To Editor Jan Seibert and our generous reviewers and community commenters:

The authors of research article EGUSPHERE-2023-1178 "Leveraging gauge networks and strategic discharge measurements to aid development of continuous streamflow records" thank you for your ideas and corrections. This manuscript has been greatly improved with your help. Please see our responses in blue below, as well as tracked changes in the resubmitted manuscript.

Editor Jan Seibert

Thanks for your detailed responses to the reviewer comments. Based on these responses, I am convinced that you will be able to address all the reviewer comments in a revised version of the manuscript. However, I do not fully agree with one of the reviewers and your response about justifying the threshold value of 0.5 for NSE. The statement made by Moriasi et al. (2015) that simulations above a certain NSE value indicate a good model simulation is problematic as the NSE values which can/should be achieved vary largely depending on hydroclimate and data quality. A longer discussion of this can be found in our commentary (Seibert et al. 2018).

Thank you. We have revised this section of the methods to read, "Note that while such benchmarks are in common use (Moriasi et al. 2015), the efficiency that any model can or should achieve varies substantially with the hydroclimate and watershed characteristics of a given site (Seibert et al. 2018). We provide all data and code for modifying the composite discharge product in accordance with alternative benchmarks as users see fit." [449-453]

RC1

General comments –

This manuscript documents work to use a "donor-gauge" concept to predict continuous daily streamflow discharge time series by developing OLS equations using a limited number of discrete, manual discharge measurements coupled with overlapping timeseries of streamflow observed at one or more nearby donor gauges. Additionally, the manuscript describes results from training multiple Neural Network models to fill in gaps or improve data that are inaccurate training on basin characteristics or basin characteristics and a process-based model, and finetuned on partial streamflow records meeting certain quality criteria.

Overall, this work is important in that it shows the feasibility of creating synthetic time series of streamflow discharge using only discrete manual measurements as long as there is at least one nearby

continuous streamgage. The paper is very well written, and the analyses are thorough and well described. I believe the paper could benefit from some reorganization and clarifying text.

Thank you for your comments.

I suggest that the results and discussion be separated into two separate sections. There is much to discuss from this work. I added comments for specific text that I thought could go into the discussion section.

Our editor shares this preference. The two sections are now separate.

The LSTMs, while thoroughly and expertly developed, are limited by a lack of transferability. This is ok, given that the primary contribution of this paper is to show that continuous streamflow discharge can be predicted at high temporal resolution by developing simple models using a limited number of concurrent measurements. That said, a lot of value could be added by holding out data at the NEON sites and showing the performance of the models on the holdout data. I understand the authors' point that this approach is not always optimal for producing the best predictions, but it is quite useful for understanding the performance of the model at ungauged sites. I don't think this analysis is required for the work to published, but I would suggest at least adding this into a subsection of the discussion that discusses future work or ideas.

Split-sample analogues for a generalist LSTM (Kratzert et al. 2019b) and a process-guided generalist (Read et al. 2019, referenced below) already feature prominently in the literature, so we have highlighted these other works as follows. In both these cases, the goal was robust out-of-sample prediction.

"See Kratzert et al. (2019b) and Read et al. (2019) for split-sample considerations in the context of a generalist and process-guided generalist LSTM, respectively." [320-322]

In my opinion, the paper is a bit misleading in that both the simple (OLS) and complex (LSTMs) models are put in the context of leveraging the donor gauge information to predict at "virtual gage" sites (for example, line 26 of the abstract). However, the LSTMs are really only being used to essentially correct the NEON time series and fill gaps. I don't think this is a big deal, as the work and results are still important, but I do think that there are some points in the manuscript that could be written more clearly. I address these in the specific comments below.

We'd like to clarify that the linear regression and LSTM methods accomplish the same goal of correcting and gap-filling the NEON time series. The two approaches supplement each other.

Another minor point is that OLS is only used at five sites for which a single donor gauge is available. Ridge regression is used at 19 sites, and segmented regression at one site. "Linear regression" is the term that we use in the paper to encompass these methods.

We address your concern about conflating the two approaches, including on line 26, in responses to the comments below.

Specific comments –

Line 26-27. If by "simple and complex" you mean OLS and LSTM-RNN, I don't think this statement is entirely true. If I understood the paper correctly, weren't the discrete manual measurements only used in the development of the OLS models? It was my understanding that the NNs were trained on the continuous timeseries data published as part of the NEON data that met quality assurance criteria.

