Interactive comment on "Continental-scale controls on soil organic carbon across sub-Saharan Africa" by Sophie F. von Fromm et al.

Note: The line numbers in our answers are referring to the updated manuscript.

Anonymous Referee #1

Received and published: 12 November 2020

General comments:

REVIEWER_01: This very interesting paper tries to explain the heterogeneity of soil organic contents in tropical soils. Pedogenic, climatic and land cover properties were related to SOC contents to identify the main factors that control SOC. It mobilized numerous original data from a network of soil measurements through Africa. Two soil depths (0-20 cm and 20-50 cm) were considered. The paper is really worth a reading as the results were rich and surprising.

ANSWER: Thank you very much for taking the time to review our manuscript and your most constructive feedback. Our detailed answers can be found below.

R_01: However, the work needs to be more explicit to be really compelling on some points (i.e. please give the main soil types studied: is there any andosols or young volcanic soils which could reinforce the role of Al, Fe extractable with oxalate in the SOC stabilization? It should be easy to dismiss that hypothesis).

A: This is a really good point. Unfortunately, we do not have the information about the main soil types. When extracting data from global and African soil-type products, none of the soil profiles was classified as Andosols. However, those products are usually not that precise and it is difficult to derive detailed conclusions. Based on the position of some soil profiles (e.g. Rift Valley in East Africa) it might be possible that some of them are Andosols or influenced by volcanic activities in the past.

Based on this comment and also on one comment from Reviewer #2 we added the following paragraph to the discussion section in the manuscript.

Line 465: "

Geographic patterns

All soils result globally from the same soil-forming factors (climate, organisms, topography, parent material and time) and are formed by similar processes (e.g. oxidation, reduction, leaching, transport and accumulation). This might explain to some extent, why similar soil and climate parameters are important to explain SOC content variation in sub-Saharan Africa as compared to other regions. However, significant differences are still visible in subtropical and tropical soils, especially in terms of mineralogy, weathering and soil formation, which are related to important differences in climate, soil age, parent material and vegetation (Buringh 1970). Such differences do occur between soils from sub-Saharan Africa, which do not only differ greatly in their soil properties and climate (Table 1), but also in vegetation, parent material, soil age and their vulnerability to degradation (Jones et al. 2013). However, due to the lack of precise geospatial products for those parameters for sub-Saharan Africa, we can only discuss them here qualitatively.

The AfSIS sites (Figure 1) are mainly derived from two parent material types: i) metamorphic rocks and ii) volcanic rocks (Hartmann and Moosdorf 2012; Jones et al. 2013; Schlüter 2008). Metamorphic rocks are most commonly found in West Africa, Southern Africa and Madagascar. These regions are characterized by old cratons, except for Madagascar, which is influenced by Mesozoic volcanism (Schlüter 2008). Most of these soils are classified as Ferralsols according to the WRB soil classification system (Jones et al. 2013). This partly explains, why the AfSIS soils from those regions are usually highly weathered with low pH_{H2O} values. In contrast, soils derived from volcanic rocks are mainly found in East Africa in the Great Rift

Valley. These soils are usually younger and less weathered (Buringh 1970), which is also mirrored in soil properties within the AfSIS data set. These soils are characterized by lower CIA values and higher Al_{ox} and Fe_{ox} concentrations compared to other soils in the AfSIS data set. Outside of the influence of volcanic rocks, Ca^{2+} rich soils are frequent in East Africa and are dominated by a high concentration of Ca_{ex} and high pH_{H2O} values. Since Al_{ox}, Fe_{ox}, and Ca_{ex} were important predictors of SOC in our analyses, the SOC content is usually also higher at AfSIS sites in East Africa compared to sites in West Africa and Southern Africa.

Although certain soil properties, in combination with climate variables, are important to explain SOC concentration variation, different soil forming factors, such as parent material and soil age, are important to understand, which soil properties will dominate – at least at this large-scale approach. To link those two aspects more quantitatively on continental-scales might be a direction for future studies."

R_01: Discussion on the limit of your study and on the differences with the previous studies could deepen the significance of your results. The clay and fine silt contents seem to have poor impact on the soil organic content on the contrary of previous results. OK but previous results have related SOC to the clay+silt size fraction at 0-20 µm, you have related SOC to the clay+silt size fraction at 0-8 µm. Please discuss that point?

A: Thank you for pointing out this important aspect. We realized that we have to be more explicit on why we used the rather unusual cut-off of <8 µm (clay + fine silt), instead of the more standard <2 µm (clay) or <20 µm (clay + silt). Our initial idea was to use clay content only. This particle size fraction is commonly used in soil and biogeochemistry studies to link SOC stabilization with clay and clay minerals (e.g. Rasmussen et al. 2018; Quesada et al. 2020; Bruun et al. 2010). However, the reproducibility for duplicates in our data set was sometimes poor for particles <2 µm, but much more reliable for particles <8 µm (Figure A1). This is the main reason why we decided to use the cut-off <8 µm. As this is closer to the pure clay fraction (< 2 µm), it allows us to better link our results to the stabilization capacity of clay content/clay minerals. A larger cut-off would include more silt-sized aggregates, which might result in additional and different C stabilization mechanisms, as discussed in more detail in Six et al. (2002a).

