A New Concept for Simulation of Vegetated Land Surface Dynamics: The Event Driven Phenology Model Part I

By V. Kovalskyy and G. M. Henebry

Response to comments from anonymous Referee #3

The paper addresses relevant scientific questions within the scope of BG. While the paper seems to present a novel modelling concept, it is not clear how different this actually is from other approaches (see below). I find the conclusions not particularly substantial as the model seems to have a lot of parameters; however, since the authors are not specific about this, it is difficult to say. The model description should be much more detailed. Results are mostly sufficient to support the interpretations and conclusions. Descriptions do not allow reproduction. The language is fluent, but in parts not precise enough - again this is related to the insufficient model description. Some definitions are missing or symbols are not clear. The supplementary material needs to be improved to allow reproduction.

Partially done. Despite the lack of specific points to be clarified in the model description, we included additional formulas and explanations in the main MS and in the appendices. We believe that provided information is sufficient for non-verbatim reproduction. We agree that an exact reproduction is hardly possible here but the details of it are unnecessary for this paper and here is why. In addition to formulas, we used mostly object oriented programing techniques and data communication schemes in the functioning of the EDPM. These elements are absolutely crucial for the implementation of the event driven framework, but those do not present any interest to the audience of BG. Neither do these details contribute to the substance of the paper. We are currently working on the Developers Guide for the EDPM that will contain all these details along with the C++ source code of the model and training procedures for complete reproduction.

The manuscript presents an interesting concept for an event driven phenology model. This is clearly important as some phenologically relevant forcings can be considered as events, e.g. frost. I argue, however, that it is not clear that anything is gained by choosing this form as opposed to a more truly continuous one. As the authors say: "The transformation of continuous factors into events relies on partitioning that depends on canopy responses." (p 5285, l 21) Hence a set of thresholds needs to be determined for this transformation. In a continuous formulation one can use integrals and step functions, which also allow to indicate events. Again some thresholds are required. I am therefore not convinced that there is really a difference between an "event driven" and "continuous" approach.

The way events are used in this EDPM implementation is indeed similar to the use of multiple linearized step functions. Yet for certain event types, the functions can be triggered not only in the time of the event but also in several subsequent steps. Temporal relevance of events and possibility to have multiple response functions that vary in time; this constitutes the main difference with a "continuous" approach in which the response functions are triggered only in the time-step at which the threshold is passed. Also, the framework can support event types that do not require proxy variables and thresholds, but we did not have access to planting, harvesting or other agricultural management records to demonstrate that in this paper.

I also find the following statement not really appropriate:
"This modeling approach opens the door to representing an ecological understanding of interactions of multiple drivers/events that drive the temporal variability of canopy characteristics..." (p 5285, l 25) While this may be true, it is also true of other truly continuous approaches, and not specific to this model. In addition, however, nothing in the work presented here shows more ecological understanding than other models nor is it shown how this would work. In fact the "learning" aims at reproducing the TNDVI which is a very integrated measure of canopy growth and not of detailed specific ecological processes. I would expect that there multiple instances of the model presented here that lead to rather similar, and in fact statistically indistinguishable, TNDVI dynamics. As such it remains to show how an approach like the presented one actually helps with the understanding of the effects of multiple drivers.

The EDPM is empirical and has the ability to learn patterns when pointed at them. The model is able to **represent** these patterns and build seasonal trajectories based on representation, but it **cannot explain** them or improve our understanding of drivers. We agree that the TNDVI is an integrated measure of canopy state and if we had a suitable alternative available (e.g., systematic daily records of FaPAR or LAI), then we would try it instead. The TNDVI, though, gives an advantage of straightforward and unambiguous data assimilation and validation in addition to possibility of training the model on consistent records.
Rephrased: *This modeling approach enables representing interactions of multiple drivers/events that drive the temporal variability of canopy characteristics (Seastedt and Knapp, 1993; Knapp and Smith, 2001; Zhang et al., 2010; Schwalm et al., 2010).*

I should like to ask the authors to modify their text to consider the above points. My main reservation about the model presented here, is that it seems to have a very large number of parameters and that it is not discussed how well these parameters are actually determined in the light of errors on TNDVI. So equation (7) is used to train the sensitivities. If there are errors one Anext, what does this mean for the sensitivities? I am afraid it is not clear to me how the model is trained. How do the observed TNDVI data enter the optimization? What does the index i stand for in (7)? Please explain more clearly what you do with the observed TNDVI and how your training procedure works. Maybe give an example, or full equations.

Done. The formula and explanation for equation (7) now reads as:

$$J_i = \frac{1}{et}\sum_{k=1}^{et}\left(A_{next} - (s_i i_k)^{\frac{1}{n}} A_{current}\right)^2 \qquad (7)$$

$$\min\{J \mid 0 < s_i \le s_{\max}\}$$

where i is intensity, s_i is given canopy sensitivity from a range of considered s values ($0<s_i<s_{max}$), et denotes total number of events of one type that were considered for training, n is number of events occurred in the same day as the event of interest, k is the sequential event index, and $A_{current}$ and A_{next} are the consecutive observations of modeled canopy property (TNDVI) The value of s_{max} was simply 2.5 which is more than double of maximum daily step change coefficient seen in observations from Mead, NE; increment step for iteration of s was 0.0001.

As things stand I do not feel I can fully comment on the model and certainly I would not be able to reproduce it. I think this should be possible, though. Please give a full list of parameters and equations for the finalized model.

We feel that our changes in the revised version will be more clear and accessible, but please refer to our response to your first comment.

Minor points
P5306,l6. What is the "canopy driven factor"? This is not used/explained before. The sentence does not seem to make sense. **Done.** It now reads “phenology driven factor”.

Fig. 1 Explain clearly what the different modules are doing. This might help understanding how this model is set up. **Done.**

Fig. 3. Where, when, for which crop, etc is the TNDVI shown? **Done.**

Fig. 4. How are the "mean seasonal errors" calculated? What are they exactly? **Done (Explained in the caption).**

Generally, provide more explanation in your table and figure captions about what is shown. **Done in appendices.**