Biogeosciences Discuss.,
doi:10.5194/bg-2016-48-AC3, 2016

Interactive comment on "Dryland vegetation functional response to altered rainfall amounts and variability derived from satellite time series data" *by* Gregor Ratzmann et al.

Gregor Ratzmann et al.

gregor.ratzmann@fu-berlin.de

Received and published: 26 May 2016

Dedicated Responses to Anonymous Referee #3

We would like to thank Anonymous Referee #3 for providing a constructive and very helpful review of the manuscript. We believe that the comments helped improving the overall quality of the manuscript. Please find following dedicated responses to each point raised by the referee.

Anonymous Referee #3 (AR #3): The analytical approach involves: 1). shifting time series analyses run per grid cell with a proxy of vegetation productivity as the dependent variable, and rainfall as the independent variable, ...

Response: We would like to stress at this point that the study does not involve time series analysis. We believe that clarifying this point is crucial as also Anonymous Referee #1 raised concerns about possible effects of temporal autocorrelation. However, temporal autocorrelation by definition is a phenomenon which is limited to time series analyses (relying on parametric methods). Thus, we would like to underline here (and throughout this response) that temporal autocorrelation does not affect our analyses at any point (see also the specific response to this point raised by Anonymous Referee #3).

AR #3: However I perceive some problems with the OLS analysis as outlined below, which I think should be addressed as a priority, since this may affect the conclusions of the paper. In addition, the conclusions of the paper should make further consideration of the anthropogenic factors in each of the study sites, at least in terms of better explaining how the patterns observed (if the modelling is robust to the potential autocorrelation problems) might also be moderate by human behaviours, particularly in areas with many crops.

Response: Please see our specific responses to the respective referee's comments on autocorrelation as mentioned above. We agree that a short coming in the discussion is the potential effect of land use. We have added a dedicated section in the discussion considering those effects and how they might affect results observed. Please see the specific response to this point at the respective referee's comment.

AR #3: Overall with respect to the writing, whilst the aims of the paper are quite clear, the readability of the paper is hampered by a) unnecessarily convoluted and confusing language and sentence constructions b) some undefined terms and c) the use of different terms to describe the same parameter. This unfortunately detracts from the science undertaken.

Response: We greatly appreciate this observation and the examples given below. Confusing and complicated language can indeed largely hamper effective communication.

We have completely revised the manuscript paying particular attention to possibly misleading and complicated language. Moreover, we have revised the use of acronyms and fixed terms ensuring that they are used consistently throughout the manuscript.

AR #3: 1. The overall variable of interest, beta. Initially this is defined as vegetation response to rainfall, in other places it is described as the 'beta response'. This must be standard- ised throughout the paper, e.g. with the use of a subscript.

Response: We understand that – as β itself is defined as response to rainfall – reading "β response" can be confusing. This statement in each instance it appears (3 in total) refers to a response function of β to another variable (e.g. MAP). Thus, we decided to replace the term "β response" by "β response function".

AR #3: 2. Abstract L26: ' we conclude that higher. . ..' This sentence is confusing. 'Rainfall plasticity' doesn't really make much sense, especially when the paper later on contains precise terms concerning measures of precipitation. As such I think some alternative terms would be better in this paper overall, and particularly the abstract to help the reader.

Response: Thank you for pointing this out. As already mentioned in the responses to Anonymous Referee #2 (who pointed to this sentence as well) plasticity does not refer to rainfall but to vegetation response to rainfall. However, as this seems to introduce major ambiguities we decided to now include the acronym β already in the abstract and consequently replacing vegetation response to rainfall plasticity by β plasticity.

AR #3: 3. Abstract L23: 'interannual rainfall amount variability' – vs L31 'rainfall variability'. Then on Page 3. L.21, 'absolute rainfall amounts' used. So we have a series of different terms i.e. 'rainfall', 'rainfall amount', 'precipitation' and 'absolute rainfall amount' which I think are all describing the same physical parameter. Better to choose one precise term such as 'total precipitation (mm)' and be consistent throughout, modifying it as necessary e.g. coefficient of variation of precipitation.

Response: This indeed might lead to confusion. We have now consistently named any reference to annual rainfall amounts "annual rainfall". Thus "interannual variability of rainfall amounts" now reads "interannual variability in annual rainfall". However, in certain sentences (such as in the Abstract page 1, line 12) we deem it necessary to mention the term "rainfall amount" to make a clear distinction from variability.

AR #3: In another case, on page 7, L8, you have 'some beta sensitivity to W for absolute values'. However THE parameter you are discussing the absolute values of is not stated.

Response: Thank you for pointing out this ambiguous sentence. It is now rewritten.

AR #3: 4. Page 4. L27. W (window) and WA (west Africa). It would make the reader's life easier to differentiate these two abbreviations. For the Window parameter, the units (years) should also be added (e.g. W7years).

Response: Indeed, the two acronyms read confusingly similar. We have replaced the window length acronym W with L. Regarding the units of W: We have added this information where missing.

AR #3: 5. Page 3 L38. GIMMS is not defined before being introduced in the text.

