Science, Technology & Education Advisory Committee (STEAC)

MEETING REPORT
(12/6/2022)

The STEAC met for a hybrid meeting on December 5th and 6th, 2022. The in-person meeting was held at NEON Headquarters in Boulder, Colorado and included virtual members via videoconferencing. The following members attended the meeting: Meghan Avolio, Henry Bart, Sarah Bevins, Michael Dietze, Richard Fiorella, Shannon LaDeau, Sparkle Malone, Jackie Matthes, Kim Novick, Steve Petruzza, Sydne Record, Daniel Rubenstein, Shawn Serbin, and Adrienne Sponberg. The following NEON-Battelle staff attended: Paula Mabee, Kate Thibault, Katie Jones, Kirsten Ruiz, Chris Florian, Kim Nitschke, Chris McKay, Claire Lunch, Christine Laney, Cove Sturtevant, Bridget Hass, Steve Jacobs, Bonnie Meinke, Courtney Meier, Chris McKay, Keli Goodman, Peter Weishampel, David Mitchell, and Ty Lindberg. The following Battelle staff attended: Storm Woods, John McCracken. Nico Franz attended. The following staff from the Environmental Data Science Innovation & Inclusion Lab (ESIIL) attended: Jennifer Balch, Chelsea Nagy. Several NEON Technical Working Group (TWG) members also joined for their relevant sessions, as detailed below, including Drs. Sara Spaulding, Marco Cantonati, Alison Donnelly, and Michael Cosh. The NSF NEON Program Director, Dr. Charlotte Roehm participated in an hour meeting with the STEAC.

The meeting was hybrid, and the following topics were discussed: I. TOS Sampling - Long-Term Strategy, Communication Plans, II. Community Empowerment, III. Google Cloud Platform for NEON users, IV. ESIIL and NEON, V. Terrestrial Instrument System Optimizations, VI. Meeting with NSF NEON Program Officer, Dr. Charlotte Roehm, VII. Algal Taxonomy Data Product Optimization Proposal, VIII. Plant Phenology Data Product Optimization Proposal, IX. Recruiting, X. Biorepository Sample Use, and XI. STEAC Business.

The National Ecological Observatory Network has been in collecting and providing data for more than 3 years. Data is being collected and used throughout the ecological community and NEON is working to connect with other networks, agencies, and centers. Currently, NEON is considering the impact of flat budgets and increasing costs, while adapting to advances in science. Designed to be a community resource, the community must decide how to balance science innovation and support for current and well used techniques. These decisions are largely made by the community through members of the TWGs and evaluated by the STEAC. Decisions are also heavily impacted by funding.

The members of the STEAC applaud NEON's commitment to the scientific community. Recognizing that NEON is responding to staffing and recruitment challenges in addition to budget and resource limitations, the members of the STEAC encourage NEON to ensure that Operation Phase II sampling includes a long-term strategic plan that fully addresses budget limitations while also allowing for NEON to build in the capacity needed for enhanced adaptability. The STEAC recommends that NEON be proactive about technological advances, in particular image recognition and artificial intelligence

and machine learning (AI/ML), that might reduce both costs and data latency in the long-term.

The members of the STEAC encourage NEON to pursue partnerships with organizations that share science, technology and justice interests. The partnerships can be designed to increase the impact and potentially share the financial burden of sampling innovation, data collection, outreach, and community engagement. In addition, partnerships can expand the NEON audience beyond the scientific community.

I. TOS Sampling - Long-Term Strategy and Communication Plan

NEON presented a thoughtful data-driven plan to reduce Terrestrial Observation System (TOS) sampling at select gradient sites to produce 70% cost savings in response to budget short falls projected for 2023. The proposal was limited to 2023 because Battelle's agreement with the NSF only lasts through October 2023 and, due to the current competition to manage NEON Operations, funding levels for NEON Operations are not yet known. The plan includes the suspension of TOS sampling for one of the two terrestrial gradient sites in each of five Domains (D02, D05, D08, D09, and D19). The sites were selected because there is high potential redundancy, determined by land cover similarity. Other gradient sites were not considered because of the lack of similarity in the national landcover database between the core and gradient sites.

