence software showing Posterior Distribution of the eagles example for scenario 1, analyzing search data.

## 14 Evidence of Absence Software User Guide

Graphs can be saved by right-clicking on them and either copying them to the clipboard or to a file. Two formats for saving to clipboard are bitmap and metafile. Metafiles (.emf) are scalable and designed for inserting into Microsoft® Office documents. The .emf files tend to be higher quality and not as large as bitmaps. (The .emf files are sharp and clear in Word® or Powerpoint® but may look rough, pixelated, or blurry in other programs.)

The full posterior distribution is reproduced in a table (table 4) on the **Parameters** sheet that appears next to the parameter set associated with the posterior. Simple summary statistics also are given.

Note: The effects of uncertainty in parameter estimates for searcher efficiency and carcass persistence are accounted for through simulation, so output may vary slightly from what is shown.

$P(\text{observe}|\text{arrive})$  is the overall probability that a fallen carcass is actually observed. The **credible maximum** is the upper bound of the  $100(1-\alpha)\%$  credibility interval. Thus, from table 2, we say, “We can assert with at least 95 percent credibility that there were not more than 5 fatalities.”

The parameter set and results (the posterior distribution) are temporarily stored in the **Parameters** sheet. They can be permanently stored in .csv files by clicking **Save Parameters and Results to .csv**.

**Table 4.** Full posterior distribution results from Parameters sheet for the eagles example in scenario 1, analyzing search data.

Results				
Posterior distribution				
$g=P(\text{observe} \text{arrive})$	0.394	95% CI:	0.373	0.411
95% credible maximum	5.			
m	$P(M = m)$	$P(M > m)$		
0	0.3937	0.6063		
1	0.2388	0.3675		
2	0.1448	0.2227		
3	0.0879	0.1348		
4	0.0533	0.0815		
5	0.0324	0.0491		
6	0.0197	0.0295		
7	0.0119	0.0176		
8	0.0073	0.0103		
9	0.0044	0.0059		
10	0.0027	0.0032		
11	0.0016	0.0016		
12	0.001	0.0006		
13	0.0006	0.		

### Scenario 2: Optimizing the Search Protocol (Eagles)

We are designing a monitoring plan for eagle carcasses for the coming field season. What search protocol will we need to be able to make the case that, if no carcasses are found, no more than three eagles were killed during the season? We are satisfied with our searcher efficiency of 0.85 and have little means of improving it. We focus on tradeoffs in the searched area and the search interval, knowing that coverage does not change linearly with searched area. We are interested in a credibility level of 0.95.

Optimize the management variables by clicking **Design Tradeoffs** on the **Parameters** sheet. Enter **Threshold** ( $\tau$ ) = 3, and **Credibility level** = 0.9. The Searcher efficiency ( $p$ ) will be held fixed by clicking on Fixed and entering  $p$ : of 0.85. Select *Variable* for coverage ( $a$ ) and search interval ( $I$ ). Allow  $a$  to range from 0.25 to 1 and search interval to range from a minimum of 1 day and maximum of 30 over a 182 day span (fig. 5).

Enter the same ecological parameters as with the previous example, namely, Weibull persistence distribution with meanCP = 40 days and probability that a carcass persists until a search is  $r = 0.85$  for a search interval of 14 days, uniform arrivals, uniform prior, and  $k = 0.4$ .

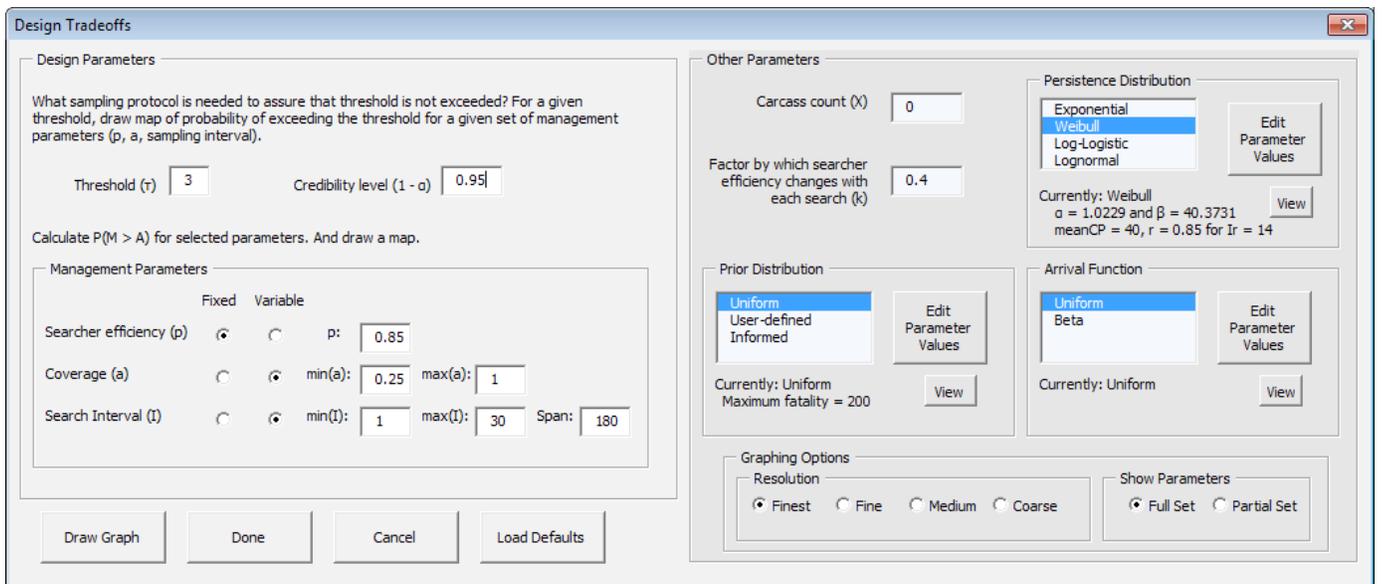


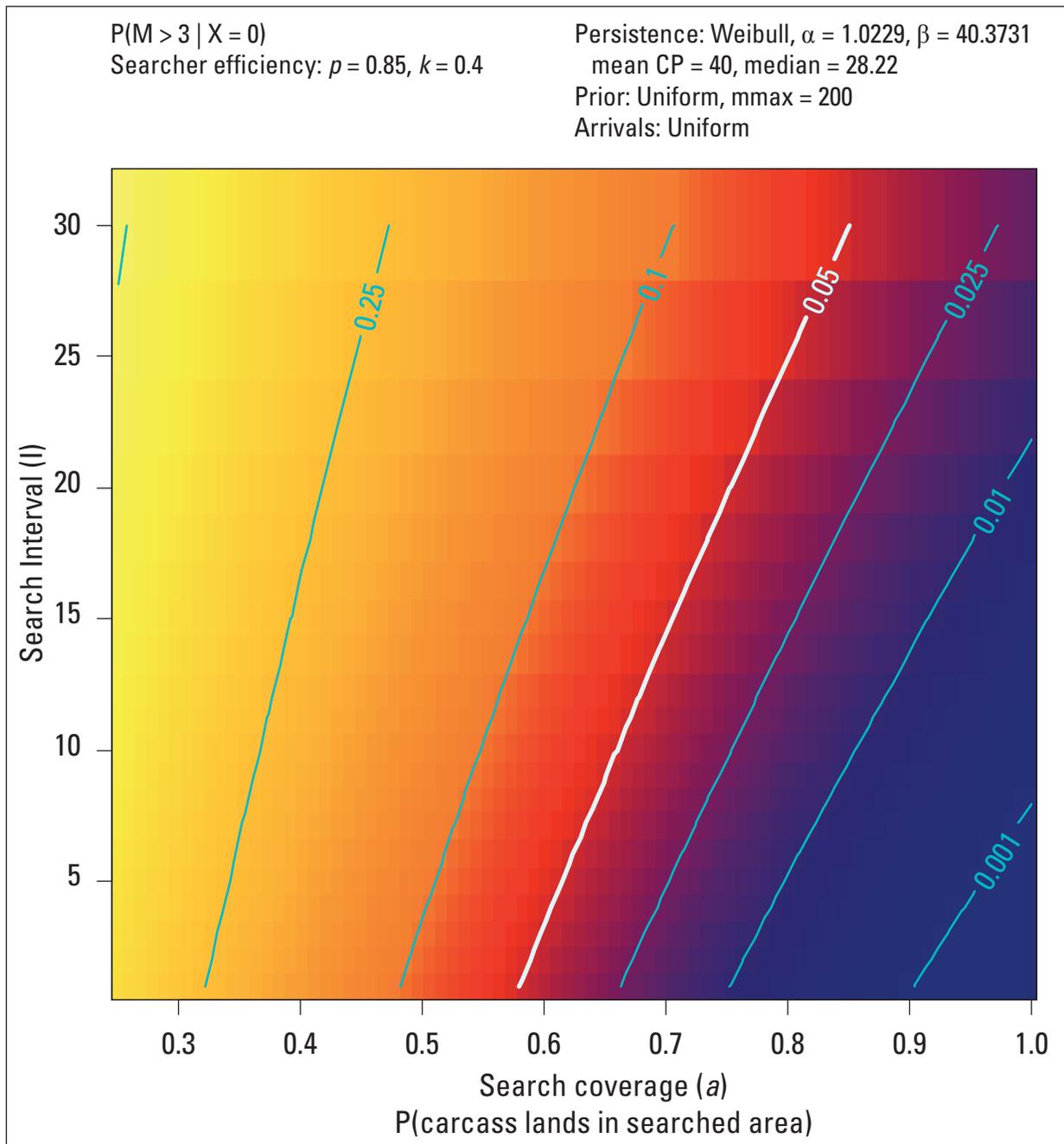
Figure 5. Screen capture of Design Tradeoffs form for eagles example for scenario 2, optimizing the search protocol.

