

Title: Sources of interannual yield variability in JULES-crop and implications for forcing with seasonal weather forecasts

Authors: K. E. Williams and P. D. Falloon

Journal: Geoscientific Model Development, 2015

General Comments:

1) Unfortunately, the main objective(s) of the paper is (are) not clearly stated in the introduction section. From this part, it seems that there are two main questions followed by this paper, both centered on “data availability” issue:

- (Page 4603, Line 9) “to investigate whether the yield variability can be modelled sufficiently well if only a subset of the forcing variables are taken from the seasonal forecast and the others set to climatology, or if the model is forced with daily meteorological data and disaggregated internally to the model time step”
- (Page 4603, Second paragraph) Due to significant practical challenges to estimate the initial conditions variables, such as soil moisture, on the start date of seasonal forecast runs, “we investigate the loss in predictability of yield if the run is started on the sowing date of the crop in that gridbox and initialised by the climatological values for that date”

I suggest authors provide an explicit statement including the objective of their study and provide a more critical literature review so the readers can clearly understand the gaps/needs in the literature, and the contribution of this paper. While doing this, I also suggest avoiding using confusing sentences while talking about the objectives and challenges addressed by your work (see my detailed comments).

Also, the introduction section is heavily built on using seasonal forecasts as the input to crop models, while, as far as I have understood, this is not the main target of this study, and is in fact an implication of this study.

2) Experimental Set-up:

- Given that the main difference between model results from set up 3.1 and 3.3 is due to JULES disaggregator, it might be interesting to show to what extent JULES disaggregator can reconstruct the 6-hour climate data, particularly for precipitation and temperature.
- Please elaborate more on set-up 3.4 and Table 1; specifically, what exactly do you mean by “we created daily climatologies of each driving data variable in the full disaggregated run”? How did you create “climatologies”? Is it the long-term average of climate variables for each day of the year? Then, did you run your model only for one year? If so, then how are you able to comment on inter-annual variability (this is also the case with set up 3.5, line 23: “Each run lasted 1 year”)?
- Please note that, you should assume that readers are not familiar with the model, scenarios you have considered, assumptions you have made, methods you have employed, etc. Therefore, I highly recommend to consider revising the writing of important parts of the paper (for example section 3) to ensure readers to a high extent understand important details of your work. I really cannot comment that much on your results and analyses based on the last two scenarios since I cannot understand the underlying assumptions.

3) I have not read the Osborne et al. (2015) or any other earlier JULES and JULES-crop papers. However, I am really concerned about model “calibration” and “validation”, and I would appreciate if

you can comment on my concerns. While working with large scale models (of any kind), especially if the time interval spans over several decades, relying on simple correlation factors to assess the performance of the model is very misleading. It is particularly important to investigate model performance (especially if it is calibrated using different year types, i.e., wet, dry, normal) during dry seasons, because outliers usually are not paid enough attention when dealing with “average” values over a long time span.

- 4) Do you think it is enough to discuss the performance of the model (and consequently talk about the impacts of different parameters on interannual variability) which simulates 50 years of crop yield globally just by using a correlation coefficient? This is indirectly related to the previous comment, but in that comment, I was just concerned about model validation. Here, I really want to know how each different scenario is compared with the baseline. Is there any difference between different set-ups in dry, normal, wet years?
- 5) Page 4609, Last paragraph (Also Figure 1): I can see that in many parts of the world where rainfed agriculture is dominant (e.g., sub-Saharan Africa), correlation coefficient between season precipitation and crop yield is very low. How can you justify this result?
- 6) Figure 1 (results related to *disagg*): In this set-up there is no irrigation. How is it possible that the model can simulate maize yield very well even in regions where irrigation is heavily practiced (India, Central US)?
- 7) Page 4612, Line 19: “since obtaining accurate initialisation data on the timescales needed for seasonal forecast runs is a particularly significant practical challenge”. You also mentioned this in the introduction section as a kind of motivation to study the impact of initialization on yield variability. As I also mentioned in another comment below, this is confusing to me. If seasonal forecasts are to be provided say one month before the crop growth season, you can still run your model, say couple of months in advance using observational data to initialize the soil moisture, and then seasonal forecasts data will be fed into the model. Now, if seasonal forecasts are to be provided on the sowing date, you still need to run the model in advance using observation data to initialize the soil moisture. I really cannot understand this argument. I hope in the next version you can clearly explain your point.
- 8) Page 4612, last paragraph: Although seasonal forecasts in high temporal resolution (sub-daily) are not available everywhere or provided by all models, but there are many models that provide seasonal forecasts of all climate/weather variables in sub-daily time scale. Moreover, one of your arguments in the beginning that using high resolution data requires high storage, which is correct, does not seem to make much difference in the near future given the emergence of cloud and high performance computing. Given these conditions, where else do you see the contribution and importance of your work?

