NEON SCIENCE DATA QUALITY PLAN

PREPARED BY	ORGANIZATION	DATE
Cove Sturtevant	TIS	09/26/2023
Cody Flagg	TOS	01/30/2018
Claire Lunch	DPS	09/26/2023
Elizabeth Knapp	EDU	09/26/2023
Josh Roberti	TIS	01/30/2018
Tristan Goulden	AOP	09/26/2023
Beth DeRego	Quality	09/26/2023
Janae Csavina	CVAL	09/26/2023
Nora Catolico	AQU	09/26/2023
Kaelin Cawley	AQU	09/26/2023
Kate Thibault	SCI	09/26/2023
Christine Laney	SCI	09/26/2023
Tanya Chesney	DPS	01/30/2018
Sarah Elmendorf	DPS	01/30/2018
David Garcia	SCI	10/30/2017
Nathan Leisso	AOP	01/30/2018

APPROVALS	ORGANIZATION	APPROVAL DATE
Kate Thibault	SCI	10/10/2023

RELEASED BY	ORGANIZATION	RELEASE DATE
Tanisha Waters	CM	10/10/2023

See configuration management system for approval history.

The National Ecological Observatory Network is a project solely funded by the National Science Foundation and managed under cooperative agreement by Battelle. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Change Record

REVISION	DATE	ECO #	DESCRIPTION OF CHANGE
A	08/23/2018	ECO-05580	Initial release
B	10/10/2023	ECO-07044	• Significant revisions throughout each section of the document to reflect updates to tools, processes, and systems. • Add new sections on Data Publication and Data Versioning and Revisioning. • Eliminated Appendix of future looking plans. • Updated NEON logo

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

1 DESCRIPTION

1.1 Purpose

The purpose of this Data Quality Plan is to describe current and developing Quality Assurance and Quality Control (QA/QC) processes relevant to the collection and publication of NEON data. This Data Quality Plan outlines the procedures for quality assurance and quality control activities supporting the management and dissemination of ecological data and information collected by the Observatory.

1.2 Scope

This Quality Plan applies to all NEON systems (IS, OS, AOP) producing data published to the NEON portal and API throughout Operations. The Data Quality Plan covers data workflow activities and requirements for event planning, technical training, preventative maintenance, calibration and validation, field sample and data collection, sample analyses, data quality assurance (from internal and external sources), data verification, database operations, data transitioning, and publication.

2 RELATED DOCUMENTS AND ACRONYMS

2.1 Applicable Documents

Applicable documents contain information that shall be applied in the current document. Examples are higher level requirements documents, standards, rules and regulations.

AD [01]	NEON.DOC.000001	NEON Observatory Design
AD [02]	NEON.DOC.050000	NEON Operations and Maintenance Plan
AD [03]	NEON.DOC.002651	NEON Data Product Numbering Convention
AD [04]	NEON.DOC.002652	NEON Data Products Catalog

2.2 Reference Documents

Reference documents contain information complementing, explaining, detailing, or otherwise supporting the information included in the current document.

RD [01]	NEON.DOC.000008	NEON Acronym List
RD [02]	NEON.DOC.000243	NEON Glossary of Terms
RD [03]	NEON.DOC.050005	Field Operations Job Instruction Training Plan
RD [04]	NEON.DOC.001271	AOS/TOS Protocol and Procedure: DMP – Data Management
RD [05]	NEON.DOC.004764	NEON Science Availability Plan
RD [06]	NEON.DOC.004825	NEON Algorithm Theoretical Basis Document (ATBD): OS Generic Transitions
RD [07]	NEON.DOC.001025	TOS Protocol and Procedure: PLT – Plot Establishment
RD [08]	NEON.DOC.003162	AOS Protocol and Procedure: Wadeable Stream Morphology
RD [09]	NEON.DOC.005512	Sensor Calibration, Validation System Design
RD [10]	NEON.DOC.004978	Instrumented Systems (IS) Algorithm Quality Assurance Document
RD [11]	NEON.DOC.000785	TIS Level 1 Data Products Uncertainty Budget Estimation Plan
RD [12]	NEON.DOC.011081	NEON Algorithm Theoretical Basis Document (ATBD) – QA/QC Plausibility Testing
RD [13]	NEON.DOC.000783	NEON Algorithm Theoretical Basis Document (ATBD) – Time Series Automatic Despiking for TIS Level 1 Data Products – QA/QC
RD [14]	NEON.DOC.001113	NEON Algorithm Theoretical Basis Document (ATBD) – Quality Flags and Quality Metrics for IS Data Products
RD [15]	NEON.DOC.001973	AOP Procedure: Flight Operator Training
RD [16]	NEON.DOC.001517	AOP Calibration Plan
RD [17]	NEON.DOC.004792	AOP STF Calibration Uncertainty Manual
RD [18]	NEON.DOC.004445	AOP NIS Calibration procedure
RD [19]	NEON.DOC.001290	NEON Algorithm Theoretical Basis Document (ATBD) – Imaging Spectrometer Geolocation Processing
RD [20]	NEON.DOC.001515	AOP Flight Season Management Plan
RD [21]	NEON.DOC.001980	AOP Procedure: Flight Planning
RD [22]	NEON.DOC.001984	AOP Flight Plan Boundaries Design
RD [23]	NEON.DOC.002890	AOP Level 0 Data Quality Checks
RD [24]	NEON.DOC.004652	NEON AOP Data Catalog Interface Control Document

RD [25]	NEON.DOC.004653	NEON AOP Data Processing Pipeline Database Interface Control Document
RD [26]	NEON.DOC.003652	NEON AOP Digital Camera Orthorectification Level 1 Processing Procedure
RD [27]	NEON.DOC.003314	NEON NIS Level 1 Processing Procedure
RD [28]	NEON.DOC.003315	NEON NIS Camera Model Procedure
RD [29]	NEON.DOC.003316	NEON Discrete LiDAR Level 1 Processing Procedure

2.3 External References

External references contain information pertinent to this document but are not NEON configuration-controlled. Examples include manuals, brochures, technical notes, and external websites.

ER [01]	Fulcrum. [Computer software]. Spatial Networks, Inc. (2023). https://www.fulcrumapp.com/.
ER [02]	Shiny. [Computer software]. RStudio. (2022). https://shiny.rstudio.com/.
ER [03]	ServiceNow. [Computer software]. ServiceNow. (2022). https://www.servicenow.com/.
ER [04]	International Organization for Standardization. (2015). Quality management systems—Requirements with guidance for use (ISO Standard No. 9001). https://www.iso.org/iso-9001-quality-management.html.
ER [05]	International Organization for Standardization. (2018). Guidelines for auditing management systems (ISO Standard No. 19011). https://www.iso.org/standard/70017.html.
ER [06]	Joint Committee for Guides in Metrology. (2008) 100: Evaluation of measurement data – Guide to the expression of uncertainty in measurement.
ER [07]	Taylor, J. R. (1997) An introduction to error analysis: the study of uncertainties in physical measurements. 2nd ed. Sausalito, Calif.: University Science Books.
ER [08]	Sturtevant, C., DeRego, E., Metzger, S., Ayres, E., Allen, D., Burlingame, T., Catolico, N., Cawley, K., Csavina, J., Durden, D., Florian, C., Frost, S., Gaddie, R., Knapp, E., Laney, C., Lee, R., Lenz, D., Litt, G., Luo, H., … SanClements, M. (2022). A process approach to quality management doubles NEON sensor data quality. Methods in Ecology and Evolution, 13(9), 1849–1865. https://doi.org/10.1111/2041-210X.13943.
ER [09]	Metzger, S., Ayres, E., Durden, D., Florian, C., Lee, R., Lunch, C., Luo, H., Pingintha-Durden, N., Roberti, J. A., SanClements, M., Sturtevant, C., Xu, K., & Zulueta, R. C. (2019). From NEON Field Sites to Data Portal: A Community Resource for Surface–Atmosphere Research Comes Online. Bulletin of the American Meteorological Society, 100(11), 2305–2325. https://doi.org/10.1175/BAMS-D-17-0307.1.
ER [10]	Docker. [Computer software]. Docker, Inc. (2023). https://www.docker.com/.
ER [11]	GitHub. [Computer software]. GitHub, Inc. (2023). https://github.com/.
ER [12]	Quay Container Registry. [Computer software]. Red Hat, Inc. (2023). https://quay.io/.
ER [13]	Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, A. J. G., Groth, P., Goble, C., Grethe, J. S., Heringa, J., ’t Hoen, P. A. C., et al. (2016). The Fair Guiding Principles for Scientific Data Management and Stewardship. *Scientific Data*, *3*(1), 160018. https://doi.org/10.1038/sdata.2016.18.

2.4 Acronyms

Acronym	Definition
AOP	Airborne Observation Platform
ATBD	Algorithm Theoretical Basis Document
CI-Dev	Cyberinfrastructure Development team
CLA	Collections and Laboratory Analysis
CVAL or CAL/VAL	Calibration, Validation and Audit Laboratory
DIWB	Data Ingest Workbook
FE	Full-time Field Ecologist
TFT	Temporary Field Technician
HQ	NEON Headquarters
IACUC	Institutional Animal Care and Use Committee
IS	Instrumented Systems
LFT	Lead Field Technician
NICL	NEON's Ingest Conversion Language
OS	Observation Systems
PDR	Processed Data Repository
QA/QC	Quality Assurance and Quality Control
RFP	Request for Proposals
SCI	NEON HQ Science team
SOM	IS Science Operations Management
SOP	Standard Operating Procedure
UID	Unique identifier
TCPP	Training Committee of Protocols and Procedures
LMS	Learning Management System
KBA	Knowledge Base Article (in ServiceNow software)
IPT	Integrated Product Team
PIM	Payload Integration Mount

2.5 Terminology

The use of common names for NEON software applications can vary across departments and Domains. These applications have one technically accurate name, and at times one or more "common" names describing the same item. This section aims to clarify and associate "common" names with the technical names herein.

SYNONYMOUS AND COMMON NAME(S)	NEON TECHNICAL REFERENCE NAME
Fulcrum	Mobile Data Entry Application Platform
Magpie	Field Data QC Shiny Application
the Editor	OS L0 Data Editor
the Deleter	OS L0 Data Deleter
Trouble Ticket	ServiceNow Incident Tracking System
Clockwork	Field Sampling Schedule Shiny Application
NEON database	Processed Data Repository (PDR)

3 INTRODUCTION TO THE QA/QC FRAMEWORKS AND ASSURANCE PLANS

This Data Quality Plan describes the Domain Training Program (Section 4) and the data frameworks that are currently in place. The NEON systems are grouped based on the common QA/QC frameworks that they share. The three QA/QC frameworks described in this plan are:

1. Terrestrial and Aquatic Observation Systems (OS) Framework (Section 5)
2. Terrestrial and Aquatic Instrument Systems (IS) Framework (Section 6)
3. Airborne Observational Platform (AOP) Framework (Section 7)

Each of these sections includes a description of the existing quality framework, including data QA/QC activities that occur throughout the various stages of the data lifecyle. The Plan concludes with a description of processes and approaches to Data Publication (Section 8) and Data Versioning and Revisioning (Section 9).

4 DOMAIN TRAINING QUALITY FRAMEWORK

Field staff are critical to data quality through manual sampling, analysis, sensor and infrastructure maintenance, remote monitoring, and data entry. The Field Operations Job Instruction Training Plan (RD [03]) describes the process used to design, develop, and deliver training for field staff responsible for (a) sampling as defined for the OS and (b) instrument maintenance as required for the IS. The Job Instruction Training Plan addresses the need for training temporary field technicians (TFTs) and fulltime staff (Field Ecologists (FEs) and lead field technicians (LFTs)) on protocols and procedures associated with OS and IS systems. The training program includes animal care and use modules, as reviewed and approved by Battelle's Institutional Animal Care and Use Committee (IACUC).

Training is primarily domain-based and consists of classroom presentations and practicums in the field and laboratory. Following the completion of these components, newly trained field technicians work directly with experienced field staff on implementing procedures. Each TOS training module has instructions for trainers that provides a framework for effective delivery of the curriculum. AOS trainer instructions are thematically grouped based on target media of protocols: chemical, biological, or physical. For IS, there is one overarching set of instructions for trainers document, as this system requires integrating training topics across content areas. Trainer instructions and supporting materials include specific learning objectives, preparation, time and materials required, assessment checklists, guidance on lesson development, and answers to quizzes. Annual training/retraining is required.

NEON uses a train-the-trainer model that provides both domain-based training of hundreds of temporary staff dispersed throughout the country each year and virtual or in-person cross-domain training of fulltime staff. FEs at each domain have ultimate responsibility for the proper implementation of procedures and for training LFTs and TFTs.

The major components of the Training Quality Framework include:

1. Designated Trainers
2. Training new trainers
3. Curriculum control, distribution, and updates
4. Tracking

4.1 Designated Trainers

The quality of training depends on FE knowledge of procedures as well as being an effective trainer. To train fulltime staff, NEON launched a Certified Training Program in 2019 to ensure consistent training of fulltime staff and to minimize drift in implementation of procedures across domains. The Certified Training Program defines what training is required by system and position, when and how best to deliver training (e.g., self-guided vs. instructor led), and the expertise required for each of the three levels of trainers (Peer, Protocol, and Certified, described below).

Each fulltime staff person, upon onboarding or changing roles, develops an individual Training Plan with their Domain Manager that specifies what needs to be learned and by when based on their duties. The Training Plan is a standardized document built from a template that has a comprehensive list of training topics, identifies what type of trainer is required, and recommends when the training should be provided. Training on protocols and procedures typically begins 2 weeks after the staff member's start date and is completed within 1 year. Based on the Training Plan, Domain Managers determine what type of trainer(s) is needed and then submits the Training Plan and a request for trainer(s) to a cross-departmental committee that manages the Certified Trainer Program, the Training Committee of Protocols and Procedures (TCPP).

4.1.1 Types of Trainers

Peer Trainer: Peer trainers provide training at their home domain. They are familiar with site-specific procedures, business systems and other general processes such as vehicle checks.

Protocol Trainer: Protocol trainers are recognized subject matter experts for a specific activity. They have experience providing classroom and on-the-job training for the NEON program, have completed at least one sampling season for OS topics and six months for IS topics. Protocol trainers may be fulltime or lead temporary staff. Trainings are preferentially provided in-person at the trainer's domain. Domain Managers determine who is qualified to be a protocol trainer based on meeting the minimum requirements, performance, and interest.

Certified Trainer: Certified trainers train all fulltime Field Ecologists and select Lead Field Technicians on OS/IS protocols and procedures. Certified Trainers are recognized subject matter experts on a suite of protocols, diagnostic maintenance actions, and safety procedures. In addition, they are experienced users of the NEON training curriculum and can articulate the broader scientific rationale and design of the NEON program, including how the various components integrate to support program goals. Certified Trainers lead individual or small-group trainings (2-4 trainees) depending on the need. Training is preferably provided with a hybrid model of remote sessions followed by in-person training, but training may be entirely remote and rely on the trainee creating video recordings of procedures for the trainer to review and assess competency.