The discrete discharge measurements were leveraged differently by the two broad model categories. For the linear regressions, they were used for fitting. For the LSTMs, they were used for evaluation-- After training LSTMs on diverse, continuous discharge data, including at the target site(s), we needed the manually measured discharge data to determine whether the models were any good.

We have removed this potential confusion from the abstract by simply stating, "Here, we investigate the potential to estimate continuous discharge when discrete streamflow measurements are available at the site of interest" [51-52].

Line 89. I'm having a hard time following the sentence that begins with "Where co-location…". I assume the author is referring to co-location of the study (or site of interest), but because the previous paragraph is discussing gauge placement, I could also see this referring to co-location of low-cost monitoring equipment. Depending on which meaning the authors intend, I think the implications and direction of this paragraph are changed. I suggest adding a bit of text to clarify. Possible rewrite:

"Where co-location of the site of interest with an existing stream gauge…" or "Where co-location of low-cost, rapid-deployment monitoring gauges with an existing stream gauge…"

The first meaning is what we intended, and we have gladly borrowed your clarifying text. Thanks!

Line 110. "wide range of flows" is subject to interpretation. I think it would be helpful if the authors provided more information on what is meant by "wide range of flows". I think this information would also provide valuable insight into how transferrable this method is to other locations. If laid out thoroughly, I think this could allow readers to have a greater understanding of how well sampled a

location needs to be to use the methods described in this manuscript to reconstruct or predict streamflows. For example, how important is it that the flows used to develop the equations or train the models captures the annual extremes? How does predictive performance change as a function of the relative distribution of flows represented by the training data? At the very least, I think some text is required to discuss the implications of the representativeness of the training data.

Lines 626-634 of the discussion address this more thoroughly, but we now provide more detail in the introduction by saying "In this study, we use the term [virtual gauge] to describe sites at which discrete discharge observations can be used to fit or evaluate models that generate continuous flow. For accurate results, field measurement campaigns should prioritize characterizing the distribution of possible flow conditions, rather than achieving any particular threshold number of observations." [160-164]

Line 150. Figure 1. The resolution of this figure is poor. I suggest replacing this figure with a higher resolution figure. I also think that this figure would benefit greatly from having the NEON sites labeled to give the reader some geographical context.

This figure resolution has been improved, and NEON sites are now labeled.

Lines 164-173. Suggest moving this text up into a new section titled "2.3 Donor-gauge Selection" as it is a crucial component of the analysis and risks being overlooked by including it in the section on model selection.

Thank you for this suggestion. Those lines have been moved to a new "Donor gauge selection" section.

Lines 169-170.

Are the three criteria provided in parentheses the only criteria used to determine donor-gauge candidacy? If so, suggest removing the "e.g.", as you are not providing examples of criteria, but the exhaustive list of criteria. If so, I suggest listing all the criteria used and not just providing examples.

For the criteria used in determining geographic similarity: what data and methods were used to determine whether a site was in an urban area or downstream from a reservoir?

There are other conceivable disqualifiers, such as tidal influence (not applicable in our case) or water withdrawal (harder to identify), but these were the three criteria we used, so that "e.g." now serves better as an "i.e." We have made clearer that visual inspection of interactive maps was the only method used.

Additionally, see the technical comment for lines 167-170 provided below.

Line 268. Authors state that the models were "trained, validated, and tested", but the methods only describe training and validation. Were the LSTMs test on holdout data? If not, suggest removing "tested".

Thanks. This wording was a holdover from a previous draft and has been amended.

Lines 362-365. Great that the authors noted the limitations of the evaluation metrics due to potentially limited representation of flow in the evaluation data. If possible, providing more information (beyond the statement on disproportionate representation of low flows) would be very helpful to the reader in understanding where there might be more or less uncertainty in the model predictions.

We now clarify that uncertainty is generally higher for high-discharge estimates.

Line 374. Figure 2. Do the bars show the mean KGE of the 27 sites? The figure caption does not state what the value represented by the bars is. Please include that in the caption.

Some of these bars represent means, which we do mention: "...displayed KGE is averaged over 30 ensemble runs...." The rest are simply KGE computed on all available observation-estimate pairs. We have appended ", computed on all available observation-estimate pairs" to the first sentence of the caption.

