Biogeosciences Discussions

Biogeosciences Discuss., editor comment EC1
https://doi.org/10.5194/bg-2021-133-EC1, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on bg-2021-133

Andreas Ibrom (Editor)

Editor comment on "Representativeness assessment of the pan-Arctic eddy covariance site network and optimized future enhancements" by Martijn M. T. A. Pallandt et al., Biogeosciences Discuss., https://doi.org/10.5194/bg-2021-133-EC1, 2021

The manuscript is centered around the question on how representative a network of greenhouse gas (GHG) flux observation sites in the Arctic is. This is a valid and relevant question and the authors give several examples of research questions for which representativeness of flux sites for larger entities is an issue. The analysis can guide decisions where to optimally place future investment in additional field stations and into upgrading existing stations. With the metadata base on arctic flux stations, including those that are not yet part of any flux network, and the online data base tool to make this information generally accessible, the work provides a service to the arctic flux community and the users of its products.

The manuscript is well written and the presentations of the results work well.

A specific limitation of the manuscript is that the largely theoretical approach builds on a number of untested propositions, assumptions and hypotheses (see below), and thus the results remain speculative, i.e. mere results of the input. This sounds worse than it is, because from a practical point of view there might not be much to be done about it and this method may encompass one of the better available approaches, if one 'believes' at all in the possibility to extrapolate from as little as 120 + stations to an entire global zone (see section 4.3). Well, technically it is possible, as important publications have demonstrated, and for this reason alone this work might be important, to understand the risk of doing so.

- The main scientific problem is to estimate representativeness from a sparse sample. The ideal statistical approach would be working with a luxuriously large sample size and then estimating the minimum sample size at any given spatial distribution to achieve a defined acceptable uncertainty of the extrapolated numbers, but this is not possible for sparse samples. In this situation, the Authors make an elegant move, to analyse representativeness from ample but different dataset that represent the spatial variability at high resolution (1 km^2). But strictly speaking, this approach estimates representativeness for the 18 variables and their combinations and not the GHG fluxes. Whether the assumption that the ecological land cover classification is valid also for the GHG fluxes, is neither tested nor shown, even questioned in their section 4.3). The whole results rest on this untested hypothesis that spatial patterns of GHG flux variability are the same or at least similar to the spatial variability of the 18 variables as aggregated by their approach. A geostatistically rigorous approach would start with

showing the quantitative relationship between the site factor proxies (the 18 variables) and the observed fluxes and than use this to determine and minimize the upscaling error. This was not attempted.

- Using the *absolute* term "well represented" for the highest category of representativeness can indicate that the number, time series lengths and completeness of EC data is large enough to estimate the average flux of an ecoregion with sufficient certainty. But in fact this has not been shown. As it looks to me "well represented" means "well represented" relative to the upper 25% of the realized representativeness of EC stations and data (lines 164-169). But what does this mean for absolute representativeness? Is this amount of data sufficient for reasonable extrapolation/upscaling? I recommend using a relative term "higher relative representation", "lower relative representation" and discussion on, what this classification means for absolute representativeness. A possible result for the whole study might be that the network design should opt for increasing "relative representation" but this does not mean and does not be confused with sufficient representativeness for a certain purpose (e.g. upscaling).
- The number of ecoregions is subjectively chosen. This choice has a very high influence on the results. If that number of regions is very high, the representativeness of a given set of stations will be more extreme and vice versa. If there is no clear theory or data, one needs to simply choose a value, but then one needs to show, what the sensitivity of the results (and conclusions) to this choice is. I suggest a sensitivity study, where the number of regions is being gradually changed, maybe they find a domain where the calculated representativeness is relatively stable?
- If the authors agree that there are still large uncertainties in the outcomes from their study, I would like to ask them for presenting this message clearly in the text: a section in the discussion challenges for extrapolation to the artic through sample sparseness and prospects for continental and global extrapolation from eddy covariance flux tower networks in the Arctic and probably the globe, too. This uncertainty should also be mentioned as a main result from the study, in the conclusions and the abstract /and even in the title, if possible).

