

SRC Comments
On M21, Revised Draft Monitoring Report, APWRA Bird Fatality
Study, December 2009

Alameda County Scientific Review Committee

The Altamont Pass Wind Resource Area (APWRA) Monitoring Team produced a draft report, Altamont Pass Wind Resource Area Bird Fatality Study (M21), in December 2009. Written comments on the draft report submitted by individual members of the Alameda County Scientific Review Committee (SRC) are compiled in this document. This document was updated as comments were received.

[Joanna Burger](#)

[Jim Estep](#)

[Sue Orloff](#)

[K. Shawn Smallwood](#)

[Julie Yee](#)

Joanna Burger

January 11, 2010

OVERALL

1. I recognize that a tremendous amount of work went into this report, that the rules and requests changed constantly, and that the data were difficult to QA/QC, as were the data on wind power output.
2. And it is a pleasure to see the report finally, and to begin to examine the overall trends. The MT has obviously made a valiant first attempt, which should be so noted.
3. And given all our discussions, I believe the conclusions are probably correct, but more caveats are needed.
4. I was surprised by a number of things I thought we had agreed upon that were not done, and that were clearly in the notes from our in-person meetings
 - a. We had discussed at great length making sure that the company capacity data was accurate, and reflective of reality, and requested that differences be explained (which was not done). Further, discussions of why we used particular data sets should be included.
 - b. We made a big point of talking about bird-years, rather than calendar years, and this has major implications for the analysis.
 - c. We had agreed to use three methods of adjusting the data and to adjust for feather piles.
 - d. We had agreed not to use the Set 2 turbines because they were searched only twice.

- e. We had agreed that the MT would look at the month or two prior to shutdown and following shutdown in each year to determine if this could tell us something about bird abundance in those years.
 - f. Differences in the removal curves the SRC saw need to be accounted for and discussed.
 - g. We had made it clear that we wanted explanations of why particular methods were used so the public could clearly see how and why particular methods were applied, and this was not uniformly implemented.
 - h. Years where no monitoring data are available should not be included
 - i. A table of the overall design, with treatments was to be included
5. There is sometimes a lack of any discussion of findings, particularly in light of either the biology of the four focal species, previous studies at Altamont, or the literature in general. One might think that few other studies existed. The SRC had several discussions over the years about these issues.
 6. Overall, the report needs a summary, abstract or executive summary, and the objectives, need to be stated earlier, and more clearly.
 7. The methods are insufficient to allow transparency, as are some of the other descriptions of procedures (someone unfamiliar with the Altamont should be able to follow the procedures from the report)..
 8. I am still wondering whether the report should focus on the four species right from the start, and put the “all fatalities” in an appendix.
 9. In the future, it would help if there were line numbers on the report so we could refer to particular lines on a given page.
 10. I think the personnel at ICF Jones & Stokes should be listed, and someplace there should be an indication of roles of each, including a list of field personnel.

SPECIFIC COMMENTS

Page 1-2: It is not clear that the 50% was to refer to all raptors, but that the monitoring related primarily to four species.

Page 1-2. We had specifically requested a table be put in that listed the conditions during the baseline and the study period (including when there was winter shutdown, and the BO/K study). Although the methods are described, they should be easy to see in a table.

Page 1-3. At the end of this section there should be a clear statement that sets out the objectives and format for the rest of the report.

Page 2-1. Before differences in study methods are given, there should be a general statement of the methods to orient the reader.

Page 2-1 – There should be a glossary that identified all parts, such as Tres Vaqueros and so on (some interested stakeholders will not be so familiar with the whole thing). Some of this could be done in a table that the reader could refer to.

Page 2-1 (bottom), Here and elsewhere, it would be useful to refer to the papers from which the data were taken. The baseline study was not done by this consulting firm.

Page 2-1 (bottom). I would be happier if numbers were accompanied with an indication of variance (e.g. $5.8 \pm ??$ times). This goes for table 2-1 also. This particularly true for average search interval.

Fig. 2-3: Baseline and current study periods should be defined in the figure legend.

Page 2-2: Here and elsewhere you need to decide on the tenses.

Page 2-2: Methods should be described in more detail. This report will be read by the public without much prior knowledge (as well as some with a very deep knowledge-base), and the methods are not in enough detail for it to be clear. Further, someone from elsewhere could not repeat your methods from this description. How many people, how long per day, what was recorded and so on.

Page 2-2: where is the evidence for decreasing search interval each year?

Page 2-2 bottom: Does the WRRS data reflect mortalities all over the APWRS or only those samples areas? Is any attempt made to exclude mortalities from strings that were NOT searched by the monitoring team?

Page 2-3 top; When is the baseline period mentioned in table 2-2? (NREL and CEC???)?

Page 2-3 top. What does “proportion of WRRS records was small mean?”

These pages – it might be useful to have some of these tables in terms of raptors as well (especially 2-4)

Figure 2-2 seems unnecessary – why not have a table so the actual data are available?

Page2-5 It might help to have a table of the different studies for adjusting fatalities (and see comment above about the three analyses the SRC expected).

Page 2-6 (top). There is a general lack of reference to the literature, and certainly, when one says “it is becoming common practice” a couple of references from the published literature are required.

Page 2-6 top: Again, it would be useful to make a small table of the definitions of the different megawatt capacity designations. It would just be easier for the reader.

Page 2-7. I understood that we were also going to look at the mortalities in the month (or two) before the start of shut down to see if there seemed to be years when bird use may have been higher than in other years.

Page 2-7 (bottom). There need to be some caveats about comparing different turbine types (e.g. In repowering), particularly since there are some mortality differences in turbine types within the smaller types..

Page 3-1 top : Any way to compare the numbers statistically in tables 3-1 and 3-2 (Goodness of fit??)

Figs 3-1 and 3-2. Again, should there be some indication of variance indicated. They should not really be a bar but a number with variance (since there are several years in the study).

Again, since we lump the data for all years, it is not possible to see if there might have been differences due to year.

For all species, maybe the bars (if you use them) could indicate the percent that were feather spots.

Page 3-2: The text is too brief and should summarize the findings more clearly for each species, with a few lines about each species.

Page 3-2 and following graphs. The way the data are presented, it is not really possible to see whether there might have been differences in bird abundance (we have only yearly data, and not monthly data by year).

It might help to graph the mortality by month (for each year) on top of each other for a given species. We could then see what was actually happening each year with mortality.

Page 3-3 middle. There was also a similar decline in the baseline years, and as we discussed, this should be mentioned as it shows further variability in the data.

Page 4-2 (top). When making statements like – would lessen the difference between the two studies, but would also result in fatality estimates that cannot be considered realistic – requires more explanation and discussion

Page 4-2 (just before assessment.. section). This one sentence paragraph needs more explanation and clarity. You need to suggest that potential increases in the populations of raptors in the area may simply result in higher mortality whether there have been management measures or not. (I am still uncomfortable not using any of the bird data from the region. Even though it has many of the same problems as the mortality data set, it is still indicative).

Page 4-3: I am not convinced that we adequately studied the winter shutdown since we had two treatments: 2month/2month and continuous. Given the inherent variability in the mortality data (and the abundance data which must be involved), the time period was not long enough. Finally, we DID NOT implement a four-month winter shutdown all these years, but a combination of things which was not consistent from year to year.

Jim Estep

January 11, 2010

General Comments

Overall, I think the report is well-done, concise, and relatively easy to wade through. It is incomplete due to the lack of description and/or analysis of some issues, and it does not address several of the recommendations made by the SRC. But in its current form, it still reasonably achieves two important functions:

- 1) it makes the required conclusion regarding the 50% reduction, and
- 2) it addresses the benefits of repowering, which may provide the only means of achieving reasonable mortality reduction with continued operation of the APWRA.

In addition to addressing specific comments, the final report should discuss in greater detail the following:

- the known and potential biases that result from comparison between the baseline and current studies and how these may affect results. The discussion is good in this regard, but this issue may require a separate section that is more complete with the various caveats to the analysis.
- inconsistent implementation of management measures (e.g., shutdown, turbine removal) and how these may affect results. It should also note the lack of other available management measures that could potentially reduce mortality. Thus, the only substantive changes made in the APWRA for which the baseline and current study differed in order to evaluate a 50% reduction were the winter shutdown and turbine removal – both of which were inconsistently applied.
- either a discussion of alternative analyses that the SRC recommended but were not used or the addition of these analyses in the final report.

The following are two major elements lacking from the report that were important aspects of the Settlement Agreement and/or the original study plan:

- There is no description, analysis, or conclusions reached regarding turbine removal and relocation. There is a brief mention in the Introduction in the description of the Settlement Agreement, but no further follow-up in the report. Along with the seasonal shutdown, this was the only other major management action that was implemented during the study and was in part required under the Settlement Agreement. This was a major action on the part of the Settlement companies and the SRC that included both Settlement-required actions and additional recommendations for removal by the SRC. While it would be difficult to isolate and assess the specific effects of this issue, it should be described as one

of the management actions implemented to reduce mortality and from which comparisons between the baseline and current study could be based.

- There is no discussion of bird use or abundance in the APWRA. Correlation between bird abundance and mortality is another important issue that has been discussed by the SRC since its inception, and is part of the study plan for the project. Data have been collected as part of this study, and while not available for use, it should at least be identified as a deficiency in the report.

Finally, the report also lacks an analysis or at least a description of several other issues that may influence bird mortality, particularly the effects of geography/topography. This may be beyond the scope of this report, but a more thorough description of the potential causes of bird mortality in the APWRA along with the various key references should be included in the introduction.

It seems that while I like the conciseness of the report, to address these issues it may require additional detail including more background in the Introduction and more explanation in the Methods.

