Estimators of wildlife fatality: a critical examination of methods

Manuela Huso, Oregon State University

November 5, 2008

Objectives in estimating mortality

Understand impacts of turbines

Effects on populations of animals, particularly rare and endangered

Explore relationship between

pre-construction activity and

post-construction mortality

Understand relationship between

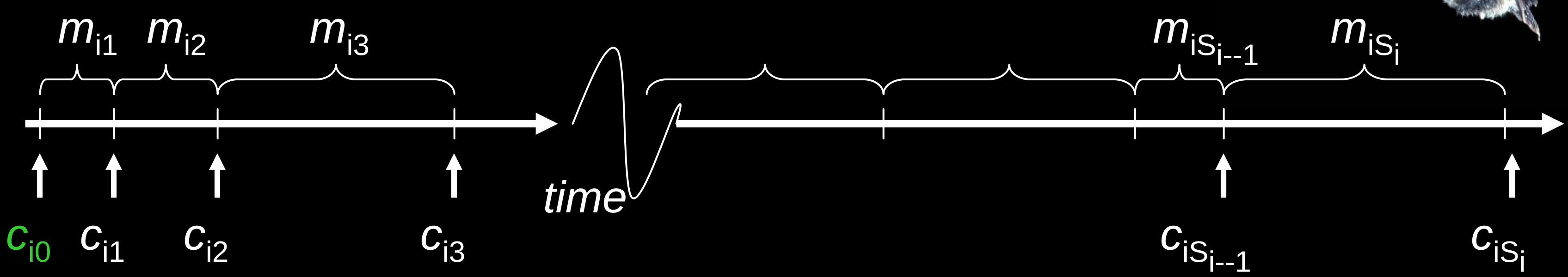
nightly activity, environmental conditions (wind speed, temperature,…) and mortality

Evaluate efficacy of curtailment, deterrents or other mitigation attempts

Overview

- Develop conceptual model of fatality
- Outline current protocol
- Review current estimators
- Propose new estimator
- Compare estimators
- Evaluate effect of precision of parameter estimates
- Discuss limitations and constraints
- Suggest next steps

Conceptual model of mortality



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

m_i = total number of dead animals at turbine *i*

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

Why use estimators instead of a direct count?

Because some fatalities won't be found and others will be scavenged before the observer visits the site.

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

Where:

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

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

Why use estimators instead of a direct count?

If the total number of carcasses we find on the site is not how many actually died…

how do we use how many we count to figure out how many were killed?

Conceptual model of carcass count

$$m_{ij} r_{ij} p_{ij} = c_{ij}$$

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

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

r_{ij} = **proportion** of animals that died in the interval (*j-1, j*] that **remain unscavenged**, i.e. observable

p_{ij} = **proportion** of animals that died in the interval (*j-1, j*] and remain unscavenged that are **actually observed**

Conceptual model of carcass count

$$20 * 0.75 * 0.33 = 5$$

$$m \quad * \quad r \quad * \quad p \quad = \quad c$$

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

r_{ij} = **proportion** of animals that died in the interval (*j-1, j*] that **remain unscavenged**, i.e. observable

p_{ij} = **proportion** of animals that died in the interval (*j-1, j*] and *remain unscavenged* that are **actually observed**

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

Conceptual model of mortality

To estimate mortality, *m*,

rearrange: $$m_{ij} r_{ij} p_{ij} = c_{ij}$$

to give: $$\hat{m}_{ij} = \frac{c_{ij}}{\hat{r}_{ij} \hat{p}_{ij}}$$

$$\mathbf{20 = \frac{5}{0.75*0.33}}$$

Conceptual model of mortality

$$\hat{m}_{ij} = \frac{c_{ij}}{\hat{r}_{ij}\,\hat{p}_{ij}}$$

Count : carcasses on the ground, *c*

Estimate : proportion remaining, *r*

: proportion detectable *p*

Does not depend on interval (very much ☺)