The members of the STEAC appreciate the data-driven approach taken to develop this plan in response to budget and resource limitations projected for 2023. The STEAC discussed the methods used (land cover and similarity to reduce redundancy). While the members of the STEAC support the suspension of TOS sampling at 5 gradient sites in response to resource limitations, the members of the STEAC encourage NEON to consider a plan that reflects the full cost savings necessary for 2023 and to be more proactive about planning for 2024 and beyond. The STEAC encourages NEON to identify a sampling strategy that is manageable for the long-term, recognizing that the NEON budget may remain flat and that costs will continue to rise.

Recognizing that NEON answers to a large community, the members of the STEAC are interested in understanding the community preference for incorporating technological advancements into the NEON sampling design that might support a higher sampling frequency. While the current operating plan does not call for a relocation of the gradient sites, the members of the STEAC and TWGs have expressed support for changes to the sampling frequency at gradient sites in response to resource limitations. A community event (e.g., virtual town hall) to provide input on the future sampling may be insightful.

In the long-term, NEON might consider working with other networks and agencies to reduce the cost of field sampling. Encouraging other networks and agencies to use NEON protocols and getting their assistance with collecting data within NEON sites

might lower the cost of data collection, or provide training opportunities. NEON might also look to LTER sites that have been very successful in leveraging additional non-NSF funding (ex. Florida Coastal Everglades Long-term Ecological Research) to establish strategic partnerships at the Domain level.

II. Community Empowerment

Dr. Bonnie Meinke gave an update about community empowerment and engagement and then led the STEAC in an active brainstorming exercise about future NEON engagement strategies. The NEON Domain Outreach Liaison activities and the partnership with the AGU Thriving Earth Exchange program have led to productive local activities within NEON Domains. The STEAC were pleased to hear that the NEON Ambassador pilot program has been largely successful. The first cohort of NEON Ambassadors chose to focus on two core areas: sharing NEON resources and training graduate students in ways that create tangible products (e.g., workshops, open course materials). These types of activities and products will help to further engage the community in using NEON data.

The STEAC brainstorm exercise facilitated by Dr. Meinke focused on three main questions: 1) What audiences should NEON be working to engage? 2) What methods can NEON use? and 3) Which metrics can be used to assess engagement?

The STEAC identified academic researchers as a key primary audience but emphasized that NEON should work towards engagement with a diversity, equity, and inclusion (DEI) lens to engage communities not currently represented. Methods for outreach could include a larger social media presence which can be used to increase the visibility of NEON community resources, including their online form where authors who have used NEON data can notify NEON and the wider research community about new papers, talks, presentations, and key findings. NEON should promote the use and increase the visibility of the NEON Code Hub, where co-derived data product sharing resources are made available to avoid duplication of effort and to facilitate broader use of NEON data. We did not have substantial time to discuss metrics, which is an important topic for future conversations between members of the STEAC and NEON.

III. Google Cloud Platform for NEON Users

NEON Cyberinfrastructure staff including Steve Jacobs and Drs. Bridget Hass, Christine Laney, and Claire Lunch gave an update on NEON's data storage and compute migration to the Google Cloud.

Overall, the members of the STEAC were highly supportive of this move and the opportunities that it presents across the full spectrum of NEON users, from entry-level users who may not have access to large computational resources to power users who are primed to take advantage of cloud-native data.

The members of the STEAC encourage NEON to build platform-agnostic solutions (e.g., workflows, data structures, data access approaches). This could be very valuable

if NEON ever needs to migrate the system to a different cloud provider, allows users to leverage general (rather than platform-specific) tools, and reduces unexpected user costs. Staying platform-agnostic should also facilitate interoperability with ESIIL / CyVerse. We acknowledge that this recommendation cannot be implemented universally (e.g., overall system monitoring), but where optional (e.g., integrating data into Google BigQuery) we think it makes sense to hold back on adoption until there is explicit community demand/interest, particularly in light of overall budgetary limitations. That said, Google Earth Engine provides an example of where there is wide adoption of the platform by the community and demand for AOP on that platform.

