NEON STEAC Fall 2021 Advisory Report

The members of the STEAC met on October 5 and 7, 2021 to discuss the status of NEON and new innovations on the horizon. The following members of the STEAC attended the meeting: Lilian Alessa, Emily Bernhardt, Sarah Bevins, Frank Davis (10/5), Mike Dietze, Anne Giblin, Peter Groffman, Karen Lips, Sparkle Malone, Jackie Matthes, Kim Novick, Steve Petruzza, and Shawn Serbin. The meeting also included 10 NEON-Battelle staff (Darcy Gora, Zoe Gentes, Christine Laney, Michael SanClements, Kate Thibault, Chris McKay, Gene Kelly, and Paula Mabee), Roland Roberts from the NSF, and Nico Franz from ASU.

The meeting was virtual, and the following topics were discussed: I. Sample use policy, II. Science staff hiring plan, III. NEON Postdoc program, IV. Ambassador program & ESA, V. Engagement, VI. Research Data Alliance, VII. STEAC Conversation with NSF, and VIII. STEAC business.

I. Sample use policy

The NEON Biorepository is a unique and important component of NEON. To date (three years) there have been 81 requests for samples and 25 loans have been made, with estimates that the facility could fulfill a loan action once every 1-2 weeks. The Biorepository is working to develop a formal sample use policy, but one challenge is the large number (70) and diversity of different types of samples (a wide range of biodiversity collections, biogeochemical samples, genomic material, etc.). Nico Franz and his group are searching for/developing protocols, working deliberately with Battelle, and the NEON Technical Working Groups (TWGs) when necessary. This appears to be a sound approach. Key to this is that NEON Biorepository already possesses an informed and adapted repository staff able to engage with the community in a positive and productive way to develop protocols and respond to requests.

A major point of discussion is how requests should be prioritized and especially if commercial uses should be allowed/prioritized. It will be important for NEON to consult with NSF for guidance. Given that, the draft policy appears to be sound when it states that “when there are conflicting sample requests, the priority for loans will generally be as follows: 1) NSF-BIO sponsored research, 2) other NSF sponsored research, 3) other, non-NSF research funding sources, 4) educational applications, 5) for-profit research companies.” This priority of for-profit requests could be increased if significant public benefit can be shown, especially around data

sharing. The members of the STEAC stand ready to help reconcile conflicts if NEON thinks that would be useful.

II. Science Staff Hiring Plan

There are exciting employment opportunities at NEON: Microbial Bioinformatician (Data Scientist III), Biorepository and Informatics Scientist (Data Scientist II), Instrument Systems Data Scientist (II), Remote Sensing Scientist (Environmental Scientist II), and a Remote Sensing Data Scientist (Data Scientist II). The members of the STEAC commend NEON for being proactive and planning to ensure a diverse candidate pool.

Members of the STEAC encourage NEON to recognize that their positions require a unique and broad combination of skills, including research, community engagement, algorithm development, instrumentation expertise, and ensuring the integrity of data pipelines. There were concerns that in some cases it may be difficult to find single individuals who possess such a wide mix of skills. The STEAC urges NEON to consider contractors to provide expertise that is not present in the candidate pool. NEON could also leverage the intellectual strengths of the community for research needs in place of requiring this expertise in an otherwise technical position. To potentially increase the diversity in the candidate pool, NEON can consider hosting a virtual career fair, zoom information session, or sending position announcements to diversity programs (ex. ESA SEEDS, SACNAS, AmeriFlux DEI (daagarwal@lbl.gov), AGU DIAC, CREST), NSF INCLUDES National Networks, and coordination hubs (ex. EDSIN)). Targeting diverse and underrepresented groups through social media might also help to further diversify the candidate pool (@UConnSEEDS @earth_jobs @earthinstitute @SACNAS, @LatinoOutdoors, @500womensci, @BlackAFinSTEM, @BLACKandSTEM @NativesInSTEM @ESA_SEEDS @GeoLatinas @LatinXinBME @GeoSpaceLatinx @LatinasinSTEM @sisterSTEM @Reclaiming_STEM @VanguardSTEM @BEcologists @HarvardForest).

III. Postdoc Program

The members of the STEAC were excited to hear about the work being done by NEON Postdocs Dr.Kelly Aho and Dr.Kelly Hondula, who are both using NEON data to address novel research questions about scaling up critical aquatic biogeochemical processes. The committee also congratulates the third NEON postdoc (Dr. Alesia Hallmark) for their success in obtaining a new professional position.

