USGS science for a changing world

Introduction to GenEst
Generalized Estimator of Mortality

Manuela Huso
Research Statistician
Forest & Rangeland Ecosystem Science Center, Corvallis, OR
NWCC/WREN Webinar August 23, 2018

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

This information is preliminary and is subject to revision. It is being provided to meet the need for timely best science. The information is provided on the condition that neither the U.S. Geological Survey nor the U.S. Government shall be held liable for any damages resulting from the authorized or unauthorized use of the information.

Development Team

- Daniel Dalthorp
- Juniper Simonis
- Lisa Madsen
- Manuela Huso
- Paul Rabie
- Jared Studyvin
- Jeffrey Mintz

Consultation with:
Robert Wolpert
Fränzi Korner oikostat.ch

USGS 2

Financial Support

- USBLM
- USGS
- DoE - NREL
- WEST, Inc.
- Bat Con. Int.
- AWWI
- Avangrid

Steering Committee

- Taber Allison
- Cris Hein
- Jerry Roppe
- Dave Cowan
- Danna Small
- Paul Rabie
- Dan Dalthorp
- Manuela Huso
- Ian Evans

USGS 3

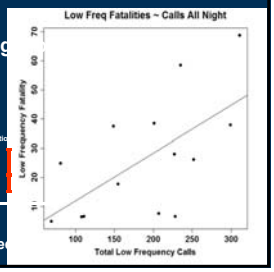
Alternative Energy Development

- What is the cost to wildlife?
- Are some species differentially affected?
  - Cave-roosting bats after WNS
  - Migratory tree bats
  - Golden eagles
- Can we avoid/reduce impacts?
- Need *accurate* estimates
- Need *measure of uncertainty*
- Need to *correctly account* for all we don't see

USGS 4

Why GenEst is Needed

- Accuracy of estimator critical for
  - Meaningful comparisons
  - Sites (and site properties)
  - Estimating trends
  - Evaluating efficacy of mitig



USGS

Monitoring widely differing

- Countries - States/provinces - Counties
- Monitoring required at…
  - All … some… none
- Public access to data
  - All publicly available…
  - some proprietary…
  - none need be public
- Adjustments for non-detection vary widely
  - Different estimators adjust differently

USGS 6

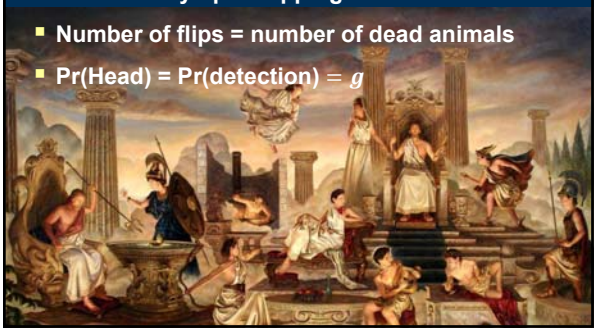
Monitoring widely differing

USGS

7

Analogy

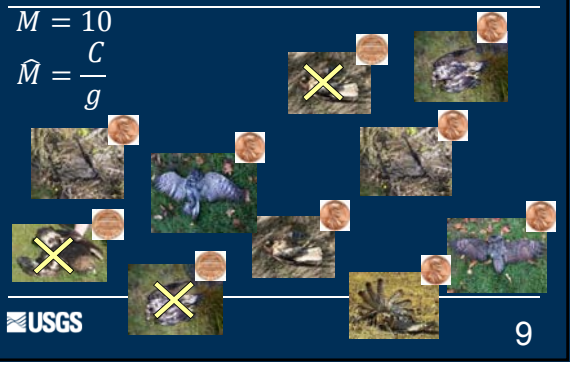
- Gods on Olympus flipping coins
- Number of flips = number of dead animals
- Pr(Head) = Pr(detection) = g