Detailed Comments:

1. Page 4600, Line 24: “However, existing studies of crop model performance focused on seasonal forecast applications show considerable variation depending on the region, scale, processes and crops involved”. Isn’t this something that you would expect? I mean regardless of seasonal forecasts, crop simulation does depend on region, scale, processes, crops, etc.
2. Could you please elaborate more on this part (Page 4601, Line 11): “Using a statistical approach to assess the reliability of global-scale seasonal crop failure hindcasts, Iizumi et al. (2013) found that within-season hindcasts generally reproduced inter-annual variability in observed yields in major wheat exporting countries ($r^2 = 0.56\text{--}0.61$) better than pre-season hindcasts ($r^2 = 0.43\text{--}0.59$).”

- What exactly do you mean by seasonal crop failure hindcast?
 - Also what exactly do you mean by with-in season hindcast and pre-season hindcast?
3. What is the baseline when you phrase (page 4601, Line 23) “Palmer et al. (2004) and Cantelaube and Terres (2005) also found that downscaling seasonal hindcasts improved crop model performance”? Do you mean comparing to not-downscaled seasonal hindcasts?
 4. Page 4602, Line 14: Can you provide one or more examples of initial conditions here?
 5. Page 4602, Line 25: I suggest you briefly define “crop failure” at some point in the introduction so that readers better follow this section.
 6. There are several parts in the one to the last paragraph of the introduction section that, in my view, requires serious clarification, specifically because, as stated earlier, this is one of the important questions addressed by the paper:
 - Is there any difference between start date of the seasonal forecast runs and JULES-crop runs? And how are they different from sowing date (is sowing date assumed to be equivalent to the beginning of the crop growth season)?
 - What does “run” exactly refer to in “if the run is started on the sowing date”?
 - What does it mean by “Starting the run before or on the sowing date means that the initialisation of crop variables (e.g. height) is trivial”?
 7. Page 4604, Line 5: I suggest not using “Sect.” as a short form of “section”.
 8. Page 4606, Line 9: You never identified the “main run” in “The run was from 1960 to 2009 and spun up by repeating the first 10 years five times, before starting the 10 main run”. So does it the main run is from 1970 to 2009? Also, is it necessary to run the first 10 years five times? If so, please elaborate.
 9. Section 3.2: Is there any implicit assumption that “there is no constraint on water availability”?
 10. Page 4608, Paragraph 2: What are the correlations between different scenarios and the baseline (I think you have used FAO reported yield values)?
 11. Page 4608, Paragraph 2: Related to your discussion about Brazil, so are you saying that using daily data followed by model disaggregation in the model outperforms using 6-hour data in the model in Brazil?
 12. Page 4609, Paragraph 1: In the second set-up, based on your argument here, can we say that removing soil moisture stress from the model is equivalent to no water availability (either green or blue water) constraint? Then, I am wondering what the relevance and importance of this set-up would be. Does it provide the readers with any interesting insights? To what extent did you use the insights in your discussion/conclusion?