

Review “A comprehensive land surface vegetation model for multi-stream data assimilation, D&B v1.0”

General remarks

The manuscript presents the D&B model. The D&B model was designed to assess carbon and water cycle dynamics and improve the understanding of underpinning processes via assimilation of a wide range of earth observation data at different scales. The authors evaluate their model by comparing simulations results of several variables against observational data at two sites.

I agree with the authors, that there is a need for data assimilation especially at large to global scales using earth observation data and I believe that the D&B model can contribute to fulfilling that need. Overall, the manuscript is well structured but requires a thorough check for language errors and consistent use of terminology. Additionally, in my opinion, the introduction (major comments 1-3) and the results (major comment 5-7) sections could be improved.

I hope that the specific comments below help to improve the quality of the manuscript.

Major comments

1. Introduction: The introduction does not provide an overview of currently available comparable models and whether such models already fulfill some of the requirements identified from BETHY studies (L54-59). I believe this is needed to fully evaluate the novelty of the D&B model.
2. L22-24: While I agree with the authors' statement on the reliable characterization of carbon and water fluxes, I think that their justification “we currently lack a robust, spatially and temporally explicit knowledge of sources and sinks of CO₂, or of the drivers of those variations” could be more concise. It is not clear if the authors refer to sources and sinks of CO₂ in general or which is what I would expect considering the scope of the article, only those related to vegetation dynamics.
Additionally, the statement should be underpinned by an overview of the current knowledge. Specifically, it would strengthen their statement to provide information on the following:
 - The spatial and temporal resolution and coverage of state of the art models and data products.
 - An overview of known and suspected drivers.
3. L28-35: This seems to be a crucial section motivating the development of D&B. However, I do not understand how the need to include primary earth observation data in data assimilation frameworks provides added value for indirect or secondary earth observation data products (L30-35).
4. Section 2.3: I am missing how DALEC and therefore D&B constrain the ratios between the biomass of different plant organs (e.g. leaf to root ratio) and also if allometric relationships are considered somehow. Is this handled implicitly via the regional-scale calibration? This should be included in the model description and potentially the discussion of model limitations.
5. Section 4.2: The section is missing the description of the initial conditions for vegetation biomass and soil carbon. If these do not need to be initialized how are initial values determined? Generally, an overview table in the main text or SI containing all variables for which initial conditions need to be prescribed and the respective initial values would in my

opinion improve the clarity of the model setup description.

Additionally, the description jumps between sites which reduces reading flow. I would kindly ask the authors to describe the sites after each other.

6. Section 4.3: The description of the evaluation metrics is not detailed enough. For example it is not clear what is meant by multi-year averages of the annual cycle and the multi-year mean and how these differ. From following sections it becomes clear the second is the average of the annual sum but the explanation is missing. The authors should also provide equations for all evaluation metrics.
7. Section 5.1 and 6.3: The authors explain the intra-annual fluctuations of FAPAR by the LAI seasonality (L414f). However, I would expect that the two PFTs evergreen coniferous forest and evergreen shrub should not have a strong LAI seasonality which is confirmed by the observations (L423f). The authors briefly discuss this in the limitations section and relate it to phenology but do not provide a detailed explanation of model behavior (L523-528). I understand that the authors cannot provide a calibrated version of D&B at this point but would like to see how their results relate to eq. 147-149 to fully explain this behavior.
8. L541-548: This paragraph is quite generic and in my opinion applies to process-based models in general. It could in my opinion be extended to highlight how this is different for D&B.