Estimating Mortality: Analogy

$M = 10$

$\hat{M} = \frac{C}{g}$



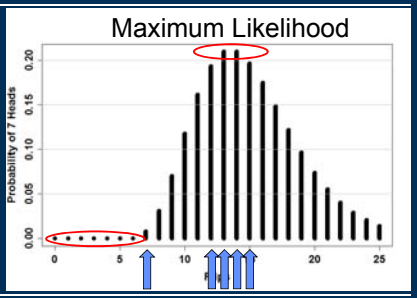
USGS

9

Probability of 7 Heads with fair coin

$M \geq 7$

$\hat{M} = 14$



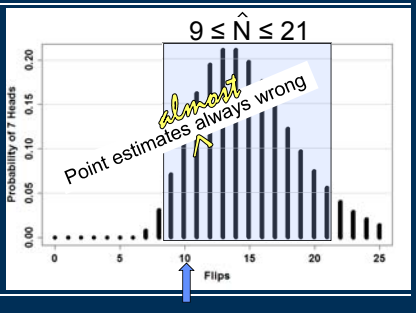
USGS

10

Probability of 7 Heads with fair coin

$\hat{N} = 14$

$N \geq 7$

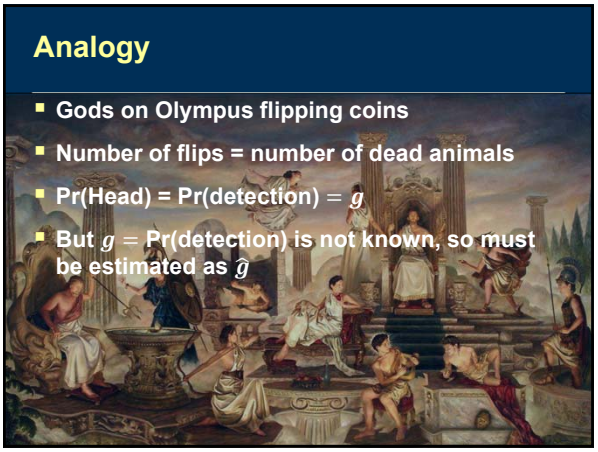


USGS

11

Analogy

- Gods on Olympus flipping coins
- Number of flips = number of dead animals
- Pr(Head) = Pr(detection) = g
- But g = Pr(detection) is not known, so must be estimated as $\hat{g}$



Current Protocol – Estimating $\widehat{M} = \frac{C}{\hat{g}}$

Estimating g is notoriously difficult

Different assumptions → different estimators:

- Persistence patterns
- Probability a searcher sees a carcass
- Fraction in searchable area

Assumptions aren't necessarily wrong…

they're just not always right.

USGS

Current Protocol

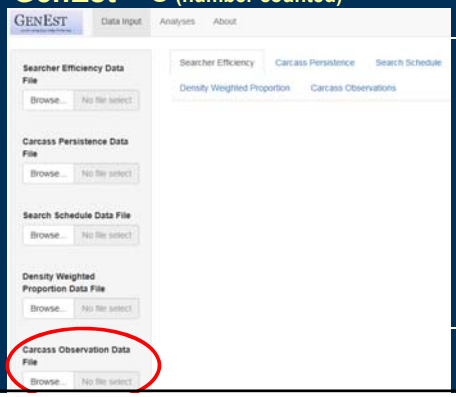
USGS

Current Protocol – C (number counted)

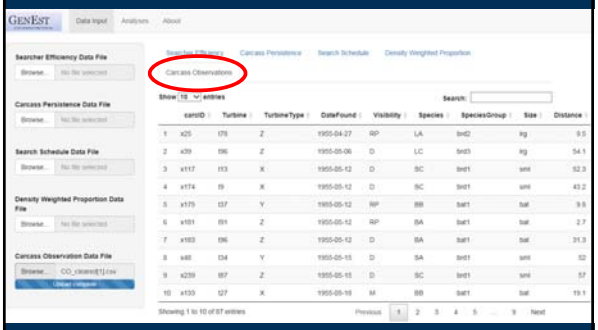
- Send crews out on regular basis to search for carcasses beneath turbines
- Count the number of "heads" observed
- Search protocol similar for all estimators

USGS

GenEst – C (number counted)



GenEst – C (number counted)

