Overview:

Understanding crop yield response to environmental changes is crucial for food security. Statistical crop models are easier for calculation but the projection capability is constrained by the range of current conditions. The process based crop models aim to capture the yield response to different environmental changes but computational expensive compared to statistical crop models. This study developed statistical emulators for 9 process based crop models using GGCMI phase II simulations. The author well validated the statistical emulators and discussed the caveats and the potential usage, such as provide an alternate approach for impact assessment. The manuscript is generally well written and I only have several minor comments on the method and results.

Minor comments:

The whole section 2.2 discussed why there are differences in climatological and year-to-year response. This part is very interesting but somehow could divert the readers who are eager to know how the study uses the training data to develop emulators. It could be a better flow if put the section into the discussion section or supplementary.

The authors need to refine the section 3.1 to give more information on Y and regressors (what temporal and spatial scale). Line 161 mentioned that "Emulating at the grid cell level". So I think equation 1 was fitting at grid cell level. My understanding is that Y is a vector of 30-year averaged crop yields across different uniform changes scenarios (a total of 756 scenarios?) in one GGCMI model. There are 34 terms in equation 1, aren't there over fitting problems when you have a small number of Y (some models did not done all required scenarios) but a large number of regressors. Please comment.

Line 138: "in the the". Double the here.

Equation 2. Some terms are gray in equation 2. Are those the dropped terms? If so, just delete them.

Figure 10 caption. "the five GCCMI Phase II crops", the authors used this terms several times, but this sounds like there are five special crops that was created by GGCMI Phase II. I think just say five crops is fine. They are common crops. And how many individual models are incorporated here? I guess it is nine. But there are not nine color lines, is that because some lines are underneath the black thick line? If so, please mention that.

Figure 11. In the figure legend, the uniform T sounds like each process model was forced with global uniform T. But I think it means the uniform increase of T, uniform ΔT is better.

Figure 11. In the caption, "Circles are emulated yearly global production changes", those are dots, not circles.

Figure 11. Why there are no open squares on plot b? And in plot c, open squares for 2 and 4 increasing of T is missing. All the three plots showed emulated uniform T lines, why not show emulated uniform T+W for plot b, and emulated uniform T+W+C for plot c?

In the SI:

Page 2. First line "is not uniform tn the GGCMI Phase II", what is tn? Should be in?

Figure S6: there are no gray lines (Ontario), why? I want to know if Ontario has the same failure in A1 as in A0.

Figure S21: The simulated RCP8.5 (open triangle) were not found on the graph.