General Comments:

This study attempts to establish crop sowing and harvest dates in India that are tailored to the seasonal monsoon – a regional climate system that dictates crop cycles for the majority of India's farmers. The authors motivate this study by 1) describing the several inconsistencies and limitations of larger-scale estimates of sowing/harvest dates and crop calendars, 2) summarizing the need to incorporate higher resolution products as we move towards more advanced, regional climate-crop assessments, and finally 3) detailing the sensitivities of Indian cropping systems – specifically rice-wheat systems – which requires rapid plot preparation and turnover between crops for the successful production of both crops. The authors also put forth an important point that it is unrealistic to hold crop planting/harvest constant in crop model simulations of future climate impacts, as we expect farmers to adapt as weather and climate patterns change.

In general, this study is of good quality and is highly warranted (as the authors point out) to facilitate improved climate-crop model experimentation and to incorporate autonomous adaptation. The motivation and methods the authors provide are entirely reasonable. Methods like the ones presented here could be of great utility to the wider crop modeling community in that a) it could prove a more realistic representation of dryland farming systems and regions that are highly dependent on rainfall and b) the methods appear straightforward enough and flexible to facilitate wide implementation across models and similar monsoonal agricultural systems.

However, I do think certain clarifications in the methods are warranted to provide the reader with additional understanding as to why some metrics were selected and, importantly, to contextualize the results (e.g. In Table 1, what constitutes a reasonable amount of time between planting and harvest, as derived using model simulations and in light of the need to better represent these rice-wheat systems?). I think the authors might briefly include measures or notes of what constitutes a "reasonable" representation in the Methods section and apply these in the Results. There might be an added caveat as well on how these Methods really demand a "good" representation of the monsoon system (and what that means – with respect to Aphrodite or another product), and that therefore there are limitations in applying this to the wider range of climate models, which have historically had issues in simulating the monsoon components (including onset and intraseasonal variability).

As such, I recommend that this manuscript undergo revisions. I think if the authors can sufficiently address the following concerns, the manuscript would be of benefit and utility to the larger community.

Specific Comments:

Line 111: I appreciate the discussion of the limitations of the Sacks et al dataset to your applications, but is the Sacks dataset actually used in your Methods at all? It looks to me that the Methods and Results are relative only to the Bodh (2015) dataset. If this is the case, then I think it can be cut from the figures and the section trimmed a bit (after it

briefly stating the Sacks et al issues and that it's not appropriate for regional assessment needs).

Section 2.2: Developing metrics for defining site or sub-region specific onset/withdrawal periods can be a rather arbitrary procedure in defining "critical" thresholds. I think it's fine that the authors selected the NPPI – in looking at Figure 4 and roughly eye-balling IMD contours (http://www.imd.gov.in/pages/monsoon_main.php?adta=JPG&adtb=1), there appears to be consistency. Figures 5 and 6 also show consistency between onset and sowing dates, which is good.

That said, I think a bit more discussion here on why this particular index was selected could be useful, particularly to people who work in the region and want more monsoon-tailored agro-climatic assessments. The original reference Zeng and Lu (2004), upon which the Lucas-Picher (2011) index is based, did indeed create the NPPI for a consistent way to evaluate important global monsoon statistics across datasets and models (despite the 0.618 threshold, which apparently did not have a physical basis).

However, given that this study is framed with respect to the importance of tailoring crop management information to the *physical monsoon system*, how does the NPPI provide that advantage to you over, say, another metric and/or threshold (Zeng et al. 2004; Li and Zhang 2009; Saini et al. 2011; Dong et al. 2015; Guimberteau et al. 2012; Zhang; Wang and Lin 2002; Goswami et al. 2006)? Were other indices considered and discarded? If so, why? The authors mention a few others in the beginning of the section but do not motivate the NPPI over the others, aside from saying what it is.

