We thank the referee for providing a review of the manuscript and agree that the suggested changes and clarifications improve it. We have made the changes outlined below in the revised manuscript. Each item starts with the reviewer's comment followed by the changes to the manuscript. The text in blue is a re-written or new paragraph/sentence which was added to the manuscript. The page and line numbers of where changes have been made to the updated manuscript are included at the end of each reply.

Comments

- **General comments; I can't understand the novelty of this manuscript. I agree that the novelty is the performance confirmation of JULES. Please rethink why the authors would like to show others the original results via this manuscript. And, for all things, if the content is related to just JULES unique performance confirmation, it might not directly help the reader's scientific knowledge. At such times the authors need to improve the explanation by changing the standpoint. Please rewrite the manuscript to serve to help the readers in getting maximum benefit from what the authors revealed.**
  This study provides an evaluation of JULES at global and regional scales and provides details on which ecosystems/regions to focus on for future improvements of the model. Changes have been made to the manuscript as suggested and we have added extra text to the manuscript in order to improve and clarify it.

- **Please organize all the information of the model introduction. The authors wrote them in 1. Introduction section and 2.1 Model description. Naturally the 2.1 section should be included contents directly related to this study's discussion, and omit the explanation that had little to do with this study. For example, the authors wrote the interminable explanation for the GPP calculation method, but the reader can understand several author statements in discussion section without such knowledge; there is no explanation about spatial resolution as model structure . . . etc.**
  The model description section has been re-written in order to include only the contents directly related to this study and to explain the effect of the meteorological data on photosynthesis and thus GPP (Pages 3–4). The following text was added:

  JULES is driven by the downward shortwave and longwave radiation fluxes, rainfall and snowfall rates, surface air temperature, wind speed, surface pressure and specific humidity. The downward shortwave and longwave radiation fluxes play an important role in the surface energy balance, where the downwelling radiation fluxes must equal the outgoing fluxes of sensible heat, latent heat, ground flux, reflected shortwave radiation and upwelling thermal energy, and the calculation of photosynthesis (Best et al., 2011; Clark et al., 2011). GPP is the total C used by plants in photosynthesis at the canopy scale with potential (without water and ozone stress) leaf-level photosynthesis calculated as the smoothed minimum of three limiting rates: (1) Rubisco-limited rate (determined using surface air temperature and atmospheric CO_2 concentrations), (2) Light-limited rate (determined using downward radiation fluxes) and (3) Rate of transport of photosynthetic products (C_3 plants) and PEP-Carboxylase limitation (C_4 plants) (determined using surface air temperature and pressure) (Clark et al., 2011). By taking soil moisture stress into account, leaf-level photosynthesis is calculated by multiplying the potential leaf-level photosynthesis by a soil moisture factor (determined using mean soil moisture concentration in the root zone and thus, precipitation).

In JULES, there are two options available for radiation interception and the scaling of photosynthesis from leaf-level to canopy-level: (i) big leaf approach and (ii) multi-layer approach. For all model simulations performed in this study, the multi-layer approach was used which takes into account the vertical gradient of canopy photosynthetic capacity (decreasing leaf nitrogen from top to bottom of canopy) and includes light inhibition of leaf respiration (Option 4 in Table 3 of Clark et al. (2011)). Canopy-scale fluxes are estimated to be the sum of the leaf-level fluxes in each canopy layer, scaled by leaf area. LAI is calculated for each canopy level (default number is 10), with a maximum LAI prescribed for each PFT.

- **Please more explain why the authors used different climate dataset. What of the JULES GPP estimate do the authors reveal? Why did you examine just sensitivity to each dataset? (why didn't you choose the sensitivity to each meteorological parameter?). Please add the comparison among three climate datasets into results. I can't understand the impact of climate dataset on GPP (e.g., fig. 2, 3 . . .), because I don't know the difference of the climate dataset specific feature related to this study. Moreover, please add the explanation of the relationship between JULES and the meteorological parameters in 2.1 Model description section. It means, the reader would like to know the model structural interpretation in discussing what types of calculation approach to choose.**

Reasons for why we used different meteorological datasets to drive JULES were added to the experimental design section (Page 4, lines 22–25).

A general overview is provided of how sensitive JULES GPP is to the meteorological dataset used at global scales rather than for each meteorological variable. By analysing the models sensitivity to each meteorological dataset, different analyses of the global climate are compared and therefore a multi-factor analysis of combined changes in meteorological variables can be performed.

The sensitivity to meteorological parameters was performed in Chapter 6 of the PhD thesis of Darren Slevin(Slevin, 2016). A brief summary of this sensitivity study is provided in section 4.4 (Page 15, lines 23–32).