I do not say that this approach isn't useful, but, honestly, I wonder, what the scientific progress is beyond some practical value. A clear prospect from this study is that it provides guidance (even best possible based on available information) for the choice of locations of additional sites and especially the evaluation of alternative choices between establishing new sites versus upgrading existing sites. But the guidance rests on untested assumptions or hypotheses that cannot be tested. I can understand, why the authors do what they do, and they have certainly chosen one of the best possible and most careful approaches, but whether this is good enough to guide investment into research remains an open question. In order to not leave readers with unrealistic expectations, this risk should be discussed and mentioned at all visible places, including the abstract and the conclusions.

Detailed comments:

Abstract

24-26: Can you briefly add representativeness for what – for the driver field that you have, all same important, or for the results from the network (which, CO2, GHG, energy radiation etc.). Later you write "capture the spatio-temporal variability of surface atmosphere exchange fluxes across the pan-Arctic ecosystem distribution." (lines 77-78)

In general: Address the limitation of the approach that it is based on the unproven assumption and why it is still the best that we can do.

Introduction

39 please try to find a more accurate term than "vast"

64-67: Please elaborate either here or in the discussion, which conditions must be fulfilled for a site being representative for a complete region or class of ecosystems and how the coarseness of the classification system affects this representativeness (see section 4.3).

70 Please add "high precision concentration" in front of "tall tower networks" and note the fundamental difference in the approaches. BTW Scientifically, the definition of representativeness is much more straight forward for an atmospheric concentration network (concentration footprint coverage, improvement of posterior probability of fluxes estimated with model inversion) than for an ecosystem flux network with much more local representation. Contrasting these two approaches would be a very good example to illustrate the challenge with (and the limitations of) the upscaling of flux data to the large spatial scale.

79-80 "We quantify representativeness here based on the similarity in key ecosystem characteristics of any location in our domain to those of the EC sites." without showing that this prior assumption is true. This is a key limitation. to alleviate this, please provide evidence (or at least theory) that this approach is (likely) valid and test it with the data available (e.g. measured GHG fluxes in ecoregions are statistically different from measured GHG fluxes in contrasting ecoregions).

89-81 "This concept is similar to producing gridded products by upscaling localized flux data to a larger region". Please explain to make sure what is similar and what is different between your approach and the common upscaling approaches. It is not like Kriging or other geostatistical approaches. A mathematically rigorous geostatistical approach would, e.g., relate the quantitative relationship between the site proxies (i.e. your 18 variables) and the fluxes (e.g. confidence band for prediction) and then estimate the upscaling error in bootstrapping approaches etc. . If I am right, the proposed approach lacks this mathematical rigor. Please comment.

136-139 Please list and define the variables here and describe for each variable the relevance for determining the GHG fluxes. Then argue for the choice of so many variables, why 18? A sensitivity study for the choice of the variables (importance of the variables for establishing a stable spatial ecoregion patterns) could, e.g., support the choice.

141 Give the pixel size (1 km^2, if the same as in the reference)

BTW- in Tab. A1: GlobPermafrost doi: 10.1111/j.1461-0248.2004.00694.x and Saxon et. al. 2005 doi: 10.1111/j.1461-0248.2004.00694.x have the same DOI, is that an error?

146-164: In this reasoning the choice of k=100 is subjective. Translating the "gut feeling" terms of "truly coherent" or "would not grant much improvement" into clear scientific (statistical) concepts would alleviate this limitation. I am not an expert, but wouldn't a sensitivity study to relate (small) changes in K to the possible existence of a stable frequency distribution of representativeness values be of some help here? The interpretation would be the resolution matching the scale of natural variability.

159-163 A very good move to compare the results to independent information. Later I

suggest a small results section where you can give some statistical similarity figures between the ecoregions and CAVM. For interpretation of the comparison, please add to which degree the definition of CAVM units uses similar or different information compared to the proposed approach.

164-168 and following down to 178. Please, see my point 2 above. Note that requiring a minimum of 5 sites is a subjective choice and should be substantiated by statistical uncertainty parameters. Please specify the quality requirements for the 5 stations (CH4 fluxes, wintertime, length of observation period, etc.).

2.4 Many of these choices make a lot of sense to me, especially including the station quality and consequently contrasting upgrading versus new locations. For the role of the random sampling versus guided sampling (262-269), I wonder whether the improvement of the guided sampling over the random sample isn't a bit of a circular argument, as the output of the analysis is the consequence of the approach and the initial assumptions. But the degree, on how the difference between those two sampling modes are and how they develop with increasing samples, characterizes the approach and also the relationships in the data. Maybe one could help the reader by explaining this more explicitly.

