Altamont Simulations

Julie Yee
May 30 – June 7, 2011

Purpose: Demonstrate the ability to estimate fatality rates based on data collected from the Monitoring Program including its double-searcher study design (Detection Probability Monitoring, M53) with a measurable level of accuracy. This is a simulation demonstration which allows for variation due to

1. Irregular search intervals. Most intervals vary between 30-40 days long. The exact number of strings searched each day is subject to variation. Moreover, searches occur 4 days per week (Mon-Thu), which create opportunities for 4-day swings in the date in which carcasses near the end or beginning of the week can be detected.

2. Temporal and spatial heterogeneity in
    - Bird fatality rates
    - Scavenger removal rates
    - Searcher efficiency rates

Other sources of variation not addressed in these simulations include: variation in searcher efficiency among searchers, and variation in fatality rates among turbine type. These can be incorporated later if needed.

Contents: The program is written using the open source R programming language to simulate 3 main components of the Monitoring Program:

1. Turbine Sampling – simulating variable searching effort

Simulate a set of search dates representative of the daily search pattern for the monitoring study.

Algorithm:

1) Simulations are generated according to the following inputs:
   a) `NPrimary` = # of turbines to be searched.
   b) `AvgTurbsPerStr` = Average # of turbines per string.
   c) `AvgStrPerBlob` = Average # of strings per blob.
   d) `AvgStrsPerDay` = Average # of strings searched per day by a team.
   e) `NRots` = # of rotations that turbines will be searched
   f) `NSSGrps` = # of rotations needed for all primary strings to be secondarily searched
   g) `PPSProp` = Proportion of searched turbines which have pre- and post-searches.
2) Generate a sample of individually numbered turbines. Distribute these turbines into various sizes of strings, and distribute the strings into various sizes of blobs.
3) Generate a primary search schedule for strings based on searching different numbers (randomly generated) of whole strings daily. Assume strings are searched in a fixed order.
4) Extend the search schedule for a pre-specified number of rotations.
5) Generate a secondary search schedule as follows:
   a) For the first rotation, randomly select blobs to have a secondary search.
   b) For each blob selected, all strings are secondarily searched on random dates.
   c) The numbers of days between primary and secondary searches are randomly generated to average ~6 days and range 1-21 days.
   d) Repeat for subsequent rotations, until all blobs are secondarily searched.
   e) Repeat the secondary searches starting with the same set of blobs selected for the first rotation, and continue throughout the remaining rotations. The string search order for secondary searches within each blob is randomized every rotation, unlike the primary searches in which the search order is repeated each rotation.
6) For each rotation, randomly select a set of strings to have pre- and post-searches.
   a) The numbers of days between pre- and primary searches are randomly generated to average ~2 days and range 1-3 days.
   b) The numbers of days between secondary and post-searches are randomly generated to average ~3 days and range 1-6 days.

Example of simulated search dates:

When using the following inputs:

```
> NPrimary <- 1300
> AvgTurbsPerStr <- 8
> AvgStrPerBlob <- 6
> AvgStrsPerDay <- 8
> NRots <- 10
> NSSGrps <- 4
> PPSProp <- 0.15
```

Then most search intervals fall between 30-40 days. The search dates are stored in arrays such as the one below, for string #125. The 10 rows represent the 10 rotations, and the 4 columns represent Pre, Primary, Secondary, and Post search dates respectively. Each rotation, 1/NSSGrps = 25% of strings receive a secondary search. As a result, there is a secondary search every four rotations for each string. This particular string has a secondary search during the 3rd and 7th rotations, and pre- and post-searches in the 3rd rotation.

```
> QAQCArray[125,,]

      [,1] [,2] [,3] [,4]
 [1,]   NA   29   NA   NA
 [2,]   NA   58   NA   NA
 [3,]   90   92  101  103
 [4,]   NA  129   NA   NA
 [5,]   NA  158   NA   NA
 [6,]   NA  205   NA   NA
 [7,]   NA  241  259   NA
 [8,]   NA  275   NA   NA
 [9,]   NA  309   NA   NA
[10,]   NA  344   NA   NA
```

2. Detection Probability Study – simulating detection outcomes

This section addresses the unknown probability of detecting a carcass that has been deposited in the Altamont. There are two general types of carcass deposits – those that are experimentally placed by humans and those that are placed naturally regardless of the cause of fatality which could include turbine strikes, predation,…etc. For convenience, I will refer to these as *experimental* and *natural placements* respectively. Experimental placements are primarily intended to study detection probability, however the potential for double-searcher detection data on naturally placed carcasses to provide additional information on detection probability will also be considered.