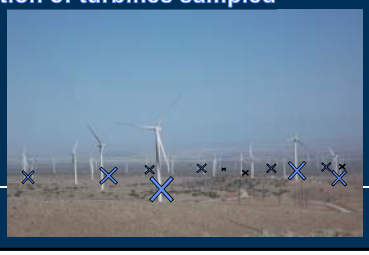


USGS

Major sources of imperfect detection

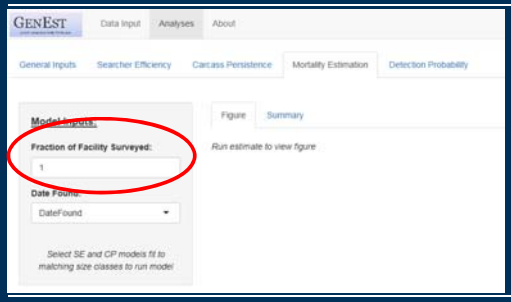
$$M_{ij} f \neq C_{ij}$$

f = fraction of turbines sampled



USGS

GenEst – *f* (fraction of facility surveyed)



USGS

Major sources of imperfect detection

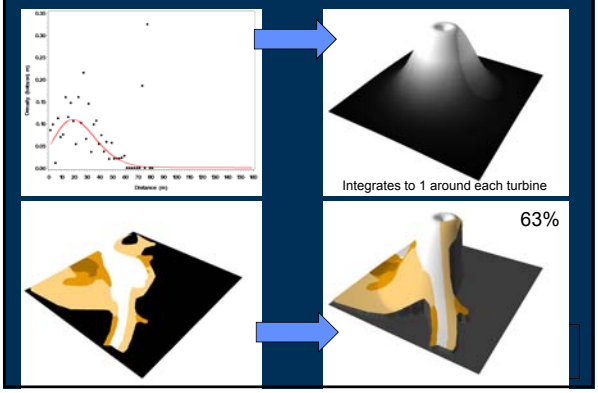
$$M_{ij} f a_i \neq C_{ij}$$

f = fraction of turbines sampled
a = fraction of carcasses in searched area

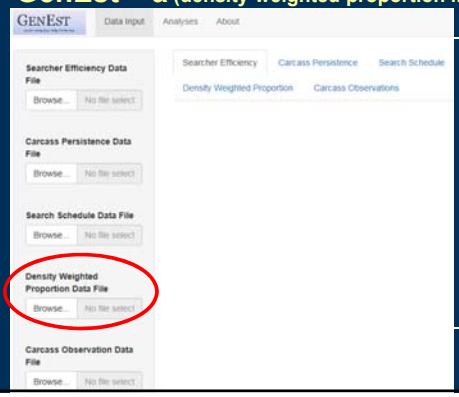
What fraction of carcasses was in the searched area?
Not what fraction of plot was searched?

USGS

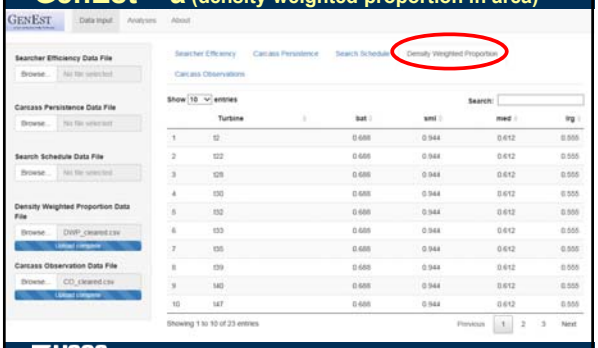
Current Protocol – *a* (proportion in search area)



GenEst – a (density-weighted proportion in area)



GenEst – a (density-weighted proportion in area)

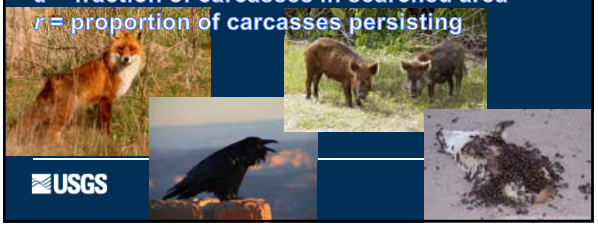


USGS

Major sources of imperfect detection

$$M_{ij} f a_i r_{ij} \neq C_{ij}$$

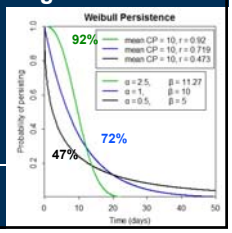
f = fraction of turbines sampled
a = fraction of carcasses in searched area
r = proportion of carcasses persisting