Results

The section 3.1. is very well especially the combination of Figure 2 presents the temporal development of the data quality and quantity in a very intuitive way.

280 Figure 1: Although to me the high resolution map is esthetically very appealing, wouldn't a map with ecoregions as a background (same colour code as stations) include more useful information in this context?

326: "excellent data coverage", this raises a lot of good feelings, but isn't it just the "highest data coverage"? Is it also excellent? if yes, why?

Between section 3.1 and 3.2 I miss a section that presents the geostatistical results, e.g. a map of ecoregions, maybe including the comparison with the CAVM distribution (if you consider this, move 268-371 to this subsection).

Section 3.2 does a very good job to present the strengths of the approach, i.e. providing a spatial data base that is an excellent basis for both spatially resolved presentation and for calculating summary statistics. Much of what is presented is a logical consequence of the approach (e.g. the differences between ER1 and ER2 standards) but the concrete application to the data sets brings about some interesting features and quantifications that turn out to have some plausible possible causes.

345 it is interesting, that elevation (and topography) are not directly part of the site factor variables. It looks, as if this would be good and commonly available candidates.

373 replace "representatives" by "representativeness"

Section 3.3. is very interesting to read. You have tested a lot of plausible settings and compared the results, which gives the impression that, despite the different possibilities, the main results where to place new stations and where to upgrade them are relatively

uniform. This qualifies the robustness of your approach! (But is does not necessarily qualify the study to be robust, because of the hence unavoidable risk of systematic errors that an optimized network design improves the representativeness, but still does not achieve the level needed for an sufficiently accurate flux extrapolation to the Arctic.)

398: Fig. 5:

- please explain what "pre optimized" means. It doesn't show in the text.
- Mention the number of added stations are (n = 5, if I am right) and that the location of these new stations are optimized regarding the parameter of interest.
- How would the selection look like, if one optimizes for representativeness of all combined, maybe, weighted to optimally represent the global warming effect?

Otherwise fun to see, how higher representativeness areas develop around the newly proposed stations.

450 replace "manuscript" by "this study"

467 reconsider using 'higher' (relative term) instead of 'good' (absolute term) coverage, as it is not sure for what the coverage is sufficiently good.

482-497 Here you discuss rightly the limitations of your study, make sure that this is explained at the most visible parts of the manuscript (conclusions and abstract), too.

484 -486: Please show, how "the maps can be utilized as a measure representing the extrapolation uncertainty" … . Given what is written from lines 492 including "specific fluxes provided by the eddy covariance tower network based on these data layers must largely remain qualitative, since no clear quantitative linkage between the bioclimatic controls and the fluxes for CO2 and CH4 can be considered.", I would even doubt it.

489 either use "a priori" or "prior"

Section 4.3 The role of small scale variability is a very important, if not crucial aspect, for your study, because it questions one major assumption to this approach, i.e. that data from one site does represent data from other sites and even entire regions. Please also note that you define the meaning of subscale variability that is unaccounted for in the land-cover classification by the choice of the k value. As long as there is no clear assessment on how many towers are subject to small scale effects, the representativeness value of the tower network is highly speculative. I would also include here the length of the observation period and whether or not winter fluxes and CH4 fluxes are being measured. The required minimum length of observation periods should be estimated from typical time scales of interannual variability. Please make sure that readers get this information on the very substantial and serious challenges of upscaling from tower networks to large surface units.

527: Please explain, why can you assume something, when the information to substantiate any possible assumption is simply lacking? If I am right, such an assumption has no scientific value. Rather discuss which alternatives exist and what the lack of information means to the risk of drawing false conclusions from your results.

sub-section 4.4

good and practically important – no other comments

Conclusions

Please limit to conclusions on novel and important results and interpretations from your study. Most of what is said does not address these. The underrepresentation of Siberia and Canada was already noticed in other publications. I miss anything on consequences from the strengths and limitations of the proposed new approach and, equally important, what it means for the (humbling) prospects on extrapolating data from 120 flux sites to a global region.