<u>Experimentally placed carcasses - algorithm</u>

1) Generate a sample of individually numbered carcasses to deposit at pre-searched turbines.
2) Since not all carcasses are obtained fresh, then allow for fresh and non-fresh carcasses by assigning each one to a randomly generated age, with a minimum of 0 days, mean of 2 days, variance of 4 days, and no maximum.
3) Randomly distribute the carcasses among pre-searched turbines at pre-search dates, and simulate their fates at every search (primary or secondary) as follows:
    a) Randomly generate the status of whether or not the carcass is remaining. Apply a Bernoulli random number generator to generate the outcomes 0 (removed) or 1 (remaining) where the probability of remaining varies as a function of carcass age, rotation and blob, which are proxies for time and space.
    b) The function defining the probability of remaining is assumed to be based on the Smallwood 2007 model: $R = \left(a_r + b_r \log(age+1)\right)/100$, where *age* is age in days, and the coefficients a_r and b_r are inputs (currently set at 115 and -20 respectively, roughly in the range of values reported in Table 4 of Smallwood 2007). When the function exceeds 1 or dips below 0, then R is truncated to remain constrained between 0-1.
    c) Spatial and temporal variations are incorporated into R by including random effects for rotation and blob. Random effects are normally distributed with mean 0, and which perturb R up or down depending on whether the effect is <0 or >0. For a blob with a random effect r_{blob} during a rotation with a random effect r_{rot}, then modify the function by taking R to the power of $\exp\left(r_{blob} + r_{rot}\right)$.
    d) For a carcass that is remaining at a string being searched, randomly generate the outcome of whether or not the carcass is detected. Analogous to part a, apply a Bernoulli random number generator to generate the outcomes 0 (undetected) or 1 (detected) where the searcher efficiency varies as a function of carcass age, rotation and blob. Assume primary and secondary searchers are blind to each others' findings, and that they share the same searcher efficiency.
    e) The function defining the searcher efficiency is assumed to be a logistic function: $D = \dfrac{\exp\left(a_d + b_d \times age + d_{blob} + d_{rot}\right)}{1+\exp\left(a_d + b_d \times age + d_{blob} + d_{rot}\right)}$, where a_d and b_d are coefficients and d_{blob}

and d_{rot} are normally distributed random effects on detection, analogous to those described in parts b and c.

4) Generate post-search detection outcomes the same way, except assume a higher post-search searcher efficiency because the searcher is already aware of the carcass:
$D = \frac{\exp(a_d + b_d \times age + d_{blob} + d_{rot} + d_s)}{1 + \exp(a_d + b_d \times age + d_{blob} + d_{rot} + d_s)}$, where d_s is >0.

5) Assuming carcasses are not removed during the post-search, then they could remain available for detection during subsequent rotations. We might see fit to use detection data from later rotations, especially when they are detected by searchers who were unaware of them previously.

Note: My simulations currently assume all primary and secondary searches to be blindly conducted (i.e. without pre-existing knowledge that a carcass exists). In real life, this assumption is likely violated when searchers repeatedly detect the same carcass across multiple rotations. I can modify the simulations to apply an inflated efficiency rate to repeat searchers for those carcasses they have previously detected, much like how the supervisor efficiency is inflated.

Naturally placed carcasses - algorithm

Naturally placed carcasses are simulated using similar steps as the experimental placements with the following exceptions:

- Modify steps 1) and 3) by generating a set of individually numbered carcasses, separate from the experimental carcasses, to deposit *at any of the sampled turbines on any date*.
- Modify step 2) by allowing each carcass age to be 0 days on the date it is placed.
- Modify step 4) by treating pre-search and post-search detections in the same way as primary and secondary detections, because supervisors should start out blind to any carcasses that they did not naturally place on their own.