Line 385. Figure 3. This is a great figure! A great presentation of a lot of info. Adding an x-axis showing time, even if the interval is in years, would add a lot of value. Additionally, if the stations could be ordered in the same manner as they are on the x-axis of Figure 2, it would help in the assessment of performance at individual sites and add to the flow of the paper.

Thanks! We appreciate these suggestions. A re-ordering of sites (and explanation of said ordering in the caption) will be useful. As for including the x-axis, we'd like to identify a tradeoff: observations begin anywhere from March 2016 to Oct 2018, depending on the site. That means there would be a lot of white space on the left side of this figure, and that the already dense detailing would be squished into as little as 3/5 of the available horizontal space. You've identified a clear value-add (visualizing when sampling begins and when gaps occur), but one that goes beyond the intent of the figure, which is to illustrate the relative contribution of NEON/reconstruction estimates to the full picture, and to allow

potential data users to visually assess (in)congruity between the black and gray lines. So, we have provided the plot you suggest, with x-axis included, as an appendix figure.

Lines 419-426. This text could go into a discussion section.

Accepted

Lines 436-441. This text could go into a discussion section.

Accepted

Lines 455-474. This text could go into a discussion section.

Accepted

Line 515. Figure A1. Do the bars show the mean KGE of the 27 sites? The figure caption does not state what the value represented by the bars is. Please include that in the caption.

See our solution to the comment on line 374.

Technical comments –

Line 99. "site location selection" is redundant. Suggest changing to "site-selection issues".

Accepted

Line 100. Suggest changing "high-quality" to "higher quality".

The quality of data gaps is indeterminate, so the comparative wouldn't work here (assuming this refers to line 104, rather than 100).

Line 105. Suggest removing the comma after "series,"

Accepted

Line 108. Change "virtual gauges." to "virtual gauges" by removing the period.

Good catch!

Lines 167-170. For conciseness and clarity, suggest the following rewrite: "For each target site, up to four donor gauge candidates were selected based on spatial proximity (within 50km) and geographic similarity (not in an urban area; not downstream of a reservoir). Generally, no more than four gauges met these criteria, but for one target site (MCRA) there were 10 nearby candidate gauges to select from, all of which were associated with the Andrews Experimental Forest in western Oregon, USA."

This rewrite does make the text more concise, though geographic similarity encompasses terrain and landcover broadly, and is a continuum, whereas artificial flowpaths and control structures grossly alter flow relationships, and do so in idiosyncratic ways. We have provided the following clarification:

"Barring gauges on reaches that are subject to overt human influence, the exact methods used to choose donor gauges are of little consequence, so long as informative donor gauges are not overlooked. In practice, there will usually be just a few, if any, potential donor gauges available for a given location. If multiple donor gauges are included in a regression, L2 regularization (ridge regression) should be used to account for their covariance (see Sect. 2.4)" [227-231]

Lines 172-173. Suggested rewrite: "…chose three candidate sites representing A) a catchment upstream of the target site (GSWS08), B) downstream of the target site on the MCRA mainstem (GSLOOK), and C) downstream on a tributary of MCRA (GSWS01)."

This becomes a bit of a parenthesis soup, especially confusing when some of them are unpaired.

Line 198. Change "glmnet" to "Ridge-regression".

Accepted

Line 455. Change "low flow conditions" to "low-flow conditions"

Accepted

RC2

This paper describes an effort to reconstruct and correct streamflow data series from NEON (National Ecological Observatory Network) stream gaging sites, based on limited field measurements at the NEON sites and information from nearby stream gages.

For anyone who has attempted to use NEON streamflow data, this contribution is a gift. The methods are clearly described, code is provided, and the resulting datasets are made available. Thanks to the authors for their efforts.

Thank you!

Comments:

1. Abstract: please describe how the linear models performed compared to the LSTM models, and why.

We now state "...with linear regression generally outperforming deep learning approaches due to the use of target site data for model fitting, rather than evaluation only." [56-58]

2. A number of datasets and models are used in the analysis. It would be helpful to provide a table of these and their hyperlinks with a description of each, in the Appendix. For example, CAMELS, NHM, etc. And in addition, can you provide links to the "donor" datasets that you used? Perhaps this is provided and I missed it.