The eligibility requirement to be a certified trainer is at least 2 years (24 months) of continuous fulltime employment on the NEON program for applicable OS/IS protocols and procedures. There is an annual enrollment window for all systems during which written applications are submitted and reviewed. Applicants answer questions pertaining to training experience and technical questions to demonstrate their understanding of NEON science design and their problem-solving skills. Applications are reviewed by relevant Science staff and the Curriculum Developer for appropriate training experience and understanding of science and engineering concepts. In addition, reviewers consider the applicants work over the preceding 12 months, specifically, the ability of the applicant to proactively anticipate issues and propose solutions as well as diligence in reporting and handling problems.

4.2 Training New Trainers

For each training assignment, Certified Trainers work with a Curriculum Developer to create the training schedule and strategy based on the trainee(s) Training Plan. Each assignment entails training on procedures as well as how to use the curriculum and ensure the quality of field and lab work performed by others. A critical component of all training is the performance-based assessment in which a trainer observes learners implementing procedures. The restrictions on travel imposed by COVID-19 required remote observations for cross-domain trainings, in which trainees video recorded themselves implementing procedures for the trainer to then review. A benefit of this method is that these videos can be shared through the trainers and staff to facilitate alignment on good and poor practices. Segments of the video recordings are also integrated into select online trainer quizzes in which the trainer must identify what was done incorrectly.

4.3 Trainer Maintenance and Development Activities

Each year, trainers complete the learner and trainer components of the training curriculum to ensure they are up-to-date on procedures and how to use the curriculum.

Field Ecologists/trainers also participate in several programs to reduce procedural drift and ensure that best practices are propagated throughout the network. These include:

- Annual all-domain trainings involving Field Science and Science staff in which FEs share effective methods for training, develop deeper understanding of NEON science, and collaborate on improving methods.
- Pre-season seminars in which protocol authors discuss with FEs procedural changes, including the scientific bases for the changes. Changes to the associated curriculum are also discussed.
- Biweekly system-specific (i.e., TOS, TIS, Aquatics) remote meetings with FEs, Science, and the Curriculum Developer to discuss current issues and/or challenges implementing procedures. Training is a standing agenda item during which the Curriculum Developer often provides suggestions and guidance on using specific training components or concerns.
- Office hours with Science staff, opportunities for field staff to discuss site-specific details with protocol authors (Science staff) as well as colleagues from other domains.

Another component of trainer development is self-reflection and feedback. There is a form for trainers to complete after each training that asks them to document how they used the training materials, what went well, and what they would do differently next time. Their responses are sent to themselves, as well as to their manager and the Curriculum Developer with the goal of revisiting the notes before training again on the same (or different topic) and to guide discussions on improving training effectiveness.

4.4 Curriculum Control, Distribution, and Updates

Curricular materials are initially developed by the HQ Curriculum Developer based on OS protocols and IS preventative maintenance documents. Content consists of presentations (recorded and instructor-led), instructional videos, worksheets, written and online quizzes, assessment checklists, and instructions for trainers. All content is reviewed and approved by protocol authors and/or science and engineering staff with relevant subject matter expertise prior to use by field staff.

The training materials are not included in NEON's document management system; however, a date and the protocol or procedure document version on which the training is based is included in the materials. The Curriculum Developer is responsible for maintaining current versions in the learning management system (LMS) Field staff have read-only access to all training materials in the Training Center, the online interface with the LMS. Trainers and trainees can submit feedback about training materials at any time via a standardized form in the Training Center. The Curriculum Developer is notified via email each time a form is submitted. If warranted, the issue is addressed within 48 hours and/or is addressed when the training materials are revised.

Updating the training curriculum is a collaborative effort involving the Curriculum Developer, Field Science, and Science. Training materials and reviewed and updated each time a new revision of the associated protocol/maintenance document is released. All content has a final review by the protocol author/technical subject matter expert prior to availability in the Training Center.

Occasionally, there are procedural changes that occur between document revisions. Knowledge base articles (KBAs) document these interim procedural changes. To ensure that these changes are incorporated into training materials, the Curriculum Developers are reviewers of KBAs. As soon as a KBA is published, references to it are added to the relevant training materials.

4.5 Tracking Training

The LMS tracks completion of each training component by each individual and is used to ensure that individual Training Plans are completed. For self-guided presentations and readings, the learner certifies that they have completed the required content. Instructor-led components, such as presentations or performance-based assessments, require the trainer to document completion. A custom reporting tool allows staff to view the status of training, specifically, which components are assigned, completed, or in progress and what competencies the learner has gained. Using this tool, managers can ensure that tasks are assigned to appropriately trained personnel.

Several other quality measures that the LMS enables include:

- The ability to know who trained specific staff reduces the propagation of errors. For example, if a staff member is implementing a procedure incorrectly, we can determine who trained that person and then reach out to other staff trained by the same person to ensure that the same mistakes are not being made.

- Providing immediate feedback to learners when they have an incorrect answer to an online quiz question to facilitate learning and increase the probability that, when quizzed again, they can answer the question correctly.
- Facilitation of synchronous instructor-led, cross-domain trainings that help ensure consistent implementation across different domains.
- A personal dashboard for each learner where they can see their assigned courses and associated progress. From the dashboard, learners can also easily review completed content, which is particularly helpful for procedures that are implemented episodically.
- The ability to set permissions that enable separation of learner and trainer content (**Figure 1**).

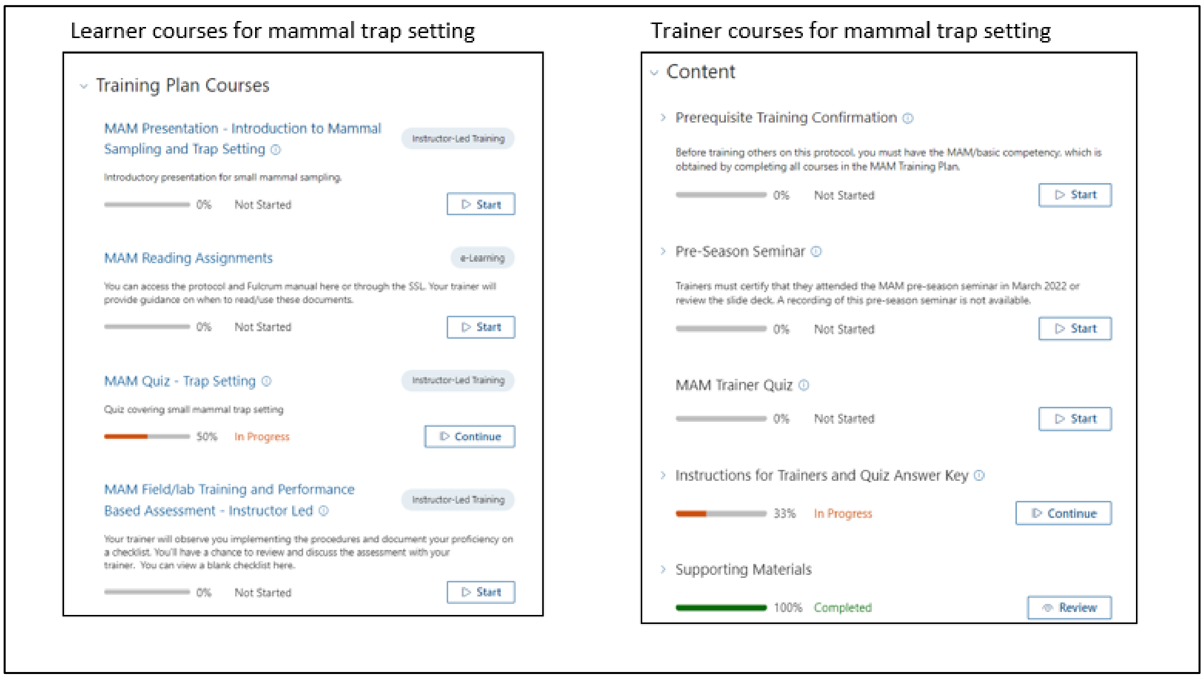

Figure 1. Example of field staff trainee (left) and trainer (right) views in the NEON Training Center.

4.6 Video Assessment Tool (VAT)

Developed during the COVID-19 pandemic when cross-domain travel was restricted, video assessment entails field staff video recording themselves implementing field or lab procedures, after which they and NEON staff with expertise in the content area (including peers, protocol authors, and trainers) review the footage. Through this process, we can quickly and efficiently uncover and address data collection and processing issues as well as best practices and share these learnings with the individuals filmed, all field staff, and our trainers.

The VAT is a structured process for identifying which procedures to target for filming and formalizes the review process and feedback mechanisms. All domains participate in the program, though assignments of what to record varies depending on site type, logistics, procedural complexity, and/or reported issues. Starting in 2023, we plan for each data collection system (i.e., AIS, AOS, TIS, and TOS) to have at least one procedure per year being included in an Observatory-level video assessment. The Observatory-Level Process involves selecting target activities and protocols based on priorities selected through evaluation and scoring of procedures by Science staff, Domain Managers, Certified Trainers, Safety personnel, and Curriculum Developers. We also plan to use this tool in individual domains when the need arises to have "eyes-on" sampling (e.g., mid-season check in). The Domain-Level Process enables access to the VAT whenever and wherever they are needed to understand issues and best practices in the field.

5 TERRESTRIAL AND AQUATIC OBSERVATIONAL SYSTEM (OS) QUALITY FRAMEWORK

The Terrestrial and Aquatic Observation Systems (OS) are reliant on detailed manual sampling of biological, chemical, and physical parameters in the NEON Domains and subsequent analysis and/or curation and archiving of samples. Consequently, the primary source of errant or non-conforming data is human error, and thus quality control in the OS systems is focused on preventing errors during data collection, entry, and ingest. The quality of OS data is highly dependent on training of field staff and assurance that OS protocols are executed consistently across the Observatory, as described in section 4. In addition to training, data quality tools have been developed to limit the amount of errant data entering the data stream and to recognize errant data prior to final publication. Processing the raw data (Level 0; L0 (AD [03])) is also an important quality process necessary to publish data products (Level 1 and higher) in a consistent format.

The major components of the OS Quality Framework (**Figure 2**) for OS data include:

1. Standardized and change-controlled sampling protocols
2. Standardized and change-controlled training materials used in domain-specific training plans
3. Vetting and auditing of external analytical laboratories
4. Data entry validation (via mobile applications developed in the Fulcrum platform (ER [01])
5. Manual quality checking and data review (performed by Field Science and Headquarters (HQ) Science staff)
6. Initial data review (via custom Shiny (ER [02]) applications, known collectively as 'Magpie')
7. Automated data ingest quality checks and validation (via the automated Cyberinfrastructure routines executed during the ingest of data, known as the 'Parser')
8. Data processing (i.e., the transitioning of data from raw to higher level data products, as described in Algorithm Theoretical Basis Documents (ATBDs; see Appendix 2))
9. Automated data QC
10. Data editing
11. Data publication
12. Data revisioning or versioning

In addition, this section briefly outlines how NEON ensures and monitors the quality of data from external laboratory services.

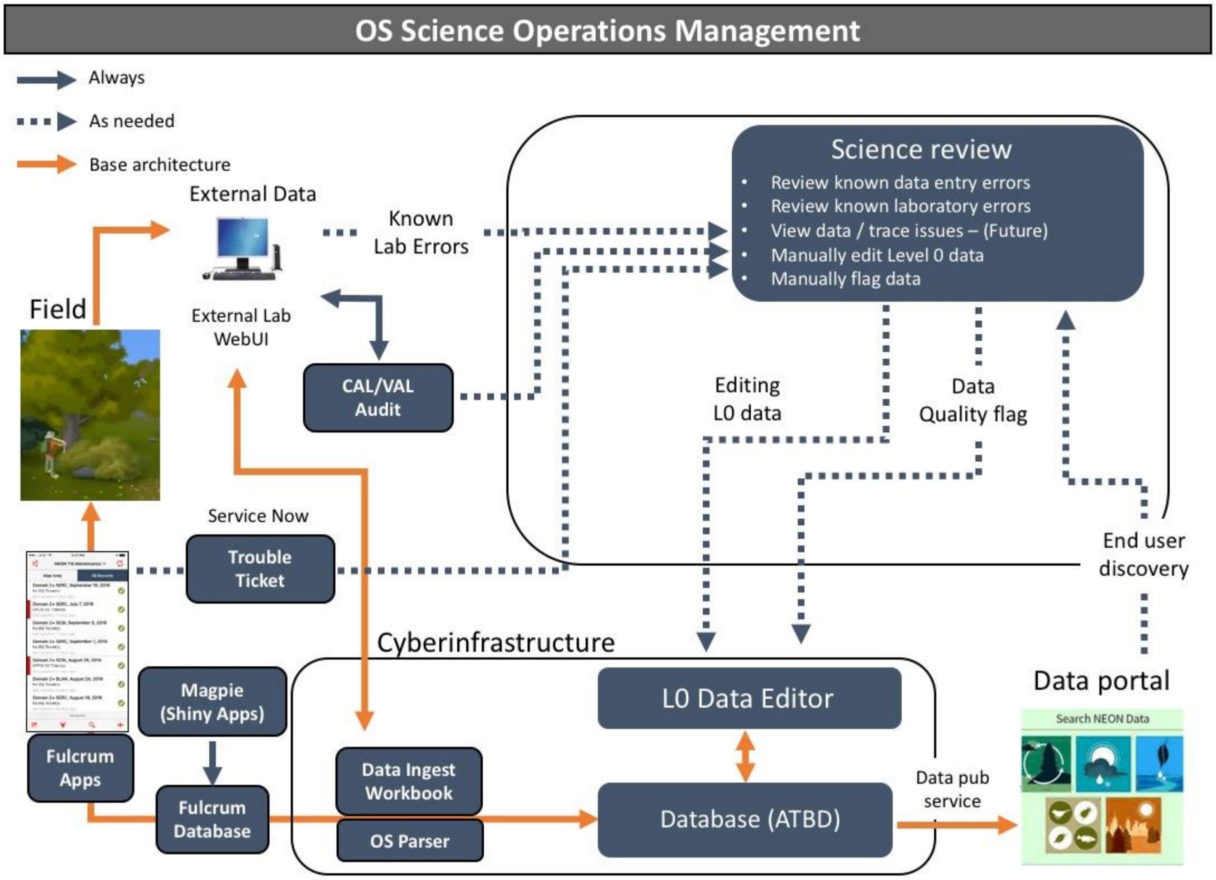

Figure 2. Overview of the OS data QA/QC framework.