Depend on time since death, => search interval

Many estimators currently used

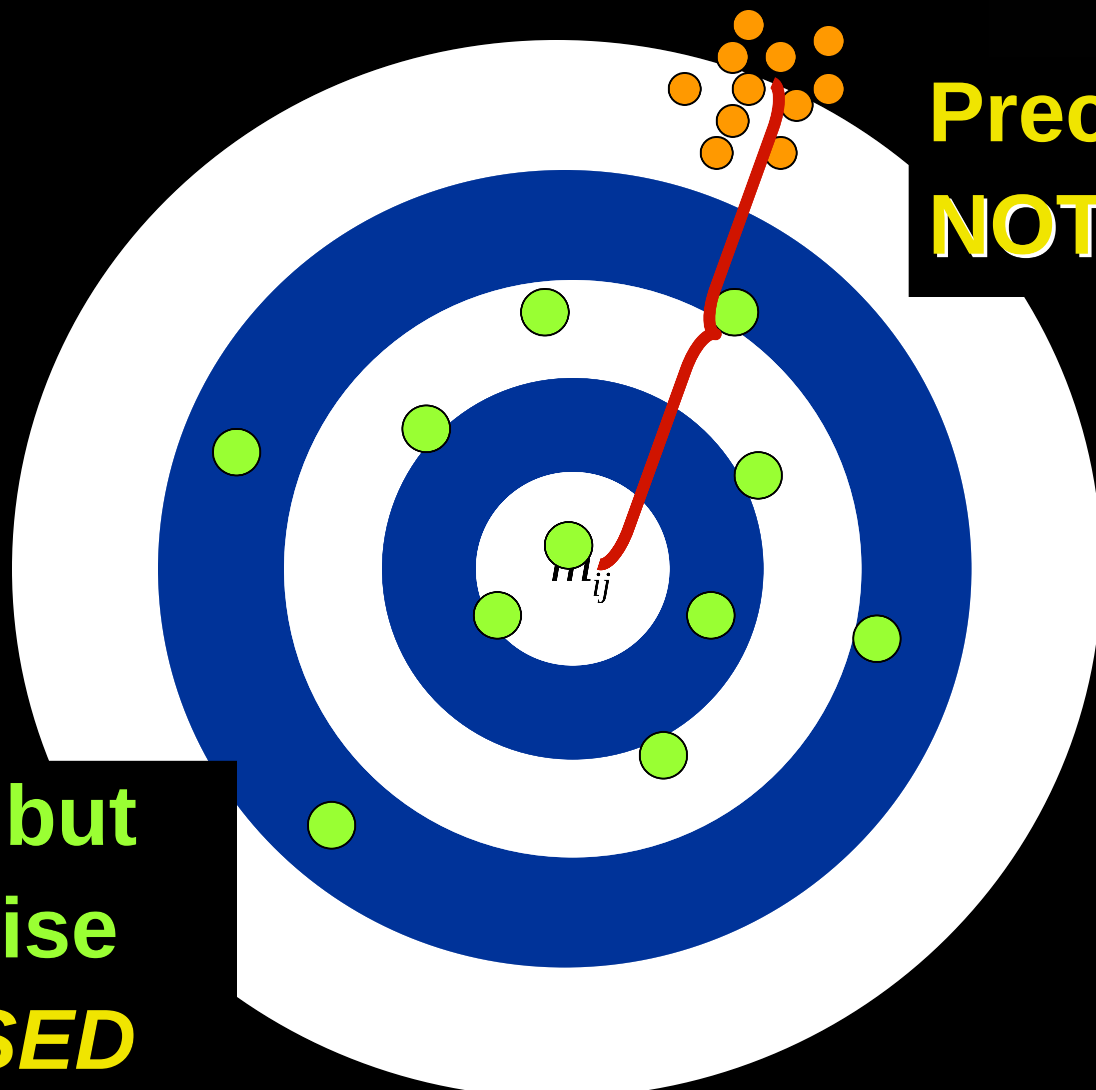
How were estimators derived?

Are they unbiased?

How precise are they?

Can they be improved?

Properties of a good estimator



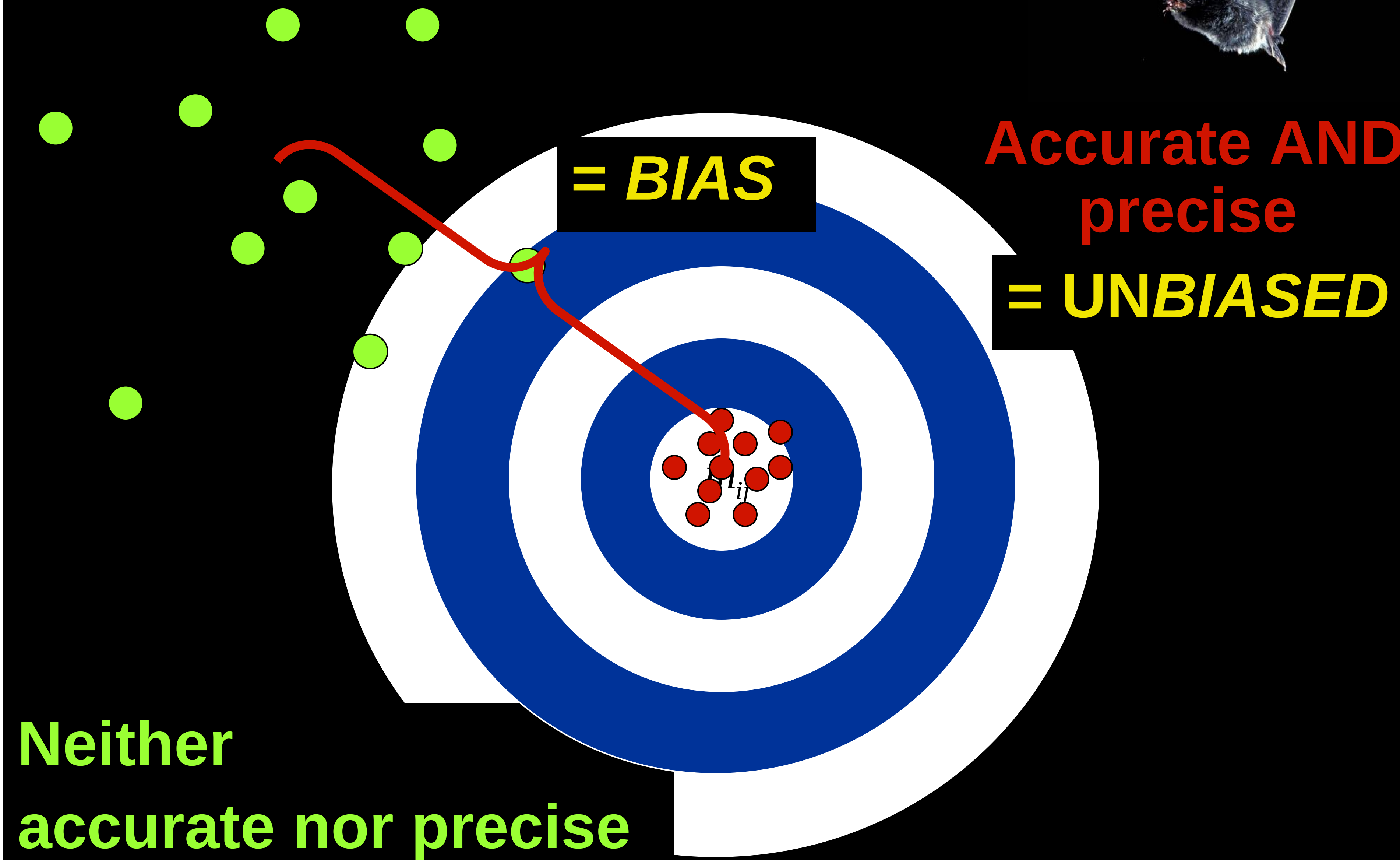
Precise but NOT accurate

= *BIAS*

Accurate but NOT precise

= UN*BIASED*

Properties of a good estimator



Many estimators currently used

Properties of a good estimator:

Why unbiased (accurate)?