The NEON postdoctoral fellows expressed enthusiasm about the level of independence and flexibility that came with their position. Though currently limited by COVID-19, opportunities for better connection with external research communities (e.g., through their external mentor's institution) were highlighted as desirable in the future. Growing interpersonal networks of collaborators and peers is critical for early career scientists, and ensuring that these opportunities exist for the current and future NEON postdocs is critical. The experience of the first cohort of postdocs was strongly influenced by the COVID-19 pandemic and associated limitations on in-person gatherings and travel. In the likely event that these restrictions continue into the near future, NEON staff and the NEON postdoctoral fellows could discuss additional opportunities for discipline-relevant networking, perhaps by linking up with related research groups located in the many academic and governmental institutions co-located in Boulder.

Both NEON postdocs found that much of their first year was dedicated to data organization, quality control, and synthesis of multiple NEON data streams into usable data products. The members of the STEAC thank the postdocs for doing this much-needed work, and appreciate their efforts to return derived products back to the community (e.g., through deposition into EDI). Looking forward, transforming raw NEON data streams into usable formats will continue to be a pressing need to ensure success of the NEON project. This issue is addressed in some detail later in the report (see Section VII).

The members of the STEAC recognize that postdoctoral scientists could play an important role in this process. However, care must be taken to ensure their contributions are properly rewarded, for example with opportunities to collaborate with future end-users of the derived data products, and through proper citation of data publications. NEON staff may also be in a unique position to identify instances of duplication of effort among early career scientists (e.g. multiple early career scientists developing the same derived data product), and may want to think about strategies for ensuring collaboration, and minimizing competitions, among research projects. NEON can play an important role in supporting these goals by engaging with discipline-specific networks and community platforms where early career scientists can share their work on social media platforms (slack, twitter, instagram, facebook). Previous discussions have acknowledged the constraints on NEON running platforms. Beyond that, the STEAC recognizes that it is difficult to fully assess the postdoc program for its successes, and opportunities for improvement, until COVID-19 pandemic restrictions have been eased. The members of the STEAC also note that the NSF Center for Advancement and Synthesis of Open Environmental Data and Sciences may be part of conversations on derived data products and postdocs in the

future. For this reason, the STEAC recommends that no major changes be made to the structure of the program at this time.

Finally, NEON is currently providing traditional research oriented postdoctoral experiences. NEON might consider offering Postdoctoral fellowships for those interested in research infrastructure management, instrumentation, cyber infrastructure and other closely aligned fields. These experiences might support long-term workforce needs and foster collaborations with other communities.

IV. Ambassadors & ESA

The call for interest for the ambassador positions received 51 impressive applications of which 20 were invited. The applicant pool had a good mix of mid and early-career researchers, as well as diversity in gender, race, and expertise. The process identified applicants' potential activities, outcomes, barriers, and contributions aligned with the interests and goals of NEON. The large numbers of applicants was encouraging, and indicates community support for NEON activities, and opportunities for greater engagement. The members of the STEAC advise NEON to be prepared to invest resources in supporting the Ambassadors and their work, such as providing guidance for effective engagement and access to networks supporting similar efforts. NEON could co-create a role for the Ambassador that can support their skill set.

The Ecological Society of America is one of the many societies NEON interacts with and provides short courses, workshops and support for SEEDs and faculty development. Ambassadors could leverage this society and their programs by becoming mentors in some of the data and critical skills workshops. Overall, the priority should remain building communities and research applications.

V. Engagement

NEON is engaging in impressive new collaborations with AGU's Thriving Earth Exchange (TEX) and NCAR, and working with Dimensions to design a community-accessible page that tracks publications and other products derived from NEON data. The new collaboration with NCAR is leading to exciting data products (e.g., gap-filled meteorology) and the STEAC is impressed by the experimental data products portal. The members of the STEAC encourage NEON to focus on collaborations with other networks and agencies that have expertise in providing products

similar to those of NEON and can use NEON data for derived data products. These collaborations are likely to lead to the derived data products the community needs. A focus on community engagement that also serves data needs is a great strategy and would allow NEON to focus on providing high quality data products. For example, there is a strong opportunity for collaborations with the USGS to further develop the aquatic data sets.

