Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2018-264-AC2, 2019

Geoscientific Model Development Discussions

Open Access

Interactive comment on "Optimizing shrub parameters to estimate gross primary production of the sagebrush ecosystem using the Ecosystem Demography (EDv2.2) model" *by* Karun Pandit et al.

Karun Pandit et al.

karunpandit@boisestate.edu

Received and published: 8 April 2019

General comments This article is about the optimization of sagebrush parameters based on GPP in the EDv2.2 model and in Great Basin.

The development and optimization of specific vegetation - here a specific shrub – are currently a key research area to increase model adequacy with observations and enable the simulation of the future development of our ecosystems. However, contrary to that suggested by the actual title, in this article there is no presentation of GPP results

Printer-friendly version

Discussion paper

estimation but rather only some "optimization and validation" of parameters. Moreover, the article does not present some general case but more a specific situation: a very small zone (200m2 simulated, 2 years of observation in 2 points). I suggest to change the title to make it more explicit. The model used here is EDv2.2, which seems interesting for medium scale simulation. But the methods used limit the scope (and so the interest) of this study. The two differents sites of observation are located close to each other but differ in the type of sagebrush present (a small specie and a big specie). However, only one allometry parameterization is proposed. Your choice to have a dynamic vegetation is curious considering you work on two very specific and well documented sites for one unique year (one for optimization and one for validation). Some of the methods used (as the use of sensitivity index) rely on strong hypotheses, that have been presented only in the discussion. Some deeper bibliography could have made it possible to anticipate errors. The purpose of the article to estimate GPP is thus more a local application of the optimization of parameters in order to simulate (not here) the GPP. Due to the small data set and the validation performed without any statistical test (and one of the two cases that seems not so adequate), there is no insurance that the method could be applied for other years (to predict) or in other sites. As no specific development was presented here, except an adaptation of parameters for sagebrush allometry, the relevance of this article for publication in GMD can be questioned.

We have revised the title of the manuscript to better match the content. The study is primarily focused on the development of shrub (sagebrush) PFT parameters to use in EDv2.2, and to observe the performance of the model for the newly developed sagebrush PFT (and wherein we used GPP as variable to conduct these comparisons). We agree that allometric relationships for different sites could not properly capture the fine scale heterogeneity in the ecosystem. For this study, we limited our objective in developing general sagebrush parameters, without trying to separate uniqueness of different sagebrush species. We used simple sensitivity and optimization analysis methods, to constrain the selected parameters. In further studies, we intend to capture the non-linear dependencies among these parameters to better constrain them for model estimates; however this is outside the scope of the present study.

Globally, considering the 14 detailed comments presented below, the editorial and figure quality of the present manuscript, I consider that in this state this article lacks of consistency and does not reach the standard quality expected for GMD.

Please see our responses below.

Specific comments 1) Not only simulations or field observations can be used to quantify GPP (p.1 l.6). A third essential data set comes from satellites and remote sensing, providing continuous values (spatially and over time). There is for example the GPP from the FLUXCOM project Tramontada et al., 2016 https://doi.org/10.5194/bg-13-4291-2016 and Jung et al, 2017 https://doi.org/10.1038/nature20780) or from a linear relationship with the Sun-Induced Fluorescence (Su et al., 2017 http://resolver.caltech.edu/CaltechAUTHORS:20171016-145548969). Of course the problem of isolating the GPP for a PFT remains . . . as is the case for the observations used in this study. Moreover, this GPP data can be used (if you know the vegetation distribution) to do more efficiency optimization and/or validation (largest temporal and/or spatial scale).

Thank you for the suggestions - we agree that additional data are ideal to quantify GPP. Given the context of this paper (please see comments above), we are limiting our analysis to the flux towers and future work will incorporate the remotely sensed data products and should be useful to assess GPP in broader spatial terms.

2) As indicated in the article (p.2 l.19), it could be difficult in models to represent and parameterize specific ecosystems and they are historically not well simulated. But this is currently a major point of development in land surface models, as for tundra (mosses, shrubs,. . .) which are now more and more represented. The sentence “Semi-arid, nonforest ecosystems provide an excellent example of this limitation” (p.2 l.20) has to be more documented. More generally, a short review of the current state of what is done in different models would be necessary in this article. Nevertheless, it is probable

that these models are not yet sufficient to reproduce specifically the sagebrush.

Thank you, we agree and have additional references cited P.4.l.1.

