Author's response to reviews

Title: Accounting for Seasonal Patterns in Syndromic Surveillance Data for Outbreak Detection

Authors:

Tom Burr (tburr@lanl.gov)
Todd Graves (tgraves@lanl.gov)
Richard Klamann (rmk@lanl.gov)
Sarah Michalak (michalak@lanl.gov)
Richard Picard (picard@lanl.gov)
Nick Hengartner (nickh@lanl.gov)

Version: 3 Date: 27 October 2006

Author's response to reviews: see over
Dear Editor,

As requested, this cover letter is a point-by-point response to the 2 reviewers. I thank both reviewers for comments leading to improvements to the paper.

Regards,

Tom Burr
Reviewer 1:
Major Compulsory Revisions

1. The informal expression “one season fits all” is now defined in the Methods section of the Abstract, and in the 3rd paragraph of the revised Background section, page 4.

2. Performance claims based on real vs simulated data. The revision makes it clear that the 2 cited references [4] and [5] based performance claims strictly on simulated data that followed a nonhierarchical model having seasonal terms (first sentence of Study description, page 11) rather than on the real data. Our view is that these 2 references could be improved by simulating from an appropriate hierarchical model and/or by using the real data. We agree that performance can and should be based on real data when there is sufficient data (such as our 9 years of real data). One of the best approaches in SS is to inject synthetic outbreaks into real background data. However, even when there is sufficient real background data, there are at least two reasons to also inject synthetic outbreaks into simulated background data. First, any differences that arise (in performances on simulated versus real data) can lead to improvements to the simulation model. An effective simulation model is useful as a model-based data summary at the very least. This paper demonstrates that the hierarchical model gives better performance predictions than does the nonhierarchical model; therefore, at least in that context, it is a better model, and hence, a better summary of the data. Second, an improved simulation model can also lead to an improved forecasting method (see discretionary revision # 2)

3. Comments regarding moving average methods performances in Table 1.

We have stated in the text and we have shown (via the Table 1 results and Figures 1 and 2) that the real data violates the “one season fits all” assumption. So, there is violation of the “one season fits all assumption” both for the real data and for the data simulated from the hierarchical model. In contrast, the data generated from the nonhierarchical model obeys the “one season fits all” assumption. There are three main aspects of overall performance (DP, tendency for DP bias, and tendency for FAP bias); the reviewer comment focuses on DP so we address DP here, and also in the revised Conclusion in the main text. As an aside, this cover letter also addresses DP bias and FAP bias.

Regarding DPs, the moving average methods are particularly good (relatively) when applied to the real data, also quite good (relatively) when applied to the hierarchical data, and roughly the same as Method 1 (EWMA is 91, Method 1 is 92, but the other two moving average methods are 85 and 84) for nonhierarchical data. We concur that the relative performance of the moving average methods is stronger in the real data than in the hierarchical model-generated data, so have modified the Conclusion section accordingly to avoid possible confusion.

Regarding DP bias (if claims were based on data simulated from the nonhierarchical model as in 2 of the cited references):
The paired overall average differences for Methods 1-8 are 8, 0, -6, -3, 9, 12, 9, and 12, respectively, for nonhierarchical minus real data DPs and are 7, 3, 0, 3, 6, 5, 4, and 7, respectively for the nonhierarchical minus hierarchical data DPs. Note that Methods 2-4 have the smallest average differences in both cases, although the reviewer points out here that the effect is smaller when comparing nonhierarchical to hierarchical data DPs.

Regarding FAP bias, we propose no changes to the last paragraph in the Discussion section. Again, moving average methods did not exhibit false alarm probability bias, but the other methods did.
Regarding the short Conclusion in the Abstract, we made the following change:
… and (b) moving average methods based on relatively short amounts of training data are competitive in all three cases, but are particularly competitive on the real data and on data from the hierarchical model, which are the two data sets that violate the “one season fits all” assumption.

4. This is similar to # 3, and our revised Conclusion makes it clear that moving average methods showed the best overall performance for the real data and for hierarchical model data. Performance on nonhierarchical data was also surprisingly good for the EWMA method when compared to Method 1. Presumably, if there were a stronger seasonal and/or day of week effect, then Method 1 would do much better than EWMA on data generated from the hierarchical model.

5. “Hierarchical model” is now more clearly defined in the revised paragraphs 2 and 3 in the section on Hierarchical Modeling, and reference [12] is added.

Minor Essential Revisions
1. How the Hierarchical model is fit to the data. Response: revised/expanded paragraph beginning “YADAS was used..” in the section on Hierarchical Modeling.
2. mild serial correlation in the middle plot of figure 2. Response: The revised figure 2 caption mentions that there is very mild (but statistically negligible) correlation in this middle plot. On the basis of repeated fits with a smooth curve fit to the residuals, we confirm what we expect: there is zero serial correlation in the residuals if the same model used to simulate the data is used (after getting parameter estimates from the simulated data) to predict the data. Therefore, the particular realization depicted in the middle plot shows mild but negligible (statistically insignificant) serial correlation.
3. The goal of the study was not clearly indicated due to word choice in the Background. Response: Our revised wording in the Background (Abstract section and main text section) makes it clear that we do not propose a new forecasting method. Note that we must also continue to steer readers away from thinking that we are trying to rapidly detect each year’s seasonal peak, which is a different goal. In our context, the annual seasonal peak is actually a nuisance, so the revised text (paragraph one of Study description, but omitted from Background section in Abstract) continues to steer readers away from thinking that we are trying to rapidly detect each year’s seasonal peak, which is a different goal.
4. Figure 2 has 3 cases and 1 case per subplot: real data, data from the nonhierarchical model, and data from the hierarchical model
5. A description of each of the 8 forecasting methods is added to the Table 1 legend.

Discretionary Revisions
1. Why the term “hierarchical” is used to describe the new model: response: this is done in the revised paragraphs 2 and 3 in the section on Hierarchical Modeling.
2. Adapt the hierarchical model for forecasting. Response: We agree. A paragraph in the Hierarchical Modeling section still mentions the computational cost. The revised paragraph 4 of the Conclusions describes why it might be promising to do so.

Reviewer 2:
Major Compulsory Revisions
None.
Minor Essential Revisions

1. Edits to Background section of main text: done, including the last sentence, plus other revisions as noted by both reviewers.
2. Definition of syndromic surveillance was too narrow. We agree and corrected it in the first paragraph of the Background section.
3. Methods section: added 3 sentences to describe BSafer dataset.
4. Make more clear that simulations were based on daily respiratory counts Consider presenting daily data. Response: Paragraph 4 in the Study description makes it clear that we work with daily counts. We believe that plots of daily data are too “busy” and that the seasonal effect is too hard to see using daily data. Concerning use of weekly fits to determine whether the daily fits are adequate: the revision makes it clear that if the daily fits are adequate, then the weekly fits based on summing daily fits will be adequate. Because this is not the case (on the basis of Figures 1 and 2), we can conclude that the daily fits are also not adequate (the logic is: if A implies B then not B implies not A). We concur that the revision needs to be clear about this logic – thanks. See the first sentence in revised paragraph 3 in Nonhierarchical modeling section under Methods.
5. Moved “study description” to Methods.

Discretionary Revisions

1. page 14, last sentence—we agree and dropped the reference to BioSense here.
2. page 5, Figure 1 … we agree and made the change.
3. Figures 1 and 2 legend – we agree and made the changes to the legend. We believe that using month/year on the horizontal axis is too crowded and does not contribute needed information, so we kept the horizontal axis labels as they were.
4. citations—we think 1 and 2 are acceptable. Editor’s decision.
5. Author contributions. We gave more detail.