We thank the reviewers for their comments. Our response to each comment is given in *italics*, and additions or changes to the text are given in ***bold italics***. In general, where a comment is repeated (e.g. as a main point and as a detailed comment) we only provide one response but indicate that this has been dealt with elsewhere.

Comments from F. Hartig (Referee):

General points

1) FOCUS OF THIS STUDY

The presentation of the ms is in general very clear, but I was missing the "big picture". What was the objective for doing this analysis? I can't find anything about this in abstract / info (apart from the information about what you do). It seems to me that the most interesting result of this study is the CO2-dependent allocation, but this message is absent in the title, and in the analysis you afford comparatively little space to it. Instead, you analyze in detail the model fit under constant CO2, i.e. under conditions that never existed in reality. Of course you can do this analysis, but you should explain why. What can we learn from evaluating the model with wrong drivers (i.e. constant CO2)?

The reason why we begin with modeling tree growth with fixed CO_2 is because it has been repeatedly claimed that CO_2 has no impact on tree radial growth and we cite the papers about this in our introduction. Thus, we begin with an analysis that assumes that this claim is true. The title points out that we are interested in both climatic and CO_2 effects on tree growth; we discuss the impact of CO_2 on growth and allocation explicitly in the abstract. We agree that we don't mention allocation in the title, and we propose to change the title to "A model analysis of climate and CO_2 controls on tree growth and carbon allocation in a semi-arid woodland". We agree that it would be good to put this aim explicitly at the beginning of the abstract and we have modified the abstract accordingly as follows:

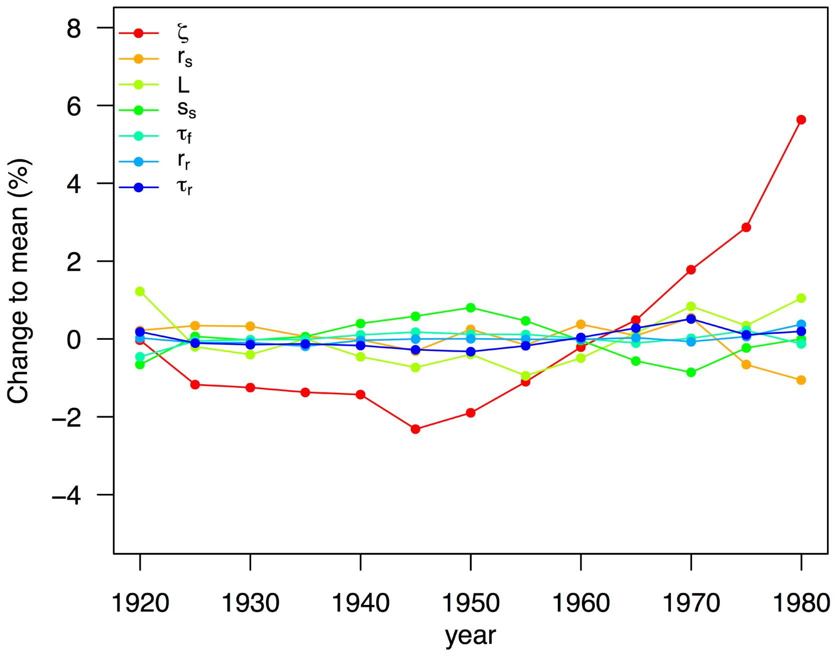
The absence of a signal of increasing [CO_2] has been noted in many tree-ring records, despite the enhancement of photosynthetic rates and water-use efficiency resulting from increasing [CO_2]. Detection of a [CO_2] effect should be easiest in semi-arid climates. To examine this issue, we used a light-use efficiency model of photosynthesis coupled with a dynamic carbon allocation and tree-growth model to simulate annual growth of the gymnosperm Callitris columellaris in the semi-arid Great Western Woodlands, Western Australia, over the past 100 years.

2) HOW CERTAIN ARE WE THAT ALLOCATION SHIFTS AND WHY

How strong is the evidence for time-dependent allocation to fine roots from your results? A big concern here is that the **time-variable parameters are simply compensating some other structural error of the model** (i.e. that the model is not able to create constant growth with increasing CO2). **As very few parameters are under calibration, it may well be that there was no other option to get the predictions close to the data than increasing allocation to roots**. The most convincing thing for me would be to put more parameters under calibration and show that you couldn't get a proper fit by changing other parameters (within reasonable bounds for these parameters of course).

We made the decision to calibrate only those parameters for which we did not have good estimates from the literature, on the grounds that one should use observations when there are *observations. However, we have now re-run the calibration including all the T model parameter,: this results in a slightly smaller LAI (1.48) and other parameter values change by less than a 1%*

The time-dependent calibration including all the T model parameters shows the same pattern as before – only the ratio of fine roots to foliage area changes notably with varying CO_2. We have not made any change to the main text, but we include a new version of the figure here.



