NEON SCIENCE COMMISSIONING PLAN

PREPARED BY	ORGANIZATION	DATE
Greg Wirth	SCI	02/07/2017
Mike Stewart	SE	02/07/2017

APPROVALS	ORGANIZATION	APPROVAL DATE
Kathy Kirby	DPM	06/16/2017
Rick Farnsworth	PMO	06/01/2017

RELEASED BY	ORGANIZATION	RELEASE DATE
Jennifer DeNicholas	CM	06/16/2017

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	06/16/2017	ECO-04621	Initial release

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

1 DESCRIPTION

This document, the NEON Science Commissioning Plan, summarizes the strategy for completing the science commissioning stage of construction at the National Ecological Observatory Network (NEON), a project sponsored by the National Science Foundation (NSF) and managed under cooperative agreement by Battelle.

1.1 Purpose

The NEON Science Commissioning Plan describes the design decisions and strategies governing the NEON science commissioning activities. The purposes of the commissioning stage are:

1. To demonstrate that the observatory system is performing according to system and/or subsystem requirements and the benchmarks, and
2. To demonstrate that the observatory system is achieving data quality and consistency over time and space, ultimately enabling the goal of producing continental scale science in operations.

1.2 Scope

The present document outlines those aspects of the science commissioning process applicable to all of the NEON data-generating subsystems.

- Section 2 lists the NEON documents related to this report and indicates key abbreviations.
- Section 3 defines key concepts associated with commissioning, explains the specific objectives of commissioning, defines the scope of commissioning, and relates commissioning to other phases of NEON construction.
- Section 4 describes the commissioning process, covering execution of commissioning tests, the FOCR review that concludes the commissioning process for each subsystem, and the FOOCR review that concludes the Observatory commissioning process.
- Section 5 explains the strategy behind the commissioning tests, including the attributes of these assessments, the various types of assessments, their key inputs and the scope of tests.
- Section 6 catalogs the various verification methods that can be used to confirm data quality.
- Section 7 describes the strategies that NEON will employ to allocate resources optimally during the commissioning process.
- Finally, Section 8 reviews the documentation that records the specific commissioning tests NEON will perform and the results of such tests.

This document does not describe the detailed commissioning procedures applicable to the various science subsystems; such procedures will be described in the separate subsystem commissioning plans as indicated in the list of Reference Documents (RD [03-08]).

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.004260	NEON Commissioning Strategy and Framework
AD [02]	NEON.DOC.001052	NEON Acceptance Plan
AD [03]	NEON.DOC.000004	NEON Configuration Management Plan (CMP)

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.002924	AOP Science Commissioning Plan
RD [04]	NEON.DOC.003777	AIS Science Commissioning Plan
RD [05]	NEON.DOC.003778	AOS Science Commissioning Plan
RD [06]	NEON.DOC.003779	TIS Science Commissioning Plan
RD [07]	NEON.DOC.003780	TOS Science Commissioning Plan
RD [08]	NEON.DOC.003781	MDP Science Commissioning Plan
RD [09]	NEON.DOC.004322	NEON Science Commissioning Data and Document Management Plan
RD [10]	NEON.DOC.004323	NEON Science Commissioning Anomaly Resolution Process

2.3 Acronyms

AOP	Airborne Observation Platform
AIS	Aquatic Instrument System
AOS	Aquatic Observation System
TIS	Terrestrial Instrument System
TOS	Terrestrial Observation System
MDP	Mobile Deployment Platform
FOCR	Final Operational Capabilities Review
FOOCR	Final Observatory Operational Capabilities Review
SCA	Science Commissioning Archive

3 COMMISSIONING OVERVIEW

3.1 Definitions

The following terms pertaining to the commissioning process merit definition:

- *Commissioning.* The process of bringing the NEON system (or a given subsystem) to a fully functional state.
- *Assessment.* A well-defined test that determines whether a given subsystem is operating at a prescribed level of performance.
- *Requirement.* A formally-adopted specification that a given NEON component (or system or subsystem) must satisfy in order for NEON to be considered fully functional. These are typically maintained in a formal database controlled by the NEON Project Systems Engineering team.
- *Criterion.* A specification that a given NEON component (or system, or subsystem) must satisfy in order to earn a passing mark on a commissioning assessment. Typically based on or identical to an existing requirement, a criterion may be defined for commissioning purposes when a requirement does not exist or has not been written in a verifiable form.

