



The US National Ecological Observatory Network and the Global Biodiversity Framework: national research infrastructure with a global reach

Katherine M. Thibault^{1*}, Christine M. Laney¹, Kelsey M. Yule², Nico M. Franz² and Paula M. Mabee¹

¹National Ecological Observatory Network, Battelle, Boulder, CO 80301, USA

²Arizona State University, School of Life Sciences, Tempe, AZ 85287, USA

ARTICLE INFO

Received October 25, 2023

Revised November 15, 2023

Accepted November 15, 2023

Published on December 14, 2023

*Corresponding author

Katherine M. Thibault

E-mail kthibault@battelleecology.org

The US National Science Foundation's National Ecological Observatory Network (NEON) is a continental-scale program intended to provide open data, samples, and infrastructure to understand changing ecosystems for a period of 30 years. NEON collects co-located measurements of drivers of environmental change and biological responses, using standardized methods at 81 field sites to systematically sample variability and trends to enable inferences at regional to continental scales. Alongside key atmospheric and environmental variables, NEON measures the biodiversity of many taxa, including microbes, plants, and animals, and collects samples from these organisms for long-term archiving and research use. Here we review the composition and use of NEON resources to date as a whole and specific to biodiversity as an exemplar of the potential of national research infrastructure to contribute to globally relevant outcomes. Since NEON initiated full operations in 2019, NEON has produced, on average, 1.4 M records and over 32 TB of data per year across more than 180 data products, with 85 products that include taxonomic or other organismal information relevant to biodiversity science. NEON has also collected and curated more than 503,000 samples and specimens spanning all taxonomic domains of life, with up to 100,000 more to be added annually. Various metrics of use, including web portal visitation, data download and sample use requests, and scientific publications, reveal substantial interest from the global community in NEON. More than 47,000 unique IP addresses from around the world visit NEON's web portals each month, requesting on average 1.8 TB of data, and over 200 researchers have engaged in sample use requests from the NEON Biorepository. Through its many global partnerships, particularly with the Global Biodiversity Information Facility, NEON resources have been used in more than 900 scientific publications to date, with many using biodiversity data and samples. These outcomes demonstrate that the data and samples provided by NEON, situated in a broader network of national research infrastructures, are critical to scientists, conservation practitioners, and policy makers. They enable effective approaches to meeting global targets, such as those captured in the Kunming-Montreal Global Biodiversity Framework.

Keywords: biodiversity, biorepository, climate change, ecological observatory, NEON, open data

Introduction

The National Ecological Observatory Network (NEON) collects and provides free and open high-quality continental and decadal-scale data for fundamental research on ecological systems (Keller et al. 2008). NEON is solely funded by the US National Science Foundation (NSF) and operated by Battelle. Data and samples are collected from 81 terrestrial and aquatic field sites, distributed throughout

the continental US, Alaska, Hawaii, and Puerto Rico (see neonscience.org for more details). NEON provides >180 open datasets that capture environmental, biogeochemical, atmospheric, and taxonomic data from various ecosystems that are subject to an array of land management and other environmental impacts.

The NEON design is intended to enable scientific research that furthers our understanding of the interactions between the drivers of environmental change and its bio-



logical responses, including biodiversity (Schimel et al. 2011). NEON's sampling design is distinguished by the use of standardized methods across space and time and the coordinated, colocated collection of diverse and complementary datasets using instrumented, observational, and airborne remote sensing approaches (Fig. 1) (Meier et al. 2023). The standardized, spatially balanced designs underpinning NEON's data collection efforts are necessary for enabling regional to continental-scale analyses, as distinct from most other networks, such as the US Long-Term Ecological Research (LTER) network that was designed to investigate the mechanisms and dynamics of local-scale ecosystem processes (Hobbie et al. 2003). The differing approaches of NEON and other networks like LTER are highly complementary and critical to informing global solutions to ongoing ecological challenges.

NEON provides both data and samples that are suited to addressing many aspects of biodiversity and that span all taxonomic domains of life (Fig. 2). Indeed, biodiversity was one of the 'Grand Challenges in Environmental Sciences'

(National Research Council 2001) that served as the foundation for NEON's design. In aquatic freshwater ecosystems (represented by 34 field sites), NEON collects periphyton, phytoplankton, plants, macroinvertebrates, zooplankton, fish, and microbial samples from benthic and surface water that are sequenced for archaea, bacteria, and fungi (Parker and Utz 2022). Terrestrial sampling at 47 field sites targets plant phenology, diversity, biomass, and productivity, as well as the abundance, diversity, and (for select groups) pathogen status of birds, small mammals, ticks, mosquitoes, ground beetles, and soil microbes (Thorpe et al. 2016). Remotely sensed observations capture structural and spectral features of landscapes at 1 m resolution over the field sites (Musinsky et al. 2022).