Response: Thank you for pointing this out, we have added the missing information.

AR #3: 6. 'Sub-pixel land cover frequency' : I think there is a sub-pixel land cover distribution as a result of the resampling procedure. Is this correct?

Response: The sub-pixel land cover frequency is a result of the MODIS land cover classification procedure for MCD12C1 and provided along with the products. It reports the relative frequency of all present land cover classes within one pixel with the most frequent one being assigned to the pixel.

AR #3: 7. Other points on language that should be addressed involve more careful checking of the text e.g.: Abstract L17. 'as explanatory variable' –change to as an/the explanatory variable.

Response: Thank you, this has been taken care of. We have revised the manuscript (as indicated earlier) to improve readability and language.

AR #3: 8. ‘Hydroclimate period’ – probably easier to use this term once and thereafter say ‘wet’ and ‘dry’ seasons. Keep the language as simple as possible, allowing the reader to focus on content.

Response: We agree that the term hydroclimatic period is somewhat unhandy. We have exchanged it now by “wet” and “dry ” where applicable.

AR #3: Page 4, L9. A month is wet season if >20mm precipitation. Is this a recognised threshold in the literature ? Please cite a reference. This is an important threshold and analytical step because on L34 the data is partitioned into binary classes of wet and dry seasons- changing the threshold will therefore affect the partitioning.

Response: We have added some further explanations on the derivation of this threshold. However, we note that Anonymous Referee #3 is further referring to a procedure (line 34) that is not affected by this threshold. The rainfall threshold (20 mm) determines whether a given month belongs to the rainy season or not whereas the dry/wet criterion indicates whether a period over which a β coefficient is derived has below or above average (MAP) rainfall.

AR #3: Page 4. L20. With respect to the analyses conducted, the principle tool used is ordinary least squares regression. However, given that the regression analyses are conducted over time and space, the analyst should immediately flag the risks of temporal and spatial autocorrelation. If present, such autocorrelation will violate model assumptions of error independence, and hence may cause problems in the interpretation of the results. Apologies if I have missed this somewhere in the SI, but I do not see any noting of either of the autocorrelative problems being acknowledged. If it is the case, it would be a significant omission in the consideration of the analysis, and I think

is the âĂŤ-major analytical issueâĂŤ to be addressed following review. If error correlation over space and time ultimately do not represent an analytical challenge, then the analysis leading to this conclusion should be included (e.g. by presenting the results of a Moran’s I analysis).

Response: We believe that this point raised by Anonymous Referee #3 includes two potential issues: i) temporal autocorrelation and ii) spatial autocorrelation. Following, we will address each point separately. i) We agree that temporal autocorrelation is an important concern in time series analysis conducted using parametric methods. However, the present study does not do time series analysis. We compute temporally shifting linear models using OLS techniques, hence we are using parametric methods. Those models, however, use annual rainfall as independent variable and growing season vegetation productivity proxies as dependent one. Thus, neither at the stage of computing those models nor at a later stage time is involved (as variable being used in modelling) in the methodological process of this study. Thus we conclude that temporal autocorrelation is not of concern at any of the analytical steps involved. ii) Spatial autocorrelation is an important issue for analysis relying on gridded data especially when larger objects or homogenous areas are comprised of several pixels. However, there are two reasons why spatial autocorrelation can be assumed not to be an issue in the present study. Firstly, the spatial resolution of the used NDVI data (approximately 8 km) makes it rather unlikely that several pixels comprise a larger body of structures or processes showing typical spatial autocorrelative attributes. Secondly, before performing the analysis we average all β values over 1 mm MAP steps, which removes any spatial information possibly leading to autocorrelation. Thus, spatial autocorrelation can be expected to be neither an issue.

AR #3: Page 4. L35. Authors bin the beta values – was this using a mean function?

Response: Thank you for pointing to this. We indeed averaged over 1 mm steps, this is now mentioned where missing.

AR #3: Page 7. L6. On a separate point, in the discussion the text states: "higher GAM R2 scores in SWA indicate an overall stronger effect of MAP on shaping beta compared to WA". Sensu strictu statistically: the coefficient of determination tells you how much of the variation in the dependent variable is explained by the independent variable; whereas, the effect size is the magnitude of the coefficient on MAP.

Response: Thank you for pointing to this shortcoming. We have adjusted this sentence by removing the reference to effect size.

AR #3: Page 7. L15. The W parameter: the purpose of the inclusion of the different W sizes should be better explained, especially given the authors' conclusion that effects of W tell you about the statistical impact of averaging over different time spans, and losing differences between wet and dry periods, rather than any ecological significance. To reveal this as being a statistical artefact in the discussion seemed to undermine the inclusion of this aspect of the analysis. A more positive way to describe this result would probably be that it highlights the importance of partitioning the analysis of responses into dry and wet-season responses.

Response: We agree that there is some ambiguity in including this parameter in the analysis. However, since the methodological approach as presented here is novel we deemed it necessary to report all parameters which have to be specified before the analysis (such as W) and their effects on results. Thus, besides the ecological information contained within this study we perform an initial application of shifting linear regression models and report on the effect one of the required input parameters has on the study outcome (which we believe to be rather small over the range of Ws considered here).