Moreover, I didn't understand why you didn't validate the **time dependent model** in the same way as the other model versions (Figs. 5,7).

We agree that it would be useful to show the regression plots for the time dependent model, and we have now added these as a new Figure Fig 10

We have also added a sentence describing these results at the end of Section 3.2 as follows:

The regression analysis shows that the impact of the CO_2 effect on simulated tree growth is very much weakened (Fig. 10) compared to simulations in which ζ is not allowed to vary (Fig. 7).

…. can we find out which driver is responsible for the change? Time can't be the reason, so it should be CO2, some other climatic variable, or something else. The references to literature in the discussion are good, but can you do anything more here? For example, you could try to make the dependence to CO2 or some other variable explicit in the model through a parameter and optimize this, or analyze how the inferred time-dependent parameters correlate to CO2 and other environmental variables (in a way, the latter is in Fig.8, but only for CO2. And here I was wondering if you want to suggest with this figure that CO2 is causal - if you look at the decrease of fine root allocation in the 20s - 50? with increasing CO2, is it sensible that he roots go down and the up again?)

The second tuning is done with time-varying CO_2 and time-varying climate. We realize that the caption to Figure 8 is somewhat confusing, We have changed this to read "Time variation of the values of parameters ……"

We suggested that the time-variation in root allocation largely reflected the influence of CO_2 because of the strong upward trend in CO_2 (Figure 3). But, the decrease in the fitted value in root allocation between 1920-1950 corresponds to a temporary increase in cloudiness (which can be seen in Figure 3 as a decrease and an increase in soil moisture α). It is a good idea to try and disentangle the impact of CO_2 and climate, and we have now done this using variance partitioning on the fitted values of ζ. This shows that CO_2 can explain ca 80% of the variance and changes in soil moisture (α) ca 20% of this variance. We have added text to the end of section 3.2 to describe this result as follows:

Variance partitioning (Chevan and Sutherland, 1991; analyses made using the hier.part package in r) shows that ca 80% of the variability in ζ can be explained as a consequence of changes in CO_2 and ca 20% as a consequence of changes in soil moisture availability, as indexed by α.

*We have also modified Figure 8 by adding in the 30-year running mean of **α**.*

3) TECHNICAL ISSUES

Can we exclude problems with the allometric relationships and multiple stems?

Please see response to this point under detailed comments below.

The description of the model assumptions and parameters is quite short. It's clear that you can't describe everything in detail, but things that could be an issue for the main conclusions should be better explained. In particular, could changes in other parameters that were not under calibration cause a similar reaction as changes in the fine root allocation parameter? As there are uncertainties also on other model parameters, why didn't you put all parameters under calibration? If you have estimates from the literature, you could code this as a prior in a Bayesian analysis.

Both the P model and the T model have been published before, and these publications (which we refer to in our description) give greater detail of the parameters. As described above, we have re-run the calibration for all the T model variables and this does not change the results.

The calibration method is not described at all. What was the calibration objective, sum of squares? How did you assess convergence? Why use this somewhat exotic method instead of a standard optimizer?

We agree that the description of the calibration method in section 2.4 is very brief. The calibration objective was to minimise the absolute difference between modelled and observed mean tree-ring width during the period 1950-2012. The criterion for convergence was a difference of no more than ±2.5% of the mean value. We used neural network regression to determine the final optimal parameter values. We use this approach because it is suited to optimising across multiple parameters and is applicable to models that are not analytically solvable. We have expanded our description to read:

We used neural-network Bayesian parameter optimization (Jaakkola and Jordan, 2000; Pelikan, 2005), where the calibration objective was to minimise the absolute difference between modelled and observed mean tree-ring width during the period 1950-2012 and the

criterion for convergence was a difference of no more than ±2.5% of the mean value., to derive mutually consistent values of these four parameters.