Minor comments

1. L25f: The sentence contains some small language issues:
 - I suggest to change “terrestrial carbon stores” to “terrestrial carbon storage” because stores has multiple meanings.
 - Unclear what “those variations” are. I assume variations in C fluxes and storage but it should be clarified.
 - Unclear what is meant by “forcing factors”. I assume changing climatic conditions (e.g. temperature change and so on) but it should be defined.
2. L34: Change “those data” to “earth observation data” or “EO data” or “remote sensing data”.
3. L49: “is” missing after “BETHY”.
4. L62f: “fast and “intermediate-time scale” and “slower processes” are vague. The authors could add more specific information. I assume they mean hours to months and months to years.
5. L114: Should be “drive” not “drives”.
6. L129: Consider replacing “arrive at” with “calculate” or “derive”.
7. L159: Consider replacing “taking account of” to “accounting for”.
8. L150-157: I have several minor issues with understanding:
 - Is root water supply capped at field capacity? I assume yes but it is not stated.
 - “Actual stomatal conductance are then set such that transpiration is capped at the root supply rate”: First, it should be “root water supply rate” not “root supply rate”. Second, I would assume that it is capped at the minimum of root water supply rate and demand for transpiration. Can you please confirm and elaborate this.
 - Could you add a reference to the equations of the supply-demand calculation to make it easier to find.
9. Section 2.2.: Variable names are introduced in some parts of the sections (e.g. L135-145) but not in the entire section. I think this should either be consistent or the authors should explain why they introduce certain variables in the main text and others only in the SI.
10. L166 “all derive” should be “are all derived”.
11. L167: I believe there is a typo in “fixed site of PFT specific fractions”. Is it “suite” instead of “site”?
12. L189: I believe “as photosynthesis” should be “for photosynthesis”.

13. L193 “ensure” instead of “insure”. Also I do not understand where the “separate calculation of FAPAR at the correct solar zenith angle” has to be performed. Is it within the model or is it a correction of data from observations. If the first, what is the difference to using FAPAR calculations from the model run?
14. L261: Delete “bptj” and please elaborate how the parameters were chosen. Was this part of the calibration or an expert assessment?
15. L266: Change “surface layers” to “surface layer”.
16. L303, 442 and other occurrences: Here the authors refer to Sodankylä as the boreal site deviating from their so far consistent terminology. Similarly they sometimes refer to Majadas de Tietar as the savannah site.
17. L321: Change “than to” to “than” or “compared to”.
18. L323: DOY has not been introduced as an abbreviation yet.
19. L323f, 333 and other occurrences: I propose to change “underestimate” to “underestimation” and “overestimate” to “overestimation”. While the current forms are not wrong, they are less common and could be confused with verbs.
20. L332: Change “math” to “match”.
21. L360: Change “discussion” to “discussing”.
22. L361: Change “creates” to “create”.
23. L373ff: This sentence is quiet long and I do not fully understand its meaning. E.g. “[...] measured NEE [...] that are not reproduced by the measurements”. I believe one of these should refer to simulations and not measurements.
24. L383f: What is the reason for the overestimation under favorable conditions?
25. L385ff: I find this sentence hard to understand and suggest to split it by variable (GPP and TER) and not method (simulation and observation).
26. L391: This is true for GPP (2.11 vs 2.25 modelled and 3.39 observed) but not for NEE (1.88 vs -0.09 modelled and -0.05 observed). Neither the order of magnitude is similar nor is the difference to the model mean smaller. I also do not understand the significance of this results. Please elaborate.
27. Section 6: When referring to the results references to the respective sections are missing.
28. L495 Change “as” to “at”.
29. L565f: I am not sure that this can be generalized. The process model may also not be able to match observations for a specific variable within reasonable bounds of the parameter space if the process is implemented but its formulation is not universal and therefore not applicable to the context of the experiment. So you cannot say that a process is missing but only that either a process is missing or the formulation of processes used is not suitable.
30. Fig. 9-11: The x-axis ticks are not correctly labelled. The axis gives hours of the day and not the month. I suggest to move the month either above or below the plot.
31. Fig. 2: S_m is not explained in the caption.
32. Fig. 3, 4, 6 and 7: You could consider adding the 5th to 95th percentile values to illustrate inter-annual fluctuations. This would in my opinion also underpin your results where you compare inter- and intra-annual match between observed and simulated data.
33. Fig. 4, 5, 7, 12, 13, 14 and 15: Captions are missing the color scale.
34. Fig. 12 and 13: Dates in the x-axis label miss separators between year, month and day.
35. SI Title of section 1.2.4 only refers to evaporation but section describes also transpiration.