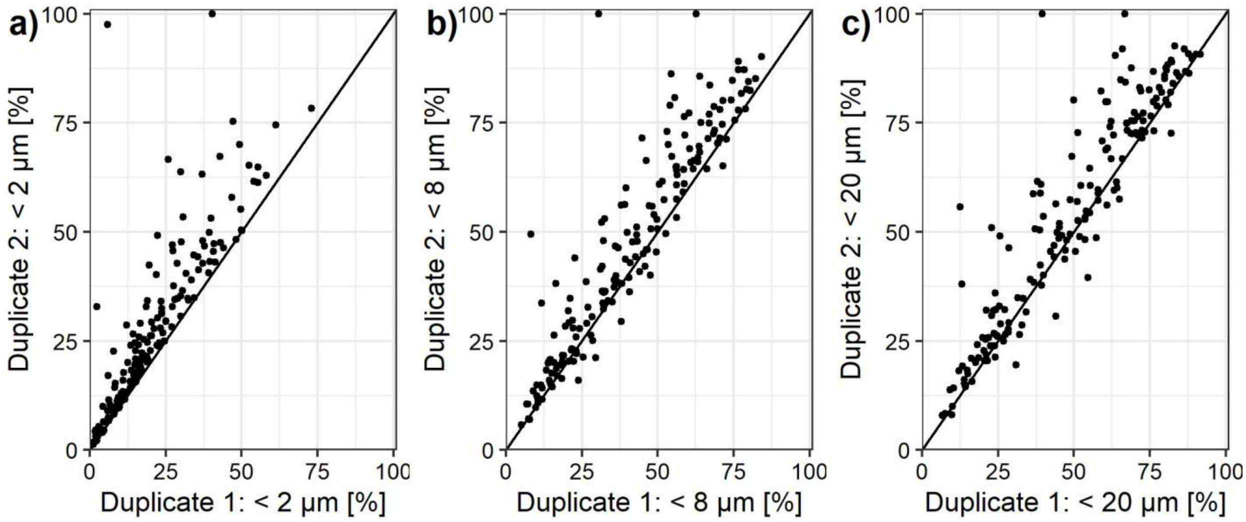

Figure A1: Duplicate measurements of the particle size distribution data [%]. Panel a) Duplicate 1 and 2 <2 µm; b) Duplicate 1 and 2 <8 µm; c) Duplicate 1 and 2 <20 µm

However, based on the comment we also re-ran our analysis with the clay + silt size fraction (0–20 µm) and compared it to the results with the clay + silt size fraction (0–8 µm). We will display and discuss the findings briefly in the following section and in the supplement (Figure A1).

When comparing the raw data, there is a slight increase in the correlation coefficients between SOC and particle size data <8 µm and <20 µm, respectively (Table A1.1 and Figure A1.2). Yet, the differences between the respective correlation coefficients are always smaller than 0.1 (Table A1.1).

Table A1.1: Correlation coefficient between SOC and particle size data <8 µm and <20 µm for all samples (n = 1,601), topsoil (0–20 cm; n = 791), and subsoil (20–50 cm; n = 810)

Samples	**<8 µm**	**<20 µm**
All	0.32	0.41
Topsoil	0.37	0.46
Subsoil	0.43	0.49

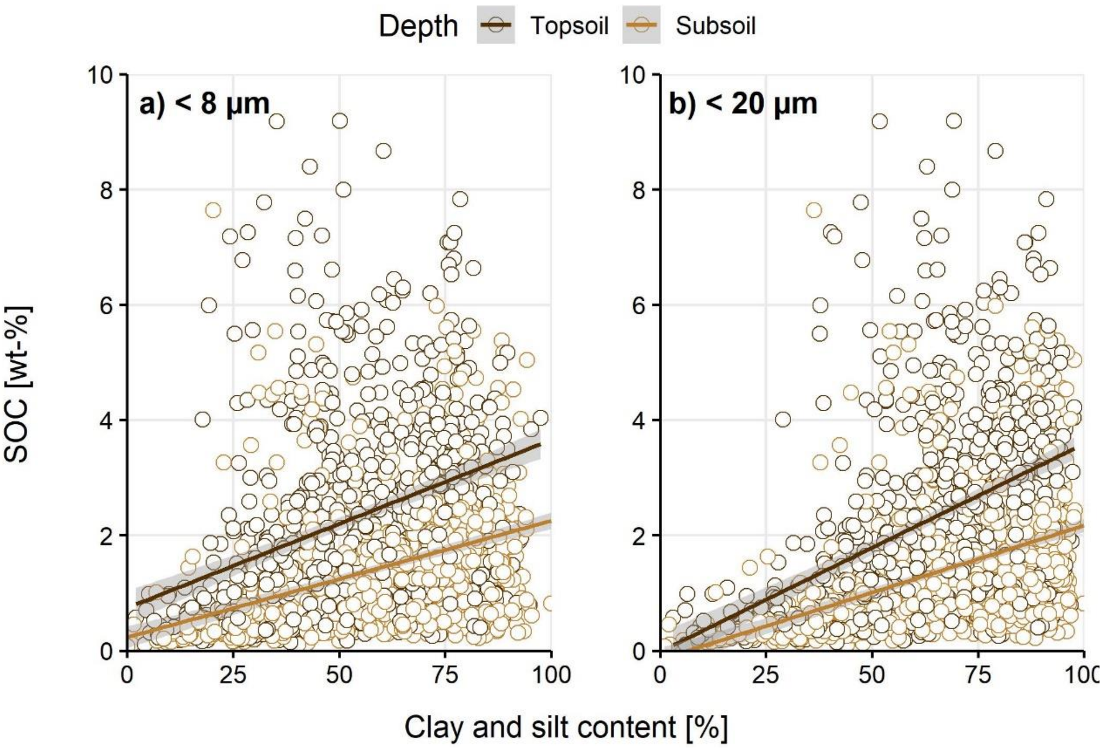

Figure A1.2: a) Soil organic carbon (SOC) content [wt-%] and clay and fine silt content <8 µm [%] by depth; b) SOC content [wt-%] and clay and fine silt content <20 µm [%] by depth.

In addition, we also re-ran all our models that only contained clay + fine silt (<8 µm) as an explanatory variable with clay + silt <20 µm. Overall, the differences between the two particle size groups (<8 µm and <20 µm) became even smaller (Table A1.3) compared to the raw data. The differences between the R^2 for the different models was always less than or equal to 0.07. Furthermore, we also substituted the variable clay + fine silt <8 µm with clay + silt <20 µm within the entire random forest models (one model for each depth layer with all explanatory variables). For the topsoil samples, the R^2 of the <20 µm model decreased by 0.01 (from 0.71 to 0.70) and stayed the same for the subsoil samples (0.72).

Table A1.3: Summary table of R² for the different models (linear mixed effects model and random forest) for the two different explanatory variables (<8 µm and <20 µm) for all samples (n = 1,601), topsoil (0–20 cm; n = 791), and subsoil (20–50 cm; n = 810)

Model	**Linear mixed model**	**Random forest (topsoil)**	**Random forest (subsoil)**
Clay + fine silt <8 µm	0.01	0.12	0.12
Clay + silt <20 µm	0.03	0.17	0.19

Based on the results of the re-analysis with the size fraction <20 µm, we argue that both groups show more or less similar results and that the overall picture does not change. Therefore, we suggest keeping the size fraction group < 8 µm, which only includes clay and fine silt and is more closely related to the pure clay content (<2 µm).

