SOIL Discuss.,
https://doi.org/10.5194/soil-2020-18-RC2, 2020

Interactive comment on “Nitrogen availability determines the long-term impact of land-use change on soil carbon stocks in grasslands of southern Ghana” *by* John Kormla Nyameasem et al.

Anonymous Referee #2

Received and published: 28 June 2020

General comments

This research assessed the effect of long-term land-use management on soil organic carbon (SOC). It then sought to explain these changes in terms of soil macro nutrients and condensed tannins, with a particular focus on plants classified into groups according to their use in the farming system.

There are significant issues around the design, analysis and interpretation to be addressed before publication (see below). These also bear on the statistical analysis and

the conclusions that can be drawn legitimately. However, the topic is relevant, the aim is worthwhile, and the paper fits within the scope of the journal, so I encourage the authors to persevere.

The paper is generally well-written (with some exceptions, e.g. lines 66-68, 180-181, 223-225), but the discussion goes well beyond the data, is frequently imprecise, and at times irrelevant (some material might be relevant in a thesis but not a research paper). Discussion needs to focus more on the findings, be less speculative, and substantially shortened. An indication of a lack of focus is the broad opening statement that the scope of the study embraces impacts on C cycling (line 10). It doesn’t - the researchers measured the outcome (SOC) of C inputs and outputs over 50 years. There are not even estimates of the current biomass production and removal, and no attempt is made to relate biomass production to soil macronutrients or accessions of condensed tannins.

Design

This is not a controlled experiment. The paper is based on opportunistic sampling on and around a research farm, from a selection of locations with supposedly different long-term management. There appears to be no classical experimental design with replicated and randomized treatments. Locations that received different management are deemed to be management ’treatments’. It appears there was no site pairing that could have been used to control error.

The fact that this is not a traditional experiment is not in itself a concern, but it does mean that extra care is needed to describe the sampling and analysis, and in particular, care is needed in drawing conclusions.

My first concern is that there are no baseline data for SOC or soil fertility in 1966, and presumably no archived soil samples dating from 1966 that could have been analyzed. To calculated changes in SOC and macronutrients over 50 years, the authors use a pseudo baseline derived from just three fields of native grassland located somewhere near the research farm. These were sampled on the assumption that the present native grassland soils and plants are exactly as they were in 1966. Each field was subsampled at only four locations and bulked for analysis. The mean of three values (fields) therefore provides the slender basis upon which the whole paper rests. We have no idea how these fields were chosen, no idea why they were not selected for development in 1966 (too poor, too good or?), no details of management and changes in management over 50 years, and no idea if the fields were representative of native grassland back in 1966. There needs to be sufficient evaluation of the assumption in a revised paper to convince readers of its validity.

Description of the experiment is vague, but it appears that 'treatments' were located on both the research farm and in the surrounding area. Apparently, in 1966 some fields around the research farm were converted from native grassland to field crops and some were sown to pasture and grazed. This allowed the researchers to select three fields of each of (i) native grassland, (ii) field crops and (iii) seeded-grazed land, located outside the research farm. These were the 'treatments'. Readers are not told where fields were located, why they were chosen, whether management remained stable over the 50-year time-frame, or how the authors know about management over this time (was anything documented?). It is quite possible, for example, that the choice of land for field crops in 1966 was based on a perception it was the most fertile land, in which case the real loss in SOC over time may have been greater than reported here. This an important point, because the authors found no change on SOC over 50+ years of arable agriculture, which is very different from the majority of studies that show SOC declining under arable agriculture (one of the driving forces behind 'conservation agriculture').

Continuing with the experiment design, the four cut-use treatments were located on the research farm: 59 small co-located plots with different species were allocated to four groups ('treatments') ranging in size from 3 to 46 species (replicates) per 'treatment'. The text says 59 plots altogether, but Fig. 2 indicates 72. Given the potentially large

differences between species within groups, one might expect large standard errors when species are replicates, but this is not the case if the analysis is to be believed. I think the analysis needs to be revisited.

Of the 59 (or 72) small plots of cut-use species, 46 are grasses that are treated as replicates of this 'land-use'. We are required to assume that these 59 (72?) plots are managed today just as they were 50 years ago - same species, nil fertilizers, same cutting regime, no differential tillage for re-establishment etc. This may be a valid assumption, but we are given too little information to test it.