Figures of carcass removal functions and searcher detection efficiency functions

The following coefficients are used for the probability functions for carcass remaining (**`a.rem`**, **`b.rem`**) and searcher detection efficiency (**`a.det`**, **`b.det`**). Normal random effects with mean 0 and standard deviation 0.25 are generated for **`NRots`** number of rotations and **`NBlobs`** number of blobs.

```
> a.rem <- 115
> b.rem <- -20
> a.det <- 0.5
> b.det <- -0.01
> RanEffRotRem <- rnorm(NRots,0,0.25)
> RanEffBlobRem <- rnorm(NBlobs,0,0.25)
> RanEffRotDet <- rnorm(NRots,0,0.25)
> RanEffBlobDet <- rnorm(NBlobs,0,0.25)
```

Figure 1. Probability of carcass remaining (a) overlaid with variations across time (b), space (c), or time and space (d),

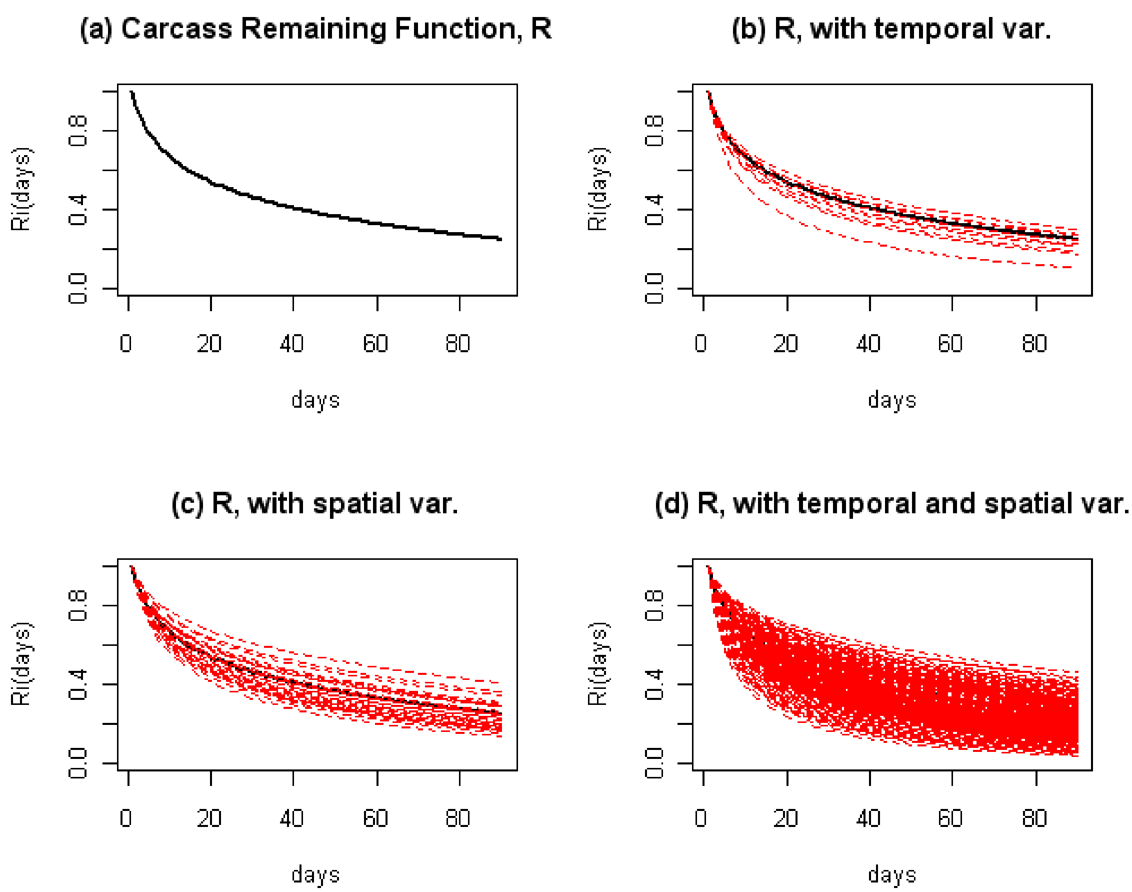
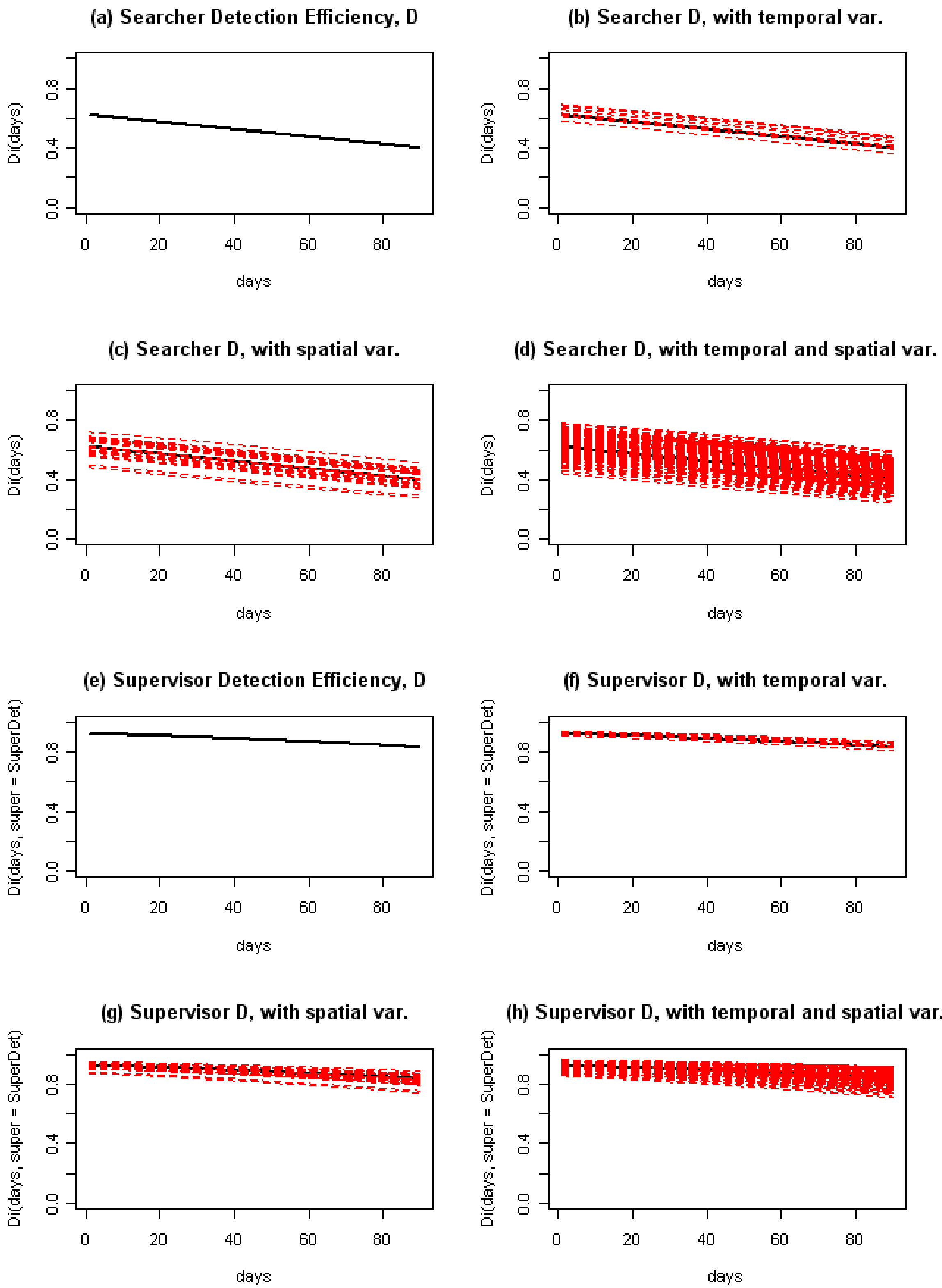