Based on the previous discussion, we made the following changes in the manuscript to better clarify the selected particle size data for the readers. In addition, we added the figures and tables shown here to the supplement of the manuscript – they are all displayed together in Figure A1.

Line 149: "We used 8 µm as cut-off to capture all clay and fine silt particles. Results were comparable to <20 µm (see SI material Figure A1), yet <8 µm was selected because it is more relevant to our interest in studying the influence of clay content on SOC concentration. Due to reproducibility issues, we could not use the fraction <2 µm (clay content) only (Figure A1)."

R_01: The separation of the two-studied soil depths is not always clear in the Ms. Can it affect the results?

A: We revised the manuscript carefully and tried to be clearer regarding the two sampling depths, and where the results vary between topsoil and subsoil:

Line 200: "In order to identify the best parameters explaining the variation of SOC, we first ran a full model, including PET/MAP, MAT, clay and fine silt content, pH_{H2O}, Al_{ox}, Fe_{ox}, Ca_{ex}, CIA, depth (topsoil vs subsoil), land cover and the interaction term pH_{H2O}*Al_{ox} as fixed effects and site/cluster/profile as random effects."

Line: 218: "Except for the two depth models (topsoil vs subsoil), we did not include land cover in those linear mixed effects models. This is justified because the distribution of the different land cover groups was too unequal among the sub-models, and land cover was not identified as an important predictor in the full model (Table B2)."

Line 228: "Regression tree analysis was applied to obtain an easily interpretable and non-linear model for the entire dataset and for both depth layers (topsoil vs subsoil) that best describes the existing data (Breiman et al. 1984)."

Line 286: "Differences between the predictors were marginal for the two depth models (topsoil vs subsoil; Figure A5)."

Line 326: "Overall, the results showed that the explanatory variables did not differ much between the depth intervals (topsoil vs subsoil), while their magnitude did."

A: It is really likely that the selected sampling depth intervals have an effect on the results. We made some changes at the beginning of the discussion section to make this clearer and also to mention it at the end of the discussion again.

Line 359: "In contrast, predictor differences were much smaller between topsoil (0–20 cm) and subsoil (20–50 cm) samples. This may partly be due to the large depth increments for each of the two sampling depths. However, since the identified SOC-controlling factors were similar for both depth layers, any differences were probably mostly driven by the fact that subsoil samples usually contain less SOC due to lower C inputs at depth (Jobbágy and Jackson 2000). Soil erosion at some sites might also dilute differences between the two depth layers, since water and wind can permanently remove surface soil."

Line 443: "Surprisingly, land cover was not an important predictor of SOC content in our models, especially in topsoils. One possibility may be that the relatively large 0–20 cm depth interval might dilute differences that would be more marked in the top few centimeters"

R_01: One of your main result is that land cover do not explain SOC content at continent scale, so in the world wide discussion on C sequestration to mitigate CO2, what do you suggest?

A: When we first did the analysis, we were also surprised that land cover appeared irrelevant to explain the variation in SOC within the data. However, this can probably be linked to the scale we are working at. It has been shown by previous studies that the effect of land cover changes is more important at smaller scales, whereas climate and soil development properties are more dominant at larger scales (Holmes et al. 2004, 2005). And this is likely also true for CO_2 mitigation potential/strategies, since they depend on site and/or regionally specific aspects. This importance of site-scale management practices makes it difficult to draw any conclusions on C sequestration to mitigate CO_2, based on our analysis. In addition, the information we have about land use at the profile level is not detailed enough. We do not know what kind of land management practice is specifically applied in a given region, nor do we know much about the land use history at those sites. Additionally, other studies in Africa have shown that historic land use can be a more important determinant than current land use in determining current soil organic carbon stocks (Vågen et al. 2006).

Although our ability to draw conclusions from this dataset is limited, given the importance of the issue, we made the following changes in the manuscript:

Line 451: "This might be due to the high variation of SOC content within the different land cover classes at this large spatial scale (Figure 2a). For example, the class cropland contains a wide variety of cultivated plots – we did not have more detailed information about land management practices. Fujisaki et al. (2018b) showed that SOC stock changes in tropical cropland soils are mainly driven by C inputs.

On the other hand, it is also known that land use changes and their impact on soil physico-chemical properties are scale-dependent and are likely to be more distinct at smaller scales (Holmes et al. 2005; Holmes et al. 2004). For example, land management and land degradation (i.e. erosion) are known to impact SOC stocks on regional scales in sub-Saharan Africa (Winowiecki et al. 2016a). However, we lacked the detailed data necessary to disentangle the impacts of different practices. Additionally, since the focus of our work was on natural soil physico-chemical and climate properties, we did not further investigate those anthropogenic factors at this large spatial scale. Future studies are needed to better understand the impacts of land management and carbon storage potential in soils across sub-Saharan Africa at different scales (Fujisaki et al. 2018a; Vanlauwe et al. 2015). Overall, our data for sub-Saharan Africa suggests that SOC content on a continental scale is better explained by stabilization potential in soils (climate, geochemistry) than by different aboveground C inputs (vegetation)."

R_01: The final sentence on soil erosion is very OK but could be strengthen.

A: We understand that soil erosion is an important aspect of SOC stocks – especially in subtropical and tropical soils – and we mention this briefly in the introduction. During the field campaign, visible signs of soil erosion by water were reported: *none*, *sheet* (uniform removal of soils in thin layers; sign of sheet erosion include bare areas, water puddling, visible grass roots, exposed tree roots, and exposed subsoil or stony soils), *rill* (shallow drainage lines less than 30 cm deep), and *gully* (channels that are deeper than 30 cm). Most of the studied profiles showed no signs of erosion (~300) or sheet erosion (~375). The other profiles showed either rill erosion (~75) or gully erosion (~25). However, we did not include soil erosion in our main analysis for several reasons. First, the focus of our work was on natural processes. However, land cover, which can be seen as a broad proxy for potential soil erosion, did not explain much of the variance in SOC. Furthermore, we think that some of the other soil properties we included in our analysis are capturing the relevant indicators for soil erosion as well, such as pH, Ca_{ex} and MAP. And lastly, similar to land cover, it is likely that the importance of soil erosion is more important at a regional scale.

To address this comment, we have added additional text on land cover and land management within the manuscript (line 451, see text reproduced in the previous comment).

R_01: Your remark on line 436 ("… on a regional scale") is also OK. Please insist a little more on that point.