Specific Comments

1. Page 1-1, Paragraph 2, Sentence 3

Include the time period for initial installation of APWRA turbines. Additional background information on the APWRA would also be useful.

2. Page 1-1, Paragraph 2

I couldn't find any reference in the report to the types or characteristics of turbines in the study (perhaps this information will be in Appendix A – Characteristics of Turbine Strings and Operating Groups). But at least a brief mention in this section noting that most of the sample turbines were 'old generation' turbines – mostly 56-100s with horizontal rotors mounted on lattice towers – would be useful. This will also help to introduce the later discussion of repowering.

3. Page 1-1, Last bullet

While this is the stated SRC role as noted on the website, it is a bit incomplete. Particularly relevant is the role of the SRC in making management recommendations to reduce mortality. For example, the modifications to the seasonal shutdown and additional removal of high risk turbines were largely the result of SRC recommendations (although our actual recommendation of a 4-month shutdown and a more robust turbine removal program were never implemented). An additional role of the SRC is to approve study methods, scaling factors, etc. (see Settlement Agreement 3aii).

The SRC workplan (P99), the SRC Charter, the Settlement Agreement, and Conditional Use Permit Exhibit G-1 provide a more complete description of the role of the SRC.

4. Page 1-2, First Paragraph, 3rd sentence

The terms ‘partial’ and ‘full’ shutdown of turbines is unclear and potentially misleading. While clarified in a subsequent statement, the winter shutdown applied only to companies participating in the Settlement Agreement, so there never has been a full shutdown of turbines in the APWRA. Also, ‘full wintertime shutdown’ implies a consistently applied shut-down and start-up of all turbines, which is not the case.

5. Page 1-2, Second Paragraph, last sentence

Add ‘the removal or relocation of turbines from high risk locations’.

6. Page 1-2, Fourth Paragraph

While the cross-over design was discontinued as per SRC recommendation, we might also note that the SRC’s recommendation for an APWRA-wide 4-month winter shutdown was not implemented, which was considered important in achieving the 50% reduction. Thus, in retrospect, retaining the cross-over design may have led to more interesting results with respect to identifying the effect of a winter shutdown, but the SRC’s focus was on mortality reduction.

7. Page 1-3, First Paragraph

It might be noted that while the shutdown was in fact extended to 3 months, this was not consistent with the SRC’s recommendation of a 4-month shutdown for 2007-08 (P49).

8. Introduction

There is no discussion in the introduction of the removal of high risk turbines. This was a key management measure in the Settlement Agreement and was an important recommendation of the SRC as it expanded the removal/relocation efforts to include additional turbine sites deemed high risk by the SRC. We might also want to include a discussion of the timing (i.e., inconsistent compliance) of the removal/relocation process and how this may have affected the ability of the monitoring program to detect changes in mortality.

9. Page 2-1, Second Paragraph

We note here some of the differences between the baseline and current study methods and acknowledge that some (e.g., search interval) are important. While this and other related issues may be addressed later in the report, perhaps the possible consequences of these potential biases resulting from the differences in study methods should be addressed earlier in the report. It seems the reader may immediately question the validity of

comparing the baseline and current study from the brief summary provided on the differences between them, and thus perhaps the report should make this acknowledgement early rather than in a summary at the end. In addition, the methods section should explain how these differences were reconciled during data analysis.

10. Page 2-2, First Paragraph

Sentence 2 and 4 use the term ‘block’. Sentence 3 uses the term ‘plot’.

11. Page 2-2, Fifth Paragraph

The description of WRRS seems unnecessary since those data are not used in the analysis. The utility of WRRS data has always been controversial due to the incidental nature of the data collection – and certainly its inconsistency with the current study protocols. Including this description is potentially confusing to the uninformed reader. Perhaps include WRRS data as ‘incidental finds’, then the description of incidental finds can include those reported by surveyors and those reported by wind company staff.

12. Page 2-3, Second complete paragraph, non-native species

While I think the proportion of non-native species is interesting and meaningful, excluding the data from the analysis entirely may be unwise. I would conduct the primary analysis using all bird data, followed by an analysis showing the effect of excluding non-native species.

13. Page 2-3, Second complete paragraph, last sentence

We note here the ‘unreliability of the backdating process’. This concept has not yet been sufficiently introduced and discussed to make this definitive statement. Why is the process unreliable and what effect does this potentially have on the results? If this information will be included in Appendix A, then cite Appendix A here. Otherwise, the Methods section should describe the backdating process and related issues and perhaps cite SRC documents that address this issue.

14. Page 2-4, Last sentence

The statement that the use of the detection probability equation is ‘common practice’ in the Altamont seems to be a fairly narrow explanation of its use in this study and leaves open the question of whether or not there are alternatives that were considered. Is the equation in common use just in the Altamont or is it commonly used in wind energy/bird mortality studies elsewhere? Because the detection probability formulation is such an important issue, perhaps the Methods section should simply report what was done and not use these types of incomplete explanatory statements. Instead, the implications of using this formulation and the debate regarding detection probability should be fully discussed in the Discussion section.

15. Page 2-5, first sentence

For clarity, following ...”detected by an observer...”, add in parenthesis (searcher efficiency).

16: Page 2-5, Searcher Efficiency header

For consistency, add *pi* after the header Searcher Efficiency and rename as ‘Searcher Efficiency Estimates’.

17. Page 2-5, Carcass Removal Estimates

Add study periods for scavenger removal trials and KB study.

18. Page 2-5, Carcass Removal Estimates, fourth paragraph

Does the report address the implications of increasing search intervals during the scavenger removal trials?

19. Page 2-6, First Paragraph

I’m struggling a bit with the use of Operating Groups to estimate mortality rates. What is the utility of this estimation? Is this a function of geography, turbine type, or turbine operations?

20. Page 2-7, third paragraph

Backdating is a key issue here with the Seasonal Shutdown. As noted, we need a clear description of this process in the methods, in an appendix, or in a cited document. The uninformed reader will question the validity of the shutdown experiment without this information.

21. Page 2-7, fifth paragraph

So here’s where the use of Operating Groups makes sense to me – on the basis of turbine type (comparing new generation turbines at Diablo Winds with older generation turbines). It’s unclear at this point how other Operating Groups can result in useful comparisons.

22. Page 3-1, first paragraph, second sentence

The proportion of the total fatalities from the current study totals 75%, not 80%.

23. Page 3-1, first paragraph and Table 3-1

My read of Table 3-1 is that the proportions of raptors vs. non-raptors are not roughly equal between the baseline and current studies, as the text suggests. In fact, Table 3-1 indicates that the proportions are opposite between the two studies (i.e., raptors in baseline is the same as non-raptors in the current study; and non-raptors in the baseline is the same as raptors in the current study).

24. Page 3-1, Table 3-2

The percentages for the current study in Table 3-2 total 108%.

25. Page 3-1, Tables 3-1 and 3-2

For consistency, include 'native' species in the title of the tables.

26. Page 3-1, Second Paragraph, first sentence, and Figure 3-1.

Does 'all birds' here refer to all 'native' birds? If so, should clarify.

27. Page 3-1, last paragraph, first sentence

Should reference the KB study here.

28. Page 3-1, last paragraph, second sentence

What's a "traditional carcass"?

29. Figures 3-3 and 3-4

Move Figures 3-3 and 3-4 up so that they occur before Table 3-3.

30. Page 3-2, first complete paragraph

It seems that because there was substantial overlap between the three curves, it may be appropriate to combine them, but because this issue may not have been entirely resolved, we may want to provide separate calculations using the different curves – or address the limitations and/or consequences in the Discussion.

31. Page 3-2, fifth complete paragraph

Why would the unadjusted mortality rate be higher using the Common Strings dataset? The Common Strings dataset is a subset of the All Strings dataset used to make clearer comparisons between baseline and current studies. The difference in rates between the baseline and current studies are fairly consistent among each of the analyses (Tables 3-3 and 3-4) – that makes sense – but I'm unclear on why there is such a substantial difference in the rates between the data sets. What is unique about the Common Strings

dataset that would result in a mortality rate that is 20% higher for raptors compared with the All Strings dataset but shows no difference for non-raptors?

32. Table 3-6

There are several addition errors in Table 3-6. See first two columns under All Strings Data Set/String Level Analysis, Total Focal and Total Raptors.

Also see Table 3-4, common strings/string level, current study, non-raptors.

33. Page 3-3, first paragraph

This may suggest that the geographic differences in kestrel use of the APWRA that may be exhibited in the Operating Group analysis may have an influence. Correlating with bird use data could be interesting.

34. Table 3-9

Indicate in the table that these data are comparisons to illustrate the differences between new and old generation turbines.

35. Figures 3-9 and 3-10

Move these to occur prior to Table 3-9.

36. Page 4-1, second paragraph, last sentence

Somewhere we need to indicate that bird use data has been collected in the APWRA, but not analyzed for this report. This is a huge issue that the SRC has been discussing since its establishment. And thus not being able to make correlations with bird use data is a significant deficiency in the report.

37. Page 4-1, forth paragraph, last sentence

What is the basis of the comment suggesting that searcher detection probability in the current study is greater than in the baseline study? It seems that in order to make this statement, we would need further explanation regarding the differences between the study methods.

38. Page 4-2, eighth paragraph

Might expand on this issue, including 1) lack of consistency with the shutdown period over all years of the study, 2) issues with application of the shutdown in terms of shut-off and start-up, and 3) no implementation of the SRC-recommended 4-month shutdown.

39. Page 4-2, last paragraph, second through fourth sentences

At this point, these all entirely speculative. There is currently no actual evidence to indicate that they are contributing factors.

40. Page 4-3, conclusions

Should there also be a statement regarding the effects of turbine removal and relocation?