Click **Draw Graph** to bring up the designs tradeoff graph (fig. 6).

The critical combination of search interval and coverage is indicated by the bold white line in the graph. For example, to assert with 95 percent credibility that no more than three eagles were killed, we would need  $P(\text{fatality} > 3 \mid X = 0) \leq 0.05$ , which corresponds to the purple and blue region of the graph. That could be accomplished by searching about 60 percent of the turbines every day ( $a = 0.58, I = 1$ ), searching about 65 percent of the turbines once per week, 70 percent every other week, or 80 percent once per month.

Notice that the contours lines on figure 6 are nearly vertical, and that there is little change in color with vertical movement on the graph. This indicates that changing the search interval without changing the search coverage would decrease the credible level only slightly. This is because for this example the carcass persistence time is long compared with the search intervals that are considered.

By contrast, colors change dramatically moving horizontally across the graph, indicating that search coverage has a marked effect in this scenario. A small increase in coverage would allow for a fairly large increase in search interval with no loss

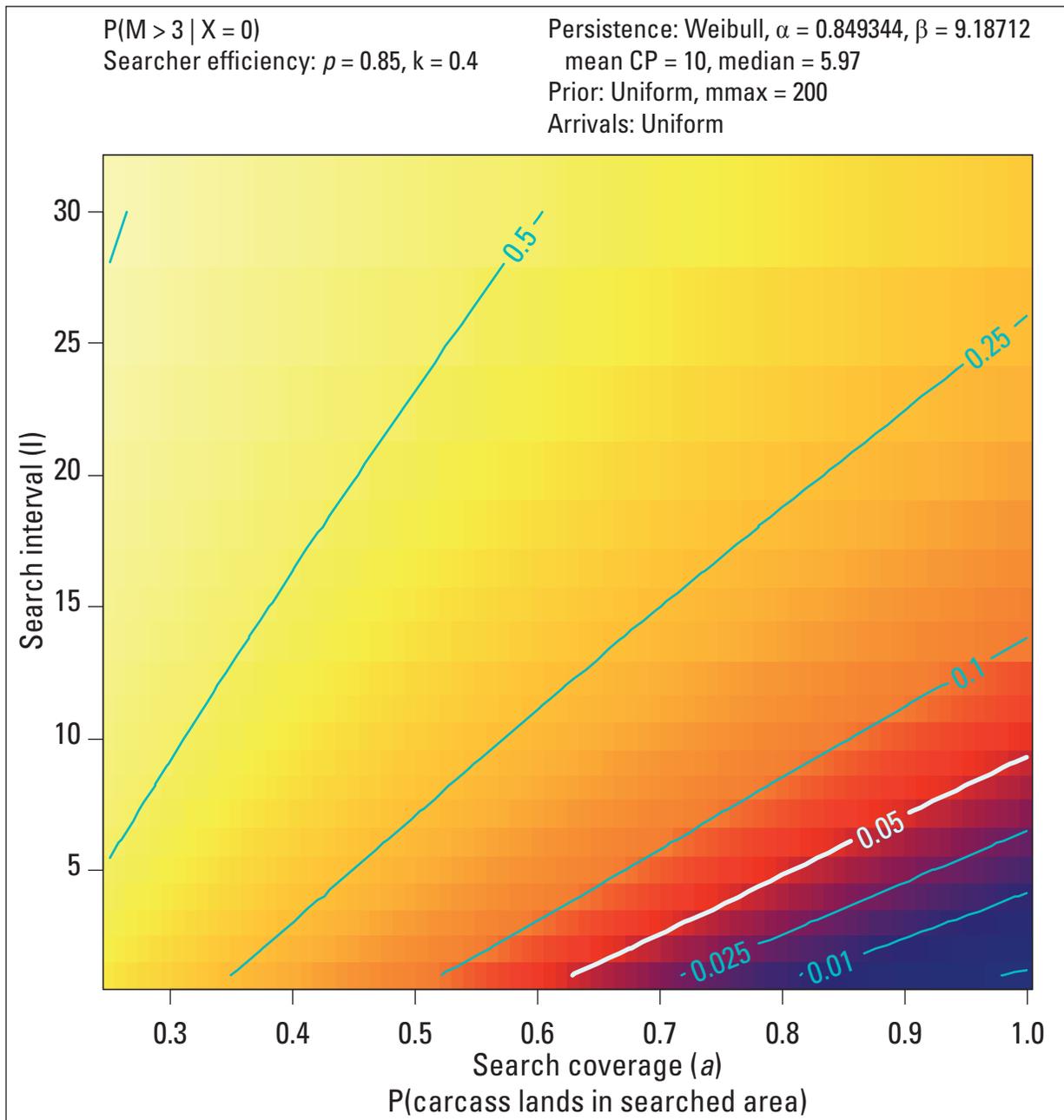


**Figure 6.** Graph generated by Evidence of Absence software showing design tradeoffs for search interval and search coverage for long carcass persistence time.

in credibility. The monetary costs of searching a larger area less frequently can be assessed and the optimal protocol from a budgetary perspective can be determined.

If the mean carcass persistence had been  $CP = 10$  days with  $r = 0.5$  for an interval of 14 days instead of  $CP = 40$  and  $r = 0.85$ , the scenario is quite different. You can change the persistence distribution from the **Design Tradeoffs** form by clicking **Edit Parameter Values** in the upper right-hand corner and editing using the form that pops up. Make the change, click **Save**, and then **Draw Graph** (fig. 7). Under these conditions where the persistence times are shorter,

the search interval has a much larger effect. If searches are done only once per month, we can never be sure that fatality did not exceed three because even with 100 percent search coverage, a large fraction of carcasses would be expected to have disappeared by the time of the search. Thus, a shorter search interval is required. For example, we could search 100 percent of the turbines every week, 80 percent of the turbines every five days, or 65 percent of the turbines every day to attain the same credibility level. Managers can choose the least expensive protocol that achieves the preferred level of credibility.



**Figure 7.** Graph generated by Evidence of Absence software showing design tradeoffs for search interval and search coverage for short carcass persistence time.

### Scenario 3: Indiana bat

The Indiana bat is listed as “endangered” under the Endangered Species Act. The bat is a small animal (5–10 grams) and lives in forested areas and bottomlands, so searcher efficiency tends to be low (on the order of 5–25 percent). The ground around the turbines in the bat habitat may not be fully searchable owing to vegetation, rough ground, or inaccessibility, so sampling coverage is frequently far short of 100 percent even when all turbines are searched.

