

1 Associate Editor

Overall, I enjoyed reading this paper. Methodological/modelling advances are proposed for managing multiple health outcomes in a spatial risk mapping context. The case study examined, looking at health related risk caused by heat stress is substantial, interesting, and well incorporated into the paper.

However, along with the two reviewers, I feel that some improvement is required regarding the presentation of methodology. As it stands, it would be hard for other researchers to reproduce the research the authors have carried out, due to some gaps in methodological description.

I note that, along with reviewer two, I also had some issues reading the keys on figures 2 and 3 some of them appear to be purely black, and if they are displaying a range of colours, do not seem to tally with the colours in the main maps.

2 Referee: 1

(a) MCMC questions

I really feel given the nature of this paper (it is not the definitive guide to heat and health in Houston) that it needs to be fully capable of reproducibility. I therefore don't feel the part of the algorithm set out in section 2.4 is sufficient. For example I had to read to find details for α_p etc., some of the other comments (the posterior is fully known) were a little glib. I think the entire sampler could be set out in the text, but if not an appendix could be used. Links to code would also be helpful (I appreciate there are confidentiality issues with some of the data).

I don't think the comments on convergence are adequate. On page 13 line 10; how does MCSE alone provide evidence of convergence? I think more is needed.

I wasn't sure about claiming IG(2,1) or the Dirichlet prior were non-informative. Using population counts sounds very informative (or was it the expected values that were used).

It would be useful to understand some of the covariate posterior information; does the 0.51 versus 0.49 for males versus females section 3.3) really mean anything? This might be a great parameter to set out the trace plots and density plots which would help provide information on convergence as well as understanding the relevance of what sounds like a modest difference to me.

(b) Why do a spatial misalignment (Brunsdon and Comber)? There are comparable numbers of census tracts and grid squares; using census tracts directly sounds like a tractable GIS problem (the denominator is the area of the census tract, points-in-polygons can count the necessary points). I couldn't see anything in the mathematics of the intensity that would prevent the census tracts being used, and you remove a layer of approximation.

(c) Heat enters as an upper level covariate, with ecological fallacy risks? Are the census data not available to study how some of these features vary with key other demographics? I appreciate all the census data are disclosure limited but some multiway information on air conditioning and age band could really limit the potential problems with confounders in the aggregate analysis.

(d) Although you make a comment in the end about the lack of suitable model fit diagnostics there is a lot more that is already available. WAIC / Gneiting and Tillmans's proper scoring rules seem appropriate given a major outcome is a risk map.

(e) From an applications point of view I'd really like to see the counts (posterior predictive density) as well as the probabilities in figure 4.

3 Referee: 2

This is a very interesting paper but the text needs to be more clear and the methods used need to be more transparent in order for the readers to follow.

The authors need to explain explicitly how to combine/connect and use data information from 2 different periods of time (heat-related emergency calls during 2006-2010 vs. mortality data between 1999 and 2006).

The prior assumptions and choices need to be justified especially for those non-informative and vague priors. Also, "Because the priors for α_p, γ_p " etc., "are all conjugate, their posterior distributions are known..."

is not clear to me that the claim is valid. It would also be helpful if the results from the MCMC are presented so to confirm convergence.

The labels of colours in figures 2 and 3 are hard to see and the 95%CI in figure 3 is not clear.