Biogeosciences Discuss., 11, C4383–C4396, 2014
www.biogeosciences-discuss.net/11/C4383/2014/
© Author(s) 2014. This work is distributed under
the Creative Commons Attribute 3.0 License.

Biogeosciences
Discussions

Interactive comment on "Identifying environmental controls on vegetation greenness phenology through model-data integration" *by* M. Forkel et al.

M. Forkel et al.

mforkel@bgc-jena.mpg.de

Received and published: 14 August 2014

We thank the anonymous Referee 1 for the very valuable comments (Biogeosciences Discuss., 11, C4000–C4002, 2014).

1 Length and complexity of the manuscript

"*The manuscript is well written and the study, although complex, very well executed. The only downfall is that the reader is bombarded by comparison after comparison showing how the GSI version of the model is better, leading to 12 figures in the main text, 21 in the appendix, which largely bury the key messages.*"

We agree with the referee that our manuscript is full of comparisons and includes many figures. However, we think these comparisons are needed and should be also part of the publication. LPJ is a widely used dynamic global vegetation model that is applied in several fields. Replacing a core module as phenology of such a model requires 1) a detailed model evaluation to quantify the impact of the changed model structure on model outputs and 2) to ensure that model results are comparable with observations. Thus, we are evaluating LPJmL against independent data streams (biomass, tree cover, ET). Additionally, we are evaluating LPJmL against seasonal, monthly, inter-annual FAPAR dynamics and trends to demonstrate the applicability of LPJmL in diagnosing FAPAR dynamics on different time scales.

We consider moving results from the main text to the appendix or from the appendix to supplementary material to improve the readability of the manuscript.

2 Model complexity and extrapolation capabilities

"*My only main comment is that, with 12 free parameters, should we not expect the GSI model to perform better than the original model (5 free parameters). What is the potential for over-fitting the model here? The main question is whether GSI is better at predicting out of sample. It would be good to see a test of the model optimized to the first half of the time series and predict the second half. Better still, seen as we are interested in future climate change applications with this model, optimized to the northern half of the distributions of each PFT, and used to predict the southern half. Tests such as these are needed to give us true confidence that a more complex model with double the parameters is truly better.*"

We identify two distinct concerns in this comment: 1) The higher complexity of LPJmL-GSI over LPJmL-OP with a risk for over-fitting and 2) the capability of LPJmL-GSI in accurately extrapolating to time periods or spatial domains that are distinct from the optimization data. We will discuss these two issues separately in the following.

2.1 Model complexity of LPJmL-GSI vs. LPJmL-OP

Indeed, the LPJmL-GSI phenology module has more parameters and thus a higher complexity than LPJmL-OP. In total, 18 parameters (12 phenology parameters + 6 productivity/albedo parameters) were considered in the optimization of LPJmL-GSI and up to 10 parameters (number depends on PFT) for LPJmL-OP. We agree with the referee that from a statistical point of view LPJmL-GSI should better reproduce observed FAPAR dynamics than LPJmL-OP but involves a risk of over-fitting. Nevertheless, not all of these candidate parameters were included in the optimization experiments. We identified some parameters as insensitive and such parameters were excluded from optimization experiments. Thus the number of free parameters in optimization experiments of LPJmL-GSI ranged between 7 and 11 dependent on PFT. This limits the potential for over-fitting. Following the statistic reasoning, one can evaluate the model fit with respect to the number of parameters by computing the AIC (Akaike's Information Criterion) (Burnham and Anderson, 2002, p.61). The model with the lower AIC value would indicate the better model because it can provide a similar fit with less parameters. We computed AIC for grid cell optimization experiments of LPJmL-OP and LPJmL-GSI, respectively. Nevertheless, these AIC values are not directly comparable because not the same set of grid cells was used in both experiments (Figure 1). LPJmL-GSI had lower AIC values in the BoNS tree PFT, and in the herbaceous PFTs (TrH, TeH, PoH). AIC values were similar in the TrBR, TeBS, BoNE, BoBS PFTs. AIC values from LPJmL-OP were lower in the TrBE, TeNE, and TeBE PFTs. This is not surprising as LPJmL-OP has a fixed phenology in evergreen PFTs (i.e. no parameters). In order to compute AIC differences (dAIC) and to investigate spatial patterns of dAIC in terms of monthly FAPAR, we estimated AIC from the global LPJmL-OP-gc and LPJmL-GSI model runs (Figure 2). Based on this analysis of dAIC for monthly FAPAR, one could select LPJmL-GSI as the better model in arctic and large boreal regions