Sue Orloff

01/10/10

I agree with the overall report conclusions that the goal of a 50% reduction in raptor fatality has not been achieved. These conclusions are further reinforced by the recent report by Shawn Smallwood (01/06/10) that reached the same basic conclusions using different methods. However, I found the full report to be deficient in many ways. The methods, analysis, and discussion supporting the conclusions are just as critical as the conclusions themselves, and need to be improved and expanded. I also think it is important to keep in mind that in its final form all aspects of this report must be sound, intelligible, transparent, and useful to a wide audience, ranging from professional statisticians to non-scientific stakeholders. I hope my comments and suggestions facilitate this improvement. My more general criticisms are below followed by my more specific comments. For the specific comments, I tried not to repeat what Shawn Smallwood has already suggested.

General Criticisms:

1. The report was so abbreviated that it was difficult to determine exactly what data or analyses were used, and for what reasons. I think the report should be detailed enough to be a stand alone document, capable of being repeated. This may be the only report that many people will read.
2. Figures and Tables should have all the needed information in the caption or chart. Many times this was missing (i.e., adjusted/not adjusted; all string/common strings).
3. The analysis in the results sometimes doesn't match what was said in the methods.
4. Many of the SRC recommendations were not followed, such as using calendar year rather than bird year, using discovery dates rather than backdates, using the

- CEC data rather than excluding this from some analyses; using the combined carcasses removal curve rather than separate curves; and not using any power output data or operating time data in the analyses. These may not make a difference in the end result, but it sets the precedence for future monitoring studies.
5. There was no section on analysis limitations which was also recommended by the SRC. I recall we were thinking there were so many caveats and limitations that it required a large separate section.
 6. High risk turbine removals, which was one of the most important management strategies the SRC recommended, was not mentioned in the methods or analysis sections. Shawn attempted an analysis of this in his recent report (01/06/10) and the SRC suggested several ways this could be looked at as well. At the very least you should reference the SRC report (P-67).
 7. There was no mention of wind industry practices that could have inhibited reduction in mortality (i.e., derelict turbines left in rows, towers used as flight diverters at the end of rows, turbine attrition leaving gaps in strings).
 8. Citations should have been used much more throughout the text as references.
 9. There was no use of statistics tests in the report.
 10. I think a glossary of common terms would make the reading much easier for most people.
 11. The report needs an executive summary.

Specific Comments:

Page 1-1, 2nd paragraph: I would add normal attrition to the list of things that vary the number of turbines over time.

Page 1-2, 4th and 5th paragraphs: The reason the sampling design was questioned and discontinued needs to be explained in much more detail. I don't think the date (2/06) for discontinuing the cross-over design is correct. I think its in 2007. Otherwise it would not have been as stated "the first two years of the current study (2005-2007)". A chart depicting the timeline of pertinent study events (shutdowns, hazardous turbine removals) would be really helpful. Then in the results section this could be used to evaluate changes in mortality rates overtime.

Page 1-3, last paragraph: This information is already in the previous footnote – so it doesn't need to go here.

Page 2-1, 1st and 2nd paragraphs: I would reverse the order of minor exceptions and major differences – the major differences should go first. I would not down play this aspect of the study. In fact, it is very important to convey how different these studies were and thus how difficult it was comparing them. This might be best summarized in a table format. Maybe expand Table 2-1.

Page 2-1, Baseline Study: It would be helpful to use a reference here to Smallwood's study (Smallwood and Thelander 2004).

Figure 2-1: This needs labels for each of the maps.

Page 2-2, Current Study: I think you need at least the same level of detail as Shawn provided in his report (01/06/10). And I think you should add the flow charts showing the filtering criteria that were developed earlier (M25-7).

Page 2-2, Analysis Methods: This first section is more about filtering the databases than analysis. Perhaps this first section should have the subheading like "Baseline and Current Study Databases". Also, there is no mention in this section of using or not using the CEC data.

Page 2-5, Carcass Removal Rate Estimates: Small birds needs to be defined. Scavenger swamping needs a reference.

Page 2-6, 1st paragraph: If it is common practice – it probably needs a reference. I am not sure that using turbine string has become common practice.

Page 2-6, 2nd paragraph: I think that installed capacity is poorly defined. It sounds the same as nameplate.

Page 2-6, 4th paragraph: It would help if you say how "year" is defined (calendar or bird year).

Page 2-6, 5th paragraph: I think the two sets of searches are for *both* the baseline and current study.

Page 2-7, Assessment of the 50% Reduction Goal, 1st paragraph: I suggest you describe all the ways the comparisons were made (all strings, operating group level, common turbines, APWRA wide).

Page 2-7, Assessment of the 50% Reduction Goal, 2nd paragraph: I think you should say that these were trends in the *adjusted estimates of fatalities* per year (so it matches what was shown in the Results Section under Trends in Fatalities over Time).

Page 2-7, Assessment of the Effectiveness of Seasonal Shutdown: I suggest you say that these analyses were done using only the common turbines (so it matches what was done in the Results Section under the same section heading).

Page 2-7, Assessment of the Effectiveness of Seasonal Shutdown: I don't know why you used discovery date in some of these analyses. I don't remember the SRC recommending this.

Page 2-7, Assessment of the Effectiveness of Repowering to Reduce Turbine-Related Avian Mortality: Although this states that the common string database was used, Table 3-9 of the results section says all monitored strings were used in this analysis.

Page 3-1, 1st paragraph: I think you should state that the 3,452 is from both studies.

Page 3-1, Table 3-1: I suggest that size class (small, medium and large) be defined in the methods section.

Figures 3-2 and 3-9, 3-10: These are both average monthly fatalities but they don't match. I think the difference is that 3-2 used discovery date and 3-9 and 3-10 used backdate. But I don't know if these were adjusted or unadjusted, or were all strings or just common strings. Again, labeling the figures would really help.

Page 3-1, Seasonal Variation in the Number of Fatalities: I can understand the spike in January after the shutdown of some turbines but why do you think there was a spike in October before the shutdown?

Page 3-1, Seasonal Variation in the Number of Fatalities, Figures 3-1 and 3-2: I suggest you compare these against the monthly baseline fatalities and don't use discovery date for either.

Page 3-1, Carcass Removal Rate Estimates: I think it would be better to put this whole section into the methods section.

Page 3-2, Comparison of Baseline and Current Study Mortality Rates, 1st paragraph: "Using the all strings dataset *and* the strings-level analysis method...." This sounds like two difference methods. I suggest saying all strings dataset with the string level analysis method. Same for the wording in the 2nd and 4th paragraphs as well.

Page 3-2, Comparison of Baseline and Current Study Mortality Rates, 1st paragraph, last sentence: I would take the word "much" out of the sentence about large raptors. There was still almost a 2 fold difference in large raptor mortality rates between the baseline and current study.

Page 3-2, Comparison of Baseline and Current Study Mortality Rates, 3rd paragraph: Maybe add that the difference between baseline and current study was also almost double in the common dataset. Why do you think the common dataset had the highest levels of mortality rates? Why do you think the operating group level had the lowest?

Table 3-4: The columns for non-raptor, common strings are added wrong.

Table 3-6: The columns for focal species, all strings, string level are added wrong.

Figures 3-5 to 3-8: I think these are adjusted rates, but putting it in the title would help.

Figure 3-6: Why is the yearly mortality pattern using operating group level so different from the other analysis levels for GE and AK? I would expect a difference in overall numbers but not in the pattern.

Page 3-3, 3rd paragraph: I don't agree with the GE fatalities tracking well with each of the three datasets (see above comment). And there is no GE spike at the beginning of the baseline study for operating group level.

Page 3-3, 5th paragraph: I would not call the 2006 all string number for GE a substantial peak, nor the following three years a decline for GE. It levels off in 2007 and then declines (which actually is only two years anyway).

Page 3-3, 7th paragraph, last sentence: The fatalities per search in Figure 3-8 do not appear flat to me. And rather than "contrary", Figure 3-8 appears to be fairly similar to the other graphs (Figures 3-5 to 3-7).

Page 3-4, Assessment of the Effectiveness of the Seasonal Shutdown, 1st paragraph: I don't know why you are using only raw fatalities for this analysis. This is not mentioned in the methods section. I would like to see adjusted fatalities as well. Again, I don't think that using discovery date is appropriate. But I would like to see Table 3-7 by year (from Baseline to Current Study).

Page 3-4, Assessment of the Effectiveness of the Seasonal Shutdown, 1st paragraph: I think it's a bit of a stretch to say there was an increase in GE fatalities from baseline to the current study if it was only an increase of one. If possible, it would be good to back statements like this up with some statistical tests.

Figure 3-9 and 3-10: Are these raw fatalities and at common turbines? I think these are backdated. Need labels. If these are raw fatalities, I would like to see the adjusted numbers as well. And I would like to see all the years shown separately for both baseline and current study.

Page 4-1, 1st paragraph: I suggest discussing some possible reasons why mortality did not decrease, such as only about 50% of the hazardous turbines rated 7-10 were removed, turbine attrition has resulted in many gaps in turbine rows, vacant towers at the end-of-rows which were used as flight diverters have not been removed, and derelict turbines are still abundant.

Page 4-1, 2nd paragraph: I suggest you add that operating time likely fluctuates each year as well and also makes it difficult to detect trends over time. And what ever happened to the operating data? I thought you had this for some of the companies.

Page 4-1, 3rd paragraph: I don't think you can say that the current study was not designed to produce estimates of APWRA annual fatalities. I believe that was one of the goals.

Page 4-1, 3rd paragraph: Where was the CEC data set excluded from the data analysis?

