Referee 1

The study presents a new model development and calibration to an interesting horizontally heterogeneous system. It is based on an impressive compiled data set of observations and derivations of relevant inputs and state variables to compare. The main conclusion is that the observed increases in SOM stocks in an agroforest system are due to higher litter input compared to an agricultural control. The modelling exercise is interesting to the soil modelling community, and to the community researching interactions of vegetation components and management. I state several main points followed by more detailed comments.

Response: We thank you for your interest in our work, we really appreciated your comments and suggestions. We tried to take into account all your comments and corrected the manuscript on the requested points.

1 Main points

1) One calibration aspect of the study that convinces me (and probably other readers as well) of the validity of the study is currently not well highlighted. The model was calibrated to the control plot only. Despite of simplifying assumptions on similarities in climate and vertical transport between the control and the agroforesty system, the model predicted the differing C-stocks in tree rows and alleys and its depth distribution well. This is a strong validation.

Response: Thanks for this rewarding comment. We better highlighted this result in the abstract "The model was calibrated to the control plot only…The model was strongly validated, describing properly the measured SOC stocks and distribution with depth in agroforestry tree rows and alleys". (P2L30-34), but also in the discussion part "Despite these simplifying assumptions on similarities in climate but also on vertical transport between the control and the agroforestry system, the model calibrated to the control plot was able to reproduce SOC stocks in tree rows and alleys and its depth distribution well. This strong validation also suggests that OC inputs is the main driver of SOC storage, and that a potential effect of agroforestry microclimate on SOC mineralization is of minor importance" (P40L732-737).

2) The quantification of the priming effect (PE) seems to be a bit complicated with running the no-PE model variant with a decomposition rate that was calibrated with the PE-model variant. To my opinion there are more straightforward quantifications already in the data (see detailed comments). I suggest highlighting the result that the priming model variant in Fig 4 was able to capture the depth distribution of C-stocks while no-PE model variant did not.

Response: We tried to better describe how the PE intensity was quantified (P25-27L548-594) (see below) and why we chose this calculation method. We also highlighted in the abstract the fact that only the PE model was able to describe SOC profiles "Moreover, only a priming effect variant of the model was able to capture the depth distribution of SOC stocks" (P2L36-37).

3) While the mathematical model is well described, information is missing on the solution of the forward model, i.e. the solution of the presented partial-differential equation given a set of parameters. Which method has been used? What was the spatial grid, the same grid as the measurements? Was this grid sufficient to represent the steep concentration gradient in the top soil? Have different grid sizes been tested?

Response: Partial-differential equations were solved using the R package *deSolve* and the *ode.1D* method (Soetaert et al., 2010) (P27L596-597). The spatial grid was as close as possible to the measurements. Due to some field difficulties, the sampling grid is not totally regular but the modelling grid is. We indeed implicitly assumed this resolution to be sufficient to represent the steep concentration but we did not deeply evaluated the effect of different grid size but if really needed we can provide an analysis in the supplementary material

4) To my understanding of the study, the increased C stocks at the walnut tree lines are explained in a big part to increase of the above-ground carbon input by the herbaceous summer vegetation between trees (Fig. 3). I would like to read some discussion on this point. Was there an organic layer?

Response: Yes, this is absolutely true, the herbaceous vegetation growing between trees in the tree rows plays an important role on SOC storage. This was very much suggested by previous works on SOC storage in these systems (Cardinael et al., 2015, 2017), but proven here with the quantification of OC inputs. We now discussed this point more into details: "The increased SOC stocks in the tree rows were explained in a big part by an important above-ground carbon input (2.13 t C ha-1 yr-1) by the herbaceous vegetation between trees. This result had already been suggested by Cardinael et al., (2015b) and by Cardinael et al., (2017) who showed that even young agroforestry systems could store SOC in the tree rows while trees are still very small. These "grass strips" indirectly introduced by the tree planting in parallel tree rows have a major impact on SOC stocks of agroforestry systems" (P41L758-764).

As commonly observed on grass strips, there was a very thin organic layer (maximum 0.3 cm thick), but not permanent during the season. Climatic conditions are very favorable for litter decomposition there, and we therefore assumed that this thin organic layer did not significantly change moisture and temperature conditions for the below mineral soil.

2 Detailed comments

L 412: Instead of interpolating parameters of several fits, I suggest fitting a single equation to the entire dataset with an additional variable "distance to tree" and parameters a and b depend on this distance. However, the simplified procedure here seems to work and this point does not affect the conclusions.

