# Interactive comment on "Ideas and perspectives: enhancing the impact of the FLUXNET network of eddy covariance sites" *by* Dario Papale

**Jason Beringer (Referee)**

jason.beringer@uwa.edu.au

I am the current director of OzFlux and I bring this perspective but these comments are personal ones and not necessarily the views of the network or NCRIS TERN.

I wholeheartedly agree in principle with the philosophy of new approach to FLUXNET is terms of ensuring timely data, interoperability and to facilitate data use. We do need a new organisational structure to make sure that FLUXNET is sustainable into the future. In the past FLUXNET collections have only been possible due to the herculin efforts and passion of a few people and we are all extremely grateful for this but it is a big sacrifice that you and others have made.

In addition to the demand for real time data that you have outlined, there is also

an emerging area of ecological forecasting, for example, The EFI-NEON Research Coordination Network that is an NSF project to create a community of practice that builds capacity for ecological forecasting using NEON. There is also potential demand from data assimilation of flux tower data into short range weather forecast (e.g. https://doi.org/10.1175/MWR-D-19-0370.1). Also real-time agricultural monitoring should be possible. Finally, data can also be used to generate regional real time evapotranspiration estimates using a fusion of flux tower, remote sensing and modelling all of which have potential for use by the public.

With respect to the proposed re-organisation from a single large database into subcollections. On the one hand this helps make the network more sustainable by delegating work and responsibility to continental clusters so that not all the work is being done by a small group. On the other hand this will require continental clusters to be functional, accessible and have open data sharing policies. I would be concerned that many clusters may not have capacity to do this or there may be a difference in data sharing between groups and individuals, such that individual sites that may want to contribute are unable because of the inability of the cluster to participate for technical, personnel or other reasons.

Following on from this I can see that there could be many, many sites (even some that were not part of FLUXNET2015) that want to contribute to FLUXNET but they are unable too because they don't have a functional continental cluster. It will be crucial to make sure there is a mechanism for them to contribute. The ONEFLUX processing code it not designed for this purpose. So I wonder if there could be an online processing tool that registered site users could upload their site data at what ever intervals they are able to do so?

I'm not entirely sure for the rationale and need to move away from a large centralised data base to Continental cluster collections? There has been a lot of effort gone into making the database and I'm not sure if it is broken in some way or has reached its capacity technically. It seems to add another layer of complexity to have Continental

cluster collections and a shuttle they queries each of them. It then relies to the ability of the clusters to maintain the data in real time in these clusters. Why not just have a continually updated big database where data is added at any time (daily or annually as it becomes available). You then rely on sites and clusters to push the data continually rather than calling and hassling for sites to submit data (current model). Users can query the data anytime too. We could build a set of tools that allow data to be accessed and queried in more sophisticated ways.

I like the idea of a persistent identifier (PID) or something similar. I would envisage that initially the processing would follow the FLUXNET2015 (i.e. ONEFLUX) pipeline and the PID would reflect that. But processing methods do evolve and a Fluxnet steering group could endorse any changes to the pipeline and periodically the PID would change to say FLUXNET2025 for example.

It is probably important to think about changes in processing and reprocessing the whole database. This may well happen in the future if we have a new pipeline and you would want to apply the new pipeline retrospectively across all the data I presume? This would be manageable under a single large database but may be difficult under Continental cluster collections.

As Papale acknowledges, it will be important full flexibility of each single network to decide what to share and when in FLUXNET and the possibility to distribute different formats and versions through their Data portals. I would envisage that in OzFlux, one could have rapid processing of near real time fluxes using the FLUXNET pipeline on a daily basis and this would have PID that differentiated these data streams as 'beta' datastreams or something. These data can be used for applications that require real time data but the data comes with the caveats of not being quality controlled by site investigators. The data should be fit for the purpose that it is being applied. As Ray Leuning would tell us, "know thy site" and investigators would still produce the finalised data sets with human skill and site knowledge. So we can have our cake and eat it too.

C3

So I think there are potential pros and cons and maybe the best way is to get a working group together and flesh out an operational model for the future. I am very keen and you can count us in.

Interactive comment on Biogeosciences Discuss., https://doi.org/10.5194/bg-2020-211, 2020.