Biogeosciences
Discussions

Open Access

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

Anonymous Referee #1

Received and published: 31 July 2014

Forkel et al. present a detailed account of the testing of different phenology schemes within the LPJml model. They compare the original LPJ phenology model with a new version that incorporates a modified version of the Jolly et al. GSI model. Both versions are optimized to multiple constraints, and tested against global NDVI data.

The study reports two interesting findings: 1) the GSI version of the model is better than the original version, and 2) the optimized GSI model can be used to indicate the relative controls of phenology and their spatial variability.

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.

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.

Minor points

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

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

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

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

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.

Page 10932: "we used"

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

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.

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.

C4002