### Post-Search Estimation: Data Entry and Calculations

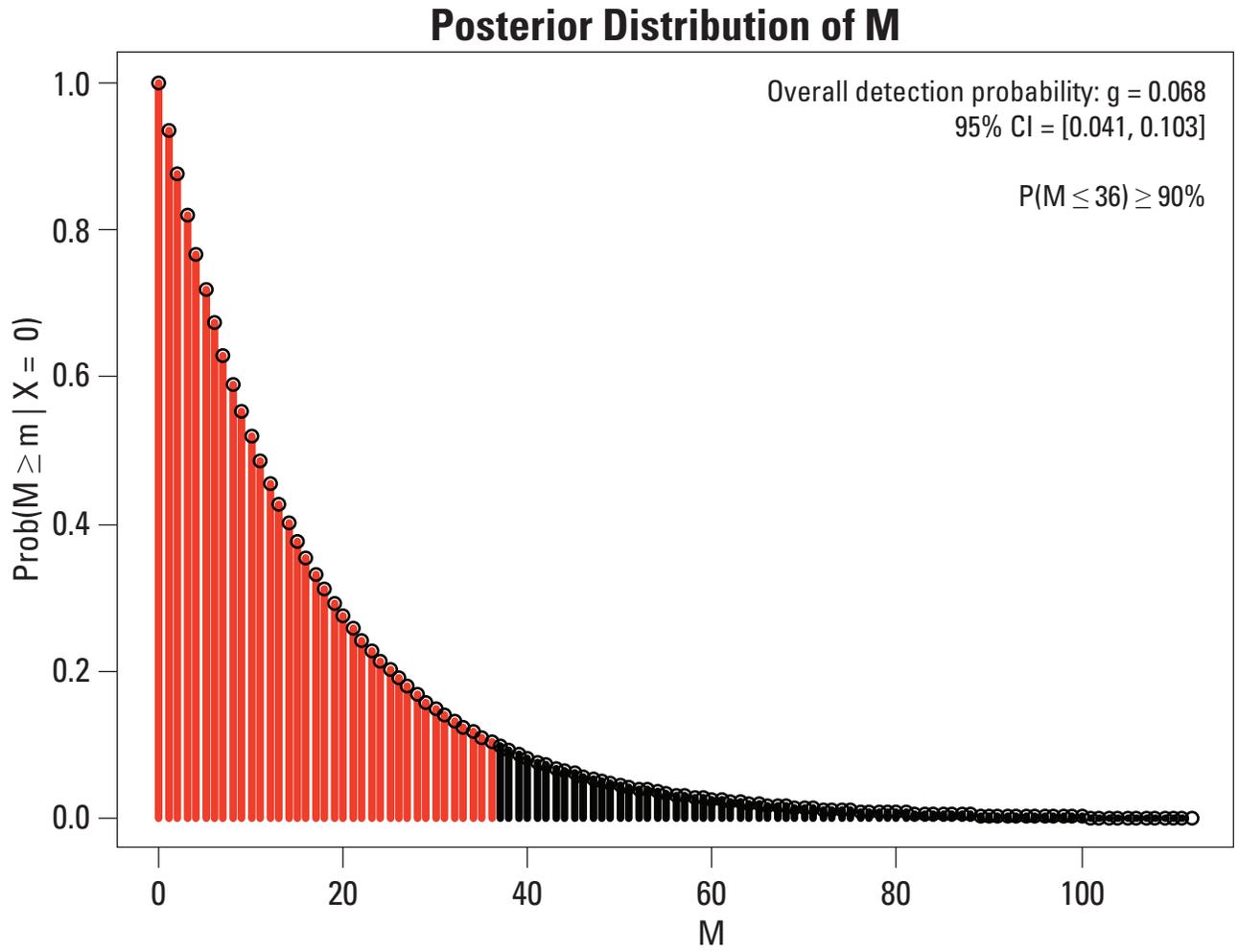
Suppose we sampled once per week for one season (half year) and found no carcasses. How many fatalities might there have been? Assume that searcher efficiency and carcass persistence trials have been conducted throughout the season and that the coverage ( $a$ ) is estimated from DWP modeling (Huso and Dalthorp, 2014) of the search areas and that we require 90 percent credibility. Enter parameters by clicking the **Edit Parameters** button. Alternatively, the user may load the parameter set by clicking the **Read Dates and Parameters from .csv** button on the **Parameters** sheet. When prompted to enter file name, navigate to the `csv files` folder and select `scenario3 input.csv` (which has the input parameters pre-entered for scenario 3) ([fig. 8](#)).

The parameter values for scenario 3 include:

- *Carcass count* ( $X$ ) = 0
- *Credibility level* ( $1-\alpha$ ) = 0.9
- *coverage* ( $a$ ) = 60 percent
- *searcher efficiency* ( $p$ ) = 15 percent, with 95% CI = [0.1, 0.2]
- *Sampling dates*: Formula with  $I = 7$ ,  $span = 182$  days
- *Carcass persistence*: Assume Weibull with mean CP = 6.5 days with  $r_7 = 0.58$  for a search interval of 7 days (95% CI = [0.47, 0.70]).
- *Prior distribution*: Assume uniform (we do not make assumptions about how many might be killed; there could have been 0 or 200 or anything in between).
- *Arrival function*: Assume constant (uniform). Bat kills tend to be more common in late summer, but arrival rate function does not have a significant effect unless we have a search schedule that changes by season.
- $k = 0.75$ : If a carcass is missed on a given search, assume it is 75 percent as likely to be seen in the next search (if it is still there). The carcasses are hard to find in the first place, so it is no surprise to miss them the first search ( $p = 0.15$ ), and, in a second pass through the field, they may be only somewhat less likely ( $0.75 \times 0.15$ ) to be found.

**Figure 8.** Screen capture of the **Edit Parameters** form showing initial search parameters for the Indiana bat example for scenario 3.

Save the parameter set and click **Plot Posterior Distribution** ([fig. 9](#)).



Carcass count:  $X = 0$   
 Coverage:  $a = 0.6$   
 Sampling:  $l = 7$ , span = 182  
 Search efficiency:  $p = 0.15$ , 95% CI = [0.1, 0.2]  
 $k = 0.75$

Persistence: Weibull,  $\alpha = 0.883$ ,  $\beta = 6.111$   
 CP: mean = 6.5, 95% CI = [4.3, 11]  
 $r$ : mean = 0.578, 95% CI = [0.47, 0.7]  
 Arrivals: Uniform  
 Prior: Uniform, mmax = 200

**Figure 9.** Graph generated by Evidence of Absence software showing the posterior distribution of M for the Indiana bat example for scenario 3 using uniform prior distribution.

## Post-Search Estimation: Interpretation

No carcasses were found, but the analysis shows as many as 36 may have been killed (90 percent credibility that  $M \leq 36$ ). There are two reasons for this:

1. Inadequate search protocol for our needs: A persistence parameter  $r = 0.578$  means that nearly half the carcasses that do arrive in the area are scavenged before a search is conducted. Also, searcher efficiency is low (0.15) and much of the area could not be searched (coverage = 0.6). With those parameters, the overall probability of observing a carcass that has arrived is only about 7 percent or 1 in 15. With such low detection probability, it is easy to miss many carcasses.
2. No prior information: This was our first season of monitoring. We could not justify a restrictive prior, so we use a uniform prior to give full voice to the data. With reliable prior information to rule out high kill rates, we could improve our statistical power.

To help manage the difficult ecological constraints (such as, small animal and difficult ground to search), the software provides tools for improving estimates by:

- incorporating previous years' data into current estimates, and
- designing a search protocol to find the optimum balance between statistical power and cost.

### Informed Prior

Use the previous year's search data to inform the current year's prior distribution and improve estimation.