A: We hope that the changes we made in line 451 (see previous comment) now help to make this statement clearer and therefore only added a few words in line 483:

Line 506: "Overall, a combination of soil pH_{H2O}, PET/MAP, Ca_{ex} and Al_{ox} seems to be an appropriate set of variables to explain variation of SOC content on a continental scale across sub-Saharan Africa. This does not imply that other variables, such as clay and fine silt content and land cover are not good predictors on a regional scale as shown by previous studies."

R_01: You have also to be clear at the end of the Ms between predict SOC content (for instance to map SOC content or stock) and predict SOC dynamics (you have not studied SOC dynamic in your study): Ln. 437 to the end are not very clear.

A: Line 511: "This would help to improve the predictive capacity of these models and to close the gap between our theoretical understanding of SOC concentration across large scales and our ability to improve terrestrial biogeochemical projections that rely on existing models."

R_01: After your result on the lack of significance between soil clay and SOC contents, what do you think about using the soil carbon saturation deficit to quantitatively assess the soil carbon storage potential in tropical soils? Could you add one or two sentences on that point.

A: This is a really interesting and relevant aspect that was brought up here by the reviewer. This topic is especially relevant for future strategies to increase the C amount in soils as proposed in the *4 per 1000* initiative. Following a literature review (e.g., Hassink 1997; Feng et al. 2013) on this topic, we have some concerns that our data and analysis will allow us to sufficiently address this topic in the presented manuscript. However, it would be a really interesting study to address this topic in sub-Saharan Africa based on a more systematic approach.

One concern is related to the fact that we do not always have a degraded site next to an undisturbed site – which is needed as reference site. In addition, the information we have about the degraded sites is not always sufficiently detailed. We do not have information, e.g., about exact land management practices, nor do we know much about land management history at those sites. Such information is crucial to put the results in a broader context since it is likely that the soil carbon saturation deficit is influenced by the type of management (e.g. till vs non-till, compost additions, enhanced residue return, …) and the time since the land was converted. Furthermore, we do not know the amount of C that is associated with the clay + silt fraction due to the fact that we only have bulk data available. However, the clay + silt fraction is of most interest regarding carbon storage potential in soils. For this, it would also be good to have more information about the specific clay minerals and their specific surface area. Last but not least, soil carbon saturation deficit assessments are likely be more relevant on a regional scale, similar to CO_2 mitigation strategies.

Based on the previous discussion we added the following in the manuscript (which is related to the changes we made regarding the CO_2 mitigation comment made earlier by the reviewer).

Line 451: "This might be due to the high variation of SOC content within the different land cover classes at this large spatial scale (Figure 2a). For example, the class cropland contains a wide variety of cultivated plots – we did not have more detailed information about land management practices. Fujisaki et al. (2018b) showed that SOC stock changes in tropical cropland soils are mainly driven by C inputs.

On the other hand, it is also known that land use changes and their impact on soil physico-chemical properties are scale-dependent and are likely to be more distinct at smaller scales (Holmes et al. 2005; Holmes et al. 2004). For example, land management and land degradation (i.e. erosion) are known to impact SOC stocks on regional scales in sub-Saharan Africa (Winowiecki et al. 2016a). However, we lacked the detailed data necessary to disentangle the impacts of different practices. Additionally, since the focus of our work was on natural soil physico-chemical and climate properties, we did not further investigate those anthropogenic factors at this large spatial scale. Future studies are needed to better understand the impacts of land management and carbon storage potential in soils across sub-Saharan Africa at different scales (Fujisaki et al. 2018a; Vanlauwe et al. 2015). Overall, our data for sub-Saharan Africa suggests that SOC content on a continental scale is better explained by stabilization potential in soils (climate, geochemistry) than by different aboveground C inputs (vegetation)."

Specific comments

R_01: Abstract: The conclusion of the abstract is not clear/strong enough.

A: Line 30: "Our findings indicate that the key controlling factors of SOC across sub-Saharan Africa are similar to what has been reported for temperate regions – except for soil texture and vegetation cover. However, the strength and importance of the controlling factors vary across the environmental gradient we studied. Changes in this gradient are mainly driven by climatic conditions, weathering status and soil pH and have an influence on the concentration of Ca_{ex}, Al_{ox} and Fe_{ox} and their role on SOC content."

R_01: Ln. 62: "broader variables such as clay content: : : " Ok but clay content implies SOC protection and association with minerals (L. 61). Please rephrase. Ln. 64 "variety of processes" Please be more explicit (aggregation, organo mineral association: : :) Ln. 64 "differ" or their relative importance differ? (different hierarchisation). Please specify

A: A similar comment was brought up by Reviewer #2. We agree that the paragraph was not really well-written. Since modeling is not the main focus of our manuscript, we would like to keep this paragraph rather short. However, we think it is an important aspect worth mentioning in the introduction. We revised the paragraph:

Line 61: "SOC stabilization is commonly conceptualized as competition between accessibility for microorganisms versus chemical associations with minerals (Oades 1988; Schmidt et al. 2011). These processes are often only considered implicitly by models (Blankinship et al. 2018; Schmidt et al. 2011). Instead, models commonly rely on broader variables such as clay content, which is used as a proxy for sorption and other organo-mineral interactions (Rasmussen et al. 2018; Schmidt et al. 2011). These more generic variables integrate a variety of stabilization processes which can be difficult to disentangle. They might even differ in their relative importance and may not adequately capture soil mineralogy and chemistry across different ecosystems and climate zones. Hence, improving the predictive capacity of such models requires not only a better understanding of the factors that control SOC dynamics, but also verification (or falsification) of those new findings in regions that are underrepresented in field studies and models."

R_01: End of the introduction: Please specify your hypothesis, you have not searched factors randomly. Please be more explicit. Your questions are too vague. Which soil properties you will focus on? What findings do you except? You can base hypothesis on soil properties, climate properties and land cover. You have everything to make a more compelling introduction. Please explicitly justify the soil properties and the soil depth you will focus on in your study.

A: We revised the end of the introduction carefully to make it more explicit and to give the reader a road-map for the following sections.

A: Line 89: "This data set used in this study covers a wide range of climatic and mineralogical conditions – from very arid to humid regions, with different pH_{H2O} values, soil texture, weathering status, exchangeable cations and extractable metals – allowing us to test different parameters to explain the variation in SOC content in subtropical and tropical soils across sub-Saharan Africa for two distinctive depth layers (topsoil: 0–20 cm and subsoil: 20–50 cm). Here, we use this continental-scale data set to address the following research questions:

The following research questions will be addressed in this study:

1. Which soil properties and climate parameters best explain variation in SOC content across sub-Saharan Africa and do these findings differ from temperate regions?

