February 29, 2008

To: Alameda County

cc: Scientific Revue Committee
Monitoring Team

From: William Warren-Hicks, Ph. D. for FPL Energy

Subject: Comments on *Bird Fatality Study at Altamont Pass Wind Resource Area, October 2005 to September 2007, Draft Report (Revised February 11, 2008).*

The following comments have been complied in response to the County Monitoring Team's recent summary report for the first two years of the ongoing research being conducted in the Altamont Pass Wind Resource Area. These comments are to replace the previous set of comments from FPLE, based on the first draft of the Monitoring Team Report (January 25, 2008).

1. p. 3, Survey Design: The discussion on p. 3 indicates that "turbines were distributed in 84 randomly selected plots stratified by geographic location and turbine size." We are concerned that the calculation methods employed throughout the report seem inconsistent with the survey design. Although not discussed on p. 3, we understand that at least three turbine size categories were censused. These classes include: very small (40-65kw installed capacity (IC)- e.g. Patterson Pass' Nordtank 65kW IC machines); small (100-150 IC- e.g. KCS 56-100's found in many areas throughout the Altamont, 100kw IC machines); medium class (≥250 kW IC- e.g. Diablo Winds' 660kW IC machines, Tres Vaqueros' Howden 330kW IC machines, and the WEG turbines- 250kW IC machines). First, we question whether the classification of the medium turbines is appropriate, given the disparity between the size and potential power output of the turbines currently grouped together. Second it is not clear that the turbine size classes were taken into account in the analyses. Also, the small (KCS 56-100's) size category was sampled based on a quasi-random selection of turbine strings (not turbines as indicated in the discussion). Implications of these issues follow:

a. Conceptual Model: A basic tenant of good statistical practice is that the methods employed for statistical inference must be consistent with the conceptual model underlying the survey design (Kalton 1983, Cochran 1977)[1]. In this case, the censused and sampled strata have differing inclusion probabilities which must be taken into account by the methods employed for statistical inference. The conceptual model, although not stated in the report, seems to be that mortality is a function of turbine size. The survey approach and associated inclusion probabilities seem to be based on this premise, therefore, the statistical inference methods must reflect this concept. Simply stated, the data from different turbine size strata cannot be simply

[1]Kalton, G. (1983). *Introduction to Survey Sampling*. Sage Publications, Beverly Hills, CA. ; Cochran, W.G. (1977). *Sampling Techniques*, Third Edition. New York: Wiley.

pooled to generate an expected mortality rate. Also, the report states that "plot" is a stratum, but the report does not address randomization within plot. We find the survey design and the randomization of sampling units confusing.

b. Calculation of an expected mortality rate: The methods employed for statistical inference should result in an expected mortality rate. The unit of this metric (i.e, the rate) must reflect the survey design (i.e., if turbine string is the sampling unit, an estimate of the expected mortality per turbine string is an appropriate statistic). The statistical methods employed in this report are not clear. For example, the equations and data used to generate the fatalities per turbine string (Table 1) are not apparent. Given that strings are not surrogates (or replicates), it is difficult to interpret the values presented in Table 1. Are the fatalities per string weighted by the number of turbines in a string?

c. Sampling unit: What is the sampling unit? The report must clarify this very important concept (see disconnect between the language on p. 3 and Appendix A). The methods employed for statistical inference must correctly incorporate the sampling unit definition, the number of units sampled, the inclusion probability associated with an individual sampling unit, and the number of replications per unit. There is little information in the report that provides us with confidence that the statistical inference equations were properly implemented.

d. Weights: The mathematical methods used to weight the resulting data are unclear. Equations should be provided, including equations for the calculation of uncertainty in the expected mortality rate. There seem to be a number of possible weights that could be employed, including the following: number of replications per sampling unit, number of turbines per turbine string (providing the sampling unit is a turbine string), amount of time turbines operate per search interval, or area surveyed per sampling unit (see discussion in Appendix A which suggests that the area surveyed per sampling unit may not be consistent for all sample units). A discussion of the choice of sampling weight, and the equations used for weighting, must be presented in detail. Also, it seems that the weighting scheme employed in this report was not applied at the level of a sampling unit for a single unit of effort, but at the level of the sum across units. If this is true, the mortality estimates will be incorrect (see above references). Also see a discussion for the proper application of weights in Sokal and Rohlf (1995 *Biometry*, W.H. Freeman and Co, New York page 133).