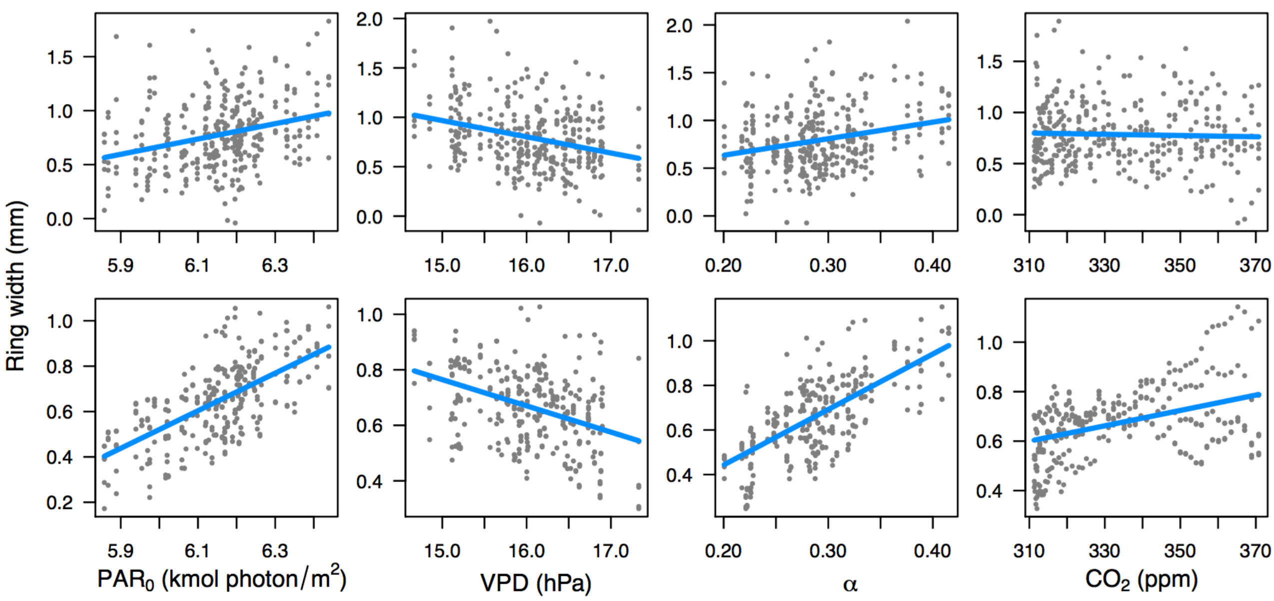
I liked the analysis of the climate-reaction of the model and the data, but I had some questions 1) Are you really using a GLM? It seems this is a LM problem.

We are in fact using an ordinary least squares multiple linear regression here, which is a special case of a GLM and we used a GLM package to derive this because of the additional flexibility that this provides for e.g. testing factors (although these are not used here). To avoid confusion we have replaced the term GLM by regression throughout the manuscript, and have replaced the initial description in section 2.6 as follows:

We investigated the optimal interval influencing carbon accumulation and tree growth using ordinary least-squares multiple linear regression.

It seems you don't account for a) temporal autocorrelation b) random effect structure due to the individuals. An analysis via a linear mixed model with a lag for temporal autocorrelation would seem more appropriate to me (use e.g. function lme in R)

As we have pointed out before in Biogeosciences Discussions (http://www.biogeosciences-discuss.net/11/C5928/2014/bgd-11-C5928-2014- supplement.pdf), it is unproductive to attempt to factor out temporal autocorrelation because a substantial part of this autocorrelation resides in the climate inputs. We have now run a linear mixed effects model, as suggested, treating the individual trees as random effects – the results are similar to not including this, see below:



3) As you seem to have strong collinearity, you should always have ALL climatic variables in the analysis. E.g. move temperature back into Fig.7. To account for a possible temporal trend caused by non-climatic factors, you could also consider taking time in as an additional predictor.

We agree that this is a good idea. We have there included temperature in Fig 7 (and also added it to the new Figure 10), and we have removed the top panels in Figure 5a as unnecessary (as suggested by the second reviewer). We have the modified the paragraph describing the results from Figure 5 to account for this changes as follows:

Regression analysis (Fig. 5, Table 3) showed that observed tree growth has a strongly positive, independent response to both PAR$_0$ and soil moisture stress (as measured by α).

and a negative response to VPD ($p < 0.1$). There is no response to MAT. These relationships can also be shown in the simulations. Although there is more scatter in the observations, the slopes of the observed and simulated response to PAR_0, α and VPD are similar in the model and the observation. The positive relationship with PAR_0 reflects the universal control of photosynthesis by light availability, and the positive relationship with α is consistent with observations that the growth of Callitris is determined by precipitation variability (Ash, 1983; Cullen and Grierson, 2009). VPD affects stomatal conductance such that increasing VPD leads to stomatal closure, with a correspondingly negative impact on photosynthesis and hence carbon assimilation.

We have also simplified the paragraph about this figure in the Discussion, as follows:

The radial growth of Callitris columellaris in the GWW is positively correlated with PAR_0 and α, and negatively correlated VPD. The response to VPD can be explained as a consequence of the atmospheric control on stomatal conductance and hence photosynthesis. Thus, both atmospheric and soil moisture deficits (the former represented by VPD, the latter by α) apparently exert independent controls on radial stem growth.

We have not added in time as a predictor because there is no way in which we could distinguish the effects of time from the effects of CO_2 which increases monotonically during the period. Furthermore we have no reason to expect a temporal trend in any factor other than those already included in the model (i.e. ontogeny) or in the drivers.

