Interactive comment on "Impacts of climate change and climate extremes on major crops productivity in China at a global warming of 1.5 °C & 2.0 °C" *by* Yi Chen et al.

Anonymous Referee #2

This paper presents a simulation study of the effect of moderate global warming on productivity of three major crops in China. Furthermore, the study quantifies the separate contributions of temperature and CO2 changes to the overall effect of climate change.

As with all simulation studies, the major problem is that one is using simulation rather than data to draw conclusions about the behavior of the system under the assumed climate. It is necessary then to present solid information to support the assumption that the trends shown by the simulations mimic the trends in the real system. In addition, since it is clear that the simulations are only an approximation, it is important to provide information about the uncertainty in the simulated results. Finally, uncertainty itself

has a relation to the real world, e.g. 80% confidence intervals should cover observed values about 80% of the time, and so it is important to justify the method of estimating the uncertainty.

Whereas several recent studies have used the median of multi-model ensembles to predict impacts, this study uses a single basic model. The authors refer several times to their model as the "well-validated" MCWLA model. The associated adjective is more an incantation than information. The authors need to give quantitative information on how well this model has been found to reproduce changes in yield due to changes in temperature, heat stress and CO2 concentration in the three crops studied here. Of particular importance is how accurately the model reproduces relative changes due to changes in temperature, heat stress and CO2, since the results of the paper focus on relative changes.

Model calibration is a major determinant of predictive accuracy. The authors say that the model is "well calibrated". What exactly does this mean? How was calibration done? Was it done for each cultivar in China or for each region and therefore some "average" cultivar per region? The authors run their model with 10 "optimal sets of parameters", in order to "reduce the uncertainties caused by crop model parameters". What is the meaning of "optimal sets of parameters". How were these derived? How will this reduce uncertainty? Normally one would predict using just one set of parameters, that results from the calibration procedure. Is this what the authors are doing, or are they using the mean or median of their ten sets of parameters?

The authors do deal with uncertainty, in both the climate projections and the impact calculations. The uncertainty in climate projections is quantified as the variability between different GCMs. The authors use four different GCMs; three provided 20 variants each and the fourth provided 10 variants. The origin of the variants (different initial conditions? different parameters?) should be noted. Also, it would seem that the first three GCMs are weighted twice as much as the fourth GCM. Why this differential weighting?

It seems that the uncertainty in the impact calculations is taken from the ten sets of model parameters. This is from a different part of the paper than the section that says that the ten sets of parameters were used to "reduce the uncertainties". It seems more logical that the ten sets were used to estimate uncertainty rather than to reduce it, but this should be clarified. In any case, it is important to clarify the origin of the ten sets of parameters, and the experimental evidence showing that this gives a realistic estimate of model prediction uncertainty.

I believe that the issues of confidence in the model results and in the uncertainty estimates, supported by comparison with data, should be addressed before this paper is published.