   We explored the relative importance of soil texture, exchangeable Ca, oxalate-extractable Al and Fe, soil pH_{H2O}, mean annual temperature, aridity index (PET/MAP), land cover and weathering status to explain the variation in SOC content on a continental scale. Based on the results from temperate regions and regional studies in tropical soils, we expect that oxalate-extractable Al and Fe and soil texture will be one of the most important predictors of SOC concentration.

2. How do controls on SOC content vary across climate regions and soil conditions in sub-Saharan Africa?

   Due to the heterogeneity of climate and soil conditions across sub-Saharan Africa we expect to see distinct controls on SOC content between regions that are drier with less weathered soils compared to regions that are humid with highly weathered soils. Exchangeable Ca will likely only be important in drier regions with less weathered soils, whereas oxalate-extractable Al and Fe will mainly be important in humid regions with highly weathered soils."

R_01: Ln. 105 "mid-infrared spectroscopy data" data or model? How have chosen the representative spectral data? Have you separate the two soil depth? Why have you finally chosen to study the lab measurements and not the predictions?

A: Thank you for this critical comment. It should be "mid-infrared spectroscopy model" and we changed it in the manuscript accordingly (line 119). The reference data set for the mid-infrared spectroscopy model was chosen using the Kennard-Stone algorithm (Kennard and Stone 1969). First, the spectra is decomposed using a PCA and selecting a subset using Kennard-Stone that gives a uniform coverage across the spectral space of the data, including also boundary cases. Terhoeven-Urselmans et al. (2010) contains more information on this and also the adaptations of the approach: "The adaptations were (i) that no samples from the same soil profiles were allowed to be split between the calibration and validation sets, and (ii) that the two extreme score values of each principal component were chosen for the calibration set and the next extreme score values chosen as starting samples for the validation set."

We decided to analyze the lab measurements only for this study since they contained more variables of interest in terms of SOC stabilization, such as oxalate-extractable Fe and Al.

Line 118: "These samples were used to calibrate mid-infrared spectroscopy models (Terhoeven-Urselmans et al. 2010) and to predict the remaining 16,255 soil samples (Vågen et al. 2016; Winowiecki et al. 2017). The calibration subset was chosen to maximize the variation of the spectral data using the Kennard-Stone algorithm (Kennard and Stone 1969). More information about this approach can be found in Terhoeven-Urselmans et al. (2010). This selection strategy results in unequally distributed samples across 51 of the 60 sentinel sites, but still captures the variation of the original dataset."

R_01: Ln. 120 why have you chosen this limit at 8 µm? What arguments?

A: We kindly refer here to the answer we gave to the more general comment from the reviewer regarding clay and clay+silt content. The changes we made in the manuscript are the following:

Line 149: “We used 8 µm as cut-off to capture all clay and fine silt particles. Results were comparable to < 20 µm (see SI material Figure A1), but < 8 µm was selected because it is more relevant to our interest in studying the influence of clay content on SOC concentration. Due to reproducibility issues we could not use particles < 2 µm only (Figure A1).”

R_01: Ln. 133 It seemed that soil texture was performed without organic matter destruction? Is that really the case? Please specify.

A: The pretreatment Calgon and ultrasonic dispersion that was done prior to Laser Diffraction measurement were done to disperse aggregates.

Line 147: “Each sample was shaken for 4 min in a 1% sodium hexametaphosphate (calgon) solution with ultrasonic energy before measuring to disperse aggregates.”

R_01: Ln. 136-140 it was not really clear to me.

A: We hope that the following changes made it easier to understand the correction we did for samples that contained inorganic carbon before calculating the Chemical Index of Alteration (CIA). According to Nesbit and Young (1982) the equation holds only for soil samples that do not contain carbonates or apatite. If so, an appropriate correction is needed.

Line 152: “Aluminum, Ca, K and Na concentrations were used to calculate the chemical index of alteration (CIA) after Nesbit and Young (1982), using the following equation:

$$CIA = Al_2O_3 / (Al_2O_3 + CaO + K_2O + Na_2O) * 100 \qquad (2)$$

where CaO is the amount incorporated in the silicate fraction. Correction is necessary for samples that contain carbonates and apatite (Nesbit and Young 1982). We adopted an approach introduced by McLennan (1993): the correction assumes that Ca is typically lost more rapidly than Na during weathering. Whenever a soil sample contained inorganic C ($C_{total} – C_{org}$; used as a proxy for carbonates and apatite) and when the CaO content was larger than that of Na_2O in the same sample (n = 476), the CaO concentration was set to that of Na2O from the same sample (Malick and Ishiga 2016). After applying the correction, no obvious correlation remained between CIA and inorganic C (Figure A2). The index increases (i.e. more highly-weathered soil) with the loss of Ca^{2+}, K^+, and Na^+.”

R_01: Ln. 145 “1,601 soil samples” is that on the 2,002 measured soil samples?

A: Line 164: “This resulted in a total of 1,601 soil samples (out of the original 2,002 samples) at 45 sentinel sites across 17 countries.”

R_01: Statistical analysis: Note for the editor: A deep analysis of the statistics has to be done by a reviewer more competent than I am in statistics.

R_01: Table 1: Please specify in the title that it is a summary for the two soil depths (0-20, n= ?, and 20-50 cm, n= ?)

A: Line 253: “Table 1: Summary statistics of all numerical soil and climate variables for the entire data set (n_{total} = 1,601; $n_{Topsoil}$ = 791; $n_{Subsoil}$ = 810)”

R_01: Figure 2: Please prefer unit (g kg-1 soil) than

A: With this manuscript we do not only target the soil science community, but also those of biogeochemistry and geochemistry. The two latter groups generally use wt-% instead of g kg^{-1} soil, since the data are compositional. If wt-% is accepted by SOIL, we prefer to keep the unit as it is. Otherwise we can change the units for SOC, Al_{ox}, and Fe_{ox} from wt-% to g kg^{-1} soil (wt-% multiple by 10) in the entire manuscript, the supplement and the corresponding R markdown file.

R_01: Ln. 250 could you give some essential data from the Table B2 in the Ms. It could be nice for the ones who do not go and see the supplement data.