3) Globally all the references of the article have to be checked. There are wrong dates in the reference list (e.g. for Bradley and Chambers), some references are missing (e.g. Skamarock et al, 2008 and Wright et al., 2004), others are never used in the text (e.g. Brabec et al,2001 and NPS, 2018) and one seems wrong (Davidson et al., 2011 about amazon forest to illustrate tundra). You also have two undifferentiated "USDA, 2018".

Thank you for pointing this out and we have updated the references throughout the manuscript.

4) In the introduction (as suggested in the title) you say that you are going to predict the GPP (p.3 l.4). This seems a little ambitious compared to what is actually done in the result section: an optimisation and validation. In my sense, prediction consists in running the model in the future and simulating the future evolution of GPP.

We have changed the title and agree that we are not predicting GPP but estimating GPP to evaluate the model performance with a sagebrush PFT.

5) At the beginning of the methods (p.3 l.16 to 23), you are doing a distinction between two types of model: "gap" or "big leaf". If the general differentiation between both is clearly understandable, some inaccuracies have to be checked and the references have to be improved / updated. (a) p.3 l.20 and l.22 you indicate that in individual based models you can have competition, coexistence and disturbance, and that it is a limit for the big leaf model. But you have also big leaf models (DGVM) with competition, disturbance,... (b) p.3 l.23 you indicate that individual-based models have problems due to computing cost, but this is becoming less and less of a problem and currently many large-scale models (initially big leaf) have developed individual based version. Moreover, in this article the small spatial and temporal scale clearly does not seem to

be a limit, and following your distinction would seem in this case most appropriate?

Thank you for pointing this out. We agree with the reviewer that there are some “big leaf” models with competition. The challenge with these models, however, is they do not capture the demographic processes such as vertical light competition, competitive exclusion, and successional recovery from disturbance. To make it more clear, we changed the word “competition” in the manuscript to “demographic processes”. Considering comments on the IBMs, we agree with the reviewer that computation time is becoming less important in these models. However ED2 is not purely an IBM, as we mentioned in the manuscript (P.3.l.18) its a cohort based model which incorporate different processes.

6) In the parameter description and associated equations (p.4 l.13 to p.5 l.11), you need to be clearer: it is difficult to follow. Directly when you list the eleven parameters I suggest that you use the same order that you use after and that you indicate directly the name of the variables used in the equations (1) and (2). For clarity, these abbreviations have to be everywhere in italics (p. 4 l.22, l.27, 28,...) and called back each time that they are used (e.g. p.5 l.4 for “CO2 concentration within the leaf boundary (Ds)”). Moreover, it is not indicated what the Cs parameter is (equation 2). I suggest also that you indicate how the “stomatal control is affected by soil moisture” (p. 5 l.3).

Thank you for pointing this out. We have added a table (P.4.l.10) to describe parameters we have used for the analysis and put it in a sequential order to match the writing in the test. We also added text to clarify how ‘stomatal control is affected by soil moisture’ in P.5.l.10. Additionally, we have provided reference (mainly Moorcroft et al., 2001; and Medvigy et al., 2009) for detailed information on equations and processes.

6) You have to take care about the quality of the figures and tables, and the associated legends (even in the supplementary). The figures have to be clearly understandable. (a) in Figure 1, the WRF grid does not make it possible to see the vegetation around the simulated polygon. I suggest that you indicate in the legend the general location

(at least "USA") and the signification of "LS" and "WBS". (b) in table 1 you indicate for the "DBH to Height" an equation with negative "b" value with a negative term "-b x DBH", so the Ht is negative. Moreover you have to give the units of variables (in cm?). (c) in table 3 you use "*" for optimized parameters and for value ranges from EDv2.2. (d) in Figure 3 and 4 you give the number of "days" in "2016". However, it seems not to correspond exactly to a year and it is never explicit: in the text "spring" is for the days"200 to 250" and in the figure 4 "2016" starts from October (2015?). Please revise the x-axis labelling. (e) in Table S1 you have to indicate clearly the dimensions for each parameter and in a consistent manner (eg "[m]"). (f) in table S2 I suggest that you indicate how the rank is done (by NSE) and that you give the dimension of the parameters. (g) in figure S1 it is not possible to see clearly the differences between simulations. Maybe you could use monthly means?