and in temperate and tropical grasslands. LPJmL-OP could be selected in temperate, subtropical ad topical forest regions.

Nevertheless, although LPJmL-OP is seemingly the better model in terms of AIC in some regions, it is definitely the worse model in terms of considered environmental processes. LPJmL-OP ignores potential drought or heat stress effects on phenology in all PFTs. Additionally, the summergreen routine in LPJmL-OP has four parameters whereas the corresponding cold temperature limiting function in LPJmL-GSI has only three parameters. Thus, LPJmL-OP is the more complex model for the relationship between cold temperature and phenology and has only a lower overall complexity because it misses effects of light and water on phenology in summergreen and herbaceous PFTs. "If a particular model (...) does not make biological sense, this is the reason to exclude it from the set of candidate models" (Burnham and Anderson, 2002, p.17). In conclusion, although LPJmL-OP has a lower overall complexity and turns out to be a better model in terms of AIC in some PFTs, it is not an alternative to LPJmL-GSI because it misses important environmental controls on vegetation phenology.

2.2 Extrapolation capabilities of LPJmL-GSI

The referee asked to demonstrate the extrapolation capabilities of LPJmL-GSI by splitting the data in temporal or spatial distinct sets for model optimization and evaluation. Such a test helps to understand the model performance in different time periods, regions or environmental conditions. We did not split the observed FAPAR time series in two time periods for model optimization and evaluation ("split-sample test", (Klemeš, 1986)) because we wanted to use the full time series length to maintain enough information about inter-annual variability and trends in model optimization experiments. More than such a test, the referee suggests splitting the data spatially (northern and southern half). We think a north-south splitting is not very insightful because especially boreal and arctic PFTs don't have an equivalent at the southern hemisphere and most

PFTs occur only in relatively small latitudinal bands. However, we split the data spatially as the referee suggested: we optimized LPJmL-GSI only in a few randomly selected grid cells and evaluated the model in all other grid cells. Optimization grid cells had a dominant cover of one PFT whereas all other evaluation grid cells had mixed PFT cover. But the referee is right, we did not explicitly show the difference in model performance between optimization and evaluation grid cells in our manuscript. To test the extrapolation capabilities of LPJmL-GSI into different regions or under different climate conditions, we investigated the relationship between model performance (expressed as the correlation between monthly FAPAR time series from GIMMS3g and LPJmL-GSI) and the distance to the closest grid cell that was used for a PFT-level optimization of LPJmL-GSI (Figure 3). If LPJmL-GSI would be not capable of extrapolation, we would expect a decrease in correlation with increasing distance from optimization grid cells.

We found no general decrease in model performance if optimized LPJmL-GSI model parameters were applied to distant geographic regions or under different temperature conditions (Figure 3). Whereas only in the Am, Cw, BSh, Df, and ET climate types significant lower correlations in distant (600-800 km) than in close (< 200 km) grid cells occurred, we found constant or even improving correlations with increasing distance from optimization grid cells in all other climate types (Figure 3 c). More important than the spatial extrapolation capability is the capability in extrapolating to different environmental conditions especially with respect to climate change applications. Consequently, we also tested if the correlation between simulated and observed FAPAR time series depends on the difference in mean annual temperature between each grid cell and the corresponding closest optimization grid cell (Figure 3 d). We did not find significant lower correlations in grid cells that were 3 to 5°C warmer than the closest optimization grid cell. This indicates that under common climate warming scenarios of 0.3° to 4.8°C (IPCC, 2014), LPJmL-GSI will likely simulate FAPAR with similar performance like during the optimization.