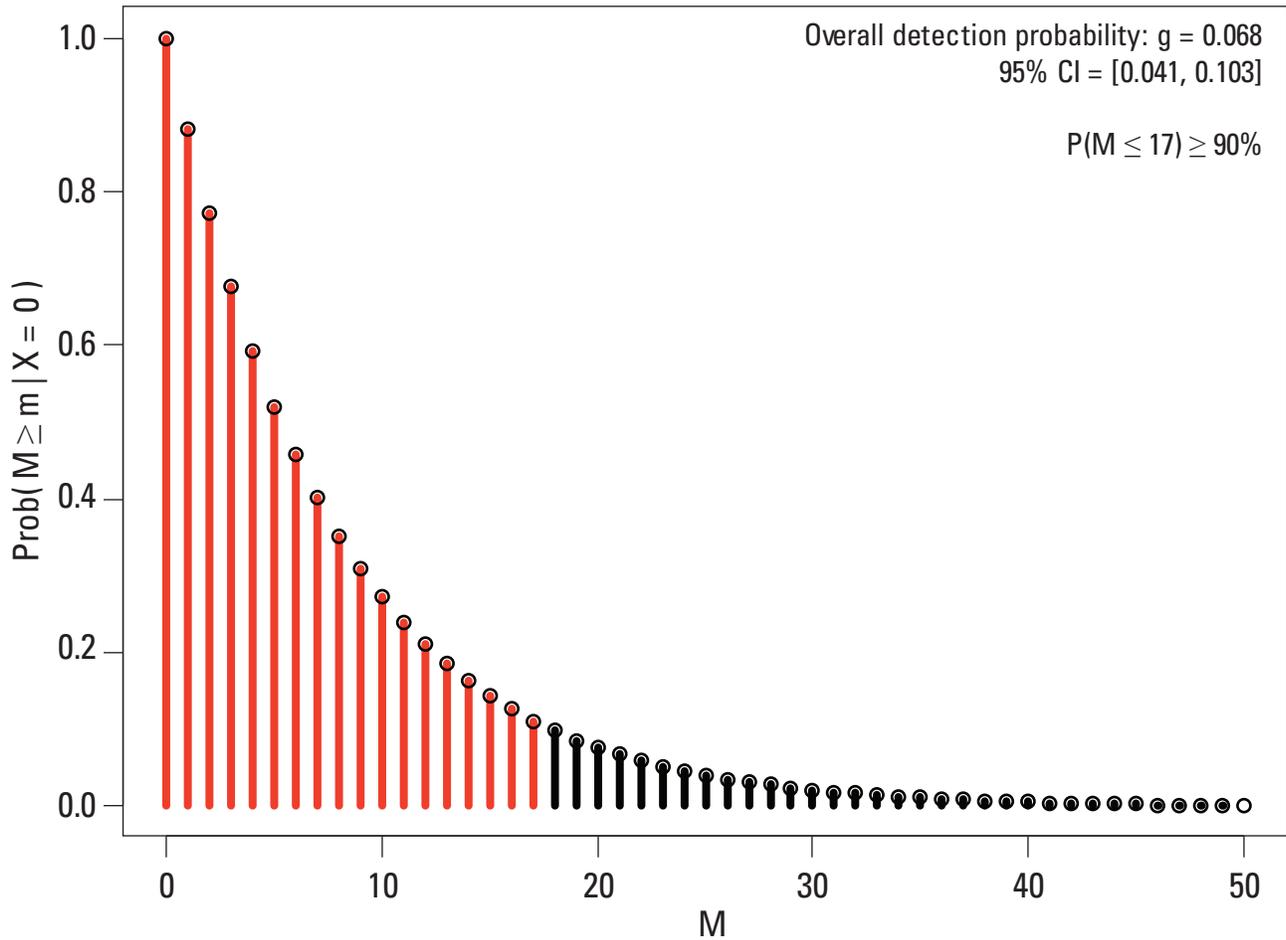
Previous year: Suppose we searched once per week for Indiana bat carcasses assuming the same management variables, prior distribution, and background parameters as in the previous example. Click **Edit Parameters** on the **Parameters** worksheet and verify that the active parameter set is correct ( $X = 0$ ,  $a = 0.6$ ,  $p = 0.15$  with CI = [0.1, 0.2],  $I = 7$ ,  $span = 182$ ,  $k = 0.75$ ; Weibull persistence with mean CP = 6.5 and  $r = 0.578$  with CI = [0.47, 0.7] for an interval of  $I_r = 7$ , uniform prior; uniform arrivals). Click **Save** to save the year 1 parameters to the **Parameters** sheet.

Current year: We need to extract the carcass count,  $P(\text{observe}|\text{arrive})$ , and bounds for the 95% CI for  $P(\text{observe}|\text{arrive})$  from analysis of the year 1 data. If the latter quantities do not appear on the right side of the **Parameters** sheet in the "Results" section, click **Estimate P(observe|arrive)**, **Estimate Mortality for Given Parameter Set**, or **Plot Posterior Distribution** to perform the analysis and list the required values. To load the year 1 results to inform the year 2 prior distribution, open the **Edit Parameters** form and double-click the "Informed" option (or single-click "Informed" and then "Edit Parameter Values") in the "Prior Distribution" section. Click "One previous year's data". The dialog box grabs values from the active **Parameters** sheet, allowing the user to **OK** the values, manually edit the values, or to retrieve previously saved values use **Read from .csv**. When the proper values are showing ( $X = 0$ ,  $P(\text{obs} | \text{arrive}) = 0.0679$ , and 95% CI = [0.039, 0.102]), click **OK** to return to the **Edit Parameters** form with the informed prior distribution selected.

Enter the second year's carcass count (again, 0 for this example), make any necessary changes to parameter values to reflect changes in search protocol (none are needed for this example because we are assuming the same search protocol and background parameters for both years), and click **Save**. Then, on the **Parameters** sheet, click **Plot Posterior Distribution** to show the Posterior Distribution of  $M$  graph ([fig. 10](#)):

There is some improvement in the estimate, which can be seen by comparing the current posterior distribution ([figs. 9 and 10](#)) with the previous ([figs. 8 and 9](#)). However, we are still short of the goal of arguing with 90 percent credibility that no more than five bats were taken. According to the summary of the distribution shown on the **Parameters** sheet, we can assert with over 90 percent credibility that  $M$  did not exceed 18 ( $P(M > 18) = 0.09$ , which can be seen in the last red bar in [figure 10](#) or read more precisely from the results section on the **Parameters** sheet). NOTE: The effects of uncertainty are accounted for through simulation, so values for the mean and standard deviation of the informed prior may vary slightly from those shown in [figure 10](#).

### Posterior Distribution of M



Carcass count:  $X = 0$   
 Coverage:  $a = 0.6$   
 Sampling:  $l = 7$ , span = 182  
 Search efficiency:  $p = 0.15$ , 95% CI = [0.1, 0.2]  
 $k = 0.75$

Persistence: Weibull,  $\alpha = 0.883$ ,  $\beta = 6.111$   
 CP: mean = 6.5, 95% CI = [4.3, 11]  
 $r$ : mean = 0.578, 95% CI = [0.47, 0.7]  
 Arrivals: Uniform  
 Prior: Informed,  $\mu = 15.224$ ,  $\sigma = 14.823$

**Figure 10.** Graph generated by Evidence of Absence software showing posterior distribution of M for the Indiana bat example for scenario 3 using informed prior distribution.

## Optimizing the Search Protocol

We can improve estimates by taking advantage of the software's design tools to optimize the search protocol prior to sampling.

1. Click **Design Tradeoffs** on the **Parameters** sheet. Within the form, enter the desired threshold, select the **Variable** option for all three management parameters, and enter a range of values to consider for each of the management parameters as shown in [figure 11](#).
2. Define the same background parameter values as in the previous example (incorporating an informed prior distribution into the estimation of Indiana bat fatalities with  $X = 0$ ,  $k = 0.75$ , etc.) using the tools on the right-hand side of the **Design Tradeoffs** form ([fig. 11](#)). Be sure that the **Prior Distribution** is **Uniform** with a maximum of 200 before clicking **Save**.

3. Click **Draw Graph** on the **Design Tradeoffs** form to create the following graphs ([fig. 12](#)).

Each graph in [figure 12](#) shows the posterior probability that the number of fatalities exceeded the threshold as a function of searcher efficiency ( $p$ ) and sampling coverage ( $a$ ) for the given search interval. The ranges of efficiency and coverage are those specified by the user in the form. The five search intervals represented are the minimum and maximum intervals specified by the user and three intermediate values. In our original search protocol, we had search interval of 7 days, 15 percent searcher efficiency, and had 60 percent sampling coverage, which put us in the yellow region just to the right of the 0.75 line in the figure for Search interval ( $I$ ) = 4 ([fig. 12](#), upper right). To get to 90 percent credibility level (red) that five or fewer bats were killed seems nearly impossible—we'd need to search every day ( $I = 1$ ) with 20 percent searcher efficiency and 80 percent coverage or with 35 percent efficiency and 60 percent coverage.

**Design Tradeoffs**

Design Parameters

What sampling protocol is needed to assure that threshold is not exceeded? For a given threshold, draw map of probability of exceeding the threshold for a given set of management parameters ( $p$ ,  $A$ , sampling interval).

Threshold ( $\tau$ )  Credibility level ( $1 - \alpha$ )

Calculate  $P(M > A)$  for selected parameters. And draw a map.

Management Parameters

	Fixed	Variable		
Searcher efficiency ( $p$ )	<input type="radio"/>	<input checked="" type="radio"/>	min( $p$ ): <input type="text" value="0.05"/>	max( $p$ ): <input type="text" value=".35"/>
Coverage ( $A$ )	<input type="radio"/>	<input checked="" type="radio"/>	min( $A$ ): <input type="text" value=".25"/>	max( $A$ ): <input type="text" value=".8"/>
Search Interval ( $I$ )	<input type="radio"/>	<input checked="" type="radio"/>	min( $I$ ): <input type="text" value="1"/>	max( $I$ ): <input type="text" value="14"/> Span: <input type="text" value="182"/>

Other Parameters

Carcass count ( $X$ )

Factor by which searcher efficiency changes with each search ( $k$ )

Persistence Distribution

Exponential  
 Weibull  
 Log-Logistic  
 Lognormal

Currently: Weibull  
 $\alpha = 0.8828$  and  $\beta = 6.1112$   
 meanCP = 6.5,  $r = 0.578$  for  $I_r = 7$