A: Table B2 (line 807) in the supplement contains the full linear mixed effects model results of the standardized and normalized fixed effect parameters for the entire data set (n = 1,601). This means that it also contains the non-significant parameters, like land cover, Fe_{ox} and clay + fine silt content, in contrast to Table 2 in the main text. The differences between the land cover groups (which can be seen in Table B2) can also been seen in Figure 2 in the main text. Therefore, we prefer not to move the content of Table B2 into the main text.

R_01: Table 2: "Depth (Subsoil)" is not clear. It is clear when reading ln 246-247, but it is not when reading the table alone. Table should explicit by itself.

A: When using categories in linear mixed effects models, one category, in our case topsoil, is used as a reference and the other category, in our case subsoil, is tested against the reference category. If it is significant, this means that the mean of the tested category (subsoil) is different from the mean of the reference category (topsoil). We agree that this is not clear when only looking at Table 2. To make it less confusing, we suggest to remove the word Subsoil from the Table (Line 282).

R_01: Figure 3: I did not get the sentence "Note that the x-axis is truncated…" At what Caex content value it should end? 76? Please specify.

A: Line 299: Figure 3 caption: "Note that x-axis for is truncated for improved visualization, which removes 3 data points (Ca_{ex} = 53.91, 54.58, and 75.66 $cmol^+/kg$);"

R_01: Ln. 286-287 and 301-302 Please make it clearer.

A: Some of the confusion was likely caused by the fact that we referred to the wrong figure in line 319: it should be Figure A6a instead of Figure 5b. We apologize for any inconvenience this might have caused.

Line 317: "The root median square error (RMSE) for the topsoil regression tree was 1.47 wt-% (range: 0.80–3.11 wt-%) and for the subsoil regression tree was 0.67 wt % (range: 0.44–2.26 wt-%); the relative RMSEs were 0.65% and 0.48%, respectively. In the topsoil regression tree (Figure A6a) Fe_{ox}, MAT and PET/MAP were the most important predictors to split and explain variation in SOC concentration."

Line 333: "The random forest models had a RMSE of 1.31 wt-% and a R^2 of 0.70 for the topsoil samples, and for the subsoil samples a RMSE of 0.87 wt-% and a R^2 of 0.72."

R_01: Ln. 304 yes we could already notice that point on Figure 3c according to the pH class (20 or 30 cmol+/kg)

A: It is true that this point could already be noticed on Figure 3c. However, Figure 3c shows the raw data, while the sentence in line 334 ("However, in subsoils, the predictive power of Ca_{ex} is reduced – the predicted SOC content is relatively uniform above Ca_{ex} concentrations of about 20 $cmol^+/kg$ (Figure 5b).") and Figure 5b refers to the random forest model results that confirm the pattern observed within the raw data.

R_01: Ln. 321-322 it is not very compelling. You have shown previously that there are slight differences between 0-20 and 20-50 cm (Ln 294, 253..) Please be clearer.

A: With this section, we are trying to discuss and explain why the differences between topsoil and subsoil samples might be so small. Based on an additional comment from Reviewer #2, we changed the structure of this section:

A: Line 349: “Here, we focus on those variables that showed the most explanatory power in terms of SOC content across all models. We then compare their explanatory power with those reported in other studies for different regions. Short-range order minerals (Al_{ox}) and to some extent Fe_{ox} explained much of the variation in SOC concentration in wet regions with acidic and highly weathered soils. In contrast, Ca_{ex} explained much of the variation in dry regions, dominated by alkaline and less weathered soils. In addition, we discuss the role of clay and fine silt content, and of land cover, since they were important in other studies. However, in our study, the latter did not explain much of the variation in SOC content, which may be due to the large spatial scale. At the end of this section, we discuss the underlying geographic patterns that emerged in the data.

Some common predictors of SOC and dependencies between predictors (MAP/PET, pH_{H20}, CIA) emerged across all modeling approaches. Number of wet months, soil pH_{H2O} and weathering status (Figures 3 and 4) occurred as key parameters in the linear mixed effects models that influence how other parameters, such as Ca_{ex}, Al_{ox} and Fe_{ox}, explain SOC content variation across sub-Saharan Africa. In contrast, predictor differences were much smaller between topsoil (0–20 cm) and subsoil (20–50 cm) samples. This may partly be due to the large depth increments for each of the two sampling depths. However, since the identified SOC-controlling factors were similar for both depth layers, any differences were probably mostly driven by the fact that subsoil samples usually contain less SOC due to lower C inputs at depth (Jobbágy and Jackson 2000). Soil erosion at some sites might also dilute differences between the two depth layers, since water and wind can permanently remove surface soil.

These findings were supported by the regression trees (Figure A6) and partial dependence plots (Figure 5), where Ca_{ex}, Al_{ox} and Fe_{ox} seemed to be more important in explaining the variation of SOC concentration compared to pH_{H2O}, PET/MAP and CIA. For example, soil pH_{H2O} was important in the full linear mixed effects model, yet it mainly influenced Ca_{ex}, Al_{ox}, and Fe_{ox} concentrations in correlation with MAP (Figure 3d); the same was true for weathering (Figure 4b). Similar relationships have been found for temperate regions, where the importance of Ca_{ex} increased with increasing pH_{H2O} and decreasing precipitation, whereas the opposite was true for Al_{ox} (Oades 1988; Rasmussen et al. 2018). However, Rasmussen et al. (2018) did not identify Fe_{ox} as an important predictor of SOC content.“

R_01: Ln. 334 Please specify the size of the clay and fine silt fraction of the previously studies and yours. Discuss with these differences in mind.

A: We tried to be more explicit in regard to the cut-offs for the clay and clay+silt fraction that have been used by other studies. However, we could not always find the exact cut-off used in those studies (e.g. Quesada et al. 2020; Rasmussen et al. 2018).