IV. ESIIL and NEON

Drs. Balch (PI) and Nagy (SP) from the University of Colorado (CU) introduced the new NSF-funded Environmental Data Science Innovation & Inclusion Lab (ESIIL). A key part of the new center's mission (full details at https://esiil.org/) is to build data constellations and novel analytics to enable open, cloud-enabled synthesis. Dr. Balch identified key partnerships with CyVerse (https://cyverse.org/) for cyberinfrastructure goals and the University of Oslo for Team Science goals. She noted potential for engaging with NEON through shared mentoring opportunities for postdocs and requested NEON help in promoting the Innovation Summit at ESIIL in May (2023). Other discussion points included the need for new bridges between computer science and ecology. Dr. Balch further noted that ESIIL aims to develop platform-agnostic cloud solutions.

The members of the STEAC encourage NEON to continue to work closely with ESIIL to look for opportunities to support shared goals.

V. Terrestrial Instrument System Optimizations

NEON has been investigating possibilities to optimize the Terrestrial Instrument System (TIS) data collection to minimize impact on scientific products of rising costs in the context of a flat budget. Dr. Chris Florian presented an analysis carried out by NEON staff members focusing on the data collected in the soil arrays at terrestrial sites, most notably the soil moisture and temperature sensors. He was joined by Dr. Michael Cosh, Chair of the Soil Sensor TWG, who provided a summary of the deliberation of the TWG as well as its overall recommendation.

The current soil temperature and moisture profiles at NEON terrestrial sites consist of two different sensors, one each for temperature and soil moisture, installed at up to nine measurement levels per profile and at five profiles per site for a total of up to 85 soil sensors per site (up to 45 soil temperature sensors and 40 soil moisture sensors). The existing sensors have become obsolete, which has motivated a reconsideration of the sensors used in these profiles as well as their arrangement. NEON is selecting a new sensor that will measure both soil temperature and water content, and it is currently expected to be installed at up to seven levels in five new profiles per site, for a total of 35 soil sensors per site. The new sensors will be installed adjacent to existing soil profiles, and existing sensors will remain in place to help promote continuity across

the transition, with the goal of producing a 1-year overlap between sensors.

Several alternative proposals for reducing the number of new sensors per site have been advanced and analyzed by NEON, and considered by the soil sensors TWG, to realize cost savings while minimizing impact on scientific quality and utility of data from the soil arrays. NEON indicated that soil temperature and water content products are fairly popular data products in their archive, with 300-400 unique users per year, which underscores the importance of preserving the scientific utility of these data products. Three options were presented: 1) removing the deepest three sensors in plots 2 and 4 at each site, 2) removing an entire profile (either plot 2 or 4) at each site, or 3) removing two entire profiles (plots 2 and 4) at each site. NEON presented an analysis on how each option impacted data quality, with all three options being qualitatively similar.
Dr. Michael Cosh of the soil sensors TWG provided a summary of the TWG recommendations in response to these three proposals from NEON, which have been summarized in the table below (Table 1).

Table 1. Summary of the soil TWG recommendations.

Proposed change	**Advantages**	**Disadvantages**
Remove the deepest three sensors at plots 2 and 4	- Minimal impact on other data products due to collocation with soil CO_2 - Maximal sampling of spatial variation in dynamic surface soils for better connections with soil biogeochemistry and remote sensing applications	- May reduce utility of NEON soil moisture data for hydrology studies - Smallest cost savings of proposed options, so future reductions may be required
Remove a single profile (plot 2 or 4)	- Maintains full vertical profile at each plot, expected to be important for hydrology studies	- Loss of collocated temperature and moisture data for soil CO_2 sensors in one plot - Reduction of spatial representativeness - May cause NEON to miss its own data quality standards as well as those of broader community standards (e.g., ICOS)
Remove two profiles (plots 2 and 4)	- Maximal cost savings - Maintains vertical profiles required for hydrology studies	- Same disadvantages as removing a single plot, but worse

The overall recommendation the TWG delivered is that option 1 represents the “least worst” option, informed by the preference to observe the fastest changing portion of the soil column (i.e., the surface) with higher resolution than slower varying portions of the soil column (i.e., the sensors at depth). *The STEAC is supportive of the TWGs*

recommendation, though also highlights and shares the TWG's concern that more instrumentation changes may be needed in the future to realize further cost savings.