Prior Distribution

Uniform  
 User-defined  
 Informed

Currently: Uniform  
 Maximum fatality = 200

Arrival Function

Uniform  
 Beta

Currently: Uniform

Graphing Options

Resolution

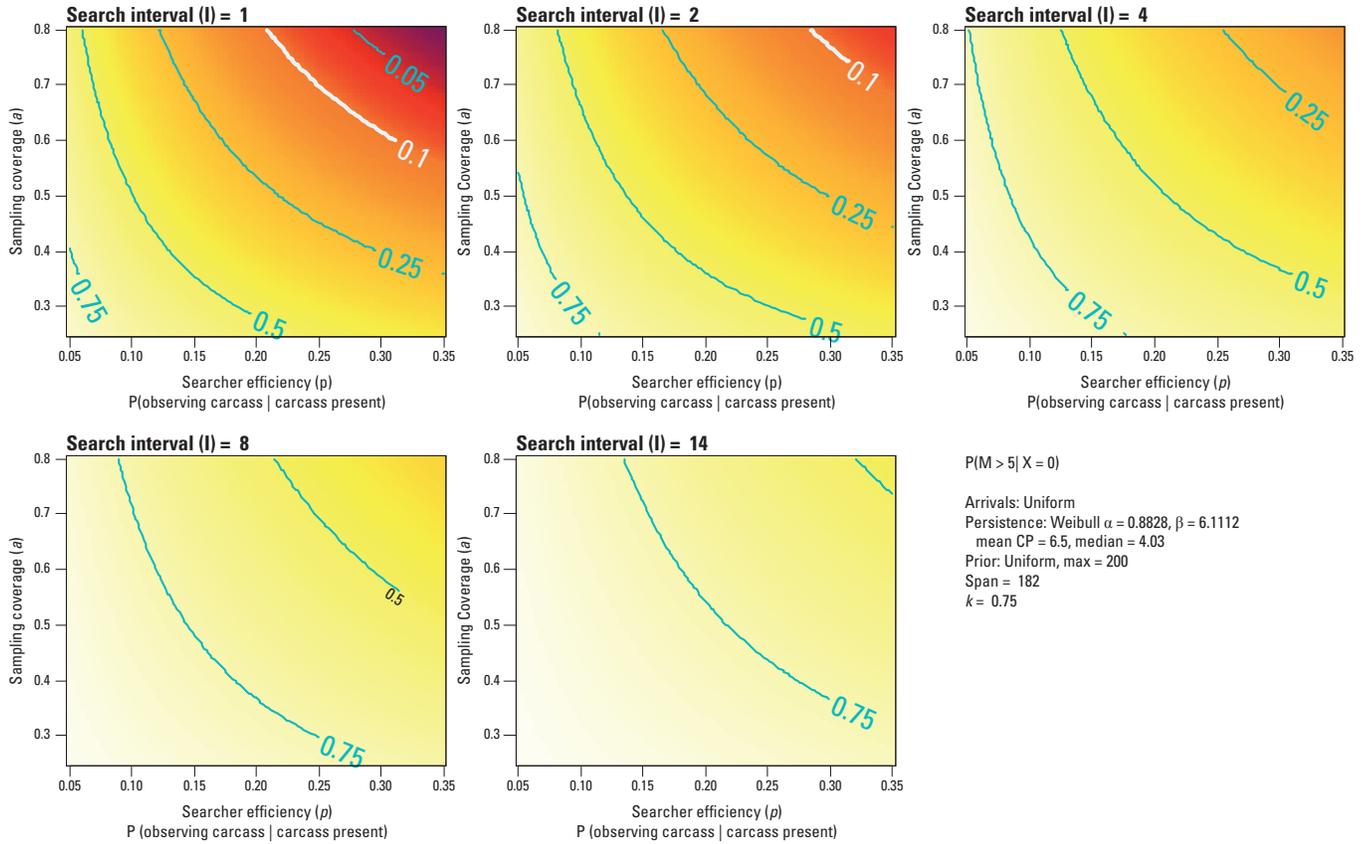
Finest  Fine  Medium  Coarse

Show Parameters

Full Set  Partial Set

Buttons: Draw Graph, Done, Cancel, Load Defaults

**Figure 11.** Screen capture showing the **Design Tradeoffs** form for the Indiana bat example.

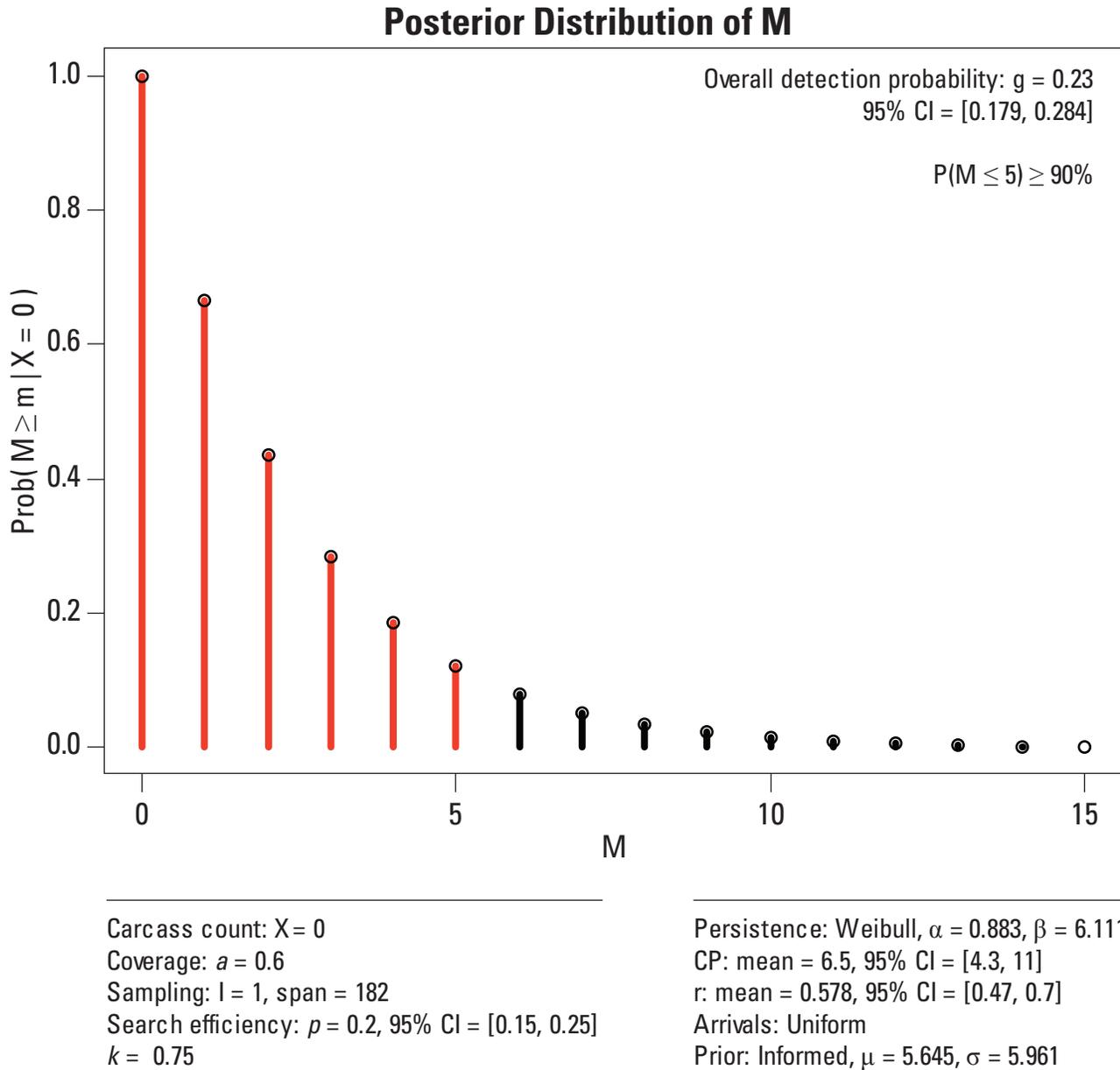


**Figure 12.** Graphs generated by Evidence of Absence software showing design tradeoffs between search coverage, searcher efficiency, and search interval for the Indiana bat example for scenario 3.

However, if we combine an intensified search (daily searches, 60 percent coverage, 20 percent searcher efficiency) with using an informed prior distribution, we do have a strong enough search protocol to conclude that five or fewer bats were killed (fig. 13).

To generate the figure, first re-run the year 1 analysis using the intensified search protocol ( $p = 0.2$  with 95% CI = [0.15, 0.25],  $I = 1$ ) with a uniform prior: **Edit Parameters** to change the parameter values, **Save**, and **Estimate Mortality for**

**Given Parameter Set.** Then **Edit Parameters** to change the year 2 prior to “Informed”. After clicking “Edit Parameter Values” for the prior distribution and “One previous year’s data”, the summary of the previous year’s search data should load into the form as the default values, i.e.  $X = 0$ ,  $P(\text{obs}|\text{arr}) = 0.231$  with 95% CI = [0.176, 0.282]. Click **OK** and **Save**. Then, **Plot Posterior Distribution** from the **Parameters** worksheet. NOTE: Calculation of  $g = P(\text{observe}|\text{arrive})$  relies on simulation, so your  $P(\text{obs}|\text{arr})$  values and associated CI may vary slightly from what is shown above.