DETAILED COMMENTS (*note that the reviewer's page numbering is wrong, so we give both their page numbers and the actual page numbers to avoid confusion*)

Title: if the main story of this paper is that an increase in fine-root biomass explains a lack of CO2 effects on growth, why doesn't this message appear in the title?

This is the same point as raised above. We have modified the title to read "A model analysis of climate and CO_2 controls on tree growth and carbon allocation in a semi-arid woodland"

Page 4769/70, Line 1. a) start abstract with an introductory sentence b) what is the knowledge gap and the motivation for this study?

This is the same point as raised above. We have text at the beginning of the abstract to explain out motivation of the study, which now reads:

The absence of a signal of increasing [CO_2] has been noted in many tree-ring records, despite the enhancement of photosynthetic rates and water-use efficiency resulting from increasing [CO_2]. Detection of a CO2 effect should be easiest in semi-arid climates. To examine this issue, we used a light-use efficiency model of photosynthesis coupled with a dynamic carbon allocation and tree-growth model to simulate annual growth of the gymnosperm Callitris columellaris in the semi-arid Great Western Woodlands, Western Australia, over the past 100 years.

We have also modified the last part of the abstract to eliminate redundancy as follows:

Our simulations suggest that the absence of increased radial growth could be explained as a consequence of a shift towards below-ground carbon allocation.

Page 4769/70Page Line 10 Simulated and observed consistent - is this basically the same statement as the previous sentence? The next sentence seems to be another repetition of the previous statement. Condense?

We agree that this is a little repetitive. We have amalgamated the two sentences to read:

Both simulated and observed responses to climate show a significant positive response of tree-ring width to total photosynthetically active radiation received and to the ratio of modeled actual to equilibrium evapotranspiration, and a significant negative response to vapour pressure deficit.

Page 4770/4771, Line 2. why despite?

We say despite, because in other parts of the world climates as dry as that of the Great Western Woodlands categorically do not support large trees. The Great Western Woodlands is indeed unique because of this, and because it has a very high biodiversity. We have made no change to the text.

Page 4770/4771, 17 not sure if everyone is familiar with the supersites.

*The fact that our study area is associated with a Supersite is not really relevant to this modelling study, but we refer to the Supersite (and gives its latitude and longitude) because it is a source of data on climate and its location and character will be familiar to Australian ecologists. However, because Supersite this might be of interest to other readers, we have added the url for the site (*http://www.tern-supersites.net.au/supersites/gwwl) *into the text.*

Page 4770/4771, 18 and following statements about temperature: where is this data coming from?

This is based on our analyses of the CRU data set, using values for the grid cell in which the GWW Supersite is located. This is not clear from the text, so we have changed the text to read:

Analysis of the CRU TS v3.22 climate data (Harris et al., 2014) for the location of the GWW Supersite at Credo (30.1°S, 120.7°E; http://www.tern-supersites.net.au/supersites/gwwl) shows that mean annual temperature has increased significantly in the last 100 years (0.139 ± 0.015 °C/decade, $p < 0.001$).

Page 4770/4771, 26 what do you mean by anthropogenic? Due to climate change, or do you question whether CC is anthropogenic?

We are not challenging the fact that changing [CO_2] is due anthropogenic emissions; nor are we challenging the fact that climate change is happening because of these emissions. However, climate variability over most of Australia is very large and under these circumstance it remains somewhat difficult to attribute regional climate changes to anthropogenic increases in [CO_2]. This is the only point we wanted to make here. We have expanded the text to read:

The changes in climate in southwestern Western Australia cannot be unambiguously attributed to anthropogenic increases in [CO_2] (Pitman et al., 2004; Cai and Cowan, 2006).

Page 4770/4771, 27 however -> in any case

We do not mean "in any case" we actually mean "however". No change made.

Page 4771/4772, line 10. ff to show what? What were your hypotheses?

This is the same issue as above. We have changed the abstract so that the issue we are addressing is clear from the beginning. However, to emphasise this here we have modified the final sentence to read:

We then use this model to explore the impact of changes in climate and [CO_2] on tree growth under water-limited conditions, and specifically to test whether we can detect an impact of [CO_2].

Page 4771/4772, line 17. provide precipitation values again, or move them from the intro. You could also highlight Fig.3, which is very useful to see the climatic characteristics of the site.

We do not give precipitation values per se in the introduction, only evidence for changes in climate variables including precipitation. So we think it is a good idea to include a brief description of the mean climate here, and amend the text to read:

The sampling site (GWW Super Site, Credo, 30.1°S, 120.7°E, 400m a.s.l.) lies in the northernmost and driest part of the GWW (Fig. 1), with a mean annual rainfall of ca 270 mm.