Page 4-1, 3rd paragraph: How does the use of the common dataset minimize the differences in sampling *intensity*?

Page 4-1, 3rd paragraph: How much smaller was the number of baseline fatalities after the parsing of data?

Page 4-1, 4th paragraph: Being more efficient doesn't necessarily mean you find more birds. And the learning curve for finding birds is probably steep in the beginning but after a while it likely tapers off. So unless the baseline search crews were changing every 6 months, I don't think detection probabilities would be that much greater in the current study.

Page 4-1, 5th paragraph: In your comparisons, I suggest you make it clear that you used the carcass removal rate for small birds that had lower rates than Smallwood (2007).

Page 4-2, 2nd paragraph: Again, I do not see the raptor fatalities per search being at odds with the adjusted mortality estimates.

Page 4-2, 5th and 6th paragraphs: These paragraphs seem better suited to the seasonal shutdown section below.

Page 4-2, 5th paragraph: I do not see the differential reduction in winter fatalities for RTH. Nor is this presented in the results section.

Page 4-2, 6th paragraph: Why wouldn't large raptors become acclimated as well?

Page 4-2, Assessment of the Effectiveness of the Seasonal Shutdown, 1st paragraph: There is no discussion as to why these are confounding factors.

Page 4-3, Assessment of the Effectiveness of Repowering to Reduce Turbine-Related Avian Mortality: Were burrowing owls also the exception with regards to reduction in annual mortality at the Buena Vista site?

Page 4-3, Conclusions: I think you need a little caveat here. Without incorporating the bird abundance and the operating data these are still preliminary findings.

I also had several editorial and grammatical comments, but these are best dealt with in track changes and can wait until the next revision.

K. Shawn Smallwood

6 January 2010

I reviewed the Alameda County Avian Monitoring Team's report of fatality monitoring in the Altamont Pass Wind Resource Area (APWRA) and whether the 50% reduction goal was reached in the Avian Wildlife Protection Program. I agree with the grand conclusions reached by the monitoring team, summarized in its Conclusions section on page 4-3. I cannot imagine how any other conclusions could have been reached. However, I also view the monitoring team's conclusions as strong evidence that the data are robust to how they are processed and analyzed, because I disagree with how the monitoring team's conclusions were reached. Consistent with what SRC members recently told the monitoring team, I believe it matters how conclusions are reached. Therefore, my comments on the monitoring team's report are critical, even though I agree with the overall conclusions. I hope that my comments are used constructively and a revised report can be produced.

Below is a summary of my general criticisms of the report, followed by comments on specific portions of the report.

SRC Recommendations: The monitoring team could have implemented more SRC recommendations than it apparently did. I found a disturbing number of examples in which the monitoring team did not heed SRC recommendations, including (a) making use of company-provided installed capacity data without taking any of the steps to ensure that these data were sound, (b) apparent use of calendar year rather than bird year in making annual fatality rate estimates, and (c) use of sampling units inconsistent with the sampling design.

I felt blindsided by reading new scavenger removal rates in the report, even after being assured as recently as 3 December 2009 (the last conference call) that the rates were essentially unchanged after recent reviews of the 48-hour search data (M32). I was surprised to read in the report that the monitoring team decided to include old carcasses in fatality rate estimates, to exclude WRRS and incidental records, to rely on carcass discovery dates rather than backdates in estimating monthly fatality rates, and to exclude non-native species, which I believe was unnecessary. I was surprised to not read of three methods used to adjust for removal rates of feather piles, because the monitoring team agreed to follow the SRC's recommendations on this approach. These surprises give me the impression that this monitoring team has little regard for the SRC as the principal "client" of the report.

Use of Scientific Literature: I was disappointed with the lack of use of the available scientific literature. Multiple peer-reviewed papers and reports are available and could have helped with the interpretation of patterns in the data. I was also frustrated with the monitoring team's misapplication of searcher detection and removal rates summarized in Smallwood (2007).

Data Analysis and Presentation: Repeating a mistake from M32, and which I previously commented on, this report represents years when no fatality monitoring occurred. It mischaracterized search intervals and study duration in the baseline period. It invented sampling units (operating groups) that differed from the sampling units in the 2005/2006 sampling design, and did not make use of the original sampling units. It also used scavenger removal rates that were much slower than can explain the number of carcasses actually found using 2-day search intervals in the KB study. The report includes factually incorrect and incomplete statements, some of which are identified in my comments below on specific sections. It also appears to have glossed over the complications in the monitoring at Tres Vaqueros, where monitoring was taken over by East Bay Regional Park District in 2006 and all monitoring at these turbines ceased in fall 2007.

Lapses in Transparency: Though I cannot tell that the operating data were used in this report, I am still troubled by the exclusion of the SRC from reviewing those data. Sharing the operating data with the SRC would not have affected the competitive advantage of the wind companies in the APWRA, so they should not have been kept confidential from the SRC. Since its first meeting, the SRC repeatedly recommended that power output or operating time data be shared with the SRC, but some data were only produced for the monitoring team, and now it looks as though those data were not even used.

I understood that the monitoring team was recording turbine status during the past four years of monitoring, yet those data were not posted on the SRC web site or used in this report as far as I can tell. Why were these data unavailable for public review and why were they not used in the report?

The revised data from the KB study were supposed to be shared with the SRC and scanned versions of the data sheets posted on the SRC web site. In fact, the data were not shared.

Specific Comments

Page 1-1, 2nd bullet at bottom: The SRC's role is broader than described in the monitoring report. It includes making recommendations to the County on measures to reduce fatality rates, as well as performing analysis of data, among other duties.

Page 1-2, 2nd paragraph: Repowering old-generation turbines with modern turbines has not been considered a "management action." Repowering would involve new projects to be preceded by environmental reviews, which differs from the manner in which mitigation measures have been handled over the past several years.

Page 1-2, 3rd paragraph: I suggest explaining the reasons the SRC recommended discontinuing the crossover winter shutdown design. One reason was that the shutdown turbines could kill many raptors when they are reactivated in the middle of winter, due to

habituation. Another reason was poor execution of the design by the monitoring team, including search intervals that exceeded the duration of shutdown periods.

Page 1-3, 2nd paragraph and Fig. 1-3: The legend of Figure 1-3 should cite the source of the data depicted. As I warned previously, I will not accept the claim that installed capacity has declined in the APWRA in the manner depicted in Figure 1-3. The SRC recommended that before the monitoring team uses these data for the purpose of extrapolating fatality rates, one or more of the following steps should be taken. One, the companies can explain the discrepancy between these data and the installed capacity data they supplied to the CEC. Two, the companies could provide to the monitoring team and SRC a written statement vouching for these data. And three, the monitoring team could use their data on turbine status to verify the alleged reduction in installed capacity. None of these steps were taken, nor were the company-provided operating data compared through time to verify the rated capacity trends claimed by the companies. I cannot accept extrapolations of fatality rates against company-provided installed capacities until one or more of the above steps are taken.

Also, it is not true that the repowering projects added to the overall megawatt capacity of the APWRA. They simply upgraded the turbines within the same MW capacity, though the Buena Vista project gave up slightly more than 3 MW in capacity.

Page 2-1, 1st paragraph: I disagree with the description of how fatality determinations were made in the baseline study. First, the fatality monitoring crew was to record all found bird remains, so that was a fundamental, simple standard. Fatality search personnel indeed made preliminary determinations of whether a carcass represented a fatality, but the final decision was made by me according to a set of standards described in Smallwood and Thelander (2004, 2005). The carcass had to have been of a bird killed within 90 days of discovery (based on an estimate of time since death), it had to have been nearby a wind turbine, and it had to have consisted of at least one bone or of multiple feathers connected by enough tissue to suggest significant trauma. I believe it is an overstatement to claim that no standard fatality definition was used in the baseline study. I also believe that any differences in *a priori* written standards had little if any influence on differences in fatality rates between the baseline and current study periods.

Page 2-1, first paragraph under Baseline Study: The report should explain that 2,548 turbines were selected systematically from the remaining pool of turbines after the first 1,526 were searched.

Page 2-1, 2nd paragraph under Baseline Study: Every statistic in this paragraph is incorrect, misapplied, or misleading. Also, the name attributed to NREL is incorrect (appears to have been a guess).

Figure 2-1: The years depicted in the figure are incorrect. Also, the representation of the sampling effort is misleading. Some of the representation makes no sense to me.

Page 2-2, 1st paragraph under Current Study: The study design was not described entirely accurately. All very small turbines (40-65 KW) were searched, all large turbines were searched (250-400 KW), and all Diablo Winds turbines were searched. The only turbines in the randomly selected blocks were medium-sized turbines (95 to 200 KW), except for turbines owned by Northwind Energy, who refused to cooperate with the study.

Page 2-3, 1st full paragraph: I disagree with the decision to exclude incidental and WRRS fatality records from the analysis. These fatalities tend to be of large raptors such as red-tailed hawks, great horned owls, barn owls, prairie falcons, and ferruginous hawks. They are the types of carcasses that likely would have remained long enough to be detected upon the next fatality search and which can significantly affect estimates of fatality rates. Not only do I disagree with this decision, I'm surprised by it because I don't recall the monitoring team ever telling the SRC it planned to make these exclusions.

Page 2-3, 2nd paragraph: I disagree with the monitoring team's decision to exclude non-native species from estimates of fatality rates. It is a simple matter to separate out the estimates of different species and species groups, so I do not understand the need to exclude non-native species.