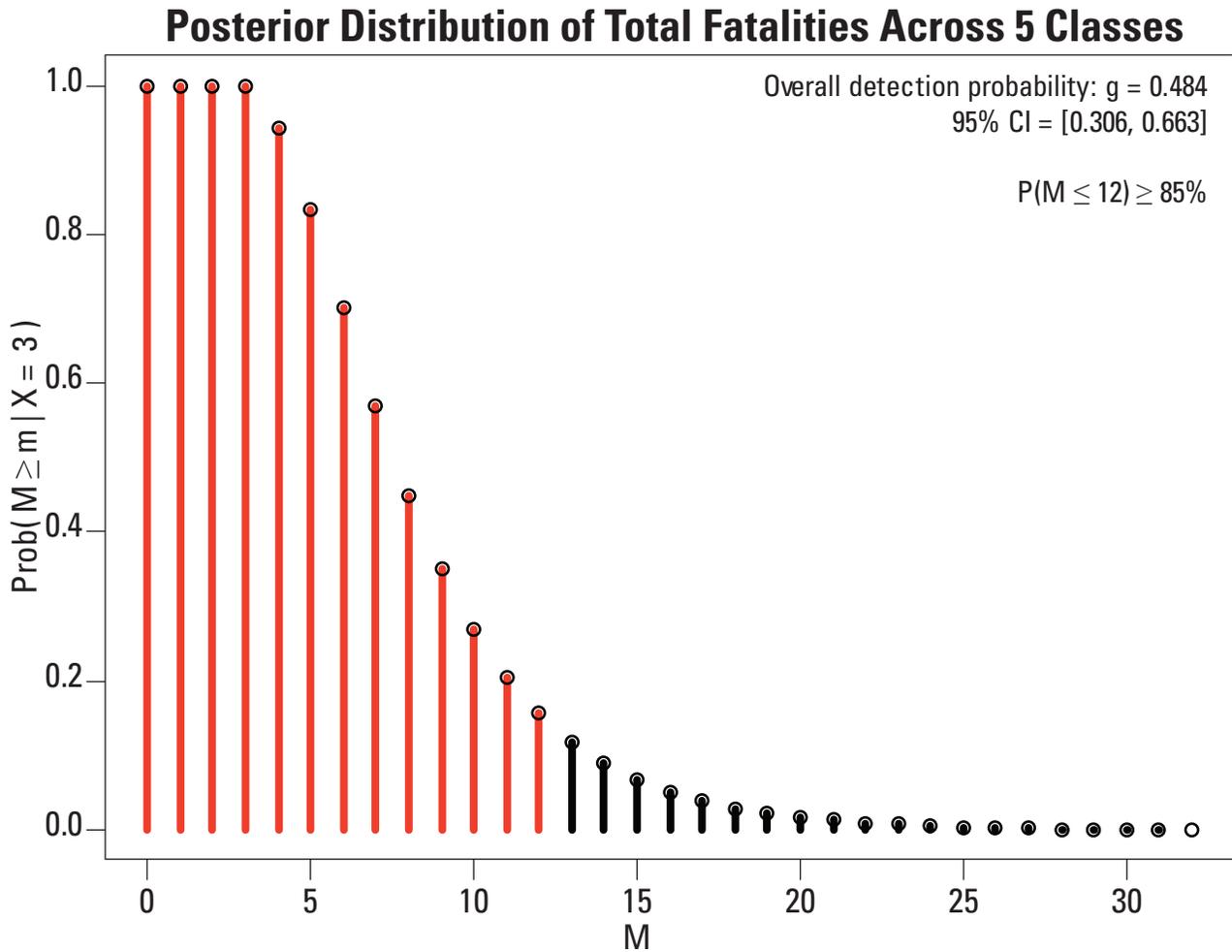


**Figure 13.** Graph generated by Evidence of Absence software showing posterior distribution for the Indiana bat example with informed prior distribution and intensified search schedule.

### Scenario 4: Multiple sites

The posterior distributions for several different hypothetical sites have been gathered into one .csv file called multi-site input.csv that is stored in the csv files folder. To run the analysis for this group of five sites, click **Multiple Classes (sites, visibility, etc.)** on the **Parameters** sheet and **New Data from File(s)** on the **Multiple Classes** sheet. Select **Combined**, click **Go** and enter the multi-site input .csv when prompted. This will generate the graph in [figure 14](#).

According to the figure, the upper bound of the 85 percent credible interval is 12. Having observed a total of 3 Indiana bats, we can assert with 85 percent credibility that the total number of fatalities at the five sites was not greater than 12. A full list of the probabilities is shown in the worksheet. The graph and (or) the table of results can now be saved, or the user may click **Return to Main Sheet**, or do another analysis of **Multiple Sites**.



**Figure 14.** Graph generated by Evidence of Absence software showing posterior distribution for multiple sites for scenario 4.

## Acknowledgments

We would like to thank T.J. Miller and Dawn Bruns of the U.S. Fish and Wildlife Service (USFWS) for their tireless support of this project. We thank Paul Rabie and Kathy Irvine for rigorous evaluation of our statistical approach. A special thanks goes to Mark Dalthorp whose clever solutions saved us many hours of analysis and greatly improved the speed of the software. 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 Ecosystems Mission Area Wildlife Program of the U.S. Geological Survey and the USFWS.

## References Cited

- Baier, T., and Neuwirth, E., 2007, Excel::COM::R: Computational Statistics, v. 22, no. 1, p. 91–108.
- Bispo, R., Bernardino, J., Marques, T.A., and Pestana, D., 2012, Modeling carcass removal time for avian mortality assessment in wind farms using survival analysis: Environmental and Ecological Statistics, v. 20, no. 1, p. 147–165, doi: 10.1007/s10651-012-0212-5.
- Huso, Manuela M.P., 2011, An estimator of wildlife fatality from observed carcasses: Environmetrics, v. 22, no. 3, p. 318–329, doi: 10.1002/env.1052.
- Huso, Manuela, and Dalthorp, Dan, 2014, Accounting for unsearched areas in estimating wind turbine-caused fatality: Journal of Wildlife Management, v. 78, p. 347–358, doi: 10.1002/jwmg.663.
- Huso, M.M.P., Dalthorp, D., Madsen, L., and Dail, D., in press, Estimating turbine-caused bird and bat fatality when zero carcasses are observed: Ecological Applications.
- Huso, Manuela, Som, Nicholas, and Ladd, Lew, 2012, Fatality estimator user's guide: U.S. Geological Survey Data Series 729, 22 p.
- Korner-Nievergelt, F, Korner-Nievergelt, P, Behr, O., Niermann, I., Brinkmann, R., and Hellriegel, B., 2011, A new method to determine bird and bat fatality at wind energy turbines from carcass searches: Wildlife Biology, v. 17, p. 350–363.
- Shoenfeld, P., 2004, Suggestions regarding avian mortality estimation: Mountaineer Wind Energy Center Technical Review Committee; Report for the West Virginia Highlands Conservancy, West Virginia, 6 p.
- Warren-Hicks, W., Newman, J., Wolpert, R., Karas, B., and Tran, L., 2013, Improving methods for estimating fatality of birds and bats at wind energy facilities: Berkeley, California, CalWEA, California Energy Commission, CEC-500-2012-086, 136 p.

## Appendix A. Accounting for Uncertainty in Searcher Efficiency and Carcass Persistence

### Overview

Searcher efficiency and carcass persistence distribution are estimated from data collected in field trials. There is often a considerable degree of uncertainty associated with estimation of these parameters. Accounting for uncertainty in carcass persistence and searcher efficiency parameters leads to increased variance in the posterior distribution and greater difficulty in ruling out high numbers of fatalities. The Evidence of Absence software is designed to account for that uncertainty by simulation. Random searcher efficiency values and carcass persistence distribution parameter values are generated according to distributions derived from user input. For each random draw, the software calculates the probability  $p$  of observing a carcass in the study area during the monitoring season. It is assumed that all individuals arrive, persist, and are detected independently, so that the number of carcasses observed ( $X$ ) is a binomial random variable conditional on the total mortality  $M$ . Thus, given prior

probabilities  $P(M = m)$ , we obtain  $P(M = m | X; p)$  using Bayes' formula 
$$P(M = m | X; p) = \frac{P(X | M = m; p) P(M = m)}{\sum_{j=X}^{\infty} P(X | M = j; p) P(M = j)}$$
 (Huso and Dalthorp, 2014). We assume  $p$  follows a beta distribution and use maximum likelihood to find the beta parameters  $\alpha$  and  $\beta$  that give the best fit to the simulated  $p$ 's; then,  $X|M$  is a beta-binomial random variable with probability mass function:

$$P(X = x | M = m; \alpha, \beta) = \binom{m}{x} \frac{B(x + \alpha, m - x + \beta)}{B(\alpha, \beta)} \quad (\text{A1})$$

where

$$B(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}.$$

From this we calculate the posterior distribution  $P(M = m | X; p)$  for the given prior distribution  $P(M = m)$ .

