Biogeosciences Discuss., 7, C1551–C1578, 2010
www.biogeosciences-discuss.net/7/C1551/2010/

Biogeosciences Discussions

Interactive comment on "No change in topsoil carbon levels of Great Britain, 1978–2007" *by* P. M. Chamberlain et al.

P. M. Chamberlain et al.

paulmchamb@yahoo.co.uk

Received and published: 25 June 2010

We would like to thank all anonymous referees for their thorough and valuable comments on the content of our manuscript and their suggestions for improving the paper. We note that many of the concerns they raise are not just applicable to Countryside Survey, but to long-term and large-scale soil sampling in general. Here we reply point-by-point to the reviewers' comments.

Anonymous Referee #1

Referee 1: This paper addresses an important issue about which there has been significant debate and contradictory findings reported in the literature. The paper benefits from a very large and carefully designed sampling strategy and this is the main strength of the work. The work certainly needs to be published with too much delay, but there are some important details that need to be addressed. Although the authors' main finding that there is no significant change in topsoil soil C stocks in GB is in line with the emerging consensus, their data are not as secure as the title and conclusion would suggest for the following reasons:

1. As with the previous high profile work in this area (Bellamy et al.), the authors have been forced into using a circular argument to overcome the lack of bulk density measurements from some of the sampling occasions. It is not their fault that bulk density was not collected earlier, but this is an important short-coming of the work.

Chamberlain et al.: We agree that the lack of BD measurements in the 1978 and 1998 Countryside Surveys is a short-coming. In fact we identified this weakness ourselves when planning for the 2007 Survey, hence we made sure BD was measured in 2007. The manuscript is clear that BD was only measured in the most recent Survey, and clearly explains how estimates of C density for 1978 and 1998 were generated.

However, instead of relying on a literature-based equation to estimate BD, work in the 2007 Survey generated a new large dataset of measured BD values, which has enabled better estimates of topsoil C density for GB than previously available. This is a major step forwards, and the data is now available for other workers to utilise.

Referee 1: 2. It is unfortunate (verging on inconceivable) that two different protocols for loss on ignition and two different analytical approaches (LOI and elemental analysis) have been used for the determination of soil organic matter during the course of what is otherwise a well-designed study set up with the purpose of long-term monitoring. The authors make a good attempt to rectify this issue by re-analysis where possible.

Chamberlain et al.: This is not strictly correct. It is true that over the three Surveys, two different LOI methods were used to determine soil organic matter, and the effects of these differences have been accounted for as far as possible. We acknowledge that this is not ideal, but after re-analysis of archived 1998 samples, only 6% of samples across

BGD

7, C1551–C1578, 2010

the three Surveys were determined by the second LOI method (which was anyway subsequently corrected for). Note that the two methods resulted in LOI differences of <2%, a very small amount.

However, once LOI has been determined, values must be converted into %C, and for this we have consistently used elemental analysis. The alternative was to assume a relationship between LOI and %C, which we deemed to be inferior to actually defining a relationship based on CS data. Elemental analysis was used to determined %C for around 1200 samples from both the 1998 and 2007 Surveys, and the relationship between LOI and %C was the same in both years, which gave us confidence that we are able to convert all the LOI data to %C. The proven relationship between LOI and %C is not a weakness of CS, but a considerable strength.

Referee 1: 3. The authors go to considerable length to describe the sampling design and the relocation of the samples in the X and Y dimensions, but provide little detailed information about the Z (vertical/depth) dimension. They say that samples were taken to 15 cm with a towel in some case and with a corer in other cases. This difference in method, although apparently trivial, is a very important detail. I have done lots of soil sampling and am well aware of the lack of precision and accuracy that can creep in at that stage. Essentially, the biggest sources of error occur at the sampling stage because of inaccuracies in volume and depth estimation. So, the questions are:

a. Was 0-15 cm actually measured in each case?

Chamberlain et al.: The protocols for 1978 recorded that 0-15 cm depth of soil was sampled from the side of a soil pit. In 1998, a 15 cm plastic core was inserted into the ground until full, if possible. However, actual lengths of soil collected were not recorded. In 2007 the same protocol was used as in 1998, and soil in the core was measured on return to the laboratory. The 2007 data suggest that 60% of cores contained 140-150 mm of soil, 29% contained 120-139 mm soil, 8% contained 100-119 mm soil, and 3% contained less than 100 mm of soil. Therefore 0-15 cm was not measured in all

BGD

7, C1551–C1578, 2010

cases. This is not surprising, given that some soils are not 15 cm deep, obstacles such as stones and plant roots sometimes stop plastic cores being inserted to their full extent, and compaction of some soils as the tube enters the ground will make it appear that less than 15 cm of soil was sampled, when in fact the complete depth was sampled. Surveyors repeated attempts to gain a full 15cm as far as possible within the constraints of the wider survey and logistical requirements.

Referee 1: b. Were all depths equally represented in the sample, i.e. were the sides of the hole left by the trowel exactly parallel?

Chamberlain et al.: Trowels were only used in the 1978 Survey, for which there are no photographs available for us to refer to. However, the 1978 work was carried out by trained soil scientists, who were responsible for digging the soil pit, recording the soil profiles present, and taking the soil sample. Some of the main scientists who carried out the 1978 Survey were involved in planning, implementing and training for the 1998 Survey so such comparability issues were considered at the time. Similarly, the main scientists involved in the 1998 Survey passed on information and trained the 2007 surveyors.

Referee 1: c. What constitutes 0 cm? This question may seem a silly one but in some classical soil sampling and survey approaches, the top of the soil is considered to be the upper limit of the mineral-rich material. Partially decomposed plant litter which has accumulated above that is either discarded or collected separately.