Analysis

The apparent lack of experimental design, and the very different numbers of 'replications', present some challenges for the analysis and interpretation of data. The authors appear to have foregone the advantages of good design which allows the experimenter to make causal inferences about the relationship between independent variables and a dependent variable, to rule out alternative explanations due to the confounding effects of extraneous variables (i.e. control), and to reduce variability within treatments, making it easier to detect differences in treatments.

The authors could usefully say a bit more to convince readers that the assumptions underlying an ANOVA have been met: the experimental errors are normally distributed; variances between treatments is equal; and samples are independent (each sample is randomly selected and independent).

Where replicate data from different treatments are combined for regression analysis, the 'treatments' with large rep numbers are over-represented, possibly giving rise to significant relationships that may not otherwise be significant. For example, the cut-use grass group of plots had the greatest TN (inexplicably), and TN is said to be the individual analyte most highly correlated with change in SOC. So, the question is, would TN still be highly correlated with change in SOC if these data were deleted from the analysis. I suggest the authors consider whether the regressions involving all data would be more appropriate if regressed as treatment means to avoid bias from under- or over-represented treatments.

The authors refer to a 16% fall in SOC due to land-use change – this appears to the mean decline of all non-grassland plots, not the mean of treatment means, and as such it is heavily biased towards the treatment with most replicates, i.e. the cut-use grasses.

Another area of concern is where the authors attribute cause and effect in a correlation when all they have is an association, e.g. lines 15-16, 205 and Fig. 3. Without more information we cannot attribute cause to either x or y, or to an unknown co-variate of either one. All Fig. 3 shows is a fairly tight C:N ratio of about 12:1, as many others have reported. In other words, Fig. 3 reflects the stoichiometry of stable soil organic matter, not cause and effect.

We might also expect a stoichiometric relationship between P and C, but this is not evident in Table 4, perhaps because the wrong fraction of P was measured (extractable no organic). This leads me to question the reliability of at least some data in Table 2 – why, for example, would cut-use grasses and legume herbs appear to deplete soil P, but not legume of non-legume trees and shrubs (numerically if not statistically) when product is removed from all of them? What could explain the apparent soil acidification under grazed-seeded grassland and cut-use legume-herb, other than errors? Or the depletion of K under legume-herb? Or the rise in K under arable land, unless K-fertilizer is applied quite heavily?

In line 19, a complex multiple regression is said to 'explain' 92% of variation in SOC stocks – the equation might mathematically account for 92% of the variation, but this is very different from a biophysical 'explanation'. The authors saw value in reporting a complex equation to account for 92% of variation in SOC, but skim over the fact that a single variable, N, accounts for 90% of the variation (Table 4). Does increasing the complexity improve our understanding of the processes?

I suggest the authors take care not to refer to means as being different when statis-

tically they are not. Line 230 refers to a sequestration rate of 31 kg C/ha in legume trees/shrubs, implying this is greater than with other land uses and a strategy worth pursuing to build SOC, but the statistical reality is that after 50 years of different land uses, no treatment differs from native grassland (Fig. 2). All you can say is that trends were evident but they were not statistically significant. Lines 231-234 state your expectation, not what you can statistically support. I think all you can say is that your trends are heading in the expected direction. You can propose a hypothesis worth testing. I recommend the authors review all data to ensure they are reported only to the number of significant digits that can be measured.

Fig. 2 shows that although there are numerical differences in SOC, statistically there are NO treatment differences (judging by the letter superscripts – I suggest you check this). Statistically speaking, there are presently no significant differences in SOC. If there are no differences in SOC, there can be no differences in the rate of decline over 50 years. Only non-significant trends.

Other

When rewriting, give the name of the nearest town to the research site, and in the Introduction give a very brief overview of how livestock are managed, to provide context.

Remove Equation 1 from the Methods and include in Results, making any other changes necessary to make this possible.

I suggest you delete Fig.2 and put the data into Table 2. This will make it easier for readers to view and relate all of the data, and make the paper shorter. Table 2 as it stands does not present 'impacts' (changes), it presents only the status of soils following 50 years of various land-uses. Only if you include the apparent change in SOC does it include an impact.

The data in Table 3 and related discussion appears to be the most original and interesting, but it's hard to interpret its significance without knowing the biomass produced.

Interactive comment on SOIL Discuss., https://doi.org/10.5194/soil-2020-18, 2020.