5.1 OS Sampling Protocol Configuration and Change Control

OS Sampling Protocols were initially developed during the Construction phase of NEON by subject matter experts on the NEON staff working closely with community experts on the Technical Working Groups (TWGs). The protocols were implemented throughout the Observatory in a rolling fashion, as sites were constructed and preliminary operations began. During the initial years, new conditions and scenarios were encountered by NEON field staff that necessitated frequent revisions to the protocols. All protocol revisions are version controlled in NEON's document management system. However, during this construction phase, the frequency and urgency of revisions resulted in changes that were not consistently vetted with TWGs and that were sometimes implemented ahead of the official new version of the protocol being released. Most OS protocols went through 10 or more versions during this period, with all versions tracked in NEON's document management system.

In Operations, OS protocols and standard operating procedures (SOPs) follow a strict process that involves structured feedback from the Field Science staff with experience in implementation and TWG and OS Integrated Product Team (IPT) review of any notable changes. Each year, approximately one third of the 50 OS procedural documents are revised, with documents selected based on known changes to be incorporated and time since last revision. If method changes are needed between planned revisions due to field or user feedback or optimization analyses, these are captured as KBAs in the Service Now system or, if significant, are incorporated into the protocol document using the Off-Cycle Revision process. The protocol revision process includes review of and revisions to the training materials (section 4), field data entry applications, and data product ingest and publication documents, as appropriate. Each OS protocol is reviewed in full by a relevant TWG at least once every 5 years.

Field Science staff feedback on these documents is captured using a ServiceNow (ER [03]) form that can be submitted at any time. Protocol authors review and provide feedback on all submitted forms. Protocol authors also garner additional feedback from the field staff on any changes they propose to incorporate using the ServiceNow Forum tool. If a proposed change involves significant modifications to equipment, temporal or spatial sampling intensity, or adds or remove fields from the resulting data product, TWG feedback is solicited and included in the review of the changes conducted by the protocol author with the OS IPT. The OS IPT is a change-control, cross-team committee of internal stakeholders that has the authority to review and approve such changes, if the change does not incur costs to NEON Operations above a specified threshold. Changes above this threshold require additional review by the Operations IPT, which includes leadership from across the Observatory. Changes that affect the scope of NEON (i.e., number and nature of data products) also require review by the Science, Technology, and Education Advisory Committee and concurrence from the National Science Foundation.

5.2 Field Data Planning and Tracking

5.2.1 Field Collection Planning

Planning of field data collection requires identification of protocols to be implemented at each field site in each year, timing of collection (e.g., growing season, peak greenness, peak biomass), and frequency of collection (e.g., bi-weekly, every 3 weeks, every 5 years). Detailed timing and frequency requirements for field collection, lab processing, and sample shipping for analysis are in the configuration-controlled protocols discussed above. To aid in planning, estimated dates for annual onset and cessation of sampling are included in protocols and aquatic site sampling design documents. For protocols that are conducted only once every 5 years, schedules are developed by balancing vegetation and climate regimes, staff availability, and coordination with AOP overflight schedules. In addition, each year Field Science Domain Managers create detailed schedules of all AOS and TOS sampling by site by day. Schedules are reviewed by protocol authors, and, following the implementation of any requested changes, annual schedules are posted on the NEON intranet for internal reference and versioning. Subsequently, any changes to the annual or inter-annual schedules are implemented only when approved by the protocol author and OS IPT, which includes representatives from Science and Field Science.

5.2.2 Tracking and Reporting of Sample Collection Bouts

Field staff record whether or not sampling was completed in Fulcrum data entry applications (see section 5.3.3). 'Sampling impractical' records can result from conditions not meeting protocol requirements (e.g., temperature threshold for sampling, hazardous site conditions (e.g., stream flow exceeds safe wading limits), or resource shortages (e.g., insufficient personnel or equipment). Interruptions to data collection are reported as they occur (or, less frequently, when they are anticipated) via the ServiceNow internal incident reporting system to provide documentation and cross-team visibility. Domain Managers provide monthly status reports on the planned vs. completed bouts per protocol. The Shiny application known as 'Clockwork' provides sampling summary information for scheduled OS protocol activities in a standardized format that allows schedule status data to be entered efficiently, consistently, and repeatably across the Observatory. To accomplish this, Clockwork queries and summarizes the data from Fulcrum. Clockwork also enables high-level QC of OS sampling by revealing tablet sync failures and mis-entered eventIDs and plotIDs.

5.2.3 OS Issue Tracking and Resolution

Field staff report field incidents through the internal incident reporting system. Generated incidence reports typically include potential and actual non-conforming sampling events and alert HQ Science to potential data quality issues. Incident report recipients are responsible for reviewing and resolving these issues where possible. Resolutions may ultimately lead to protocol clarifications, protocol revisions, or changes to the training program. The internal incident reporting system can be queried for specific types of incidents. Analysis of incidents can help to identify systemic issues that are then tracked as Problems in the system.

5.3 Field Data Entry Validation and Review

5.3.1 Field Data Collection

Field staff capture information about the collection process as samples and data are collected. This includes all details on who, what, where, when, and how the samples and data are collected, along with pertinent observations, applicable protocol version, and sample chain of custody information. Field staff are responsible for following all sampling requirements outlined within individual protocols and corresponding Standard Operating Procedures (SOPs). Field staff record data electronically on mobile devices using Fulcrum data applications whenever possible (see 5.3.3 below), but field staff always have paper datasheets available in case of hardware failure.

5.3.2 Transcription and Storage of Field and Lab Data Recorded with Paper Datasheets

Paper datasheets are change-controlled documents made available for each relevant SOP of a sampling protocol. While technicians are required to use mobile data entry applications whenever possible, adverse field conditions or equipment failure may necessitate collecting data on paper. The OS Data

Management Protocol (RD [04]) provides field staff with instructions on how to ensure that data collected on paper are properly handled from field collection to data entry. To summarize the protocol:

a) paper datasheets are scanned and saved as a digital copy upon return from the field

b) the data are entered into a data entry application

c) the digital data are reviewed (data quality check) to ensure that typos are not introduced into the final dataset

d) the paper copies stored at the Domain Support Facility are shipped at the end of the year to NEON headquarters for long-term reference (note that this step was discontinued in 2023)

Data quality checks for manually recorded data involve reviewing a minimum of 10 or 10% (whichever is greater) of digital records against paper datasheets per sampling bout per protocol. Field staff randomly select the records to be checked and compare paper datasheet values against digital records. All data values are checked, and corrections made if there are deviations between paper and digital copies. If a single correction is made to a digital record, it is noted as an error. If the total errors detected exceeds 5% of the records checked, technicians must review an additional 10% of records to verify data quality. If the 5% threshold is exceeded again, technicians are required to review every data record.

5.3.3 Data Entered Directly into Data Entry Applications

Digital data entry applications have been created for all OS sampling protocols using the Fulcrum platform (Spatial Networks, Inc.). Fulcrum allows a developer to build and deploy a single application to many different operating systems simultaneously. Field staff can therefore collect data on diverse mobile devices, as well as enter data through a web browser. Regardless of platform, the user interface has a consistent appearance, data validation rules, and workflow. Fulcrum applications implement data quality control through several methods at the user interface level (**Table 1**) at the time and point of capture. Data are prevented from entering the system, or moving on to data publication, if any validation rules are violated.

Table 1. Fulcrum application Quality Control method examples.

Method	Reason	Example(s)
Geographic Range Constraint	Field staff must select valid, geo-referenced locations for most protocols.	For collecting soil samples, technicians may only select plotID values that have the code “sls” associated with them.
Taxonomic Value Constraint	Field staff may only select animal or plant taxa that are known and that may exist within a Domain’s geographic boundaries.	For plant observations, technicians may only select taxonID codes published in the US Department of Agriculture’s PLANT Database (plants.usda.gov) for a given Domain.

Method	Reason	Example(s)
Required Fields	Certain data fields are critical to a data product; field staff therefore cannot save and submit data unless a record includes values for these fields.	Data records are rejected if they are missing a valid sampling location, a sampling date, and/or username (i.e., who recorded the data).
List of Values (LOV) Constraint	Certain data fields have a limited set of acceptable values; field staff choose the value from a drop-down menu, rather than manually typing them, to prevent typos and ensure consistency.	The position of a plant's upper canopy, in relation to other plants, can only be described with five distinct values: "Open grown", "Full sun", "Partially shaded", "Mostly shaded", or "Full shade". Ticks must be identified as belonging to one growth stage: "larval", "nymph", or "adult".
Numeric Range Constraint	Most numeric data fields have established or reasonable minimum and maximum values. The sign of the numeric value is also important for interpretation.	Soil pH values must be positive, recorded with decimals, and in the range [0, 14]. The diameter of a tree must be greater than 0 and less than 400 centimeters.
Conditional Validation	More complex scenarios necessitate comparing data values across different fields and constraining the entry into one field based on the value entered into another field.	Male mammals can never be recorded as pregnant. The mass of a dried soil sample (i.e., after oven drying) must be less than the wet mass.

5.3.4 Data Review with the Fulcrum Data QC Application

Data entered directly onto mobile devices cannot be compared to a paper copy. With the large number of automatic, direct, and "in place" data validations made at time of entry (**Table 1**), field staff can instead focus on verifying higher-level data quality standards upon return from the field. These checks currently include verifying that data sets maintain referential integrity, the absence of duplicate records (within a table), orphan records (data that are missing a valid upstream parent sample), or "barren parent" records (parent samples missing required downstream child samples). Each set of data entry applications is accompanied by a user manual and a QA/QC checklist. QA/QC checklists provide clear data quality control guidelines for Field Science to support more consistent application of QC procedures across domains, provide concrete expectations about what level of data review is reasonable, and identify priorities for reviewing data given available resources.

Fulcrum user interfaces are not designed to check or maintain referential integrity, therefore a separate QC application, referred to as 'Magpie', has been developed on a Shiny platform (ER [07]). The Magpie application is a user interface that allows field staff to query Fulcrum data tables for the presence of duplicate data records. Duplicate records are identified by information specified in the data ingest workbook (DIWB, see section 5.7.1 below for more details), and typically includes the combination of sampling location, date, and type to create a NEON unique identifier. If field staff find duplicate records, they first determine whether the record set consists of exact matches (i.e., every data value in the record set is identical) or inexact matches (i.e., one or more data values are not identical across the duplicate record set). Exact matches can be reported for deletion before data are ingested; inexact matches can either be corrected or reported for deletion, as appropriate.

5.4 External Laboratory Quality

5.4.1 External Laboratory Contracts and Audits

External laboratories conduct numerous analyses on samples collected by NEON field staff and return the resulting data to NEON HQ. The data quality requirements for external laboratory services flow down from NEON HQ Science and are detailed in the Request for Proposals (RFPs) and corresponding procurement contracts. The Statements of Work included in the RFPs are not managed in the NEON document management system, but changes are managed according to the same process as protocol changes, including TWG and OS IPT review. The procedures used by the contracted laboratories are tracked, with version information captured in the data return and procedures provided to end users on the NEON portal.

The NEON Calibration, Validation and Audit Laboratory (CVAL) participates in the evaluation and selection of labs that bid on the RFPs. Once the selection is made, CVAL performs a pre-award document audit and, if necessary, an onsite audit to ensure the offeror's quality systems, equipment, and facilities will meet contractual needs. After awards are made, CVAL performs an annual or biennial audit of the laboratory (or more frequently, if quality comes into question). At a minimum, an audit ensures that the documentation is up to date and that data submission to NEON is proceeding according to plan. Audits can also include on-site audits or analysis audits where CVAL will send a known sample (blind, if possible) to ensure quality performance. Where external laboratory data quality does not meet data requirements, Science stakeholders are made aware to discuss impact to data and potential data flagging needs. Further, if the laboratory is unable to remedy the relevant quality issues, the contract may be terminated, and a new vendor sought.

5.4.2 Sample Tracking and External Laboratory Data Validation

Upon collection, field staff assign sample identifiers to each collected sample. Identifiers consist of a human-readable ID following a specified format, a barcode label, or both. Sample IDs are applied physically to the sample container and noted in the corresponding data record. At a minimum, these IDs are unique within each NEON protocol and Domain combination. These sample IDs are critical to maintaining a chain of custody for each sample and for linking analytical data generated in the laboratory for a given sample to its associated field metadata.

At prescribed intervals throughout the field season, field staff package and ship field-collected samples to the designated analytical or archive facility. Associated with each shipment is a shipping manifest that captures all sample IDs contained within the shipment. A hard copy of the manifest is included in the package, while a digital version is emailed to the receiving facility and NEON Collections and Laboratory Analysis (CLA) staff. Sample shipping manifests are generated via the shipping application developed in Fulcrum. Confirmation of receipt by the receiving facility is done through email and uploading of the receipt data into NEON's central database via the NEON spreadsheet upload interface. External facilities upload analytical and/or custody data on a schedule as defined in each facility's contract. Uploading of data is done through the NEON spreadsheet uploader web interface, and all uploaded data are checked by the Parser (see section 5.7.1) prior to ingest with any failures resulting in rejection by the uploader. CLA and HQ Science are responsible for monitoring for deficiencies and resolving them with the contractors.

5.5 OS Field Audits

OS field audits evaluate whether the Observatory conforms to OS requirements and scientific methods set forth in OS protocols and applicable operating procedures. The audits aim to ensure these protocols are effectively and consistently implemented and maintained throughout the Observatory. These audits take into consideration the protocol training and implementation, change management of the procedures and/or processes, domain processes, and results of previous audits, including any corrective actions.

OS field audits are conducted on-site (at the Domain Support Facilities and field sites) and/or virtually, in accordance with international quality standards ISO 9001 (ER [04]) and ISO 19011 (ER [05]). Audits are completed on a scheduled basis by qualified auditors to ensure understanding and adherence to planning, protocols, processes, and procedures. Records of audits are maintained as records by the Observatory within dedicated repositories. Nonconformities discovered result in documented corrective actions and subsequent tracking of completion of approved resolution, where applicable. Audit results are presented to auditees, Observatory Leadership and stakeholders.

5.6 Post-Collection Data Validation with the OS Parser (L0 Data)

Once the data passes through the Fulcrum database or spreadsheet uploader, it enters the Parser (see section 5.7.1 below) and is quality assured through the instructions outlined in the Data Ingest Workbook, as described in section 5.7.1 below (see also **Figure 2**), prior to ingestion into the database.

Fulcrum HQ is the custom software that pulls data from Fulcrum and submits it to the OS Parser. Fulcrum HQ queries the Fulcrum cloud database daily, and data that fall into a specific date range are automatically pulled from Fulcrum. Lag times for automatic ingest of Fulcrum data are defined per protocol to allow sufficient time for field staff to review and QC the data. Data records are pulled from Fulcrum and submitted to the Parser, which applies further validation steps, particularly validation that checks data for consistency and referential integrity with data already stored in the database. Data records that pass Parser validation and are written to the database are then locked in Fulcrum, so field staff cannot modify data that have been successfully ingested. Data records that fail at this step remain unlocked, and an alert is sent to both field and HQ staff, to notify them that data corrections are needed. Fulcrum HQ continues to attempt data ingest daily, until the records are successfully ingested and locked.