AR #3: Page 7. Line 18. The authors mention here local variations in land use. This is an important factor in explaining vegetation patterns across the globe i.e. anthropogenic disturbance. It should at least be acknowledged that there may also be differing disturbance regimes in the two sites, which may be dependent

upon human density and pre- dominant modes of agricultural production and management. For instance high human population density combined with high levels of fuel-wood extraction seasonal burning may restrict the growth of perennials and development of grassland into savannah in WA whereas such anthropogenic constraints are fewer in SWA. CIESIN has gridded population data you could check: http://sedac.ciesin.columbia.edu/data/collection/gpw- v3.

Response: Thank you for pointing this out. We agree that differences in land use may lead to local deviations from the response functions along MAP gradients (as noted in the manuscript). As this point is admittedly rather short in the manuscript thus far, we added some information on the specific land use practices in both regions (for MAP > approx. 400-500 mm (semi-) nomadic livestock keeping in WA and farm-based livestock keeping in SWA and a mixture of crop farming and mainly communal livestock keeping in both regions above those values). We moreover note how this could affect results. Indeed population density might to a certain degree affect the results. We have added a sentence discussing this possible effect.

AR #3: Understanding this component of the work is essential to the reader since the derived cyclical fraction constitutes the proxy for vegetation productivity. The concept of measuring values as the integral of vegetation values above a baseline of productivity is straightforward. However, the text in the SI on the details of the work undertaken is quite confusing: "To determine the onset and the end of the CFR of any given year, a baseline is derived, which constitutes the mean upper limit of the dry (or cold) season values between two vegetation peaks. Values above this baseline are part of the CFR. The baseline is calculated using the amplitude between the mean of the four lowest values ("low level mean") between two peaks and the average of these peaks" (SI pages 4-5). Perhaps a diagram as provided in figure s3 would help the reader here.

Moreover, given the central importance of this step in establishing the dependent variable upon which the analysis depends, I would like to see some more justification of the approach used, and its appropriateness in this instance. I appreciate this is difficult given that the main citation is an article in press. I wonder whether it is possible to get an author's draft to circulate amongst reviewers?

Response: To improve readability of the SI on the derivation of productivity proxies we have simplified and shortened this section. Although it might provide less details on the procedure now, we believe it will help the reader getting the idea behind the method. Moreover, we have included a schematic depicting the constituents of a vegetation index time series leading to the phenologically-derived proxies. For a more detailed description of the procedure we kindly refer to the document describing the phenological parametrization model which is now available on ResearchGate (not in press anymore) http://bit.ly/1UfqE3v (we had to shorten the link since it did not fit into the PDF).

AR #3: For instance, given that the central question of the paper is examining responses to rainfall variability, are the authors not concerned that the linear interpolation of outliers is removing some real variability in the vegetation responses? That is, removal of outliers may be employed as a statistical sub-procedure to remove bias from parameter estimates caused by errors in data collection or data entry by researchers. However, such outlying data points are often real measurements that should be included in analyses. What is the basis for interpolation in this case?

Response: We agree that, in general, removing outliers should be a matter of strong consideration before applying such procedures. However, outliers resulting from sub-optimal measuring conditions is a particular feature of time series of vegetation index data derived from earth observation. Clouds, e.g., frequently impair the quality of any surface reflectance measured at the satellite platform (and clouds are rather common features during rainy seasons and consequently growing seasons). This challenge is partly overcome by using the NDVI as this index is less strongly susceptible to atmospheric (and cloud) effects. Moreover, the product used (GIMMS3g NDVI) accounts for potential atmospheric effects by using a biweekly maximum value composite (as particularly water has a dampening effect on NDVI values). Nevertheless, particularly

low outliers of the final composite product during a growing season can be expected to be rather artefacts than representing real variability. This is why most phenological parametrization models use some kind of outlier removal a priori using, e.g., fitted splines or interpolation. Thus we deem it not only justifiable to use an outlier removal but consider this step as required to ensure data quality. Regarding the effect outlier removal might have on overall variability we are confident that this effect is negligible. Firstly, given the above remarks, outlier removal should rather enhance the estimation of interannual vegetation productivity variability. Secondly, should a given outlier be removed under the false assumption of noise the overall effect on the estimation of a productivity proxy the corresponding year should be negligible. Thus, overall, outlier removal can be expected to improve the estimation of the true interannual variability rather than deteriorating it.

AR #3: Page.2. L18. 'arid-most parts': define with respect to rainfall as is done for the semi- arid regions on the following lines.

Response: This has been changed.

AR #3: Page 3. l23. 'characterised by high inter-annual length of the wet-season variability' : re-order sentence

Response: This sentence has been reordered.

AR #3: Supplementary information Figure S6: 'shidting linear. . ..' Spelling. Error also in S7.

Response: Thank you, the errors have been corrected.

Interactive comment on Biogeosciences Discuss., doi:10.5194/bg-2016-48, 2016.