Samples collected from all sites include environmental samples (e.g., soil, sediment), nontraditional organismal samples (e.g., ground roots, mammal feces, bulk bycatch samples), and traditional natural history voucher specimens (e.g., pinned mosquitoes, pressed herbarium samples). Most samples are curated at the NEON Biorepository,

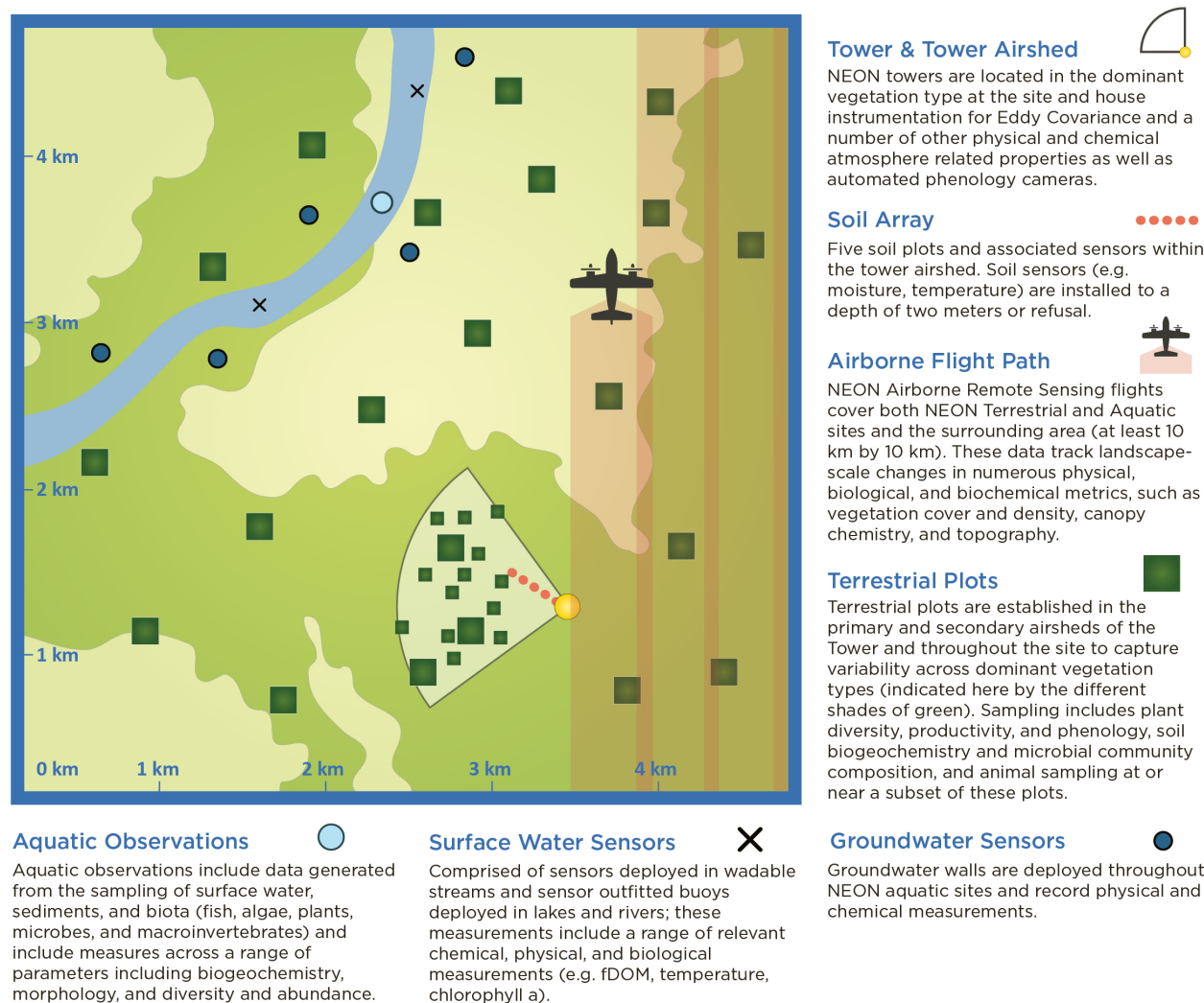


Fig. 1 Schematic of idealized colocated aquatic and terrestrial field sites within the US National Ecological Observatory Network, illustrating how the three collection systems, the observational sampling, instrument systems, and airborne remote sensing, are spatially integrated. Image credit: National Ecological Observatory Network/Battelle. By Williams C.

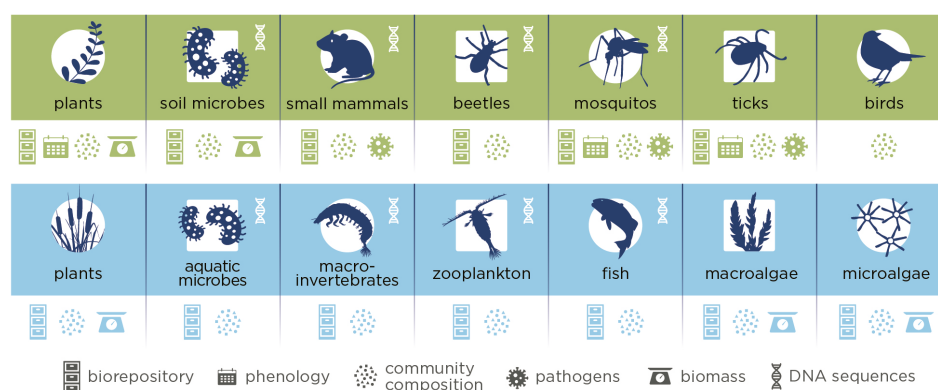


Fig. 2 Schematic of the organisms collected by National Ecological Observatory Network at terrestrial sites (in green, top row) and at aquatic sites (in blue, middle row), and the types of data and samples provided for each of the taxa represented by the smaller icons, as described by the legend at the bottom. Image credit: National Ecological Observatory Network/Bat-telle. By Williams C.

managed by Arizona State University's (ASU) Biodiversity Knowledge Integration Center in Tempe, Arizona, USA; nearly two-thirds of samples there are cryo-preserved. Unique compared to conventional natural history collections, a portion of most NEON Biorepository collections may be consumptively or destructively analyzed for research (Tazik et al. 2022). All data associated with the organismal samples curated at the NEON Biorepository are available from the Symbiota (Gries et al. 2014; <https://symbiota.org/>) portal at biorepo.neonscience.org, the Environmental Data Initiative (EDI) at <https://edirepository.org/> (Gries et al. 2023) and from the Global Biodiversity Information Facility at gbif.org (GBIF 2023b). NEON also shares its datasets for birds (Thibault et al. 2023) and ticks (Paull et al. 2022) via GBIF for enhanced discoverability and interoperability.