To be right

To allow comparability across studies

Constant bias: no problem

Non-constant bias: big problem

Why precise?

To be useful

Many estimators currently used

Properties of a good estimator:

Accurate (unbiased)

Precise

When you sample the entire population, or there are no unknowns, the estimator should <u>equal</u> the true population parameter

$$m_i = \sum_{j=1}^{S_i} c_{ij}$$

Many estimators currently used

$$\hat{m}_{ij} = \frac{I_{ij} c_{ij}}{\hat{\bar{t}}_{ij} \hat{p}_{ij}}$$

$$\hat{m}_{ij} = \frac{I_{ij} c_{ij}}{\hat{\bar{t}}_{ij} \hat{p}_{ij}} \left(\frac{e^{I_{ij}/\hat{\bar{t}}_{ij}} - 1 + \hat{p}_{ij}}{e^{I_{ij}/\hat{\bar{t}}_{ij}} - 1} \right)$$

$$\hat{m}_{ij} = \frac{c_{ij}}{\hat{s}_i \hat{p}_{ij}}$$

m_{ij} = *number* of dead animals at the i^{th} turbine in the j^{th} interval

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

p_{ij} = *proportion* of carcasses likely to be observable at the i^{th} turbine in the j^{th} interval

t_{ij} = *average number* of days a carcass will persist uscavenged at i^{th} turbine in the j^{th} interval. (Usually t_{ij} is considered the same for all turbines, but not all intervals)

I_{ij} = length (*number* of days) of the j^{th} interval at the i^{th} turbine.

s_{ij} = proportion of carcasses that are expected to persist uscavenged at i^{th} turbine in a fixed interval of (usually) 3 days

Many estimators currently used

Properties of a good estimator:

- Accurate (unbiased)
- Precise

Perfect conditions (no unkowns):

- There are no scavengers
- All carcasses are observed

$$m_i = \sum_{j=1}^{S_i} c_{ij}$$

Simple Evaluation of First Estimator

$$\hat{m}_{ij} = \frac{I_{ij} c_{ij}}{\hat{\bar{t}}_{ij}\,\hat{p}_{ij}} = \frac{c_{ij}}{(\hat{\bar{t}}_{ij} / I_{ij})\,\hat{p}_{ij}}$$

1

All carcasses are observed

t/*I* is not a proportion, i.e. **t**/*I* ≠ *r*

$$\hat{m}_{ij} = \frac{c_{ij}}{\hat{r}_{ij}\,\hat{p}_{ij}}$$

Simple Evaluation of First Estimator

$$\hat{m}_{ij} = \frac{I_{ij} c_{ij}}{\hat{\bar{t}}_{ij}\, \hat{p}_{ij}} = \frac{c_{ij}}{(\hat{\bar{t}}_{ij} / I_{ij})\, \hat{p}_{ij}}$$

1

All carcasses are observed

t/*I* is not a proportion, i.e. **t/*I*** ≠ *r*

If $\bar{t}$ = very large number, much larger than I,

(as when there are no scavengers)

then **t/*I*** > 1 and $m_{ij} < c_{ij}$!

We estimate that fewer died than we actually have in our hands!

Analytic Evaluation of Second Estimator

$$\hat{m} = \frac{I * c}{\hat{\bar{t}} * \hat{p}} \left(\frac{e^{I/\hat{\bar{t}}} - 1 + \hat{p}}{e^{I/\hat{\bar{t}}} - 1} \right)$$

What if searchers observe very few?

As searcher efficiency, *p,* becomes smaller, rather than compensating for it, the estimate of *m* approaches that of the previous estimator.

Proposed estimator

$$\hat{m}_{ij} = \frac{c_{ij}}{\hat{r}_{ij}\,\hat{p}_{ij}\,\hat{e}_{ij}}$$

Estimated number of mortalities

Proposed estimator

$$\hat{m}_{ij} = \frac{c_{ij}}{\hat{r}_{ij}\,\hat{p}_{ij}\,\hat{e}_{ij}}$$

Number of carcasses observed

Proposed estimator

$$\hat{m}_{ij} = \frac{c_{ij}}{\hat{r}_{ij}\,\hat{p}_{ij}\,\hat{e}_{ij}}$$

Estimated proportion of carcasses remaining unscavenged through the interval

A function of the search interval

Proposed estimator

$$\hat{m}_{ij} = \frac{c_{ij}}{\hat{r}_{ij}\,\hat{p}_{ij}\,\hat{e}_{ij}}$$

Estimated proportion of carcasses observed

Proposed estimator

$$\hat{m}_{ij} = \frac{c_{ij}}{\hat{r}_{ij}\,\hat{p}_{ij}\,\hat{e}_{ij}}$$

Effective interval

A function of the search interval and average persistence time, $\bar{t}$

$$\text{If } \; f(x) = \lambda e^{-\lambda x}; \; S(x) = e^{-\lambda x}$$

$$\hat{\lambda} = 1/\bar{t}; \;\; \hat{r} \approx S(I/2) = e^{-I/(2\bar{t})}$$

$$\hat{e} = \frac{\min(\, x : S(x) = 0.01, I\,)}{I}$$

Effective interval

If *average persistence time*, $\hat{\bar{t}}$, is *short*

e.g. 2 d, then any carcass will almost certainly be removed by ~9 d = t_{99}

t_{99}= number of days beyond which only 1% will remain

If *search interval*, I, is *long*

e.g. 28 d, then we have essentially no chance of counting any animals that died in the first 18-19 d