Thank you for pointing this out. We have updated these figures / tables. (a) we updated figure 1 related to study area which now shows location of LS and WBS sites in Reynold Creek Experimental Watershed (RCEW) area, (b) We removed -ve in the coefficient and provided unit for DBH, (c) in table 4 (earlier 3) we adjusted the confusion with regards to the use of '*' symbol, (d) we have updated the figures to make it more readable (e) we provided information about NSE score equation used in the ranking (Supplement P.6.l.9). (f) we have provide unit for applicable parameters (Supplement Table S1), (g) we updated the figure (Supplement Fig S2) to show average monthly GPPs to make different simulations more discernible.

7) The 2.3 section is called "Inventory and EC tower data" but is mainly about allometric equations. Moreover the approach method to describe shrub allometry can be improved. You suggest that the problem comes from the fact that the model is "originally developed for tropical forest" (p.6 l.9), when it seems to be more precisely due to allometric equations developed for trees and not for shrubs. Then, it could be appropriate to explicitly indicate that from the allometric data available, you transformed them (if I understand well) to a theoretical height considering that the shrub is a cube (?). But

more importantly, it could be beneficial if you evaluate the impact of this hypothesis, for example by showing the adequacy between “DBH to Height” results or the height simulated compared to observed height. There is also another solution: to change the allometric equation for shrubs, as is used in other models (e.g. Druel et al., 2017 https://doi.org/10.5194/gmd-10-4693-2017).

This is a good idea and we compared the predicted height from the cube root volume with observed sagebrush height using a new set of data from the Great Basin (see Supplement FigS1). We observed a good match between observed and predicted heights for sagebrush.

8) There is no overlapping between the period of station data (2015-2016) and the years used for the forecast, 2006 to 2014 (p.7 l.12). If it can be understandable to use random years for long term “spinup”, using “random years” for all simulations and optimization/validation can introduce a new bias superimposed on the parameter set. Even more important, if you use a random forecast year to simulate specifically 2015 and 2016 (validation and simulation), that means the difference between both simulations is a random year? We used corresponding years of meteorological data for simulation in the revised manuscript. We used 1 km WRF data from 2001 to 2017 for both the sites studied. This will help reduce the interannual uncertainty that may arise from using meteorological data from a random year.

9) For the initial parameterisation of the 11 parameters, you choose a sensitivity index. But there are two fundamental hypotheses to use such index: you expect that the responses to the parameters are linear and that there is no interaction between parameters. Unfortunately, you never indicate those hypotheses! It is true that at the next step (for optimization) you use a more adapted method (not requiring such hypotheses) and that in the discussion you put two related sentences, but the method is not consistent with the optimization and the hypotheses are required from the beginning. The test of the mean of the best sets of 10 parameters shows that the hypotheses were not well considered.

We agree that the chosen method assumes linear dependencies of selected parameters with the target variable. We spent extensive time on developing the shrub (representing sagebrush) PFT for the EDv2.2 model (e.g. establishing allometric relationships) and several preliminary model run-ups to match with the ecosystem conditions. We used the exhaustive (brute force) method due to computational limitations. This study was mainly intended to introduce the sagebrush PFT and its implementation in EDv2.2. We agree that additional robust optimizations (and sensitivity) should be performed. We've modified the paper to highlight this intent and the conclusions we may draw from the existing work. We have added lines to state the limitation of the applied SI method and our assumption on parameters under methods section (P.8.l.18).

10) You indicate that your "simulations were configured to allow" that other plants than shrubs can grow in the model (p.9 l.1). That means that you specifically activated the competition between species and so other plants can grow? If this is the case, you introduce new uncertainties and so probably directly biases to the optimization and comparison with GPP observations! I really do not understand why you do not use the observed fraction of vegetation in your two (well documented) stations. On the one side you work on very few observations and simulated points (in time and space), but you do not limit the variability induced by the model configuration. Why?

The study site is heterogeneous and thus we need to allow additional PFTs to grow to capture total GPP. We do not understand the question here but to clarify we used density information for initialization that has been collected at the sites.

12) The results section suffers from the limitation of the method: only one polygon is simulated, two observation sites considered, with heterogeneous vegetation inside each site (grasses and shrubs) but also between sites (Low Sagebrush /W. Big Sagebrush), and only two years of data (with one not complete for one site), one for simulation and one for validation. Thus, it is not possible to represent inter-annual or spatial variability. Likewise, no statistical tools are used to validate the optimization. We can just observe that one is coherent (WBS) and the other is bad (LS) (the value of the

differences are also missing, e.g. p.14 l.5 to 8). In conclusion it is not obvious that the values obtained for the parameters can be used for other years or sites.