## Uncertainty in Searcher Efficiency

The user provides an estimated mean searcher efficiency  $\hat{f}$  and 95% CI =  $[\hat{f}_l, \hat{f}_u]$ , which is then assumed to follow a beta distribution(appendix C) with parameters  $\alpha = \frac{16\hat{f}^2(1-\hat{f})}{(\hat{f}_u - \hat{f}_l)^2} - \hat{f}$  and  $\beta = \alpha \left( \frac{1}{\mu} - 1 \right)$ , which is equivalent to mean  $\mu = \hat{f}$  and variance  $\sigma^2 = \left( \frac{\hat{f}_u - \hat{f}_l}{4} \right)^2$ .

## Uncertainty in Persistence Distribution

For the exponential persistence distribution, the user provides point and interval estimates for the mean carcass persistence time ( $\overline{CP}$  with 95% CI =  $[CP_{Lwr}, CP_{Upr}]$ ). (Equivalently, the user may provide point and interval estimates of  $r$  = probability that a carcass persists until the first search after arriving, but these are converted to  $\overline{CP}$  and  $[CP_{Lwr}, CP_{Upr}]$  by the software for the subsequent analysis). We make the assumption that mean CP is distributed as a beta random variable that is scaled and translated to the interval  $[CP_{Lwr}, CP_{Upr}]$ . More precisely, we set  $\mu = (\overline{CP} - CP_{Lwr}) / (CP_{Upr} - CP_{Lwr})$  and define parameters  $\alpha = 16\mu^2(1-\mu) - \mu$  and  $\beta = \alpha \left( \frac{1}{\mu} - 1 \right)$  for a beta random variable  $b$ . Then, the random variable  $CP = b(CP_{Upr} - CP_{Lwr}) + CP_{Lwr}$  has mean =  $\widehat{CP}$ , variance =  $[(CP_{Upr} - CP_{Lwr})/4]^2$ , and is a scaled and translated beta RV on the interval  $[CP_{Lwr}, CP_{Upr}]$ .

To simulate uncertainty in an exponential persistence distribution, we generate random CPs following the beta distribution previously described.

For the two-parameter persistence distributions (lognormal, log-logistic, and Weibull), we simulate uncertainty by holding the shape parameter ( $\alpha$ ) fixed and assuming the scale parameter ( $\beta$ ) varies as a beta random variable with standard deviation equal to one-fourth the width of the 95% CI for  $\beta$  and mean equal to the user-provided point estimate for  $\beta$ .

## Appendix B. Choice of Representative Interval Length for Calculating $r$

The length of the interval  $I_r$  chosen by the user to define the  $r$  parameter should be “typical” for the search protocol (for example, mean, median, and mode) for ease of interpretation, but beyond that, the choice is not critical. The  $r$  parameter should be thought of simply as a parameter of the persistence distribution. Assuming a search interval of  $I_r$ , an interpretation of  $r$  is that it represents the probability that a fresh carcass persists until the first search after arriving. For intervals of different lengths and for intervals in which the arrival function is not constant,  $r$  is still a parameter of the persistence distribution but it cannot be interpreted as the probability that a carcass arriving in that particular interval will persist until the next search.

## Appendix C. Parameterizations of Distributions Used in the Software

*Weibull Distribution.*—The parameterization used for the Weibull( $\alpha$ ,  $\beta$ ) distribution in the Evidence of Absence software is identical to that used in the base package in R with  $\alpha = \text{shape}$  and  $\beta = \text{scale}$ . The probability density function (pdf) and cumulative distribution function (cdf) are given by

$$f(x) = \frac{\alpha}{\beta} \times \left(\frac{x}{\beta}\right)^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^\alpha} \text{ and } F(x) = 1 - e^{-\left(\frac{x}{\beta}\right)^\alpha}, \text{ respectively, while the mean and variance are } \mu = EX = \beta\Gamma\left(1 + \frac{1}{\alpha}\right) \text{ and } \sigma^2 = VX = \beta^2 \left[ \Gamma(1 + 2/\alpha) - \Gamma^2(1 + 1/\alpha) \right].$$

*Lognormal Distribution.*—The parameterization used for the Lognormal( $\alpha$ ,  $\beta$ ) is identical to that used in the base package in R after substituting  $\text{meanlog} = \beta$  and  $\text{sdlog} = \sqrt{\alpha}$ . The pdf is given by  $f(x) = \frac{e^{-(\log x - \beta)^2 / (2\alpha)}}{x\sqrt{2\pi\alpha}}$ , and the mean and variance are  $EX = e^{\beta + \alpha/2} = \mu$  and  $VX = (e^\alpha - 1)e^{2\beta + \alpha} = \sigma^2$ .

*Log-Logistic Distribution.*—The parameterization used for the loglogistic( $\alpha$ ,  $\beta$ ) is identical to that used by

the R statistical package `actuar` after substituting  $\text{shape} = \alpha$  and  $\text{scale} = \beta$ . The pdf is  $f(x) = \frac{(\alpha/\beta)(x/\beta)^{\alpha-1}}{\left[1 + (x/\beta)^\alpha\right]^2}$ , the cdf is  $F(x) = \frac{x^\alpha}{\beta^\alpha + x^\alpha}$ , and the mean and variance are given by  $EX = \frac{\beta\pi/\alpha}{\sin(\pi/\alpha)}$  (if  $\alpha > 1$ ; otherwise, undefined) and  $VX = \beta^2 \left[ \frac{2\pi/\alpha}{\sin(2\pi/\alpha)} - \left( \frac{\pi/\alpha}{\sin(\pi/\alpha)} \right)^2 \right]$  (if  $\alpha > 2$ ; otherwise, undefined).

*Beta Distribution.*—The parameterization used for the beta( $\alpha$ ,  $\beta$ ) that used in the base package in R  $\text{shape1} = \alpha$  and  $\text{shape2} = \beta$ . The probability density function is given by  $f(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}$  for  $0 \leq x \leq 1$  and 0 otherwise. The mean and variance are given as  $\mu = \frac{\alpha}{\alpha + \beta}$  and  $\sigma^2 = \frac{\alpha\beta}{(\alpha + \beta)^2 (\alpha + \beta + 1)}$ , respectively.

*Negative Binomial.*—The parameterization used for the negative binomial( $r$ ,  $p$ ) distribution is identical to R's with  $\text{size} = \alpha$  and  $\text{prob} = p$ . The probability mass function is given by  $p(x) = \frac{\Gamma(x + \alpha)}{\Gamma(\alpha)x!} p^\alpha (1-p)^x$ . The mean and variance are given as  $\mu = \alpha(1-p)/p$  and  $s^2 = \alpha(1-p)/p^2$ , respectively.

## Appendix D. Using Previous Year's Search Data to Inform Current Year's Prior Distribution

*The general idea:* The number of fatalities in a year is a random variable  $M$ . The search data from one year gives information about the random process generating carcasses. For example, if there is perfect overall detection probability  $g = 1$  in a given year but no carcasses are found ( $X = 0$ ), then we conclude there were  $m = 0$  fatalities in that year. It does not mean there will zero the following year, but it does give some indication that a very high number of fatalities will be unlikely unless the random process that generates carcasses is extremely variable or changes substantially.

*Specifics:* We make the assumption that the number of fatalities generated at a site in year is a random variable  $M \sim$  Negative binomial with mean  $= \lambda$  and variance  $= 5 \lambda$ . The previous year's search data is then used to estimate a posterior distribution of  $\lambda$  as:  $P(\lambda|X) = \frac{P(X|\lambda)P(\lambda)}{\int P(X|\lambda)P(\lambda)d\lambda}$ . Taking a uniform prior for  $\lambda$ , we get  $P(\lambda|X) = \frac{P(X|\lambda)}{\int_0^{\lambda_{max}} P(X|\lambda)d\lambda}$ .