Page 4771/4772, line 7 showed no obvious, or no? It should be pretty clear if there is variation or not.

We did not test for all environmental parameters in the field, and so it is possible that there were small variations in e.g. pH, soil organic or nutrient content, water content etc. This is why we prefer to say no obvious variation rather than no variation. No change made to text.

Page 4771/4772, line12 A single stem of given diameter would have another allometry than three stems of the same aggregated diameter. In the data that you used to create the allometries, was the fraction of multiple stems similar? It seems important to me that you discuss the potential issues that could arise through this decision in more detail. Is your model sensitive to the error created by this assumption?

These allometric parameter will have a effect onto the ontogenetic ageing trend, which is faster in the very beginning growing years (10-50 years), then slowing down to a more constant level of growth. These allometric parameters should not change the simulated interannual variability and basic trend of the mature trees.

Page 4773/4, line 20; at some point here you seem to move from the P to the T model, but it's not clear to me exactly when

We have created a new paragraph at the point where we start to describe the T model, and modified the text to make this clearer as follows:

In the T model, the fraction of incident PAR absorbed by the canopy (fAPAR) is estimated from the leaf area

Page 4774/5, line11 You say your model "has no free parameters", by which I understand that you want to say that the parameters should be chosen identically for all C3 species. There is clear evidence for differences in photosynthesis of C3 plants. Even if the P-model does not include the proposed mechanisms that lead to these differences, it seems unlikely to me that one couldn't get a better description of individual species' photosynthesis if one would adjust them. Moreover, with respect to the question that come up later - can you exclude the possibility that parameters in the P-model should be adjusted if CO2 changes, or would you say that this is impossible?

No "free" parameters here means none that have to be estimated for a particular data set. So yes, the P model is supposed to apply to C_3 plants.

Of course, there is a great deal of observed variation (among species, and environments) in photosynthetic capacity (V_{cmax} and J_{max} – the subject of the Evans paper) and in the response of photosynthesis to drought (the subject of the Flexas & Medrano paper). However, the P model is designed to represent the behaviour of plants that are adapted, and acclimated, to their natural habitat. Acclimation includes the well-established trend to increasing photosynthetic capacity and leaf N along gradients of increasing drought. There is also evidence that the steepness of decline in stomatal conductance and apparent V_{cmax}, both of which vary greatly among species, is systematically reduced in plants adapted to dry environments (e.g. Zhou et al. 2013).

Zhou, S., Duursma, R., Medlyn, B. E., Kelley, J. W. G. and Prentice, I. C.: How should we model plant responses to drought? An analysis of stomatal and non-stomatal responses to water stress, Agricultural and Forest Meteorology, 182-183, 204-214 (2013).

The P model is based on two explicit hypotheses: the least-cost hypothesis (Prentice et al. 2014, cited in the text), and the co-ordination hypothesis, which rests on the long-standing observation that the electron transport- and Rubisco-limited rates of photosynthesis tend to equality under typical daytime conditions. Application of these hypotheses allows us to avoid specifying values of V_{cmax} or J_{max}. The model does possess two parameters that have to be estimated from data: one in connection with the least-cost hypothesis, which has been estimated from global $\delta^{13}C$ measurements; another for determining J_{max}:V_{cmax} ratios, which has been obtained from published experimental data (unpublished results by H. Wang et al.) Other terms required for modelling photosynthesis – the intrinsic quantum efficiency of photosynthesis, the photorespiratory compensation point, and the affinities of Rubisco for CO_2 and O_2 – vary only slightly among wild C_3 plants. The model's generality allows prediction not only of GPP in different environments (T. Davis et al., unpublished results) but also the response of GPP to changing atmospheric [CO_2], which implies a modest (and realistic) downregulation of V_{cmax}. Importantly, therefore, no parameter of the P model is expected to vary with [CO_2].

Page 4774/5, 13 150 trees of what sizes?? How selected?

Out sampling was comprehensive and measurements were mace on all specimens of Callitris within a 1 km^2 plot at the sampling site. We have added as statement to this effect in the text as follows:

The measurements were made on all Callitris trees within a 1 km^2 plot.

Page 4774/5, line 21 I'm a bit confused as to which parameters are in what model. If the P-model simulates GPP, then shouldn't quantum efficiency be in there? Here it seems it's in the T model. Yield factor as well?

Quantum efficiency is in fact a parameter of the P model not the T model, so it is somewhat confusing to have it here. We have now taken this out of the the table, but inserted the value used in the text as follows:

the intrinsic quantum efficiency of photosynthesis (taken as 0.085: Collatz et al., 1998; Wang et al., 2014)

The yield factor is a parameter of the T model and we have left this in the Table.

