Advisory Report to Battelle
Science, Technology and Education Advisory Committee (STEAC)
July 2017

Overview
The global environment is changing rapidly, with effects on ecological systems occurring across large spatial extents from regions to continents[1,2]. The impact of these changes on U.S. ecosystems is expected to have unprecedented ecological, human health, and socioeconomic consequences for society. Understanding, predicting, and managing the impacts of environmental change requires a long-term, multi-scale, nationwide strategy. The National Ecological Observatory Network (NEON), currently in the final phase of construction, is the first ecological observatory to observe and enable the documentation and forecasting of the nature and pace of environmental change at the continental scale[3]. The NEON research platform is designed to address grand challenges in environmental science, with research focused on questions that are relevant to large geographic regions and that cannot be addressed solely with traditional ecological approaches[4]. NEON will provide long-term (30 years), large-scale datasets and provide a research and educational platform for investigator-initiated sensors, observations and experiments[4]. The Observatory is fundamentally unique in its coordinated strategy to provide critical biological and physical observations at sites spanning the wide ecological and climatic variability found across the US. The Observatory compliments localized research occurring in other networks (e.g., Long-term Ecological Research Network; Critical Zone Observatory Network)[3,4] and its construction coincides with the era of "big data", placing the Observatory in a unique position to lead the ecological community toward a culture that embraces transparent and reproducible data aggregation, preservation and exchange through the development of approaches and technologies for handling large, heterogeneous datasets[5].

The NEON Project is currently at a critical juncture as it transitions from construction to initial operations, and it is prudent to now reflect, evaluate, and assess what this powerful instrument can achieve and to ensure, via strategic engagement, that the ecological community has both confidence and interest in the data being delivered. To this end, the NEON Science, Technology, and Education Advisory Committee (STEAC) convened a two-day meeting June 21-22, 2017 in Boulder, CO. We focused on the following key areas: (1) guiding principles for prioritizing NEON science; (2) data quality and protocol optimization; (3) cyberinfrastructure; (4) initial operations options; (5) alternative operations models; and (6) communication and engagement. The following observations and recommendations reflect the consensus of STEAC members who participated in the meeting (16 of 20 members were in attendance).

Guiding Principles for Prioritizing NEON Science
NEON is designed to enable the understanding and forecasting of the nature and pace of environmental change at the continental scale. As NEON construction nears completion, it is critical for the ecological community to establish a clear set of criteria to guide the prioritization of research questions and to assess how much scientific value will be lost when budgetary or other constraints dictate changes to NEON infrastructure and operations. The science questions with the highest priority should drive any decisions on the type, number, and scheduling of measurements. Given that NEON is inherently an observational platform, and not a series of planned, hypothesis-driven experiments, we suggest the following guiding principles will maximize NEON utility: (1) maintenance of long-term, high quality data, (2) a focus on regional to continental-scale observations and questions, (3) robust coordination of spatiotemporally-distributed observations across sites, (4) prioritization of time-sensitive data and leveraging of future discoveries; (5) synergies with other observatories (e.g., LTER, CZO); and (6) high societal relevance.

Data Quality and Protocol Optimization

The longstanding promise of NEON is to provide high quality, research-ready data to the ecological community. It is critical that this aim be met and that the community values and has trust in all data products. This puts a premium on stringent data quality assurance and control (QA/QC). We recognize that much effort has been and continues to be put into QA/QC protocols. However, there are concerns about a real or perceived lack of alignment between protocols developed and implemented by staff at NEON central headquarters versus issues encountered by onsite Domain science staff and community users who have downloaded NEON data. There should be transparently documented QA/QC at each step along the data stream, from the moment it is collected in the field to when it is posted on the data portal. Further QA/QC efforts should be coordinated across the organization. We therefore recommend a cross-NEON data quality technical working group (TWG) that is charged with aggressively pursuing data quality issues. We further encourage Battelle to make QA/QC a key responsibility and supported function in all staff positions, from field technicians to cyberinfrastructure developers. We also note the critical need for community engagement and feedback on data quality and recommend that mechanisms be pursued to capture community feedback as data are rolled out. Finally, QA/QC is one of many areas where we see a strong need for adaptive management. There needs to be flexibility for changing protocols and reporting mechanisms based on lessons learned during initial operations.