Our *effective interval* is

$$\hat{e}_{ij} = \frac{\min(t_{99}, I)}{I}$$

e.g. 9 days ≈ 1/3 of the actual interval, so $\hat{e}$=0.33

Proposed estimator

$$\hat{m}_{ij} = \frac{c_{ij}}{1 \;\; 1 \;\; 1}$$

Each of these is a proportion, < 1

Expansion (1/rpe) will be >=1

If $\hat{r}$=1 (then $\hat{e}$ = 1) and $\hat{p}$=1

then $\hat{m}_{ij} = c_{ij}$

Compare Estimators through simulation

$$\hat{m}_{ij} = \frac{I_{ij} c_{ij}}{\hat{\bar{t}}_{ij} \hat{p}_{ij}}$$

$$\hat{m}_{ij} = \frac{I_{ij} c_{ij}}{\hat{\bar{t}}_{ij} \hat{p}_{ij}} \left(\frac{e^{I_{ij}/\hat{\bar{t}}_{ij}} - 1 + \hat{p}_{ij}}{e^{I_{ij}/\hat{\bar{t}}_{ij}} - 1} \right)$$

$$\hat{m}_{ij} = \frac{c_{ij}}{\hat{r}_{ij} \hat{p}_{ij} \hat{e}_{ij}}$$

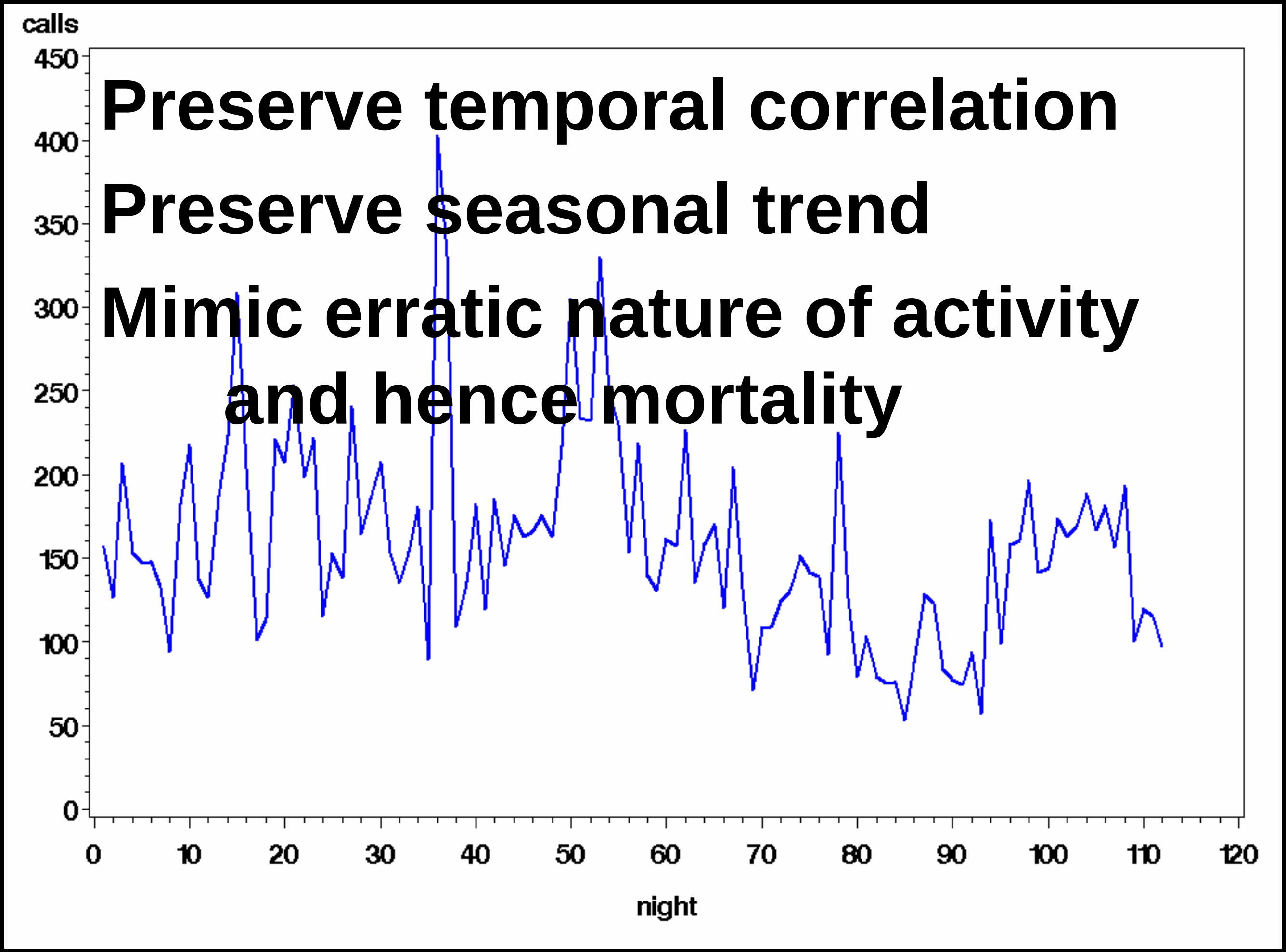
Are these estimators of mortality accurate under realistic conditions?

