The concept

To estimate the total number of fatalities at a turbine, we'll count the number of carcasses that fall on roads and pads and then divide by the estimated fraction of the total that fall on roads and pads.

Let *f* be the distribution (pdf) of carcasses as a function of distance from turbine, so that the fraction of carcasses that land within *r* meters is given by . Essentially, *f*(*r*) gives the expected fraction of carcasses falling in a 1 m annulus centered *r* meters from the turbine, and the expected fraction that fall in a 1 m quadrat at *r* meters would be *f*(*r*)/(2*r*). To estimate the fraction that would fall on roads and pads, we first calculate point densities for every point *i* on a virtual grid superimposed on the search area around the turbine and rescale so that the total sums to one, i.e. , where *R* is the search radius. The estimated fraction of carcasses falling on roads and pads is then , where is the set of indices for points that lie on roads and pads.

To estimate *a* we need ** . ** can be derived either from a distance function that gives the expected relative density as a function of distance from the turbine (a la DWP paper) or from an estimate of the carcass distribution function *f*. If we searched everywhere uniformly, we could simply fit a parametric distribution to the empirical count distribution to get *f*. With a roads and pads search, the distribution of observed carcass distances will be skewed toward the turbine, where the fraction of area in roads and pads is higher than it is at greater distances. In that case, the distribution of distances of observed counts is not a good match for *f*. Instead, the probability of observing a carcass at a given distance *r* is the probability that it lands at a distance *r* times the probability that it is observed given that it has landed, and the probability density function is , where the integral in the denominator ensures that integrates to one and represents a legitimate pdf. Given  = the set of distances of the observed carcasses, we can estimate *f* by maximum likelihood. The log-likelihood is given by:

, where *n* is the observed number of carcasses. Since *w* does not depend on **, we can ignore the first term when maximizing the likelihood.

An example

1. Roads and pads on a single turbine (turbine #3 at Fowler).

2. Lay down a virtual 0.5 m x 0.5 m grid for accurate calculation of area in roads and pads as a function of distance.

3. Calculate *w* = fraction of ground that is in roads and pads as a function of distance from the turbine. *w* is stored as an array with columns *r* and *w*(*r*).

4. Define a scenario for scattering carcasses. E.g., let the relative density of carcasses be a function of distance *x* from a turbine with *y*0 = 0.05 (as in DWP paper but without truncation). As a pdf on ring densities, it is . The gamma distribution with the same mean and variance as the *y*0 = 0.05 pdf has parameters ** = 2.04459 and ** = 19.15065. The two distributions are shown on the right (*y*0 in thick red; gamma in thin blue):

5. Proof of concept: scatter 100 carcasses according to the gamma distribution defined in step 4. Compare the empirical distribution of the scattered carcasses (black) with the target gamma distribution (thick red) plotted as a cdf. Count carcasses on the roads and pads. Calculate the empirical cdf of observed carcass distances (dotted blue). Use maximum likelihood to estimate the true distribution *f* (solid blue) from the observed carcass distribution (dotted blue).



6. Estimation: 1000 reps of scattering 100 carcass, counting on R&P, estimating *f*, and calculating estimated total.

Comparison of the weighted MLE and the linear logistic (where the linear logistic matches the underlying random process that generates the data):



Linear logistic

Mean: 110.5

StDev: 54.79

95% CI: =[107.1, 113.9]

Weighted MLE

Mean: 105.8

StDev: 57.92

95% CI: =[102.2, 109.4]

Summary statistics:



\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

A bad estimate:



Fit of gamma to H & M bat curve (large turbine)



For 100 carcasses thrown onto Fowler #3 with *y*0 = 0.05 linear logistic, the weighted gamma and linear logistic have broadly similar estimates:

wgmle: mean = 108.5, 95% CI = [104.9, 112.2]

llogis: mean = 113.4, 95% CI = [110.0, 116.9]

Both are somewhat positively biased on a linear scale.

The wgmle is somewhat more variable with:

wgmle: sd = 58.6, mean absolute error = 38.2

llogis: sd = 55.8, mean absolute error = 37.7

Logistic does better on the extremes (especially on the low end):

signif(quantile(simres$est.mle,c(.01,.05,.1,.25,.5,.75,.9,.95,.99)),4)

1% 5% 10% 25% 50% 75% 90% 95% 99%

31.60 46.38 53.81 72.44 97.42 130.20 165.80 198.50 323.80

signif(quantile(simres$est.logi,c(.01,.05,.1,.25,.5,.75,.9,.95,.99)),4)

1% 5% 10% 25% 50% 75% 90% 95% 99%

36.20 50.53 58.35 78.43 103.60 136.90 170.00 198.90 309.90

On a log scale:

wgmle: geometric mean = 97.2, 95% CI = [94.51, 100.05] (i.e., off by about 3% on average)

llogis: geometric mean = 103.0, 95% CI = [100.22, 105.76] (i.e., off by about 3% on average)



llogis estimates tended to be greater than wgmle estimates (68.3% of the time). So, when the carcass configuration was conducive to underestimates, llogis tended to be better. When the carcass configuration was conducive to overestimates, mle tended to be better. However, with enormous overestimates (> 3x, or approximately 1% of the time), llogis was consistently better.