I also disagree with the decision to include carcasses older than 90 days when estimating fatality rates. I disagree that fatality rate estimates representing a certain time period should include fatalities that could have occurred a year to several years before that represented time period. The monitoring team rationalized their inclusion of older fatalities by claiming that a higher proportion of baseline fatalities had been estimated to have died >90 days before discovery (also see Table 2-5). However, I did not find this claim to be true. After screening out fatalities for other reasons (cause of death, etc.), the proportions of old versus recent fatalities did not differ between the baseline and current study periods (see Table below). The percentage of old fatalities was 11% in the baseline period amongst the wind turbines for which comparable fatality rates could be estimated, and it was 12% in the current study.

Study	Fatalities determined to have been possibly, probably, or certainly caused by wind turbines			Total
	Missing ETD	≤90 days since death	>90 days since death	
Baseline study	0	993 (84%)	186 (16%)	1,179
Set 1 turbines only	13	822 (89%)	103 (11%)	925
Current study	1	3,728 (88%)	521 (12%)	4,250

Page 2-4, bottom: Another major factor affecting estimates of fatality rates likely includes crippling bias, but at this time we have no means of addressing this quantitatively.

Page 2-5, Searcher Efficiency: The monitoring team appears to have misapplied searcher detection rates reported in Smallwood (2007), and in so doing biased fatality estimates low for nonraptor species. The correct rates were 100% for large raptors, 75% for small raptors, which were the rates the monitoring team reportedly used for large and small birds, respectively. However, Smallwood (2007) reported rates of 80% for medium and large-sized nonraptor birds, 78% for rock pigeons, and 51% for small nonraptor birds. These were the values that should have been used for nonraptor species other than bats.

Page 2-5, 3rd paragraph under Carcass Removal Rates: I do not understand the parenthetical statement related to carcass inclusion. Specifically, I do not understand the inclusion of carcasses only if they were found after at least two previous searches. What does this mean?

Page 2-5, 5th paragraph: By using found carcasses in scavenger removal trials, the analyst does not know how many of the carcasses had already been removed by scavengers. Thus, the approach did not avoid bias.

Page 2-6, 2nd paragraph under Estimating Mortality Rates: There was no need to use installed capacity to extrapolate rates to estimate annual APWRA-wide fatalities. Fatality rates based on nameplate capacity could just as well have been extrapolated to the permitted nameplate capacity of the APWRA, i.e., to 580 MW. The basis was solid for extrapolating fatality rates to nameplate capacity, whereas the basis was weak for extrapolating rates to installed capacity. For one thing, the monitoring team could not have known the installed capacity in monitored turbine rows from year to year during the baseline study. In my opinion, the monitoring team used a flawed methodology for extrapolating fatality estimates from installed capacity.

Page 2-6, 3rd paragraph under Estimating Mortality Rates: The averaging of fatalities by operating group was inconsistent with the sampling design of the current study. The study design had divided the APWRA into four regions, and within each region selected blocks of medium-sized turbine rows randomly. Otherwise, all very small turbines and all large turbines were selected for monitoring. In this way, only one extrapolation of fatality rates was required, and that was from the randomly selected blocks to the population of medium-sized old-generation turbines. I am not averse to estimating fatality rates within groups other than the original sampling units, but why not also use the original sampling units? I thought the SRC discussed this matter with the monitoring team.

Page 2-6, 4th paragraph under Estimating Mortality Rates: What was the basis of the annual mortality estimate? Was it the calendar year or the “bird year,” as recommended by the SRC? The monitoring report does not define the year, but the results suggest to me that the calendar year was used. I thought we agreed the monitoring team would estimate annual fatalities by bird year.

Page 2-6, 5th paragraph under Estimating Mortality Rates: The first approach to estimating baseline fatality rates appears to have been inconsistent with the recommendation from the SRC. I thought we all agreed that the Set 2 turbines, which were the turbines searched only twice, would not be included in fatality rate estimates.

Page 2-7 under Assessment of Effectiveness of Seasonal Shutdown: I disagree with the decision to assume that all baseline fatalities estimated to have died >90 days before discovery actually died 90 days before discovery. This approach arbitrarily lumped fatalities into time periods 90 days before discovery even though we believed many of these birds had been dead up to one to several years prior to discovery. Old weathered bones did not represent birds killed 90 days ago. Including these finds, which are typically discovered during first searches, likely introduced noise to the analysis of the effectiveness of the winter shutdown. I don't understand the logic behind this decision.

Page 3-1 under Seasonal Variation in Numbers of Fatalities: I disagree with the decision to use carcass discovery dates rather than backdates in the monthly comparisons. We know that there is plenty of error in the carcass backdates but we also know that the carcass discovery dates are usually not when fatalities occurred, so why wouldn't one use the estimated dates of death when making monthly comparisons of fatality rates?

Furthermore, the fatality counts per month appear to be unadjusted for searcher detection error and scavenger removal error. I know from experience that the monthly graphs will appear different when the fatality counts are transformed into adjusted rates. In fact, I would suggest that monthly fatality counts are potentially misleading and of little use for assessing impacts or formulating mitigation measures.

Tables 3-5 and 3-6: I do not trust the estimates of fatality rates for American kestrel and burrowing owl. I suspect they were biased low due to the scavenger removal adjustments that were used. The removal curves in Figure 3-4 appear very different from what the SRC had been shown previously. I would like to review the data. The monitoring team agreed to share the recently modified data with me, but I have yet to see them despite my email request for the data.

Keep in mind that according to report M32, the monitoring team found 16 small raptor fatalities during the 2-day searches of the KB study and 4 during the 30-day searches at the same turbines. Accounting for the 2 that I believe were found by both search teams, the net total of 18 small raptors were found over only 4 months at 19 MW of nameplate capacity, even though those turbines were searched out to only 40 m instead of the usual 50 m. Even unadjusted for searcher detection error and scavenger removal rate, this rate of fatality discovery indicates the average annual number of fatalities had to have been considerably larger than the monitoring team estimated.

I do not know how the monitoring team developed their scavenger removal rates from the KB study because recent modifications to data processing have yet to be described, but I suggest their methods are in need of review. The original purpose of the KB study was to (1) test whether more frequent searches would detect more fresh carcasses and (2) track

the fate of fresh carcasses. Both objectives were intended to establish superior carcass removal rates (SRC meeting notes of 23 Oct 2006). The monitoring team appears to have focused on objective 2 and perhaps forgot to check the results from objective 2 against those of objective 1. The removal rates used by the monitoring team would indicate considerably slower removal than predicted by the model predictions in Smallwood (2007), whereas the actual numbers of carcasses found during the KB study would indicate faster rates than predicted by Smallwood (2007). As summarized in Smallwood (2009), when performing 2-day searches the monitoring team actually found 1.41 times the number of small raptors and 2.43 times the number of small nonraptor birds than expected based on the Smallwood (2007) removal models applied to the results of the 30-day searches. If the monitoring team actually found more birds than expected based on the removal rates in Smallwood (2007), then how is it that they developed slower removal rates after monitoring found carcasses?

Figure 3-5: The years in the graphs appear to be incorrect. No fatality searches were performed in 1997. Searches did not begin until spring 1998. Also, were the years depicted calendar years or “bird” years? The numbers of golden eagles per year are wrong (see my report), including nearly 0 eagles in the second year, when I found eagle fatalities to be higher rather than lower.

The monitoring team depicted eagle fatalities as peaking in the last year of the baseline, but I found they had declined. There are numerous discrepancies like this, suggesting to me that the fatality monitoring team estimated fatality rates per year in precisely the manner the SRC recommended they not do so – by calendar year with some turbine strings represented by a portion of a year, etc. It also appears the monitoring team included the turbine rows searched only twice in the baseline, which the SRC asked them not to do.

Figure 3-9: It seems odd representing average values as bars in bar graphs. I would have expected error bars.

Table 3-9: It would help to specify in the table legend that the estimates were derived from Diablo Winds.

Page 4-1, top paragraph: I don’t agree with the expectation and reasons for the expectation presented. I do not agree, necessarily, that installed capacity declined to the levels reported by the monitoring team, and I do not agree that the crossover shutdown design should have resulted in reduced fatalities. Further, the monitoring report made no mention of the end-of-row derelict towers used as flight diverters, and how the SRC felt that these towers increased the hazard to raptors.

Page 4-2, third full paragraph: I disagree that it is difficult to explain the inability to detect declining fatality rates in the face of mitigation actions during the current study. The mitigation actions were marginal at best, so it was unlikely to result in a noticeable effect.

Page 4-2, 6th paragraph: I don't agree with the monitoring team's suggestion that fatalities might have increased due to development in the region pushing more raptors into the APWRA. Between the baseline and current study periods, there could not have been enough land conversions to explain the increases in fatality rates. I doubt that the APWRA was any more crowded with raptors in the current study than during the baseline period. I have not seen any evidence from my 20 years of bird surveys in the Central Valley that crowding of raptors increased anywhere outside of spreading development.

Page 4-2, Assessment of the Effectiveness of Seasonal Shutdown: I disagree with the monitoring team's conclusion that it was surprising not to detect an effect of the winter shutdown. One of the reasons for the recommendation of a four-month shutdown was the need to detect an effect. It was recognized in 2005, when I first assessed the potential of this measure, that the measure was a tradeoff between saving birds and losing power generation – winter was not the optimal time to shut down turbines to save birds. The optimal time was recognized then and now as summer/fall. However, the recommended four-month winter shutdown was never implemented, so the power to detect an effect was no greater than tiny. I've been saying this repeatedly for four years. For the reasons I gave and for the reasons discussed by the monitoring team in the report, the SRC recommended to the monitoring team this past fall that it give minimal time and attention to assessing the effectiveness of the crossover winter shutdown. There should be no surprise about not detecting an effect. (That said, I feel that I did detect some effect.)