5.7 Data Ingest Validation

5.7.1 OS Parser and Data Quality Control Ingest Workbook (DIWB)

All data ingested into the OS database tables are processed through the OS Parser, whether the data originate in Fulcrum applications or spreadsheets (**Figure 2**). See section 5.3.3 above for further details about Fulcrum applications. Data in spreadsheet format are uploaded via a custom NEON web-based interface. The OS Parser is an automated software system that reads and evaluates incoming data according to rules established in the DIWB. Parser validation rules are written in a standardized syntax called NEON's Ingest Conversion Language (NICL) that is interpreted and implemented in this software system (see NEON's Document Library – data.neonscience.org/documents – for a draft document explaining the language in greater detail (https://data.neonscience.org/api/v0/documents/Nicl_Language_DRAFT). The Parser prevents data from entering the NEON database (PDR), if any rules defined by the DIWB are violated.

The DIWB is a machine-readable spreadsheet that describes all the data fields collected for a given data product and defines the Parser validation rules for each field. The spreadsheet includes information that directs the Parser on:

a) where to post the incoming data in NEON's PDR
b) measurement units and data types
c) minimum/maximum data value ranges
d) conditional data validation rules (e.g., collection date must be after setup date)
e) valid values for categorical data (a.k.a., 'list of values')
f) valid names and geographic ranges of taxonomic identifications
g) which data fields are considered the unique identifier for a sample
h) referential integrity checks (e.g., "this unique sampleID must not already exist in PDR upon ingest" or "this sampleID must have a parent primary key in a related table").

The DIWB also contains instructions for simple calculations that can be conducted on incoming data to generate data quality flags or derived values using logical or arithmetic statements. Note that calculations are only possible using individual fields within a single data record; no summaries across records are calculated using this system.

Data that enter the OS pipeline through the spreadsheet uploader or through Fulcrum are ingested by the Parser and their data values are checked against the rules and specifications described in the DIWB.

Records that pass validation move on through the ingest process to populate the Level 0 database. Records that fail validation are rejected by the Parser, preventing ingest into the Level 0 database. Rejected records generate error messages that describe what caused the failure; these messages are sent to field staff and HQ Science staff, depending on the nature of the error and the source of the data (i.e., Fulcrum vs. external laboratory). In the spreadsheet-uploader user interface, the Parser operates two checkpoints, first running all non-sample-related validations. Error messages from this checkpoint are displayed to the user in real time. The user can correct the errors and re-submit, if possible. If data pass this first checkpoint, sample-related validations are then carried out. Errors at this stage are sent to NEON staff as described above. Staff can correct data errors or seek out the appropriate contacts to make corrections; if data are determined to be correct, or workflows are at fault, data collection applications or DIWBs may be modified to allow data ingest.

5.8 Data Processing

All OS data products are transitioned (processed from incoming Level 0 values to Level 1 values posted on the public-facing data portal or API) according to the steps outlined in Algorithm Theoretical Basis Document (ATBD; RD [06]) OS Generic Transitions (https://data.neonscience.org/api/v0/documents/NEON.DOC.004825vB). The ATBD describes (a) the process by which select incoming L0 fields are verified and copied to L1 for public consumption, and metadata on geolocation and habitat attributes from definitional data tables maintained inside PDR are associated with incoming L0 data, and (b) the standardized routines for rectifying scientific names and higher taxonomically identified organisms, to include fuzzing or redaction of rare, threatened and endangered species (RT&E), as necessary.

Each datum stored in PDR is associated with a 'named location'; e.g., HARV_001.basePlot.div – following the convention of site code_plot number.plot type.sampling protocol – is one of the valid named locations for conducting plant diversity sampling (DP1.10058.001). As part of the TOS plot establishment protocol (RD [07]), the AOS Rapid Habitat Assessment and Geomorphology protocol (RD [08]) and AIS sensor installation process, structured metadata about the location are collected either via field measurements (e.g., latitude, longitude, elevation) or from geodatabases (e.g., land cover, soil types). These data are uploaded to PDR separately (usually once, though updates are possible), through an independent process from the ongoing recording of repeat field measurements of individual data streams. The transition system joins information from the two data sources to deliver a complete data and metadata package in the published data product.

5.8.1 Assigning Taxonomic information

For each type of taxonomic survey conducted by NEON (e.g., beetles, small mammals, fish), HQ science staff maintains a list of recognized names and synonyms, compiled from a variety of published or online sources (e.g., the Integrated Taxonomic Information Service - www.itis.gov). The taxon lists also include additional information on higher taxonomy (e.g., family, order). The transition system looks up taxa in incoming data streams by taxon code or scientific name, determines the currently accepted scientific name for each datum, and returns the updated name and associated higher taxonomy. For taxonomic identifications performed by NEON field staff, the end user receives only the de-synonymized names and associated higher taxonomy. For taxonomic identifications conducted by expert contractors, the end user will receive both the NEON-accepted taxonomy plus the original names and higher taxonomy applied by the taxonomist.

5.8.2 Fuzzing Taxonomy for Species of Concern

The L0->L1 processing system fuzzes (obfuscates) taxonomic identifications for species of concern (generally Federally-listed taxa) (RD [06]). Fuzzing the taxonomic identifications consists of reassigning the taxonID code and associated scientificName from an identifying taxon (e.g., *Zapus hudonius preblei*) to a non-identifying taxonID code and scientificName (e.g., *Zapus sp*.). Where all species within a genus that are found in a Domain are species of concern, NEON will fuzz taxonomy to the level of family (e.g., Dipodidae *sp.*). Note that genus sp. and family sp. taxonomic identifications are also regularly employed by technicians during regular sampling to designate individuals that were not identified to species (e.g., individuals missing pertinent morphological features required to key out the taxonomy to species). Thus, publication of a genus sp. or family sp. datum is not a definitive clue to locations of RT&E taxa. If full redaction is required for a site by species on request of the landowner (as is the case for species of concern in the Great Smoky Mountains National Park site), the entire record of that taxon is redacted from the L1 output during processing.

5.8.3 OS Publication Workbook

The OS Publication workbook serves a dual purpose for each data product:

a) to define the transition of data from L0 to L1, and
b) to define the publication of L1 data to the data portal and API.

The workbook is machine-readable, and machine-read, for both purposes. Data transition is defined by fields in the publication workbook indicating:

a) the database location of L0 data to be transitioned;
b) the type of transition for each datum: untransformed, location data, sample data, or taxonomic data;
c) fixed inputs to the transition, such as which taxonomic data to search (plants, beetles, etc.); and
d) the database location to which the L1 data be written.

5.9 OS L0 Data Editor

Editing of L0 data is a critical component of OS data quality. As noted above, quality control in the OS subsystem is focused on avoiding errors at the points of entry and ingest. However, some low rate of error is expected to persist. When errors do occur and are discovered, L0 data editing allows for correction to be made as close to the point of origin as possible. Error discovery includes both opportunistic and automated components; see Section 5.10 below for additional details on data quality review in the OS subsystem.

There are three interfaces involved in OS L0 data editing:

(1) the OS L0 Data Editor (the Editor)
(2) the OS L0 Data Deleter (the Deleter)
(3) the Sample Editor.

They serve different but complementary purposes, as described below.

5.9.1 The Editor

The Editor is a third point of entry into the OS Parser, in addition to Fulcrum and spreadsheet upload. L0 data are downloaded from the L0 Data Viewer, edited, and uploaded back to the Editor. The Editor uses the Parser to evaluate the changes that will be made to L0 data if the uploaded data are parsed, including edits to downstream data products not included in the upload. Spreadsheets of current data and pending edits are presented to the user for review and a submit button presented. If the user chooses to submit, the changes are processed by the Parser and posted to the database. This process ensures that all data edits are subject to the same rigorous validation processes as any other data.

5.9.2 The Deleter

In some cases, entire records or sets of records of OS data may need to be deleted from the database. The most common cause of this is accidental submission of duplicate data. The Parser can prevent duplicate submissions in most circumstances, but not all.

Every record of OS L0 data in the database is numbered by a Unique identifier (UID). The Deleter is a simple interface that takes an input of a list of UIDs for deletion. Like the Editor, before posting to the database, it presents the user with a spreadsheet of changes that will be made if deletes are submitted and uses the Parser to validate submissions (e.g., it will not allow deletions that would result in orphaned samples).

5.9.3 The Sample Editor

Some edits to samples are made through the Editor, because they involve edits to the relationships between samples and data on a per-record basis, potentially resulting in the generation of new samples and re-assigning data to other samples. These types of changes need to pass through the full capabilities of the Parser. However, for the unique case of changing the identifier of a sample, but nothing else about the sample, the edit is made through the Sample Editor. The Sample Editor is a user interface that allows editing of a single sample identifier at a time; editable identifiers include sample tags, barcodes, and UIDs. The Sample Editor will change the value of a sample identifier in every instance where it appears in the database, retaining the underlying sample-data relationship. If an edit will result in changes to data or sample-data relationships, the Sample Editor rejects the edit and informs the user.

5.9.4 Post-Edit Transitions

After editing L0 data, the transition to L1 data must be re-run so that edits are reflected in published data. This is referred to as reprocessing and is described below in section 9.1.

5.10 Post-hoc Quality Control

The general strategy for quality control of OS data, in addition to the front-end validation rules described above which prevent ingest of spurious values, is to craft a set of QC scripts for each data product, for more holistic evaluation of data after collection and processing. The scripts analyze three aspects of data quality: completeness, timeliness, and plausibility. They employ a combination of generic functions written for the purpose of OS data product quality review that are used by all or most data products, and custom code that is specific to an individual data product. Generic functions produce standardized summary outputs which enable aggregating reporting across products, for example the percent of records examined that passed or failed a given test. While this is advantageous, custom code is often also required to ensure confidence in the quality of OS data products.

Examples of data quality tests that fall in the three overarching categories include:

- Completeness:
  - Correct number of collection bouts per year
  - Correct number of data records per bout
  - For each upstream data record, the expected number and type of downstream data records
  - Data present in high-priority data fields
- Timeliness
  - Bouts completed during designated time windows
  - Appropriate spacing and duration of bouts
  - Samples processed within relevant time limits
- Plausibility
  - Presence of outlier values
  - Consistency in repeated measurements
  - Consistency compared to internal and external historical ranges

The goal is to implement automated execution of these scripts on a monthly or quarterly basis, depending on the frequency of data ingest, so that issues can be identified and addressed quickly. When problems are found, there are a range of possible outcomes, including: editing L0 data to fix resolvable data entry errors, adding post-hoc flagging or remarks to the data (currently a manual process, and applied only in limited circumstances), improving protocols and/or training materials, and updating data entry applications for improved front-end control. In most cases there is communication with field staff and/or contracted laboratories to understand the root cause of the data quality issues and develop solutions to prevent reoccurrence.

Post-hoc OS quality control is still evolving. A near term goal is the development of OS-wide metrics for each of the three data quality categories (completeness, timeliness, and plausibility), which will be aggregated from the data-product-specific calculations. In addition, work will focus on automating execution of the scripts and harmonizing outputs of quality tests in a centralized, cloud-based location for easy querying and reporting. In the longer term, it may be possible to implement automated data edits or flagging based on script outputs.

6 AQUATIC AND TERRESTRIAL INSTRUMENT SYSTEMS (IS) QUALITY FRAMEWORK

NEON represents one of the largest and most diverse environmental sensor networks, streaming hundreds of gigabytes (GB) of sensor data per day. The quality of IS data is highly dependent on proper functioning of the sensors and robust procedures for processing raw sensor data into data products. The major components of the Quality Framework specific to IS (**Figure 3**) include:

1. Sensor Calibration
2. Sensor Installation and Maintenance
3. Sensor Problem Tracking and Resolution
4. Data Processing
5. Continual Data Monitoring

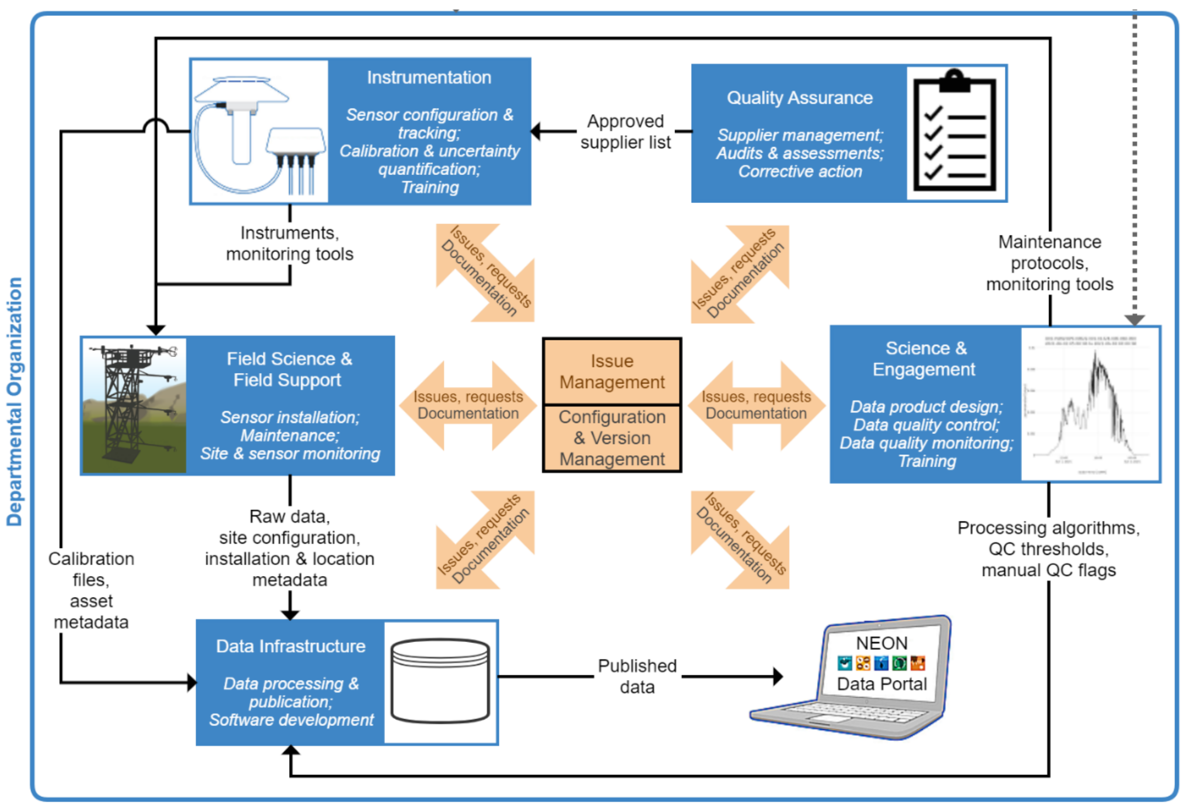

Figure 3. QA/QC framework for sensor data. Blue boxes with white text show departments and their quality-related functions. Black arrows and text depict the flow of quality-related materials among departments. Centralized issue management and configuration & version management systems are shown in light orange with brown and gray text, respectively. Wide arrows with the same color scheme depict how these systems centralize their functions. Reproduced from Sturtevant et al. (2022) (ER [08]) under CC BY 4.0 license, with the upper portion of the figure (showing the decision support hierarchy) omitted.