A simple sensitivity study of the model to changes in climate (surface (2m) air temperature, precipitation and atmospheric CO_2 concentrations) when simulating GPP at global and regional scales for 2000–2010 was performed in Chapter 6 of the PhD thesis of Darren Slevin(Slevin, 2016) . Only changes to one climate variable were made at a time due to complex interactions associated with multiple changes in climatic factors resulting in complex non-linear ecosystem responses which can be difficult to explain. JULES GPP was found to be sensitive to changes in all three climate variables with modelled LAI only sensitive to changes in surface air temperature (Slevin, 2016). At the regional scale, for model simulations with varying air temperature, GPP increased with increasing temperature in the extratropics, but decreased with increasing temperature in the tropics. Model simulations with varying precipitation at regional scales show the same trend as those at global scales with GPP increasing with increasing precipitation and decreasing with decreasing precipitation except for the magnitude of the effect observed.

Information on how differences in the three climate datasets (WFDEI-GPCC, WFDEI-CRU and PRINCETON) affect GPP simulations has been included in various parts of the manuscript (Page 11, lines 19–24; Page 15, lines 1–16). In the model description section

(2.1), a paragraph has been added which provides an explanation of the relationship between JULES and the meteorological parameters (Page 3, lines 25–30; Page 4, lines 1–5). This relationship between JULES and the meteorological parameters is also included in the discussion section (Page 12, line 30–Page 13, line 10; Page 15, lines 1–8; Page 15, line 33–Page 16, line 9).

- **The authors should organize first and second paragraph of "1. Introduction". The authors should integrate the two paragraphs into one. P1 L19-20: delete the sentence (Changes in atmospheric CO2 . . .). P2 L2 and L4: "location of" → reservoir in? P2 L3: "Changes in the land surface" is not clear. P2 L7: "models and observations (Friedlingstein" → the existing studies (e.g., Friedlingstein . . .**
  The first two paragraphs of the introduction have been combined into one with changes made to the text and repetitive text removed. The new paragraph follows (Page 1, line 18–Page 2, line 8).

  The land surface is an important component of the climate system, provides the lower boundary for the atmosphere and exchanges energy, water and carbon (C) with the atmosphere (Pielke et al., 1998; Pitman, 2003; Seneviratne and Stöckli, 2008). It also controls the partitioning of available energy (into latent and sensible heat) and water (into evaporation and runoff) at the surface (Bonan, 2008). Changes in the land surface due to human activities, such as those from tropical deforestation, can influence climate at various time and spatial scales and since the land surface is the location of the terrestrial C cycle, it's ability to act as a C source or sink can influence atmospheric CO_2 concentrations (Le Quéré et al., 2009; Pan et al., 2011; Le Quéré et al., 2013; Tian et al., 2016). The reduced ability of the land surface to absorb increased anthropogenic CO_2 emissions in the future has been shown by models and inferred from observations (Friedlingstein et al., 2006; Canadell et al., 2007; Friedlingstein et al., 2014; Sitch et al., 2015). Friedlingstein et al. (2006) and Friedlingstein et al. (2014) have suggested that a major source of model uncertainty is the land C cycle and this can affect the ability of earth system models (ESMs; also known as coupled carbon-cycle–climate models) to reliably simulate future atmospheric CO_2 concentrations and climate (Dalmonech et al., 2014).

- **The explanation relevant to data used is strange format (P5 L10-P7 L21). For example, why is parameter's unit necessary here? most explanation of "P6 L28-P7 L9" is for the Zhao's work, not this study. After downloading the data, what did the author do as the data pre-processing? The explanation directly related to this study (P7 L13-21) should be written at the start of the paragraph . . . etc.**
  The units of the meteorological variables used to drive JULES has been tidied up in the Data section (Section 2.3) (Page 5, lines 26–28 and lines 32–33). The information provided regarding Zhao's work has been shortened (Page 6, lines 14-26). Information on how the data was pre-processed has been included at the end of the paragraph (Page 6, lines 24-26). The paragraph regarding CARDAMOM was structured in such a way that general information on the framework was put first followed by the model output used in this study (Page 6, lines 27–34; Page 7, lines 1–4).

- **P10 L19-21: The statement does not match with fig. 3. It is significant mistake.**
  The statement now reads (Page 8, lines 24–26)

This value is greater than that estimated by MODIS, FLUXNET-MTE and CARDAMOM with annual average global GPP estimated to be 112, 130 and 114 Pg C $year^{-1}$, respectively, for the same period (Figures 2a, b and d).

- **P15 L13-14: "In general, CARDAMOM was better at simulating GPP than JULES.". Please present factual evidence if the statement is correct. The dataset is created with ground observations, and the empirical method is used to expand it from point to spatial data; CARDAMOM may include some significant error.**
  The statement "In general, CARDAMOM was better at simulating GPP than JULES." was used since global GPP simulated by CARDAMOM GPP (Figure 2) and the pattern of zonal means of total annual model simulated GPP (Figure 5) was between that of MODIS and FLUXNET-MTE. However, we removed this sentence from the Conclusions section (Page 17, lines 1–7). All GPP estimates have errors, but these are not always quantified and provided.