We have included the following table in the appendix. Note that links for NEON data are instead given as citations, to ensure that users find their way to the complete link, doi, and timestamp reference corresponding to the datasets we used. Instead of links to the donor gauge datasets, we provide a link to the landing page of the National Water Information System, and an example of a page for a single donor gauge. There are 54 such gauges involved in this study, and we do not recommend that anyone download these individually, but instead use the cited dataRetrieval package. We also inserted the sentence "NWIS gauge ID numbers are provided in cfg/donor_gauges.yml at the GitHub and Zenodo links below." on lines 196-7.

Table A2. Model input data used in this study.

Resource	Description	Source/Link
NEON discharge field collection	Discharge measurements from field-based surveys	NEON 2023b, NEON 2023c
NEON continuous discharge	Discharge calculated from a rating curve and sensor measurements of water level	NEON 2023a
User-focused evaluation of NEON streamflow estimates	3-tier classification of the reliability of NEON continuous discharge by site-month	https://www.nature.com/articles/s41597-023-01983-w

CAMELS dataset	Catchment Attributes, Meteorology, (and streamflow) for Large-sample Studies	https://ral.ucar.edu/solutions/products/camels
National Hydrologic Model (NHM)	USGS infrastructure that, when coupled with the Precipitation-Runoff Modeling System, can produce streamflow simulations at local to national scale	https://www.usgs.gov/mission-areas/water-resources/science/national-hydrologic-model-infrastructure
MacroSheds	A synthesis of long-term biogeochemical, hydroclimatic, and geospatial data from small watershed ecosystem studies	https://portal.edirepository.org/nis/mapbrowse?scope=edi&identifier=1262
Daymet	Gridded estimates of daily weather parameters	https://developers.google.com/earth-engine/datasets/catalog/NASA_ORNL_DAYMET_V4
HJ Andrews Experimental Forest stream discharge	Stream discharge in gaged watersheds, 1949 to present	https://portal.edirepository.org/nis/mapbrowse?packageid=knb-lter-and.4341.33
USGS National Water Information System	Streamflow and associated data for thousands of gauged streams and rivers within the USA	https://waterdata.usgs.gov/nwis, e.g. https://waterdata.usgs.gov/monitoring-location/06879100/

3. Results/Discussion. A number of products including code, model results, and datasets are produced. It would be helpful to provide a table of the results created by the analysis and their hyperlinks with a description of each, in the text.

The following table has been added to the appendix.

Table A3. Products of this study.

Product	Description	Link
Data archive landing page	Figshare page linking to each of four archives described below	https://doi.org/10.6084/m9.figshare.c.6488065
Composite discharge timeseries	Analysis-ready CSVs combining the best available discharge estimates across linear regression and LSTM approaches from this study, and NEON's published data	https://doi.org/10.6084/m9.figshare.23206592.v1
Composite discharge plots	Interactive plots of our composite discharge product	https://macrosheds.org/data/vlah_etal_2023_composites
All model outputs and results	Complete predictions from all linear regression and LSTM models, run results, and diagnostics	https://doi.org/10.6084/m9.figshare.22344589.v1
All model input data	Donor gauge streamflow, training data for LSTMs, model configurations, etc.	https://doi.org/10.6084/m9.figshare.22349377.v1
All code associated with this paper	Zenodo archive of GitHub repository	https://doi.org/10.5281/zenodo.7976251
All figures associated with this paper	High-resolution images of all figures from the main body and appendix	https://doi.org/10.6084/m9.figshare.23169362.v1

4. Results/Discussion. I could not find metadata in the materials provided at the hyperlinks. Please provide measurement units (are these unit area flows in mm?) and other metadata for the composite series linked at
https://figshare.com/articles/dataset/Composite_discharge_series_for_each_NEON_stream_river_site_/23206592

5. Results/discussion (l. 390-400) and conclusions: (l. 495-500). Please provide a slightly expanded explanation for why the LSTM models performed considerably less well than the linear models. It appears that it was not the model structures themselves but their data requirements that limited the performance of the LSTM models relative to the regression models. Can you help readers better understand these limitations more generally? Were watershed characteristics or other data shown in Table 2 relevant to the performance of the linear vs. LSTEM models? How did the differences in drainage areas of the CAMELS data set relative to these watersheds (l. 25-251) influence model performance?