6.1 Sensor Calibration Program

The sensor calibration system design document (RD [09]) details the calibration's traceability to national or international standards, including both documented and physical standards. Often, the physical standards' calibrations are outsourced to certified metrology facilities unless transfer of standards are maintained or acquired by CVAL. The initial frequency of calibration for each type of sensor is documented in RD [09] based on the science requirements for sensor performance as captured in the NEON requirements database. Calibration frequencies are re-evaluated and adjusted periodically after conducting observatory-wide analyses of sensor drift and component failure rates associated with shipping and handling.

At their specified re-calibration intervals, most instruments are swapped out in the field with freshly calibrated sensors and returned to CVAL for re-calibration and re-deployment. In some instances, a sensor may return to the factory for factory calibrations, if more economical than developing the calibration capabilities in CVAL (e.g., DustTrak™, TSI, Inc.) or if a particular sensor fails CVAL validation (e.g., 3D Sonic Anemometer). Some sensors (e.g., EXO2 multisonde, YSI, Inc.) also require periodic calibrations in the field. Instrument metadata are verified upon receiving the instrument back at NEON CVAL. In other cases, instruments may stay in the field and transfer standard/s are used on-site to calibrate and validate the instruments, as per sensor-specific requirements (RD [09]). Comparisons with secondary transfer standards may be regularly scheduled or made on a case-by-case basis pending incident tracking reports. Specific Standard Operating Procedures for the calibration and operation of each type of sensor are captured in change-controlled calibration documents.

The product of calibration is a calibration file which includes information needed for data processing (as defined by ATBDs). These files are automatically incorporated in data processing. Calibration files also contain sensor and calibration metadata for traceability to standards and conditions. Finally, calibration files provide a valid date range for when the calibration can be applied to a measurement stream as defined by RD [09] for frequency of calibrations, which may be read into the data processing algorithm to generate quality flags if a sensor measurement stream is non-compliant with its required calibration frequency.

6.2 Sensor Installation and Maintenance

Knowing each sensor's location is important not only for managing NEON's assets but also for connecting sensor data to the physical location it represents and other relevant metadata needed for processing. When installing or exchanging instruments for any reason (e.g., calibration or repair), Field Science staff follow version-controlled standard operating procedures detailing installation, physical verification, and geolocation recording. Accurate metadata on a sensor's physical location on the geoid, and spatial relationship to other sensors, to various degrees of precision is critical for many key environmental analyses. For example, hydrologic analyses of groundwater flow direction and magnitude within a NEON aquatic site require centimeter-scale distances and elevation differences between pressure transducers. NEON sensors and infrastructure are installed according to Science-derived installation tolerance requirements that trace all the way back to NEON's scientific objectives (Metzger et al., 2019 (ER [09])). Robust geolocation data, including uncertainties, are collected via GPS and total station surveys and stored as properties of each configured location. As geolocations change over time, a complete history is maintained.

In addition to physical installation, each sensor is virtually installed by linking its unique identifier to that of a location in an audited relational database. The layout of the entire network is represented by a configuration-managed hierarchy of locations that represent its structure (domain -> site -> tower -> measurement level -> boom -> sensor). Each configured location is associated with properties that are used to validate and restrict sensor installation according to the site design as well as store important metadata for processing and dissemination, such as a complete history of its geolocation.

Preventative maintenance of instrument assemblies is typically performed every other week (or as prescribed). Complete instructions for performing preventative maintenance are provided within Preventative Maintenance Procedures (PMPs) specific to each instrument assembly as well as site infrastructure. These PMPs are change-controlled and provided in electronic format, making them easily accessible on field tablets while on site. Field staff are trained on PMPs and other operating procedures as part of the comprehensive training program described in Section 4. Preventive maintenance activities are documented with a Fulcrum mobile application (Spatial Networks, Inc.).

Before arriving on-site, field staff remotely view sensor values, diagnostics, and automated monitoring reports to devise a general plan of what issues need to be addressed. Depending on corrective maintenance needs (instrument failures, etc.) and environmental conditions, it is often not possible to perform the full suite of preventive maintenance activities. In this scenario, field staff use a prioritization hierarchy (**Figure 4**) to decide the most important maintenance to complete in the time available. Although corrective maintenance activities are generally prioritized over preventive ones, the prioritization logic ultimately aims to minimize the overall negative impact to NEON's scientific objectives, considering the redundancy of a measurement, risk of failure, and the scope of impact of a particular maintenance task. Corrective maintenance is reported in the central issue management system and conducted in consultation with subject matter experts across the Observatory.

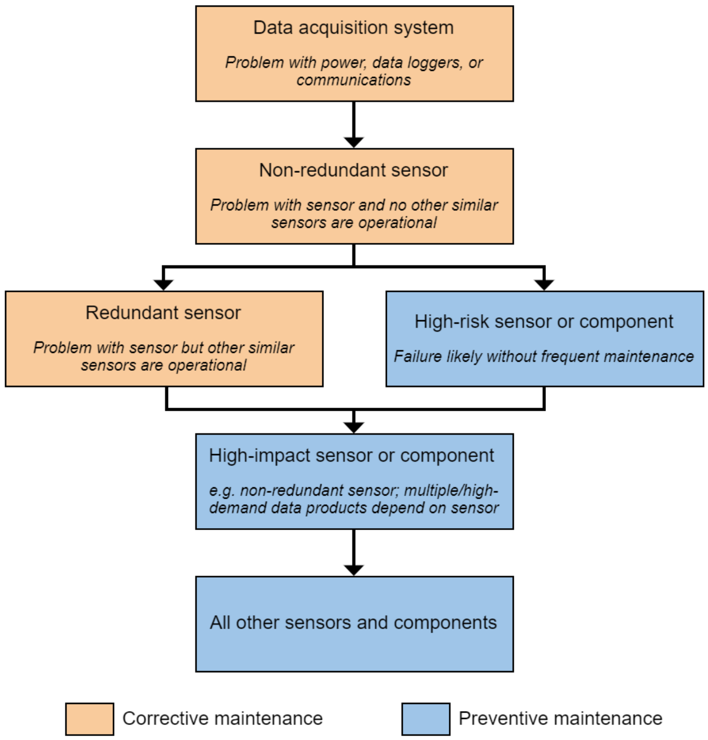

Figure 4. General prioritization hierarchy for conducting corrective and preventative maintenance with limited time. Tasks nearer to the top are higher priority. The hierarchy is filled with specific sensors according to the site design. Figure modified from Sturtevant et al. (2022) (ER [08]) under CC BY 4.0 license.

6.3 Sensor Problem Tracking and Resolution

Proper handling and timely routing of sensor health status information is important for: (1) avoidance or correction of instrument problems/errors, and (2) quality control and flagging of data products when the instrument is in an error state. While the status information used for (1) and (2) may be similar or identical, the allowable latency period is different. For the purposes of data quality control, sensor health information need only be processed into quality flags (see the *Automated Data Quality Control* section 6.4.3 below) sometime prior to publishing the data, which occurs about one month after collection of raw sensor measurements. On the other hand, fulfilling NEON's operational availability requirements necessitates near-immediate awareness of sensor malfunction so that appropriate personnel are notified, and corrective action taken to minimize sensor downtime. Two categories of incident tracking differentiate the information source/type as well as downstream handling:

1. Sensor health status warning/error
2. Field-identified measurement interference

6.3.1 Sensor Health Status Warnings and Errors

In addition to reporting scientific measurements, many sensors report their health status. Most often this is in the form of status codes that correspond to normal vs. error conditions. Alternatively, ancillary measurement streams may be used to warn of impending error status, such as when the tank pressure of a calibration gas is low. Finally, assessments of the measurements themselves may indicate poor sensor health. Sensor health status is monitored via several methods and access points in near-real-time so that appropriate action may be taken to resume normal sensor operation.

Where applicable, the location controller (LC) at each site monitors sensor health status in real time, as specified in the Command, Control, and Configuration (C3) document for each sensor assembly, to automatically attempt to improve sensor health or protect sensor and infrastructure assemblies from damage. Examples include turning on a heater to melt ice buildup or shutting down a pump when an inlet is blocked.

After data are received and stored at NEON HQ, at a typical latency of one day, error statuses are monitored via automated algorithms. These are complemented by algorithms that make more complex assessments for common quality concerns, such as a swapped install of incoming and reflected solar radiation sensors. Reports from this system are displayed in an application, referred to as Sensor Health, for human review and alerts generate daily reports emailed to field staff to take action and submit incident reports as needed (**Figure 5**). Field staff also use software to manually connect to the LC to discover other sensor issues. If error status is present, the staff either corrects the error immediately or creates an incident report for tracking and resolution.

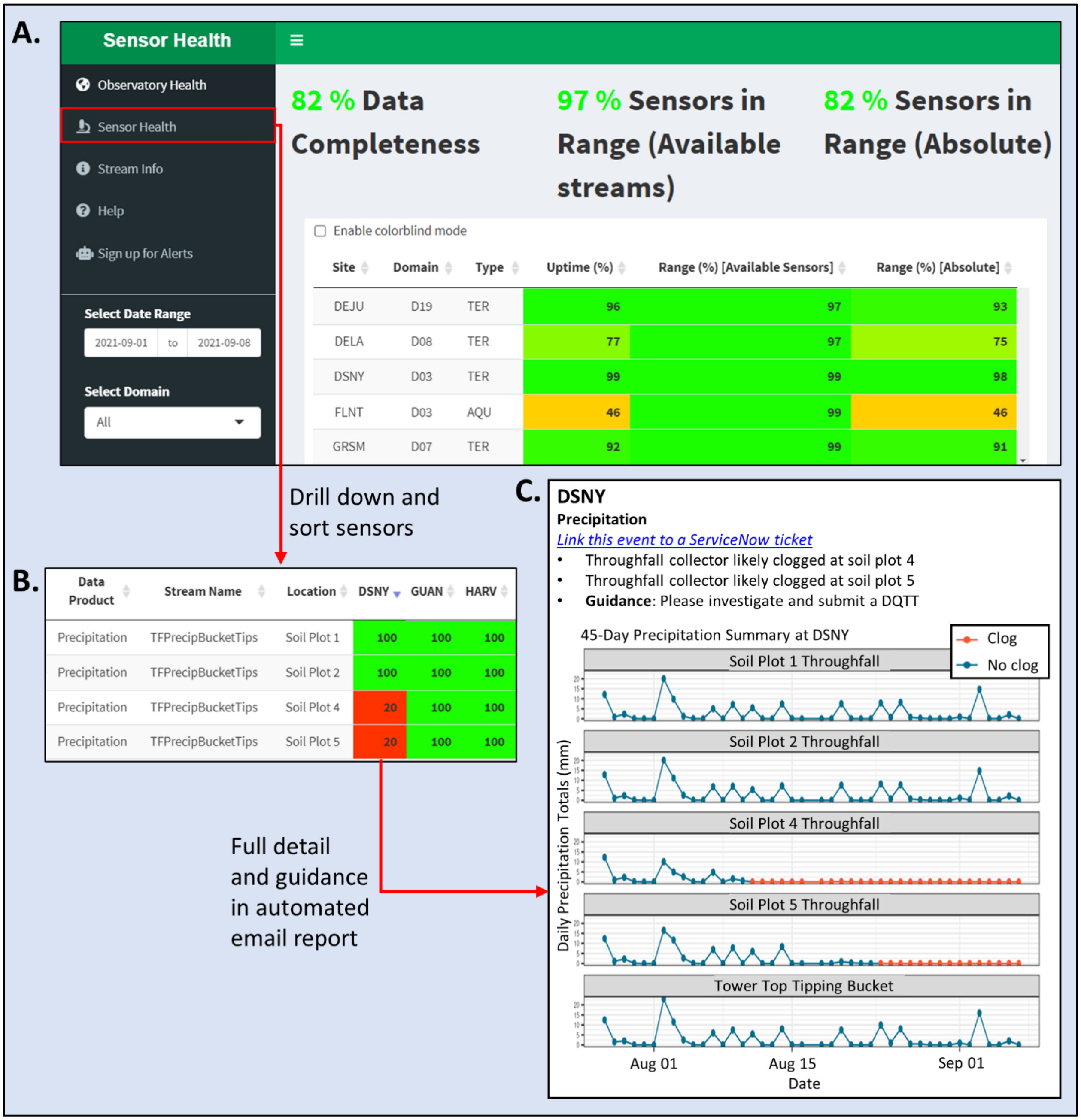

Figure 5. The NEON sensor health monitoring dashboard shows (A) color coded, high-level summaries of sensor data availability and quality. Users can drill down and sort sensors (B) to identify problematic data streams. Automated email alerts (C) provide supporting information, the ability to link to a ticket in the centralized issue management system, and pre-determined guidance for how to handle issues. Reproduced exactly from Sturtevant et al. (2022) (ER [08]) under CC BY 4.0 license.

6.3.2 Field-Identified Measurement Interference

Field staff often encounter the interference of a measurement assembly by some known or unknown cause in which the sensor continues to operate nominally but data are adversely affected for a period before, during, or after identification of the problem. For example, in early July of 2016, field staff visited the Smithsonian Environmental Research Center site and found several throughfall precipitation collectors clogged by leaves and animal nests. The sensors were operating normally but recording no rainfall events. Field staff cleared the blockages, but the data between the previous site visit and the date the blockages were cleared were of poor quality. Using the incident management system, field staff may submit a “Data Quality Trouble Ticket” (DQTT). Each DQTT typically includes the following information:

a) Site
b) Observer
c) Date
d) Issue title and summary
e) Affected sensors and their location
f) NEON asset tag (if available)
g) Photos of the problem
h) Action taken in the field

Upon receipt of a DQTT, Science staff review data relevant to the quality concern and interact with Field Science staff to discuss the concern. If warranted, the data are manually flagged (indicated as suspect; see Manual Flagging section 6.5.1 below).

6.4 Data Processing

Raw sensor data streaming from field sites to NEON headquarters requires processing to yield products useful for research. Quality assurance of data processing rests on a foundation of minimizing any transformation of raw data before it is stored centrally within data infrastructure, enabling the correction of calibration, installation, location metadata, or software errors when they are inevitably discovered. Built upon this foundation are quality assurance and control procedures implemented throughout the data processing pipeline, including algorithm validation, uncertainty estimation, automated and manual data quality control, and quality monitoring of published datasets.