We will consider including this analysis in the appendix or supplementary material of

the revised manuscript.

3 Minor Points

"*Introduction: Page 10919, line 18: Give citation for these claimed browning trends.*"

We suggest to add the following citations: (Baird and Verbyla, 2012; Bi et al., 2013; de Jong et al., 2013).

"*Page 10919, line 21: Why focus on just the boreal browning here? What about the other regions?*"

We suggest to add references to studies that investigated relations between browning trends, droughts and land use change in subtropical regions (Cook and Pau, 2013; van Leeuwen et al., 2013).

"*The introduction is in general too long and could be edited to improve flow.*"

We will try to shorten and improve the introduction as soon as we will get other referee comments.

"*Page 10925, line 9: Please give the values of k used, and cite references.*"

Parameter values of the light extinction coefficient were optimized in optimization experiments. Prior and posterior values can be found in Tables D2-D5.

"*Page 10928, line 21: It is not clear to me how the fact that LPJ here is a prognostic model makes it impossible to use a running window averaging approach.*"

The running window averaging approach in the original GSI model computes the actual daily GSI value as the mean value from daily iGSI indicator values for time period of 21 days (Jolly et al., 2005, p.622). In LPJmL-GSI, the actual daily phenology status needs to be estimated from the phenology status of the previous day to ensure that the model is fully applicable in a prognostic mode. Otherwise, one needs first to run the LPJmL-GSI phenology module, then apply the running window average on the daily phenology status, and then compute photosynthesis and all other process modules of LPJmL afterwards. This does not correspond to the common modelling approach in LPJmL. An alternative implementation would be to compute the daily phenology status dependent on the average status of the previous x days. We will test this alternative implementation in future studies when we will apply and optimize LPJmL model parameters against site measurements of daily eddy covariance fluxes.

"*Page 10932: "we used"*"

We changed this.

"*Page 10935, line 4: Is Table D2 all optimized parameters, or all relevant parameters. Please clarify in the main text.*"

Table D2 lists all parameters of the LPJmL-OP phenology model and all parameters that were addressed in optimization experiments of LPJmL-OP. Except the parameter GDDbase all of these parameters were included in grid cell-level optimization experiments of LPJmL-OP.

"*Page 10939, Line 19: "In temperate broadleaved evergreen forests, the GIMMS3g FAPAR dataset 20 might have a wrong seasonality." This is quite a bold statement, given that GIMMS3g and its predecessor have been extensively used in temperate broadleaved evergreen forests. I hate to ask for a figure to back up this statement, given that the authors have already included so many in the manuscript, but it would seem one is warranted.*"

The difference in the mean seasonal cycle between GIMMS3g, GL2 and LPJmL-GSI FAPAR in temperate broadleaved evergreen forests is already shown in Figure 8 of the main text. Although the mean seasonal FAPAR cycles from GIMMS3g and GL2-VT are negatively correlated (r = -0.48), the correlation is not significant and GIMMS3g is

within the uncertainty of the GL2 dataset. Thus, we consider using a weaker wording.

"*Page 10954, Line 6: "water availability is regulated through seasonal thawing and freezing of the active permafrost layer". Seasonal thawing co-varies with temperature, suggesting temperature could be used as a driver (and is sure to have lower error propagation than going through modeled soil moisture). I would suggest reconsidering your interpretation.*"