NEON is also considering changes to or the elimination of the dust and particulate matter size distribution and mass data products. However, there was insufficient time to cover NEON's analysis of this data product at the December meeting. The STEAC anticipates this data product will be discussed in greater detail at an upcoming meeting and takes no position on the dust and particulate matter data products at this time.

VI. Meeting with NSF NEON Program Director, Dr. Charlotte Roehm

The STEAC appreciated the opportunity to interact with Dr. Roehm from the National Science Foundation. Dr. Roehm stressed the importance of the continued use of and relevance of the data products for a diverse group of users. Dr. Roehm noted that the diversity of use of the data has been encouraging as evidenced by >500 publications using NEON data. After emphasizing that the STEAC's input is critical to moving forward successfully, Dr. Roehm invited questions.

Major themes to these questions resonated with many of the challenges members of the STEAC grappled with during this two-day meeting. We discussed flat budgets in light of increasing costs, how flat budgets impact retention of early career scientists, and how the NSF can help to support projects with budget shortfalls beyond their control. Dr. Roehm noted that all programs are being negatively impacted by flat budgets and inflation and that NSF has no concrete solutions at the moment. Dr. Roehm also stressed that the NSF tries to keep its major investments going, while at the same time emphasizing that partnerships that build larger networks and more users might provide greater short-term flexibility in mitigating budgetary challenges.

Another major theme of the discussion pertained to how NEON will be able to remain nimble to novel advances in technologies (e.g., artificial intelligence and machine learning for gathering data from tick drag cloths) and instruments over its projected 30-year collection period. Dr. Roehm acknowledged the disconnect between tools available now and those used by NEON and gave a concrete example of using Assignable Assets to add new instrumentation to towers. Dr. Roehm also mentioned that the STEAC and user community may help to identify new needs and that partnerships could play a role in addressing this challenge, too. This spurred a question about how to get the broader technological community engaged in updating and upgrading NEON's sampling technology. There was discussion around Research Coordination Networks (RCNs) to bring the community together around NEON's current capabilities and future directions.

Another major theme of questions had to do with community engagement with NEON, if NSF is happy with the levels of success of NEON proposals, and who is responsible for the use of NEON data. Dr. Roehm suggested that there are many possibilities for expanding community awareness and building capacity across communities (e.g., synergies with the ESIIL, RCNs, STEAC). Regarding funding within the NSF, Dr.

Roehm said there were no plans for any new NEON-specific calls and reminded the members of the STEAC that other non-NSF agencies can also leverage use of NEON samples and data. Overall, the NSF is seeing an increased use of NEON data and samples in a diversity of proposals and publications, which is good.

NEON is working to remain relevant, which is evident by the increase in their diverse user community, growing partnerships, and the updates to sampling designsthat are considering new technologies. The STEAC recognizes the importance of the community's sense of ownership and engagement with NEON infrastructure. To increase community awareness on ways to engage with and advance NEON infrastructure the members of the STEAC discussed strategies that included adding an additional category for review panels to consider if proposals use NEON data, having specific RFPs (e.g., EAGERs), highlighting NEON more whenever there is a connection to the Observatory in NSF activity (e.g., in convergence accelerators), and we discussed ways to make the Observatory more visible to NSF writ large (e.g., NEON days that happened at NSF before the pandemic).

VII. Algal Taxonomy Data Product Optimization Proposal

Dr. Kate Thibault presented an analysis carried out by NEON staff members focusing on the micro-algal taxonomy data collected at aquatic sites. She was joined by Drs. Sarah Spaulding and Marco Cantonati of the Algal Taxonomy TWG, who represented their TWG recommendations and provided the necessary expertise throughout the discussion. In summary, there are challenges to ensure consistency of taxonomic algal data that is processed across 3 labs contracted by NEON. To respond to consensus discrepancies that result from the use of different voucher flora libraries across labs, NEON's Algal Taxonomy TWG provided an optimization proposal. The proposal included a 4-step approach that improves data quality and aligns with existing national-scale methods.

