Webinar on Estimating Fatality at Wind-Power Plants

Manuela Huso
Research Statistician
Forest & Rangeland Ecosystem Science Center

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

Overview

- Background
  - Why is estimating fatality so difficult?
- Whom the software is for
- What the software does and does <u>not</u> do
- User responsibilities
- Input and Output
- Future Steps

USGS

Why is estimating fatality so hard?

- Need to estimate abundance (morbid)
- Different circumstances than other studies?
  - Open population
  - Super population estimate
  - Nonhomogeneous population with greatly varying probabilities of detection
  - Nonconstant density within plots below turbines
  - Recruitment rate variable for some species (migrants) perhaps not for others (residents)
  - Loss rate variable among species and locations

USGS

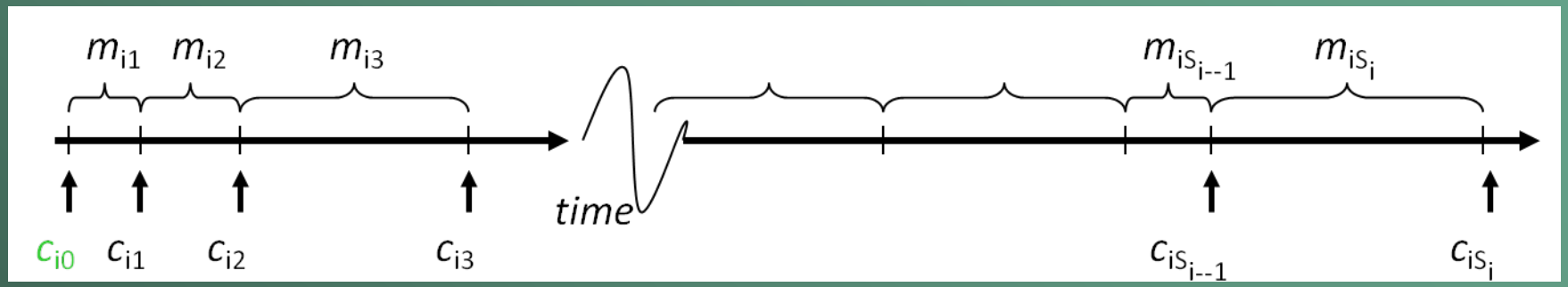
Why is estimating fatality so hard?

- Imperfect detection
  - Not all turbines searched -> sampling fraction
  - Carcasses removed -> proportion persisting
  - Searchers miss some -> proportion seen
  - Inaccessible areas in search plot -> proportion of carcasses in search plot expected to be within searchable area
  - Some carcasses will fall outside designated search plot -> proportion <u>total</u> carcasses expected to be within potential search plots

Density Weighted Area

USGS

Why is estimating fatality so hard?



$$\hat{m}_i \neq \sum_{j=1}^{S_i} c_{ij}$$

m_{ij} = number of dead bats at turbine *i* in interval *j*

c_{ij} = *number* of carcasses counted at turbine *i* at the end of the interval *j*

USGS

Why is estimating fatality so hard?

<1

$$s\ a_i\, p_{ij} e_{ij} r_{ij}\, f\ m_{ij} = c_{ij}$$

- Not all turbines searched -> sampling fraction ***f***
- Carcasses removed -> proportion persisting r_{ij} , e_{ij}
- Searchers miss some -> proportion seen p_{ij}
- Inaccessible areas in search plot -> proportion of carcasses in search plot expected to be within searchable area a_i
- Some carcasses will fall outside designated search plot -> proportion <u>total</u> carcasses expected to be within potential search plots ***s***

USGS

Why is estimating fatality so hard?

- Rearrange to estimate m
- *f* is easy to calculate: usually *n/N*
- Usually *s* ~ 1, i.e. search plots are large enough to encompass most of carcasses

$$\hat{m}_{ij} = \frac{c_{ij}}{\hat{s}\,\hat{a}_i\,\hat{p}_{ij}\,\hat{e}_{ij}\,\hat{r}_{ij}\,f}$$

USGS

Why is estimating fatality so hard?