The crux of the difference in performance is given on lines 563-568. We now elaborate with

"Furthermore, whereas the LSTM models must parameterize each day of prediction individually, the regression models need only parameterize relationships between flow regimes. Still, if given enough training data, including examples of watersheds and streams similar to each of those modeled in this study, the LSTM approaches would eventually close the performance gap." [568-71]

6. Results/discussion. For each site, could you summarize the record lengths over which you were able to reconstruct the 5-minute data vs. developing a daily estimate? And if the daily estimates were based on the LSTM models, which have a median NSE of 0.32 to 0.47, why are the confidence intervals for daily discharge in the resulting composite data series so narrow?

Figure 3 provides just such a summary. The hashed pink and purple sections have only daily reconstruction estimates.

As for narrow confidence intervals on inaccurate estimates, this implies that the corresponding models were confidently incorrect! Confidence intervals are quite variable for the LSTM models, and are often similar in relative magnitude to those of the linear regressions. For the LSTMs, these intervals were generated by building ensembles of 30 models and computing the 95% quantiles of each point estimate. The optimization of a neural

network is a stochastic process, so a different solution is reached for every iteration. Optimization problems with a diversity of solutions will naturally produce a wide range of answers (resulting in wide confidence intervals), whereas other problems might be more constrained. Here, for example, inadequate training data for a particular type of stream might result in a model that consistently predicts off the mark for that stream type, but does so with a consistent bias (low accuracy, high precision, in a sense). Users of the composite discharge product must therefore be mindful of both uncertainty and accuracy when evaluating our estimates for a particular site.

7. Conclusions. Do you have any recommendations for NEON to improve their stream gaging procedures, data collection, and information management processes? Will it be necessary for users of NEON data to use your code in the future?

We have communicated with NEON in regards to this and similar efforts. As lines 47-48 and 100-102 indicate, measuring discharge, especially in new systems, is hard, and NEON products are showing clear improvement over time, especially as their gauge placements and rating curves mature. Primarily we think a thoughtfully-engaged user community can help improve the use and reuse of NEON data. Ideally, it will not be necessary for NEON data users to rerun this code on future data releases. Still, we've taken great pains to make it reusable with minimal effort.

CC1

The manuscript "Leveraging gauge networks and strategic discharge measurements to aid development of continuous streamflow records" uses data from the National Ecological Observatory Network (NEON) to develop models for predicting continuous streamflow from discrete measurements. The authors also provide a dataset that combines published NEON data with gap-filled and corrected data derived from their models. Both the manuscript and dataset will be extremely useful for NEON data users and their process may be useful for other researchers attempting to build continuous streamflow records at other sites. We have some substantive and editorial comments outlined below that will strengthen the manuscript and increase the reproducibility of their analyses.

Thank you so much for taking the time to provide this feedback!

Substantive comments:

Abstract:

The authors do an excellent job of summarizing the challenges to developing continuous streamflow records and highlighting the importance of overcoming these challenges for the research community. The abstract is a bit narrowly focused on the implications of their analyses for the NEON dataset when they may be useful for a broader group of researchers.

Good point. We zoom out at the end of the intro and discussion, but not here in the abstract. We have added, "The success of this effort demonstrates the potential to establish "virtual gauges," or sites at which continuous streamflow can be accurately estimated from discrete measurements, by transferring information from nearby donor gauges and/or large collections of training data." [61-64]

Methods Section:

Section 2.1: Based on the text in the Methods section, it appears that the authors applied the tier classifications from the outputs produced by the analysis in Rhea et al. (2023). That manuscript used a combination of RELEASE-2022 and Provisional NEON data available at that time. For best alignment, we suggest that the authors apply the classification framework from Rhea et al. (2023) to the RELEASE-2023 and provisional NEON data used in this manuscript. These new outputs should be included in the archived dataset associated with this manuscript for completeness. If this is already done, the authors should update the Methods section with additional detail to make this clear. Also, any Provisional data used in the analysis should be archived, if possible, by the authors to maintain reproducibility.

We used the updated version of Rhea 2023 (the dataset accompanying Rhea et al. 2023), published 4/11/23. This update repeats the original analysis using RELEASE-2023 data. We did not use provisional continuous discharge data, but did use provisional field Q, which is included in the archived dataset associated with this paper. All of this has been made clearer in our methods.