Chamberlain et al.: The CS2007 Soils Manual (Emmett et al. 2008, referred to in the manuscript) states that the surveyors must ‘move vegetation and loose litter to gain access to the soil surface’, which is essentially what the reviewer describes. This is what was done in 1978 and 1998 as well. Thus 0 cm in CS is ‘the upper limit of the mineral-rich material’ in all three Surveys. A line has been added to Section 2.1 to make clear that further details of sampling can be found in the Soils Manual.

Referee 1: d. How have the authors accounted for shrinking and swelling of soil associated with soil wetness? The authors comment on the effect of wetness on soil volume. They are correct that when volume is not constrained the errors can be large as the soil will expand and contract in all directions. In the field however, expansion in the X and Y dimensions is constrained to some extent, but upward expansion and downward contraction are important considerations. This may mean that the mass of dry soil within a 15 cm depth soil will vary with soil wetness. The authors argue that the sample size is sufficiently large to nullify this effect, but this argument only holds if there was no significant difference in the antecedent rainfall regime across GB in the years the samples were collected.

Chamberlain et al.: We are aware that changes in bulk density associated with rainfall are possible, leading to differences in the mass of dry soil in a collected core. Moisture-driven changes in BD will effectively raise or lower the topsoil surface level, effectively changing the depth to which soils are sampled. The ideal way for this to be countered would have been for a metal plate to have been inserted into the soil at 15 cm depth in 1978, giving a datum from which all other change could be measured. However, this was not done in CS, and is not routine in any other surveys of which we are aware.

In 1998 and 2007 (the Surveys for which soil moisture data are available), mean soil moisture contents were 45.1 and 45.7%, which suggests that soil moisture was not significantly different in the two years, and hence for these years no correction should be necessary. We have no data for 1978 and so are unable to make such a comparison, and have made no attempt to correct for these possible moisture-induced changes for 1998 and 2007 due to small mean moisture change for those years.

Changes in BD due to moisture will feed into estimates of C density, and soils with higher moisture contents may have a lower apparent C density in the top 15 cm. However without additional information this effect cannot be quantified – such information will be obtained when CS is repeated, as a new LOI-BD relationship would be produced, and deviation from the CS2007 relationship may indicate moisture-induced changes in BD.

Referee 1: E.g. how have the authors corrected for the effect pounding a corer in to the soil will have had on soil compression?

Chamberlain et al.: Prior to the 2007 Survey, field work was carried out to determine the best way of measuring topsoil BD. This field work included testing two soil sampling methods; the first involved taking soil samples using a variety of differently sized and shaped plastic cores, while the second involved digging a pit, lining with plastic and filling with water to ascertain the volume. Five different soil types were tested: sandy clay, stony loam, sand, peat, and woodland loam.

The estimated bulk densities (which are unpublished) produced using both plastic cores and the pit method were almost all within the ranges of typical literature values of soil BD, and the plastic core method did not produce BD estimates consistently higher or lower than the pit method. It was therefore concluded that the plastic core sampling method yielded reasonable BD estimates across the whole range of soils encountered in CS. In light of this, we have not corrected for any compaction of soil due to the core being driven into the ground, as this previous work suggested that such corrections are unnecessary.

Referee 1: All of these questions are important details any one of which can contribute a 5-10% error in the estimate of soil C stock. Given that the differences discussed are of this order, detailed and accurate information is essential.

Chamberlain et al.: The reviewer is correct in asserting that these effects can contribute a 5-10% error in the estimates of soil C but only for individual samples. Errors in summary measures are not just determined by the errors in individual measurements but depend crucially on the level of consistent bias in the samples, and reduce with increasing sample size. The national and other averages reported in the paper would not be subject to anything like this level of inaccuracy unless the effect was consistent (i.e. in the same direction) for all samples. Since most of the possible errors discussed above will be random the level of error in the overall results will be small. Further-

BGD

7, C1551–C1578, 2010

Interactive Comment

more estimates of change will not be affected by such factors unless their extent varies from year to year and this is not generally the case. For example, measurements of soil moisture, as stated above, show no overall difference between sampling years for which we have data so errors in estimates of change due to moisture effects will reduce the significance of overall averages but not their size.

We are aware of the possible effects described by the reviewer but would point out that these are largely generic effects which affect all soil sampling projects, not just CS.

Referee 1: 4. I do not think I saw any comments on how or whether the authors corrected for inorganic C in the soils. Many GB soils were limed in the 1970s and 1980s and others are naturally calcareous. Small amounts of carbonate in the soil will have affected the C analysis in different ways depending on the method. The lower temperature LOI is unlikely to be affected, but the higher one might have been, and the elemental analyser approach will certainly detect inorganic C. Please can the authors clarify and comment.

Chamberlain et al.: There are few naturally calcareous soils sampled in Countryside Survey (in 2007, there were only 10 samples recorded as being in the calcareous grassland Broad Habitat, for example). The LOI method now used consistently in CS uses 375°C, low enough that carbonate will not decompose, as the reviewer suggests. However, the reviewer is correct that elemental analysis would detect carbonate-C. We did not pre-treat any of the samples with acid, but can see no evidence of carbonate-C affecting the elemental analyser results. For example, the relationships between LOI and %C for soil samples with pH < 7 (soils unlikely to contain carbonate) and pH > 7 (i.e. potentially carbonate-rich soils) are almost exactly the same (%C = 0.54LOI and %C = 0.55LOI, respectively). On the basis of these results, inorganic C does not appear to significantly contribute to the LOI values determined in CS, and do not therefore need to be accounted for it in analytical procedures.

Referee 1: 5. The arguments about how much C could be stored in the soils if they were

BGD

7, C1551–C1578, 2010

managed to maximise this attribute is statistically robust, but mechanistically rather naive. Well-established ecological and environmental factors (vegetation, climate, altitude etc) affect total C storage and these seem to have been adequately captured by the stratification. However, for mineral soils, the clay content is the key soil property which influences total C storage and this has not been considered.