The data and samples provided by NEON and other global networks are key to yielding the depth of understanding required to achieve the ambitious goals of the global community to 'to halt and reverse biodiversity loss' as most recently codified in the Kunming-Montreal Global Biodiversity Framework (GBF) adopted by the 15th Conference of Parties of the United Nations Convention on Biological Diversity in December of 2022 (Convention on Biological Diversity 2022). The purpose of this study is to review the composition and use of NEON resources to date as a whole and specific to biodiversity, as an example of the impact of national-scale research infrastructure on global efforts such as the GBF.

Materials and Methods

We compiled data to summarize NEON's contributions to biodiversity science to date and four categories of use of NEON: web portal visitation, data downloads, NEON Biorepository sample loans, and scientific publications. Here we describe the methods used to generate the usage metrics.

Web portal visitation

Google Analytics (GA) was used to track web portal us-

age across NEON's three public-facing websites: (1) <https://www.neonscience.org> supplies all general information about NEON, including news, events, partnerships, and data collection and management documentation; (2) <https://data.neonscience.org> is the data portal for accessing all of NEON's data products as well as data visualization tools; and (3) <https://biorepo.neonscience.org> is a Symbiota portal dedicated to the NEON Biorepository collections hosted by ASU. As metrics of web portal visitation, we tallied both the mean number of unique IP addresses of web portal visitors per month and the mean page views per month for each of the portals in 2022. As GA does not provide actual IP addresses, it is not possible to verify if users are unique across the three portals. To examine the global reach of this US infrastructure, we also assessed the distribution of unique IP addresses visiting the portals from October 1, 2022, through September 30, 2023, across the countries recognized by GA. Google Analytics - Universal Analytics was used for October 1, 2022 through a) November 30, 2022 for www.neonscience.org, b) April 30, 2023 for biorepo.neonscience.org, and c) May 31, 2023 for data.neonscience.org. Google Analytics 4 - GA4 was used for all other dates through September 30, 2023. Although users can intentionally manipulate their IP addresses (e.g., via use of a virtual private network, or VPN) to yield inaccuracies in the associated location data, this metric is the best available. Note that GA uses the International Organization for Standardization (ISO) 3166-1 standard for country names (ISO 2020).

Data downloads

Users can download NEON data from NEON's data portal (data.neonscience.org), through NEON's application programming interface, directly from storage buckets, or through various externally hosted repositories. As the external repositories use different approaches for packaging data and tools for tracking use, here we focus only on the direct NEON download options (i.e., all options except the external repositories). We use Google Cloud Storage bucket logs to determine the mean number of download requests per month, mean unique IP addresses associated with data downloads per month, and the mean volume of data

downloaded per month for Oct 1, 2022, through September 31, 2023. Data file transfers associated with known IP addresses from internal servers or Google services were filtered out.

Sample loan requests

Samples and specimens at the NEON Biorepository have been available for research and educational use since 2019. NEON Biorepository staff record all inquiries and requests for sample loans as they are received. To measure use of this unique resource, we tallied the total number of researchers who have engaged in sample use requests (i.e., have borrowed or used samples, are actively using samples, or are awaiting funding to use samples), as well as the total number of samples involved in these requests, since 2019.

Scientific publications

We compiled data on scientific publications from two sources. First, the Dimensions dataset of publications that use NEON resources (obtained on 7 September 2023; Digital Science & Research Solutions, Inc. 2018) is based on a query for either the phrase, “National Ecological Observatory Network,” or one of the digital object identifiers associated with NEON’s released data products in the abstract, acknowledgements, or references within their database. Second, the GBIF dataset of publications obtained on 28 September 2023 includes publications that reference datasets containing NEON occurrence records obtained from GBIF (and are not captured in the Dimensions data) (GBIF 2023a). Occurrence records document the presence of a taxon in a location at a specified time (Wieczorek et al. 2012). We analyzed the distribution of countries represented by the author organizations within these two datasets. We also manually tagged the publications within the Dimensions dataset with the NEON data product identifiers (IDs) included in the publication and other applicable keywords (https://www.zotero.org/groups/2665225/neon_publications). To determine the number of publications related to biodiversity science, we then tallied the number of publications in the Dimensions data that had been tagged with one of the biodiversity-relevant product IDs (see Table S1) or one of the following keywords (“phenocam”, “biomass”, “vegetation”, “biodiversity”, “tree”, “forest”, “NEON Samples”).

Results and Discussion

NEON’s contributions to biodiversity science

Standardized methods and data

First and foremost, NEON provides open and freely available data products, producing an average of 1.4 million records and over 32 TB per year (based on 2020–2022 data).