1) Align benthic sampling with the Environmental Protection Agency (EPA) National Rivers and Streams Assessment (NRSA) methods, combining multiple field samples into one composite sample per site and date.
2) Follow EPA NRSA and EPA National Lakes Assessment (NLA) laboratory analysis, with 600 valve counts for diatoms only (no soft algae) in benthic samples from all sites and 300 cell counts that include soft algae and diatoms, often identified to higher taxon ranks than species via this method, for phytoplankton in lakes and rivers (n = 10 sites).
3) Implement a voucher flora approach (e.g., Alers-Garcia et al. 2021) for taxonomic consistency using the NRSA diatom flora being generated for the 9 EPA ecoregions.
4) Archive diatom slides from benthic samples and preserved soft algae from phytoplankton, as well as a new composite sample that would enable metabarcoding analysis in the future, once a reference library is developed by the greater scientific community.

Additionally, NEON will archive existing samples stored at NEON Domain Support Facilities from AY2021-AY2022, and possibly AY2023 depending on the timeline for availability of the NRSA diatom voucher flora, without conducting taxonomic analysis.

These changes were suggested by the 2021 Algal Taxonomy TWG, are supported by the 2022 Aquatic Biology TWG, and will help ensure that NEON is using its funds in an efficient manner that produces quality data for the user community.

The members of the STEAC appreciate the deep discussion of the limitations of storing samples and the necessary changes in the sampling frequency and intensity needed to ensure that the taxonomic data is compatible across samples processed at 3 labs. The members of the STEAC discussed at great length the challenges and advantages of metabarcoding. Since only 1/5th of the species have reference keys at this time, moving to metabarcoding would lead to the loss of taxonomic information, which is less than ideal in the taxonomic community, though this "loss" would be backwards recoverable as genomic libraries continue to grow. The TWG communicated to the STEAC that there creating species-specific barcode libraries would be a large and extremely expensive effort *so a move to metabarcoding would move the community in a direction of taxonomic free analysis in the future. Encouraging research in this area is needed before NEON can incorporate metabarcoding into the sampling. The members of the STEAC also discussed the potential for image analysis of diatoms to be used. Digital images of diatom slides are being archived and the STEAC encourages the collection of digital biological collections, especially in light of concerns about the long-term preservation of algae within samples, as well as the potential for reanalysis once image analysis workflows have been developed. While members of the TWG indicated that some labs are actively working on developing these technologies, image processing would also require advancements before it would become feasible for NEON to incorporate image processing into the sampling scheme. The STEAC supports the Algal Taxonomy Data Product Optimization Proposal and encourages NEON and the TWG to look for and encourage opportunities to advance the science of metabarcoding and image analysis in the algal taxonomy community.*

VIII. Phenology TWG

Dr. Katie Jones presented an analysis carried out by NEON staff members focusing on the plant phenology data collected at terrestrial sites. She was joined by Dr. Allison Donnelly, Chair of the Terrestrial Plant Diversity and Phenology TWG, who provided a summary of the deliberation of the TWG as well as its overall recommendation. The Phase I ground-based phenology records provided by NEON (representing monitoring of 30 individuals of 3 dominant species at each terrestrial site) are already a valuable resource that are perhaps unparalleled for their completeness and degree of standardization. Moreover, the members of the STEAC were encouraged by the preliminary data from Phase II sampling (5 individuals of up to 20 species) which reveals the extent of phenological trait diversity. However, the members of the STEAC recognize that the 3x/week sampling during periods of rapid phenological change carries major labor costs, especially in more remote gradient sites. The use of Bayesian modeling to explore the uncertainty associated with modified sampling scenarios was appreciated. The members of the STEAC also enjoyed hearing from representatives of the Terrestrial Plant Diversity and Phenology TWG about their review of, and recommendations concerning, the modified phenology sampling proposals. Overall, it

seems like the mechanisms for review and feedback between NEON and this TWG are working well.