Figure 2. Searcher detection efficiency overlaid with temporal and spatial variations (a-d) for blind searchers, and supervisor detection efficiency with variations (e-f).



Examples of simulated records of detection outcomes

The simulated fates of carcasses are stored in arrays such as the ones below. These arrays show the presence/absence status and detection outcome for each carcass at each search after it was placed. Each row represents a search date or a placement date, beginning with the placement date. Note that the detection outcome is "NA" for experimental carcasses placed on a pre-search date, because confirmation of its presence does not count as a detection outcome.

The first column is just an index that can be ignored. The next six columns are string number, rotation number, survey type (1=Pre, 2=Primary, 3=Secondary, 4=Post), survey date, blob number, and fatality date. The next three columns are probabilities: `PrRem` is the probability *R* of the carcass remaining based on its age and location and the rotation in which it is searched. `CondPrRem` is the probability of the carcass remaining, conditional on that it was still remaining as of the previous search. `CondPrDet` is the probability that the searcher detects the carcass, conditional on that it has remained at its location. Note that the probability of detection is higher during the post-search because the supervisor had first-hand knowledge of this carcass' location. The last two columns contain the presence and detection outcomes for this carcass at each search, and these values are sequentially generated randomly from the conditional probabilities.

<u>*Example 1:*</u> *Carcass undetected by primary and secondary searchers, although confirmed by post-searcher to have been present*

This array shows Experimental Carcass #174 which was placed on String #134 during its 4th rotation pre-search on day 147. It was fresh at the time it was placed, and it remained on the string for at least 9 days. During this time, it went undetected by primary and secondary searchers although it was confirmed to remain present by the post-searcher. The carcass was removed before the primary searchers could come around again in the following rotation.

```
> Fates[[174]]
[[1]]
     Str Rot SType SDate Blob FDate      PrRem CondPrRem CondPrDet Present Detected
5333 134   4 1-Pre   147   22   147 1.00000000        NA        NA       1       NA
5334 134   4 2-Pri   148   22   147 1.00000000 1.0000000 0.6699445       1        0
5335 134   4 3-Sec   155   22   147 0.58105774 0.5810577 0.6542863       1        0
5336 134   4 4-Pos   156   22   147 0.55392003 0.9532960 0.9326379       1        1
5338 134   5 2-Pri   186   22   147 0.24463915 0.4416507 0.5812566       0        0
5342 134   6 2-Pri   221   22   147 0.13724236 0.5609992 0.4944838       0        0
5346 134   7 2-Pri   260   22   147 0.07923885 0.5773644 0.3984169       0        0
5350 134   8 2-Pri   298   22   147 0.04662840 0.5884538 0.3117253       0        0
5351 134   8 3-Sec   305   22   147 0.04211940 0.9032993 0.2969079       0        0
5354 134   9 2-Pri   331   22   147 0.02824540 0.6706030 0.2456283       0        0
5358 134  10 2-Pri   365   22   147 0.01535723 0.5437072 0.1881516       0        0
```

<u>Example 2:</u>   Carcass detected by primary searchers, but unconfirmed by post-searchers

```
> Fates[[158]]
[[1]]
     Str Rot SType SDate Blob FDate      PrRem CondPrRem  CondPrDet Present Detected
4969 125   3 1-Pre    90   21    88 0.93712143        NA         NA       1       NA
4970 125   3 2-Pri    92   21    88 0.84410548 0.9007429 0.50905627       1        1
4971 125   3 3-Sec   101   21    88 0.65286372 0.7734386 0.48656050       1        0
4972 125   3 4-Pos   103   21    88 0.62762718 0.9613449 0.87283144       1        0
4974 125   4 2-Pri   129   21    88 0.44138527 0.7032603 0.41732338       1        1
4978 125   5 2-Pri   158   21    88 0.33638684 0.7621161 0.34892429       1        0
4982 125   6 2-Pri   205   21    88 0.23108039 0.6869484 0.25090885       1        1
4986 125   7 2-Pri   241   21    88 0.17375411 0.7519206 0.18942189       0        0
4987 125   7 3-Sec   259   21    88 0.14934730 0.8595325 0.16331443       0        0
4990 125   8 2-Pri   275   21    88 0.12937834 0.8662918 0.14261102       0        0
4994 125   9 2-Pri   309   21    88 0.09103995 0.7036723 0.10585753       0        0
4998 125  10 2-Pri   344   21    88 0.05567220 0.6115139 0.07700376       0        0
```

In this example, Experimental Carcass #158 was 2 days old when it was placed during the 3rd rotation on String #125 on day 90. Although this carcass existed on days 88 and 89, it would have been unavailable for detection even if there had been searchers on that string on those days. It was detected by primary searchers on days 92, 129, and 205, but missed by primary searches on day 158 and completely missed by both the secondary and post-searches on days 101 and 103. The carcass disappeared before day 241.