For a given  $\lambda$ ,  $P(X|\lambda) = \sum_m P(X|M_\lambda = m)P(M_\lambda = m)$ . The terms  $P(X|M_\lambda = m)$  are beta-binomial probabilities, specifically,  $X|M$  is binomial( $M, p$ ), where  $p$  is the probability of observing a carcass that arrives in the area of interest, and it is assumed that  $p \sim$  beta( $\alpha, \beta$ ).  $\alpha$  and  $\beta$  are derived from the estimated mean and 95% CI of  $p$  from the first year's search. The terms  $P(M_\lambda = m)$  are negative binomial probabilities  $\lambda^m e^{-\lambda} / m!$ , from which we calculate  $P(\lambda|X)$ . The prior distribution for  $M$  is then given as  $P(M) = \int P(M|\lambda)P(\lambda)d\lambda$ .

## Appendix E. Using Several Previous Years' Data to Inform Current Year's Prior Distribution

Let  $M_i$  be a random variable for the number of fatalities in year  $i$ , for  $i = 1, \dots, n+1$ , where we are interested in estimating the posterior of  $M_{n+1}$  using data from previous years to inform the prior distribution. Let the carcass count and detection probability for year  $i$  be denoted by  $X_i$  and  $g_i$ . We have  $X_i \sim \text{binomial}(M_i, g_i)$ . Make the simplifying assumption that  $M_i \sim \text{i.i.d. negative binomial with mean } \lambda \text{ and } \sigma^2 = 5\lambda$ . Then,  $(M_{n+1}|X) = P(X | M_{n+1}) / \left( \sum_m P(X | M_{n+1}) P(M_{n+1}) \right)$ , where  $P(M_{n+1})$  is the prior distribution for  $M_{n+1}$  informed by  $\{X_i : i = 1, \dots, n\}$  and  $\{g_i : i = 1, \dots, n\}$ .

This prior can be calculated (estimated) as follows:

$$P(M_{n+1}) = \int P(M_{n+1}|\lambda) P(\lambda | X_1, \dots, X_n) d\lambda \quad (\text{E1})$$

where  $P(\lambda|X_1, \dots, X_n) = \frac{P(X_1, \dots, X_n|\lambda) P(\lambda)}{\int P(X_1, \dots, X_n|\lambda) P(\lambda) d\lambda}$ , and  $P(\lambda)$  here is a prior distribution for  $\lambda$ .

Taking  $\lambda \sim \text{uniform}$ , this simplifies to  $P(\lambda|X_1, \dots, X_n) = \frac{P(X_1, \dots, X_n|\lambda)}{\int P(X_1, \dots, X_n|\lambda) d\lambda}$ . Because the  $X_i$ 's are independent, the numerator becomes  $\prod_i P(X_i | \lambda)$ , with each term  $P(X_i|\lambda) = \sum_m P(X_i | M_\lambda) P(M_\lambda)$ .

Let  $n$  be the number of years for which data (counts, detection probabilities, and relative weights) are available. We estimate the total number of fatalities by calculating a pooled overall detection probability as the weighted average of detection probabilities through the years and then applying the Bayesian formula to the collection of years considered as a single entity. The method for calculating a pooled  $g$  is described below.

## Appendix F. Estimating Total Number of Fatalities Over Several Years (or Several Sites)

Let  $G_i$  be a random variable for the conditional detection probability of a carcass that arrives at site  $i$  and let  $EG_i = g_i$  and  $VG_i = \sigma_i^2$ .

Let  $A_i = M_i / M$  be a random variable for the proportion of the total carcasses arriving in site  $i$ , where  $M = \sum_i M_i$  and  $M_i \sim \text{Poisson}(\lambda_i)$  are independent.

Then, let  $G$  be a random variable for the probability of detecting a carcass that arrives at one of the sites during the study period.

We have:  $G = \sum_i G_i A_i = \sum_i P(\text{observe carcass} | \text{carcass arrives at site } i) P(\text{carcass arrives at site } i)$ .

Define  $g = EG = E[\sum_i G_i A_i] = \sum_i E[G_i A_i]$ . But  $E[G_i A_i] = EG_i \times EA_i$  because  $G_i$  and  $A_i$  are independent. Define  $EG_i = g_i$  and note that  $EA_i = E[M_i / M] \cong \frac{\lambda_i}{\lambda} \left(1 + \frac{1}{\lambda}\right) - \text{cov}(M_i, M) / \lambda^2$ ; but  $M = M_i + M'$ , where  $M' = \sum_{j \neq i} M_j$  and the  $M_j$  are independent. Thus,  $\text{cov}(M_i, M) = \text{cov}(M_i, M_i + M') = \text{cov}(M_i, M_i) + \text{cov}(M_i, M') = VM_i + 0 = \lambda_i$ , so  $EA_i \cong \frac{\lambda_i}{\lambda}$ , which we will define as  $a_i$ , or the expected proportion of carcasses arriving at site  $i$ .

Therefore, we have:

$$EG \cong \sum_i g_i \lambda_i / \lambda \quad (\text{F1})$$

The variance of  $G$  is given by  $VG = V[\sum_i G_i A_i] = V[\sum_i G_i M_i / M] \cong \sum_i a_i^2 \sigma_i^2 + 1 / \lambda \sum_i a_i (g_i^2 + \sigma_i^2) - g^2 / \lambda$ , which can be shown in several steps. For  $n = 2$ ,  $V[A_1 G_1 + A_2 G_2] = V\left[\frac{M_1 G_1 + M_2 G_2}{M}\right] \cong (\mu_X / \mu_M)^2 \left(\frac{\sigma_X^2}{\mu_X^2} + \frac{\sigma_M^2}{\mu_M^2} - \frac{2 \text{cov}(X, M)}{\mu_X \mu_M}\right)$ , where  $X = M_1 G_1 + M_2 G_2$  and the  $\mu$  and  $\sigma^2$  terms represent the means and variances of their subscripted variables. The components of the expression are calculated here:

1.  $\mu_M = \sigma_M^2 = \lambda$
2.  $\mu_X = \lambda_1 g_1 + \lambda_2 g_2$
3.  $\sigma_X^2 = V[M_1 G_1 + M_2 G_2] = V[M_1 G_1] + V[M_2 G_2]$ , because carcasses both arrive independently and are observed independently at different sites, and  
 $V[M_i G_i] = E^2 M_i \times V g_i + E^2 g_i \times V M_i + V g_i \times V M_i = \lambda_i^2 \sigma_i^2 + g_i^2 \lambda_i + \sigma_i^2 \lambda_i$  for  $i = 1, 2$ , so  
 $V[X] = \sum \lambda_i^2 \sigma_i^2 + \sum \lambda_i (g_i^2 + \sigma_i^2)$
4.  $\text{cov}[X, M] = \text{cov}[M_1 G_1 + M_2 G_2, M] = \text{cov}[M_1 G_1, M_1 + M_2] + \text{cov}[M_2 G_2, M_1 + M_2]$   
 $= \text{cov}[M_1 G_1, M_1] + \text{cov}[M_1 G_1, M_2] + \text{cov}[M_2 G_2, M_1] + \text{cov}[M_2 G_2, M_2]$ .

The middle two terms are zero because the sites are independent. The first and last terms are calculated as

$$\text{cov}[M_i G_i, M_i] = E[G_i M_i^2] - E[M_i G_i] E[M_i] = E G_i \times E M_i^2 - E G_i E^2 M_i = E G_i (E M_i^2 - E^2 M_i) = g_i \lambda_i, \text{ so}$$

$$\text{cov}[X, M] = g_1 \lambda_1 + g_2 \lambda_2.$$

Combining the terms into the original expression gives:

$$\begin{aligned} VG &\cong \left( \frac{\lambda_1 g_1 + \lambda_2 g_2}{\lambda} \right)^2 \left( \frac{\lambda_1^2 \sigma_1^2 + \lambda_2^2 \sigma_2^2 + \lambda_1 (g_1^2 + \sigma_1^2) + \lambda_2 (g_2^2 + \sigma_2^2)}{(\lambda_1 g_1 + \lambda_2 g_2)^2} + \frac{\lambda}{\lambda_2} - 2 \frac{g_1 \lambda_1 + g_2 \lambda_2}{(g_1 \lambda_1 + g_2 \lambda_2) \lambda} \right) \\ &= \frac{\lambda_1^2 \sigma_1^2 + \lambda_2^2 \sigma_2^2 + \lambda_1 (g_1^2 + \sigma_1^2) + \lambda_2 (g_2^2 + \sigma_2^2)}{\lambda_2} - \frac{g_2}{\lambda} \end{aligned}$$

which can be rewritten as:

$$VG \cong \sum a_i^2 \sigma_i^2 + \frac{\sum a_i (g_i^2 + \sigma_i^2) - g^2}{\lambda} \tag{F2}$$

By mathematical induction, the sums can be taken over  $i = 1, \dots, n$ .

Publishing support provided by the U.S. Geological Survey  
Science 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  
9777 NW 9th St., Suite 400  
Corvallis, Oregon 97330  
<http://fresc.usgs.gov>