"Officially released (NEON 2023c) and provisional (NEON 2023b) field measurements were used..." [172-3]

"Continuous discharge data (NEON 2023a) were also retrieved via neonUtilities. We used RELEASE-2023 and *not* provisional data in this case" [184-5]

"Evaluation results used to distinguish likely reliable vs. potentially unreliable subsets of NEON's RELEASE-2023 continuous discharge time series, per site-month, were provided by Rhea et al. (2023a) and accessed through HydroShare (Rhea 2023)." [187-9]

Section 2.5: The authors should provide citations and/or a rationale for the NSE and KGE benchmarks that they apply to the various datasets since those are different for each model. Some potential examples include Knoben et al. (2019), Moriasi et al. (2015), Golmohammadi et al. (2014), and Chiew and McMahon (1993).

Thank you for these citations. We have cited the 0.5 NSE threshold recommended by Moriasi et al., but note Editor Jan Seibert's observation that this benchmark is problematic as, "the NSE values which can/should be achieved vary largely depending on hydroclimate and data quality."

Results and Discussion Section:

Line 455: While practical considerations such as site access, timing, and safety do indeed limit the collection of higher-flow discharge measurements, low-to-baseflow conditions are typically present throughout the majority of annual water year, particularly in smaller stream systems. Stream monitoring networks that include small streams with unstable low-flow hydrologic controls must balance the accurate characterization of the low to mid-flow regime with measuring discharge during high-flow events, which are relatively infrequent. Streams with unstable low-flow hydraulic controls (e.g., those with unconsolidated bed material) exhibit highly variable stage-discharge relationships, both within a given water year and across longer timescales. As such, multiple gaugings are required each year in order to accurately characterize the stage-discharge relationship in the low-to-middle segments of the rating curve. In contrast, high-flow gaugings can often be used in rating curves over longer periods of time due to the relative stability of high-flow controls (e.g., banks, valley walls, etc.).

This is a crucial point that we have overlooked. We have added, "This often occurs for practical reasons such as site access and technician safety, but may also reflect a need to characterize the low-flow variability of the stage-discharge relationship in streams with unstable low-flow hydrologic controls, such as unconsolidated bed material." [612-15]

The Discussion section focused a lot on NEON, without much discussion of how the methods could be applied more generally. This was an important point brought up in the Introduction, but it was not extensively revisited in the Discussion. Limiting most of how the results apply specifically to NEON to the Results section, and discussing how the methodology could be applied to un-gaged watersheds or watersheds with incomplete records more generally, would likely increase the amount of interested readers beyond just those who are using NEON data.

We have separated the Results and Discussion, and taken the opportunity to expand on how this methodology can be applied in other settings.

Editorial comments (primarily about citing NEON data to improve reproducibility):

Throughout the text of the paper: it is not clear from the citations whether NEON provisional data (see definition at https://www.neonscience.org/data-samples/data-management/data-revisions-releases) were used in assessment of the linear regression and LSTM models (e.g. lines 116-117). Please clarify in the text and citations in terms of which data were used in which part of the analysis.

We have clarified the text by explicitly stating that provisional data were *not* used for the continuous discharge data product. The citations have been clarified by addressing your comment on line 117 below.

Line 106: The readme for the macrosheds repository linked here reports an incorrect data product ID (reports DP4.00230.001, should be DP4.00130.001).

Good catch. We've revised the readme.

Line 117: Separate citations are needed for released and provisional datasets of the Discharge field collection data product (DP1.20048.001). These are two separate datasets. See https://www.neonscience.org/data-samples/guidelines-policies/citing for NEON citation guidelines.

Thank you for pointing this out. We now include the provisional data citation for field Q in section 2.1. In-line citations and references have been corrected as per the following comments.

Lines 118, 126: The date the data were accessed should be in the references entry, not the inline citation, unless specifically requested by the journal. Regardless, the accessed date should be included in the references on lines 671 and 669, respectively. See https://www.neonscience.org/data-samples/guidelines-policies/citing for NEON citation guidelines.

This has been corrected.

Lines 669-671: NEON data citations need to be updated to match the NEON citation guidelines detailed in https://www.neonscience.org/data-samples/guidelines-policies/citing.

These references now match the official guidelines.

Line 326: Please change '2023-release' to 'RELEASE-2023' to be consistent with NEON terminology.

Accepted.

Line 509: Please change 'DP4.00230.001' to 'DP4.00130.001' as the incorrect data product ID is currently being used.

Egad! This mistake has been corrected.

Line 579: In the acknowledgements section, please include the standard NEON data acknowledgement (see https://www.neonscience.org/data-samples/guidelines-policies/citing).

The standard acknowledgement is now included.