A: Line 419: “Clay and fine silt content (<8 µm) was never an important predictor of SOC concentration within our different models. This is in contrast to some earlier studies that indicated that total clay content explains a large proportion of SOC storage and stabilization due to the sorption of soil organic matter to surfaces of clay minerals and building of aggregates (e.g. Amelung et al. 1998; Kahle et al. 2002). The relationship between SOC and total clay content is used in various models to describe turnover and storage of SOC. However, this simplified correlation may not account for the different stabilization mechanisms related to various clay minerals, e.g. 1:1 vs 2:1 clay minerals (Oades 1988). There are contradictory results on whether clay content explains the variation in subtropical and tropical soils or not. For example, Bruun et al. (2010) showed for various tropical soils that clay mineralogy, Fe_{ox} and Al_{ox} are better explanatory variables for SOC content than clay content (<2 µm) alone. In contrast, Quesada et al. (2020) found a strong relationship between clay and SOC content for highly weathered soils in the Amazon Basin that are dominated by 1:1 clay minerals such as kaolinite, whereas soils in the same system, dominated by 2:1 clay minerals, showed stronger relationship between SOC and Al species. In a comparison between tropical and temperate soils, Six et al. (2002b) found that less C was associated with the clay and silt fraction (<20 µm) in tropical soils than in temperate soils. Even though different studies used various cut-offs to define the clay (<2 µm), clay and fine silt (<8 µm), and clay and silt fraction (<20 µm), they all show that the relationship with SOC can be complex in subtropical and tropical soils.

Due to the broad spatial scale, soils in the AfSIS data set contain different clay minerals (Butler et al. 2020). No clear relationship between clay and fine silt content (<8 µm) and SOC concentration was observed in the models, although the raw data indicates overall a positive trend between clay and fine silt content (<8 µm) and SOC concentration (Figure 2b). This overall positive relationship does not hold across all sites (Figure 2c) and variable relationships across individual sites (Figure A4) may explain the low predictive power of clay and fine silt content in this data set. Instead, variables that better capture the different behavior of clay-sized minerals, e.g. Ca_{ex}, Fe_{ox} and Al_{ox}, are likely more suitable soil parameters to explain the variation of SOC content – even in highly weathered soils across sub-Saharan Africa. This is supported by the fact that a clay and fine silt-only model resulted in a very small R^2 (linear mixed effects model: 0.01; random forest: 0.12; Table B8)."

R_01: Ln. 360-361 Yes it is OK, but please end the §by your results and not by the results of other scientists. Your result seemed to show that this mechanism seem to be stronger to explain SOC stabilization than straight organo-mineral association in quite different tropical environment. There is no words on particulate organic matter in your discussion. Have you any data to inform the role or the proportion of this organic matter pool in the total SOC content in your soil samples.

A: Unfortunately, we do not have any data about the particulate organic matter pool and thus, cannot discuss it in our manuscript. However, it is likely that some of those processes are integrated within the topsoil samples due to the relatively large sampling depth interval (0–20 cm). We switched the last two sentences to finish this section with our own results:

Line 391: "In conclusion, the important role of Ca_{ex} in our data set can probably linked to the bridging of clays and organic material by Ca^{2+} to form organo-mineral complexes that are physically, chemically and biologically stable (Muneer and Oades 1989; Oades 1988; Rowley et al. 2018). Whenever Ca_{ex} was present in the soil system, it had a strong and positive relationship with SOC, independent of acidity and weathering status of the soil, as well as the available amount of water in the region."

R_01: Here is some bibliography you could read in relation with your study. Perhaps that could help in the discussion:

- Barthès et al. (2008) in Geoderma (to discuss about Al, Fe oxides in tropical soils)
- Beare et al. (2014) in Biogeochemistry (to discuss about oxides and clay content)
- Fujisaki et al. (2018) in Geoderma and in AEE (to discuss about clay content, aridity index and SOC content and SOC storage in tropical soils specially)
- Wang et al. (2016) in Biogeoscience (to discuss about aridity index)
- Chenu and Plante (2006) in European Journal of Soil Science (to discuss about clay and fine silt content at 0-8 µm and 0-20 µm).

A: Thank you very much for all literature suggestions. We read them with great interest and added some of them (i.e. Barthès et al. 2008; Fujisaki et al. 2018a; Fujisaki et al. 2018b) to the introduction and discussion section. Based on the suggested literature we found an additional reference that is also relevant for our manuscript, i.e. Vanlauwe et al. (2015).

References

Amelung W, Zech W, Zhang X, Follett RF, Tiessen H, Knox E, Flach K-W (1998) Carbon, Nitrogen, and Sulfur Pools in Particle-Size Fractions as Influenced by Climate. *Soil Science Society of America Journal,* 62(1): pp. 172-181, doi: 10.2136/sssaj1998.03615995006200010023x.

Barthès BG, Kouakoua E, Larré-Larrouy M-C, Razafimbelo TM, de Luca EF, Azontonde A, Neves CSVJ, de Freitas PL, Feller CL (2008) Texture and sesquioxide effects on water-stable aggregates and organic matter in some tropical soils. *Geoderma,* 143(1): pp. 14-25, doi: 10.1016/j.geoderma.2007.10.003.

Breiman L, Friedman J, Stone CJ, Olshen RA (1984) *Classification and Regression Trees.* Taylor & Francis, pp.°368.

Bruun TB, Elberling B, Christensen BT (2010) Lability of soil organic carbon in tropical soils with different clay minerals. *Soil Biology and Biochemistry,* 42(6): pp. 888-895, doi: 10.1016/j.soilbio.2010.01.009.

Buringh P (1970) *Introduction to the study of soils in tropical and subtropical regions.* Centre for Agricultural Publishing and Documentation: Wageningen, Netherlands, pp.°99.

Butler BM, Palarea-Albaladejo J, Shepherd KD, Nyambura KM, Towett EK, Sila AM, Hillier S (2020) Mineral–nutrient relationships in African soils assessed using cluster analysis of X-ray powder diffraction patterns and compositional methods. *Geoderma,* 375: pp. 114474, doi: 10.1016/j.geoderma.2020.114474.

Feng W, Plante AF, Six J (2013) Improving estimates of maximal organic carbon stabilization by fine soil particles. *Biogeochemistry,* 112(1): pp. 81-93, doi: 10.1007/s10533-011-9679-7.

Fujisaki K, Chapuis-Lardy L, Albrecht A, Razafimbelo T, Chotte J-L, Chevallier T (2018a) Data synthesis of carbon distribution in particle size fractions of tropical soils: Implications for soil carbon storage potential in croplands. *Geoderma,* 313: pp. 41-51, doi: 10.1016/j.geoderma.2017.10.010.

Fujisaki K, Chevallier T, Chapuis-Lardy L, Albrecht A, Razafimbelo T, Masse D, Ndour YB, Chotte J-L (2018b) Soil carbon stock changes in tropical croplands are mainly driven by carbon inputs: A synthesis. *Agriculture, Ecosystems & Environment,* 259: pp. 147-158, doi: 10.1016/j.agee.2017.12.008.