- Usually search interval is short, so *e* ~ 1
- Sometimes density in searchable area, *a* << 1
- Sometimes *p* and/or *r* << 1.

$$\hat{m}_{ij} = \frac{c_{ij}}{\hat{s}_i \hat{a}_i \hat{p}_{ij} \hat{e}_{ij} \hat{r}_{ij} f}$$

USGS

Why is estimating fatality so hard?

p **assumed to be binomially distributed**

Many options for distributions of persistence times, from which we calculate *r*

Simplest: Exponential with constant removal rate

$$\hat{r} = \frac{\int_{x=0}^{I} e^{-x/\hat{\bar{t}}} dx}{I}$$

USGS

Why is estimating fatality so hard?

Alternatives from survival analysis literature:

Weibull

$$\hat{r} = \frac{\int_{x=0}^{I} \exp(-(x/\hat{\bar{t}})^{1/\phi})\,dx}{I}$$

LogLogistic

$$\hat{r} = \frac{\int_{x=0}^{I} 1/(1+\exp(x/\hat{\bar{t}})^{1/\phi})\,dx}{I}$$

LogNormal

$$\hat{r} = \frac{1-\int_{x=0}^{I} \Phi(x/\hat{\bar{t}})^{1/\phi}dx}{I}$$

USGS

Why is estimating fatality so hard?

- No "closed form" solution
  - i.e. can't easily write an equation for the variance
- Some authors cite Goodman 1960
  - Var of product, not inverse of product
  - Result: Confidence Limits <0
- Solution: Bootstrap (Efron and Tibshirani 1986)

$$\hat{m}_{ij} = \frac{c_{ij}}{\hat{s}\ \hat{a}_i \hat{p}_{ij} \hat{e}_{ij} \hat{r}_{ij} f}$$

USGS

Goodman, L. A. 1960. On the exact variance of products. *Journal of the American Statistical Association* **55:** 708-713.

Efron, B., and R. Tibshirani. 1986. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science* ***1: 54-77.***

What software DOES - DOES NOT do

Does:

- **Gives (1-alpha) level confidence limits on estimated SE*, CP* and mortality****
- **Saves all bootstrap samples for further use and archiving**
- ***Can* be used to determine best CP distribution**
- ***Can* be used to select best SE and CP models**

*** all confidence limits for SE, CP between 0,1**

**** all confidence limits for mortality >= 0**

USGS

What software DOES - DOES NOT do

Does NOT:

- Determine best CP distribution for user
- Determine best CP or SE model for user
- Estimate fatality using fewer than 10 SE or CP trial carcasses in any factor combination
- Determine Density Weighted Area adjustment

USGS

Who should use this software?

- Managers with some statistics background
- Consultants with some statistics background
- Researchers with some statistics background

$$\hat{m}_{ij} = \frac{c_{ij}}{\hat{s}\ \hat{a}_i \hat{p}_{ij} \hat{e}_{ij} \hat{r}_{ij}\ f}$$

USGS

User Responsibilities

- Use adequate trial carcasses for SE & CP
- Measure factors related to SE & CP
- Select models of SE & CP (software can help)