6.4.1 Algorithm Quality Assurance

Data products are designed by scientists according to established scientific methods and in consultation with NEON’s scientific technical working groups and other experts. Developing and maintaining IS data products involves extensive collaboration between the Science, Engineering, Calibration & Validation

Laboratory, and Cyberinfrastructure teams. This collaboration results in a set of cross-team validated and change-controlled documents that provide the recipe for generating data products and maintaining their provenance. An overview of these documents is provided in **Table 2**; additional details of the development process can be found in RD [10].

Table 2. Overview of documents required to generate IS data products.

Deliverable	Description
Command Control and Configuration Document (C^3)	A document prescribing the configuration (e.g., output streams and reporting frequency) and command and control (e.g., heater controls) for a specific sensor type
Ingest Workbook	A machine-readable document containing the details (e.g., term names, descriptions, units) of L0 (raw) data streams received from sensors and used in processing algorithms
Calibration File	A machine-readable document containing applicable calibration coefficients and related calibration metadata for a specific sensor. These documents are traceable to a sensor's unique identifier.
Algorithm Theoretical Basis Document (ATBD)	A document detailing the algorithm used for creating NEON L1 or higher data products, including the underlying measurement theory, approximations and/or assumptions made, required data and parameter inputs, quality control methods used, and a detailed exposition of uncertainty resulting in a cumulative reported uncertainty for the data product.
Constraints	Machine-readable document providing quality control test thresholds and other parameters required by the data product algorithm.
Publication Workbook	A machine-readable document describing the contents (data streams, descriptions, units, etc.) and format of a data product as packaged for publication

Following the recipe provided by these documents, software is generated to transform the inputs into data products. To minimize errors, new or updated processing software is validated by a collection of tests that execute during the software build process and must pass prior to generating a deployable artifact, such as a Docker (ER [10]) image. Tests make use of "golden" and "tarnished" datasets to cover the spectrum of potential errors and logical scenarios that may be encountered. Static code analysis alerts developers to potential code issues, confusing or misleading code patterns, and missing or incomplete unit test code coverage. All software and deployable artifacts are version-controlled using repository managers (e.g., GitHub (ER [11]), Quay (ER [12]). A software systems architect ensures a holistic view of NEON's software systems to minimize undesired impacts of system changes.

6.4.2 Uncertainty Estimation

The objective of a measurement is to estimate the value of a particular quantity known as the measurand. Because uncertainty of measurement is inevitable, measurements should be accompanied by a statement of their uncertainty for completeness (JCGM 2008 (ER [06]); Taylor 1997 (ER [07])). Quantifying the uncertainty of IS measurements provides a measure of the reliability and applicability of individual measurements and their associated IS data products. The TIS Level 1 Data Products

Uncertainty Budget Plan (RD [11]) describes the philosophy and rationale for assuring that estimates of IS measurement uncertainties are traceable to nationally and internationally accepted standards. For all purposes, the processes by which NEON evaluates and quantifies measurement uncertainties emulates those proposed by JCGM (2008). All methods are transparent to the end-user via ATBDs provided with each data product.

6.4.3 Automated Data Quality Control

Data products undergo a suite of automated QC tests that assess and communicate the operational validity of each reported value. As described in RD [12] and RD [13] and specified in each ATBD (see Appendix 1. COMPLETE LIST OF ALGORITHM THEORETICAL BASIS DOCUMENTS (ATBDs)), these include a standard suite of plausibility tests applied at the original measurement frequency. In addition, each ATBD specifies relevant sensor-specific quality tests, such as evaluating status codes reported directly from the sensor (see Appendix 1. COMPLETE LIST OF ALGORITHM THEORETICAL BASIS DOCUMENTS (ATBDs) for a complete list of ATBDs).

QC test constraints (thresholds) are typically generated using data-driven methods, scientific judgment, and/or applicable resources from the World Meteorological Organization (WMO), National Oceanic and Atmospheric Administration (NOAA), other governing body, and information from NEON's CVAL. QC test constraints are intended to be refined over time on a site-by-site or instance-by-instance basis as the data record grows. To facilitate this evolution, constraints are organized hierarchically from realm (observatory-wide) down to a specific instance of a sensor within a site. The finest available constraint applicable to each product instance is used in processing. For example, if range constraints are not specified for air temperature at the first measurement level at site CPER, the site-level air-temperature constraint for CPER is be used.

As data are aggregated into averages (e.g., 30-minute) for publication, so too are the results of QC tests. Quality metrics (RD [14]) summarize the proportion of native-resolution values that passed, failed, or could not be evaluated over the aggregation interval of the data product (e.g., for each 30-minute average value). The quality metrics for individual quality tests are then aggregated further into a final quality flag for each output value, representing a binary "good" vs. "suspect" indicator. The final quality flag is included with the basic download package of the data product, and the results of the individual quality metrics are provided in the expanded download package. A manual override of the final quality flag is available for issues that escape automated detection (see the Manual Flagging section 6.5.1 below).

6.5 Post-hoc Quality Control

6.5.1 Manual Flagging

Subject matter experts review reports of measurement interference by field staff (see the Field-Identified Measurement Interference section 6.3.2 above) as well as any other quality concern reported through the central issue management system (including user reports submitted through the NEON website). If warranted, a dedicated science review flag is raised, which in turn raises the final quality flag. Records of science review flags are created via a custom Shiny application (referred to as 'DQ Blizzard') and stored in NEON's central database in machine readable format and are automatically applied as data are (re)published. In most circumstances, the data remain available but indicated as suspect. Under extreme circumstances data may be removed if determined to be unusable under any reasonable circumstance. Details and justification for all science review flags are provided to end users in publication metadata.

6.5.2 Published Data Monitoring

Central to an efficient and robust science operations management framework is the application of algorithms that monitor data quality as data are processed. If a high proportion of data are flagged, it could indicate that quality test thresholds need adjustment or that a sensor needs attention. Post-processing tests can also apply more complex analyses useful in detecting data that are within plausibility limits but do not reflect true environmental variation. For example, soil temperature at progressive depths should be correlated. If not, one or more temperature sensors likely need replacement.

Monitoring of published data quality occurs monthly and reports are viewable via an internal application, referred to as "DQ Blizzard". "Eyes on the data" is an important component of QA/QC, but it is not feasible to have human eyes on every data product all the time. By providing summary statistics and alerts for human review, human review may be efficiently applied.

The following is a non-exhaustive list of monthly published data monitoring checks:

- Quality flag duration/proportion
- Complete time-series check (missing time stamps)
- Consistency tests
    - Within-product (correlated profiles or systems)
    - Cross-product (different sensors measuring fundamentally similar quantities)
- Validation of science requirements
    - Data completeness and validity meet or exceed defined thresholds (RD [05])

6.6 IS Field Audits

IS field audits evaluate whether the Observatory conforms to IS requirements and scientific methods set forth in applicable operating procedures. The audits aim to ensure these procedures are effectively implemented and maintained throughout the observatory. These audits take into consideration training and implementation, control for change management of the procedure and/or processes, domain processes, results of previous audits, and corrective actions.

IS field audits are conducted on-site (at the Domain Support Facilities and field sites) and/or virtually, in accordance with international quality standards ISO 9001 (ER [04]) and ISO 19011 (ER [05]). Audits are completed on a scheduled basis by qualified auditors to ensure understanding and adherence to planning, processes, and procedures. Records of audits are maintained as records by the Observatory within dedicated repositories. Nonconformities discovered will result in corrective actions and be a part of the records for the audit. Concluding results are provided to auditees, Observatory Leadership, and stakeholders.

7 AIRBORNE OBSERVATION PLATFORM (AOP) QUALITY FRAMEWORK

The AOP consists of aircraft-mounted remote sensing instruments that will provide long-term, quantitative information on land use, vegetation structure, and biophysical and biochemical properties over the NEON sites at regional scales, in addition to supporting Principal Investigator-directed research and targets of opportunity. The remote sensing payload consists of three primary sensors integrated into a common airborne frame along with the associated support equipment required for operation and data collection. The AOP payload consists of the NEON Imaging Spectrometer (NIS), Waveform and Discrete LiDAR, and a high-resolution Digital Camera. The payload also includes two Global Positioning System receivers (GPS) and Inertial Measurement Units (IMUs) which provide the high-quality positioning and orientation required for orthorectification and the NEON Imaging Spectrometer (NIS) observation timing.

The quality of AOP data is highly dependent on proper payload operation, maintenance, and calibration. This begins in the AOP Sensor Test Facility with the maintenance and calibration of the payload and associated instrumentation. Airborne Sensor Operators are thoroughly trained on operating the instrumentation as well as flight collection decision-making in the presence of inclement weather. Since weather is a critical aspect of the overall data quality, it is imperative that the data are collected under the required weather conditions with full weather documentation recorded and included in the metadata.

Data quality continues with extraction and backup of the data from the flight disks onto a portable computer system termed the ‘Hotel Kit’. During this process, L0 data quality assurance and quality checks are performed for each sensor prior to shipment to the NEON Data Center for long-term archival storage. After ingest into the NEON Data Center, file checksums are verified, and further L1+ QA/QC checks are performed during transition of raw data to higher level products that include automatically generated reports that are emailed and reviewed by AOP Science staff. The AOP quality framework discussed in this document includes a description of the AOP training program and a description of the quality processes currently in place. These processes include L0 data quality checks, payload maintenance and calibration, data management, and L1+ QA/QC.

7.1 Training for AOP Personnel

The flight operations crew consists of the subcontracted pilot and co-pilot (currently Twin Otter International), and a team of two AOP Airborne Sensor Operators (ASO) who rotate through defined duties on a daily rotation. Each flight day one ASO is responsible for operating the NIS, while the second is responsible for the operation of the LiDAR/camera system. Both share in providing ground coordination and support at the Fixed-base operator (FBO) in addition to monitoring the GPS units for failure, providing GPS data download support and monitoring weather conditions. General pilot flight training is managed by the flight services sub-contractor. NEON relies on the flight services subcontractor to verify the qualifications, certifications, and insurance for both the pilots and the aircraft. The AOP training described in this plan refers only to the Airborne Sensor Operators (ASOs) and the instructors. We note that specifics of NEON flight requirements are communicated to subcontracted pilots consistently throughout the flight season, with particular attention paid to pilots new to the NEON program.

Airborne Sensor Operator Instructors are selected based on their experience and qualifications. The NEON Flight Operations Lead designate select ASOs as Instructors. Instructors must have completed the qualification standards prior to training others. The qualification standards are outlined in the Flight Operator Training Procedure (RD [15]) and include all the procedures to operate the aspects of the AOP payload safely and effectively.

7.1.1 ASO Classroom Training

Prior to hands-on training with the flight hardware, the ASOs receive classroom training consisting of self-guided reading material and lectures. This begins with the written procedures for payload operation and daily support activities as guided by the overall training plan (RD [15]). Training materials include a mixture of online maintained documents accessed through Sharepoint and an AOP Apache Subversion™ versioned repository of all training materials and procedures. These are synchronized with the latest official versions released through the NEON document management system but are maintained separately for ease of access while the ASOs are deployed.

AOP Science staff also train the ASOs through a series of lectures on the purpose and methodology of the nominal flight collection and instrument operations to improve ASO understanding of the overall collection philosophy and how this relates to the resulting NEON data products.

7.1.2 ASO Sensor Test Facility and Payload Installation Facility Hardware training

After classroom training, ASOs operate the full payload in the AOP Sensor Test Facility (STF) under guidance from senior ASO and STF personnel. Typically, this training is done in conjunction with the checkout of the payload from the STF to the hangar for scheduled flight operations. NIS operation training consists of, but is not limited to, monitoring and controlling the NIS environmental health, connecting components such as the external vacuum, collecting prototype data, and recording and storing metadata. LiDAR training consists of the operation of the discrete and waveform LiDAR systems, the operation of the digital camera systems, operation of the GPS/IMU system, and projecting the planned flight lines to pilots. The final component covered is the operation of the Hotel Kit, the hardware and software package that automates the extraction and L0 QA/QC of the flight data. Typically, flight disks are saved from the previous flight season to facilitate training in the extraction and QA/QC of actual flight data.

7.1.3 ASO Flight Training

The final training step consists of flight training on the AOP instrumentation. At this point, the ASO trainee should be fully versed in operating the full payload and support equipment in the lab as an integrated system. Following installation of the checked-out payload into the Twin Otter, the ASO trainees operate the payload in the aircraft. This enables them to test the full installation as well as gain an understanding of the conditions and subtleties of payload operation in the aircraft. It also enables full operation of the power system, including transfer from 'ground' power to 'survey' power, which is used for flight operations.

Flight Operations training initially consists of the ASO Trainee observing the ASO Instructor while the ASO Instructor is conducting one of the three ASO duties (NIS, lidar/camera, ground support). Each duty assignment is allotted three hands-on training stages with progressively increasing independence: 1) operation with assistance, 2) operation with minimal assistance and 3) operation with no assistance. With successful completion of all three stages, the ASO Instructor signs off on the ASO Trainee as being certified to operate the AOP Payload during flights. Additional training activities may be included if the ASO Trainee demonstrates the need for additional training to operate the instruments successfully and safely.

Annual re-certification is required for each year and is conducted by the ASO Instructors. The ASO Instructors are recertified by the Flight Operations Lead or a currently certified ASO Instructor, as they are considered the experts at conducting the required flight operations, including operating the AOP instrumentation.

7.2 Payload and Sensor Maintenance

The AOP Payload and companion Hotel Kit is the primary responsibility of the Remote Sensing Systems team. When the payload is stationed in the STF during the off-season, routine maintenance, calibration and verification of the sensors are conducted. After the payload and Hotel Kit have been acceptably calibrated and verified for functionality, the payload will be checked out by the Flight Operations team for the flight season. After completion of the flight season, the payload is checked back into the STF and the offseason maintenance, calibration and verification cycle will restart. NIS calibration is performed at least once in the off-season and may be repeated if issues are discovered in final calibration parameters. Issues in the payload are typically first identified prior to the flight season when the payload is integrated and tested, and the NIS calibration is verified. During the deployment, ASOs may support resolution of any issues with the payload or Hotel Kit that arise, but the Remote Sensing Systems team retains primary responsibility for maintaining the payload in operating condition.

7.2.1 Sensor Test Facility Off-Season Maintenance

The off-season maintenance of the AOP Payloads begins with the return of the payload to the STF. During the check-in, any issues with operating the sensors and integrated payload are noted and prioritized for repair during the off-season. After the payload is checked in and potential issues that prevent operation are repaired, the sensors are calibrated, as detailed in Section 7.3. Maintenance of the payload consists of testing and verifying its performance and replacing components that have fallen outside of specifications or are approaching end-of-life. These changes are logged and tracked to ensure components are replaced as required. Typically, STF personnel conduct all maintenance required by the NIS and Payload Support systems. The LiDAR system maintenance is primarily conducted by the vendor.