<u>Example 3:</u>   Carcass undetected by primary and secondary searchers, and unconfirmed by post-searchers

```
> Fates[[187]]
[[1]]
     Str Rot SType SDate Blob FDate      PrRem CondPrRem CondPrDet Present Detected
6017 151   5 1-Pre   189   25   188 1.00000000        NA        NA       1       NA
6018 151   5 2-Pri   191   25   188 0.84111912 0.8411191 0.7210556       1        0
6019 151   5 3-Sec   200   25   188 0.56369456 0.6701721 0.7025987       0        0
6020 151   5 4-Pos   205   25   188 0.49152566 0.8719716 0.9431979       0        0
6022 151   6 2-Pri   227   25   188 0.32417770 0.6595336 0.6432971       0        0
6026 151   7 2-Pri   262   25   188 0.20414441 0.6297299 0.5596406       0        0
6030 151   8 2-Pri   303   25   188 0.12868569 0.6303660 0.4575285       0        0
6033 151   9 1-Pre   335   25   188 0.09010525 0.7001963 0.8190173       0        0
6034 151   9 2-Pri   337   25   188 0.08806791 0.9773893 0.3751240       0        0
6035 151   9 3-Sec   342   25   188 0.08313343 0.9439696 0.3634789       0        0
6036 151   9 4-Pos   347   25   188 0.07841442 0.9432356 0.8005449       0        0
6038 151  10 2-Pri   368   25   188 0.06067204 0.7737358 0.3057007       0        0
```

In this example, Carcass #187 was 1 day old at the time it was placed during the 5th rotation on String #151 on day 189. It remained present at that string at least 2 days until day 191 when the string was searched by primary searchers, but it went undetected. The carcass was removed before any other searches could occur there.

3. Estimating the number of fatalities

Data collected from the detection study can be used to derive the detection probability function, which in turn can be used to derive an estimate of the number of fatalities placed in the Altamont. Although we know what the detection probability function was that generated these simulated data, we will not have this knowledge with real life data. We need the ability to estimate detection rates, and equivalently fatality rates, with a reasonable accuracy by analyzing the data. In this section, we test this ability in our simulations by performing a statistical analysis based only on the observable data, which includes detection events but not true presence/absence status or age. For example, here are the observable data for the three experimental carcasses from the preceding examples:

Example 1: Carcass #174

	Str	Rot	SType	SDate	Blob	Detected
5333	134	4	1-Pre	147	22	NA
5334	134	4	2-Pri	148	22	0
5335	134	4	3-Sec	155	22	0
5336	**134**	**4**	**4-Pos**	**156**	**22**	**1**
5338	**134**	**5**	**2-Pri**	**186**	**22**	**0**
5342	**134**	**6**	**2-Pri**	**221**	**22**	**0**
5346	**134**	**7**	**2-Pri**	**260**	**22**	**0**
5350	**134**	**8**	**2-Pri**	**298**	**22**	**0**
5351	**134**	**8**	**3-Sec**	**305**	**22**	**0**
5354	**134**	**9**	**2-Pri**	**331**	**22**	**0**
5358	**134**	**10**	**2-Pri**	**365**	**22**	**0**

Example 2: Carcass #158:

	Str	Rot	SType	SDate	Blob	Detected
4969	125	3	1-Pre	90	21	NA
4970	**125**	**3**	**2-Pri**	**92**	**21**	**1**
4971	**125**	**3**	**3-Sec**	**101**	**21**	**0**
4972	**125**	**3**	**4-Pos**	**103**	**21**	**0**
4974	**125**	**4**	**2-Pri**	**129**	**21**	**1**
4978	**125**	**5**	**2-Pri**	**158**	**21**	**0**
4982	**125**	**6**	**2-Pri**	**205**	**21**	**1**
4986	**125**	**7**	**2-Pri**	**241**	**21**	**0**
4987	**125**	**7**	**3-Sec**	**259**	**21**	**0**
4990	**125**	**8**	**2-Pri**	**275**	**21**	**0**
4994	**125**	**9**	**2-Pri**	**309**	**21**	**0**
4998	**125**	**10**	**2-Pri**	**344**	**21**	**0**

Example 3: Carcass #187:

	Str	Rot	SType	SDate	Blob	Detected
6017	151	5	1-Pre	189	25	NA
6018	151	5	2-Pri	191	25	0
6019	151	5	3-Sec	200	25	0
6020	151	5	4-Pos	205	25	0
6022	151	6	2-Pri	227	25	0
6026	151	7	2-Pri	262	25	0
6030	151	8	2-Pri	303	25	0
6033	151	9	1-Pre	335	25	0
6034	151	9	2-Pri	337	25	0
6035	151	9	3-Sec	342	25	0
6036	151	9	4-Pos	347	25	0
6038	151	10	2-Pri	368	25	0

For naturally placed carcasses, the data are similar except sparser. We will not know placement dates and we will not even be aware that a carcass exists until it is found the first time. Completely undetected carcass (such as naturally-placed versions of #187) will not be represented at all in the data. Basically, for the naturally-placed counterparts of the experimental carcasses above, the data would only consist of the records in bold.