  In the conclusions, the following paragraph was added which discusses the sources of error in the three benchmarking datasets (Page 17, line 30–Page 18, line 4).

  The three benchmarking datasets all contain sources of error. Since observations of GPP do not exist at global scales, the MODIS and FLUXNET-MTE datasets are referred to as observation-based estimates of GPP as they are generated using observations and models. CARDAMOM may contain significant error from the assimilated data and model structure (number of pools, fire resilience of ecosystems), but so do the empirically based FLUXNET-MTE data (up-scaling of a partitioning algorithm) and MODIS GPP (a model based on PFT specific light-use efficiency). The advantage of CARDAMOM is that it is a process-based model and it ensures that the whole ecosystem functioning is coherent, while the observation-based datasets are only empirically based representations of GPP. In Figure S4 of the Supplementary Information of Bloom et al. (2016), there is a detailed study of the sensitivity of CARDAMOM to these various factors at 4 selected pixels representing temperate, boreal, wet and dry tropical ecosystems. Overall, there is not much difference in retrieved parameters because of the large error/uncertainty terms used when computing the likelihood.

- **Fig. 7: As everybody knows, accuracy of the satellite observations is essentially not good at low latitude because of bad observed condition by cloud cover. The authors should represent the difference of GPP in not only low latitude but also other region. Since the evaluation data is global scale, you can do the comparison at global scale. If you keep the way to compare your results with others at just low latitude, please explain the reason.**
  The reason for only examining the difference in GPP fluxes between 15°N–30°N (Figure 5) was to find out which region contributed most to this difference and the possible reasons behind it. We suggest that this difference in GPP was due to incorrect simulation by JULES in Mexico (Figure 7). Even when JULES was driven with multiple meteorological datasets, it was unable to simulate GPP in this region (Figure 5) (Page 12, lines 14–29; Page 15, lines 6–8).

- **Abstract; L6: delete "it was found that" L8: delete "fluxes" L9: delete "It was found that L9: between → among L9-11: this sentence is not clear. L12: what is the meaning of "no impact" → Please add the quantitative interpretation.**
  The words were deleted as suggested (Page 1, lines 5–7; Page 1, line 8; Page 1, lines

9–10). The sentence at lines 9-11 was re-written (Page 1, lines 9–10). A quantitative interpretation was added to Line 12 (Page 1, lines 10–12).

Bibliography

M. J. Best, M. Pryor, D. B. Clark, G. G. Rooney, R .L. H. Essery, C. B. Ménard, J. M. Edwards, M. A. Hendry, A. Porson, N. Gedney, L. M. Mercado, S. Sitch, E. Blyth, O. Boucher, P. M. Cox, C. S. B. Grimmond, and R. J. Harding. The Joint UK Land Environment Simulator (JULES), Model description–Part 1: Energy and water fluxes. *Geoscientific Model Development*, 4:677–699, 2011. doi: 10.5194/gmd-4-677-2011.

A. A. Bloom, J.-F. Exbrayat, I. R. van der Velde, L. Feng, and M. Williams. The decadal state of the terrestrial carbon cycle: Global retrievals of terrestrial carbon allocation, pools, and residence times. *Proceedings of the National Academy of Sciences*, 113:1285–1290, 2016. doi: 10.1073/pnas.1515160113.

G. B. Bonan. Forests and Climate Change: Forcings, Feedbacks, and the Climate Benefits of Forests. *Science*, 320:1444–1449, 2008. doi: 10.1126/science.1155121.

J. G. Canadell, C. Le Quéré, M. R. Raupach, C. B. Field, E. T. Buitenhuis, P. Ciais, T. J. Conway, N. P. Gillett, R. A. Houghton, and G. Marland. Contributions to accelerating atmospheric CO_2 growth from economic activity, carbon intensity, and efficiency of natural sinks. *Proceedings of the National Academy of Sciences of the United States of America*, 104: 18866–18870, 2007. doi: 10.1073/pnas.0702737104.

D. B. Clark, L. M. Mercado, S. Sitch, C. D. Jones, N. Gedney, M. J. Best, M. Pryor, G. G. Rooney, R. L. H. Essery, E. Blyth, O. Boucher, R. J. Harding, C. Huntingford, and P. M. Cox. The Joint UK Land Environment Simulator (JULES), model description–Part 2: Carbon fluxes and vegetation dynamics. *Geoscientific Model Development*, 4:701–722, 2011. doi: 10.5194/gmd-4-701-2011.