The data catalog includes 85 products that represent taxonomic or other organismal information (Table S1). These data products span NEON’s three data collection systems, observational, sensor-based, and remotely-sensed. As of 15 October 2023, published NEON data includes 13,708 unique taxonomic names, excluding the archaea, bacteria, and fungi contained within NEON’s microbial data products that lack Linnaean binomials. Associated with these data is a suite of open, standardized field sampling protocols, sensor maintenance procedures, analytical laboratory procedures, and documents that describe the data processing algorithms NEON uses (see <https://data.neonscience.org/documents> for the complete library). These documents are freely available to the global public, and NEON encourages adoption of these methods by other researchers, institutions, and networks to expand the comparability and interoperability of data around the world.

Specimens and samples

Each year, NEON collects and archives approximately 100,000 biological, genomic, and environmental samples and specimens from its 81 terrestrial and aquatic sites. These samples and specimens complement the field observations and automated measurements collected at these same field sites. As of September 2023, NEON has collected and curated more than 503,000 samples and specimens; a total of more than 3 million are expected to be collected during the Observatory’s anticipated 30-year lifespan. These samples include individually curated voucher specimens representing more than 3,600 taxa; however, the samples also include pooled individuals collected as bulk or bycatch. Thus, NEON samples actually contain many more taxa than have been identified. We estimate that the terrestrial invertebrate bycatch likely will contain 10,000 to 15,000 arthropod species and 20+ million individual specimens over the 30-year lifetime of NEON.

Community development

Throughout its construction and ongoing operations, the broader impacts of the Observatory include development and adoption of information and data management standards, development and delivery of data skills training resources (<https://www.neonscience.org/resources/learning-hub>), and supporting interoperability with data provided by partner networks. For biodiversity data, these partners include GBIF, EDI, the Barcode of Life Data System (BOLD) (Ratnasingham and Hebert 2007), the Phenocam Network (Brown et al. 2016), and the USA-National Phenology Network (Elmendorf et al. 2016). Interoperability efforts include sharing NEON data on other platforms; for example, GBIF contains nearly 1M records provided by NEON as of October 2023, with growth in these contributions planned throughout NEON’s lifespan, enabling NEON data to be discovered and used alongside the bil-

lions of other GBIF records. Interoperability efforts also include building of tools to enable harmonization of NEON data with other datasets, such as the ecocomDP data model and R package that harmonizes community ecology data from NEON, LTER, and EDI (O'Brien et al. 2021).

Regarding standards development and adoption, NEON has developed a broad and well-described internal vocabulary for ingested and published data streams, and it also employs community supported vocabularies (e.g., Darwin Core [Wieczorek et al. 2012]), Minimum Information about any (x) Sequence (MixS) (Yilmaz et al. 2011). Meta-data are structured using community schemas (e.g., Ecological Metadata Language [Jones et al. 2019], schema.org). Through the genomic standards consortium, NEON provided input to the development of the MIxS environmental package. NEON has also partnered with the Environment Ontology (Buttigieg et al. 2013) to describe land use terms and is leveraging this standard to share NEON metagenomics data with the US National Microbiome Data Collaborative (Eloe-Fadrosh et al. 2022).

Use of NEON resources to advance science

Leading indicators

Web portal visitation, data downloads, and sample use requests are considered leading indicators, or impact measures that are predictive of eventual scientific and educational outcomes that are key to the success of NEON. Tables 1, 2 report recent values for the primary leading indicators that NEON tracks. Ideally, these values could be interpreted relative to similar metrics from equivalent organizations or activities. However, the diversity of NEON data products is unparalleled within the environmental research network community, and organizations do not typically share these usage data publicly (with GBIF as a notable exception). Even in the absence of benchmarks, it is clear that there is substantial interest from the global community in NEON, with more than 47,000 unique IP addresses visiting its three portals each month (Table 1), more than 1.7 TB of data requested on average each month, and over 200 researchers engaged in sample use requests from the NEON Biorepository (Table 2). Moreover, approximately 61% of the IP addresses visiting NEON portals are affiliated with

Table 1 Usage metrics of the three web portals for accessing NEON resources Oct 2022–Sep 2023

NEON portal	Mean number of unique web portal visitors per month	Mean sessions per month
www.neonscience.org	47,152	67,235
data.neonscience.org	4,847	7,446
biorepo.neonscience.org	1,287	1,654

NEON: National Ecological Observatory Network.

countries outside of the United States (Fig. 3).

Lagging indicators

The Dimensions (Digital Science & Research Solutions, Inc. 2018) dataset includes 590 publications that use NEON resources, with increasing rates of publication observed since NEON entered its full operations phase in May 2019 (Fig. 4). The Dimensions dataset also provides citation counts for each publication, which also shows increasing rates in recent years, with 79% of publications in the dataset cited at least once and a cumulative total exceeding 17,000 citations. Within this dataset, we identified 265 publications (45%) that used at least one of NEON's biodiversity data products (Table S1), samples, or other NEON resources to address a biodiversity-relevant question. These publications are complemented by the GBIF dataset of an additional 288 publications that reference datasets that include NEON occurrence records (GBIF 2023a) and are also biodiversity research products. These counts are likely underestimates of all product use, as data users often fail to cite data properly (Vannan et al. 2020).