As stated above, we have simulated both sites with respective ecosystem and atmospheric conditions to address variation between the sites. In our revised analysis, we could use two years of data for calibration and another year for validation. We agree that these are not sufficient to capture inter annual variability but we were mostly limited with the available observation data from the sites. We agree that the values cannot be used for other years and sites until further optimization is performed. We have stated this in the Conclusion (P.17.l.18).

13) Not being a specialist of optimization, I cannot say something precisely on this part. But, the choice of the optimization method is not justified or discussed. There exists currently other methods less computationally costly (such as genetic algorithms) and it is possible to extract statistic values to evaluate the efficiency (such as the variability fraction explained before and after the optimization).

We agree there additional optimization tools could improve the results and provide robust information on sagebrush PFT parameters (Please refer to answers to Q.9 for more).

14) The discussion allows to go further, but showing mostly the limits of the methods used for the study, which should have been stated earlier in the methods (e.g. the nonlinear dependence among parameters). This shows also the gap between the objective indicated (to predict the GPP of sagebrush) and the results (not really validated, even in very restrictive conditions).

Good point, we have tried to clarify the objectives and the results and how our study has contributed to the overall modeling of shrub-steppe (P.2.l.30). We have stated limitation of our tools (P.8.l.18) and potential improvements we would achieve with different methods (P.17.l.2)

Printer-friendly version

Discussion paper

Specific comments

p.1 l.10. Suggested change: "one of the most critical" to "one critical" The text has been changed. p.1 l.28-31. Suggested change. Remove from "we expect that. . . " (to put in the conclusion?) As suggested, the lines were removed. p.2 l.3. Need for a reference for "anthropogenic CO2 emissions" The text has been removed. p.2 l.4. Suggested change: Add a small definition for "photosynthesis" We have updated the text. p.2 l.10. Suggested change: "distinct ecosystems" to "distinct ecosystems at large scale" We changed the texts p.2 l.20. There are currently two spaces after "ecosystems". We corrected. p.2 l.27. How do they suppress fire? Removed the text 'suppress fire' p.2 l. 34. After "Great Basin", indicate the density of station (or indicate if there are only two stations. . .) The text is revised.

p.3 l.13. This section (2.1) could gain in clarity if you distinguish (a) the general model presentation (p.3.14 to p.4 l.8) and (b) the presentation of parameters used in this study and their related equation(s). We tried to differentiate the information in the section through different paragraphs . We added a table showing parameters used in the study followed by brief descriptions and controls of the parameters. p.3. l.18 "plant function type" abbreviation is already defined just above (p. 3 l.6). We updated the text accordingly. p.3 l.23. You use acronym "IBMs" which is not defined. Please define it l.21. We corrected as per your suggestion. p.4 l.13. Suggested change: parameters. These included" to "parameters:" The text has been revised. p.4. l.18-19. Suggested change: "here we are trying to describe the ones related tothe parameters we have use in this study" to "here describe the ones related to the parameters used in this study" We made suggested change. p.5 l.8. It could be important to state from where the "allometric allocation" comes from, and maybe indicate that they are in Table 1 ? We referred Table 1 for the allometric equation referred in the text P.5.l.18 p.5 l.16. Please clearly indicate where it is (country, state). We updated with region and Country. p.5 l.16 to 18. If I understood well, you have to indicate that the "200 m x 200 m polygon" is the simulated area in this study (using the 3km resolution WRF forecast). Likewise, in the legend of Figure 1 (p.6 l.2) change “study polygon” to “simulated polygon”. We updated the Figure 1 showing study area. p.6 l.16. Suggested change: Add a line break before the “GPP data. . .” We updated with a line break to separate two types of data sources. p.8. l.4. If the sagebrush parameters come only from bibliography, put the citation l.2. We updated the text to appropriately represent the procedure P.8.l.3. We also updated supplement Table S1 with all PFT parameters to clearly state the source/reference of different parameters. p.8 l.10. Indicate why “370 ppm” or to which year that corresponds (2000?). we updated the text as suggested (P.8.l.11) p.9 l.29. Change “Fig. 2b and d” by “Fig. 2b, c and d”. We made necessary edits as suggested. p.12 l.21. Change “Table 5” to “Table 6”. We made necessary correctons. p.15 l.4. Suggested change: “was observed” to “was obtained” Changed the text P.16.l.10. p.16 l.16. Suggested change: Add a line break after the “GPP.” We made the edits as suggested. p.16 l.20. I am not sure that you can say “quite well”. Text has been updated.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-264, 2018.