USGS

Example

3 input files

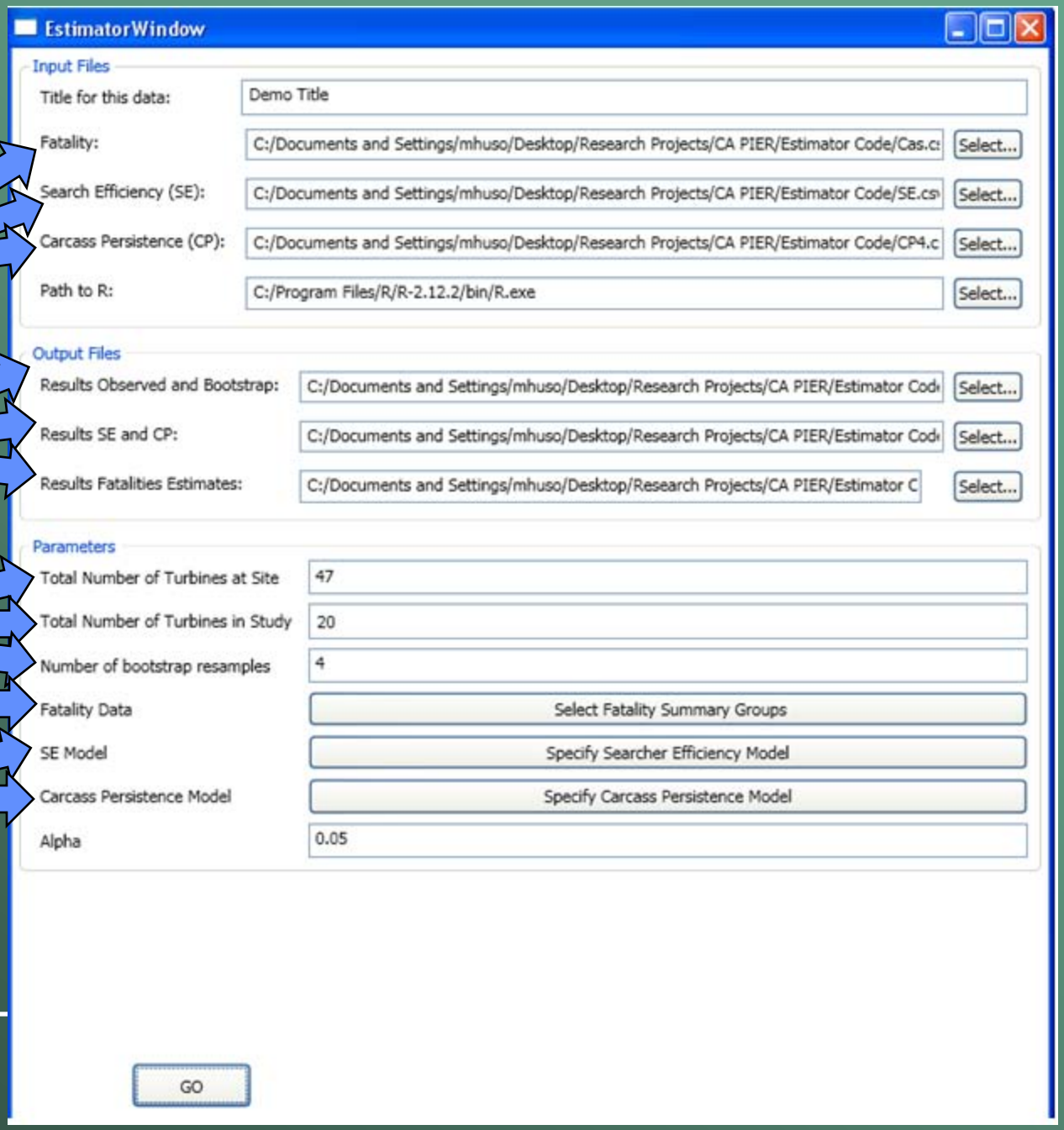
3 output

Other in

Output groups

SE & CP

models



USGS

Example Estimator Input Data

Fatality Data input

Date	LastSearch	Turbine	Season	Size	Vegetation	DWArea	PropSite	BatBird	Group	MW
7/22/2009	6/19/2009	B1	Summer	S	Open	0.73	0.96	Bat	tree	1.5
10/6/2009	9/18/2009	B4	Fall	L	Open	0.87	0.96	bird	raptor	2.3
9/3/2009	8/14/2009	B7	Fall	S	Heavy	1.00	0.96	bat	non-tree	1.5

. . .

Date – Date carcass was found
LastSearch – Date of previous search (combined with Date to calculate interval)
Turbine (necessary for bootstrapping)
Season, Size, Vegetation – optional factors affecting SE and/or CP
DWArea – Density Weighted Searchable Area of Plot
PropSite – Density Weighted Proportion of Site represented in plots
TaxonGroup, SpeciesGroup, KWH – factors for which fatality will be summarized

USGS

Example Estimator Input Data

Searcher Efficiency Trial Data input

Date	Vegetation	Size	Turbine	Found
5/22/2009	Open	L	B7	1
7/16/2009	Moderate	S	G3	1
10/2/2009	Heavy	S	D1	0
. . .				