Research highlights

After only a few years of full operations for NEON, the resulting literature includes several impactful and different uses of NEON resources. From a biodiversity perspective, the findings generally fall under four themes. First, NEON data and samples have been used as a baseline against which to track changes. These include leveraging both the unplanned disturbances, such as hurricanes (e.g., Kenney et al. 2021) and wildfire (e.g., Choi et al. 2023), and planned disturbances, such as cattle grazing (e.g., Gaffney et al. 2021), that NEON sites have experienced to date. Second, NEON data and samples have yielded discoveries of new species (e.g., Ciugulea et al. 2019; Will and Liebherr 2022) and range expansions for known species (Gorris et al. 2021). NEON data have played a key role in advancing the emerging field of ecological forecasting, including efforts to forecast changes in species abundance and community composition through time (Averill et al. 2021; Quinn Thomas et al. 2023). NEON samples have aided both the development of machine learning methods to characterize

Table 2 Usage metrics of NEON data products and samples archived at the NEON Biorepository

Category of use and time period	Usage metric	Value
NEON data downloads per month (Oct 2022–Sep 2023)	File downloads	853,000
	Unique IP addresses/hosts	1,331
	Data download volume (TB)	1.77
Biorepository loan requests (Jan 2019–July 2023)	Number of researchers	205
	Number of samples	13,673*

NEON: National Ecological Observatory Network.
*Count does not include several thousand NEON samples loaned prior to the initiation of the Biorepository.

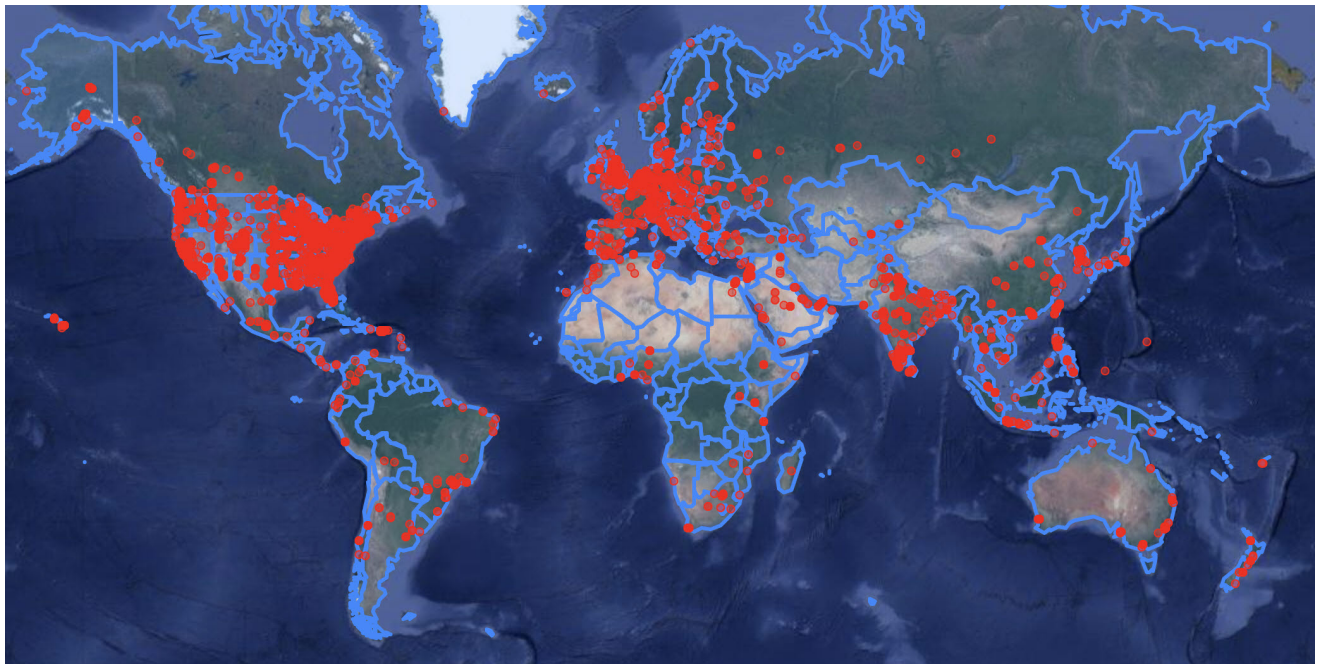


Fig. 3 Map of locations derived from IP addresses requesting data downloads from data.neonscience.org, Oct 2022–Sep 2023, using Google Analytics.

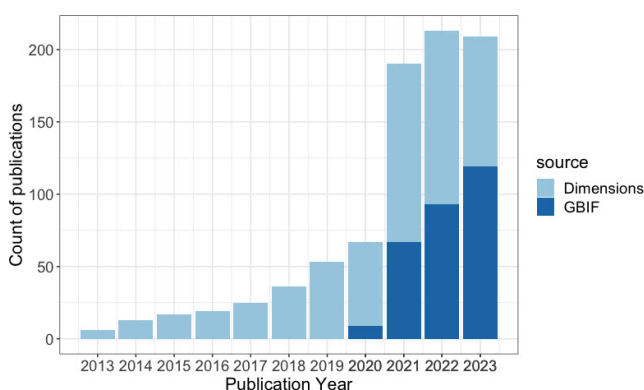


Fig. 4 Count of publications produced each year from January 2013 through September 2023 that use National Ecological Observatory Network resources, as tracked by each of two sources. GBIF: Global Biodiversity Information Facility.

biodiversity (Blair et al. 2020) and the elucidation of mosquito evolutionary history (Soghigian et al. 2023). Finally, NEON data and samples have been used to reveal large-scale patterns of species richness, sometimes with unexpected results (e.g., Carrasco et al. 2019; Weiser et al. 2022).

