Reply to: *Interactive comment on "Modelling spatial and temporal vegetation variability with the Climate Constrained Vegetation Index: evidence of CO2 fertilisation and of water stress in continental interiors" by Anonymous Referee #2*

Comment 1: *The author proposes an empirical model on NDVI based on temperature and precipitation. I am not convinced the method presented in the manuscript is of high interest. My main concern is the very poor skills of the model. The model does an OK job simulating the broad spatial distribution of NDVI and its climatological seasonal cycle, but it fails to simulate interannual variability (figure 6 showing a poor fit between the model, CCVI and the observation, FASIR). The author shows that CCVI seems better than an earlier model, RVI, at least for interannual variability. Nevertheless, CCVI has a very poor correlation (r=0.255). I was quite surprising to read in the introduction "the RVI reproduces spatial, seasonal and interannual variability well (Los, 2013)", when Figure 9 gives a coefficient of correlation of only 0.188. This is also clear from figures 8 and 9: both models (RVI and CCVI) do capture the mean seasonal cycle of NDVI, hence a good temporal correlation on figure 8, but completely fail to simulates anomalies from year to year.*

Response:
– leaf seasonal cycle / spatial variability: The present study was carried out to address an important shortcoming in current land-surface models; the poor simulation of leaf seasonality (Randerson et al. 2009, Richardson et al 2012, Los et al 2013, Kelley et al 2013). The CCVI model simulates leaf variability as seen in the NDVI with greater skill than current models evidenced by higher spatial correlations and higher temporal (seasonal) correlations, and simulating correct timing of maximum leaf development (current land-surface models are criticised for not doing so). I would therefore like to take issue with the qualification "OK job", the CCVI model clearly provides more realistic simulations of leaf seasonality and leaf spatial distribution than current models.

- Fig 6 a poor fit: The correlation between mean annual global FASIR NDVI and CCVI with CO2 (Fig 6) is 0.77 and between FASIR NDVI and CCVI control is 0.72. I disagree that these qualify as poor fits.

- Interannual variability: One has to bear in mind that the interannual variability in NDVI is a small signal and in many cases a close to zero signal with a relatively large error. For example, in mid-to-high latitudes a typical seasonal range in NDVI is about 0.6, a large interannual signal has deviations rarely larger than 0.1. A representative error in NDVI is between RMSE=0.02-0.05 (0.01 residual calibration error, 0.01 residual cloud contamination, 0.01-0.02 residual atmospheric effects; 0.02 – 0.05 residual directional effects), there are also errors in the precipitation and temperature data which will add to the overall uncertainty. Thus for an area with a large interannual signal around 0.1 and an error around 0.05 we expect, in the absence of other factors, slightly over half of the variance explained by climate (r around 0.7-0.8). Various parts of the globe have a very small interannual signal (tropical forests, deserts) or have a very small interannual signal during a large part of the year (winter in mid to high latitudes, dry season in arid and semi-arid regions) or show small deviations for years close to the average seasonal cycle, which should be the rule rather than the exception. We expect a low correlation where the variance in both model and observations is low. Thus there are good reasons to expect low correlations between observed and modelled interannual variability for parts of the globe. This is reflected in the spatial distribution of the coefficients of correlation; values are near zero (positive or negative) in deserts and tropical forests (which they should be) and are around 0.5—0.7 in areas with large interannual variability, which is slightly below the best case scenario. There are also areas with high correlations during 1982-1999 that are low in 2000-2006 and vice versa which indicates that large deviations, associated with higher correlations, occur during different times. The median correlation coefficients mentioned by the reviewer were reported to compare the RVI and CCVI using one number, they are not a reflection of the overall performance of the model. To summarise: large deviations (which are the ones that matter most) tend to be captured by the CCVI since higher correlations occur in areas with higher interannual variability. The CCVI was derived from 1982-1999 data and was tested on satellite data from 2000-2006; about 50 years of phenology data observed since the 1950s and was qualitatively evaluated for the response to known extremes during the 20^{th} century (dust bowl in the US in the 1930s and 1950s; 1984 drought in the Sahel and big dry and wet periods in Australia – which will be added to the analysis). In all cases simulations appeared realistic.

