Responses to comments from anonymous referee #2

Main comments
The manuscript describes a validity test of the decomposition model Yasso07 for tropical conditions. Litter bag experimental data is used to calibrate the model in order to find new more appropriate parameterization for SOC change estimates in tropical ecosystems. The manuscript aims to contribute to an improved applicability of the Yasso07 model, which is of great interest in for ex. GHG fluxes for national reporting. Especially for those countries that use a low level method (Tier 1), have the possibility to improve their SOC change estimates with a newly parameterized model. The paper is clearly a calibration exercise. However, the authors do not show any of the "calibration results", which can help to clarify the large differences in the decomposition rates found.

Authors' response:
The model was fitted to new set of data and model parameterized with global+Benin data and with Benin data only resulted in consistent results on mass loss and on loss of specific chemical fractions (Figs. 3-7). However, high confidence intervals of the parameter values obtained with Benin data only (Table 2) showed that this model version is more uncertain than Y07A. See also responses to comments from referee #1. |
| The author show that the quality of litter and the parameter values between global/Benin data varies. This is not enough to analyze the new parameterization found in the calibration with Benin data. You need also to show how (pdfs prior and after calibration) the parameter sets vary and try to explain better why these things vary. Here you could analyze the correlation between parameters and also the density of the parameter prior and after the calibration. Also the decomposition rate between the Yasso07 with Benin only (Y07A) and Global+Benin data (Y07B) should not vary so much, at least not according to the results showed in the figures and also stated in results and discussion: Y07A and Y07B have similar mass loss.

Authors' response:
Attached the posterior densities of the model parameters for the global Y07, although we do not consider them very useful (especially as they are a bit crap plots) when the basic information has already been presented in Table 2. Unfortunately we don't seem to have correlation plots available, although there would be more than 400 of them for the 22 parameters in Table 2, so we do not fully understand why they are very necessary. We think even showing the 22 distributions in the paper would be simply waste of space, not to mention showing them for all three solutions. |
| Another thing is why do the authors not show uncertainties of the model estimates? They state early in methods that the calibration method is especially good because it renders confidence intervals. In addition, an evaluation of a model needs to include model residual analysis. The authors state only the magnitude of the residuals in the results. What distribution do the model residuals have? This is important in order to trust the confidence limits coming out of the model. Or did you make some assumptions about the model errors in your Bayesian calibration?

Authors' response: please see responses to comments from referee #1 and also response to previous comment. |
| The large litter quality variation should be discussed more thoroughly in order to understand the applicability of the new parameterization in tropical conditions. Are the quality values in the range of other studies? Outside?

Authors' response: |

to line 15, p.3014, the text added:
The chemical fractions data on AH and WS reported by these studies were lower than the lower limit of the data range originating from the tropics used in Y07A while data on ES and lignin were within the range. These observations indicate that even Y07A does not incorporate the full range of litter quality from the tropics and show the needs to further improve the model parameters and applicability in tropical conditions.

Minor and technical comments:

p. 3004

l. 5 Unclear if it is the modeling experiment or the field experiment. l. 14-16 This is not specially clear from the previous sentence. Perhaps write that the litter quality varied between the species, which were only captured by the recalibrated parameterizations. l. 22 Is it mandatory to report? And do all tropical African countries report? According to my knowledge it is voluntary to report to the climate convention (UNFCCC). To the Kyoto protocol each country can choose to not report, but then you need to show that the soils are a sink. African countries do not have the same demand on reporting as for ex European countries.

Authors' response:
p. 3004 l. 5, text revised to:
'The field experiment was conducted….'

l. 14-16, text revised
'The findings of this research showed that the high between species variation of litter quality in tropics resulted in the large variation of the rate of decomposition, which was only captured by the recalibrated parameterizations of the model'.

l.22: text revised to:
"The SOC changes are reported under the Climate Change Convention as a part of the national greenhouse gas inventories of the forestry sector in line with decision 17/CP.8 of the Conference of Parties of the UNFCCC (UNFCCC, 2003, 2008, 2010), but majority of the countries in tropical Africa either report no changes in SOC stocks or apply default stock change factors of the Intergovernmental Panel on Climate Change methodologies (par ex. IPCC, 2003, 2006) together with rough estimation of the land use and land use change."

p. 3005

l.1-5 Check language. Maybe a verb is missing. l. 6-11. Revise the sentence. Here you could also add what the changes are? Do all report the same trend? l. 18. It is difficult to make prediction in future with ecosystem models too. Many assumptions need to be made …l.21 Mostly? There exists many models and many studies, but the models need data (measurements) to be developed. So, I do not agree with this sentence.