Chamberlain et al.: Clay content has not been measured in Countryside Survey to date; we simply could not measure everything of interest. While we are aware that our estimates are 'mechanistically naïve'; we aimed to provide them so that some upper limit is placed on the potential of GB soils to sequester and store C. However, since both reviewers 1 and 2 wish them to be removed from the paper, they will be.

Referee 1: 6. On a stylistic matter, even though the amount of data is large, the number of table and figures ((and 8, respectively) seems excessive for what is a relatively simple message.

Chamberlain et al.: We acknowledge that the paper is long and does contain a large number of tables and figures - and that this number has gone down on the basis of the referees' comments. However, we are also aware of the scientific and political importance of this work, given that the only other major study of soils in this geographical area (Bellamy et al., 2005) came to a very different conclusion (i.e. that C was being lost from some topsoils at a huge rate). Given the prominence and effects of this previous report, and the contrasting findings of our work, we have had to explain very carefully and clearly what we have done and why our results are so different. The level of detail in the reviewers' own comments on our manuscript emphasises the scrutiny this work is undergoing. A paper of this length is therefore necessary in order to substantiate and explain our own results, and to lay our methods and our working before the scientific community.

Anonymous Referee #2

Referee 2: The manuscript "No change in topsoil carbon levels of Great Britain, 1978-

BGD

7, C1551–C1578, 2010

2007" will be a very important contribution to the discussion of the role of soil carbon for the terrestrial carbon balance and is in the scope of BG. It's a unique data set with repeated sampling in 1978, 1998 and 2007 on 635 plots (all three surveys) and more than 2500 additional sampling plots. It is written very well even though it needs some considerable shortening and pruning especially in the discussion section and for the tables and figures. The data analysis and statistical modeling seems to be done carefully using sophisticated mixed model approaches in combination will traditional statistics. The uncertainty of the dataset and the estimated soil C stock change in the UK has been discussed in his manuscript. However, some major uncertainties and unexplained results remain and should be further addressed.

Referee 2: 1.) Additional information on the soil sampling should be provided: Has there been only one sample per sampling plots? Has this sample been at least a composite sample? Has the forest floor or other organic layers been excluded or included?

Chamberlain et al.: Yes, there was only one sample per plot taken. It was not a composite sample; for our reasons for this and the scientific robustness of our approach, see our comments to Reviewer 3.

For the removal of organic layers, see the comments to Reviewer 1.

Referee 2: The increasing soil C densities from 1978 to 1998 (mainly in woodlands) and thereafter decreasing C densities are insufficiently explained, which puts a question mark to the whole data set.

Chamberlain et al.: There were only a few significant changes in C densities, but more in C concentration. We wonder if the reviewer means C concentration, in which the biggest changes certainly were in woodlands? We do not have evidence for a causal driver for the increase in C concentration in the period 1978 to 1998, and then the decrease after that. These are often difficult to pin down in complex surveys and there are dangers of correlative associations. However, we have examined possible reasons for the woodland results in the manuscript (page 2278 lines 1-9). The fact that we

BGD

7, C1551–C1578, 2010

Interactive Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

cannot sufficiently explain these differences does not put a 'question mark' over the whole data set but rather suggests factors at work which we (and others) have not as yet considered.

Referee 2: It should be estimated if this up and down or accelerated C loss maybe explained with sampling and analytical errors and biases (e.g. by using different methods at the different surveys).

Chamberlain et al.: We have accounted for all differences other than the change in soil sampling method between 1978 (pit, bag) and 1998/2007 (core). Any changes between 1998 and 2007 cannot be the result of changes in methodology, because there were none. The sampling method difference between 1978 and 1998/2007 does remain; however both sampling methods used a similar approach (removal of loose organic layers and vertical sampling of the 0-15cm layer), and it is difficult to see how this could produce both the overall differences and the differential effects and patterns observed across the different habitats.

Referee 2: Especially in land use systems with steep C gradient with depth such as forests and grasslands, a small error in the sampling depth or not similar separation between mineral soil and forest floor may lead to a bias in the whole data set.

Chamberlain et al.: Sampling depth remains an issue – see our comments to Reviewer 1.

Referee 2: If age class effects and harvesting are considered as reason for increasing densities (p.2286 l. 20) data on forest harvest should be shown to substantiate this explanation. I can only agree and repeat the comment 3 of reviewer 1 on the sampling procedure – it should be taken seriously.

Chamberlain et al.: Since such data are not available for the CS plots, we are not in a position to be able to demonstrate that changes in forest ages and harvesting practices are responsible for our observed changes in topsoil C. However, it is valid to

BGD

7, C1551–C1578, 2010

Interactive Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

suggest possible mechanisms for the changes we report (for example, just as Bellamy et al. (2005) did in proposing climate change as a possible cause of the changes they reported). This seems to us to be reasonable in the discussion section of a paper, and the papers we reference are both relevant and thorough treatments of the subject. We have therefore left our original comments intact.

Referee 2: 2.) A chapter “Causes of change in topsoil C concentration and density” in a paper with the title “No change in topsoil carbon levels of Great Britain, 1978-2007” does not really fit. I agree with reviewer 1 that the title should be changed taking into account the uncertainty which is inevitable combined with the presented data set.

Chamberlain et al.: The point is taken. We suggest changing the overall title of the paper to “No evidence of large-scale changes in topsoil carbon levels of Great Britain, 1978-2007” and the chapter heading to “Cause of change in topsoil C concentration and density within vegetation classes”.

Referee 2: Even though it is a unique dataset, the failure to detect significant soil C density changes can be also a result of the relatively weak sampling scheme (missing replications, missing bulk density for 1978 and 1998) than of missing regional scale C density changes.