USGS

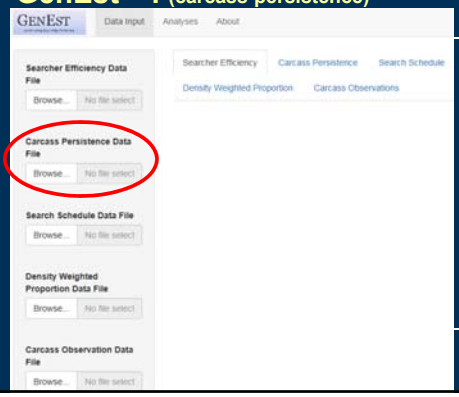
Current Protocol – *r* (proportion persisting)

- Place trial carcasses in field
- Check often to determine fate
- Model persistence <u>pattern</u> not just average time
- Estimate probability of persisting to next scheduled search



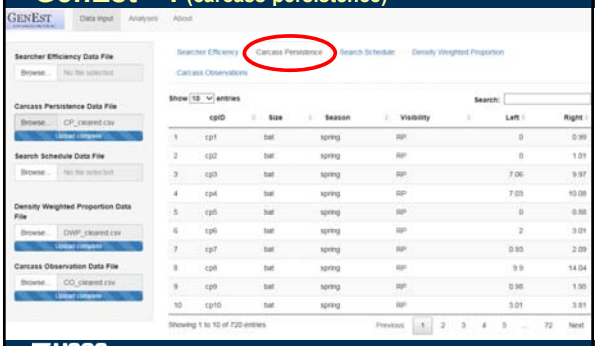
USGS

GenEst – r (carcass persistence)



26

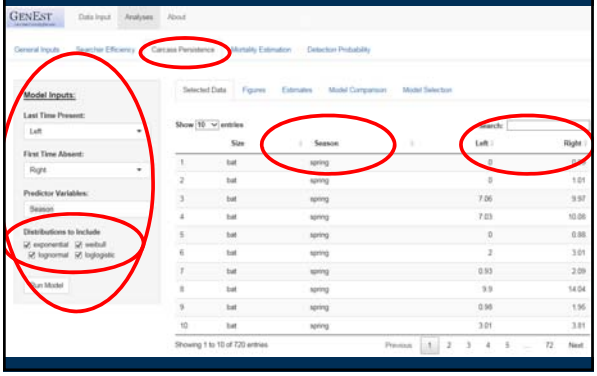
GenEst – r (carcass persistence)