We agree with the referee that seasonal freezing and thawing of upper active layer in permafrost soils is too a large extent driven by temperature changes. As a consequence it is likely possible to explain phenology in arctic and boreal ecosystem only by temperature changes. Nevertheless, air temperature and soil thawing are not completely synchronized because soil temperature depends also on topography, substrate, and the insulating effects of the snow, litter and vegetation cover (Jorgenson et al., 2010; Shur and Jorgenson, 2007; Zhang, 2005). Soils might be still frozen if air temperature is already positive or vice versa. Indeed, we did not find a completely synchronized temporal dynamic of the cold temperature and water limiting functions for phenology (Figure 12 of the manuscript) which suggests that water availability might affect phenology independently from temperature seasonality. Also experimental studies highlighted the role of permafrost-regulated soil moisture on phenology and productivity in boreal and arctic ecosystem (Natali et al., 2012; Schuur et al., 2007). Thus, we assume that temperature might be enough to explain average spatial patterns of phenology in boreal and arctic regions but variations in snow or vegetation cover that affects soil temperature and thus moisture might be important factors in explaining inter-annual variations of land surface phenology.

4 References

Baird, R. A. and Verbyla, D.: Browning of the landscape of interior Alaska based on 1986-2009 Landsat sensor NDVI, Can. J. Of, 1382, 1371–1382, doi:10.1139/x2012-088, 2012.

Bi, J., Xu, L., Samanta, A., Zhu, Z. and Myneni, R.: Divergent Arctic-Boreal Vegetation Changes between North America and Eurasia over the Past 30 Years, Remote Sens., 5(5), 2093–2112, doi:10.3390/rs5052093, 2013.

Burnham, K. P. and Anderson, D. R.: Model selection and multimodel inference a practical information-theoretic approach, Springer, New York. [online] Available from: http://site.ebrary.com/id/10047705 (Accessed 17 December 2013), 2002.

Cook, B. and Pau, S.: A Global Assessment of Long-Term Greening and Browning Trends in Pasture Lands Using the GIMMS LAI3g Dataset, Remote Sens., 5(5), 2492–2512, doi:10.3390/rs5052492, 2013.

IPCC, I. P. on C. C.: Climate Change 2013 - The Physical Science Basis - Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge., 2014.

Jolly, W. M., Nemani, R. and Running, S. W.: A generalized, bioclimatic index to predict foliar phenology in response to climate, Glob. Change Biol., 11(4), 619–632, doi:10.1111/j.1365-2486.2005.00930.x, 2005.

De Jong, R., Verbesselt, J., Zeileis, A. and Schaepman, M.: Shifts in Global Vegetation Activity Trends, Remote Sens., 5(3), 1117–1133, doi:10.3390/rs5031117, 2013.

Jorgenson, M. T., Romanovsky, V. E., Harden, J. W., Shur, Y. L., O'Donnell, J., Schuur, E. A. G., Kanevskiy, M. and Marchenko, S.: Resilience and vulnerability of permafrost to climate change, Can. J. For. Res. - Rev. Can. Rech. For., 40(7), 1219–1236, 2010.

Klemeš, V.: Operational testing of hydrological simulation models, Hydrol. Sci. J., 31(1), 13–24, doi:10.1080/02626668609491024, 1986.

Van Leeuwen, W. J. D., Hartfield, K., Miranda, M. and Meza, F. J.: Trends and ENSO/AAO Driven Variability in NDVI Derived Productivity and Phenology alongside the Andes Mountains, Remote Sens., 5(3), 1177–1203, doi:10.3390/rs5031177, 2013.

Natali, S. M., Schuur, E. A. G. and Rubin, R. L.: Increased plant productivity in Alaskan tundra as a result of experimental warming of soil and permafrost, J. Ecol., 100(2), 488–498, doi:10.1111/j.1365-2745.2011.01925.x, 2012.

Schuur, E. A. G., Crummer, K. G., Vogel, J. G. and Mack, M. C.: Plant Species Composition and Productivity following Permafrost Thaw and Thermokarst in Alaskan Tundra, Ecosystems, 10(2), 280–292, doi:10.1007/s10021-007-9024-0, 2007.

Shur, Y. L. and Jorgenson, M. T.: Patterns of permafrost formation and degradation in relation to climate and ecosystems, Permafr. Periglac. Process., 18(1), 7–19, 2007.
Zhang, T.: Influence of the seasonal snow cover on the ground thermal regime: An overview, Rev. Geophys., 43(4), 2005.

