Science, Technology and Education Advisory Committee (STEAC) for the National Ecological Observatory Network (NEON)

September 2018 Advisory Report

I. Overview

The STEAC met in Boulder, CO on Sept 26-27, 2018 with 15 of 20 STEAC members present, 12 in person and three via WebEx. The committee is grateful for the presentations that NEON staff prepared and for the time they took to interact with us. STEAC had a wide-ranging and productive discussion during our two-day meeting, the details of which are provided in the meeting minutes. Here we highlight the many positive NEON developments over the past six months, along with our observations and recommendations in five specific areas: Organizational Structure and Decision-making (Section II); Cyberinfrastructure (Section III); NEON Staff Professional Development (Section IV); Data Availability and Training (Section V); and Field Staff Training (Section VI). NEON is poised to make a significant contribution to the ecological/environmental sciences, including addressing several of NSF's 10 Big Ideas (e.g., Harnessing the Data Revolution, Understanding the Rules of Life, Growing Convergence Research, and INCLUDES--Enhancing STEM through Diversity and Inclusion). The STEAC is looking forward to providing continued advice as the NEON Project transitions into full operations and becomes an open data organization.

Significant NEON developments over the past six months:
- The completion of construction is a major accomplishment that the organization and ecological community should take a moment to celebrate. We acknowledge the tremendous effort by Battelle to organize, manage and complete the construction of NEON infrastructure.
- NEON staff have been highly responsive to our feedback. We appreciate the detailed response to our April 2018 meeting report, and we are pleased to see the efforts that NEON science and education staff have made to address previous STEAC recommendations. Notable examples include: continued emphasis on engagement; focusing on consistent branding; addressing employee and temporary staff retention and career development issues; and tackling concerns about data quality and availability.
- The engagement plan has evolved into a thorough and useful document that provides strategic direction for the NEON project.
- NEON staff have released an R package for eddy covariance processing that will help establish best practices in land-surface research across the user community.
- NEON staff have led or participated in a number of successful research, infrastructure, and meeting proposals that better position the organization to carry out its mission and to connect effectively with the user community (e.g., NSF INCLUDES and RAPID awards; NSF grant for cyberinfrastructure).
- STEAC enjoyed hearing about the strong representation of NEON activities at ESA. There is growing enthusiasm about NEON in the user community, which is poised to increasingly utilize NEON data products going forward. Travel awards and networking activities for early career scientists at ESA are the types of initiatives NEON should continue to offer to ensure wide community buy-in. NEON staff are also positively affected by these advances in community use of NEON data and participation in NEON-sponsored activities.
- We were pleased to see fifty nominations to our call for new STEAC members -- an illustration that many in the user community are keen to be involved. Four of the five new

members, selected from these nominations, were able to participate in our September meeting, despite learning of their appointment only shortly before the meeting.

II. Organizational Structure and Decision-making

Overview. As NEON moves from Construction to Operations, major changes in organization and decision-making will be necessary. A new Chief Scientist/Observatory Director is now in place, and the "client" for the NEON enterprise has shifted from NSF to the ecological/environmental science user community. This is a non-traditional client, with complex interactions with NSF. The current organizational chart that was provided to us (Figure 1) appropriately shows the Observatory Director as the organization lead. This is essential, since all decisions going forward must be driven by the science. ***However, it is unclear to STEAC that the Observatory Director actually has authority over financial and personnel decisions that will be required to successfully carry out the scientific mission of NEON.*** We note that the PMO, Domain Support and other boxes do not report directly to the Director. Moreover, it is not clear if there are layers of authority above the Observatory Director that are not shown on the chart. Finally, there are high-level science staff who direct much of the science effort yet are not clearly identified in this chart. The reporting mechanisms to NSF are also not indicated on the chart.

This is a critical time for clarifying these structural and decision-making issues. Battelle has succeeded in accomplishing the first critical task -- Observatory Construction. Now they are poised to succeed in the second critical challenge—developing a productive, collaborative relationship between the engineering and corporate processes/structures and the scientific enterprise. To aid in this success, we have several specific recommendations.

Specific recommendations:

- ***The status of the Observatory Director as PI of the NSF award needs to be clarified and finalized.*** The Observatory Director needs to be recognized and seen as the Chief Scientist internally and as the face of NEON externally to the user community and NSF.
- ***A clear chain of command for scientific, financial, and personnel decisions needs to be specified and made transparent.*** This is especially critical when difficult issues arise where a final decision needs to be made. For example, when a site is affected by a severe natural disaster, who will make decisions about how to balance the need to rebuild quickly with the opportunities to do high-impact, "opportunistic" science? In Operations, ***decision-making must be driven by science***, and the Observatory Director must be empowered to *lead* that decision-making process.
- ***The lines of communication between NEON and NSF need to be clarified.*** As PI of the NSF award, the Observatory Director should be responsible for all communications with NSF.
- ***The lines of communication between NEON and other administrative structures at Battelle also need to be clarified.*** Although the organizational chart shows the Chief Scientist at the head of the organization, our understanding is that the Chief Scientist, CEO, and Program Manager all report to different people at Battelle Memorial and it is unclear what implications this has for the decision-making process and ultimately for NEON science.
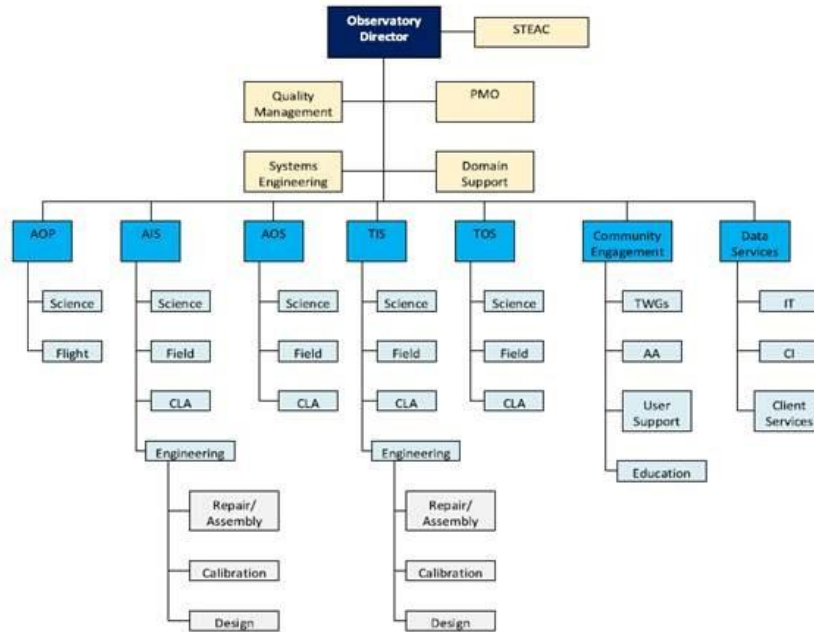
Figure 1. Current NEON organizational chart provided to STEAC on September 27, 2018.

II. Cyberinfrastructure

Overview. STEAC recognizes that the NEON cyber-infrastructure staff are a talented, dedicated, and hardworking group dealing with some very challenging aspects of the NEON enterprise. They are clearly making progress; for example, STEAC was impressed with the accessibility and aesthetic functionality of the new data browser. The new API and data portal are excellent additions, and the recent $3M NSF grant to support CI in enhancing data portal capabilities is a major achievement.

The STEAC sees a clear need for a comprehensive and sustainable plan for NEON cyberinfrastructure, with personnel leadership to support and organize these efforts across the Observatory. At this point of transition from Construction to Operations, the organization and development of a transparent and functional cyberinfrastructure system is critical. Cyberinfrastructure is a need that crosses all NEON functions, from data collection, QA/QC, data release to the community, and user feedback. New developments in the beta version of the data portal have improved the discoverability and interpretation of NEON datasets that are coming online. However, areas for improvement include: clear designation of personnel to organize cyberinfrastructure across the Observatory; more coordination among divisions of NEON that develop and use cyberinfrastructure; and collaboration with groups outside of NEON that share common issues and can communicate appropriate expertise via an advisory board, a CI technical working group (TWG), and/or informal meetings. NEON CI staff are making good use of cloud native/aligned technologies (kafka, docker etc.) to develop their ETL pipelines. These products are being widely utilized by many projects outside of NEON and discussing/sharing of best practices is essential.

Specific recommendations:
- *Designated Information Officer*. Designate an individual who is responsible for all aspects of IT, including data, technical and scientific computing, and security.

- *Advisory Board*. Establish a CI advisory board and TWG.
- *Open-source-driven development*. All NEON code (science, Cal/Val, technical etc.) should be on Github or equivalent. Such transparency of your roadmap/features is critical for the Observatory.
- *Increase visibility of CI staff contributions.* Highlight their work via blog posts, webinars, and newsletter items. The user community is looking to learn and collaborate.
- *Leverage NSF CISE investments.* Connect with existing NSF centers and CISE (Computer Information Science and Engineering) faculty to bring in best practices from other projects outside of NEON.
- *Training for NEON software products.* Provide regular training/workshops/webinars for NEON created software tools.

IV. Science Staff Professional Development

Overview. There are substantial opportunities within Battelle for professional development that can advance the skill set and growth of NEON science staff. However, we perceive a lack of clear communication regarding the nature of these opportunities and their availability to NEON staff. A widely disseminated understanding of career paths available for employee growth across the different staff positions appears to be lacking. Additionally, we continue to hear from staff that there is little opportunity (due to time constraints) for science-oriented NEON employees to advance their own research interest and skills. This kind of scholarship is key to making NEON science staff attractive collaborators to the broader user community.