What search interval is optimal to achieve acceptable precision?

What trial sample size is necessary for precise and accurate estimation of searcher efficiency and carcass persistence?

Simulation: First simulate count

Started with activity data from PA



Simulation: First simulate count

Simulate number dead (average 10% of activity)
at each tower on
each night

For each carcass, varied 3 factors:

- Average Carcass Persistence:
  7 levels: $\bar{t}$ = 1, 2, 4, 8, 16, 32, 64 days
- Searcher efficiency:
  3 levels: p = 20%, 50% and 80%
- Search interval: constant with
  6 levels: I = 1, 2, 4, 7, 14, 28 days

Have one ‘known’ total number dead and 126 (7*3*6) different simulated *counts* of observed carcasses

Simulation: Estimate mortality

To each *count*, applied the three estimators using <u>known</u> values of searcher efficiency, *p* and average carcass persistence time, $\bar{t}$

Know how many died, and

know how many were counted.

Compare estimates from 3 estimators to see which one does the best job given this count.

Simulation: Estimate effect of precision of $\hat{p}$ and $\hat{\bar{t}}$

Simulated searcher efficiency <u>trials</u>

with n=10, 25, 50, and 100 carcasses

Simulated carcass persistence <u>trials</u>

with n=10, 25, 50, and 100 carcasses

To each count (same as earlier), applied the 3 estimators

using the same <u>estimated</u> searcher efficiency, $\hat{p}$
and <u>estimated</u> average carcass persistence time, $\hat{\bar{t}}$,
based on <u>trials</u> that used n=10, 25, 50 or 100

Simulation:

Compared estimators using

Percent Relative Bias:

$$PRB = \left(\frac{\hat{m} - m}{m} \right) * 100$$

Multiplying factor: the number by which you need to multiply the estimate to give actual mortality

$$Factor = \left(\frac{1}{1 + PRB/100} \right) = \frac{m}{\hat{m}}$$

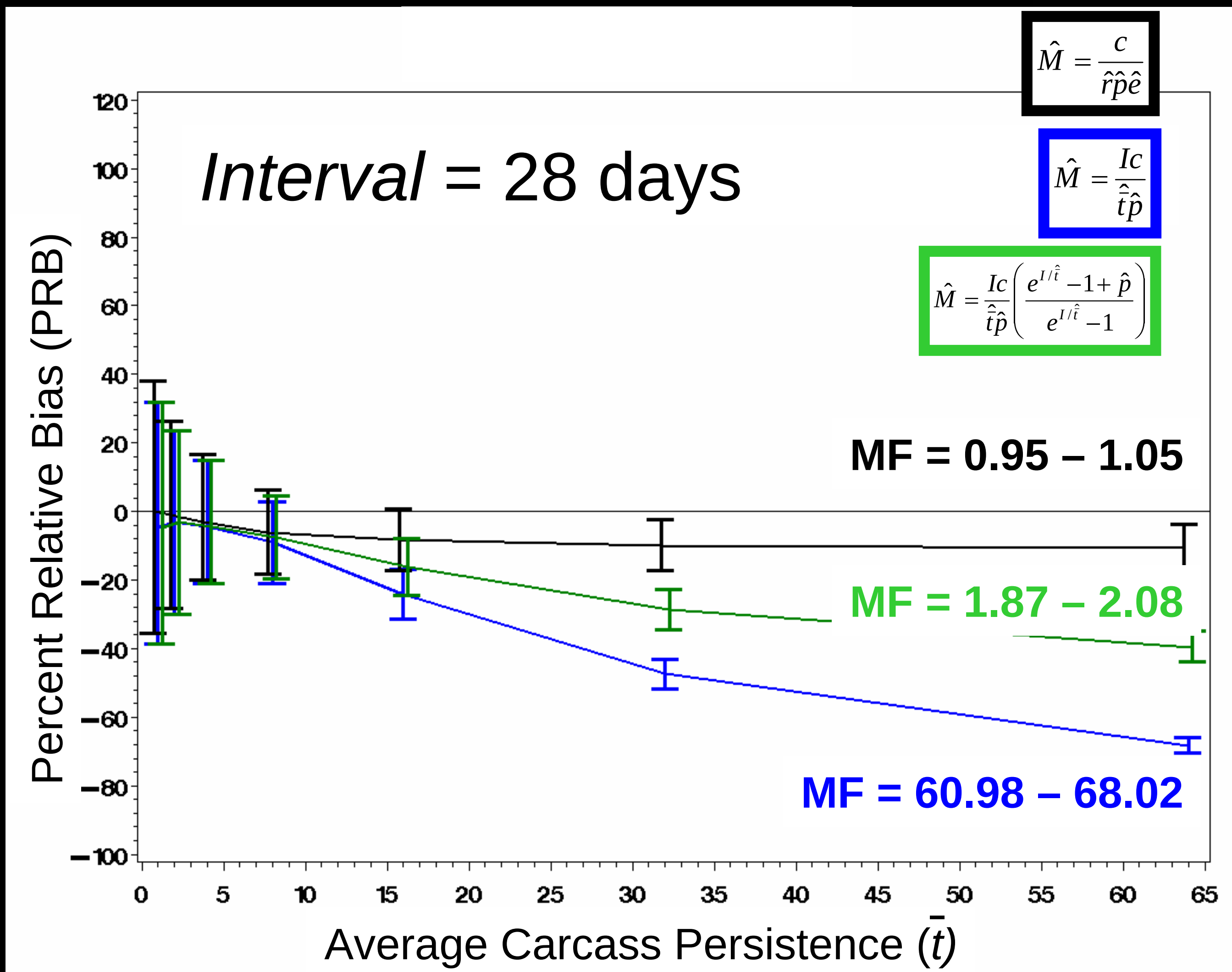
Simulation:

Repeated **1000** times

Calculated mean, 2.5th, and 97.5th quantiles of PRB and MF to form empirical 95% confidence interval

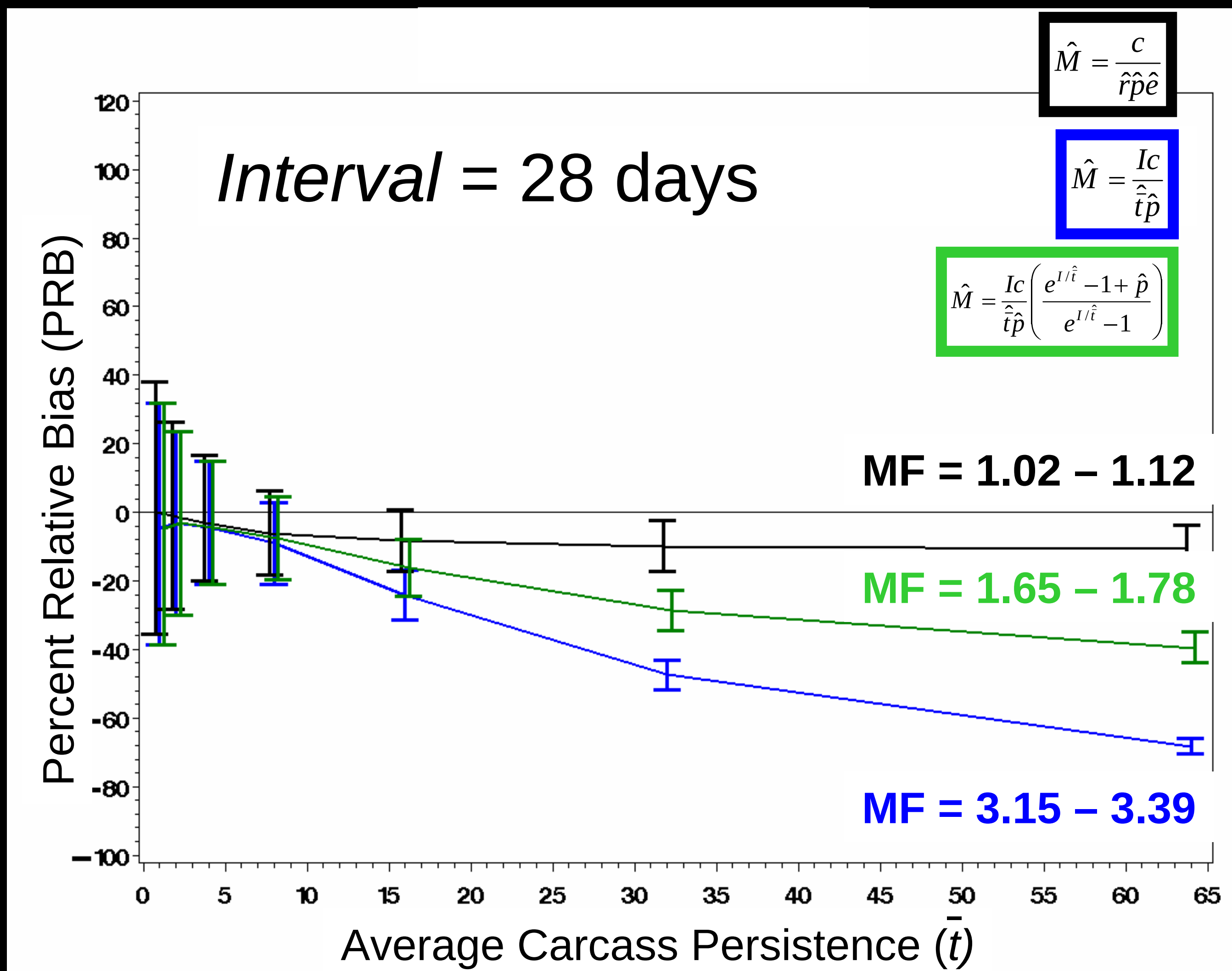
Simulation:

Results: SE=50%, parameters known



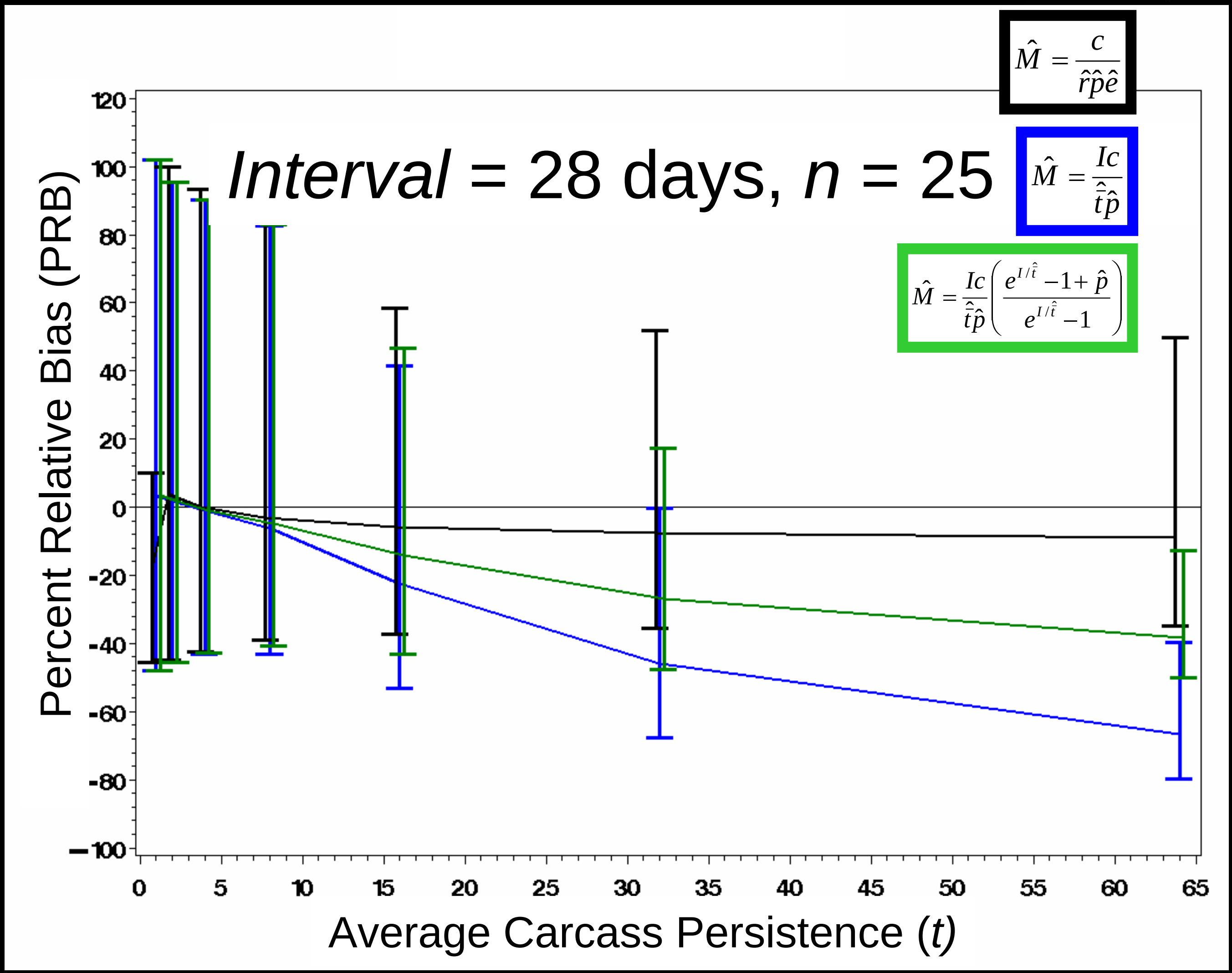
Simulation:

Results: SE=50%, parameters known



Simulation:

Results: SE = 50%, trial *n*=25



Precision

$$\hat{m}_{ij} = \frac{c_{ij}}{\hat{r}_{ij}\hat{p}_{ij}\hat{e}_{ij}}$$

Anything that reduces precision in variables in the equation will reduce confidence in the overall estimate

Expansion factors for searcher efficiency: 1/p

Average Searcher Efficiency	Trial Size	Lower CI	Upper CI	
20%	5	1.40	19.23	5
	10	1.80	15.15	
	25	2.46	10.75	
	50	2.97	8.70	
	100	3.44	7.46	
	Trial Size	Lower CI	Upper CI	
50%	5	1.17	6.85	2
	10	1.23	3.82	
	25	1.46	3.19	
	50	1.55	2.67	
	100	1.66	2.45	
	Trial Size	Lower CI	Upper CI	
80%	5	1.01	2.09	1.25
	10	1.03	1.81	
	25	1.07	1.56	
	50	1.11	1.46	
	100	1.15	1.39	

Expansion factors for carcass removal: 1/r

Search Interval = 1

Average Carcass Persistence Time	Trial Size	Lower CI	Upper CI	
4 DAYS	5	0.98	1.31	1.133
	10	1.04	1.24	
	25	1.08	1.19	
	50	1.09	1.17	
	100	1.11	1.16	
16 DAYS	5	1.00	1.07	1.032
	10	1.01	1.05	
	25	1.02	1.05	
	50	1.02	1.04	
	100	1.03	1.04	
32 DAYS	5	1.00	1.03	1.015
	10	1.00	1.03	
	25	1.01	1.02	
	50	1.01	1.02	
	100	1.01	1.02	

Expansion factors for carcass removal: 1/r

Search Interval = 14

Average Carcass Persistence Time

Average Carcass Persistence Time	Trial Size	Lower CI	Upper CI	
4 DAYS	5	0.77	43.03	5.75
	10	1.68	19.75	
	25	2.80	11.83	
	50	3.50	9.46	
	100	4.07	8.14	
16 DAYS	5	0.94	2.56	1.55
	10	1.14	2.11	
	25	1.29	1.85	
	50	1.37	1.75	
	100	1.42	1.69	
32 DAYS	5	0.97	1.60	1.24
	10	1.07	1.45	
	25	1.14	1.36	
	50	1.17	1.32	
	100	1.19	1.30	

