Biogeosciences Discuss.,
doi:10.5194/bg-2016-495-RC1, 2016

Biogeosciences
Discussions
Open Access

Interactive comment on "Water, Energy, and Carbon with Artificial Neural Networks (WECANN): A statistically-based estimate of global surface turbulent fluxes using solar-induced fluorescence" *by* Seyed Hamed Alemohammad et al.

Anonymous Referee #1

Received and published: 20 December 2016

Review on BG-2016-495 General comments: The authors proposed a new global product of GPP, ET and H by using ANN. The manuscript is well written and the topic falls on to the scope of the journal. I do have several concerns.

First, the authors highlighted the use of SiF as input data. I see SiF was the only input data related to vegetation. Therefore, with/without SiF in WECANN must give different flux estimates. What happens if the authors use EVI or NDVI instead of SiF? Any significant difference in WECANN performance?

Printer-friendly version

Discussion paper

Second, what is the significant contribution from this work? Spatial (1 degree) and temporal (monthly) resolutions are too coarse. The approach is on the similar family of other machine learning methods (e.g. see Tramontana et al 2016 Biogeosciences). Stress the novelty of this manuscript. If there is any new discovery, then highlight it.

Tramontana, G., Jung, M., Schwalm, C.R., Ichii, K., Camps-Valls, G., Ráduly, B., Reichstein, M., Arain, M.A., Cescatti, A., Kiely, G., Merbold, L., Serrano-Ortiz, P., Sickert, S., Wolf, S., Papale, D. (2016) Predicting carbon dioxide and energy fluxes across global FLUXNET sites with regression algorithms. Biogeosciences 13, 4291-4313.

Third, the authors used MPI-BGC product as a training dataset while testing the product against FLUXNET data. As MPI-BGC product was trained against FLUXNET dataset, the approach is self-correlated. Why not evaluating the product against independent datasets from MPI-BGC? E.g. water balance derived ET in basin scale.

Fourth, the spatial domain should be clearly defined. The authors said it is global product, but it did not include Antarctica and Greenland. Given the coarse resolution (100 km), most islands are likely uncovered but the global map (Fig 2) showed fluxes in some islands. How did it happen? Also, how to treat with water fraction for each 1-degree pixel?

Fifth, I recommend showing global uncertainty maps for GPP, LE, H. I think one of strengths in WECANN is its ability to quantify uncertainty. Show the uncertainty map and discuss where and why uncertainties are high. Also quantify uncertainties in global values (e.g. XXX PgC yr-1 +- Y PgC yr-1).

Sixth, test global more carefully. When I look at Fig 2, I found higher ET in mid to south east South America (e.g. cerrado) compared to other global ET products. Also, your ET in this region is relatively very high compared to your GPP map. So, water use efficiency will be very low in this region, which is unlikely. See global distribution of C4 maps. Higher proportion in C4 in this area is likely to lead higher water use efficiency. It is notable that your ANN did not consider C4 information.

Printer-friendly version

Discussion paper

Specific comments: P6: why only 21 FLUXNET sites were used? More than 150 sites data are open to public

P6 L23-24: The authors explained that target data is used for training, validation, and testing. I am confused with the terminology of validation and testing. How do they differ? Also, in L36, "after training, ... was evaluated". Here, does "evaluation" indicate validation or testing? I recommend clearly defining each term, and use them consistently across the whole manuscript.

P6 L30: NN -> ANN (?)

P7 L9: Please define "multiple datasets." Is this training dataset?

P7 L12: What is "this" in "this prior distribution"?

P8 L20: Is this "target estimate" from 3.2?

P8 L22: Add another unit for GPP as PgC yr-1, which could be easily compared to the other studies. Same for LE (km3).

P9 L29: I was surprised to see the reduction of GPP in the Saharan Desert after removing SiF. How to interpret this as we know there must be zero GPP? Also, exclusion of SiF in LE made mixed tendencies in this region. As we are confident LE and GPP are close to nil in this area, it will be interesting to test the impacts of inclusion/exclusion in SiF on LE and GPP here.

P10 L3: All three R2 looks too similar, so it is hard to tell 0.96 is higher than 0.94.

P10: The authors compared WECANN to FLUXNET-MTE, ECMWF, GLEAM and MODIS-GPP which were the training data for WECANN. I feel there should be self-correlation, so I am curious whether this is a reasonable approach.

P10 L8: I know there are few eddy flux tower data in India, so FLUXNET-MTE might involve higher uncertainty. However, this is the same situation for WECANN as it used FLUXNET-MTE and others, which are all uncertain as training dataset.

P10 L4: Be quantitative. Report bias.

P10 L20: Define “G”

P11: Many contents in this page should move to Methods.

P12 L5-6: Then why not removing this site given obvious deficiencies?

P12- : As the authors well recognized, I feel it is odd to compare 1 degree WECANN to several hundred meters in flux towers. All discussion from this comparison seems too subjective. I think “validation” of 0.5-degree product is unlikely possible. As your products are too coarse, I would recommend evaluating at larger scales. For example, look interannual variability of global GPP (PgC yr-1), ET (W m-2), and H (W m-2) and compare to atmospheric inversion estimates. Test whether your product could capture big climate extreme events such as Russian heatwave, Texas drought etc. Compare to other existing global land surface products which were not used as input/training dataset in WECANN.

Interactive comment on Biogeosciences Discuss., doi:10.5194/bg-2016-495, 2016.