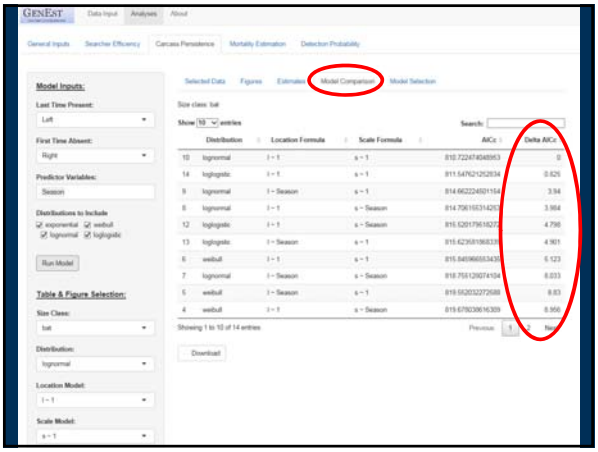


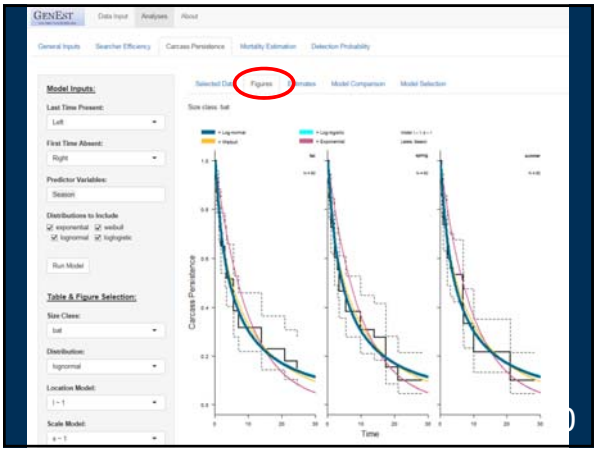
USGS

27

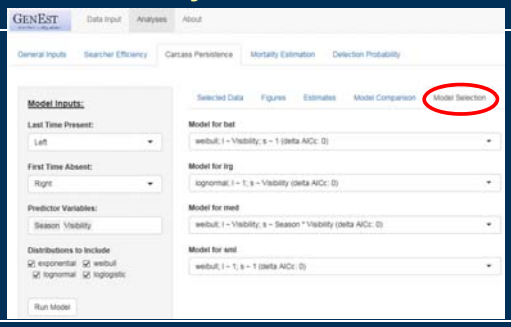
GenEst – Analyses – Carcass Pers







GenEst – Analyses – Carcass Pers



USGS

Major sources of imperfect detection

$$M_{ij} f a_i r_{ij} p_{ij} \neq C_{ij}$$

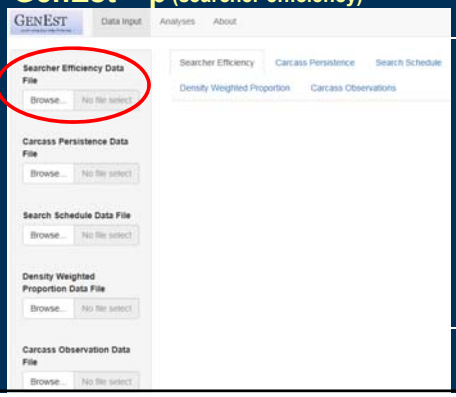
f = fraction of turbines sampled
a = fraction of carcasses in searched area
r = proportion of carcasses persisting
p = proportion of carcasses found

Current Protocol – *p* (proportion observed)

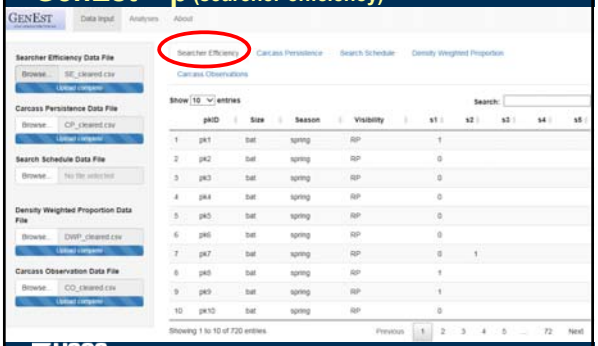
- Place trial carcasses in field
- Determine if observed by searcher or not
- If not, leave for next search
- Estimate probability of searcher observing carcass
- Estimate decline in probability (*k*) with each search

USGS

GenEst – p (searcher efficiency)

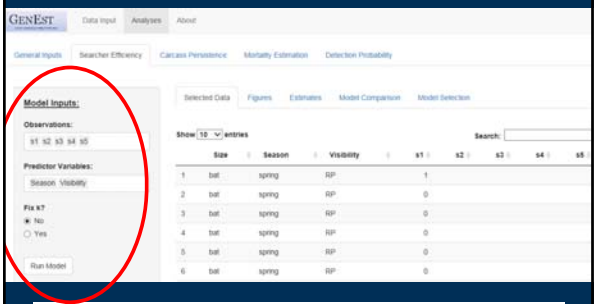


GenEst – p (searcher efficiency)



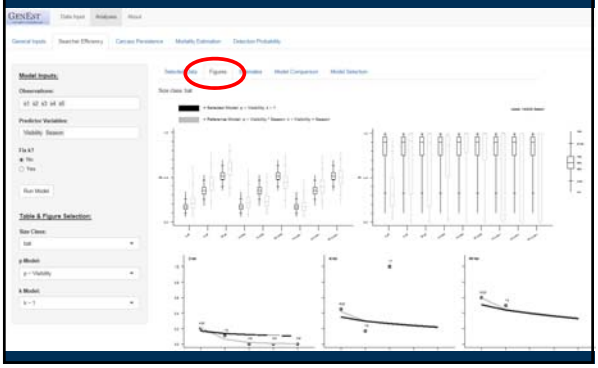
USGS

GenEst – Analyses – Searcher Eff.