Date – Date carcass was placed
Size, Vegetation – optional factors affecting SE
Turbine (necessary for bootstrapping)
Found – Indicator of whether searchers found the carcass

USGS

Example Estimator Input Data

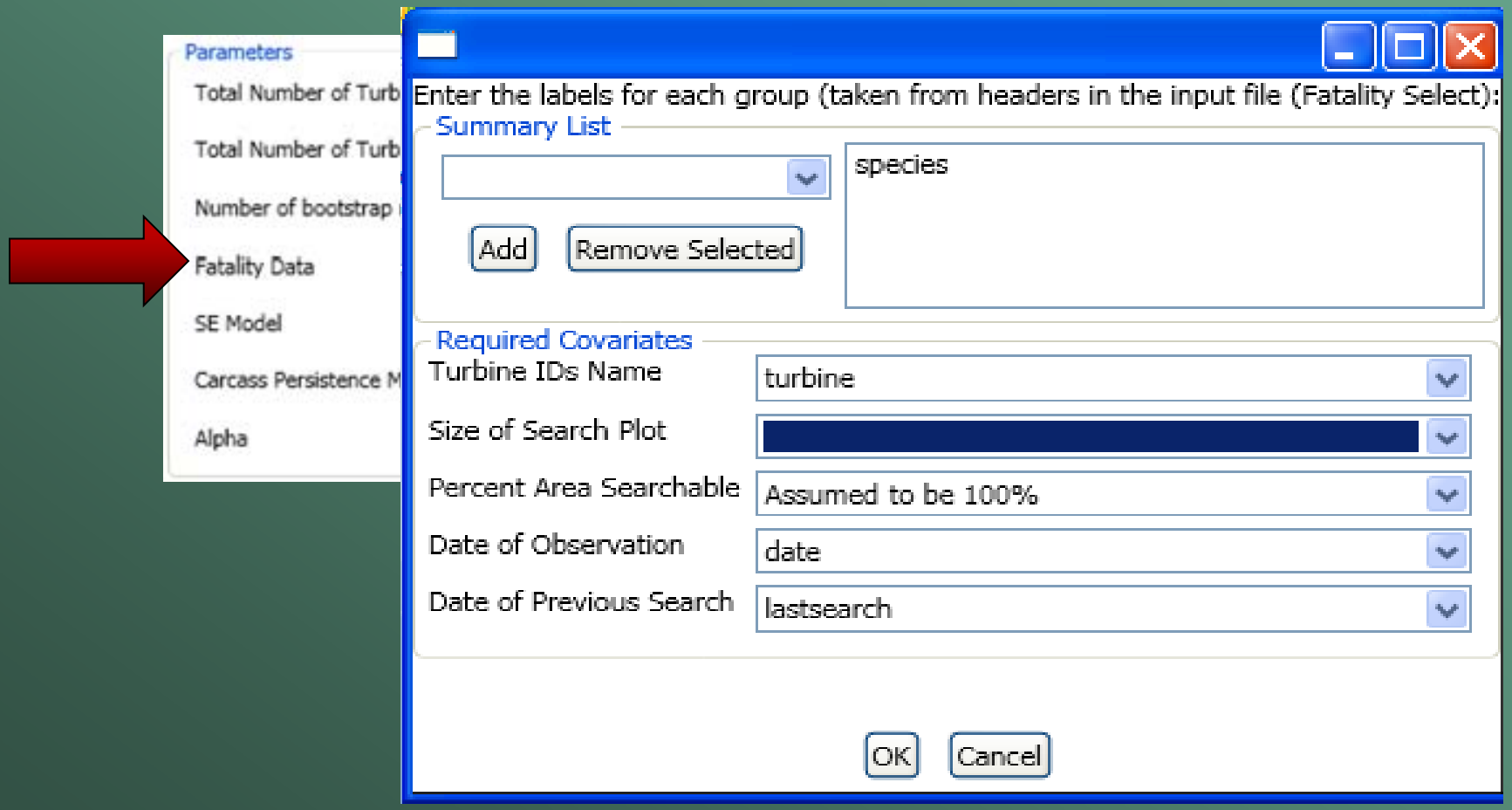
Carcass Persistence Trial Data input

Date	Season	Size	Turbine	Left	Right	Event
7/12/2009	Summer	S	B5	1	1	2
7/16/2009	Summer	S	F3	7	10	3
7/23/2009	Summer	S	D4	35	35	0
. . .						