Page 4775, line 1. I appreciate that one can probably get this information from the cited publications, but as this is pretty central to the further story, it would be useful if you could explain shortly why these parameters are so influential, and what the underlying assumptions behind that are (physiological / ecological), and if the model response caused by these parameters (specifically fine root allocation) could also originate from some other parameter that wasn't varied in the analysis. Looking at Table.1, there seems uncertainty also in the other parameters. So why didn't you put them all under calibration, potentially with constraints given by the uncer- tainty ranges that you had here?

We have already published a comprehensive sensitivity analysis, which is discussed in the cited references. Furthermore, we have now allowed all the parameters to vary (see response above) and shown that the model response in terms of fine root allocation is not an artifact of only calibrating "unknown" parameters.

3 Why this algorithm, and not a simple optimizer? What was the objective function for the optimization?

Please see response to this question above.

11 I'm not sure if I get it correctly - when you say linear interpolation, do you mean that you take the monthly precipitation, and distribute it evenly across all days of a month? Would your model not not react differently when you compare two scenarios 1) evenly distributed rain 2) unevenly, according to typical precipitation patterns for this region, which should have a mix of rain and dry days. Same problem for other variables potentially.

It is correct that the climate variables are lineally interpolated between mid-month points. The input to the T model is annual GPP. We have compared simulations run with actual daily climate inputs with simulations using daily input interpolated from the monthly values, and shown that this has a negligible impact on annual GPP.

Page 4776/7, line 1 About this whole section: it was unclear to me why you do this, and how the growing season enters your model/analysis. Is the growing season a parameter of the T model?

No the growing season is not a model parameter, but it is necessary to define the period over which carbon is accumulated in the formation of a ring.

Page 4776/7, line 15 1) Why do you use a GLM? The response seems normal. 2) What kind of GLM did you use, i.e. which distribution 3) OK, you regress climate against growth, vary the interval to do this, and look at p-values and Rˆ2. But what is the argument that tells you that the best Rˆ2 is the optimal growing season for your analysis? I see absolutely no reason for this. Couldn't it just be that you average away noise on larger time scales, hence the better Rˆ2 for a longer period, independent of what one would really think of ecologically as the growing season? Also, for sure you will get much better predictions for shorter time scales if you include nonlinear relationships and lags. Or let's put it the other way around: is your definition of growing season: the time span that allows the best prediction of growth with a regression using only linear terms of climate data as predictors?

Please see response above to these comments.

Page 4778/9, line 4. You state that the model captures the dynamics realistically (Fig.4) ... well, the mean growth is fine all right, but if we look at the variation of the mean, I doesn't look so great to be honest. This is a bit surprising because the univariate responses to climate seem indeed fine. Do you have an explanation for why the observed and predicted time series are so seemingly unrelated? Line 6 1) to state that r=0.37 is high is quite optimistic. A simple linear regression of 2-year PAR, MAT and alpha had an R2 of 0.3, not so much worse (Table2). 2) See main comments: I wonder how sensible it is to give p-values on this because data is not independent (individuals + temporal) 3) supplement Fig.4 by a predicted vs. observed plot

Our definition of the growing season is an implicit way of including time lags resulting from carbon storage dynamics. This is a simple approach which avoids the need to simulate these dynamics explitly, and our approach to estimating the growing season is rough but we selected the interval length that gave a peak in R^2 and significance level of the predictor variables. We state this in the last paragraph of this section. We suspect that the deviations may be connected with variability in carbon storage. The consistency between statistical relationships with climate shown by the the model and the observations, provides the key evidence that the model is capturing key elements of the variation. We have responded to the comment about autocorrelation and the comment about individuals + temporal non-independence above. Please see response there.

Page 4778/9, line 8 so, you exclude the possibility that it could be a problem of the model? 11 again, I wonder what kind of GLM you are using here.

Please see response to this comment above.

Fig. 5 separation of a / b is very easy to spot. Separate the panels visibly Lines18ff I would say that the fact that the temperature effect disappears or even changes sign (observation) due to the apparent collinearity of temp with VDP basically suggest that we should disregard Fig.5a, because the temp correlations are spurious, and rather concentrate on Fig.5b. Here, however, the model / data comparison doesn't look so convincing any more. There seems to be a slightly positive temperature reaction in the data, but not in the model, and there seems to be a difference in the VDP reaction as well. I would guess that the former is not significant, but the latter seems to be. What is your interpretation here - is there a discrepancy between model and data, and if so, could it be that some parameters in the P-model would need to be adjusted?

We have removed the upper panels as unnecessary and modified the text describing this Figure accordingly. Please see text above, describing the changes we have made. We have added a new Figure (Figure 10) which shows the results from the full model, and this shows a slightly positive result with temperature for both observations (albeit non-significant) and in the model.