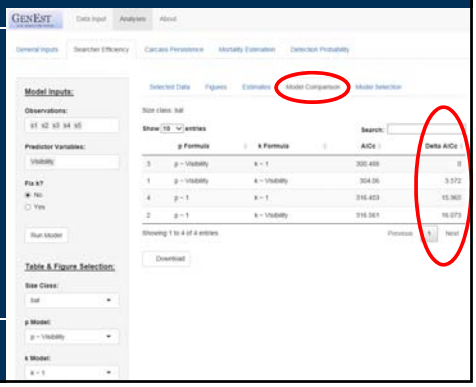


USGS

GenEst – Analyses – Searcher Eff.

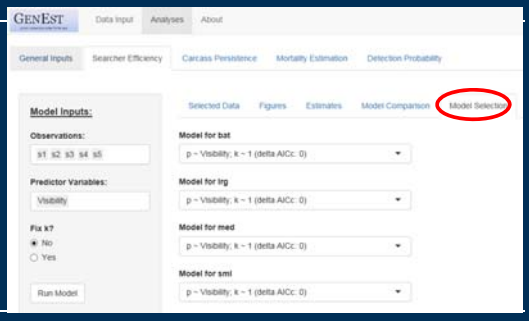


GenEst – Analyses – Searcher Eff.



USGS

GenEst – Analyses – Searcher Eff.



USGS 39

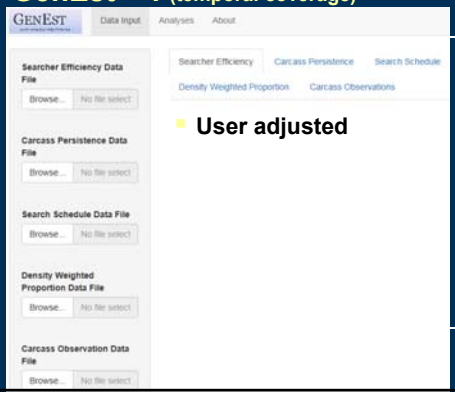
Major sources of imperfect detection

$$M_{ij} f a_i r_{ij} p_{ij} v \cong C_{ij}$$

f = fraction of turbines sampled
a = fraction of carcasses in searched area
r = proportion of carcasses persisting
p = proportion of carcasses found
v = proportion of annual mortality arriving during monitoring period

USGS 40

GenEst – v (temporal coverage)



- User adjusted

41

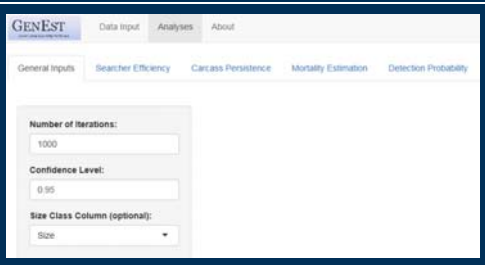
GenEst – *g* overall pr(detection)

$$\underbrace{\sum_j \sum_h \left(\prod_{l=0}^{j-h-1} (1-k^l p) k^{j-h} p \right) \int_{t_{h-1}}^{t_h} S(t_j - t) v(t) dt}_{g_{ij}}$$

f = fraction of turbines sampled
a = fraction of carcasses in searched area
r = proportion of carcasses persisting
p = proportion of carcasses found
v = proportion of annual mortality arriving during monitoring period