Authors' response:
p.3005, l 1-5: text revised to:
"Under the UNFCCC mechanism "Reducing Emissions from Deforestation and Forest Degradation in Developing Countries (REDD)", countries have economic incentives for the conservation, the sustainable management and the enhancement of their forest carbon stocks and thus robust methods leading to confident and verified carbon stock estimates are increasingly required."

l. 6-11: text revised to:

"A number of studies have reported SOC stocks estimated from only sporadic soil sampling and digital maps (for example Manu et al., 1991; Prudencio, 1993; Volkoff et al., 1999; Henry et al., 2009) and the reported values in the existing databases in Africa (Batjes, 1996, 2002, 2005, 2006; FAO, 2008) were global estimates without clear indications on distribution of SOC and its changes according to ecosystems."

p.3005, l. 18 & l. 21. Yes, agree. Text on lines 20-23 revised as follows:
'Thus, processes taking place in soil and SOC stock changes are studied with the help of decomposition models that are based on empirical measurements from the target area (Coleman and Jenkinson, 1996; Currie and Aber, 1997; Kurz and Apps, 1999; Chertov et al., 2001; Liski et al., 2005; Sierra et al., 2012).

p.3006

l.6 What model?

Authors' response:
Misleading text corrected to
However, only a few litter types from tropical tree species were included in the dataset that was used for the parameterization of the widely applied soil carbon model (Tuomi et al., 2009) and many soil models are parameterized only for temperate and boreal conditions (Coleman and Jenkinson, 1996; Chertov et al. 2001).

p. 3007

l.6 Table 1. Maximum and minimum temperatures should be included?
Or did you use a version without the temperature amplitude? l. 16-17 It might
be easier to understand what type of forest rather than stating who carried out certain descriptions. l. 25 dry organic matter

Authors' response:
l.6: Minimum and maximum temperatures will be included in Table 1. However, mean monthly climatic data were used as input to the model.

l. 16-17: text changed to:
"Küppers et al. (1998) confirmed the tree species richness of the Lama forest described by previous studies (Akoègninou, 1984, Mondjannagni, 1969, Paradis and Houngnon, 1977) and reported 67 families based on an inventory carried out in 1998."

l.25: text changed to:
"In the Lama forest, the amount of leaf litter fall ranges from 26 to 42 dry organic matter yr^{-1};"

p.3008

l. 14 Repetition.

Authors' response:
Text on l. 13-15 revised to avoid repetition
'In total, 30 litterbags were placed in each plot from where five bags were collected every four weeks (between February and July 2010). The remaining litter was dried in open air and in the oven at 75 °C to constant weight.'

p. 3009-p. 3011

l.25 Is it meaningful to write the equations again? They are all already published before.

Authors' response:
Equations are shown in order to make table of parameter values understandable for a reader.

p. 3011

l. 9-10. Uncertainty information or values are not needed to run the model. The model can be run without this information. If you use the uncertainty information in the model simulations the results will be estimated with a confidence interval. l. 13 Figures are results? Or are they already published data that can be written about in the methods? l.14. Predictions are for future estimates. Do you mean simulations? l.16 New data is confusing. Is there additional data from Benin. l.19 Here you state that the model is calibrated to two different data sets. In results (table and figures) there is also the original parameterization showed. You should state that the new parameterizations are compared to the original one. l.16-23 Clarifiy. And there are some repeated sentences. l.24-25 What do you mean? l. 26 How do you sample the posterior when you are to produce the posterior? Do you mean prior?

Authors' response:
l.9-10: Data together with uncertainty are used as input of the model and the results are with confidence interval.
l.13: figures are results.
l.14: text changed to:
"The simulations of Yasso07 were compared with the observations…"
l.16: text changed to:
"Then, Yasso07 was fitted to a dataset where data from Benin were merged…"
l.19: add to line 23, the text:
"The new parameterizations were compared to the original one".
l.16-23: text changed to:
"Then, Yasso07 was fitted to a dataset where data from Benin were merged with global data including data from Europe (Berg et al., 1991a,b, 1993) and from Northern and Central America (Gholz et al., 2000; Trofymow et al., 1998). Yasso07 was also fitted only to data from Benin. This resulted in two new versions of Yasso07: Y07A and Y07B to refer to global + Benin data, and Benin data respectively. The predictions of these two versions were compared with the observations using mean residuals and standard deviations thereof and the new parameterizations were compared to the original one."
l.24-25:
l. 26: "We assessed the convergence of the Markov chains by checking the Gelman-Rubin diagnostics (Gelman & rubin, 1992) that compares the variance of the parameters within chains to that between chains. Although we only calculated three chains for each case, as the computational requirements of such samplings were rather high, we could easily see that the chains had consistently the same stationary distributions enabling us to state that the chains were indeed sufficiently close to convergence. This was also visually clear as the chains did not jump between different states but stayed around the MAP estimate at all times, which is indicative of good mixing properties of the chains."
Our samplings had typically acceptance rates between 0.2-0.3, which, we believe, is what the ref. #1 means by the question "what proportion of chains were accepted". This is a typical acceptance rate for the adaptive Metropolis algorithm in a case where the posterior is unimodal and resembles multimodal Gaussian density."

p. 3012

l.3-5. What do you use the Metropolis Hasting for? What are your moving chain criteria? The whole section of calibration needs to be clarified. l.19 Are these residuals for one particular species? During whole period? It is unclear.