Expansion factors for carcass removal: 1/r

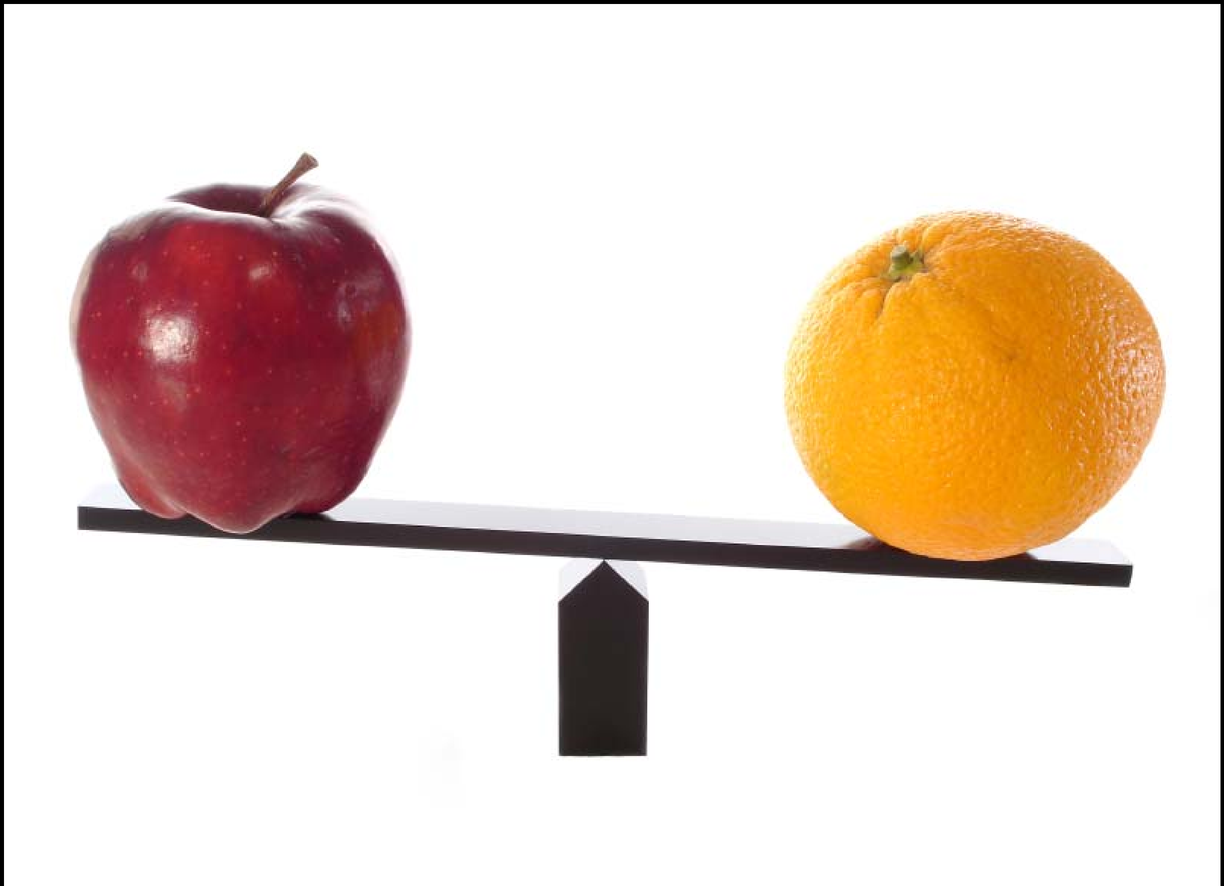
Search Interval = 28

Average Carcass Persistence Time

	Trial Size	Lower CI	Upper CI	
4 DAYS	5	0.59	1851.21	
	10	2.81	389.98	
	25	7.83	140.00	33.1
	50	12.25	89.50	
	100	16.54	66.31	
	Trial Size	**Lower CI**	**Upper CI**	
16 DAYS	5	0.88	6.56	
	10	1.29	4.44	
	25	1.67	3.44	2.40
	50	1.87	3.08	
	100	2.02	2.85	
	Trial Size	**Lower CI**	**Upper CI**	
32 DAYS	5	0.94	2.56	
	10	1.14	2.11	
	25	1.29	1.85	1. 5
	50	1.37	1.75	
	100	1.42	1.69	

Effect of imprecise estimates of p and r

Balance cost of effort
with accuracy and precision



Search interval (*I*) can be selected in design

Searcher efficiency (*p*) cannot be controlled, but it can be improved (training, vegetation control, selection of sites) AND precision in estimate can be increased through selection of sites (reduced heterogeneity) and approaches used for searcher efficiency trials

Carcass persistence (*r*) cannot be controlled, but precision in estimate can be increased by approaches used for estimating scavenging rate

Summary

Contributions of proposed estimator:

Distributional assumption of carcass persistence allows estimate of proportion remaining, *r*, for <u>any</u> interval

Definition of 'effective' interval allows

correction when $I >> \hat{\underline{t}}$

Summary

- Commonly used estimators can be highly biased under realistic conditions
- Bias is not constant, but differs with carcass persistence and search interval as well as searcher efficiency, **making direct comparison across sites with different conditions meaningless**
- None of the estimators is unbiased under all realistic conditions, but we **can control bias by designing shorter search intervals**
- Can control variability of estimates by estimating *r* and *p* using **large trial sample sizes**

Practical Issues

- Use unbiased estimator
- Use shorter search intervals, max 7 -10 days
- Use adequate sample size in searcher efficiency and carcass persistence trials

  At least 50 carcasses per factor,

  e.g. (large, spring) or (small, winter)
- Bootstrap variance based on resampling carcass persistence and searcher efficiency trials, **only IF the trial sample size is adequate**

Next steps

- Refine estimator?
- Derive variance of estimator (Birkes)
- Develop more efficient models of
  p and *r* using covariates
- Develop software for estimator and bootstrapped CI and make publicly available
- Further research into design: tradeoffs of # turbines vs optimal search interval

Thank you!

Questions? Comments?