5 Extended figure captions

Figure 1: Distribution of AIC based on the total cost (FAPAR, albedo and GPP) from single grid cell optimization experiments of LPJmL-OP and LPJmL-GSI grouped by the dominant PFT of each grid cell. 530 single grid cells were used for optimizing LPJmL-OP and 348 for LPJmL-GSI. Only 71 grid cells were used in both optimization experiments. These 71 grid cells were distributed in the BoNS, BoBS and TeBS PFTs. Thus, the sample size is too small to compute AIC differences (dAIC) between LPJmL-OP and LPJmL-GSI from grid cell level optimization experiments. The dAIC values in this plot refer to the difference in the median AIC.

Figure 2: AIC differences defined as LPJmL-GSI minus LPJmL-OP based on the cost for FAPAR. AIC was computed based on the sum-of-squared error (SSE) of monthly FAPAR and the active number of LPJmL parameters per grid cell. The SSE was computed between monthly FAPAR time series from GIMMS3g and from the LPJmL-GSI or LPJmL-OP-gc model runs, respectively. The number of active parameters per grid cell does not only depend on the use of LPJmL-OP or LPJmL-GSI but depends also on the number and type of established PFTs per grid cell.

Figure 3: Extrapolation capabilities of LPJmL-GSI in terms of monthly FAPAR dynamics. (a) Correlation coefficient between monthly FAPAR time series from LPJmL-GSI and GIMMS3g (1982-2011). Areas without vegetation, with more than 50

Interactive comment on Biogeosciences Discuss., 11, 10917, 2014.

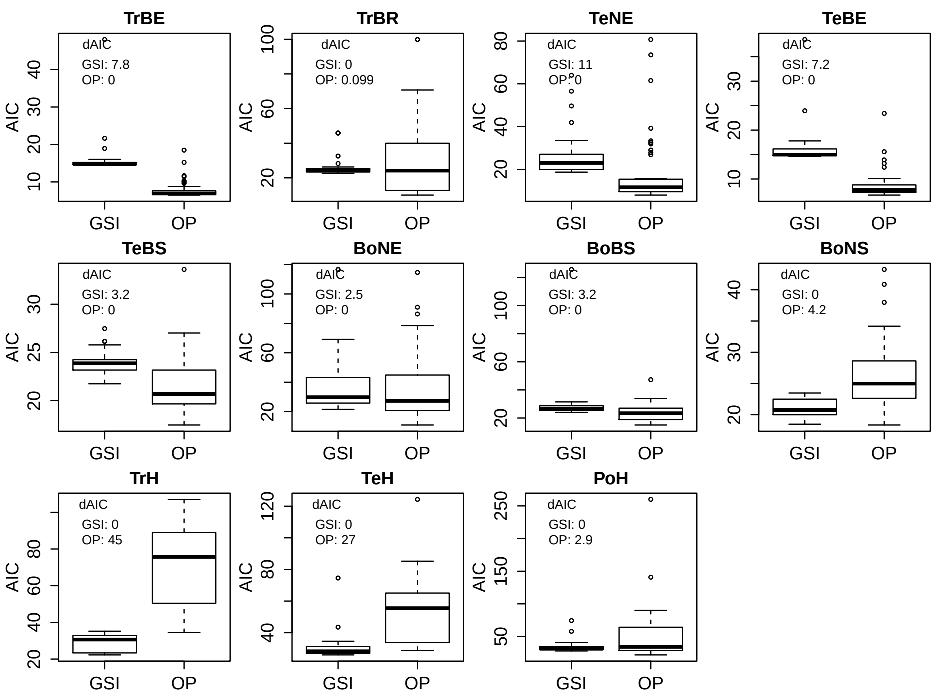

Fig. 1. Distribution of AIC based on the total cost from single grid cell optimization experiments (see extended figure caption at the end of our reponse).