Page 4-3, Conclusions: Whereas I disagree with the methods used by the monitoring team and whereas I disagree with multiple statements in the report, I cannot agree more strongly with the monitoring team's final conclusions. Our coming to the same overall conclusions indicates that the data are robust to methodology and interpretation. I also want to take this opportunity to applaud the monitoring team for working so hard in recent months to prepare the fatality, search history, and turbine attribute data. I feel that the data from the current study have been greatly improved since a couple of years ago.

References

- ICF Jones & Stokes (no authors identified). 2009. Draft Report M32: Altamont Pass Wind Resource Area 48-hr search interval bird fatality study. Report to Alameda County Community Development Agency, Hayward, California.
- Smallwood, K. S. 2007. Estimating wind turbine-caused bird mortality. *Journal of Wildlife Management* 71:2781-2791.
- Smallwood, K. S. 2009. Smallwood's review of M32. Alameda County SRC document P-111. 6 pp. http://www.altamontsrc.org/alt_doc/p111_smallwoods_review_of_m32.pdf
- Smallwood, K. S. and C. Thelander. 2004. Developing methods to reduce bird mortality in the Altamont Pass Wind Resource Area. Final Report to the California Energy

Commission, Public Interest Energy Research – Environmental Area, Contract No. 500-01-019. Sacramento, California, USA.

Smallwood, K. S. and C. Thelander. 2005. Bird mortality in the Altamont Pass Wind Resource Area, March 1998 – September 2001 Final Report. National Renewable Energy Laboratory, NREL/SR-500-36973. Golden, Colorado, USA.

Julie Yee

January 11, 2010

SUMMARY. Many important elements of the previous version of the M21 report (June 2008 draft) have been omitted in the current draft. While I don't believe it is necessary to repeat content across multiple reports, it would be good to bring some those elements back into the final version to make this a self-contained report. I think it's particularly informative to include raw counts by focal species and other species groups, such as Table 4 of the June 2008 draft.

I was skeptical of the findings for no shutdown effect, and was not sure how to interpret the comparison between baseline and current survey period without having any error estimates. Therefore I performed my own analyses on the fatality dataset. I have not had time to fully write up these analyses, but it is essentially based on the same approach I wrote up in document P55 and illustrated in documents M16, except modified in light of recent discussions during the past 2 years. I would also need to consult with the monitoring team about the data in the database before I finalized the results. In brief, I did find a shutdown effect for red-tailed hawks and American kestrels. The golden eagle species had too few fatalities to make any inferences, and the burrowing owl followed a completely different pattern than the other focal species. I further modified the model to examine differences between the baseline and current survey periods. My findings concur with those in M21, showing large increases, however these increases were not statistically significant. I attribute this to the large interannual variation within periods, especially in the Current Survey period. In fact, I applied my model to estimate the difference between biennial subperiods 2005-2006 versus 2007-2008 and found statistically significant decreases for both the burrowing owls and red-tailed hawks in the range of a 50-60% decrease. I feel rather confident that a real decrease occurred in those years. However, given the short duration of the subperiods, and large interannual variation, it is also conceivable that this decrease occurred in those years by random. Further monitoring will be critical for determining long-term decline. My analyses require proofing but I feel that these statistical significant findings are not likely to change even after I resolve my data questions.

I have listed my comments in order of appearance in the M21 report.

Page 1-2, paragraph 3. As a clarification, the report should note that the crossover study reversed the treatment assignments in the 2nd year, so that turbines receive opposite treatments (shutdown or no shutdown) in the two consecutive years (hence the term "crossover").

Page 1-2, paragraph 4. As a minor correction, the crossover experiment was designed to last two winters and end in February 2007, which it did. The SRC did not begin meeting

until September 2006. Later, the SRC called into question the effectiveness of the study and recommended not to continue it further.

Page 2-1, paragraph 2. Appendix A was not attached.

Page 2-1, paragraph 5. Always include the denominator, such as *per year*, for annual statistics wherever it is otherwise ambiguous. For example, I think here it was supposed to read “strings were surveyed an average of 5.8 times *per year* with an average search interval of 52 days *per year* over a period of approximately 55 months.” And for the CEC phase of the study, which was a 9-month survey period, is it correct to say that those statistics are per the 9-month period? There is a similar type of statement on p 2-2, paragraph 4.

Page 2-2, Table 2-1. What do the minimum and maximum values of the Strings Searched range represent? Different calendar years? Different bird years? Different months?

Page 2-2, paragraph 1. 4th sentence: Typo “plot” instead of “block”? 5th sentence: Are those operating groups listed (Diablo Winds, Tres Vaqueros, and Altech) equivalent to the “medium” and “very small” turbine strata which document R29 previously described to be completely rather than partially surveyed? This is important.

Page 2-2, paragraph 2. The reference to transects may be confusing to unfamiliar readers who don’t know that the surveys are conducted by walking along transects within the search area.

Page 2-2, paragraph 3. For clarity, I would edit or omit the last phrase “from birds found during previous monthly searches.” Double-counting can occur when there are repeated finds of the same fatality from current as well as previous searches.

Page 2-2. Analysis Methods. Identify range of dates analyzed. The previous sections provided this information somewhat, however the database appears to contain records outside of the range of dates provided for surveys. I’m especially uncertain whether these analyses only pertain to data through September 2009, or beyond.

Figure 2-1. The dates at the top of the figure adds to the confusion about dates analyzed. I assume that these are bird years, so that 1997 represents Oct 1997 through Sep 1998. Is that correct? If so, then why is 1997 included? I thought the report was to analyze full bird years starting with 1998 (Oct 1998 – Sep 1999).

Page 2-3, paragraph 3. The description of the fatalities >125 m from turbine might be misleading because, although not explicit, the wording is suggestive that the exclusion was due to having small and similar proportions in Table 2-4. However, it was decided to exclude those fatalities mainly because of the variable and highly uncertain detection coverage in those ranges. The fact that the proportions are small only provides an assurance that not too much data are excluded.

Tables 2-2 through 2-5. These tables are great. It is very helpful to see how the various filters whittle down the fatality sample size.

Figure 2-2. The caption is not very clear. Does the pie represent only non-raptor fatalities? If so, I suggest following the example on captioning in Table 2-3, i.e. “Proportion of Non-Raptor Fatalities Comprised of Native and Non-Native Non-Raptor Species”

Page 2-5, Carcass Removal Rate Estimates. I recommend noting that the traditional applications of survival time analysis is being used in this analysis not for the usual purpose of analyzing *survival* time of anything, but rather for the remaining (or persistence) time of fatality evidence.

Page 2-5, paragraph beginning with “In the 48-hour Search Interval Study”. It could be misleading to say that any data were censored because of discontinued monitoring. In censored survival time analysis, it is the survival times that can be censored, but the available data up to the censor time is completely included in the analysis.

Page 2-5, next paragraph. Elaborate a little more on the “regular but increasing intervals”

Page 2-5, last paragraph. The common terminology and spelling is “Kaplan-Meier product limit estimator”

Page 2-5, last paragraph. I still have the same criticism of this survival analysis as in my previous review of M32. A least squares regression on survival point estimates, K-M or not, is probably not appropriate because the survival estimates are not independent from each other. My concern is that this approach allows a regression to be drawn with an arbitrarily high or low amount of error depending on the number and location of survival point estimates used. To handle the independence issue, it would be better to derive the curve and its standard errors directly from an appropriately fitting parametric survival time model and bypass the least squares regression altogether.

Figure 3-3. What happened to these curves to change it so much since the last draft of M32???

Page 2-6, paragraph 2. These are very useful and important definitions for nameplate capacity and installed capacity. The definition of *installed capacity* only includes megawatt capacity of *installed* turbines, correct? This is not explicitly stated. Also, I’m confused by the definition description for how these two types of capacities were used to calculate mortality rates and extrapolate to the entire APWRA. As a general rule, the denominator which the rates are based upon (report says nameplate capacity) should match how the extrapolation is performed (report says installed capacity). I thought the extrapolation was going to be by nameplate capacity. After performing the extrapolation, then it may be desirable to divide by the total installed capacity to calculate the fatality rate per installed MW capacity. i.e.,

$$\text{Fatality rate} = \frac{\text{fatalities at surveyed turbines}}{\text{nameplate capacity at surveyed turbines}}$$

$$\text{Total APWRA fatalities} = (\text{Fatality rate}) \times (\text{Total APWRA nameplate capacity})$$

$$\text{Rate of fatalities per installed MW} = \frac{\text{Total APWRA fatalities}}{\text{Total APWRA installed capacity in MW}}$$

Page 2-6, paragraph 3, 3rd sentence. At the string level, when averaging the fatality rate across strings, was the average weighted? Weighting by the nameplate capacity for each string would be appropriate. Did the same kind of averaging and expansion occur at the operating group level?

Page 2-6, paragraph 3, last sentence. Detailed information about how operating groups were paired should be provided, preferably in the form of a description and table.

Page 2-5, paragraph 5. Some further details on the subsets of data would be useful here. In particular, the numbers of fatalities for each species groups for the *all strings* and *common strings* datasets. Also, the description for the *common strings* dataset is not very precise. For example, I can't tell how were strings handled when they were monitored with a mixture of <60 day and >60 day intervals over the course of three consecutive but incomplete years? I'd like to see a more technical description. It would also help to have the *common strings* identified in the database.

Page 2-7, paragraph 1. These rates were calculated as the *most conservative* in what sense?

Page 2-7, paragraph 2. Was this index calculated using unadjusted or adjusted fatalities? Is it correct to assume that "fatalities" always means unadjusted fatalities? I would rather see it explicitly stated.

Page 2-7, paragraph 3. I could not find backdate data in the database. If the details are omitted, then at least provide sufficient references.