NEON as a global resource

NEON is fully funded by the US NSF and collects data and samples from US field sites, but its mission is global, to engage the international community to address the grand environmental challenges of our time. To that end, access to NEON data and samples is open to the global public, and NEON has active partnerships with many global networks and valuable insights are gained from international researchers who contribute to NEON's advisory groups.

NEON has forged formal partnerships with, for example, the International LTER Networks (iLTER), Integrated Carbon Observation System, Australia's Terrestrial Ecosystem Research Network, and Research Data Alliance, among others. NEON is a key member of international projects, including FLUXNET, the Global Lakes Ecological Observatory Network, GBIF, BOLD, the Forest Global Earth Observatory, Biodiversity Information Standards, the European Union's (EU) Copernicus Programme, the Omic Biodiversity Observation Network of the Group on Earth Observations (Meyer et al. 2023), and the Global Carbon Project. Partnership activities include data sharing, standards and methods co-development, research collaboration, conference support, and advisory services.

Further evidence of NEON's global relevance can be found in the portal visitation data and the author affiliations across the Dimensions and GBIF publication datasets (Digital Science & Research Solutions, Inc. 2018; GBIF 2023a). 232 countries, territories, or areas of geographical interest are represented in the user base for www.neonscience.org, for the period October 1, 2022, to September 30, 2023, with 55 countries associated with more than 1,000 users. Across the publication datasets, authors are affiliated with organizations representing 84 countries. The most represented countries among the organizations include the US (59%), the UK and Germany (4% each), and Canada and China (3% each).

In conclusion, the future health, welfare, and security of the planet depends upon diverse ecological systems that are undergoing poorly understood transitions due to climate change. NEON's development and application of

standards across diverse ecological regions and types of environmental data are unique in the field of ecology. These standardized biological and environmental data and samples enable unprecedented interoperability, which in turn enables a comprehensive and predictive understanding of ecosystem functions in relation to change. Because many of these standards are developed by and/or shared with national and international partners, NEON data are actively integrated and aggregated by a variety of resources to serve a global user base. NEON has democratized access to these environmental data through the application of FAIR data principles (Wilkinson et al. 2016), with cyberinfrastructure that provides free and open access to data and analytic tools, thus empowering a community of practice to develop open-source code, analytic tools, standardized protocols, and derived ecological data products. Because data and/or metadata are shared regularly between NEON, its global partners, and third-party aggregators, worldwide use of the data is enabled and science further accelerated.

NEON is presented here as an exemplar of the global research infrastructure needed to achieve the ambitious targets set by the global community through the GBF, a key part of the “new biodiversity science landscape centered on big data integration” (Heberling et al. 2021). However, NEON must be complemented by many more global networks that collect and provide free, open, and standardized data to achieve our shared goals. NEON has served as a founding member of the Global Ecosystem Research Infrastructure (GERI) network to facilitate the necessary collaboration and interoperability (Loescher et al. 2022). Still in its early phase of development, this group is actively working to strengthen existing connections, pilot the potential to address a global use case centered on ecological drought, and expand the reach of the network to additional countries. Such a strong global network is key to unlocking the as yet unrealized potential of NEON and the broader GERI network, with new frontiers in artificial intelligence, microbial gene discovery, synthetic biology, and sensor technologies ahead.

Supplementary Information

Supplementary information accompanies this paper at <https://doi.org/10.5141/jee.23.076>.

Table S1 NEON biodiversity data products that include taxonomic or other relevant organismal information (n = 85).

Abbreviations

BOLD: Barcode of Life Data System
EDI: Environmental Data Initiative
EU: European Union

FAIR: Findable, Accessible, Interoperable, and Reusable
GBIF: Global Biodiversity Information Facility
GBF: Global Biodiversity Framework
GERI: Global Ecosystem Research Infrastructure
ISO: International Organization for Standardization
LTER: Long-Term Ecological Research Network
MixS: Minimum information about any (x) sequence
NEON: US National Ecological Observatory Network
NSF: US National Science Foundation

Acknowledgements

The National Ecological Observatory Network is a program sponsored by the US NSF and operated under cooperative agreement by Battelle. This material is based in part upon work supported by the US NSF through the NEON Program.

Author's contributions

CML collated and analyzed data on web portal visitation and data downloads. KMT collated and analyzed data on publications and author affiliations. KMY collected and analyzed data on sample loan requests. KMT, PMM, NMF, KMY, and CML all contributed to writing the manuscript. All authors read and approved the final manuscript.

Funding

Funding for this work was provided by the US NSF Award #1724433. The NSF did not participate in any aspect of the study; the authors are solely responsible for the content.

Availability of data and materials

The datasets used and/or analyzed during the current study are available online (Digital Science & Research Solutions, Inc. 2018; GBIF 2023a) or available from the corresponding author on reasonable request.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

- Averill C, Werbin ZR, Atherton KF, Bhatnagar JM, Dietze MC. Soil microbiome predictability increases with spatial and taxonomic scale. *Nat Ecol Evol.* 2021;5(6):747–56. <https://doi.org/10.1038/s41559-021-01445-9>.
- Blair J, Weiser MD, Kaspari M, Miller M, Siler C, Marshall KE. Robust and simplified machine learning identification of pitfall trap-collected ground beetles at the continental scale. *Ecol Evol.* 2020;10(23):