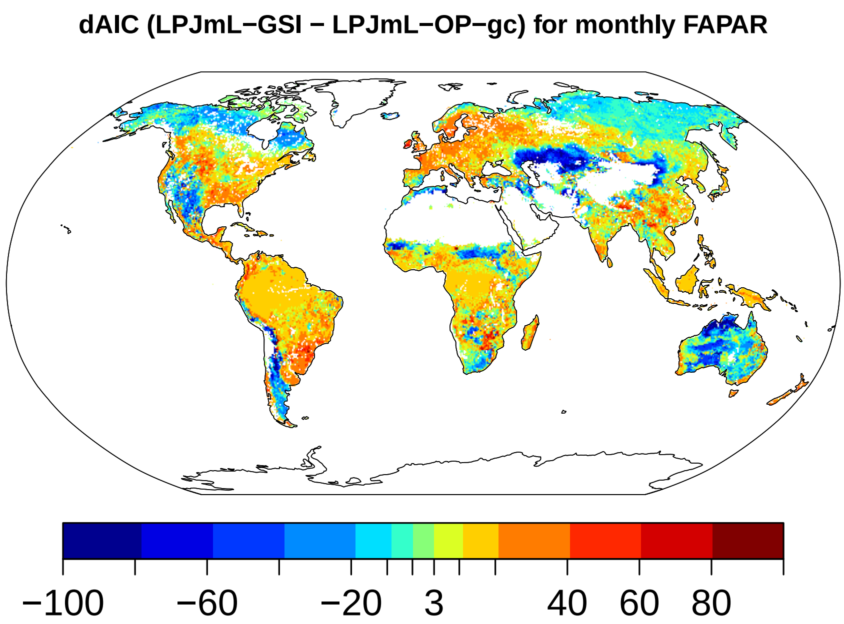

Fig. 2. AIC differences defined as LPJmL-GSI minus LPJmL-OP based on the cost for FAPAR. (see extended figure caption at the end of our reponse).

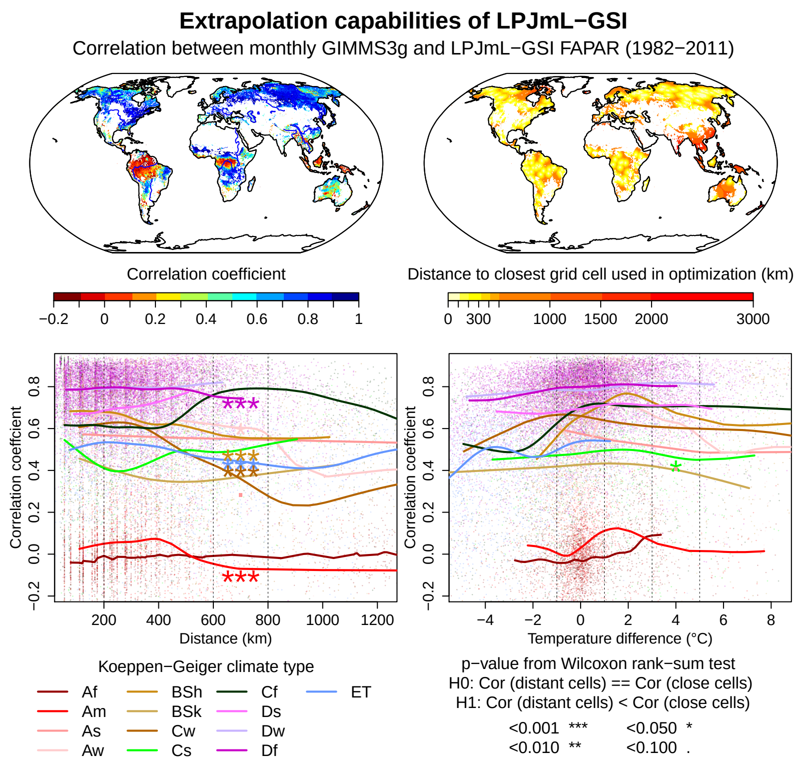

Fig. 3. Extrapolation capabilities of LPJmL-GSI in terms of monthly FAPAR dynamics. (see extended figure caption at the end of our reponse).