

Second Review “Developing an eco-physiological process-based model of soybean growth and yield (MATCRO-Soy v.1): Model calibration and evaluation”

General remarks

I thank the authors for their responses to my comments and the amendments to the manuscript. The quality of the manuscript has improved substantially and I have only a few minor comments and only one issue that requires clarification. I provided the specifics in the comments below.

Major comments

1. I thank the authors for their response and changes regarding my question about the segmented linear model. I do, however, have two follow-up questions. First, I ask the authors to clarify what they determined manually: only the breakpoints, or also the lines within each segment? If only the breakpoints were chosen manually, how were the segments fitted – e.g. by ordinary least squares (OLS)? If so, this should be stated explicitly. Second, if instead both breakpoints and lines were determined manually, I am uncertain whether the term *segmented linear model* is appropriate in the statistical sense, and whether it would be clearer to describe these instead as *segmented linear functions*.

Minor comments

1. L167f: Replace “[...] simulated a function [...]” with “[...] is a function [...]” and “[...] which observed [...]” by “[...] who observed [...]”.
2. L178: Replace “[...] no and high (h) [...]” with “[...] no (l) and high (h) [...]”.
3. L190: “[...] value of growth rate of dry weight [...]” sound strange. Do you mean the cumulative growth rate expressed as dry weight?
4. L216: Replace “[...] stop growing, [...]” with “[...] leaves stop growing, [...]”.
5. Eq. 22: Why are the glucose reserves considered when calculating the LAI?
6. L245 and Eq. 25: Is SR a parameter or a variable? The text says parameter but it also says at harvest time suggesting that SR is generally dependent on time. Please clarify. Also why do you only state that Yield is a function of four variables but do not provide the actual equation or a reference to it?
7. L247-250: This sentence is quite long and complicated. Consider splitting it into two.
8. L299: Delete “based on the”.
9. L301: “broad planting densities” sounds a bit odd why not use “low” since you use high later in L302.
10. Sec. 5.1: Are the areas provided by MIRCA comparable to the areas reported to FAO? Even though the authors do not use FAO harvested area data I assume that FAO yield and harvested area data are not independent. Similarly, the GDHY is a product derived from census and remote sensing data containing assumptions and uncertainties. It could be worth to mention this in the discussion as an additional reason for the discrepancies between MATCRO-Soy and observations unrelated to MATCRO-Soy but uncertainty of the available data.
11. Fig. 8: The resolution is quite low making it hard to read the labels.
12. L640-646: Here the authors introduce new results which may be better suited for the results and not the discussion section.