Two general methods of data analysis will be performed on the simulated data to develop estimates of total fatalities.

Method 1: Adjustment factor method. This is generally the method used in Smallwood (2007), also known as a Horvitz-Thompson estimator, and similarly proposed in the Monitoring Team study plan (M53) which is currently implemented today. Some substantial differences exist between the Smallwood and M53 implementation, however, in that M53 utilizes pre-, secondary, and post-searches interspersed at random intervals among the routine (primary) searches. This creates an opportunity to develop carcass removal and searcher efficiency functions that coincide spatially and temporally with the monitored fatalities.

Methods that use these data to develop new estimates of carcass removal and searcher efficiency functions have not been fully explored. These simulations can be a tool for exploring how well some approaches that are discussed in SRC meetings might perform, at least empirically:

> Method 1a: Regression approach. Apply the Smallwood (2007) approach to the M53 experimental trials data to develop new, simultaneously varying correction factors. Apply the derived correction factors to the number of naturally placed carcasses detected.
>
> Method 1b: Simple proportions approach. Compute simple proportion of experimental carcasses detected within 90 days of placement. Apply this proportion as a correction factor (divisor) to the number of naturally placed carcasses detected < 90 days old. This approach assumes that searchers can correctly determine when carcasses are < 90 days.

Method 2: Data augmentation method. Royle et al (2007) introduced the data augmentation approach for estimating population sizes N (total numbers of fatalities) when subjects (carcasses) are distributed according to spatial factors, and detected with imperfect probabilities that can vary spatially and temporally. Their Bayesian framework is a unified approach which simultaneously estimates detection probabilities and N, and provides the ability to compare the statistical evidence for competing models for N – something which has previously been highly problematic (Link 2003). Since its seminal publication, the data augmentation approach has been increasingly cited in the Bayesian ecological literature and merits serious consideration (Royle and Dorazio, 2008; Link and Barker, 2010; and Kery and Schaub (draft)).

References:

Kéry, M. & Schaub, M. (2010 draft) *Bayesian population analysis using WinBUGS/OpenBUGS – a hierarchical perspective*

Link, W.A. (2003) Nonidentifiability of population size from capture-recapture data with heterogeneous detection probabilities. *Biometrics* **59,** 1123-1130.

Link,W.A. & Barker, R.J. (2010) *Bayesian inference with ecological applications*. Academic Press, London.

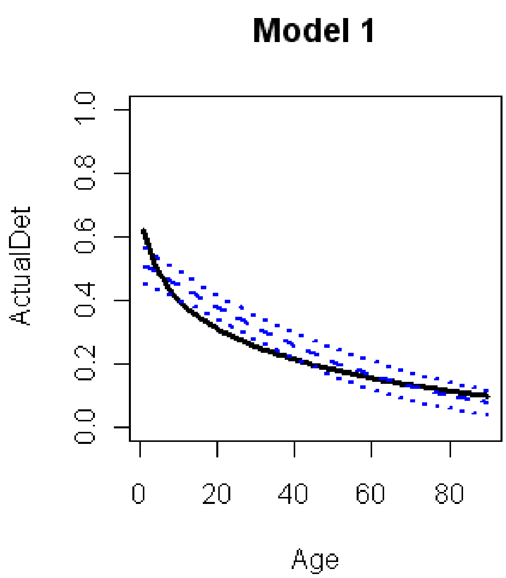
Royle, J.A. and R.M. Dorazio. (2008) *Hierarchical Modeling and Inference in Ecology: The Analysis of Data from Populations, Metapopulations, and Communities*. Academic Press, San Diego, CA. xviii, 444 pp.