Chamberlain et al.: Although C density is important, it will always be an estimate based on the combination of C concentration and bulk density. The reviewer is correct that missing BD for the earlier Surveys is a problem, but we have sought to address that issue as much as we are able. In doing so, we have produced BD data for one sampling which covers a breadth of soils never before reported at a national scale. However we consider that it is the CS topsoil C concentration results which are even more important than C density, since they are direct measurements made on the soil samples, and it is on changes in C concentration that previous reports of losses of soil C (e.g. Bellamy et al., 2005) have been based.

We are not sure what the reviewer mean by ‘weak sampling scheme’. A recent review

BGD

7, C1551–C1578, 2010

of sampling strategies compared two model-based schemes (grid and optimised grid) (as deployed in National Soil Inventory/Bellamy et al. 2005) and two design based schemes (stratified random and clustered stratified random). The stratified random scheme (as deployed in CS) was found to be the most suitable option for the specific questions being addressed, particularly in terms of the assessment of status and change in soil organic carbon (Black et al., 2008). Also, missing replications over time are not the result of a ‘relatively weak sampling scheme’, but are the result of the reality of repeated country-scale soil and vegetation surveys, in which plots are lost because of (for example) access restrictions and urbanisation, and plots are gained because the size of the Survey expands with changes in funders’ priorities and scientific considerations. No long term Survey such as CS is immune to such changes, and the advanced and robust statistical approach to this increase in samples over time has been praised by an independent statistical advisor recently requested to review the CS and NSI approaches, and was considered not to be the source of differences in results.

Referee 2: 3.) Chapter 3.6 should be completely omitted together with p.2268 l.22-25 and p. 2289l. 23-25. The maximum C density of a land use class is a week indicator for a potential to increase C densities. Maximum soil C densities are mainly constraint by environmental parameters such as the clay content, water availability and productivity of a site and only partly by human management. In order to estimate the management impact on soil C densities more data on the recent and historical management practice are needed. These data seem to be not available in this soil survey. Thus, any speculation on the manageability of soil C stocks lacks an adequate data source.

Chamberlain et al.: These passages have been removed, since both Reviewers 1 and 2 wished them to be so.

Referee 2: Further comments and suggestions:

p. 2272 l.6: How was the volume and plant material (roots?) determined?

BGD

7, C1551–C1578, 2010

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

Chamberlain et al.: The volume of non-sieved debris (largely plant roots and stones) was determined by displacement of water. A sentence to state this has been added to the methods section.

Referee 2: p. 2272 l. 18 ff. and p. 2273 l. 2 ff: The function used to estimate C densities from C concentrations including its error should be displayed.

Chamberlain et al.: The following text has been added to the manuscript, and satisfies the request from the reviewer:

The functions used to estimate C density and its sampling error from LOI were purely empirical and derived from the 2007 data. In each case a cubic polynomial was found to provide the best fit to the data. C density was estimated using the equation:

Log(C density) = 0.0414logit(LOI)3 - 0.1491logit(LOI)2 - 0.0146logit(LOI) + 0.9577 (Equation 4)

and its error by

Variance(C density) = -0.000055041LOI3 + 0.004878LOI2 + 0.14404LOI - 0.0906 (Equation 5)

The log transformation was used in Equation 4 to avoid estimates below zero and the logit transformation was used to improve fit at low and high levels of LOI. Simulated sampling errors were generated and applied on the log(C density) scale as this was the scale at which data were approximately normally distributed. Logarithms are to base 10. Note that the functions use LOI not C concentration, a minor point given that LOI and C concentrations are simple multiples of one another.

Referee 2: Equation 3: The unit on the left site of the equation does not fit to the unit on the right site (left site: mass per area; right site: mass per volume).

Chamberlain et al.: This is incorrect. The left hand side is mass per unit area in the top 15cm, which makes it a volume. The use of the word ‘topsoil’ frequently throughout the

BGD

7, C1551–C1578, 2010

text is intended to remind readers of this.

Referee 2: Chapter 2.6: The reporting categories (BH, AVC) are not in line with the stratification classes of the CS (ITE). Please explain this shortly.

Chamberlain et al.: The CS stratification system (the ITE Land classification) is a purely statistical device created to increase the efficiency of estimation. It was not intended to be used as a reporting tool. The CVS vegetation classification, of which the AVCs are the high level components, was, in contrast, specifically created from CS data as a reporting and classification tool. The BH classification was developed, some considerable time after the first CS, as part of the UK Biodiversity Action Plan. The Broad Habitats are the framework through which the Government is committed to meet its obligations for monitoring in the wider countryside and is used for reporting in CS specifically because of that.

Referee 2: p.2278l. 28 ff: You conclude that land use change does not play a role for soil C stocks. However, at p. 2281 l.11 ff an effect of land use change on soil C densities is assumed: You write that the areas that are converted from one land use class to another are expected to adjust to a new soil C density level. Maybe this needs some clarification.

Chamberlain et al.: The reviewer is right to point out this potential discrepancy. We have changed the first statement (page 2278 line 28ff) and it now reads:

Trends in these plots (Table 3) were broadly consistent with those observed for the whole dataset. This could suggest that shifts in AVC (i.e. land use change) are unlikely to be a major factor in the observed changes in soil C concentration, but could also arise if the majority of plots that change AVC have sufficient time between surveys to adapt to their new status.

Referee 2: Page 2282 l. 21-25 should be shifted to the introduction or omitted.

Chamberlain et al.: They have been omitted.

BGD

7, C1551–C1578, 2010

Referee 2: Chapter 4.2 and 4.3 should be pruned considerable or omitted. Most points raised here are repeated somewhere else.