We note that the integration of Battelle Ecology with Battelle Memorial Institute makes sense and is a positive development. However, science staff at several levels indicated that there has been a lack of clear communication and transparency regarding the ramifications for their benefits, vacation time, and intellectual property development. Another critical challenge is merging the corporate and academic cultures that are an inherent part of NEON. Continued efforts to clearly communicate the nature and extent of these changes will help attract and retain personnel at all levels.

Specific recommendations:
- *Provide a clear and specific plan for professional career development and advancement for the population of NEON employees transitioning to Battelle Memorial.*
- *Encourage training/professional development opportunities that exist outside the Battelle organization.* In addition to providing details on Battelle-specific training, encourage NEON staff to develop connections with the broader user community. It is essential to develop a culture within NEON where scientific collaboration is valued and strongly encouraged. This will require that staff be given time to devote to these activities and not be made to feel that such endeavors are counter to their ongoing work or that these activities need to be done on their own time.
- *Ensure that NEON staff have access to relevant journals.* We learned from staff that there is a Battelle library, but that it does not provide access to journals required for their environmentally/ecologically-focused work (e.g., ESA journals). NEON is an ecological project, and staff scientists require access to ecological and other relevant journals to be successful in their positions.

V. Data Availability and Training

Overview. STEAC recognizes and applauds NEON's increased commitment to make data available to the public. It is exciting to see the data coming online, and STEAC is encouraged by the positive

reports from ESA and elsewhere that the community is starting to use NEON data. STEAC is also enthusiastic about some of NEON's "Prototype Data Products", for example, the new Eddy4R application -- an important proof-of-concept of the scalability of data processing algorithms. STEAC also recognizes that NEON staff are striving to achieve very high standards of data quality and availability. To this end, NEON staff recently created a document identifying thresholds of acceptability for release of high-quality data at the site and observatory levels. STEAC reviewed the document and while we appreciate NEON's aspirational goals of achieving very high levels of data accessibility, we have some recommendations for communicating these goals both internally and externally.

<u>Specific recommendations:</u>
- *Make data review by science and domain staff a top priority.* We continue to hear from NEON staff that they do not have time to review data and that this is not seen as a priority. Staff review of data and testing of data products is critical to ensure high data quality going forward; the Observatory will not be successful otherwise.
- *Develop realistic, incremental benchmarks for availability of data products.* Starting at 80-95% availability, as written, for the majority of data products within and across sites, is a great aspirational goal, but unlikely to be attained in the early operational stages of the Observatory (and possibly even later, based on the success rates of other PI-driven ecological projects). STEAC believes it is important both for internal and external communications to have more realistic metrics, such as a rate of improvement over time (e.g., a 5% increase in availability each year). NEON staff scientists should consult with their TWGs or other experts in the field to develop realistic goals for the highly diverse data streams.
- *Re-evaluate communication of allowable and unallowable reasons for data loss.* The STEAC had difficulty interpreting Table 1 in the change-control document and had concerns that some of the "un-allowable" reasons for data loss were beyond their control (e.g. they really were 'unavoidable.'). We would like to see this revised and relaxed to more easily understood and realistic metrics, which will help set up NEON science and scientists for success.

VI. Field Staff Training
Overview. NEON trains ~300 field staff per year, and the quality of NEON data depends on the quality of training that these field staff receive. NEON has developed, and continues to develop, the training program for domain field staff. The field audits being developed now, if designed well, will provide valuable feedback about the structure of the overall training program and effectiveness of the individual training modules. Because the measurements being made across domains are widely made more generally in the field of ecology, additional investment in the quality of these training materials (e.g., videos) could broadly benefit the user community, contributing to NEON's overall mission related to community engagement.

<u>Specific recommendations:</u>
- *Short-term investment in training materials.* A short-term commitment of additional staff to produce high-quality training materials (e.g., written and video materials) would benefit both NEON and the broader user community.
- *Design field audits to provide actionable information.* It is our understanding that field audits were part of the site commissioning process (i.e., transitioning sites from Construction to Operations phases), but these may have been dropped. Are there currently planned field audits? If not, there should be. If yes, are these audits designed to identify critical training

issues and to differentiate whether identified problems are associated with individual training modules, experimental protocols, management, field schedules, or broader aspects of training such as the amount of training, the hiring wave in which staff were added, the length of service of staff members, etc? This information is critical to effectively assess the success of the training program and to address in a timely fashion any issues that will affect data quality and availability.

- *Consider adding a dedicated trainer at NEON HQ for training new full-time domain staff.* Data quality and availability are obviously directly tied to a highly trained workforce. We recommend that a dedicated trainer of new full-time field staff be hired to provide uniform, high-quality training to not only prepare staff to make measurements themselves, but also to properly train new field staff at individual domains (i.e., training the trainers). This position could provide a new career ladder opportunity for high-performing, experienced domain field staff.