Page 2-7, paragraph 4. There is an inconsistency here in the definition of winter. It says Nov-Jan, whereas I thought we agreed on Nov-Feb to cover the range of seasonal shutdown months. Also, the results cite Nov-Feb.

Page 2-7, paragraph 4. Intuitively it doesn't sit right with me to use fatalities backdated 90 days when the search interval range is restricted to 60 days. There is a history of applying 90-day backdates and a reason for restricting to 60 day search intervals, so I am not faulting the Monitoring Team on this. But I'm not sure the combination of the two rules makes sense, or what amount of error it could cause. What if the fatality was discovered in a <60 day interval but then backdated 90 days into a >60 day interval? Is the fatality then excluded? And what if the fatality was discovered in a larger (>60 day)

interval but then backdated into a smaller, say 30 day, interval? Is it then appropriate to adjust the fatality based on the smaller interval? I'm interested in better technical details on how the Monitoring Team handled these types of situations, and how many fatalities fell into this category. It would also be helpful to have a description of the typical circumstances in which this occurred (i.e. was it typically a certain season, certain years, certain geographic area, or certain operating groups?).

Page 2-7, paragraph 5. The description of the analysis to assess the effectiveness of repowering does not make sense. It reads that the rates were calculated using the *common strings* dataset for the *current study*. Was the analysis restricted to just the current study? (No years are provided). If so, then why restrict the dataset down to just the common strings? I can understand if it was a way to control for geographic variation (i.e. Diablo Winds turbine strings tending to occur in more geographically similar locations as other strings when using the *common string* set rather than the *all strings* set). But then Table 3-9 in the Results section refers to using *all monitored strings*. So now I really don't know what period or strings these rates represent.

Page 2-7, paragraph 6. It's very ambiguous how an APWRA-wide annual fatality rate was calculated by using just the Diablo Winds operating group rates.

Page 3-1, paragraph 1. The descriptions are inaccurate. The proportions of raptors to non-raptors are not very "roughly equal" between the two studies (58-42 for Baseline versus 43-57 for Current study). Small birds do not make the majority in either study period – in fact, none of the bird size classes is a "majority" because they are all fewer than 50%. It is at least true that small bird class is the greatest size class in the current study, but that is not true for the baseline study when large birds was the greatest class. Also, there is an error in Table 3-2 (Large birds in current study is 1020/2588 and should be calculated as 39%).

Page 3-1, paragraph 2. The report tries to attribute the spike in fatalities in January as a post-seasonal shutdown effect. This is an incomplete association because January is the post-shutdown month only for one of the four winters (2007-08). This is not to say the association is inaccurate, but the other years need to be considered separately. Also, since the north and south geographic areas had different shutdown periods during the first two winters of the current study, then that would also have to be factored into the analysis as well.

Figure 3-2. I'm trying to understand why these figures look different than the figures in Figures 3-9 and 3-10. Is it because different datasets were used (perhaps all strings data for Figure 3-2 and common strings data for Figures 3-9 and 3-10)? If so, this should be indicated in the captions. Also, are these unadjusted or adjusted fatalities?

Page 3-2, 3rd complete paragraph. There is a reference to Figure 3-4 in relation to Shawn's curve, so I recommend including Shawn's curve on the Figure.

Page 3-2, Comparison of Baseline and Current Study Mortality Rates. The recommendation by the SRC was to report the unadjusted rates but not to use them as estimates for rates of birds killed. This section reports unadjusted and adjusted rates, describing them equally as estimates of fatality. This can lead to misinterpretations by readers who may try to use unadjusted values as one of the estimates of fatality. For example, the lower range for burrowing owls killed is cited in the last sentence as 62, which is based on an unadjusted rate.

Tables 3-3 through 3-6.

- a) Indicate what the error values represent (after the “±” symbol), such as standard deviations or standard errors.
- b) Given all of the discussion we’ve had about the differences in search interval sizes between baseline and current study, and their potential for biasing the comparison between the baseline and current surveys, I’m very surprised that the proportional effect of the adjustment on the fatality rates was so similar between the two studies and even among the four species. For example, if I were to take the ratio of adjusted fatalities over unadjusted fatalities in the “All Strings Dataset / String-Level Analysis,” then I get ratios ranging from 1.8 to 2.2 across the four focal species and both periods. I would have expected the ratio to be larger in the Baseline period than in the Current Study period, and also larger for small raptors than for large raptors. However, there are no such patterns. I repeated the same exercise for the other two combinations of analysis type and again there was no variation in the ratios (“All Strings Dataset / Operating Group-Level Analysis”: ratio ranged 2.6-3.2 and “Common Strings Dataset / String-Level Analysis”: ratio ranged 1.9-2.1).
- c) Why is there not an Operating Group-Level Analysis using the Common Strings Dataset?
- d) There are numerical typos in Table 3-6 for the “Total focal species” and “Total raptors” under “All Strings Dataset / String-Level Analysis.” (If adding down the column, then 341 should be 541, 792 should be 925, 309 should be 456, and 728 should be 882).

Page 3-3, paragraphs 3 and 7. The report tries to attribute spikes in figures 3-5 through 3-7 to clearing searches at the beginning of the CEC and current studies. However, the spike at the beginning of the current study appears for (bird year) 2006, which is actually the second year for most of the turbines in the current study. I don’t see how those spikes can be attributed to clearing searches. I’m skeptical about the value of including fatalities found during clearing searches in the analysis.

Page 3-4, paragraph 2. I don’t understand why the report states a decline from November through December. Where is this evidenced?

Page 3-4, Seasonal Shutdown. I was surprised not to see any evidence of a fatality reduction due to seasonal shutdown, especially since I saw strong indications of shutdown effects in a model-based analysis I performed in 2007 on preliminary data. The possible explanations could be: 1) there is no shutdown effect and my model was incorrect, 2) there is a shutdown effect and the method in the M21 analysis is too crude to identify, and/or 3) neither analysis is incorrect but two years of updates to the dataset have reversed the patterns indicated by the old version of the data.

To investigate these possibilities, I reran my model using the current dataset (Current Study only and restricted to intervals < 60 days) and again found significant results including a reduction in fatalities during shutdown status for American kestrels and red-tailed hawks, no reduction for burrowing owls, and insufficient data for golden eagles. This eliminates the 3rd explanation.

I believe the 2nd explanation (i.e. M21 shutdown analysis is too crude) is likely because the data in Tables 3-7 and 3-8 aggregate the winter period (Nov-Feb) which includes varying mixtures of searches during operating and non-operating statuses.

To investigate the 1st explanation, I tried other models which produced consistent results. Since different seasonal variations among years could confound the result of a model, I repeated my model using just winter data buffered before and after by non-shutdown months (altogether, October through March). I also repeated this model by restricting to two year subsets (bird years 2005-2006 and 2007-2008). These model variations produced the same results for red-tailed hawk (strong shutdown effect) and burrowing owl (no shutdown effect), but the restrictions to the data led to an incomplete analysis of American kestrels; the restricted model estimated a shutdown reduction but could not construct a standard error or significance level.

Finally, I constructed simple descriptive graphics by taking unadjusted fatality counts from each string search in which the string was shut down either 0% or 100% of the interval (Figures 1-4). These figures are consistent with the modeling results.