Date – Date carcass was placed
Turbine (necessary for bootstrapping)
Season, Size, Vegetation – optional factors affecting CP
Days (hours) since being placed that carcass was last known to be present
Right – Days (hours) since being placed that carcass was first known to be absent
if carcass still present when monitoring discontinued, then Left=Right
Event – censoring type: 0 => right, 2=> left, 3 => interval

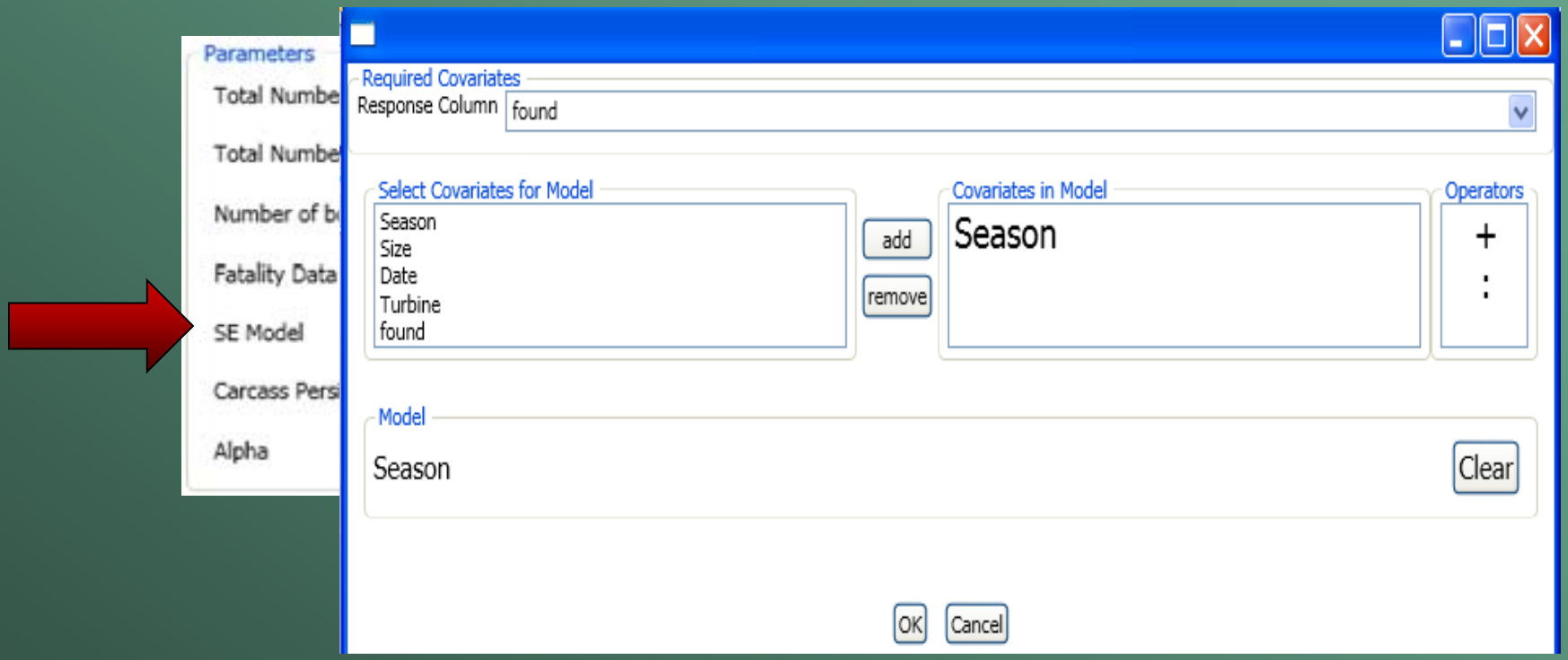
USGS

Example Estimator Interface



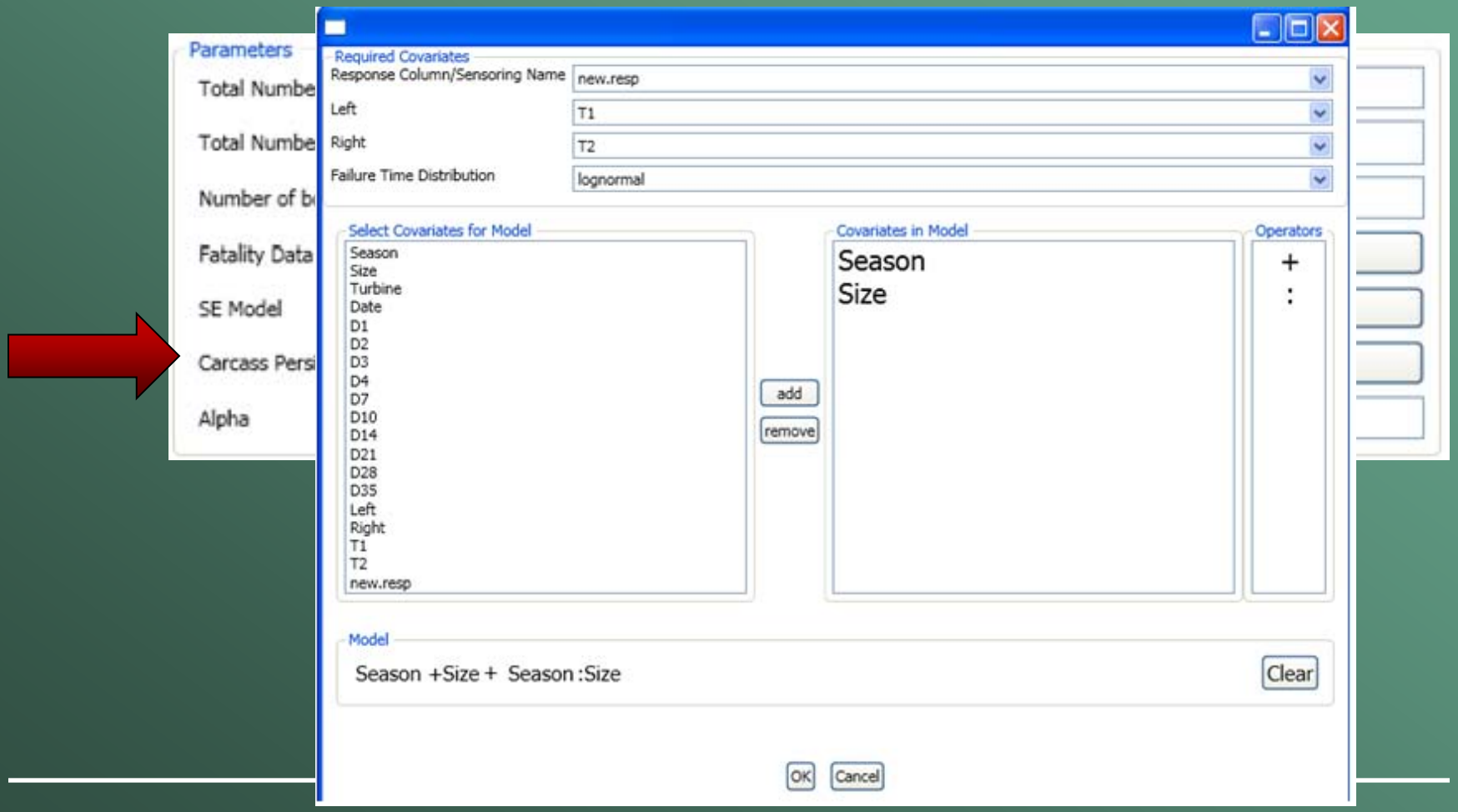
USGS

Example Estimator Interface



USGS

Example Estimator Interface



USGS

Estimator Process

1) *From original sample of CP trial carcasses determine best persistence distribution and model*
2) *From original sample of SE trial carcasses, determine best SE model*
3) *Input SE trial, CP trial and Fatality data into estimator*

NOTE:

- **Sampling unit is the <u>turbine</u>**
  - **Estimate total fatality *<u>at the turbine</u>*, through modeling**
  - **Estimate total fatality *<u>at the site</u>* through sampling**

USGS

Estimator Process

Fatality Estimator generates 3 bootstrap samples:

1) Draw bootstrap sample of turbines
-> bootstrap sample of carcasses

2) Draw bootstrap sample of CP trial carcasses
-> estimate r_{ij} based on user-defined covariates and distribution

3) Draw bootstrap sample of SE trial carcasses
-> estimate p_{ij} based on user-defined covariates