Page 4779/80 Line 6 What was unclear to me - with the varying CO2, Fig. 6/7, did you recalibrate the model, or did you use the calibration done for Fig.4/5

We did not recalibrate the model, because as shown Table 1, the calibrations are virtually identical with CO_2 fixed or varying. To make this clearer, we have modified the description of the second experiment as follows:

To examine the impact of changing [CO_2] on tree growth, we made a second simulation with the same parameter values but using the observed annual [CO_2] between 1901 and 2013 (296–389 ppm: Fig. 3).

8 For the Regression, why did you exclude temperature? We see in Fig. 3 that they are collinear, it seem crucial to include this in the regression.

We have re-run the regression including temperature, as previously mentioned.

Page 4780/1, Line18 Are you explaining the model results or the data?

We are actually describing relationships that are apparent in both the observations and in the simulations. To make this clearer, we have add a sentence saying:

These relationships are seen in observations and reproduced by the model.

4781/2, Line 4 As the climate was changing at the same time, I would formulate this a bit more carefully.

This result comes from the regression analysis, where the partial residual plots for the observations show no independent relationship with CO_2 (Figure 7). Since this is the case, we have not modified the sentence.

4781/2, Line 6 it may be that tree-ring studies find that, but on the other hand inventory and satellite data does seem to support effects of CO2 fertilisation. What is your take on this?

One point here is that the different communities are measuring different things: the satellite data are measuring greenness or total biomass and the tree-ring people are measuring radial growth only. Our results suggest that the two lines of evidence are compatible if increased productivity from CO_2 fertilisation results in a change in allocation.

4781/2, Line 12 wouldn't we expect increase in WUE also without fine root changes? –> so what evidence does this add to fine root changes?

Yes we would expect an increase in WUE without fine root changes. The increase of allocation to fine root is our explanation for the discrepancy shown by van der Sleen et al.

Figs 4,6,9: I assume the model error bars originate from the different conditions at the different sites? Clarify.

The model "error bars" represent the standard deviation of the ten individual trees that we simulated. Each of these individuals was started at a different size, corresponding to the size of the actual tree in the first year of the simulation. This is explained in the first paragraph of section 2.3.4. We have clarified the captions for these three figures.

Reviewer 2

Overall, I think that a very heavy revision is needed to make it a smooth paper for BG.

We have made a substantial revision of the paper in response to Florian Hartig's detailed comments.

1. pg 4775, L26 onwards. Key parameters of the allocation process have been derived from Bayesian optimization using only tree-rings. I'm not so familiar with Bayesian techniques, but how reliable is an 'allocation routine' (i.e. C partitioning between different tree organs and functions) when data for only 1 organ (stem) and 1 function (growth) are used as constraints? Some processes are very difficult to measure and few data are available, but I would expect that a calibration of an allocation routine should comprise at least 'something' also about roots and leaves (root and leaves biomass for instance) and not only stem. If I'm wrong please explain why.

The strength of Bayesian calibration is that it is absolutely not necessary to have observations for every parameter that one wishes to calibrate. We have repeated the calibration, including all of the parameters in the T model (see response to reviewer 1 on this) and have shown that we obtain identical results. A strong indication of the robustness of the techniques is that we obtain coherent results for different time periods, and the fine root/foliage area ratio is the parameter that emerges as responding to changing CO_2

Also I note that the tree-ring data used for calibration are also used for validation. Have you tried to split the dataset and used different datasets for calibration and validation? It might also be instructive (as your parameter variability shown in Fig. 8).

Regression analysis and associated significance testing do not require separation of validation and calibration data sets. We have not tried splitting the data set, but our results do include analyses based on separate sub-sets which show robust results.

2. the performance of the model is only evaluated against tree-ring data. However, exact simulation of stem growth might be the results of different biases: e.g. overestimation of GPP and overestimation of respiration. An analysis of the overall modeled fluxes (GPP, respiration, stem growth, leaf growth, fine root growth) will say more about the model, even if the fluxes can not be validated because lack of measurements.

The P model has been validated against flux derived GPP data – this is described in Wang et al (2014). So, if the modelled respiration were wrong, then the simulated tree rings would be wrong. We have added a sentence about the validation of the P model in the section describing that model (section 2.3) as follows:

The P model has been shown to reproduce observed geographic patterning in the magnitudes of observed GPP reasonably well (Wang et al., 2014).

3. The fact that radial growth has not responded to increasing [CO2] is recent decades is interesting but the story about increased allocation to fine roost is highly speculative. It can be mentioned but I do not see any strong indication supporting this. The only support is that the parameter 'fine-root mass to foliage area' increased by about 14% from 1950 to the end of the period. . .. But this parameter was not calibrated against fine-roots or leaves but just stem (see above). I do not find this convincing . . .. Lines as the three ones at beginning of pg 4783 are not-supported. It is reasonable that trees with more carbon resources invest more in fine roots in infertile or dry sites to favor uptake of the limiting resources. But also all other possibilities should

be mentioned and then eventually excluded: e.g. increase in respiration, increased allocation to leaves (are not your data of Fig 5 (ring vs PAR), indicating some light limitation?) or even increased allocation to belowground transfer to mycorrhiza/exudations . . .