Authors' response: please see responses previous comments.

p. 3013

l.5 You should present the model residuals. I is difficult to see the residuals from the figure. In addition the shape is of great interest to be able to check if the model is applicable or not. l.26 These sentences are contradictory.

Authors' response:
l.5: please see responses previous comments.

p.3014

l. 6-10. Something is missing in the sentence. l. 10 According to figure 4 it is around 150 mg g-1 for D glauca. Check that the numbers are correct. It is unclear if you mean the difference between the highest and lowest value or just the value. In any case the numbers are not the same from the figure: : :A-values are around 550 in beginning. l.16 which Yasso07. The original version?

Authors' response:
l.6-10: text changed to:
"The litter quality of species studied differed from the litter quality of *Drypetes glauca* (Gholz et al., 2000) the only tropical tree species whose data on litter quality (leaf, fine root) were included in the dataset used to initially parameterize Yasso07 (Tuomi et al., 2009)."

l.10: the concentration of *D. glauca* is not included in figure 4.

p. 3015

l.7-8 Reference.
l. 10 What are the problems of the model? l.18 Adjusted model or adjusted parameters?
Since you use the word Validity in title I assume that you do not change/adjust the model. l. 21 Uncertainty in estimates regarding Y07A and Y07B. Mean residuals?
It is crucial to have the residual comparison to be able to compare the results of the different parameterizations of the model. l. l. 17-23 Preformed similar. If they preform similar they should have similar parameters? Y07B had much larger decomposition rate, which should give much faster mass losses. This is not seen in the results.

Authors' response:
l.7-8: text changed to:
"It is likely that simple components, easily decomposable, for example waxes, ethanol-soluble, become resistant or give resistant products during the process (Nömmik and Vahtras 1982, Berg 2000)"
l.18: "adjusted model versions" to be replaced with "re-parameterized model versions"

p.3017

l.3 Is it the model that is refined or is it the models' parameters that are tuned?
l.8 Which Yasso07 model? A/B or both? Or the original version? In general in the
conclusions there is no conclusion whether the decomposition process in the tropics is

affected by the high variation of litter quality. l.20. I think this needs to be discussed more. Different litter types have different effect on SOC changes. This is of importance when assessing SOC changes.

Authors' response:
l.3: text changed to:
"This study showed that the decomposition process in the tropics is affected by the high variation of litter quality and provided evidence that the soil carbon model Yasso07 did not incorporate the chemical diversity of tropical litters, even if the climatic variation was accounted for in the structure of the model. It demonstrated that the model parameters can be refined by integrating available data from the tropics."

l.8: text changed to:
"This Y07A version is suitable for estimating …"

Manuscript evaluation Criteria:

4.
Are the methods and assumptions valid and clearly outlined? This should be improved in the next revision. The text is poorly written and there are some information missing that is needed to follow.

Authors' response:
Text clarified and missing information added. See responses to previous comments.

5. Are the results sufficient to support the interpretations and conclusions? Yes. But the core of the paper: that the variation in litter quality is the major driver of decomposition in the tropics, needs to be lifted more. This needs to be discussed further by for ex. Comparing the parameterizations found in the paper with old parameterizations. There is a very large difference in decomposition rates which should be high lightened after a thorough discussion and presentation of calibration results.

Authors' response:
Text to support that "the variation in litter quality is the major driver of decomposition in the tropics" is added. See responses to previous comments.

6. Is the description sufficiently complete and precise to allow their reproduction by fellow scientists (traceability of results)? In the case of model description papers, it should in theory be possible for an independent scientist to construct a model that, while not necessarily numerically identical, will produce scientifically equivalent results.
Model development papers should be similarly reproducible. For MIP and benchmarking papers it should be possible for the protocol to be precisely reproduced for an independent model.
Descriptions of numerical advances should be precisely reproducible. Mostly. There are some information left out for example, the model needs max and min temperatures, the part of calibration description is poor.

Authors' response:
Text on the use of climatic data in the model is added. See responses to previous comments.

7. Do the authors give proper credit to related work and clearly indicate their own new/original contribution?
Mostly, the discussion part lacks references. The last paragraph has some statements that needs references but include none.

Authors' response:

References added. See response to the comment.
9. Does the abstract provide a concise and complete summary? Needs language revision. And check that the sentences are linked. It is difficult to understand the conclusions from the information given in abstract. **Authors' response:** Text in abstract changed. See responses to previous comments.
11. Is the language fluent and precise? No. The manuscript needs language revision. Throughout the whole text there are many words missing in sentences, many repetitions etc. **Authors' response:** Missing words are added and repetitions corrected. The language will be revised in the paper.
14. Are the number and quality of references appropriate? More references in the discussion are needed. Especially when the authors state possible explanations of their data. **Authors' response:** References are added in the text.