

General comments:

The paper uses NDVI-derived GPP to estimate environmental controls on the greening of terrestrial vegetation across northern Eurasia. Such a work involves two important aspects: 1) GPP modelling and validating (if GPP not directly available), 2) linked GPP with environmental variables (statistical analysis on trends, correlation, and etc.) These two aspects are both the current research hotspots, and therefore this paper will be interesting to many readers. However, the paper failed to describe relevant methods and models clearly. And some concepts are not clear in the paper either. I would suggest a substantial revision of the paper. General comments are as follows.

1. My major concern is how GPP data are modelled from GIMMS3g NDVI and VIP NDVI datasets. The authors did not say clearly in the paper.

From P9126 L13 “GPP is estimated in a manner similar to ...”, I assume the satellite derived GPP data have already been produced by other researchers and were used in this paper (as far as I know, there no such GIMMS3g derived GPP available now). But from L19 of the same page “In order to estimate FPAR ..., NDVI was temporally interpolated...”, I assume the authors did the satellite GPP modelling work here in the paper. The changed verb tense confused me where the NDVI-derived GPP are from.

Suggest a careful check of verb tense in the paper to make clearly what is the work done in this paper and what is used directly here.

2. LUE model is the most important part in this paper, and all the consequent trend, correlation, and attribution analyses are based on how the LUE model performs. However the authors failed to give detailed model equation, parameterization, and proper and convincing validation presentations. I cannot judge if the LUE GPP models for two NDVI datasets, and parameters are proper and sufficient to draw the consequent conclusions about trends and correlations in the paper.
3. Section 2.1.2 is not clear. If the long term NDVI derived GPP data are collected from somewhere and used here, you probably need to describe the data sources, references, authors and downloading websites clearly. None of your references in this section can be linked to any available long-term GPP used in your paper. Then I guess you did the LUE modelling exercise here. If so, this section should be rewrite and moved to Section 2.2 Methods
4. Data used for GPP modelling with LUE algorithm. What data were used for the model from ERA-Interim dataset? How were the differences in spatial resolution handled? Where and how did the authors get PAR data?
5. The model results (parameterization) should be presented in this paper, better before the validation section. Also, if you used the GPP data of 10 flux site for parameterizing the LUE models, be careful if you used them again for validation.
6. Section 3.1 the discussion about the validation is in sufficient. The authors showed the GPP model based on LUE algorithm is better in spring and worse in autumn, but they did not discuss why. A Nash–Sutcliffe Efficiency coefficient above zero only shows the model is better than using observation mean, it does not show the model is accurate

enough to draw some convincing conclusions. The last paragraph in Section 3.1 relates this paper results to other researches, but these researches seem irrelevant to the GPP validation of this section. I do not understand what the author meant “LUE algorithms, similar to the one used for the generation of the GIMMS3g dataset”. The generation of GIMMS3g (NDVI) dataset never need an LUE algorithm. There were tons of researches about GPP models using LUE concept, with satellite data as input, but the authors neglected them and did not mention in the validation discussion. Again, it is no meaning to me to talk about validation without showing some scatter plots about modelled GPP vs. tower GPP.

7. Ensemble mean

In statistics, there is a clear definition about “ensemble average”. I don’t think it is proper to call “ensemble mean” of two GPP values modelled from GIMMS3g NDVI and VIP NDVI. Maybe I am wrong, but I would expect an explanation from the authors.

8. Captions of tables and figures are too long and are very distractive. I always get lost and don’t what the authors want to show in a table or figure. Suggest a concise caption and move explanations to relevant sections.

9. I am not sure if +ve and -ve are allowed in a formal publication. For a non-native English speaker like me, it took me a while to find what they mean (here is the first time I read them in a publication).

10. De-trended correlation analysis.

If I understand correctly, you were using annual or seasonal data in your correlation analysis. That means, for example, you have 27 spring GPP data and 27 spring temperature data for a pixel. In this way the data have been de-seasoned already, how can you further do a de-trended processing before doing correlation analysis? You will only have noise left after de-trending and de-seasoning.

11. In correlation analyses, how are different spatial resolution handled?

Detailed comments:

P9126L14 mentioned twice “similar to the MODIS MOD17”.

P9126L19-20 I don’t understand why the authors needed daily NDVI, so that the biome-based NDVI-FPAR relationship can be used. This sounds illogical to me.

P9127L14-15: If you call those “derived using alternative GIMMS3g and VIP NDVI inputs using a LUE model” observation-based GPP data, what are model-based GPP data?

P9130L14. I don’t understand the sentence “Strong trends in the time series examined introduces the issue of collinearity”. Do you mean “trends introduce ...” or “time series introduces ...”. I am lost in your grammar and I cannot see the logic here either.

P9130L26. “closer” than what?

P9132L5, “normalized difference vegetation index” for NDVI has been given already. No need to repeat here.

P9132L16, What is “above ground GPP”? Do you mean there is also “underground GPP”? Please be sure about ecology concepts? Again you might need to check P 9139 L10 “GPP refers only to above ground carbon exchange...”. Such a saying about GPP sounds not correct to me.

P9132L25, If the authors say the uncertainty of the mean GPP curve is caused by the difference of two GPP values (from GIMMS and VIP respectively), how do they compute the std from TWO value? Which pixel is used to calculate the mean and uncertainty from two GPP datasets? If the authors pooled all the spatial GPP data together to calculate the mean and std for a certain year, such an uncertainty is not caused by the difference of two datasets, instead it is spatial uncertainty. Further clarification is expected.

P9133L11 An increase of $34.6 \text{ g C m}^{-2} \text{ yr}^{-1}$ from 1982 to 1998 (17year) is lower than $2.4 \text{ gCm}^{-2} \text{ month}^{-1} 10\text{yr}^{-1}$ of this paper result. The authors’ claim of “higher than our estimate” is not correct.

P9133L14-15 I did not see the logic between “the higher GPP trend in summer” and “the vegetation is predominantly cold constrained” (intra-annual). Trend is an inter-annual variation, and higher summer growth activity is an intra-annual concept, no link with the trend.

P9133L22 to the end of the paragraph, the authors discussed the differences between GIMMS NDVI and VIP NDVI and claimed VIP NDVI dataset is better. However such a discussion is irrelevant to the authors’ results in the paper. The authors showed in Table 2 that GPP modelled from VIP is worse than from GIMMS in general, contrary to the authors claim based on other studies.

P9134L8 “More than half of the region is affected by a significant positive trend (Fig. 4a).” I cannot see “more than half of the region” from Fig. 4a.

P9134L22, Suggest cite AR5 and give a proper reference.

P9134L24, also need to indicate the time span for the trend from IPCC.

P9135L3, the period of 1982-2008 spans 27 years. Not 26. That of 1997-2008 is 12 years, correct!

P9137L5 “The cause... are...” verb form might be wrong.

P9137L6-7. The positive relationship between precipitation and clouds seems contradicts Table 3, which shows more positive precipitation trends but less positive cloud trends. Why so?

P9151 Table 1 “whose GPP data has been used”, notice plural verb form

P9152 Table 2. The table presentation is less informative than a scatter plot for model validation purpose. The authors provided a correlation table to show the validation of their GPP model from each of the two different NDVI datasets and their mean GPP from the both datasets. I cannot see how many samples are used in the correlation analysis and their spread, therefore I cannot judge the LUE model for GPP is good or bad (see general comments).

P9159, Fig 4 a, c, and d (except b for precipitation), for all the y –axes, I don't understand month⁻¹ in the trend unit. Where is it from, and how are the trends of environmental variables calculated?