[Please note, the references provided on defining the onset are not comprehensive, and there is much more work here. Admittedly, some of the variables are data-limited or cannot be applied on the scales you seek. This is just to note there is a lot out there, so some added motivation would be good.]

One way to better couch this section might be to simply move Section 3.1 to the Methods, rather than in the Results. Then you can say, for example, that you evaluate the NPPI across different models, and found the metric to be consistent with observed sub-regional monsoon onsets and with observed planting periods. You could then start the Results roughly from Section 3.3. Just a thought.

Lastly, Lucas-Picher (2011) note on their page 859 that the NPPI was used specifically as a standard with which to compare models, and that these periods may not specifically represent the beginning or end of the actual monsoon. Additionally and incidentally, they also note large range in NPPI across models, which led them to conclude that key physical monsoon processes may be missing from many model environments. How do the authors interpret these issues for their work here? Perhaps a statement should be added about the careful need to use models that are most representative of monsoon processes (not just that RCMs are better), and what the best measures are to evaluate this.

Section 2.3.1 Line 204: I may just be unable to find this information, but how was the "area averaging" applied to the $RelMonsoon_{croprules}$? While the gridbox heterogeneity may be incorporated into the initial calculations of $RelMonsoon_{croprule}$ (as the authors mention in line 217), I wonder how the area-averaging actually impacts the authors' goal of capturing some of the spatial variation, which to me seemed a big goal of the study.

Section 2.3.1 Line 222-225: I would be careful to caveat this statement in relating it to the ASM, as the authors found that the methods needed to be adjusted for the northeast monsoon period in southern India (which will also affect Sri Lanka as well).

Figure 3: It would be helpful in Figure 3 to put "NPPI calculation" into box 2.3 and "$RelMonsoon_{croprule}$"/*MonsoonDerivedCropDate* into box 2.4, to make those linkages explicit.

Section 2.4 Line 229-231: This SWLs versus time period is a bit confusing given that regional temperature can exceed the SWLs, as noted. Perhaps it would be less confusing if the authors just picked one type of scenario, based either on time period *or* SWL.

Section 3.1 Line 261: How are you judging "compare well with observations"? Might a skill score or some other quantitative metric be included here? In fact, there are a couple of areas in the manuscript (e.g. Line 269 comparing the monsoon onset with the sowing date) where it is stated that the results look reasonable and compare well, but there is not much context given to what is meant by this.

I recommend moving Figure B1 to Figure 4, as it's a bit difficult to distinguish between the colors and having the differences with Aphrodite on hand would be more useful and quicker to interpret. Leave Aphrodite in the absolute colorbar however!

Figure 5: This is showing a difference in days, correct? So sowing in the lightest blue areas in the northwest of India are ~15 days ahead of the onset? If this is the case, then the colorbar title should be changed to "difference in days" or similar, not "Day of Year".

Section 3.2 Line 277: Table 1 seems to suggest that rice is sown 19 days before the onset as calculated using the Aphrodite dataset, if I'm reading that correctly (or at least there's an ~19 day gap between onset and sowing). How does this align with the comments in the introduction (Line 84), which said that "rice is usually sown with the first rains"?

I ask because this is also relevant to judging how "well" the onset and sowing periods overlap, and can be used to assess how well the method works. If the cropping systems are so tightly scheduled as they are described in the Introduction (Line 90-97), then this difference between sowing and onset could be impactful, no? What constitutes too much of a gap between sowing and onset?

And along these lines, I'm a bit curious about the ERAint and ECHAM5 results (Looking at Table 1 and Figure 6, for example). With respect to rice sowing, the India average

results are much closer to the onset it seems (3.4 and 10.1, respectively). How do we qualitatively compare these results across models and observations?

In Figure 6 anyway, the coverage of blue makes ECHAM5 look - dare I say it – better than Aphrodite?! This is of course assuming that the Bodh et al dataset tells us something physical about the monsoon onset, in that farmers plant very close to or upon the onset (and thus sowing is a good proxy for the onset). A little guidance on how to interpret this would be appreciated, particularly with respect to the Aphrodite results (which I presume would be the most realistic?).