D. Dalmonech, S. Zaehle, G. J. Schürmann, V. Brovkin, C. Reick, and R. Schnur. Separation of the Effects of Land and Climate Model Errors on Simulated Contemporary Land Carbon Cycle Trends in the MPI Earth System Model version 1. *Journal of Climate*, 28:272–291, 2014. doi: 10.1175/JCLI-D-13-00593.1.

P. Friedlingstein, P. Cox, R. Betts, L. Bopp, W. Von Bloh, V. Brovkin, P. Cadule, S. Doney, M. Eby, I. Fung, et al. Climate-Carbon Cycle Feedback Analysis: Results from the C4MIP Model Intercomparison. *Journal of Climate*, 19:3337–3353, 2006. doi: 10.1175/JCLI3800.1.

P. Friedlingstein, M. Meinshausen, V. K. Arora, C. D. Jones, A. Anav, S. K. Liddicoat, and R. Knutti. Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks. *Journal of Climate*, 27:511–526, 2014. doi: 10.1175/JCLI-D-12-00579.1.

C. Le Quéré, M. R. Raupach, J. G. Canadell, and G. Marland. Trends in the sources and sinks of carbon dioxide. *Nature Geoscience*, 2:831–836, 2009. doi: 10.1038/ngeo689.

C. Le Quéré, R. J. Andres, T. Boden, T. Conway, R. A. Houghton, J. I. House, G. Marland, G. P. Peters, G. R. van der Werf, A. Ahlström, R. M. Andrew, L. Bopp, J. G. Canadell, P. Ciais, S. C. Doney, C. Enright, P. Friedlingstein, C. Huntingford, A. K. Jain, C. Jourdain, E. Kato, R. F. Keeling, K. Klein Goldewijk, S. Levis, P. Levy, M. Lomas, B. Poulter, M. R. Raupach, J. Schwinger, S. Sitch, B. D. Stocker, N. Viovy, S. Zaehle, and N. Zeng. The global carbon budget 1959–2011. *Earth System Science Data*, 5:165–185, 2013. doi: 10.5194/essd-5-165-2013.

Y. Pan, R. A. Birdsey, J. Fang, R. Houghton, P. E. Kauppi, W. A. Kurz, O. L. Phillips, A. Shvidenko, S. L. Lewis, J. G. Canadell, P. Ciais, R. B. Jackson, S. W. Pacala, A. D. McGuire, S. Piao, A. Rautiainen, S. Sitch, and D. Hayes. A Large and Persistent Carbon Sink in the World's Forests. *Science*, 333:988–993, 2011. doi: 10.1126/science.1201609.

R. A. Pielke, R. Avissar, M. Raupach, A. J. Dolman, X. Zeng, and A. S. Denning. Interactions between the atmosphere and terrestrial ecosystems: influence on weather and climate. *Global Change Biology*, 4:461–475, 1998. doi: 10.1046/j.1365-2486.1998.t01-1-00176.x.

A. J. Pitman. The evolution of, and revolution in, land surface schemes designed for climate models. *International Journal of Climatology*, 23:479–510, 2003. doi: 10.1002/joc.893.

S. I. Seneviratne and R. Stöckli. *Climate Variability and Extremes during the Past 100 Years* , chapter The Role of Land-Atmosphere Interactions for Climate Variability in Europe, pages 179–193. Springer, 2008.

S. Sitch, P. Friedlingstein, N. Gruber, S. D. Jones, G. Murray-Tortarolo, A. Ahlström, S. C. Doney, H. Graven, C. Heinze, C. Huntingford, S. Levis, P. E. Levy, M. Lomas, B. Poulter, N. Viovy, S. Zaehle, N. Zeng, A. Arneth, G. Bonan, L. Bopp, J. G. Canadell, F. Chevallier, P. Ciais, R. Ellis, M. Gloor, P. Peylin, S. L. Piao, C. Le Quéré, B. Smith, Z. Zhu, and R. Myneni. Recent trends and drivers of regional sources and sinks of carbon dioxide. *Biogeosciences*, 12:653–679, 2015. doi: 10.5194/bg-12-653-2015.

D. Slevin. *Investigating sources of uncertainty associated with the JULES land surface model.* PhD thesis, School of GeoSciences, University of Edinburgh, 2016. URL `http://hdl.handle.net/1842/18757`.

H. Tian, C. Lu, P. Ciais, A. M. Michalak, J. G. Canadell, E. Saikawa, D. N. Huntzinger, K. R. Gurney, S. Sitch, B. Zhang, J. Yang, P. Bousquet, L. Bruhwiler, G. Chen, E. Dlugokencky, P. Friedlingstein, J. Melillo, S. Pan, B. Poulter, R. Prinn, M. Saunois, C. R. Schwalm, and S. C. Wofsy. The terrestrial biosphere as a net source of greenhouse gases to the atmosphere. *Nature*, 531:225–228, 2016. doi: 10.1038/nature16946.