7.2.2 Issue Resolution During Deployed Flight Campaigns

Issues encountered during the flight season are reported in the daily flight logs. These logs are routinely tracked by all AOP personnel and provide the initial communication method for raising issues impacting collection. For long term tracking, issues are also recorded in NEON's incident management reporting system, as this provides a searchable record of the issue and the steps taken to resolve it. These reports may be used to identify systemic issues (i.e., problems) and address those issues between flight seasons.

Hardware issues are typically routed directly to STF personnel. Initial troubleshooting occurs with the STF personnel guiding ASOs through the relevant steps required to isolate and correct the issue. If remote trouble shooting is unsuccessful, STF personnel will be deployed along with support equipment as needed. If the payload (or sensor) is unable to be repaired while deployed, the aircraft and installed payload will be returned to the HQ hangar for additional troubleshooting. If necessary, the payload will be removed from the aircraft and returned to the STF for serious issues. The return of the payload to the STF mid-season is the option of last resort as it causes the payload to miss a large portion of the scheduled flight season.

Data quality issues identified by the L0 QA/QC checks, the L1+ QA/QC checks or by manual inspection by AOP Scientists are typically routed through the Science Subject Matter Expert (SME). The SME attempts to determine if it is a processing error or was initiated with a hardware error. Fully identifying the cause of the issue may require combined consultations between the AOP Science, STF, and Flight Operations staff. These processes are detailed below in the *AOP Data Management* section 7.5.

7.3 Sensor Calibration Program

The AOP Calibration Plan (RD [16]) defines the methodology for calibration of the sensors included on the NEON payloads and how the calibration results will be verified. The calibration plan includes the determination of the baseline calibration that is conducted in the STF. The STF maintains light sources that allow sensor traceability to the National Institute of Standards and Technology (NIST). The uncertainty of the calibration is detailed in the AOP Calibration Uncertainty Manual (RD [18]). Independently, the STF sensor calibration is verified through a series of engineering and vicarious calibration flights occurring prior to deployed flight operations, as well as the long-term verification of the sensor calibration in the field.

The Calibration Plan emphasizes a description of the methods required to determine the NIS radiometric and spectral calibration as well as further sensor characterization such as a) dynamic range and linearity, or b) signal-to-noise. The test sets used in the calibration process are described, and expected uncertainty given, when determined. Independent methods to verify the calibration determined in the laboratory calibration are discussed. Operationally, the collected imagery is also verified to ensure that the instrument is operating as expected and producing high-quality data.

The NEON Imaging Spectrometer Calibration procedure (RD [18]) describes the systematic procedure for the user to follow in creating a NIS calibration cube. It also includes a procedure for verifying the radiometric calibration through vicarious methods after installation of the payload in the aircraft.

7.3.1 Sensor Test Facility Calibration

The AOP Sensors are calibrated in the STF during the off-season after check-in of the payload. Typically, the NIS sensor is calibrated once during the off-season. NIS calibration consists of three main components, the characterization of the NIS Focal Plane Array (FPA), the spectral calibration of the NIS, and the radiometric calibration. While collection of the calibration data may occur in any order, it is important to characterize the FPA prior to processing the spectral calibration, and to complete spectral calibration before completing the radiometric calibration.

Specific Test Sets in the AOP Sensor Test Facility are utilized to collect the required calibration data. In addition, NIS On-Board Calibration (OBC) data are collected during these calibration collections on the AOP Test Sets. The OBC system is part of the overall NIS sensor package and is used to monitor changes in the sensor performance between the calibration work in the AOP Sensor Test Facility and the flight data collected during deployment. The OBC data are heavily used during the L0 QA/QC process and may also be used to dynamically adjust the NIS calibration to account for sensor changes from the lab calibration.

The AOP STF Test Sets are designed to be traceable to the National Institute of Standards and Technology (NIST) or to first principles. For the spectral calibration, emission and laser lines are used that are very well defined. For the radiometric calibration, a NIST FEL lamp is used as the primary source with an included stated uncertainty. Although not currently implemented, it is planned to leverage the uncertainty in the calibration to propagate the NIS calibration uncertainty throughout the AOP data product processing pipeline.

The LiDAR system vendor typically calibrates the LiDAR sensor. This work requires that the subframe be de-integrated from the Payload Integration Mount (PIM) frame. As this work typically takes one to six weeks depending on the LiDAR system, it is typically scheduled in December to avoid delays in integrating the payloads prior to the flight season. At the time of LiDAR calibration, the vendor also performs any required maintenance on the LiDAR and digital camera system. Once maintenance and calibration are complete, the subframe is integrated back into the PIM and full payload testing is conducted prior to payload checkout.

7.3.2 Flight Calibration

A series of calibration flights after installation of the payload into the aircraft are performed prior to each flight season. These flights are designed to verify the STF calibration, verify LiDAR performance after calibration by the vendor, and to determine calibration parameters that are not able to be determined in the lab.

A vicarious in-situ calibration flight is conducted to verify the radiometric performance of the NIS. This requires collecting reflectance data of a well characterized area, modeling the atmosphere, and predicting how much light is propagated to the NIS. This is compared to the reported radiance to verify agreement.

The LiDAR performance is verified through a planned flight over the Boulder, CO, Airport runway. High-quality GPS data has been collected across the runway such that the overall runway surface is positioned to a high accuracy (< 5 cm) in three dimensions. The LiDAR is cycled through the various collection settings to ensure vertical data quality across the range of nominal operating conditions. In addition, a flight is conducted over NEON HQ. The NEON HQ buildings have also been surveyed to a high horizontal accuracy (< 5 cm) at building corners allowing definition of the building footprint. The building edges are compared to the reported LiDAR locations of the same features. This allows an assessment of the horizontal quality of the collected data. We note that these tests do not provide a full description of the lidar system uncertainty under any collection scenario (e.g., heavy forest, highly sloped terrain), but do provide assurance the sensor is properly calibrated and installation parameters have been correctly determined.

Finally, two other flights are flown to determine the geolocation/orthorectification of the LiDAR and NIS data. The first of these is to verify the timestamps assigned to the LiDAR and NIS data. It is important to understand any difference that may stem from hardware or software differences between the sensors to ensure accurate co-registration of the independent data streams. The second flight is used to determine the roll, pitch, and yaw misalignment of the installed sensors to relative to the GPS/IMU (see RD [20]).

7.4 AOP Flight Season Management Plan

While AOP may be thought of as an instrumented system in that the payload primarily consists of remote sensing instrumentation, it also contains an aspect of observation-based systems in that specific collection times are determined by the ASOs within required schedule constraints that ensure the quality of the AOP data. Collections are constrained at multiple time scales including seasonal periods of vegetative peak greenness, daily periods of acceptable solar illumination geometry, and daily to sub-daily weather windows to minimize weather impact on the collected data.

The purpose of the Flight Season Management Plan (RD [20]) is to define the framework and associated guidelines for conducting AOP flight operations acquiring airborne remote sensing data across all NEON Domains on an annual basis through deployment of two science payloads. The Flight Season Management Plan focuses primarily on Flight Campaign Planning, including:

1) the basic criteria for defining the AOP survey area over each NEON terrestrial and aquatic site;

2) creation of flight plans and supporting documentation used during airborne surveys;
3) scheduling of the annual airborne campaign for each payload;
4) FBO selection; and
5) permitting and FAA requirements.

Flight planning is also critical to the overall quality of the collected AOP data. It is important that the planned flight lines adequately cover the survey area with enough overlap between adjacent lines to minimize coverage gaps. In addition, local topographic variations may result in large variation in coverage width and must be accounted for in the planning process. Flight plans also include instrument parameters that are specific to the landscape being observed. Ideally, the entire survey area would be collected under optimal conditions; however, constraints on available time and inclement weather often prevent all areas from being collected in ideal conditions. Therefore, AOP prioritizes sites within a NEON Domain, and areas within each NEON site. This enables the Flight Operations team to concentrate on the highest priority areas, such as tower airsheds and TOS sampling regions, while also striving to collect the entire survey area.

The purpose of the Flight Plan Procedure document (RD [21]) is to provide systematic guidance on the creation of standard plans for use with the LiDAR systems deployed on AOP flight campaigns over NEON sites, including the definition of the instrument parameters that align with the planned flight lines.

This document covers creation of flight plans for NEON core and gradient terrestrial and aquatic sites based on the flight survey area designs defined in the Flight Plan Boundaries Design document (RD [22]). It guides the prioritization of the NEON sites and survey areas within the sites (e.g., tower airshed, TOS sampling boundary). The result of the flight planning process are the flight plans for the season along with summary documents that guide the ASOs in implementation of the flight plans according to the scientific goals behind the collection philosophy.

7.5 AOP Data Management

AOP Data Management begins with the collection of the data by the Flight Operations team at a NEON collection site. The data are recorded onto a series of independent hard-drives with one hard-drive (or a series of Raided hard-drives) allocated to each AOP instrument or collection component. After the flight is completed, the flight disks are removed from the flight instrumentation and stored in Pelican cases and transported to the Hotel Kit. The Hotel Kit is used nightly throughout the flight season to automatically extract the flight data, implement the L0 QA/QC algorithms, and augment data included in the AOP flight databases. After the data have been backed-up, they are exported to an external RAID set. One RAID disk is maintained in the field while the other is shipped to the NEON Data Center where the data is ingested into the NEON database.

Once in the NEON database, data are downloaded by AOP Scientists and processed through the AOP Processing Pipeline where additional metadata and QA checks are generated. During the processing, L1+

QA/QC reports are automatically generated that are reviewed by AOP Scientists to ensure the data are of the expected quality. These reports contain information on the total coverage of a site, whether the range of values for each data product fall within expected limits, and a comparison of each data product with collections from previous years. Lidar data reports also contain a simulated total propagated uncertainty based on component errors of the sensor sub-systems. If the data are acceptable, the data are published to the Data Portal for download by external users. For each site each year, the data management process is documented from extraction through the NEON Data Center ingestion using the NEON incident and service management system to ensure all data are processed and published.

7.5.1 L0 Data Ingest and QA/QC

The L0 Data QA/QC procedures are documented in the AOP Level 0 Data Quality Checks document (RD [23]). The purpose of this document is to describe the L0 raw quality assurance and quality checks to be performed for each sensor on the NEON AOP. The quality checks determine, to the extent possible, that the raw data are within range and acceptable for producing higher-level data products. The document also guides AOP personnel through the procedure of implementing the code used to perform these checks.

In the field, the data must be extracted from the sensor systems prior to the implementation of L0 quality checks. Currently, this is performed through the Hotel Kit using custom software. The ASOs extract the L0 data and perform the quality checks in-field at the end of each day of acquisition. The extraction of L0 data and associated quality checks is ideally completed prior to the beginning of the following day's acquisition, imposing a maximum time for data extraction and execution of quality checks of 12 hrs. Any issues noted in the extraction or QA of L0 data is communicated by the ASOs to a Scientist on Duty (SOD). The SOD is a member of the AOP science team with a thorough understanding of the data QA processes and may assist with issue resolution. The SOD provides final verification of the suitability of any flights that occur within their review period. If issues arise, it is the responsibility of the SOD to properly document these issues and ensure they reach a suitable resolution before data are shipped to the NEON Data Center. If no resolution can be found, a consultation regarding a re-flight will be made with the AOP Lead, Flight Operations lead, and affected ASOs.

The AOP Level 0 Data Quality Checks document details the quality checks for each individual sensor systems including 1) the GPS / IMU, 2) the spectrometer, 3) the waveform LiDAR, 4) the digital camera, followed by a section 5) on assessment of spatial coverage for the LiDAR, camera and spectrometer. Detailed within each section are the in-range criteria imposed for success, raw data required for verification, algorithm design, and, if required, potential limitations and future improvements.

7.5.2 L1+ Data QA/QC

The AOP data are processed through the AOP Data Processing Pipeline. This automates a series of processing steps that generates the data products from all sensors. Generally, the Processing Pipeline is an automated independent chain for each sensor (GPS/IMU, discrete lidar, waveform lidar, spectrometer, camera) and requires manual processing steps for initiation of each processing chain, and some manual steps within the processing chain. The manual and automated processing steps are detailed in a series of documents (RD [25-29]). During the Pipeline processing steps, a series of L1+ QA/QC reports are generated. The reports provide information on processing decisions and quality of the resulting data. They also provide an easy mechanism for AOP Science Staff to review metrics relating to the data quality of the relevant data product. These reports are reviewed by the SME and are also maintained with the completed data and serve to inform the external user of the decisions that were made in the processing workflow.

8 DATA PUBLICATION

NEON's observational, instrumented, and airborne observational data are processed into more than 180 data products available for end users, and all data, samples, and associated documentation are delivered in a publicly discoverable and accessible manner. The Findable, Accessible, Interoperable, and Reusable (FAIR) Principles (Wilkinson et al., 2016; ER [13]) are used to guide planning and decision making for continual improvement. NEON data products and prototype datasets are freely available in the public domain (CC0 license) for the lifetime of the Observatory via its websites and its public API to support data re-use, re-distribution by others, and the production of derivatives. The only exception to this license applies to data related to rare, threatened, or endangered taxa per Government regulations. To measure published data integrity, a cryptographic checksum of all published data is stored in the databases along with the metadata.

8.1 Publication of OS and IS Data Products

The publication process is shared for all OS and IS data products, except for eddy covariance data products. The publication software extracts data from the database (PDR) and packages it for download, based on the criteria defined in the publication workbook. Publication criteria defined in the publication workbook include:

a) the database location of data to be published;

b) field names, which are used both as identifiers in the database and as column headers in published data;

c) data type, units, and measurement scale;

d) data modifications to be made in publication process, such as rounding of numeric values, the spatial and temporal resolution for publication of each data table (i.e., whether data are specific to a site or to a laboratory and how user date query should be interpreted); and

e) basic vs. expanded data package designations (i.e., which tables and fields should be provided in each package).

Data products published by this process are published in a non-proprietary, tabular format (comma-delimited). In addition to data files, the publication package includes additional metadata files providing provenance and additional quality information (**Table 3**).

Table 3. Overview of documentation included in IS and OS data product publication packages.

Terms	Description
Package metadata (EML)	Machine readable file in Ecological Metadata Format providing publication metadata, including data provider, publication date, data license, etc.
ReadMe	A .txt file that contains high level information about the data product, including sensor information, study area, etc.
Issue Log	Part of the Data Product ReadMe file. Contains high level descriptions of issues and resolutions (e.g., manual flagging for data at a given site location and/or date range).
Sensor Positions	A .csv file with geolocation history for the primary sensor source(s) of the data product (IS data products only).
Variables	A .csv file containing variable definitions, including data type and units.
Science Review Flags	A .csv file containing manual adjustments to quality flags included in the data package.