3.2 Objectives

The specific objectives of the commissioning process are as follows:

- Demonstrate, through the use of verifiable, quantitative criteria, the ability of each NEON data-generating subsystem to deliver quality data consistently over time and space.
- Confirm the scientific validity of the data published on the NEON data portal.
- Identify and mitigate issues that affect NEON's ability to gather and disseminate quality data.
- Establish an initial performance baseline for the NEON subsystems against which future performance can be compared.

These objectives drive the corresponding commissioning activities that we describe below.

3.3 Scope

The NEON Science Commissioning effort encompasses:

- the design and execution of a set of tests for each of NEON's data-generating subsystems addressing process quality, data quality, and (when applicable) system performance;
- a Final Operational Capabilities Review process to assess the results of all commissioning tests across all sites (or payloads) associated with a NEON subsystem; and,
- a Final Observatory Operational Capabilities Review process to assess whether the NEON observatory is complete and ready to transition from construction to operational status.

3.4 Relation to other NEON construction phases

Science commissioning represents the final major stage of NEON construction at a given site or payload. Figure 1 illustrates the idealized stages of project completion. As indicated on the left side of the figure, NEON's design phase generated increasingly detailed requirements on the project, beginning with an initial conceptual design stage that generated the ecological "Grand Challenge" questions NEON intends to address. The Grand Challenge questions led to the observatory's top-level science requirements, depicted here as "measurement objectives" that dictate what data NEON must collect in order to address the Grand Challenges. These measurement objectives drive corresponding system requirements indicating what instrumentation and procedures NEON will use to generate the data. Finally, detailed system design plans describe how NEON will be constructed.