Cyberinfrastructure

The success of NEON in acting as a "single instrument" delivering complex ecological data fundamentally rests on robust computational infrastructure serving open, near real-time observations in accessible and reproducible workflows. We recognize the significant progress made on development of the data portal, use of mobile platforms for standardizing and improving data collection, QA/QC activities, and development of Application Programming Interfaces (API) for data access. We also note progress on data science outreach, including educational modules, data workshops, and community-developed packages for accessing and analyzing data. However, concerns remain about how data will be delivered in a useable format and how to initiate and encourage community involvement in data QA/QC. Of core importance will be efforts to quantify who is using the data and how it is being used; these metrics will be needed to justify long-term cyberinfrastructure evolution. Attention is also needed on open standards, visibility on social coding platforms such as GitHub, and strong adherence to FAIR (Findable Accessible Interoperable Reproducible) principles of transparency. Providing visibility into all phases of the data lifecycle is essential for ensuring confidence in final and derived products and should be given high priority. For example, the versions of scripts/tools used to verify/reformat raw data should be viewable in GitHub (or any open source code repository), with appropriate version numbers and the options/parameters utilized being included in the data bundle manifest. Engaging with eco-informatics researchers who actively practice pragmatic FAIR adoption will facilitate best practices for NEON data.

Initial Operations Options

The NEON Initial Operations Options Whitepaper (Jan 2017) provided alternative scenarios for reducing the operating costs of the Observatory. Three options were outlined: a $65M option submitted to NSF in December 2016, an alternative $65M budget, and a $60M budget. The proposed options come down to a choice between (1) reducing the number of total observing sites (while maintaining the full suite of measurements at the remaining sites), (2) reducing what is being sampled across all sites, or (3) something in between. Our current understanding is that Battelle is proceeding with construction of the full set of sites and that NSF has expressed support for this approach. Since it would be a mistake, from both a scientific and community engagement perspective, to construct all sites and then defer operations at a subset of them (i.e., "mothball" sites; alternative $65M option), that leaves the initial $65M (Dec 2016) option as the most viable alternative. We generally support the idea of implementing all sites, as

this creates the best scenario for the future function and development of the network. Deferring all operations at some sites would not have wide community acceptance and since substantial investment has already been made to construct all sites, the cost savings from this approach would likely be minimal and disproportionate to the impact on science and public buy-in for the Project.

The great challenge that STEAC now faces is how to evaluate the effect of the measurement reductions that have been proposed under the initial $65M (Dec 2016) option. We appreciate the great effort that went into detailing these reductions as described in the Whitepaper. However, since the design and scope for the NEON Project was initially approved, scientific questions and opportunities have changed, as have the methods used to quantify the nature and magnitude of ecological change. Thus, it is very important to take an adaptive management approach going forward, with continuous evaluation of the impact of the reductions on data quality and quantity. Indeed, implicit in the original NEON design was the expectation that the Observatory would evolve over time as ecological science evolves and matures. At the same time that the NEON Project moves ahead with "bottom up" implementation and adaptive evaluation across all sites, there should be a higher-level "top down" evaluation of whether the measurements are achieving the overarching science goals. It may be useful to base this evaluation on the Science Traceability Matrix, though we note this document is incomplete and should be reviewed and revised as the program evolves.

Alternative Operation Models

One alternative operating model presented to STEAC by Battelle was a planned pilot project in which some NEON field measurements would be conducted by scientists at institutions located near NEON installations (i.e., colleges and universities, Natural Heritage Programs, state and local government agencies, NGOs). Following an extended discussion of this topic, we strongly recommend that Battelle explore this possibility through funding of 2-3 prototype sites. At this time, we are not suggesting that this be the primary model for operating NEON, but rather a model that should be investigated as a potential mechanism to reduce cost, increase efficiencies, bring in site-specific expertise that could benefit NEON data quality, and enhance community engagement. The latter is especially critical given that the current Operations budget is perceived by many in the ecological community to be shrinking core NSF-DEB grant programs. One drawback might be the associated paperwork for establishing contracts and the time commitment needed for oversight. Yet, once protocols and expectations are established, these issues should be minimal. We recommend that, if this model is more fully implemented, that the funding mechanism be via a contract (not a grant) with explicit expectations, milestones and deliverables, and that the work be conducted in close collaboration with the respective NEON Domain staff.