Response: Yes, this is indeed another possibility. As we were able to well reproduce root profiles with this simplified method, we think it is not really necessary to look for another equation as it would indeed not change the conclusions.

L 444: Please specify exactly which observations and which predictions have been used for calibration.

Response: We used SOC stocks measured in 2013 in the control plot (observations) and predicted SOC stocks (predictions) for the calibration. These stocks were considered at equilibrium (P25L551-553).

Table 7: The prior knowledge in eq. 19 was specified as normal distribution. Table 7 instead reports a range of values instead of a mean and a variance (xb and diagonal of Pb in equation 1). Moreover many ranges span several orders of magnitudes suggesting that the parameters should be log-transformed before estimation. Where does the variance of the posterior come from? And what is the meaning of "prior values" in the posterior column?

Table 7: Where did the prior information come from? Are these uninformative priors or does it affect the results if you take different priors?

Response: We acknowledge that this point was not clear enough. The optimization procedure that we used is sensitive to local minima. We therefore performed 30 optimization procedures starting with different parameter prior values to check that the results did not correspond to a local minimum. The prior range presented in Table 7 represents the range in which prior values were sampled for the 30 optimizations, it is therefore normal that they span several orders of magnitudes. The prior values presented in brackets in the posterior column represent the prior values that minimized the **J(x)** value. The variance of the posterior is based on Santaren et al., 2007 (GBC 21, GB2013). The BFGS algorithm does not directly calculate variance of posteriors. To obtain them, we quantified the variance using the curvature cost function at its minimum once it was reached.

We clarified it in the text: "To determine an optimal set of parameters which minimizes J(x), we used the BFGS gradient-based algorithm (Tarantola, 1987). For each model variant, we performed 30 optimizations starting with different parameter prior values to check that the results did not correspond to a local minimum. As the BFGS algorithm does not directly calculate the variance of posteriors, they were quantified using the curvature cost function at its minimum once it was reached (Santaren et al., 2007)." (P26L571-576), and in the Table 7 (now Table 5) footnote: "The prior range represents the range in which prior values were sampled for the 30 optimizations per model variant. The prior values presented in brackets in the posterior column represent the prior values that minimized the J(x) value (Eq. (34))." (P32L660-661).

Eq 21: Please explain the derivation. Usually the BIC = ln(n)k - 2log(L), which involves the Likelihood instead of the mean squared deviation. From a Bayesian perspective -2log(L) α Jdata(p) , where Jdata is the first term of J of eq. 19 (excluding the prior term).

Response: Here, we used the MSD to estimate the maximum likelihood. This is indeed not the classical BIC. This approach is similar to Manzoni et al., 2012 (SBB 50, 66-76) who used the residual sum of square to estimate the maximum likelihood. We rephrase to clarify: "where *N* is the number of observations, *MSD* is the mean squared deviation used to estimate the maximum likelihood, and *k* is the number of model parameters" (P26L585-586).

L 478: Please, clarify terminology of spin-up vs model calibration. To my understanding you calibrated 4 or 5 parameters depending on the three model variants so that equilibrium stocks, i.e. simulations after 5000 years, were close to observed C-stocks (n=?) of the control plot in 2013. I suggest putting this content to the calibration section.

Response: We moved this paragraph to the optimization procedure section and we clarified the terminology of spin-up vs model calibration: "These four or five parameters were calibrated so that equilibrium SOC stocks, i.e. after 5000 years of simulation, equaled SOC stocks of the control plot in 2013. The associated uncertainty was estimated with the 93 soil cores sampled in the control plot (see section 2.2.1). Due to a lack of relevant data, we assumed that the climate and the land use were the same for the last 5000 years, and that SOC stocks in the control plot were at equilibrium at the time of measurement. Therefore, SOC stocks at the end of the 5000 years of simulation equaled SOC stocks in the control plot. Three different calibrations were performed, corresponding to the three different models that were used: one calibration with the two pools model without the priming effect, one calibration with the two pools model with the priming effect, and one calibration with the three pools model" (P25L548-557). "SOC pools were initialized after a spin-up of 5000 years in the control plot. At t_0, SOC stocks in the agroforestry plot therefore equaled SOC stocks of the control plot" (P27L592-594).

L 508: This derivation of the effect of priming is hard to grasp. To my opinion its more straightforward is compare predictions of the PE-variant model versus the non- PE variant; each consistently calibrated and applied for prediction:

- *Effects of litter inputs: predictions of no-priming variant only: agroforestry stocks vs control stocks*
- *Combined effect: prediction of the priming model variant only: at agroforestry plot versus the control plot*
- *Effects of priming only: prediction of the priming model variant versus the predictions of the no-priming variant for the agroforestry system*

Since the profile was not matched well with the no-priming model one can focus on sums.