We have been very careful not to overstate the significance of our findings, but we feel that this modelling result (which is not strengthened by the fact that a full calibration yields as similar conclusion) could help to explain an apparent paradox in the literature about CO_2 fertilisation. Ww hope that this work will promote more field investigation of the issue, because this is clearly necessary to definitively test the hypothesis that the response to CO_2 is primarily through changing allocation.

Analysis of flux data show that on monthly (or longer) times scales, light limitation of GPP is ubiquitous. However, we can now rule out increased allocation to leaves over the period considered based on our re-calibration of the model with other variables (including LAI) allowed to vary – see detailed description of these results above.

It is possible that part of the belowground transfer is to mycorrhiza/exudates. Measurements suggest that both increase together with an increase in fine roots. In our model, we do not treat mycorrhiza/exudates separately, so in a sense the simulated increase represents an overall change in belowground allocation. We have added some text on this issue in the Discussion as follows:

It is possible that this belowground allocation could in part represent increased carbon export to mycorrhiza or the rhizosphere (Godbold et al., 2015).

4 I find results of Fig 5 interesting and clear. However, I do not find necessary to know how you arrive to understand that MAT is not important. For me, Fig5b is the only one you need to show . . . Similarly, all the section starting from L18 of discussion is a repetition of what already said . . .

We have taken the suggestion to remove Figure 5a, and have simplified the text associated with describing the results in both the Results and Discussion sections. Please see detailed response above

5. you say in results that "T model captured the amplitude and interannual variability of Callitris tree growth in the GWW realistically (Fig. 4)". However, there are many deviations and in many years growth is underestimated or overestimated by ca. 50%. Why some growth peaks are very well captured and others (as around 1970) are not captured? What about the deviations around 2005-2010? And what about the period 1920-1930? Similar (last line of Results) ".... In particular, this simulation does not produce an overestimation of ring widths in recent years compared to observations": it is true that there is an improvement but still a large overestimation around year 2005. Please be accurate in results.

We agree that the model-data agreement is modest, but as discussed above the good similarity between the regression results based on model and observations that the model is capturing key features of the climate response. We have rephrased the statement to be more accurate as follows:

In particular, this simulation does not produce the large systematic overestimation of ring widths in recent years compared to observations that is seen in the simulation with observed [CO_2] and fixed ζ.

6. I think the study need a better structure: (i) what are your detailed objectives? For example, you mention "growth response to the shift in precipitation regimes in the 1960-70s" in both abstract and discussion; is this a goal? Definitions of goal might help you to have a clear discussion (ii) I would present first the GLM results and then the results of the other model (e.g. about Fig. 4 and similar). In doing so, maybe your GLM results can be used to understand and explain the issues I mentioned above in my point 5. (iii) Abstract should be re-written accordingly to what said above.

We have restructured the abstract to make it clear at the beginning what our main objective is in this study, and we hope that this addresses the reviewer's concern. The purpose of the regression analysis is to demonstrate similarity of behavior between model and observations, and this cannot be dealt with separately or before the results of the tree-growth modeling.

Minor remarks

Paragraph 2.6. I understand your point about "Growing season"; but many people define "growing season" as something else (period of measurable diameter growth); so your text is confusing. Actually you do not need whole this info about growing season. Delete?

It is necessary to include the information about growing season because otherwise the description of our methodology would be incomplete. We define what we mean by growing season in the first paragraph, so this should make it clear that we are not talking about the period of measurable diameter growth.

Table 1: yield factor (y): if you re-ran the model you could use a better estimate derived from data (see Vicca et al 2012 Ecology Letter); I guess that you are using now a simulated value of y . . ..

The reviewer has misunderstood what the yield factor is. The yield factor refers to only the growth component of respiration. Vicca et al (2012) provide data on the ratio of NPP to GPP and the ratio of biomass increment to GPP; the former includes both growth and maintenance respiration, and the latter also includes carbon lost e.g. through exudation. The value we use for the yield factor is not simulated but rather prescribed, and is typical of values found in the literature.

4776 L23. And what about correlation with the other meteo stations?

We provide the two correlations as examples. These are the worst correlations between CRU and the meteorological station data for each variable.

Why not joining Fig 4, 6 and 9 in only one figure?

We prefer to keep these as separate figures, so that they can be referred to at the appropriate place. There is too much information involved to plot them in a single panel, and a three-panelled figure would take as much space as the current three figures do.