REGEN
REGISTRY

Expert Peer Review for Ecometric

Methodology for Methodology Proposal for Soil Organic Carbon Estimation in Regenerative Cropping and Managed Grasslands Ecosystems

Date: December 20th, 2022
Reviewer: Prof Stuart Marsh

Summary of Internal Review Process

The intent of the Expert Peer Review is to ensure methodologies meet the integrity expected by our community and ensure the methodology will work to regenerate ecosystems. The task of an Expert Peer Reviewer is to act as an ally to methodology developers by providing critical feedback to help facilitate an understanding of how to improve the methodology to best serve Earth Stewards while maintaining scientific and community integrity.

The *Methodology Proposal for Soil Organic Carbon Estimation in Regenerative Cropping and Managed Grasslands Ecosystems* has been reviewed and feedback has been provided in two ways:

1) **Overall Reflections**: To provide feedback to your methodology, comments and reflections are recorded in this document.

REGEN
REGISTRY

Expert Peer Review:

General Comments:

Overall, a lot of thought has gone into developing a workable methodology to measure SOC in a consistent fashion. The proposers stand a reasonable chance of generating some interesting results if they follow the defined methodology closely. There are some areas of detail that could be looked at so as to improve the outcomes. But before getting into that, here are some general comments to think about in terms of the annual cycle and the processing methodology.

You have defined an annual monitoring process – how will you pick up on seasonal effects? SOC can be expected to vary through the year – we see that in peatlands, for example, which appear to "breath" – so, how will you know which part of the annual cycle you are sampling? Always at a high point? A low point? Transitional? What is the SOC behavior through the year?
The inclusion of 12-month time intervals covers this, as even if sample timing hit a seasonal effect, the remeasurement timing would be at a comparable point and so the SOC change would be unaffected. Although our temperate latitude preference is for winter sampling when the soil is more dormant, we cannot state this as a defined criteria as it will not work closer to the equator. Our judgement is that by sampling at the same time of year 12-months apart the delta will be unaffected by seasonal changes.

Still on the subject of the annual cycle, you are then taking a +/- 4 month period, which gives a total 8-month window. This means that some imagery could come from early spring (e.g. March) and some from autumn (e.g. November) – this is the majority of the year and it encompasses a wide range of soil conditions and vegetation growth stages. Do you need to tighten it up a bit? I assume that this length of time has been taken so that you can guarantee acquiring cloud free imagery, but have you got the balance right in terms of the spectral quality of the data?
The inclusion of image selection as close as temporally possible to the sampling date shows our intention, while the +/- 4 months is simply to give flexibility during periods of prolonged cloud cover. We are working on active sensor inclusion to mitigate. This is something we will monitor by recording imagery and sampling date.

You give a comprehensive review of methodologies and then pick ML? Do you know how this compares with all the other methods? Was it chosen because it is the best or for other reasons? This wasn't quite clear to me. Which method has performed the best historically and, if you have not chosen it, then why not? Is there some other advantage to ML that trumps its performance, such as repeatability, automation, etc? I think this is a good section, but you can be a bit clearer about exactly how you came to the choices that you did. This is important, because everything else is founded on it, and so you really need to have clarity here.

The choice of ANNs methods was made by research into their advantages and disadvantages and by comparison of results from other methods. Paragraphs on advantages and disadvantages have been added to 3.3.

Finally, I wonder if there is a contradiction in the sampling strategy avoiding vegetation but the remote sensing strategy utilizing spectral information, which is largely driven by the vegetation response? Yet you use the samples to drive the ML algorithm, tying the two together…
A dependency of vegetation health is soil type and condition, including organic matter which influences the SOC levels.

