Abstract :

Line 13 : explain in full words the term LCCC

LCCC explicitly defined as Lin's concordance correlation coefficient

Introduction :

Lines 48-49 : state (if relevant) that an initial low soil water content increases slaking.

Added "and soils of low initial water content more prone to rapid and explosive slaking"

Lines 76-78 : add that in the paper by Annabi et al., 2017 the method used to measure soil aggregate stability is the normalized method(ISO/DIS 10930, 2012), which is time and cost consuming, which is not the case of the SLAKES approach.

Added "Tools that make aggregate stability quantification accessible, such as the SLAKES application, may facilitate the production of such maps." Detractions of wet-sieving and simulated rainfall techniques were added at line 59 at the request of RC1.

Methodology

Lines 93-95 : please refer to the WRB soil classification as the Australian classification is unknown by most of readers.

Australian Soil Classification has been removed and text changed to: The soils of the floodplain area at L'lara are classified as Vertisols according to the World Reference Base for Soil Resources, with some expression of calcic horizons (IUSS Working Group WRB, 2015). The sand hill area is represented by Luvisol, Lixisol, Solonetz, Leptosol and Regosol soil groups.

Lines 93-100 : it would be interesting to present the soil and landuse maps of the study area, as they are primary drivers of soil aggregate stability. These maps would be very useful to help the reader interpret the SI maps you present later in the paper. These data are moreover used for soil sampling as input parameters.

A soil type map was produced but unfortunately only covers L'lara and it is in the Australian Soil Classification and we have been requested to use WRB. MrVBF, NDVI and land use maps have been added to Figure 1. The MrVBF map gives a good indication of the distribution of Vertisols versus other soil with a sandy topsoil.

Lines 108-119 : the reading of this paragraph is not straightforward, as the sampling strategy is quite complex. I think the 108 samples described lines 108 to 116 should be introduced by a short sentence line 108, such as for example : "A training set of 108 samples and a test set of 50 samples were defined. The training set comprises 58 on- and 50 off-farm samples."

Text has been modified for clarity.

Lines 112-113 : why are the input parameters for the sampling strategy different for off-farm samples ? Is it due to the fact that a soil map is not available ? This could be mentioned.

Correct, the soil map was only available on-farm. Text has been adjusted accordingly.

Lines 113-114 : I do not understand on which sampling set the K-means clustering is applied, and for what purpose.

K-means was the stratification method for stratified random sampling to identify off-farm samples. Text has been updated for clarity.

Line 130 : why are 20 to 30 soil aggregates necessary for the slaking test, as only 3 aggregates are necessary for the test, and the test is repeated three times at most ?

Text has been changed to “12 to 15” aggregates. While we did not need to repeat the test more than three times, however the application did crash sometimes and the test could be compromised if a shadow was inadvertently cast over the sample while analysing so it is recommended to have some spare aggregates to run additional tests. Note this has been moved to section 2.3 at the request of RC1.

Lines 144-146 : I think it is important to provide information on the repeatability of the measurements, e.g. to ensure the average value calculated for the SI is representative of the whole sample SI. Indeed, the SI is calculated on 3 aggregates, which could be considered as a low number. It is therefore important that you provide at least a graph with the distribution of the differences in SI values for the 108 samples, including 'outlier readings'. In that respect, and to further explore the representativity of the measured aggregates, it would be interesting to present the values of the 'a' coefficient for each aggregate that is tested.

Modified text: “An additional reading was required for approximately 20\% of samples and was more commonly required for soils with higher slaking index values compared to samples which exhibited minimal slaking. When additional readings were taken the outlier reading was discarded and remaining readings averaged to provide the final SI for each sample.”

The graph you mention would be great to have but unfortunately the data was collected by different people over a number of months. Some reported every scan taken including replicates and outliers

for each sample, others only the final two replicates used, and others only reported the final averaged value. I will ensure that all scans are recorded and look to include such a graph in future publications, but I am reluctant to publish the incomplete dataset here.

The version of the app used reports the slaking index for each aggregate after the 10 minute analysis time but not the 'a' coefficient for each aggregate, this may be introduced in later versions of the app though.

Line 145 : I do not understand what are these 'outlier readings', and on what basis they could be discarded.

When an additional reading was taken it was always within one unit of one of the original duplicates. The additional sample and the duplicate within one unit were then averaged to give the final slaking index value and the other duplicate was treated as an outlier and not used in the calculation.

Line 175 : what is the unit of the aspect ? How did you go around the circular nature of the variable ?

Degrees symbol added to the table. The variable was not found to be a significant predictor when left in degrees or when aspect was investigated as a cardinal direction factor.

Results :

Line 209 : you state that some aggregates "increased in size by 730%". As I understand it, it is not the actual increase that is measured after 10 mn of immersion at the end of the SLAKES experiment, but rather a final aggregate size using the Gompertz function at t=∞.

Correct. Added "is projected to increase".

Line 210-211 : you mention that all SI values are below the maximum theoretical value of 7.8 suggested by Fajardo et al. (2016). What about the 'outlier readings' you mentioned line 145 ? This should be clarified.

All reasonable results were below this threshold. At times when a shadow was inadvertently cast over the petri dish values of >1,000 were reported but these were discarded.

Line 246 : just to make sure, you mention average SI values, is it an average or a median value ?

It is average value. The value returned from the app is the average of the three aggregates analysed and then we average the value from duplicate tests to achieve the final value.

Line 261 : make reference to Table 2.

Reference to Table 3 added

Lines 302-304 : is there a way to account for the uncertainty due to the (relatively weak) regression applied for the mapping ?

Points added to the discussion – “Another contributing factor for the improved validation metrics under increased OC scenarios is due to the SI values being based on modelled data from which unexplained error has been removed. Future efforts should account for the error of the underlying regression equations and quantify the uncertainty of the resultant maps by bootstrapping and applying random error based on the the prediction variance of the underlying regression equations.”

Lines 340-341 : this assumption is not straightforward, and requires to provide a soil and landuse map.

Land use, MrVBF and NDVI maps have been added to Figure 1 to facilitate interpretation.

Lines 345-346 : the same deals for MrVBF : a MrVBF map would help the reader.

Land use, MrVBF and NDVI maps have been added to Figure 1 to facilitate interpretation.

Lines 362-363 : I do not think the mapping of SI change is the main result that "shows" the benefit of increasing soil OC on SI values. This was shown by the results leading to Figure 4. Here, the mapping allows to precisely locate where there is a real benefit to increase soil OC to increase aggregate stability.

Correct. Sentence changed to “The produced maps highlight areas that are expected to have lower SI when OC levels are increased”.

Figures, tables :

Figure 1 : the black lines (bold and not bold) on the map are not defined in the legend.

Description of lines has been added to the Fig. 1 caption as well as for Figs 6 and 7.

Table 3 : for readability, emphasize in bold characters the correlations that are significant at a given confidence level.

Bold font has been used to indicates correlation with significance at $p < 0.05$ and the table caption updated accordingly.

Minor edits :

Line 200 : "[...] on SI has been investigated"

Amended

Line 240 : "[...] natural vegetation (Fajardo et al., 2016 ; Flynn et al., 2020)."

Amended

Line 244 : remove "In a review of"

Amended

Line 267 : remove "3)"

Amended. This was an incomplete reference to Fig. 3

Line 283 : remove one "been"

Amended

Line 354 : remove"under"

Amended

Line 356 : remove "be"

Amended

Line 381 : "through the use of"

Amended