USGS 42

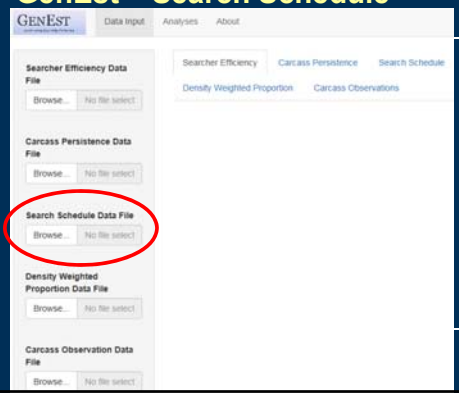
GenEst – Analyses – General Inputs



USGS

43

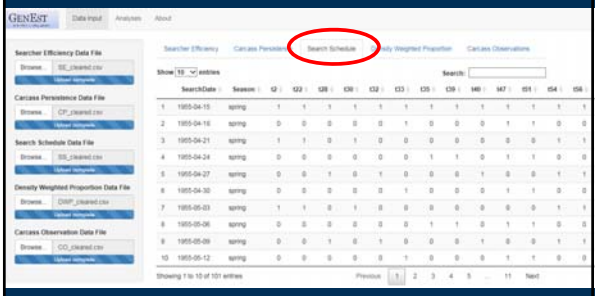
GenEst – Search Schedule



USGS

44

GenEst – Search Schedule



USGS

45

GenEst – Analyses – Mortality Estimate

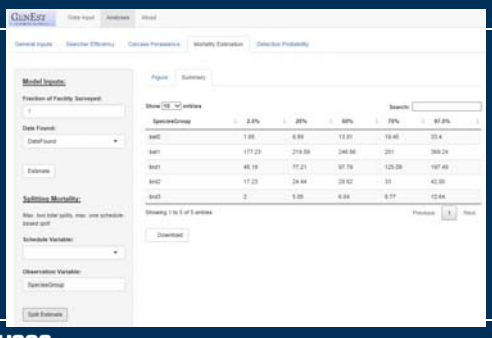
USGS

46

GenEst – Analyses – Mortality Estimate

USGS

GenEst – Analyses – Mortality Estimate



USGS

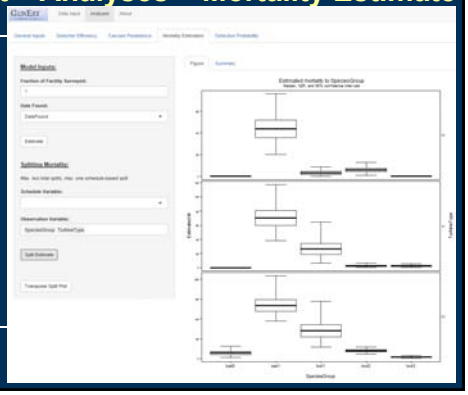
48

GenEst – Analyses – Mortality Estimate

USGS

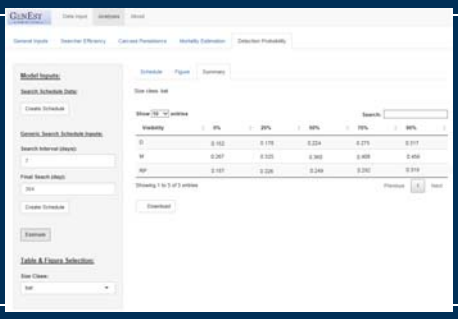
49

GenEst – Analyses – Mortality Estimate



USGS

GenEst – Analyses – Pr(Detection): g



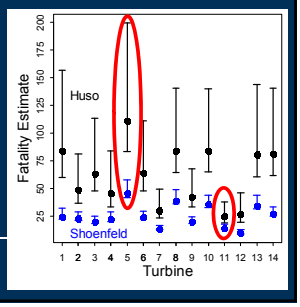
USGS

51

Current Protocol – Estimating *M*

Simplifying assumptions re:

- Persistence patterns
- Proportion observed
- Proportion in search area



USGS

GenEst

- Natural evolution of methods –
  each new estimator offered improvements
- Relaxes restrictive assumptions
  - Persistence patterns
  - Proportion found on 1st, 2nd, 3rd … search
  - Proportion in search area
- Information with which to evaluate assumptions
- Accurate estimate of uncertainty
- Data collection process little changed

USGS

53

GenEst is…

- GenEst R package for Windows, Mac, Linux
  with GUI and command line implementations
- GenEst Statistical Models
  technical manual "USGS Techniques and Methods"
- GenEst User Guide
- R Vignettes:
  - GenEst - A Tutorial with Wind Examples
  - GenEst - A Solar Example with Carcass Size
  - GenEst - Command Line Walkthrough

