

“eddy4R: A community-extensible processing, analysis and modeling framework for eddy-covariance data based on R, Git, Docker and HDF5”

by S. Metzger et al.

We thank Anonymous Referee #3 for the valuable feedback on this manuscript. To ensure the comments result in the intended improvements of the manuscript, we outline below our plans for addressing them in a revised version. The comments by the reviewer are recited in italics, followed by our reply in upright font.

We would value further feedback and specification, and are happy to incorporate additional suggestions.

Reviewer comments

This study present a radically new way to process eddy-covariance data. It combines R-coded EC software that are wrapped in a portable Docker image that can be used on various platforms. It is meant to be scalable and to make use of parallel processing of large quantities of data.

Author intentions for revision

Many thanks for this succinct summary.

Reviewer comments

Major comments

In line with the other reviewers, I think that the paper currently lacks a clear scientific question. I could image that for GMD a clear description of a software environment would suffice, but this paper seems to describe “work in progress”.

Author intentions for revision

As stated by the reviewer, the aim of manuscript is to introduce the novel eddy4R-Docker software framework to address a methodological rather than scientific question: the portable, reproducible and extensible processing of eddy-covariance data. For this reason, the GMD journal was chosen, and three examples of geoscientific applications are provided in favor of a single in-depth scientific survey. One core component of [GMD model description papers](#) is “...evaluation against standard benchmarks...” which is addressed in Sect. 3.3. To

demonstrate completion of the v1.0.0 development stage we intend to include an application example as suggested below by the reviewer.

Reviewer comments

I am a big fan of Docker and directly downloaded the Docker image. I was disappointed in the fact that the image did not contain clear examples (e.g. the three examples outlined in the paper). I could see that the eddy4R.base and eddy4R.qaqc packages were part of the Docker image. I think it is a missed opportunity not to provide examples of (simple) data processing and plotting. Now the advantage of Docker images remains untraceable to the readers and remains rather theoretical.

Author intentions for revision

We could not agree more with the reviewer in that an application example would add much value for the reader and potential user. For this reason, we intend to include an [R-vignette](#) example of a (simple) data read-in, processing and plotting workflow. This example will utilize the functionality of both R-packages presented here, eddy4R.base and eddy4R.qaqc.

Reviewer comments

For instance, the HDF5 section (2.4) is clear but a rather standard description that is available on internet (meta-data, directory structure, self-documenting). Again, this is a missed opportunity to guide users through an example (download raw data, process the data, and HDF5 output and visualization of results). You want to convince the “traditional ASCII” community.

Author intentions for revision

Agreed. The R-vignette example will include HDF5 read-in, write-out examples.

Reviewer comments

Section 2.5 presents the way NEON wants to deploy Docker images. Again, this remains rather high level, while the stated goal is to “empower the Science community at large by putting the key to the scientific algorithms into the hand of scientists”. Again, a clear running example in a Docker container would convince these scientists more than a NEON brochure.

Author intentions for revision

We intend to address this concern through the R-vignette example.

Reviewer comments

Section 2.6 would be an ideal starting point for further “Docker-assisted” data analysis, but unfortunately stops at a reference to the eddy4R wiki pages.

Author intentions for revision

In response to the reviewer suggestion, we intend to introduce the R-vignette example “Docker-assisted” data analysis in Section 2.6.

Reviewer comments

In section 5 there is a reference to the raw data, but again unfortunately no examples are given in which a Docker image automatically reads, processes, and presents results. In the remainder of the paper, three examples are given, which is basically fine, but without a traceable and “hands-on” exercise does not add much. It is (and should be) part of the standard software testing.

Author intentions for revision

We intend to address this concern through the R-vignette example.

Reviewer comments

In summary, I very much like the concept presented in this paper. However, without more in depth possibilities for potential users of the software, the papers seems more suitable for internal documentation than convincing readers that this is a promising way for the community to process eddy covariance data.

Author intentions for revision

We intend to address this concern through the R-vignette example.

Reviewer comments

Minor comments

Page 1: line 34: mention where the NEON site is and also where the aircraft data were collected.

Page 1, line 38: “streaming generation of science-grade EC fluxes”: please explain better what this means.

Page 6, line 185: current recent

Page 6, Figure 3, introduced at line 191. This hardly adds anything. A link would do here. Also figure 4 and figure 7 seem illustrations that do not add much.

Page 7, line 231: CI?

Author intentions for revision

We will certainly work to address all the smaller comments, but we wanted to reply to your bigger concerns before beginning the work to ensure our plan of action is satisfactory.