

Firstly, we would like to thank both of the anonymous referees for their reviews and the encouraging comments. The odd constructs or grammatical errors that they note have been fixed.

We take seriously the problem of interpolation/extrapolation raised in one of the reviews. One of the more difficult challenges in this analysis was producing a gridded observational dataset to be compared with models, using the scattered point observations. After rigorously testing and evaluating multiple interpolation techniques, we feel that we have developed a method that provides a sound representation of the domains investigated. By choosing to define the quality of prediction based on the effective number of point observations that went into the interpolation, we created an objective mask, which even included some gaps over Montana in the US gridding. This quality of prediction is a measure of the interpolation error (as shown with several tests in the paper), but it still does not define a formal error. To that end, we remain stuck. We were pleasantly surprised/encouraged that the higher quality grid cells were more accurately hindcast by the model.

We are pleased that the referees concur on the advantage of defining extreme events based on the return time rather than exceedance of a fixed threshold. This definition does, however, present challenges regarding the non-stationarity of ozone in response to local emissions and the extension of the technique to future climate simulations. With that in mind and based on discussions of our paper with colleagues (which unfortunately were off-line in spite of our efforts to encourage the ACPD venue for non-review discussions), we have revised the layout of the final discussion section to clearly address these challenges.