USGS

54

What GenEst Does

- Accounts for all sources of non-detection
- Unbiased* → Meaningful…
  - Comparisons of sites (and site properties)
  - Estimation of trends
  - Evaluation of efficacy of mitigation treatments
  - Estimates for different
    - Locations, e.g., power lines vs buildings
    - Groups, e.g., passerines vs raptors

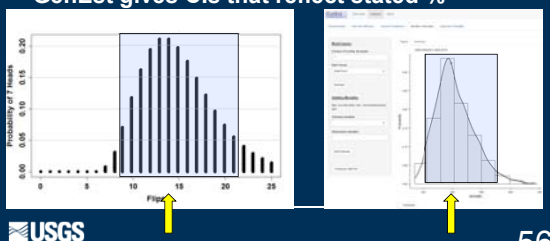
  … from same study, same software

* As unbiased as possible when g is not known

USGS

What GenEst Does

- Variance is critical
  - Single point estimate almost certainly wrong
- GenEst gives CIs that reflect stated %



USGS

What GenEst Does Not Do

- NOT make decisions for the user
  - User must input valid data
  - Select best model
- NOT interpret results
- NOT estimate when 0 carcasses observed

GenEst is NOT EoA

(EoA and GenEst estimates converge when >>0 carcasses are observed)

USGS

Current Status

- First Beta-test reviews
  - Returned and currently being reconciled
    - Statistics specialists
    - Users very experienced in mortality estimation
  - Anonymous unless waived by reviewer
- Publish Spinoffs:
  - `cbinom` package – Dalthorp
  - $\bar{X}$ paper – Madsen, Dalthorp, Huso
  - *p* and *k* simultaneous estimation – Dalthorp, Madsen, Schirmacher, Huso

USGS

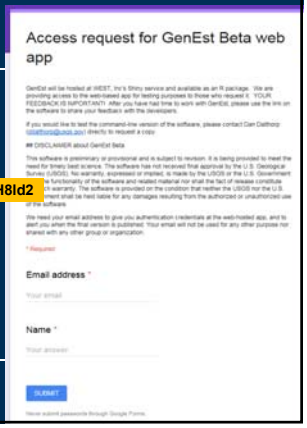
Current Status

Request reviews from larger community

You!

https://goo.gl/forms/bYKOWA5vELhOH8Id2

Will reconcile as we receive them



USGS

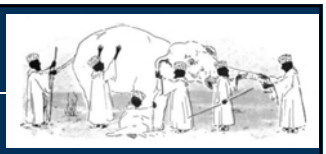
Future Work

- Training workshops around the US
- Simulation analysis comparing GenEst to other estimators
- With funding…
  - Incorporate Design Module
    - Optimize:

    search intervals vs spatial coverage vs searcher eff
  - Develop DWP Module
  - Other?

USGS

Summary



- Collaborative effort
- Peer reviewed
- Open source
- No decision regarding "which estimator"
- Unbiased*
- Comparability among estimates
- We can all see the same elephant!

USGS

* To the level any estimator can be unbiased when detection probability is unknown

61

THANK YOU!

Daniel Dalthorp	USBLM	Scott Darling
Juniper Simonis	USGS	Christy Humphrey
Lisa Madsen	DoE - NREL	Guillaume Marchais
Paul Rabie	WEST, Inc.	Dan Maynard
Jared Studyvin	Bat Conservation Int.	TJ Miller
Jeffrey Mintz	AWWI	Laura Nagy
Fränzi Korner-Nievergelt	Avangrid	Guillaume Péron
Robert Wolpert	Pattern Energy	Scott Pruitt
Taber Allison	~	Andrew Ryckman
Cris Hein	Ed Arnett	Michael Schirmacher
Jerry Roppe	Joana Bernardino	Linda Schuek
Dave Cowan	Aurélien Besnard	David Tidhar
Danna Small	Regina Bispo	Gillian Vallejo
Ian Evans	Mitch Craig	Sara Weaver
		Anonymous reviewers

Explore/Comment:

https://goo.gl/forms/bYKOWA5vELhOH8Id2