The members of the STEAC applaud initiatives within NEON to leverage external expertise together with in-house efforts, specifically the NEON-AmeriFlux and the new NEON-NCAR partnerships. Specifically, the STEAC members remain strong in their support of the eddy covariance data pipeline for NEON data that now leverages the AmeriFlux data processing pipeline such that these data products will be consistent with the AmeriFlux data products. It was felt this would more easily enable larger synthesis efforts. Similarly, the partnership to leverage NCAR meteorological data gap filling was also seen as a major advancement, as the STEAC mentioned a high proportion of users would be looking for gap-filled versions of this data. Along these lines, the STEAC also suggested that other similar partnerships be pursued or enhanced. For example, it was mentioned that NASA JPL has been advancing the tools for atmospheric correction of airborne remote sensing data, and that the NEON AOP, should reach out to JPL (if not already) to leverage these open-source tools to build in more consistency in datasets to simplify the use of NEON and other similar datasets (e.g. NASA AVIRIS) together. Overall, the STEAC suggested that NEON continue to find ways to reach out and leverage existing and community-supported data pipelines. Data usability is going to be critical to the success of NEON.

The members of the STEAC also applaud NEON for working with Dimensions to track the publications and grants that utilize NEON data. This is essential to measure the impact of NEON and identify which communities are utilizing NEON data, and which communities might require more engagement.

VI. Research Data Alliance

The members of the STEAC were glad to hear (from Christine Laney) that NEON has re-engaged with the Research Data Alliance (RDA) as this is an important community-led effort to address cutting-edge issues in cyberinfrastructure. There is high potential for NEON to both contribute to and benefit from this group. However, participation by NEON with the large number of RDA interest groups and working groups is limited by time and resources. We

endorse the approach where NEON personnel participate in particularly useful and relevant efforts as needs emerge and time allows, e.g., NEON is benefiting from Laney's participation in the Data Granularity Working Group. There may be particularly useful synergies with other networks, e.g., LTER, Ameriflux, through RDA.

VII. STEAC Conversation with NSF

STEAC members discussed success stories and continued challenges related to NEON data product development, usage, as well as ongoing user challenges with Dr. Roland Roberts, the program director for NEON. One example of remaining challenges for users focused around the communication of existing "known issues" surrounding NEON data quality and historic data updates. In previous NEON updates to STEAC, it was indicated that a majority of NEON data access occurs through the API, but often the reporting of proper data usage, data quality and known issues are not easily found by end users, e.g., buried within README files. The members of the STEAC recommend that both "known data issues" and historical data updates or "change logs" are more clearly communicated to end-users, particularly those leveraging the NEON API, directly or through external tools. Specifically, these "known issue" logs and "change logs" could be pulled out to a higher level and provided as a separate text file, in addition to being found within the larger data README.

Similarly, the STEAC discussed ways to increase data usage, particularly by providing more "analysis-ready" datasets. The members of the STEAC conveyed that in the long run, the useability of NEON data will be critical to the overall success of the Observatory and at the same time, creating more processed data will minimize the duplication of external efforts where different groups essentially develop the same products from NEON data.

The members of the STEAC also noted cases where multiple NSF macrosystems grants were working to get the same data products derived and functional. The duplication of effort is not efficient, but if the different groups collaborate, it raises questions on how credit is then divided between multiple early career researchers. The members of the STEAC note that continued conversations with NSF about accelerating or optimizing data products that the community has a need for, is important. The members of the STEAC also recognize that the new Dimensions webpage will show grants, which may help create awareness around data efforts in the community and reduce duplications.

VIII. STEAC business

The members of the STEAC approved term limits for STEAC members (3 years with the option to stay on one additional year) and made changes to the bylaws with a unanimous vote. The Members of the STEAC also thanked Michael Dietze for his service as chair of the STEAC and welcomed Sparkle Malone as the new Chair. Kim Novick will continue to serve as Co-Chair and the Secretary position will be filled at the next STEAC meeting. The members of the STEAC are brainstorming a range of ideas of where we need expertise on the STEAC and will develop a plan over the coming months. The members of the STEAC from institutions that are competing for the Competition for the Management of Operations and Maintenance of the National Ecological Observatory Network (NEON) (nsf21603) recused themselves until the end of the competition.