Chamberlain et al.: The text from Chapter 4.2 has been removed. Chapter 4.3 has been retained, but all references to Bellamy et al. (2005) have been removed from this section (see the instruction below). Chapter 4.3 continues to compare our results with those of other W. European studies (with the exception of Bellamy et al. 2005), since it is important to put our results in this wider context.

Referee 2: All comments on the study of Bellamy et al. (2005) should be combined in the chapter 4.6 in order to structure the comparison with this study. The following party should be merged into this chapter: p. 2283 l. 17-20, 2284 l. 1-5, 2284 l. 12-28, p. 2285 l. 1-7, p.2289 l.5-6.

Chamberlain et al.: All comments on Bellamy et al. (2005) are now in Chapter 4.4 (the chapters have been renumbered), as requested.

Referee 2: Table 2 and 3 may be omitted, since very similar results are displayed in table 5 and 6. It seems to be enough to display densities since you used one function to derive C densities from C concentrations. The concentration data may be added in an appendix.

Chamberlain et al.: We have not done this, because C concentrations are our actual measurements – they are what each CS has determined on all soil samples. Hence we consider C concentration to be the fundamental measurement we make, and as such it should not be relegated to an appendix.

Additionally, the C density values reported here cannot be simply derived from C concentration since the scatter observed in 2007 was added into density estimates in all three years (page 2272 lines 21-27). Equations 4 and 5 (above, and in updated manuscript) link C concentration and C density, but not simply.

These considerations have led us to retain both C concentration and C density in the

BGD

7, C1551–C1578, 2010

paper, and put neither in an Appendix. We hope this is acceptable.

Referee 2: Table 6 and Figure 7 contain almost the same information and data. One of them should be omitted.

Chamberlain et al.: Figure 7 has been removed.

Referee 2: Large parts of Table 7 and 8 are redundant. Both tables should be merged.

Chamberlain et al.: Table 8 has been removed anyway, since Reviewers 1 and 2 wished the estimates contained within it to be removed from the paper. We have kept able 7 as it was, because we do not consider that any of it is redundant – it sets out our working so that readers can see how changes due to changes in area and C stock have been apportioned. CS provides one of the most accurate estimates of changes in habitat areas (used by many other organisations) and not to combine these area changes with the soil C results would be a major lost opportunity.

Referee 2: In table 9 you may want to include additional national surveys as they are reported from Ciais et al., 2010 (Global Change Biology, Vol. 16/5, page 1409-1428) and the model derived soil C stock change estimated in this publication.

Chamberlain et al.: Thanks to the reviewer to alerting us to this paper, the data from which has been added in to Table 9.

Referee 2: Figure 4 should be omitted since it is mainly are visualisation of the information that are already included in the tables.

Chamberlain et al.: Figure 4 is not a visualisation of data in the Tables, it shows differences in C concentration in individual repeat plots, while the tables show the estimated means for the whole dataset using the statistical model. It is important that differences in individual plots are shown, as the large differences observed in a small number of repeat plots appears to be a feature of country-scale soil sampling (the same pattern is seen in Lark et al. (2006) for the NSI). Including this figure enables the reader to understand the contents of the CS soils dataset, and is part of our attempt to communicate

BGD

7, C1551–C1578, 2010

very clearly how the CS data were collected, what the dataset contains, and how we have analysed it.

Referee 2: Figure 7 should get the legend of the vegetation class into the graph like it was done already for “Crops and weeds” and for “Fertile grassland”. That makes figures much easier to read.

Chamberlain et al.: Figure 7 has been removed, at the request of this reviewer (see above).

Referee 2: Figure 8 is already included in Table 8 and thus, can be omitted.

Chamberlain et al.: Table 8 has been removed at the request of Reviewers 1 and 2, so Figure 8 has been retained.

Anonymous Referee #3

Referee 3: The main conclusion is that that there has been no net change in topsoil carbon in Great Britain in recent decades, in apparent contradiction to the findings of the National Soil Inventory (NSI) of England and Wales, reported by Bellamy et al., and much of the paper is devoted to speculation on the reasons for the differences between the two studies. There are many possible reasons for the differences, as discussed in the paper and further below. To understand them properly requires a full statistical comparison.

Defra has recently commissioned a project (Defra Project SP1101) to make such a comparison, involving representatives of the Countryside Survey (CS) (including the second and third authors of this paper) and the NSI, and an independent statistician. The project will report in September this year. It seems highly premature to publish the current paper, with its rather speculative discussion, before that study has reported.

Chamberlain et al.: We disagree. For such a significant and policy-relevant topic it is clearly important for the scientific community to see and consider all the evidence as soon as possible as currently only one source of evidence (that of the NSI) is available to the community. While welcoming the Defra (Defra is the British Department for Environment, Food and Rural Affairs, which co-funds Countryside Survey and the NSI) project as one avenue for exploration it is entirely possible that the differences in findings between the two surveys are not statistical in nature and that the Defra project will not be able to resolve the issue. Indeed the independent statistician brought into examine the data from both Surveys, quickly concluded the statistical design differences between the two studies cannot explain the different results and reported as such to Defra at the initial meeting. Indeed in another study funded by a range of organisations comparing the statistical design approaches of CS and NSI (amongst others), a stratified random method (as employed by CS) was found to provide "the most suitable option for the specific questions being addressed, particularly in terms of the assessment of status and change in soil organic carbon" (Black et al. 2008). Also as part of the Defra project, re-analysis of stored sampled by NSI has been requested, as has re-analysis of their results using the CS approach. While all this is ongoing, the researchers who hold the NSI data continue to publish papers using them (e.g. Kirk and Bellamy, 2010). Furthermore, we believe discussion by the scientific community may throw up further relevant areas or data for consideration and this should not be delayed, particularly as there already has been much debate about the original NSI study results, aspects of which have been questioned by the scientific community and published. Public debate in the refereed literature is surely the most open way forward?