Hartmann J, Moosdorf N (2012) The new global lithological map database GLiM: A representation of rock properties at the Earth surface. *Geochemistry, Geophysics, Geosystems,* 13(12): pp., doi: 10.1029/2012gc004370.

Hassink J (1997) The capacity of soils to preserve organic C and N by their association with clay and silt particles. *Plant and Soil,* 191(1): pp. 77-87, doi: 10.1023/A:1004213929699.

Holmes KW, Kyriakidis PC, Chadwick OA, Soares JV, Roberts DA (2005) Multi-scale variability in tropical soil nutrients following land-cover change. *Biogeochemistry,* 74(2): pp. 173-203, doi: 10.1007/s10533-004-3544-x.

Holmes KW, Roberts DA, Sweeney S, Numata I, Matricardi E, Biggs TW, Batista G, Chadwick OA (2004) Soil databases and the problem of establishing regional biogeochemical trends. *Global Change Biology,* 10(5): pp. 796-814, doi: 10.1111/j.1529-8817.2003.00753.x.

Jobbágy EG, Jackson RB (2000) The vertical distribution of soil organic carbon and its relation to climate and vegetation. *Ecological Applications,* 10(2): pp. 423-436, doi: 10.1890/1051-0761(2000)010[0423:Tvdoso]2.0.Co;2.

Jones A, Breuning-Madsen H, Brossard M, Dampha A, Deckers J, Dewitte O, Gallali T, Hallett S, Jones R, Kilasara M, Le Roux P, Michéli E, Montanarella L, Spaargaren O, Thiombiano L, van Ranst E, Yemefack M, Zougmore R (eds., 2013) *Soil Atlas of Africa.* Publications Office of the European Union: Luxembourg, pp. 176.

Kahle M, Kleber M, Jahn R (2002) Carbon storage in loess derived surface soils from Central Germany: Influence of mineral phase variables. *Journal of Plant Nutrition and Soil Science,* 165(2): pp. 141-149, doi: 10.1002/1522-2624(200204)165:2<141::Aid-jpln141>3.0.Co;2-x.

Kennard RW, Stone LA (1969) Computer Aided Design of Experiments. *Technometrics,* 11(1): pp. 137-148, doi: 10.1080/00401706.1969.10490666.

Malick BML, Ishiga H (2016) Geochemical Classification and Determination of Maturity Source Weathering in Beach Sands of Eastern San' in Coast, Tango Peninsula, and Wakasa Bay, Japan. *Earth Science Research,* 5(1): pp. 44-56, doi: 10.5539/esr.v5n1p44.

McLennan SM (1993) Weathering and Global Denudation. *Journal of Geology,* 101(2): pp. 295-303, doi: 10.1086/648222.

Muneer M, Oades J (1989) The role of Ca-organic interactions in soil aggregate stability .III. Mechanisms and models. *Soil Research,* 27(2): pp. 411-423, doi: 10.1071/SR9890411.

Nesbit HW, Young GM (1982) Early Proterozoic climates and plate motions inferred from major element chemistry of lutites. *Nature,* 299: pp. 715-717, doi: 10.1038/299715a0.

Oades JM (1988) The retention of organic matter in soils. *Biogeochemistry,* 5(1): pp. 35-70, doi: 10.1007/BF02180317.

Quesada CA, Paz C, Oblitas Mendoza E, Phillips OL, Saiz G, Lloyd J (2020) Variations in soil chemical and physical properties explain basin-wide Amazon forest soil carbon concentrations. *SOIL,* 6(1): pp. 53-88, doi: 10.5194/soil-6-53-2020.

Rasmussen C, Heckman K, Wieder WR, Keiluweit M, Lawrence CR, Berhe AA, Blankinship JC, Crow SE, Druhan JL, Hicks Pries CE, Marin-Spiotta E, Plante AF, Schädel C, Schimel JP, Sierra CA, Thompson A, Wagai R (2018) Beyond clay: towards an improved set of variables for predicting soil organic matter content. *Biogeochemistry,* 137(3): pp. 297-306, doi: 10.1007/s10533-018-0424-3.

Rowley MC, Grand S, Verrecchia ÉP (2018) Calcium-mediated stabilisation of soil organic carbon. *Biogeochemistry,* 137(1): pp. 27-49, doi: 10.1007/s10533-017-0410-1.

Schlüter T (2008) *Geological Atlas of Africa.* Springer: Heidelberg, Germany, pp.°307.

Schmidt MWI, Torn MS, Abiven S, Dittmar T, Guggenberger G, Janssens IA, Kleber M, Kögel-Knabner I, Lehmann J, Manning DAC, Nannipieri P, Rasse DP, Weiner S, Trumbore SE (2011) Persistence of soil organic matter as an ecosystem property. *Nature,* 478: pp. 49, doi: 10.1038/nature10386.

Six J, Conant RT, Paul EA, Paustian K (2002a) Review: Stabilization mechanisms of soil organic matter: Implications for C-saturation of soils. *Plant and Soil,* 241(2): pp. 155-176, doi: 10.1023/A:1016125726789.

Six J, Feller C, Denef K, Ogle SM, de Moraes Sa JC, Albrecht A (2002b) Soil organic matter, biota and aggregation in temperate and tropical soils - Effects of no-tillage. *Agronomie,* 22(7-8): pp. 755-775, doi: 10.1051/agro:2002043.

Terhoeven-Urselmans T, Vågen T-G, Spaargaren O, Shepherd KD (2010) Prediction of Soil Fertility Properties from a Globally Distributed Soil Mid-Infrared Spectral Library. *Soil Science Society of America Journal,* 74(5): pp. 1792-1799, doi: 10.2136/sssaj2009.0218.

Vågen T-G, Walsh MG, Shepherd KD (2006) Stable isotopes for characterisation of trends in soil carbon following deforestation and land use change in the highlands of Madagascar. *Geoderma,* 135: pp. 133-139, doi: 10.1016/j.geoderma.2005.11.012.

Vanlauwe B, Descheemaeker K, Giller KE, Huising J, Merckx R, Nziguheba G, Wendt J, Zingore S (2015) Integrated soil fertility management in sub-Saharan Africa: unravelling local adaptation. *SOIL,* 1(1): pp. 491-508, doi: 10.5194/soil-1-491-2015.