Communication and Engagement

We recommend that Battelle embrace a comprehensive approach to NEON branding, communication, engagement, and marketing as a singular priority. Messaging confusion can arise internally and externally creating barriers to success associated with inherent cultural differences among partners and participants. For example, differing engineering and scientific expectations for Observatory performance during commissioning can impact perceptions of data quality and utility that may influence how the research community embraces the Observatory throughout its existence. This includes the issue of the role that NEON science staff play in directly addressing grand challenge questions with NEON data in collaboration with the research community. While NEON represents a unique and critical mission, the current lack of a comprehensive, compelling narrative has negative impacts on stakeholder perceptions and employee self-valuation that perpetuate programmatic and human resource challenges. Attention to these issues is imperative. STEAC is working with NEON staff to develop a set of high-level, guiding principles for a comprehensive communications and engagement strategy, recognizing that engagement includes: empowering all employees in service, education, active outreach, and engagement; developing

accountability for stakeholder interactions at the highest level; ensuring access and equity to opportunities to participate or partner; providing a "human" face to users who seek information; and facilitating professional development of students, faculty, and other members of the research community to engage in transformational science.

Battelle should embrace transparency in all aspects of budgeting, operations, and resource allocation to overcome existing challenges, such as the perceived "drain on funds" that NEON imposes on other areas of NSF-supported research. A narrative should be created to (1) describe the impact that major research infrastructure has had historically on advancing scientific disciplines and (2) that highlights the lasting and transformative impact that NEON will have on science and society, with an emphasis on human, discipline, technological, and institutional outcomes that have resulted from the visioning, construction, and commissioning of this impressive endeavor. NSF also has a role in facilitating community engagement by sufficiently funding core grant programs, along with those programs (e.g., Macrosystems) that directly support and leverage the collective NEON effort. These collective actions will set the stage for stakeholders to develop confidence in NEON data products and their own ability to address environmental and societal grand challenges.

Summary of Recommended Action Items

- Adopt guiding principles for prioritizing NEON science
- Create a cross-cutting TWG for data QA/QC and protocol optimization
- Make clear and defensible decisions about release of preliminary/intermediate data products
- Implement mechanisms for user community to provide feedback on data quality
- Enhance discussion among NEON headquarters staff and Domain managers to ensure consistent and effective development and implementation of protocols
- Develop a written roadmap describing short and long-term cyberinfrastructure development plans
- Engage with eco-informatics community on FAIR adoption
- Adopt initial $65M (Dec 2016) option as most viable alternative, but… (see next bullet)
- Utilize an adaptive management approach with continuous evaluation of how science decisions impact data quantity and quality
- Periodically review and refine the Science Traceability Matrix as the NEON Project evolves
- Implement a pilot project that contracts with local entities to collect specified NEON data
- Continue to develop a comprehensive strategy for branding, communication, and engagement

References

[1]Peters, D.P.C., H.W. Loescher, M.D. SanClements, and K.M. Havstad. 2014. Taking the pulse of a continent: expanding site-based research infrastructure for regional-to continental-scale ecology. Ecosphere 5(3): art 29.

[2]Peters, D.P.C., P.M. Groffman, K.J. Nadelhoffer, N.B. Grimm, S.L. Collins, W.K. Michener, and M.A Huston. 2008. Living in an increasingly connected world: a framework for continental-scale environmental science. Frontiers in Ecology and the Environment 6(5): 229-237.

[3]Keller, M., D.S. Schimel, W.H. Hargrove, and F.M. Hoffman. 2008. A continental strategy for the National Ecological Observatory Network. Frontiers in Ecology and the Environment 6(5): 282-284.

[4]Kampe, T.U., B.R. Johnson, M. Kuester, and M. Keller. 2010. NEON: the first continental-scale ecological observatory with airborne remote sensing of vegetation canopy biochemistry and structure. Journal of Applied Remote Sensing 4, 043510.

[5]Hampton, S.E., C.A. Strasser, J.J. Tewksbury, W.K. Gram, A.E. Budden, A.L. Batcheller, C.S. Duke, and J.H. Porter. 2013. Big data and the future of ecology. *Frontiers in Ecology and the Environment* 11(3): 156–162