Figure 1. American kestrels

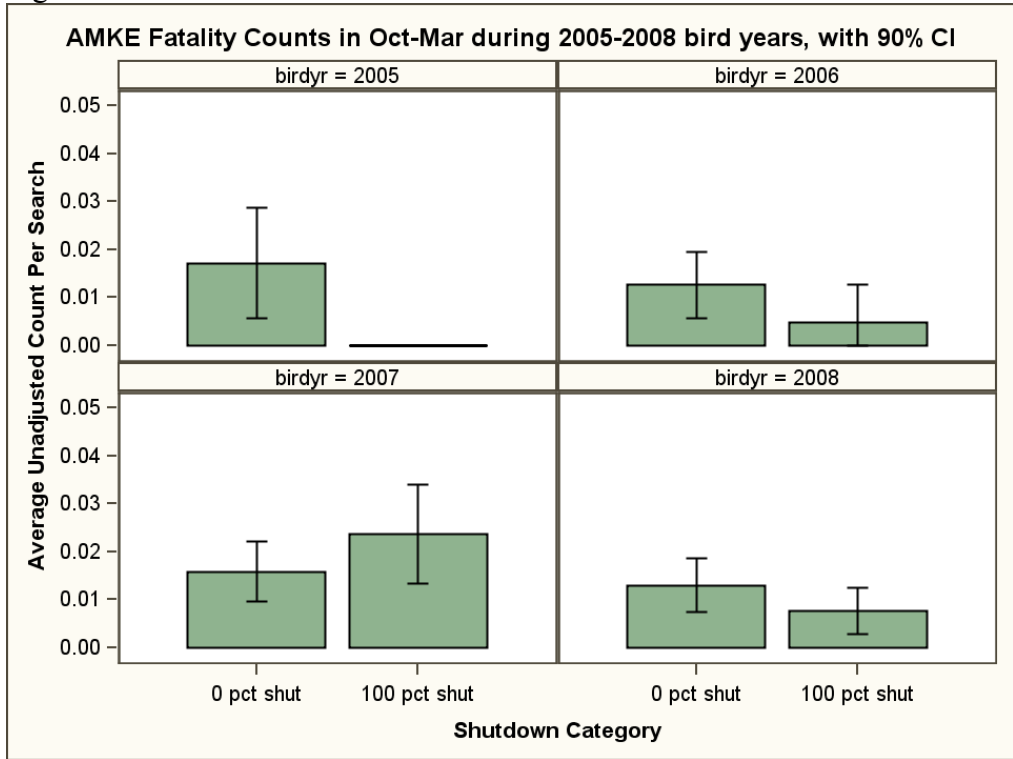


Figure 2. Burrowing owls

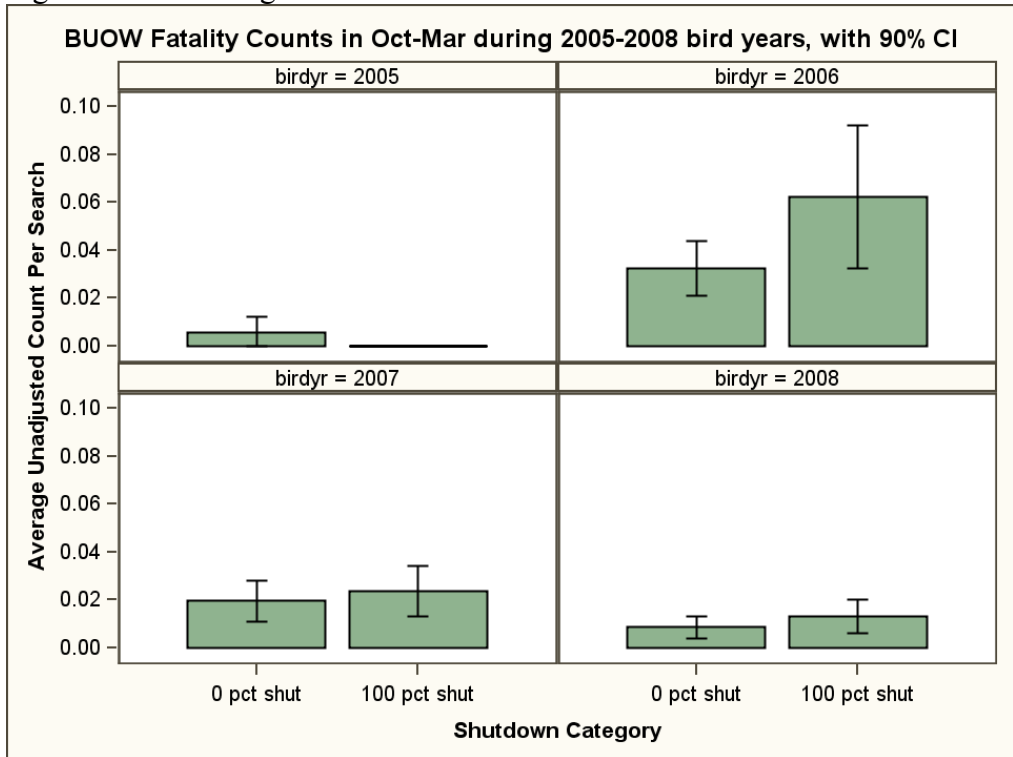


Figure 3. Red-tailed hawks

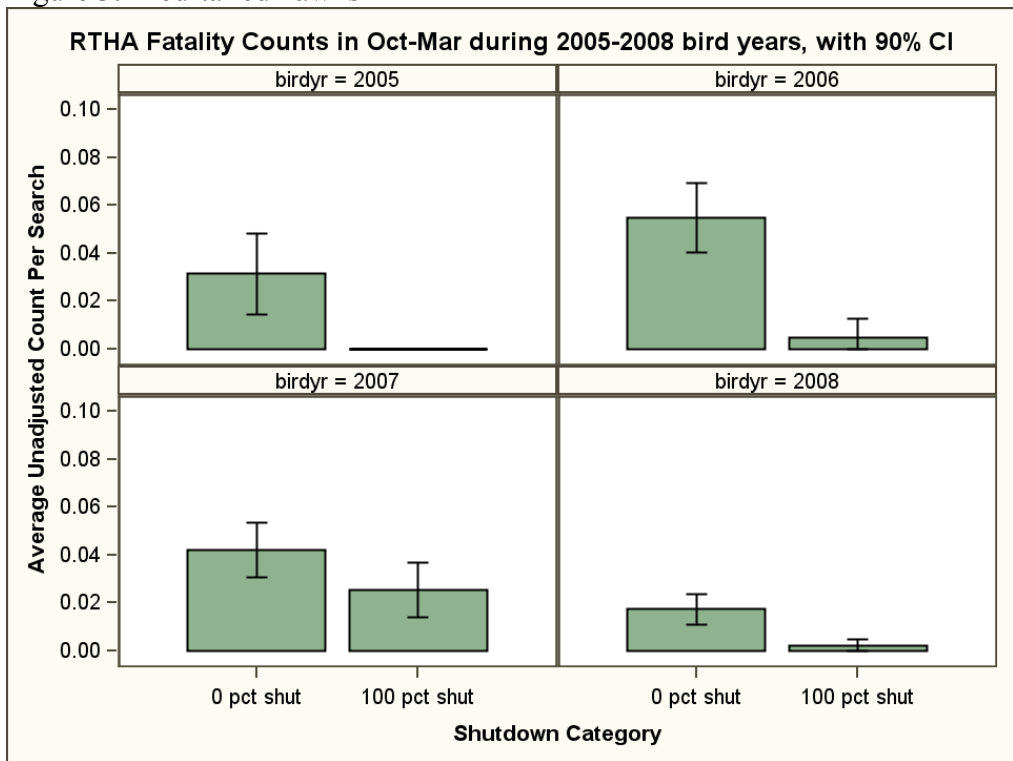
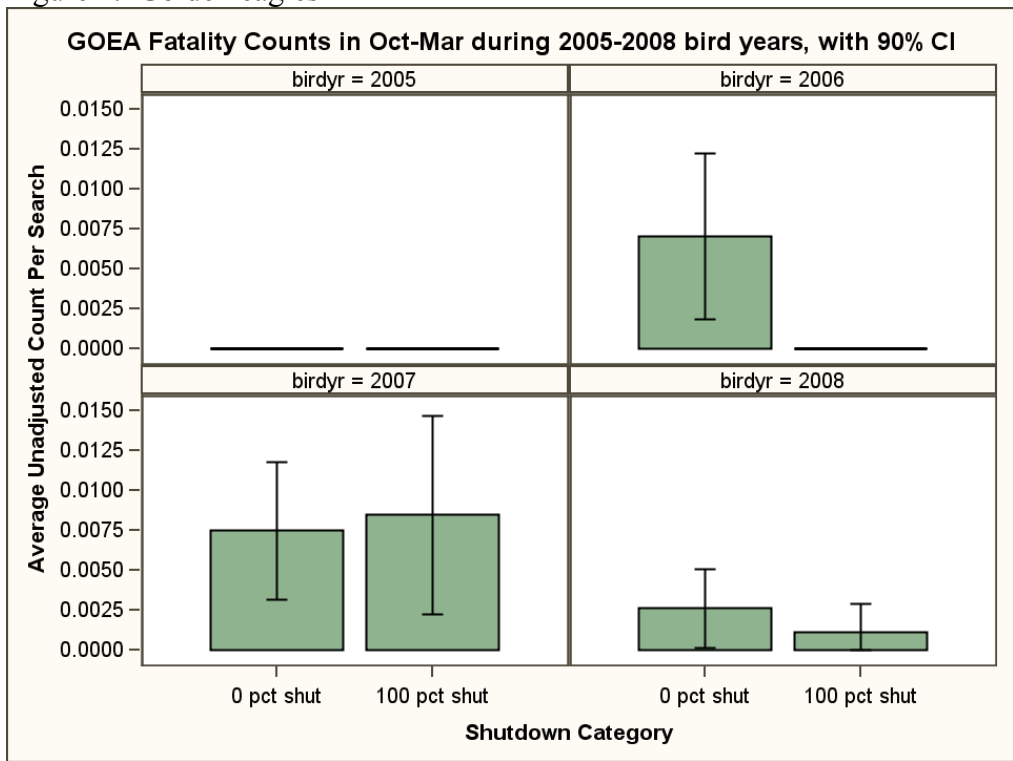


Figure 4. Golden eagles



Page 4-1, paragraph 1. There are no error estimates to help determine whether the “marked increase” from Baseline to Current Survey period is statistically significant. To examine this increase better, I converted my model for shutdown season into a model for comparing between the two periods. I expanded the dataset to include the Baseline as well as the Current Survey period. I used two versions of this data, one with all strings and the second with common strings¹. I continued to analyze focal species separately, although it was clear that only the BUOW and RTHA species had sufficient data to be analyzed this way. Both the BUOW and RTHA showed increases, however the standard errors were relatively large and none of the increases were statistically significant.

By the way, I analyzed the data for other trends including trends within the Current Survey period. I found a very interesting and statistically significant 50-60% decrease in BUOW and RTHA fatalities from the biennial period 2005-2006 to 2007-2008. This is worth discussing.

Page 4-1, paragraph 3. I agree that the selection of turbine strings in the Baseline study is not conducive to producing APWRA-wide estimates. However, I disagree that the same is true for the Current Survey. While the Current Survey has its limitations, it is not that far of a reach to obtain APWRA-wide estimates from it.

Page 4-2, paragraph 1. I don't understand where it reads in the top sentence that the baseline fatalities “detected during searches with a longer interval would be under-adjusted relative to an estimate using a more traditional carcass removal curve, thus exaggerating differences between the baseline and current studies.” What is the traditional carcass removal curve? Is it the Smallwood curve? And what is meant by under-adjusted? Does it mean that it is incorrectly too low of an adjustment? Or does it simply mean that it is not adjusted as high as the traditional curve?

¹ My set of *common strings* is probably a little different than the one analyzed in the M21 report. I could not find a *common strings* identifier in the database, so I constructed my own definition. I used strings which were surveyed at least 10 times during each the Baseline and Current survey periods, counting only the string searches that were <60 day intervals.