Referee 3: Specific points Page 2268 line 21 and page 2269 line 23

It is misleading to highlight that the losses in the NSI were up to 50%. Only a small number of sites had such losses and most were far smaller. From the spread of changes shown in Fig 4, evidently significant numbers of the CS sites also showed losses (and gains) of this magnitude.

Chamberlain et al.: We were not referring to specific sites within the NSI, but rather highly organic soils in their results generally. The equation given in Bellamy et al. (2005) for the relationship between original C concentration (Corg) and the rate of

BGD

7, C1551–C1578, 2010

Interactive Comment

change is

Rate of change in Corg (g kg-1 yr-1) = 0.6 – 0.0187 x original Corg (g kg-1)

For a soil originally containing 400 g C kg-1, the rate of change would be -6.88 g kg-1 yr-1, which over 25 years (the period of the survey reported) would be a total loss of 172 g kg-1, or 43%. The value of 50% we referred to was based on the relative rate of change, and we acknowledge that this is not the same thing. However, losses of 43% of C in organic soils are still very large indeed. The text has been modified to quote this value of 43%.

Our comment therefore is merely the extrapolation of the average annual losses reported by Bellamy et al. (2005) to the entire 25 year period of their survey, highlighting the values for organic soils because that is where their results most differ from our own. While large differences do occur at individual locations, most probably through relocation error (see below) and soil heterogeneity, in both CS and NSI data the attribution of such large losses to whole categories of soils is unexpected and we feel deserves comment, particularly since similar results were not found in CS. If, as the reviewer says only a small number of NSI sites had such losses, it is even more difficult to see why the results are reported as applying to substantial groups of soils.

Referee 3: Page 2269 lines 21-25

Bellamy et al. did not conclude climate change was the major cause of the losses measured in the NSI, they merely suggested a link. Kirk and Bellamy (2010) show that the major causes were past and continuing changes in land use and management, with a small contribution of climate change.

Chamberlain et al.: Acknowledged. We have replaced the word ‘conclusion’ on line 25 with ‘suggestion’. Kirk and Bellamy (2010) had not been published when we submitted our manuscript, but a reference to it has been added to the revised manuscript.

Referee 3: Page 2270 lines 12-19

Here it says 5 samples were taken from 5 “segments” of each of 256 plots in 1978 then each plot was resampled in 1998 – were the same locations taken within each “segment”? How were the segments defined?

Chamberlain et al.: There were 256 1km survey squares in 1978, not plots. In Countryside Survey, when a square is first visited by the surveyors, they divide the square into five segments and make sure that one of the main CS vegetation plots, from which soil samples are taken (and many other measurements are made), is in each of the segments. The segments are simply a way of making sure these plots are spread throughout the CS square, and once plots are first sampled their location is fixed. The segments are thereafter ignored. Thus in 1978 there were 1197 main vegetation, hence soil, plots spread across 256 squares.

Where a plot is recorded as repeated, the same location in the square was visited, and a soil sample collected 2-3m away from the previous sample location. See Emmett et al. (2008) for the precise protocol the surveyors use.

Referee 3: In Table 1 it says 277 plots were sampled in 1978 and 160 in 1998. Why are these numbers different? And how did only sampling 160 plots in 1998 give only 100 less samples in 1998 if only half the plots were resampled. Line 19 says 560 plots sampled in 2007 but table 1 says 1629?

Chamberlain et al.: The reviewer has misunderstood the structure of Countryside Survey, in particular the difference between a CS 1-km square and a plot within a square. Hence line 19 states that 591 1-km squares (not 560 plots) were sampled in 2007. This yielded a total of 2614 plots from which soils were returned (591 squares each with 5 plots, with a small number of samples not taken for various reasons). The difference between a CS-square and a plot is made clear on page 2270 lines 8-14.

In addition, the reviewer has misunderstood Table 1, which shows the number of plots which were sampled in 1, 2 or 3 Surveys, and is described on page 2270 lines 21-26. Thus, 277 plots were sampled only in 1978, and 160 were sampled only in 1998.

These plots have only been visited and sampled once across the 3 Surveys which have sampled soils. There are other plots which have been sampled twice (represented by the lines ‘1978 and 1998’, ‘1978 & 2007’ and ‘1998 and 2007’) in Table 1, and other plots which have been sampled 3 times.

CS is repeated around every 8 years and, although a small proportion of plots are lost each Survey (for reasons listed on page 2270 lines 21-24), each Survey has tended to be larger than the one before, as new plots are incorporated from new squares. This loss and gain of plots is simply the reality in a Survey such as CS, which has now been undertaken five times in total (soils have not always been sampled).

To make this clearer, we have added the phrase ‘Plots sampled in:’ to Table 1.

Referee 3: Page 2270 lines 17-18

What proportion of the plots was re-located by maps and what proportion by markers? Presumably it would not be possible to achieve a re-location accuracy of 2-3 m (page 2287 line 12) with maps.

Chamberlain et al.: We have no information on the proportions of methods used to re-locate plots. Plots are permanently marked as well as having their locations recorded on surveyor-drawn maps together with reference to local landmarks. Only where markers have been lost have the maps and plot descriptions been used on their own. Furthermore plots have been deleted from the database or treated as unrepeated if relocation was not considered to be of sufficient accuracy. The advent of GPS has also much improved the ability to relocate plots with accuracy in recent surveys, and it is our experience in CS that where necessary relying on well-drawn and annotated maps, it is possible to accurately (2-3 m) return to plot locations.

Referee 3: Page 2270 line 26-page 2271 line 1

If the plot locations varied over time, what does a re-location accuracy of 2-3 m (last point) mean?