Comments by Sections:

Methodology Overview:

My main comments here relate to the period being sampled. If this is to be compared year to year then it needs to have several key characteristics:

1. It needs to be consistently measurable through different years i.e. you can't measure it one year and then fail to the next due to cloud cover, for example.
   In this event we would aim to increase the measurement interval to 24 months and SOC stock change would be assessed against crop type / production type of comparable systems to allow the project team to interpret.
2. It needs to be representative of a known or average SOC state, that has a meaning in terms of the overall SOC through time
   Our report format compares geographical averages – in the UK the source data is NRM laboratory averages of farming system / habitat type from all soil sample results over the previous 12 months. As we gather more data over different soil and crop types, we will be able to subdivide averages into crop and soil type brackets. This is an important part of our interpretation support to the project / farm management team to help them make evidence-based decisions for following cropping cycles.
3. It needs to be reproducible, so that the same snapshot can be measured year on year I think you need to carefully examine some of your choices and ask whether they meet or miss these key criteria. Is 8 months too long? What are the seasonal effects? Is there a time period that has consistently low cloud cover? Can the plan be designed to be both tighter, i.e. covering a shorter period, but also having a higher chance of success?
   This is addressed by our 12-month sampling interval, the 8-month example is I assume relating to the +/-4month imagery date which hasn't yet proved a problem and we have consistently been able to source imagery within +/-1month of sampling date in UK, N Europe

and NZ. We are also looking at aggregated met data covering the 12-month interval to work out key comparability metrics.

Another possible was around the problem might be to use other optical data, so that more days are open to sensing in any given period. If you use multiple optical sensors, you could have two or more acquisitions a day – is this a better way around the cloud issue? Or does it make the ML aspects more challenging? Worth thinking about.
It is the intention to use additional sensors if further research indicates a benefit. Possible sensors might include Sentinel-1 radar imagery to address the problems of cloud cover.

Project Boundaries:

I assume that you will be selective in deciding which types of land cover you will include (e.g., a forest, a field) and which you will exclude (e.g. urban, industrial)? I think you should go into a bit more detail about exactly how you plan to define these boundaries. Then, once you come up with the masks, I assume that you will stick with them.
The methodology as titled only currently covers Regenerative cropping and managed grasslands.

The temporal boundaries seem more obvious, but then you talk about disasters and potentially changing your regime to reflect events. It is not obvious to me that this is a good idea. How will you compare sudden events to the long-term trend? I worry that you might spend a long time working out how an extreme event fits rather than what is the long-term development of SOC. Maybe retain this is a longer-term aspiration, but only to be implemented once you have got a good understanding of the year-on-year position?
A good suggestion – we will report results graphically with each annual result set adding a bar to allow visualised change vectors. This will contextualise any abnormal results and allow diagnosis in consultation with the land steward experts.

Calculating Carbon Sequestration:

You undertake quite a good review of the various remote sensing methods that have been used to calculate SOC. What I don't get from your review is any sense of which method has been the most successful. You describe several and their relative outcomes and merits, but no comparison is made between them. Then you plump for ML via ANN. Is this better than the others? I would have liked to see some kind of ranking given, and then a rationale for selection of the ML/ANN approach described. This might not be that it is the best, you may have other reasons for using this approach such as it can be easily automated. But this needs to be discussed in more detail.

The choice of ANNs methods was made by research into their advantages and disadvantages and by comparison of results from other methods. Paragraphs on advantages and disadvantages have been added to 3.3.

Once established, the selected methodology seems sound and is well described. You might like to justify why samples will be within one month of any ancillary data – is this important in your 8-month window for imagery? These dates should probably align in some way.
The soil sample dates and the sample dates for the ancillary data, where relevant, will be chosen to be temporally close.

Calculating the Creditable Carbon Change:

You assume a reduction in CO2e. Is this reasonable? Could it not go up? Would it not be better to retain an open mind about the likely direction of travel, when describing such things.
No predictive assumption is made on the direction of travel, in fact to do so would introduce a significant business risk of challenge if the predicted change didn't materialise. Our role is to report the results scientifically to the project team. It is the project team's responsibility to interpret and apply the results as evidence-based data to diagnose what has happened and why.

I like the analysis of uncertainty and you go into some detail on the sample uncertainty, but you don't explain exactly how you will assess uncertainty in the ANN? Is there a recognized method?
MAPE

Data Reporting:

No comment.

Data Storage:

No comment.

Data Verification:

No comment.

Final Decision: - Pass. I think that ecometric can pick up on the points that I have made – none are showstoppers – and address them in implementing their plan. I would be happy to take

another look in the New Year, if that would be helpful, but I don't think we should hold them up from getting on with it in the meantime.