Eddy covariance products include a similar set of metadata as that shown in **Table 3**, albeit generated by a different process. Some metadata is included directly with the data in Hierarchical Data Format (HDF5) due to the complexity of the eddy covariance data product suite.

8.2 Publication of AOP Data Products

At the conclusion of the AOP processing chain, AOP scientists manually upload the data products to the long-term archival system prior to publication. As AOP scientists can manually interact with the datasets during processing or during post-processing quality assurance, checks must be implemented to ensure unwanted files are not uploaded. These files can exist because intermediate files generated in the processing chain were not deleted, or because additional files were created during investigation of possible quality issues. To ensure the fidelity of uploaded files, a schema was generated which describes the allowed filetypes for each data product and the year-site combinations and flight-day identifiers allowed for upload. After upload and publishing, the files identified for availability on the portal are checked against files that exist in the long-term archival storage system to ensure any file marked as available on the portal is accessible. Once uploaded files are confirmed to be acceptable, AOP scientists manually initiate publication of each site-month set of AOP data products. If AOP observations of a single site span multiple months, the month with the most observations is selected to represent the entire site's collections.

Published AOP data also contains several unique cases for multiple sites published as a single site. Aquatic sites that are co-located with terrestrial sites are published under the terrestrial site code. Consequently, co-located aquatic sites are not shown to be available on the data portal, as their availability is within the corresponding terrestrial site. For example, NEON aquatic sites SUGG (Lake Suggs) and BARC (Lake Barco) are located within the AOP flight area for the terrestrial site OSBS (Ordway

Swisher Biological Station). If users are searching for AOP data associated with SUGG or BARC, it will appear to be unavailable on the data portal but can found by searching for OSBS data. AOP also collects the TREE (Treehaven) terrestrial site as a contiguous flight box with the STEI (Steigerwaldt-Chequamegon) site. AOP data available for TREE and STEI are sourced from the same dataset, however, the four-letter site code listed in data product files is exclusively STEI. Therefore, when users download data for TREE, they will receive data products that cover the AOP flight box for STEI and TREE but will not see any filename labels for TREE. The STEI dataset also contains the non-contiguous Chequamegon National Forest. AOP collects the Chequamegon National Forest as a separate flight area from STEI. As a result it is processed separately and given the unique site code CHEQ, used exclusively by AOP. AOP data downloads of STEI will contain files with the CHEQ four letter identifier which will correspond to only data collected at Chequamegon National Forest. AOP also collects terrestrial sites WOOD and DCFS as a contiguous flight box and labels filenames associated with these datasets as WOOD, and KONZ and KONA as a contiguous flight box and labels filenames associated with these datasets as KONZ, similarly to the STEI/TREE case.

8.3 Communicating Data Product Changes and Issues

Long-term data management requires effective capturing, cataloging, and communicating the inevitable quality issues that arise and the changes that are made to the data, from the method of collection through the processing lifecycle. To that end, NEON data product publication packages include an issue log (**Table 3**) that captures the known issues and changes alongside the time periods and sites to which an issue applies. For issues and changes that are systematic and impact >1 data product and/or >1 site's data for a particular data product, the Data Notifications feature on the NEON website is used. Examples of systematic issues include when the data processing algorithm has been found to contain an error, when an external laboratory provides a substantial amount of potentially erroneous data, or when the COVID-19 pandemic severely limited the sampling and maintenance that could be accomplished. Examples of changes include refining of quality control thresholds for automated flagging, introducing new fields or tables into a data product, or changing methods of data collection in such a way that does not significantly compromise the comparability of data through time. Significant changes are intended to be captured through the data revision process described below (9.3).

9 DATA VERSIONING AND REVISIONING

9.1 Data Reprocessing

The NEON CI architecture enables multiple types of data reprocessing for all subsystems. Data are reprocessed when the science team or external data user has identified and corrected issues, such as errors in data entry in the OS system, corrected by the Data Editor, or errors in calibration in the IS system, corrected by updating calibration factors. Corrections are incorporated into data published on the data portal by re-running the relevant data processing. All OS and IS data files on the NEON Data Portal include a time stamp in the file name indicating the date and time of file generation; reprocessed data can be identified as such by the new time stamp.

9.2 Provisional and Released Data

Data are initially published with a Provisional status, which means that data may be updated on an as-needed basis, without guarantee of reproducibility. Until the first Data Release was published in January of 2021, all NEON data were Provisional. Provisional data allow NEON to provide data more rapidly on the Data Portal, while retaining the ability to make corrections or additions as they are identified.

After initial publication, a lag time occurs before the data are more formally Released (**Figure 6**). During this lag time, extra quality control (QC) procedures (as described in data product documentation) may be performed. This lag time also ensures all data from laboratory analyses are available before a data Release. Additionally, the user community will have had the opportunity to use the data in scientific applications and provide quality-related feedback.

Figure 6. The lifecycle of NEON data from Provisional to Released status.

Each Release consists of a complete set of data files that will not be changed further and will remain accessible throughout the lifetime of the Observatory. AOP data is an exception to this rule as the large volume of AOP files prohibits maintaining multiple copies of the datasets in the event data are updated. All AOP data products will only be tagged with the latest release or as provisional data. Each year's Release will include all data collected for each included data product up to a system-specific lag period prior to the release date. For IS, AOP, and OS data products, the respective provisional periods are 6, 6, and 12 months (**Figure 7**). Exceptions may be made for certain OS data products that need more time to obtain and publish external lab data and for other unexpected disruptions to data availability.

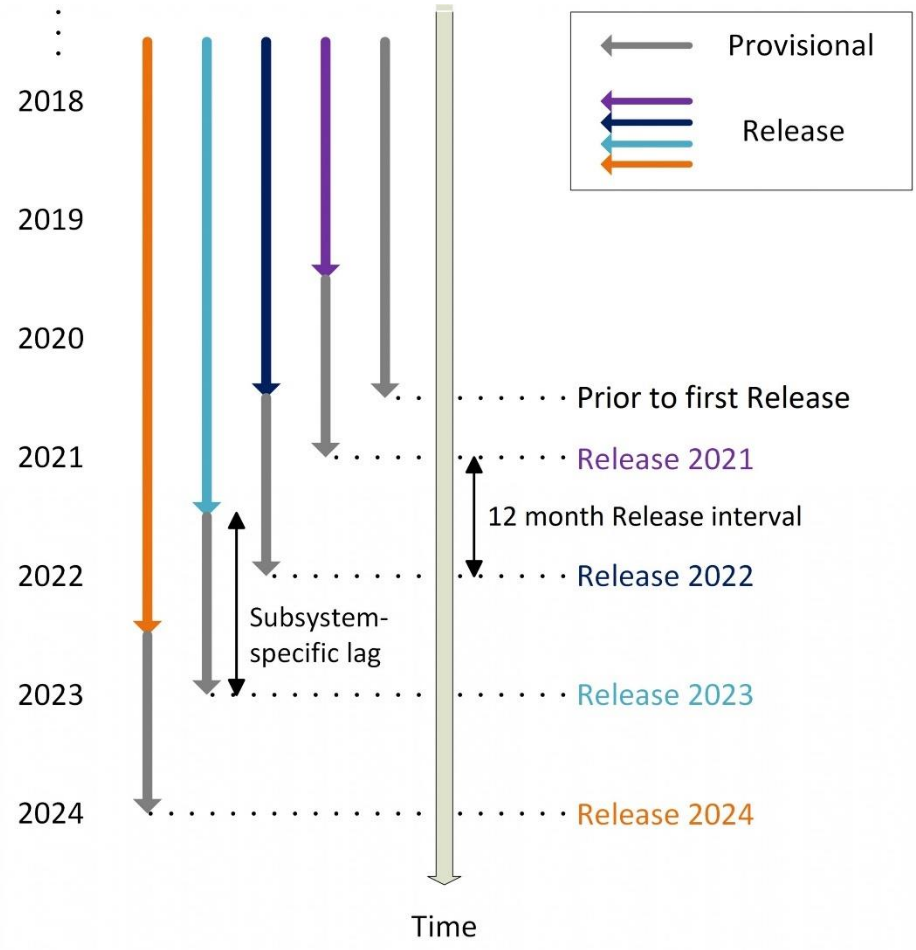

Figure 7. A projected timeline of NEON's first four Releases, with a 12-month release interval. Note that the lag time may be different among the three data processing pipelines.

Each data product within a Release is associated with a Digital Object Identifier (DOI) for reference and citation. DOI URLs will always resolve back to the dataset and are thus ideal for citing NEON data in publications and applications. Data products that are sub-products of another product and are not downloadable individually use their parent products' DOIs. Data products that are hosted fully by another repository (e.g., PhenoCam, AERONET) are not included in any release and are not assigned DOIs by NEON. Because AOP data products are associated with the latest release or as provisional data, DOIs associated with AOP data in previous releases are 'tombstoned' with each new release, rendering them unavailable after generation of the latest release.

Both Provisional and Released data have been checked to the greatest extent possible for any errors before publication and are considered fit for research. It is always possible that additional data or quality information may become available later, for both Provisional and Released data. The important difference between Provisional and Released data is that Provisional data files are subject to change at any time, without traceability, and therefore do not have a guarantee of reproducibility, while Released data files will be unchanged and past Releases will continue to be publicly accessible to users (except for AOP data products whose file sizes are prohibitively large to store multiple versions through time). Any updates or corrections to data in a Release will be reflected in a subsequent Release.

9.3 Data Product Revisions

If an instrument or protocol is significantly changed to the extent that users should be aware of potential issues with incompatibility through time, we will generate a new Revision of the data product, denoted by a change in data product identifier. Data from different revisions of the same data product are not directly comparable and should be used with caution when combining for use or analysis. Upon a data product revision, the REV field of the data product identifier will be incremented. The data product identifier takes the form DPL.PRNUM.REV, where DPL is the data product level, PRNUM is the product number, and REV is the product revision. Each data product revision will be findable in the Explore Data Product page, along with a short summary of the changes made between revisions.

APPENDIX 1. COMPLETE LIST OF ALGORITHM THEORETICAL BASIS DOCUMENTS (ATBDS)

Subsystem	Number	Document Title
AOP	NEON.DOC.001210	NEON ATBD - NEON Imaging Spectrometer Level 1B Calibrated Radiance
AOP	NEON.DOC.001211	NEON ATBD - AOP Digital Camera Image Orthorectification
AOP	NEON.DOC.001288	NEON ATBD - Imaging Spectrometer Radiance to Reflectance
AOP	NEON.DOC.001290	NEON ATBD - Imaging Spectrometer Geolocation Processing
AOP	NEON.DOC.001292	NEON ATBD - L0-to-L1 Discrete Return LiDAR
AOP	NEON.DOC.001293	NEON ATBD - L0-to-L1 Waveform Lidar
AOP	NEON.DOC.001455	NEON ATBD - Spectral Photometer
AOP	NEON.DOC.002387	NEON ATBD - Lidar Ecosystem Structure Level-2 Data Product
AOP	NEON.DOC.002390	NEON ATBD - Elevation (DTM and DSM)
AOP	NEON.DOC.002391	NEON ATBD - Vegetation Indices
AOP	NEON.DOC.003791	NEON ATBD - Elevation (Slope and Aspect)
AOP	NEON.DOC.003839	NEON ATBD - AOP Leaf Area Index
AOP	NEON.DOC.003840	NEON ATBD - AOP fPAR
AOP	NEON.DOC.004326	NEON ATBD - AOP Surface Albedo
AOP	NEON.DOC.004363	NEON ATBD - AOP Total Biomass
AOP	NEON.DOC.004364	NEON ATBD - AOP Water Indices
AOP	NEON.DOC.004365	NEON ATBD - AOP Spectrometer Mosaic
AOP	NEON.DOC.005052	NEON ATBD - AOP Digital Camera - Mosaicing
IS	NEON.DOC.000007	NEON ATBD - TIS Soil Water Content and Water Salinity
IS	NEON.DOC.000646	NEON ATBD - Single Aspirated Air Temperature
IS	NEON.DOC.000651	NEON ATBD - Atmospheric Properties and Units
IS	NEON.DOC.000652	NEON ATBD - Biological Temperature
IS	NEON.DOC.000653	NEON ATBD - Barometric Pressure
IS	NEON.DOC.000654	NEON ATBD - Triple Aspirated Air Temperature
IS	NEON.DOC.000780	NEON ATBD - 2D Wind Speed and Direction
IS	NEON.DOC.000781	NEON ATBD - Photosynthetically Active Radiation
IS	NEON.DOC.000809	NEON ATBD - Net Radiometer
IS	NEON.DOC.000810	NEON ATBD - Primary Pyranometer
IS	NEON.DOC.000813	NEON ATBD - Quantum Line Sensor
IS	NEON.DOC.000814	NEON ATBD - TIS Soil Heat Flux Plate
IS	NEON.DOC.000815	NEON ATBD - Global, Direct and Diffuse Pyranometer
IS	NEON.DOC.000816	NEON ATBD - Secondary Precipitation and Throughfall (tipping bucket)
IS	NEON.DOC.000851	NEON ATBD - Humidity and Temperature Sensor
IS	NEON.DOC.000898	NEON ATBD - Primary Precipitation (DFIR)
IS	NEON.DOC.001198	NEON ATBD - Surface Water Elevation
IS	NEON.DOC.001316	NEON ATBD - Surface Water Temperature
IS	NEON.DOC.001328	NEON ATBD - Groundwater Level, Temperature, and Specific Conductivity
IS	NEON.DOC.001571	NEON ATBD - TIS Soil Temperature
IS	NEON.DOC.001624	NEON ATBD - Homogeneity and Stationarity
IS	NEON.DOC.001789	NEON ATBD - Above Canopy and Understory/Snowpack Phenology Camera

Subsystem	Number	Document Title
IS	NEON.DOC.002181	NEON ATBD - Nitrate
IS	NEON.DOC.004388	NEON ATBD - Temperature at Specific Depths in Surface Water
IS	NEON.DOC.004571	NEON ATBD - Eddy-Covariance Data Products Composite
IS	NEON.DOC.004737	NEON ATBD - Summary Weather Statistics
IS	NEON.DOC.004738	NEON ATBD - Buoy 2D Wind Speed and Direction
IS	NEON.DOC.004931	NEON ATBD - Water Quality
IS	NEON.DOC.004968	NEON ATBD - Eddy-Covariance Storage Exchange (Profile) Assembly Raw Data
IS	NEON.DOC.011081	NEON ATBD - QA/QC Plausibility Testing
IS	NEON.DOC.011083	NEON ATBD - Soil CO2 profile
OS	NEON.DOC.004825	NEON ATBD - OS Generic Transitions