- 13143-53. <https://doi.org/10.1002/ece3.6905>.
- Brown TB, Hultine KR, Steltzer H, Denny EG, Denslow MW, Granados J, et al. Using phenocams to monitor our changing Earth: toward a global phenocam network. *Front Ecol Environ*. 2016;14(2):84-93. <https://doi.org/10.1002/fee.1222>.
- Buttigieg PL, Morrison N, Smith B, Mungall CJ, Lewis SE; ENVO Consortium. The environment ontology: contextualising biological and biomedical entities. *J Biomed Semantics*. 2013;4(1):43. <https://doi.org/10.1186/2041-1480-4-43>.
- Carrasco L, Giam X, Papes M, Sheldon KS. Metrics of lidar-derived 3D vegetation structure reveal contrasting effects of horizontal and vertical forest heterogeneity on bird species richness. *Remote Sens*. 2019;11(7):743. <https://doi.org/10.3390/rs11070743>.
- Choi DH, LaRue EA, Atkins JW, Foster JR, Matthes JH, Fahey RT, et al. Short-term effects of moderate severity disturbances on forest canopy structure. *J Ecol*. 2023;111(9):1866-81. <https://doi.org/10.1111/1365-2745.14145>.
- Ciugulea I, Burroughs S, Defrancesco C, Spitale D, Charles DF, Lange-Bertalot H, et al. *Adlafia neoniana* (naviculaceae), a new diatom species from forest streams in Puerto Rico. *Plant Ecol Evol*. 2019;152(2):378-84. <https://doi.org/10.5091/plecevo.2019.1616>.
- Convention on Biological Diversity. Decision adopted by the conference of the parties to the convention on biological diversity. 15/4. Kunming-Montreal Global Biodiversity Framework. 2022. <https://www.cbd.int/doc/decisions/cop-15/cop-15-dec-04-en.pdf>. Accessed 7 Sep 2023.
- Digital Science & Research Solutions, Inc. Dimensions [Software]. 2018. <https://neon.dimensions.ai/discover/publication>. Accessed 7 Sep 2023.
- Elmendorf SC, Jones KD, Cook BI, Diez JM, Enquist CAF, Hufft RA, et al. The plant phenology monitoring design for The National Ecological Observatory Network. *Ecosphere*. 2016;7(4):e01303. <https://doi.org/10.1002/ecs2.1303>.
- Eloe-Fadrosch EA, Ahmed F, Anubhav, Babinski M, Baumes J, Borkum M, et al. The National Microbiome Data Collaborative Data Portal: an integrated multi-omics microbiome data resource. *Nucleic Acids Res*. 2022;50(D1):D828-36. <https://doi.org/10.1093/nar/gkab990>.
- Gaffney R, Augustine DJ, Kearney SP, Porensky LM. Using hyperspectral imagery to characterize rangeland vegetation composition at process-relevant scales. *Remote Sens*. 2021;13(22):4603. <https://doi.org/10.3390/rs13224603>.
- Global Biodiversity Information Facility (GBIF). What is GBIF? 2023a. <https://www.gbif.org/what-is-gbif>. Accessed 28 Sep 2023.
- Global Biodiversity Information Facility (GBIF). Research uses and citations of NEON-provided biodiversity information accessed through the Global Biodiversity Information Facility's global infrastructure. 2023b. <https://www.gbif.org/resource/search?contentType=literature&publishingOrganizationKey=e794e60e-e558-4549-99f8-cfb-241cdce24>. Accessed 28 Sep 2023.
- Gorris ME, Bartlow AW, Temple SD, Romero-Alvarez D, Shutt DP, Fair JM, et al. Updated distribution maps of predominant *Culex* mosquitoes across the Americas. *Parasit Vectors*. 2021;14(1):547. <https://doi.org/10.1186/s13071-021-05051-3>.
- Gries C, Gilbert EE, Franz NM. Symbiota - a virtual platform for creating voucher-based biodiversity information communities. *Biodivers Data J*. 2014;(2):e1114. <https://doi.org/10.3897/BDJ.2.e1114>.
- Gries C, Hanson PC, O'Brien M, Servilla M, Vanderbilt K, Waide R. The Environmental Data Initiative: connecting the past to the future through data reuse. *Ecol Evol*. 2023;13(1):e9592. <https://doi.org/10.1002/ece3.9592>.
- Heberling JM, Miller JT, Noesgaard D, Weingart SB, Schigel D. Data integration enables global biodiversity synthesis. *Proc Natl Acad Sci U S A*. 2021;118(6):e2018093118. <https://doi.org/10.1073/pnas.2018093118>.
- Hobbie JE, Carpenter SR, Grimm NB, Gosz JR, Seastedt TR. The US long term ecological research program. *BioScience*. 2003;53(1):21-32.
- International Standards Organization (ISO). ISO 3166-1:2020. 2020. <https://www.iso.org/standard/72482.html>. Accessed 15 Sep 2023.
- Jones M, O'Brien M, Mecum B, Boettiger C, Schildhauer M, Maier M, et al. Ecological Metadata Language (EML) version 2.2.0. KNB Data Repos. 2019. <https://eml.ecoinformatics.org/>. Accessed 11 Oct 2023.
- Keller M, Schimel DS, Hargrove WW, Hoffman FM. A continental strategy for the national ecological observatory network. *Front Ecol Environ*. 2008;6:282-4.
- Kenney G, Staudhammer CL, Wiesner S, Brantley S, Bigelow S, Starr G. Hurricane Michael altered the structure and function of longleaf pine woodlands. *J Geophys Res Biogeosci*. 2021;126(12):e2021JG006452. <https://doi.org/10.1029/2021JG006452>.
- Loescher HW, Vargas R, Mirtl M, Morris B, Pauw J, Yu X, et al. Building a global ecosystem research infrastructure to address global grand challenges for macrosystem ecology. *Earth's Future*. 2022;10(5):e2020EF001696. <https://doi.org/10.1029/2020EF001696>.
- Meier CL, Thibault KM, Barnett DT. Spatial and temporal sampling strategy connecting NEON Terrestrial Observation System protocols. *Ecosphere*. 2023;14(3):e4455. <https://doi.org/10.1002/ecs2.4455>.
- Meyer R, Davies N, Pitz KJ, Meyer C, Samuel R, Anderson J, et al. The founding charter of the Omic Biodiversity Observation Network (Omic BON). *GigaScience*. 2023;12:giad068. <https://doi.org/10.1093/gigascience/giad068>.
- Musinsky J, Goulden T, Wirth G, Leisso N, Krause K, Haynes M, et al. Spanning scales: the airborne spatial and temporal sampling design of the National Ecological Observatory Network. *Methods Ecol Evol*. 2022;13(9):1866-84. <https://doi.org/10.1111/2041-210X.13942>.
- National Research Council. Grand challenges in environmental sciences. Washington, D.C.: The National Academies Press; 2001. 107 p.
- O'Brien M, Smith CA, Sokol ER, Gries C, Lany N, Record S, et al. eco-comDP: a flexible data design pattern for ecological community survey data. *Ecol Inform*. 2021;64:101374. <https://doi.org/10.1016/j.ecoinf.2021.101374>.
- Parker SM, Utz RM. Temporal design for aquatic organismal sampling across the National Ecological Observatory Network. *Methods Ecol Evol*. 2022;13(9):1834-48. <https://doi.org/10.1111/2041-210X.13944>.
- Paull SH, Thibault KM, Benson AL. Tick abundance, diversity and pathogen data collected by the National Ecological Observatory Network. *Gigabyte*. 2022. <https://doi.org/10.46471/gigabyte.56>.
- Quinn Thomas R, Boettiger C, Carey CC, Dietze MC, Johnson LR, Kenney MA, et al. The NEON ecological forecasting challenge. *Front*