Chamberlain et al.: The location of any given plot has never changed. The relocation accuracy refers to the accuracy of returning to plots which have been visited over more than one survey. Obviously if new plot locations are used, there is no relocation accuracy as they are new. Plots which were lost (for example due to urbanisation) were sometimes replaced elsewhere but such replacements are always treated as new plots. To avoid ambiguity, we have changed 'Plots were re-located using maps and/or markers placed in previous Surveys' (page 2270 line 17) to 'Surveyors located plots visited in previous Surveys using maps and/or markers placed during those Surveys' and have replaced "the changing number and location of plots" (page 2270 line 26-page 2271 line 1) by "the changing number of plots". The use of plots in CS and the reasons for changing numbers and locations are clearly stated in Section 2.1.

Referee 3: Page 2271 lines 1-4

The difference in sampling method will affect the results; presumably sampling by trowel from the side of a soil pit would not give a uniform sample over 0 -15 cm.

Chamberlain et al.: See our response to Reviewer 1.

Referee 3: But a more important limitation is that only one sample was taken per segment. There is large short-range variability in soil carbon, even at a range of 2-3 m. This severely constrains the reliability of the estimate of the "plot" carbon content. By contrast, in the NSI, 25 cores were taken at each site on a 20-m by 20-m grid centered on the target point and bulked for analysis, and the same procedures were used for both samplings. The problems with short-range variability and re-location error are evident in Fig. 4 of the paper. What is notable about the CS histograms of change are the heavy tails which show that the CS values for change are subject to substantial re-sampling error (e.g. if a peaty patch of soil in the plot was sampled in the baseline survey, but not at re-sampling, or vice versa). Data with a histogram like that should be analysed using robust statistical methods (e.g. Winsorized means).

Chamberlain et al.: It is difficult to answer this comment because it appears to be based

BGD

7, C1551–C1578, 2010

Interactive Comment

on a misunderstanding of the nature of variability and of re-location error. The reviewer is confusing the implications of the variability of individual measurements with that of the variability of estimates derived from sets of measurements. Bulking samples from the same location, as in the NSI, will of course increase the precision of individual measurements but this is only useful if individual locations are of interest in their own right, which is not the case with CS. In terms of comparison of groups of sites, improving the precision of measurements at individual locations has a very limited effect, governed by the ratio of between-to-within-location variance. The absolute maximum theoretical improvement in precision obtainable in this fashion is 30% and in practise is highly unlikely to be even as much as half this for soil carbon data. For the same number of locations therefore there will be little difference in discriminatory power between CS and NSI data. In any case bulking, because it relies on mechanical rather that mathematical averaging, may not fulfil its potential to increase variability if mixing of samples is inadequate. With regard to relocation errors bulking using samples from a grid, as in the NSI, will not eliminate or reduce the effects of such errors when there is a gradient in soil properties and the relocation errors of the NSI data are much greater ($\leq$20m or $\leq$50m on open land) than for CS ($\leq$3m). As the reviewer says above some NSI plots showed losses of carbon of up to 50% (see Lark et al. 2006 Fig. 2), a level of C difference as great as in CS, suggesting that the bulking of samples does not address this problem.

Winsorization and similar robust techniques were devised to protect measures of central tendency against outliers in data and not to analyse non-normal data such as those with heavy tails as in Figure 4. Relocation errors will, by definition, be symmetrical about zero and as such will increase the variability of estimates but will not introduce bias. Winsorized means, however, are biased unless the distribution of sample values is symmetric. Furthermore techniques such as Winsorization which affect the extremes of distributions are inappropriate when looking for effects occurring in those extremes. For soil C the suggestion from the NSI data is that soils with the most carbon are changing most, i.e. an implication for the extremes of the C distribution. In any

BGD

7, C1551–C1578, 2010

Interactive Comment

case the bootstrapping approach used by CS to estimate significance allows for such non-normality of the data.

Referee 3: Page 2271 lines 16-18

The use of regression predictions in place of a quarter of the baseline data needs to be taken into account in the subsequent analysis. If, as is likely, the "missing" 26% of the 1998 values were distributed substantially differently from the others (e.g. disproportionately on mineral soils), then the robustness of the method is compromised. More importantly, the variability of these 26% of observations will be smaller than the real data, and this will have inflated the significance level. This should be accounted for in the bootstrap re-sampling.

Chamberlain et al.: The reviewer's comments are not correct. We suspect that they are thinking in terms of prediction of one year's values from another which is not the case. The adjustment (rather than prediction) is derived from two different LOI measurements on each of the 76% of 1978 samples which contained sufficient soil to do this. This calibration set included samples from the complete range of CS soils and the resulting adjustment works equally well for samples distributed like those for which it is used as for the remainder. It is robust to the sample distribution. The very small residual variation about the calibration equation means that the variability of the adjusted observations is not smaller than the unadjusted so that significance levels are not inflated.

Referee 3: Page 2287 line 12

Relocation distance in the NSI was $\leq$ 20 m (equal to the side of the plot sampled) on enclosed land and $\leq$ 50 m on open land.

Chamberlain et al.: The text has been modified to make this point clearer.

Referee 3: Page 2287 lines 15-16

The statement that Walkley-Black is not appropriate for measuring soil carbon in soils

BGD

7, C1551–C1578, 2010

Interactive Comment

with > 80-150 g kg-1 is misleading. All analytical methods have their pros and cons. Walkley-Black is more accurate than LOI across most of the range of soil carbon contents, but it becomes less accurate at large carbon contents, greater than approx. 150 g kg-1. The paper by Vos et al. cited does not appear in the references.

Chamberlain et al.: The Walkley-Black method is not appropriate for measuring soil carbon in soils containing >80-150 g kg-1, because at such concentrations the small amounts of soil required are unrepresentative. Although we did not specify why Walkley-Black is not appropriate, both the papers we reference do so, and both recommend using other methods for such soils. It is therefore unclear why our statement is misleading, and we have left it in.