Response: We agree that the calculation was not straightforward and we clarified it in the new version (see below). Nevertheless, we consider our calculation as the most correct even though it is a bit complex to understand it. Indeed, we can not directly compare the different versions of the model to calculate priming because the decomposition rate of a classical first order kinetics takes implicitly into account a fixed fraction of decomposition due to priming. In all situations, there are regular inputs inducing priming and when we optimized the decomposition rate parameter in the control plot we implicitly represented this priming but at a fixed rate. Therefore comparing the different versions of the model would not estimate the priming in the agroforestry plots.

“Furthermore, at equilibrium state (i.e. when the input rate is constant) the decomposition rate of a first order equation (Eq. (6)) takes *PE* implicitly into account. Indeed, when FOC enters the system, there is an induced priming, a constant FOC input rate therefore induces a constant priming. This means that when we optimized the decomposition rate parameter in the control plot, we implicitly represented this priming but at a fixed rate. When FOC inputs are modified, due to the tree growth for instance, the PE intensity is modified and this effect cannot be represented by classical first order kinetics.” (P27L607-612).

“To estimate the change of SOC decomposition rate due to priming when trees are planted, the decomposition fluxes predicted by Eq. (7) $\left(-k_{HSOC,z}\times(1-e^{-PE\times FOC_{t,z,d}})\right)$ in the agroforestry plot must be compared to the fluxes in agroforestry plot using the decomposition from the control plot calculated by Eq. (7) with $FOC_{t,z,d}$ corresponding to the FOC inputs in the control plot. Thus, to calculate the importance of priming on SOC storage when trees are planted, we used the decomposition rates calculated following Eq. (7) in the control plot and we applied this decomposition rate to the agroforestry plot as a classical first order kinetics (without the FOC control, i.e. $k_{new}=k_{HSOC,z}\times(1-e^{-PE\times FOC_{t,z,d}})$ with $FOC_{t,z,d}$ fixed constant)”. (P28L616-625).

Fig 3: Please, note that the largest above ground input comes from herbaceous vegetation. Is this an important aspect for C-stocks of the agroforestry system?

Response: Yes, this is definitely an important aspect for C-stocks in the agroforestry system. We added the following sentence to the result section: “In the agroforestry plot, the largest aboveground OC input to the soil comes from the herbaceous vegetation, and not from the trees” (P28-29L636-637).

L698 (3.4.2): Please, remind the reader that C-stocks of the agroforestry plot were not part of model calibration (that used the control plot only) but are used here for validation.

Response: As suggested, we added the following sentence at the beginning of the section: **“**As a reminder, SOC stocks of the agroforestry plot were not part of model calibration (that used the control plot only) but were used here for validation” (P33L677-678).

Fig. 4: This is a nice demonstration of priming formulation being able to match the depth-shape. Although uncertainty of the mean (standard error) is low due to the high sample number, you may add the standard deviation across 93 measurements in order to get an impression of the variability.

I would like to see a figure, where C-depth profiles can be compared between cases without being dispersed across facets. Maybe zoom in to 5 to 15 stock range.

Response: Yes, the uncertainty of the mean is extremely low for measured SOC stocks, as suggested we instead added the standard deviation of measurements (P36L693-694).

Concerning the C-depth profiles of Fig 4., this was actually our first idea. But SOC profiles are extremely close, especially between the control and the alleys, and the figure was very messy. We would therefore prefer to stick to this presentation, which is much clearer, even if we have to compare different facets.

Fig. 5: Please, use a color scale with a clear zero.

Response: We changed the color scale as requested. We also added a 2D graph of modeled control and agroforestry SOC stocks (P37).

Fig. 6 Please, add difference in measured stocks to the "Inputs+PE" column for comparison.

Response: Thanks for this suggestion, it was done (P39).

L 753: Suggest: "Despite of these simplifying assumptions, the model calibrated to the control plot was able to ..."

Response: This sentence was changed as follows: "Despite these simplifying assumptions on similarities in climate but also on vertical transport between the control and the agroforestry system, the model calibrated to the control plot was able to reproduce SOC stocks in tree rows and alleys and its depth distribution well. This strong validation also suggests that OC inputs is the main driver of SOC storage at this site, and that a potential effect of agroforestry microclimate on SOC mineralization is of minor importance" (P40L732-737).