- Ecol Environ. 2023;21(3):112-3. <https://doi.org/10.1002/fee.2616>.
- Ratnasingham S, Hebert PD. bold: The Barcode of Life Data System (<http://www.barcodinglife.org>). Mol Ecol Notes. 2007;7(3):355-64. <https://doi.org/10.1111/j.1471-8286.2007.01678.x>.
- Schimel DS, Keller M, Berukof S, Kao B, Loescher HL, Powell H, et al. National Ecological Observatory Network 2011 science strategy: enabling continental-scale ecological forecasting. Boulder: National Ecological Observatory Network, Inc.; 2011. p. 56.
- Soghigian J, Sither C, Justi SA, Morinaga G, Cassel BK, Vitek CJ, et al. Phylogenomics reveals the history of host use in mosquitoes. Nat Commun. 2023;14(1):6252. <https://doi.org/10.1038/s41467-023-41764-y>.
- Tazik D, Franz N, Thibault KM, Yule K. National Ecological Observatory Network (NEON) biorepository sample use policy. NEON. DOC.000027 version C. 2022. <https://drive.google.com/file/d/1UwUH5LcfdN7n-CLzFvxuaAQzRxo5lZp8/view>. Accessed 11 Oct 2023.
- Thibault KM, Simkin S, Benson A. NEON breeding landbird point counts. 2023. <https://www.gbif.org/dataset/516eb7c0-3586-4dfc-bc18-c1fc1e765b7f/project>. Accessed 28 Sep 2023.
- Thorpe AS, Barnett DT, Elmendorf SC, Hinckley ELS, Hoekman D, Jones KD, et al. Introduction to the sampling designs of the National Ecological Observatory Network Terrestrial Observation System. Ecosphere. 2016;7(12):e01627. <https://doi.org/10.1002/ecs2.1627>.
- Vannan S, Downs RR, Meier W, Wilson BE, Gerasimov IV. Data sets are foundational to research. Why don't we cite them? Eos. 2020;101. <https://doi.org/10.1029/2020EO151665>.
- Weiser MD, Siler CD, Smith SN, Marshall KE, McLaughlin JF, Miller MJ, et al. Robust metagenomic evidence that local assemblage richness increases with latitude in ground-active invertebrates of North America. Oikos. 2022;2022(8):e08791. <https://doi.org/10.1111/oik.08791>.
- Wieczorek J, Bloom D, Guralnick R, Blum S, Döring M, Giovanni R, et al. Darwin Core: an evolving community-developed biodiversity data standard. PLoS One. 2012;7(1):e29715. <https://doi.org/10.1371/journal.pone.0029715>.
- Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2016;3:160018. <https://doi.org/10.1038/sdata.2016.18>. Erratum in: Sci Data. 2019;6(1):6. <https://doi.org/10.1038/s41597-019-0009-6>.
- Will K, Liebherr JK. Two new species of Mecyclothorax Sharp, 1903 (Coleoptera: Carabidae: Moriomorphini) from the Island of Hawai'i. Pan-Pac Entomol. 2022;98(1):1-17. <https://doi.org/10.3956/2022-98.1.1>.
- Yilmaz P, Kottmann R, Field D, Knight R, Cole JR, Amaral-Zettler L, et al. Minimum information about a marker gene sequence (MI-MARKS) and minimum information about any (x) sequence (MIxS) specifications. Nat Biotechnol. 2011;29(5):415-20. <https://doi.org/10.1038/nbt.1823>.