The paper by de Vos et al. is in the references.

Referee 3: Page 2287 lines 16-17

The statement that a mixture of Walkley-Black and LOI was used in the re-sampling of the NSI is misleading. In the first and second phases of the re-sampling Walkley-Black was used for all samples. In the third phase (land uses other than arable and managed grass), soils with OC > approx. 150 g kg-1 were analysed by loss on ignition (LOI). This involved less than 6% of the samples. In Bellamy et al. it is mistakenly stated that all soils with OC > 150 g kg-1 in both samplings were analyzed by LOI. The editors of Nature declined to publish an erratum note on this because they considered it did not seriously mislead readers; it is however reported in Lark et al. (2009).

Chamberlain et al.: The text has been changed to refer to the Lark reference. However, the fact remains that a) Walkley Black was used for highly organic soils in the original survey when it is not recommended that this method is used for such soils, and b) that both Walkley Black and LOI were used for the NSI resampling.

To make this clearer, the relevant sentences now read:

The two studies also differ in their methods of SOC determination; CS used a consis-

BGD

7, C1551–C1578, 2010

Interactive Comment

tent LOI method in all years, whilst Bellamy et al. (2005) used Walkley-Black determinations for all samples in their original survey (Lark et al., 2009; despite Walkley-Black methods not being appropriate for C concentrations >80-150 g kg-1; Bellamy et al. 2005; de Vos et al. 2007), while their resurvey used Walkley Black for soils containing <150 g kg-1 and LOI for all other soils.

Referee 3: Page 2287 lines 24 -page 2288 line 18

The suggestion that the change in lab methods in the NSI resulted in major systematic errors is wrong. The change in method affected less than 6% of the samples, and only those with OC > 150 g kg-1. So this cannot have caused major errors. For the soils analyzed by LOI, the maximum C content was 500 g kg-1 because the LOI conversion factor was 0.5, as Chamberlain et al. say. Hence the upper limit of 500 g kg-1 for the re-sampled soils in Fig. 2 of Lark et al (2006). However, this only affected 6% of the soils, and so had no major bearing on the overall results. Only 5% of the soils in the original sampling (measured by Walkley-Black) had C contents > 300 g kg-1 (see Table 1 in Bellamy et al.).

Chamberlain et al.: We note that the reviewer does not address our central point in this section, that there were different upper limits for C concentration in the original NSI and the resurvey, which is scientifically implausible and which we suggest erroneously led to the large reported losses of C from highly organic soils.

The reviewer suggests that the change in analytical method by the NSI could not have caused the reported results because it only affected 6% of samples. However the change affected the most extreme values, those with OC > 150 g kg-1 and it is well known that statistical procedures can be markedly influenced by errors in small numbers of extreme values. The parameters of a regression line, for example, can be completely changed by the presence of a single outlier. The argument by the reviewer against effects due to changes in analytical methods is therefore not convincing.

Referee 3: Page 2285 lines 1-6

The statement that Bellamy et al.'s results may be partly explained by regression to the mean is wrong. There is a possibility of a regression to the mean in Equation (1) of Bellamy et al., which they used to interpolate the soil carbon changes at the sites that were not re-sampled for the map in Fig 1b, but did not use anywhere else in the paper. However, Lark et al. (2006) show that the bias due to regression to the mean in this equation is small. So regression to the mean does not affect Bellamy et al.'s results in any way.

Chamberlain et al.: This is not correct. With regard to Equation (1) of Bellamy et al. (2005), Potts et al. (2009) re-examined the issue and concluded that there was likely to be a non-negligible effect of regression to the mean and this was accepted in the response to this article by Lark et al. (2009). However the equation is not the only result which will be affected by this problem. Regression to the mean affects any calculations based on categorisation by an initial value and hence will apply to the estimated rates of loss given in Table 1 of Bellamy et al. which form the main results of the paper.

References

Bellamy, P. H., Loveland, P. J., Bradley, R. I., Lark, R. M., and Kirk, G. J. D. (2005) Carbon losses from all soils across England and Wales 1978-2003, Nature, 437, 245-248.

Bellamy P.H. & Kirk, G.J.D. (2009) Response to comments on 'Baseline values and change in the soil, and implications for monitoring' by Potts et al. Eur. J. Soil Sci. 60, 483–484.

Black, H. I. J., Bellamy, P. H., Creamer, R., Elston, D., Emmett, B. A., Frogbrook, Z. L., Hudson, G., Jordan, C., Lark, M., Lilly, A., Marchant, B., Plum, S., Potts, J., Reynolds, B., Thompson, R., and Booth, P. (2008) Design and operation of a UK soil monitoring network, Environment Agency, Bristol, UK.

Emmett, B. A., Frogbrook, Z. L., Chamberlain, P. M., Griffiths, R., Pickup, R., Poskitt, J., Reynolds, B., Rowe, E., Spurgeon, D., Rowland, P., Wilson, J., and Wood, C. M. (2008) CS Technical Report No. 3/07. Soils Manual, Centre for Ecology & Hydrology, Wallingford, UK.

Kirk G.J.D. & Bellamy P.B. (2010) Analysis of changes in organic carbon in mineral soils across England and Wales using a simple single-pool model. Eur. J. Soil Sci., 61, 401–411.

Lark R.M., Bellamy P.H. & Kirk, G.J.D. (2006) Baseline values and change in the soil, and implications for monitoring. Eur. J. Soil Sci. 57, 916–921.

Potts, J. M., Chapman, S. J., Towers, W., and Campbell, C. D. (2009) Comments on 'Baseline values and change in the soil, and implications for monitoring' by RM Lark, PH Bellamy & GJD Kirk, Eur. J. Soil Sci., 60, 481-483.

Interactive comment on Biogeosciences Discuss., 7, 2267, 2010.