As an aside, it can be a bit confusing over the course of reading the paper to distinguish "before" or "after" relative to the monsoon statistics and sowing/harvesting. It's just a bit hard to keep the differencing straight. If there's a way to clearly and consistently indicate sowing dates that are before or after the onset, for example by scaling or color coding them explicitly, that would be very helpful.

Figure 6: This figure has potential to be very helpful, particularly related to my comments above, but is confusing in its current form and description. I think "hit", "overlap" and "miss" need to be more clearly defined, as there's an "overlap" color and the bar itself is called "overlap". The text, Lines 270-275, do not provide any further clarity in their current form, as both the word "around" and "at least close to" are used for the blue and the yellow. These are too vague to be helpful in interpreting the maps.

Technical Comments:

While I did not catch many technical faults with the manuscript, I would say in general, I the Intro and Motivation could be condensed a bit. It reads a bit long now and some of the material can be more succinctly stated.

Also "Section 2.3.1" may not be needed as you move directly to Section 2.4

The Discussion and Conclusions section can be condensed I think, as the beginning of the Discussion section (Lines 335-343) already read like the beginning to a conclusion.

I would lastly request that the authors check the titles of all their colorbars, as some (like Figure 5) I don't think quite convey the intent of the figure.

References:

Dong, G., H. Zhang, A. Moise, L. Hanson, P. Liang, and H. Ye, 2015: CMIP5 model-simulated onset, duration and intensity of the Asian summer monsoon in current and future climate. *Clim. Dyn.*, 355–382, doi:10.1007/s00382-015-2588-z.

Goswami, B. N., G. Wu, and T. Yasunari, The Annual Cycle, Intraseasonal Oscillations, and Roadblock to Seasonal Predictability of the Asian Summer Monsoon. http://journals.ametsoc.org/doi/pdf/10.1175/JCLI3901.1 (Accessed December 4, 2017).

Guimberteau, M., K. Laval, A. Perrier, and J. Polcher, 2012: Global effect of irrigation and its impact on the onset of the Indian summer monsoon. *Clim. Dyn.*, **39**, 1329–1348, doi:10.1007/s00382-011-1252-5.

Li, J., and L. Zhang, 2009: Wind onset and withdrawal of Asian summer monsoon and their simulated performance in AMIP models. *Clim. Dyn.*, **32**, 935–968, doi:10.1007/s00382-008-0465-8.

Saini, R., M. Barlow, and A. Hoell, 2011: Dynamics and thermodynamics of the regional response to the Indian monsoon onset. *J. Clim.*, **24**, 5879–5886, doi:10.1175/2011JCLI3928.1.

Wang, B., and H. Lin, 2002: Rainy Season of the Asian–Pacific Summer Monsoon. *J. Clim.*, **15**, 386–398. http://journals.ametsoc.org/doi/pdf/10.1175/1520-0442%282002%29015%3C0386%3ARSOTAP%3E2.0.CO%3B2 (Accessed December 4, 2017).

Zeng, X., E. Lu, X. Zeng, and E. Lu, 2004: Globally Unified Monsoon Onset and Retreat Indexes. *J. Clim.*, **17**, 2241–2248, doi:10.1175/1520-0442(2004)017<2241:GUMOAR>2.0.CO;2. http://journals.ametsoc.org/doi/abs/10.1175/1520-0442%282004%29017%3C2241%3AGUMOAR%3E2.0.CO%3B2 (Accessed December 4, 2017).

Zhang, H., Diagnosing Australia-Asian monsoon onset/retreat using large-scale wind and moisture indices. doi:10.1007/s00382-009-0620-x. https://link.springer.com/content/pdf/10.1007%2Fs00382-009-0620-x.pdf (Accessed December 4, 2017).