The TWG recommends:
1. Continuing with full dynamic sampling (high frequency spring and fall) at the core site at each seasonal domain
2. Reduce frequency at gradient sites to 1x/wk, retain 1x/2wk mid-season frequency where previously scheduled
3. (if the two options above do not meet required reductions) Drop phenology observations completely at sites

Overall, the STEAC recognizes that the original protocols may embed a fair amount of temporal oversampling, and support efforts to reduce sampling frequency to 2x a week, particularly in gradient sites. While the members of the STEAC support the TWG recommended changes for 2023, we share the TWG's concerns that 1x a week has the potential to compromise the integrity of the data in some sites, and they recommend that NEON repeat the uncertainty modeling with a more random sampling interval (e.g., every 5-10 days) to reflect the fact that precise weekly sampling is often challenged by weather and other logistical factors. Updated analyses might also assess the ability of proposed protocol changes to detect phenological changes over time, not just produce accurate estimates within a single year. The STEAC also recommends that NEON consider delaying the shift in sampling protocols, at least in some sites (i.e., core sites), given that the combination of this shift and the reduction in sampling frequency produced the highest uncertainties. In the longer term, the members of the STEAC recommend that NEON investigate the potential to exploit long-term repeat digital imagery (e.g., from PhenoCams) to enhance the temporal resolution of phenology monitoring through integration of digital and manual observations.

IX. Recruiting

The members of the STEAC appreciate the information on recruitment and workforce diversity provided by both Battelle and NEON employees. It's clear that recruitment targets a variety of hiring platforms, conferences and job fairs, but that challenges remain. Many hiring challenges are site-specific and the STEAC members recognize that finding available and affordable housing in many areas remains an issue.

The members of the STEAC encourage NEON to enhance current engagement and recruitment by reaching out to MANRRS (Minorities in Ag, Nat Resources and Related Sciences) as well as SACNAS (Society for the Advancement of Chicanos/Hispanics and Native Americans in Science). The members of the STEAC recognize that the budget for engagement is very limited, preventing NEON from offering paid internships through NEON/Battelle, although it is widely recognized that well-paid internships can be a powerful DEI recruitment tool. As a result, the STEAC encourages NEON to engage with early career researchers through existing programs. For example, hosting a virtual Introduction to NEON data during the first week of REU programs.

Newly developed, aggregated demographic data on NEON employees was very useful and will help refine employee recruitment, but the STEAC is interested in having the data also include information on seasonal employees, instead of just permanent staff. Data on that portion of the workforce is key, since recruitment of seasonal field staff is one of the primary challenges that NEON Domain Managers face. The plan to take a more detailed look at employee demographics at the Domain level (instead of nationally aggregated data) will also be helpful in the future. The members of the STEAC are also concerned that, similar to challenges across academia, current pay rates and 'benefits' are not competitive with other seasonal jobs. The members of the STEAC are interested in exploring these challenges further.

X. NEON Biorepository

Dr. Nico Franz provided an update on the NEON Biorepository.

The members of the STEAC were pleased to see that there have been significant improvements in request processing since 2019 (~4 weeks), though complex requests can take substantially longer due to the time it takes to pull and extract samples with the user specified protocol. If the NEON Biorepository has the capacity to do the user specified preparations, they do it. For very complex requests, additional funding or efforts from PIs is necessary. The NEON Biorepository is providing a high level of service to ensure the wide use of samples. Current challenges include space limitations with a growing collection, improving the use and visibility of destructive samples, data portal and API challenges, and Biorepository discoverability. The members of the STEAC appreciate that the Biorepository is continuing to focus on building, promoting Biorepository services: i.e., higher accessibility and ease-of-process for researchers to attain custom-designed and -prepared sample loans and specimens. The STEAC also supports a focus on well-targeted outreach and networking to increase the impact and visibility of the NEON Biorepository.

XI. STEAC Business

The STEAC voted to approve the minutes for the 11/16/2022 meeting. Throughout the meeting the STEAC also identified topics to discuss in future monthly meetings: Phase II sampling frequency at gradient sites, plant phenology sampling optimization; community empowerment and engagement focused on enhancing community resources through grant opportunities, leveraged funding and outreach activities at the NEON Domain level, and developing partnerships for recruitment strategies in challenging Domains.