In addition I would like to make the observation that tests of leaf interannual variability in other models are usually not reported; I found one exception, the ISBA model which was calibrated to match the NDVI data and which showed similar interannual correlations with NDVI as the CCVI. In another case I could obtain model output –- i.e. for the CENTURY model which did not show any depression of leaf seasonality during the big droughts in the US in the 1930s and 1950s (as mentioned these were reproduced by the CCVI).

A fundamental premise of science is that preference is given to a model that performs better than another (or to one that is simpler but performs equally well). In this study I developed a model that results in better spatial and seasonal simulations of leaf variability than current models. I suspect it also provides better simulations of interannual variability than other models but as mentioned I have only limited means to test this because this essential information appears lacking. I disagree with the reviewer that the interannual performance of the CCVI model is poor; it provides a reasonable estimate of what is a small signal within the ranges of uncertainty. I see the lack of testing of interannual variability of leaf seasonality in other models as a weakness, not a strength and I find it curious that the reviewer seems to prefer models which interannual variability is not tested over one that is.

Comment 2: Also, as already pointed out by the first reviewer, Colin Prentice, it seems illogical to develop a NDVI model that is not based on PAR, which, with leaf temperature and water availability, is the key control on leaf growth, and hence diagnosed NDVI. One can of course always train a model using observations over a given period with temperature and precipitation only, but I have serious doubts on the predictive skill of such model, lacking physiological foundations.

Response: The logic behind the development of the CCVI model is that foliage will develop up till the constraints set by the environment (precipitation and temperature). Radiation was initially not considered for the CCVI because it did not explain much of the variability in residuals between modelled RVI and measured NDVI in a previous study (Los 2013). After comments made by Colin Prentice I analysed the importance of radiation constraints but did not find evidence that these affect leaf seasonality for current or near future climatic conditions (analysis added to revised paper). In other words there appears to be a sufficient supply of radiation for most of the globe not to limit leaf growth whereas temperature is limiting in mid-to-high latitudes and precipitation in low latitudes. Furthermore use of radiation as a limiting factor would in many cases lead to the wrong leaf seasonality, for example radiation is at a maximum in June in temperate northern latitudes, whereas NDVI and temperature tend to peak in August. There are other examples that demonstrate that temperature rather than radiation is a limiting factor in temperate regions, e.g. the gradual trend towards earlier leaf out in spring observed over the past 5 decades in these latitudes has been attributed to warmer temperatures, not to changes in solar radiation.

Had I found evidence that radiation is an important constraint on leaf seasonality, I would have included its effect in the model. I did not find evidence and therefor did not include it. To reiterate, the aim of the present study is to simulate leaf seasonality, the model is not a photosynthesis model or NPP model which would require incorporation of radiation. Photosynthesis or GPP/NPP can be modelled when the CCVI is implemented in a model that simulates these.

There is in my opinion a clear problem with current land-surface models. They can reproduce the daily photosynthesis / GPP cycle and therefore appear to capture processes important at sub-daily to daily time steps. However, their simulation of the seasonal leaf cycle is poor which indicates a problem scaling from the sub-daily time step to the seasonal cycle. We don't know how they perform in terms of interannual variability, but, based on their poor performance in simulating the seasonal cycle, I doubt they reproduce interannual variability well and have reservations as to their suitability simulating the impacts of climate change on the biosphere.

Comment 3: As a final note, I see GMD as a journal where new models are being described comprehensively, models equations being fully disclosed in such a way that a reader could, in theory, reproduce the same model. I think this manuscript is a long way from this standard. Section 3 is extremely hard to follow, and at the end, I couldn’t picture what the equations supporting the model are.

Response: I will expand and improve the explanation of the model.