Dear Referee,

First of all, thank you very much for taking time to review this paper and for the valuable comments and suggestions that you have made.

Review of Biogeo Disc / van der Voort Radiocarbon is a valuable tool in attempts to understand the formation and turnover of soil organic matter, but the difficulty and expense of making measurements have meant that the available data are relatively few. Therefore the authors' idea to assess the variability and representativeness of 14C measurements is welcome. Broadly, they conclude that there is not a great deal of spatial variation, so that depth variations in 14C are fairly consistent from site to site, and results already available have therefore been reasonably representative. I do not think this is at all a trivial conclusion – it is an important finding. My main criticism of the work is that the variable chosen to represent soil carbon was SOC concentration (%), which is not directly relevant to SOM turnover. Better would be SOC pool (g/m2), which is the natural relative of turnover rate. Therefore I suggest that either the authors justify the use of concentrations, or they reanalyse their results using C pools.

→ Thank you for the positive feedback
→ We did not include the SOC pools originally because our bulk density (BD) estimates are not considered to be very precise (as they are based on single profiles taken proximally to the plot, see Walthert et al. 2002, 2003). However, based on this suggestion, we have incorporated available BD data and now values for SOC stocks have been incorporated in our statistical analysis (Table 5, 6). However, as the C stocks (gC/cm^3) are the multiplication of C content (gC/g soil) with BD (g soil/cm^3) and soil interval length, no new additional correlations emerge in our statistical analysis.

More minor comments: Line 21 Here the results for "topsoils" are claimed to have been reported. As far as I understand it by topsoil they mean the top of the mineral soil beneath the O (LF) layer. I question whether this really is topsoil in the sense of containing organic matter undergoing turnover and being intimately connected to ecosystem processes – in other words I reckon that the LF material is functionally important and should be counted as soil. If the authors do not agree, then some discussion would be welcome.

→ Indeed, by topsoil we mean the top of the mineral soil beneath the O(LF) layer. It is correct that the LF material is functionally important, and it was also incorporated in this paper, although it receives less focus than the mineral soil. It was measured in a number of locations (Fig. 3.) to get a better idea of variability. We calculated turnover for this layer (Table 2). We also did statistical analysis on it (Variability analysis in Table 2, Spearman correlation in Table 5).

It would be of interest to know for example how much C (g/m2) is in the LF layer, and how much in the 0-5 cm at the top of the mineral soil.

→ As mentioned previously, for the 0-5 cm layer we have bulk-density estimated from single profiles proximal to the plot where the samples were taken. For the LF layer we found data that we were previously unaware of collected on the same plot, and this enabled us to determine the carbon stocks (WSL LWF). These data are now included in Supplemental Table 1.

Moreover, the numbers quoted in the Abstract (Δ14C 159, sd 36.4) appear actually to refer to the LF layer (Table 2)!

→ this was corrected - thank you.

Line 26 This last line of the Abstract is weak, if there are "important consequences" you should say what they are.

→ This formulation indeed broad, as our conclusions that apply to models are two-fold. On the one hand, large small-scale heterogeneity could be incorporated in plot-scale or catchment-scale models (CENTURY, DAYCENT etc), whilst on the other hand, for large spatial scale (Earth system models) we have a relative homogeneity. Because of this, we would prefer to keep this statement generic and provide more specific details later on in the manuscript.

Materials and methods The dates of much of the soil sampling are stated to be "in the course of the 1990s" which could mean that some samples were collected 10 years apart. Other samples were collected in 2014. It therefore is not strictly correct to compare 14C values, since they are not constant with time in soil in situ – indeed that is why 14C is a useful variable, and why the data of sampling is an important qualifier of every 14C measurement. Maybe the analysis here would not be much affected by the assumption that the 14C values refer to the same point in time, but the issue should be acknowledged and the assumption justified – perhaps the MRT values are sufficiently long that a few years' difference in sampling date is of no consequence?

→ Thank you for this suggestion; this is absolutely correct. The measurements on these sites that have been done in 2014 are included in another paper (Van der Voort et al., in prep) for which the different atmospheric signal will be taken into account. This paper only concerns data only of the period 1994-1998. Because this was unclear, we have removed the reference of this sampling campaign.

Equation (2) This doesn't look right to me – the leading Δ14C shouldn't be there. Also, is it really necessary to apply the equation only to samples with a value of R>1, which is what seems to be stated in lines 124-5? And after reading further, I realise that I do not understand the difference between R and Fm.

→ Thank you for that remark, indeed there was a typo, which has been corrected.

Line 173 I don't see why the expressions "worst case" and "best case" are used here – the facts are the facts, we should not judge them.

→ Thank your for suggestion, we have adjusted the wording to incorporate this suggestion:

Line 199: "ranges from 50 ‰ (relative highest degree of variability scenario, Podzol) to 20 ‰ (relative lowest degree of variability, Cambisol) (Table 3).

Also, I do not fully understand what is learnt by showing that the variability of 14C correlates (or does not correlate) with variables like slope, MAP etc. This is not considered in the Discussion, yet the results for variation with clay and MAP appear as conclusions.

→ This is a good point, and required clarification. We have clarified this is the results section:

line 176: “The Spearman coefficient identified few significant correlations between $\Delta^{14}C$ in samples and climatic variables”.

Further details:

(1) This paper looks at the variability on different scales, such as the plot and regional scale (Tables 2, 3 and 4). We do compare it to slope in the text, as we considered this as a causal factor in the degree of variability.
(2) We then look at the correlation between $\Delta^{14}C$ and Slope, MAP etc. (Tables 5 & 6), because we would like to know if SOM dynamics depend on these environmental variables

I could not see any information about clay contents (e.g. why not in Table 1?).

→ Thank you for this suggestion, this has been added in Supplement Table 1.

Line 234 The word “marked” here is used rather carelessly. The values of MAP and MAT admittedly vary, but within fairly small ranges in a global context. And since the soil types and geologies also vary it can hardly be claimed that variations in the site attributes have been sufficiently covered – it might be for example that a trend in MAT counters one in MAP, or in soil type or in geology, or indeed in vegetation type (as far as I can see no information on tree species is provided, certainly not in Table 1) or NPP. Although the results are certainly of considerable interest, the fact that definite trends cannot be found does not mean that there are no trends

→ This is a valid point. We tried to eliminate the site-induced variability by inserting an nmle linear mixed-effect model and taking the Site/Core variable as the random variable. However, it is correct that this dataset has its limitations w.r.t. MAT and MAP range, and we explicitly tackle this by saying:

Line 369 “While the present observations remain limited in geographic scope, the relative homogeneity of $\Delta^{14}C$ signatures observed in surface and deep soils across climatic and geologic gradients implies that the rate of C incorporation may be similar and hence relatively insensitive to changing climate conditions”.

Line 288 Is it really necessary to incorporate “factors that drive small-scale variability” into larger-scale models of SOM turnover? Is it not possible that ecosystem complexity and the costs of analysis mean that the more complex models implied here are unachievable?

→ For global models, this point is absolutely valid. There are however plenty of plot-scale models (i.e. CENTURY, DAYCENT, YASSO), for which these factors could be taken into account. We have adjusted the wording to make this point clearer:

Line 380 “The latter is essential for the use of radiocarbon to assess carbon turnover and associated processes in forest soils, especially for plot-scale modelling.”