Combine r_{ij} and p_{ij} (and other bias corrections) with bootstrap sample of carcasses
-> estimate adjusted fatality

USGS

Estimator Process

- After ~1-5000 (# determined by user) iterations, summarize fatalities per turbine by factors selected by user,
  - e.g. bats/birds
  - raptors/waterfowl/passerines
- Expand estimates to entire site using *s* and *f*

USGS

Example Estimator Output

User-defined alpha level of 0.05; all reported confidence intervals are 0.95% confidence intervals

Searcher Efficiency Estimates; AIC for model = Season+Size:207.11

SE.Cov.Ind	Season	Size	Found	Placed	Estimate	Lower	Upper
Fall.S	Fall	S	7	19	0.37	0.21	0.58
Spring.L	Spring	L	16	20	0.74	0.65	0.82
Spring.S	Spring	S	8	18	0.51	0.41	0.68
Summer.L	Summer	L	5	11	0.56	0.39	0.71
Summer.S	Summer	S	4	9	0.32	0.19	0.47

Carcass Persistence Estimates; AIC for model = Season+Size and distribution = lognormal:221

Estimates and Confidence Intervals for r based on an interval of 20 days.

CP.Cov.Ind	Season	Size	Placed	CP	Lower	Upper	r	r.Lower	r.Upper
Fall.S	Fall	S	10	17.24	10.14	32.46	0.25	0.19	0.31
Spring.L	Spring	L	10	36.67	20.33	233.58	0.32	0.26	0.43
Spring.S	Spring	S	10	9.97	5.85	19.96	0.19	0.13	0.26
Summer.L	Summer	L	10	14.71	4.74	110.5	0.23	0.11	0.4
Summer.S	Summer	S	10	4	0.96	21.04	0.09	0.02	0.27

USGS

Example Estimator Output

Results based on fewer than 2000 bootstrap resamples:
STANDARD ERROR ESTIMATES CONSIDERED EXPERIMENTAL.

Turbine search areas were not provided.

This study included 20 turbines, and these estimates have been expanded to reference the 47 at the study site.

User defined alpha level of 0.05; all reported confidence intervals are 0.95% confidence intervals
The number of observed casualties for some requested groups is less than 5,
use caution when interpreting estimates.

Requested Summaries

Factor	Level	Number. Found	PerTurbine. Estimate	PerTurbine. Lower	PerTurbine. Upper	SiteTotal. Estimate	SiteTotal. Lower	SiteTotal. Upper
sizeseason	LSpring	3	0.61	0.2	1.12	29	9	53
sizeseason	LSummer	2	0.93	0.62	1.93	44	29	91
sizeseason	SFall	12	6.32	3.36	9.93	297	158	467
sizeseason	SSpring	3	1.71	0.75	5.89	80	35	277
sizeseason	SSummer	8	8.51	3.55	13.83	400	167	650

USGS

Example Estimator Output Data

All Bootstrap Samples Data

Date	LastSearch	Turbine	Season	Size	Vegetation	DWArea	PropSite	BatBird	Group	MW	SE.hat	CP.hat	Mort	Iter
7/22/2009	6/19/2009	B1	Summer	S	Open	0.73	0.96	bat	tree	1.5	0.37	17.24	10.53	0
10/6/2009	9/18/2009	B4	Fall	L	Open	0.87	0.96	bird	raptor	2.3	0.74	36.67	4.22	0
9/3/2009	8/14/2009	B7	Fall	S	Heavy	1.00	0.96	bat	non-tree	1.5	0.51	9.97	10.56	0

. . .

Echoes original carcass input data

Iter – bootstrap iteration

Iter = 0 => original data

Iter > 0 => resampled data

USGS

Future Work

- Develop hosting site*
- Incorporate different SE models
- Incorporate new work by
  - Dail and Huso on simultaneous estimates of r, p and m
  - Korner et al. 2011, alternative estimator
- Incorporate evidence of absence calculations
- Other ?

USGS

* Available as soon as permission is granted

Acknowledgements

California Energy Commission PIER program
Nicholas Som and Lew Ladd, EcoStats, LLC
Ted Weller, USDA PSW Research Station
Ed Arnett, Bat Conservation International

Thank you!

USGS