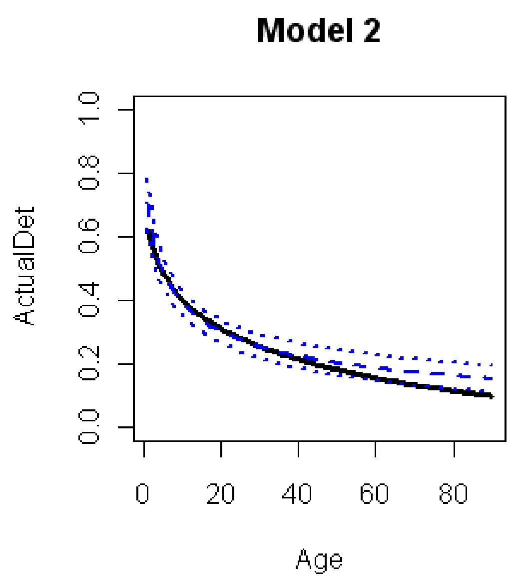
Royle, J.A., Kéry, M., Gauthier, R. & Schmid, H. (2007) Hierarchical spatial models of abundance and occurrence from imperfect survey data. *Ecological Monographs* **77,** 465-481.

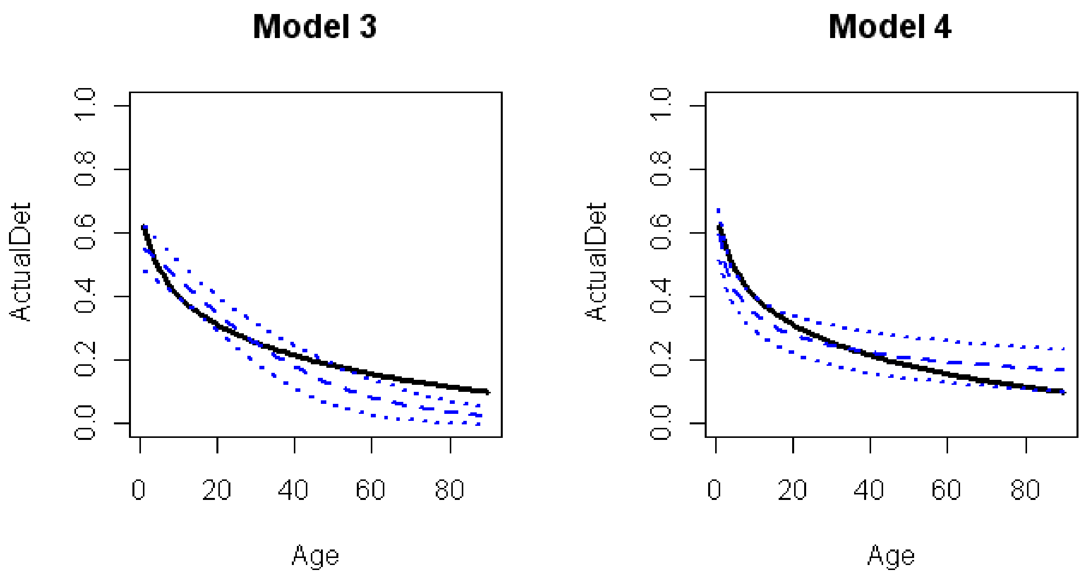
Smallwood, K.S. (2007) Estimating wind-turbine caused bird mortality. *Journal of Wildlife Management* **71**:2781–2791

Sample results of five test models:

	Model Number				
	1	2	3	4	5
Age					
known	✓	✓			✓
estimated			✓	✓	
Searches					
primary+ secondary	✓	✓			
primary			✓	✓	✓
Leave-out time					
90 days	✓	✓			✓
1 rotation			✓	✓	
Predictors					
age	✓		✓		
log(age)		✓		✓	
Method					
regression adjustment factor	✓	✓	✓	✓	
simple adjustment factor					✓
data augmentation					







Estimates based on *one simulation* of 1-year of data analyzed with models 1-4, using natural (non-trial) carcass counts aged < 90 days – appears to be biased high due to the "double-counting" of carcasses missed in its first rotation but detected in later rotations.

	NaturalFats	Count	AdjFactor	Estimate
1	538	261	0.4099343	636.6874
2	538	261	0.3697671	705.8498
3	538	261	0.3956191	659.7255
4	538	261	0.3312166	788.0041

Same analysis except using natural (non-trial) carcass counts aged < 35 days

	NaturalFats	Count	AdjFactor	Estimate
1	538	221	0.4099343	539.1108
2	538	221	0.3697671	597.6736
3	538	221	0.3956191	558.6182
4	538	221	0.3312166	667.2372

Estimate based on *one simulation* of 1-year of data analyzed with model 5 – appears to be biased low due to biased high detection probability when trial carcasses are deposited shortly before primary searches.

5	538	129	0.6450000	393.7984