2. p. 5, equation 1: We have the following issues with equation 1:

a. The numerator (the unadjusted mortality rate) seems to be incorrectly calculated (see above).

b. The denominator (R x p) assumes that the scavenging rate and observer bias are independent events and that they are inversely proportional to mortality. This is clearly not the case. Simply stated, should a bird be scavenged, it cannot then be "unobserved" by the investigator. Although in practice scavenging and observer bias studies are conducted independently, from a biological perspective these factors are not independent and can greatly misrepresent the true mortality at the site.

Take the following scenario:

Observer bias and scavenger bias studies were conducted independently at the site of interest. The result of these studies is considered truth (no uncertainty). The result of these studies is p=0.2 (observer bias) and R=0.2 (scavenger bias).

To test the validity of the 2 studies, a third blind study was conducted at the site, and 50 birds were put into the field. We have no idea if observer bias is operating or scavengers are scavenging. And, if both issues occur during the survey, we don't know the proportional number of birds missed due to each cause (i.e., the blind study replicates the conditions most often encountered in wildlife studies). During this third study, the observers follow a pre-defined SOP for transect-line searchers with a search area consistent with the studies generating the bias constants. The equation used to calculate the "true" mortality is M / (R * p), where M is the unadjusted mortality (Smallwood 2007).

The observers were able to find 10 birds. This is not unexpected or considered unusual. We already know that the true site-specific observer bias is 0.2. Maybe, there were no scavengers? And, if we ignore the scavenging bias term in the equation it is easy to find the truth by calculating: 10/0.2 = 50. So, by ignoring scavengers the equation works perfectly.

Now we deal with the scavenger bias (remember that this study is conducted independently of the observer bias study). Maybe the observers are perfect and only scavenging occurs. Again, if we ignore the observer bias term in the equation we find that the equation works perfectly: 10/.2 =50. And, given the results of the independent scavenging study the field results cannot be considered unusual.

Question: Which is operating, observer bias or scavenging bias? Or, is the combination of scavenging bias and observer bias equal to 0.2? Answer: we don't know because the study designs were not established to generate the probability of greatest interest.

The problem: In the field, both observer bias and scavenging bias maybe operating, but we don't know for sure. And, we don't know the percentage of missed birds associated with each issue. The original studies did not adjust for the interaction of these "bias factors" in the study design. This is a classic "sampling without replacement" problem.

The following discussion illustrates this issue: In most field searches the observer walks down a transect looking for a bird within a pre-defined area. If the bird is truly within eyesight and the observer misses the bird, then the "true" number of birds is miscounted and this finding could correctly be attributed to observer bias. However, if a scavenger obtains the bird prior to the observer's opportunity to view the bird, then observer bias has not occurred. In this case, only scavenging bias has occurred. The probabilities of observer bias and scavenging bias are therefore not independent. In addition, this interaction of observer and scavenging impacts on the ability to find the "true" number of birds is generally not evaluated during the independent studies used to generate observer and scavenging bias constants. Notice that once the scavenger obtains a bird, the number of birds that could be "found" by the observer during the observer bias study is reduced by one, therefore changing the denominator of the observer bias "rate."

The majority of the current equations in the wildlife literature used for adjusting "found" birds to the "true" number of birds are flawed. The assumption of an inversely proportional and independent probability of observer and scavenger bias results in a high chance of overestimating the "true" number of birds. And, the degree of overestimation is function of the magnitude of the observer and scavenger bias (which can be shown by simulation).

Most of the current mortality equations allow us to count birds multiple times. For example, after we independently adjust for observer bias and scavengers, we find 100 birds (50 + 50), which is incorrect. However, the estimate of 100 birds is consistent with the independently generated observer and scavenger study designs.

Note that using the Smallwood equation we find 10/ (.2 * .2) = 250 - not anywhere close to the truth. Neither are the terms additive: 10/.4 = 25.

Unfortunately, we need the "expected" probability of observer bias conditional on scavenging. And, we need the "expected value of the marginal" probability of scavenging. When multiplied together, the resulting joint probability of scavenging and observer bias should, on average, provide accurate estimates of the "true" number of birds. Ideally, the conditional and marginal probabilities generated from the field studies would be distributions (as a function of time). In any case, these conditional and marginal probabilities are not calculated in the typical bias-generating studies. And, I don't believe we have a mathematical way of estimating these probabilities given the study designs that are commonly used to generate observer and scavenger bias (but I will think about this a bit more).

Suggestion: Either (1) design and implement a field study to generate the required probability metrics, or (2) calculate mortality independently using observer bias, and again with scavenger bias; report both numbers, and take the worst case.

Otherwise, the Smallwood equation and others like it will run a high risk of overestimating the true number of birds.

c. Given the large influence of this equation form, and the large influence associated with the selection of a specific R and p, a comparison of mortality generated using equation 1 with the baseline value of 1300 birds is not possible unless the baseline numbers are recalculated in the same manner as is used in this document.

3. equation 2, p. 5:

a. We understand that scavenger removal trials were conducted during the survey period. Therefore, we strongly suggest that the site-specific value be used to adjust mortality.

b. On p. 6, it seems that equation 2 was used under the assumption that the bird died at the end of the previous observation period (i.e., 37 or 44 days prior). This results in a maximum scavenger removal rate over the search interval. Given the sensitivity that the selection of R has on the mortality estimate resulting from equation 1(see above discussion), the calculation of R in this manner has an extreme influence on the resulting estimated mortality. If site-specific values of R are not available (the preferable approach), we suggest using the median value of the search interval in equation 2.

c. Also, note that on p. 6 the value of R for small raptors approximates the value we used in the example above. Again, given the mathematical form of equation 1, combined with the values of R on p. 6, the estimates of mortality provided in this report are likely over estimates of the true mortality.

4. p. 6, searcher detection efficiency: We suggest that searcher detection bias be studied on-site and the results used in the mortality calculations. We believe that the degree of observer bias is strongly case specific. Extrapolating estimates of p from historical studies is not generally appropriate. Again, given the sensitivity of equation 1 to small changes in p, we suggest that a site-specific and accurate estimate of p be generated by running on-site observer bias studies.

5. Tables 1 (p. 9) and Table 4 (p.15):

a. Although it is noted in the text that bones (fatalities estimated to be over 90 days old) are not included in the total recorded bird fatalities table, it is not explicit what other kinds of incidents are included in these totals. Are these birds only from within the search plot (50, 60 or 75m)? Are birds that were injured counted as fatalities? Are all feather spots included?

b. Why are the total number of birds in Table 1 (1596) and Table 4 (1804) different? We also note that the species-specific values are different between the tables.

c. The extrapolation methods seem to suggest that the probability of species presence is uniform across the Altamont. Our understanding it that this is not true, particularly for burrowing owls.

6. Stdev, Tables 1 - 3:

a. How is the Stdev calculated? Please note that a metric such as mortality per turbine string is not normally distributed, but most likely follows a lognormal distribution. Therefore, the authors should consider presenting an uncertainty estimate of the median, or the transformed log mean. Please include all equations used for the calculations in the text. What should be reported in Tables 1 -3 is the standard error of the weighted mean. The label may be wrong, or the estimating equation may be wrong, but we can not tell which is the case because the equations are not provided.

b. Also, note that the coefficient of variation (CV = std./mean) values are larger than would be expected. For example, in Table 2 the burrowing owl has a CV of 2.6. The table does not indicate the sample size (n), which provides additional confusion. But, we suspect that the number of sampling units used to generate the weighted mortality rate is large, which generates confusion over the size of the CV. Again, please provide exact equations, sample sizes, and methods used for the calculations in the tables.

7. Table 2: If the sampling unit is a turbine string, please explain how the extrapolation to individual turbines was accomplished. Statistically, extrapolation to units of smaller size than that used in the survey (i.e., a turbine string extrapolated to an individual turbine) is difficult if not impossible to defend.

8. Appendix A: are the EnXco and Diablo turbines included in the calculations presented in Tables 1- 4. If so, were the mortality estimates adjusted for the change in search area?

SUMMARY

We have the following 3 general concerns with this document:

a. The statistical methods seem inconsistent with the survey design.
b. The equations used to calculate mortality are, by mathematical definition, over estimates of the true mortality rate.
c. The report is not transparent.