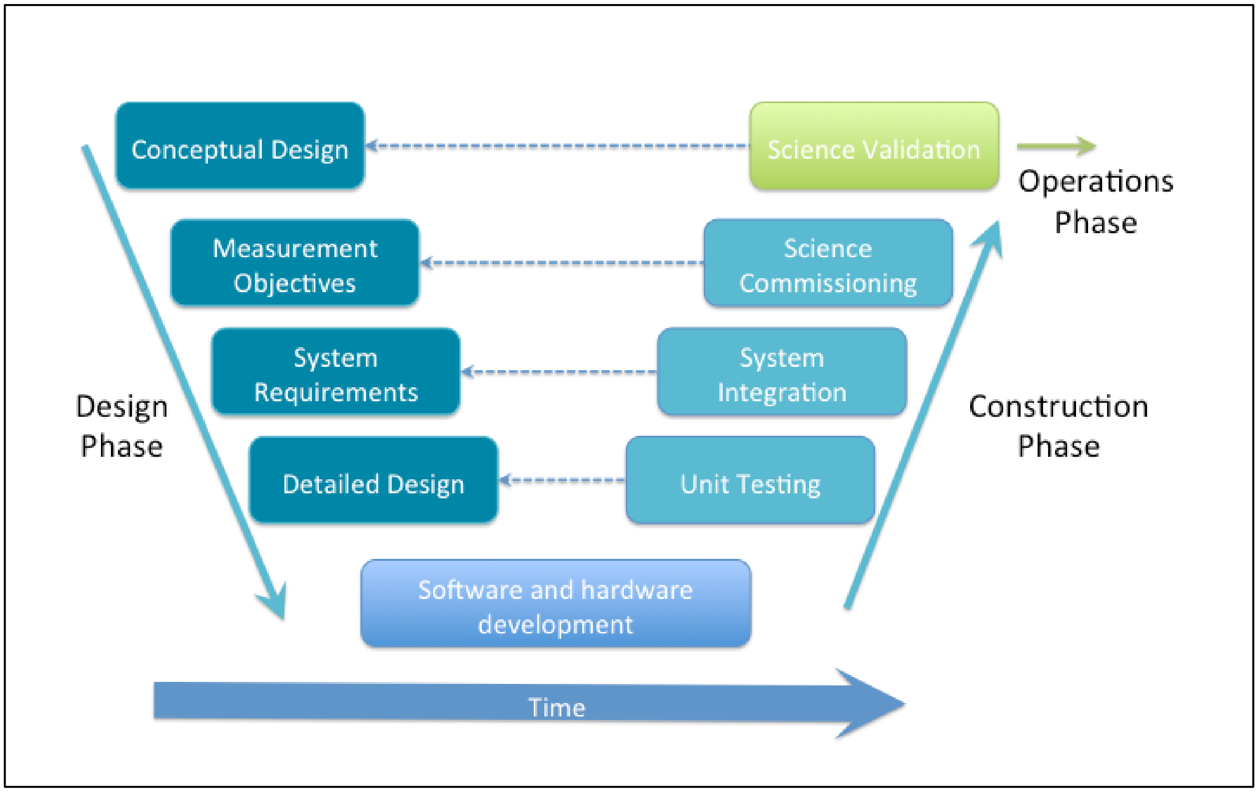

Figure 1: V diagram indicating phases of NEON design, construction, and operation

Based on these designs, the NEON staff initiates construction of the observatory by developing software, hardware, sampling protocols, data processing protocols, and other mechanisms for collecting, processing, and publishing data.

As components of the observatory reach completion, NEON applies a variety of testing in order to confirm that the completed observational and instrumented subsystems satisfy the original design goals. The dotted arrows in Figure 1 connect each level of testing with its corresponding level of design requirements. For NEON's instrumented systems, engineering staff perform unit testing of individual instrumental components to confirm that the component satisfies the detailed design requirements, then NEON's Systems Integration and Verification (SIV) team assesses whether the assemblies built from the units satisfy the system requirements. For observational systems, Field Operations staff execute the observational protocols during the Initial Sampling period to confirm the basic feasibility of each sampling protocol. Once the completed systems are ready to operate, the NEON science staff conducts science commissioning assessments to confirm that each site or payload meets the measurement objectives. The successful completion of these science commissioning tests marks the end of the hierarchy of assessments and thereby completes the NEON construction phase.

Because assessing NEON's ability to detect changes over time inherently requires data collected over significant timescales, NEON will execute a "science validation" stage under the initial years of full operations to establish whether NEON is adequately measuring changes in key ecological parameters. Science validation will include tests involving multiple sites within the same subsystem, cross-subsystem testing, and tests probing variation on timescales of six months or more. A full description of this stage lies outside the scope of the present document.

4 SCIENCE COMMISSIONING PROCESS

4.1 Subsystem-Level Commissioning

4.1.1 Scope

The NEON science commissioning process applies to each of NEON's data-generating subsystems (AIS, AOP, AOS, TIS, and TOS) as well as the first MDP unit. NEON will evaluate the performance of AIS, AOS, TIS, and TOS on a subsystem-wide basis using tests performed at individual sites (or, when appropriate, individual domains). The first MDP unit will have a dedicated commissioning process and review, as will each AOP payload (see discussion in section 4.4 below).

4.1.2 Objectives

The Subsystem-level Commissioning process has two objectives:

- *Assess subsystem performance.* NEON will employ process quality assessments, data quality assessments, and system performance assessments to determine whether a given subsystem performs sufficiently well to generate quality data consistently over space and time. Each assessment will consist of quantitative and verifiable criteria indicating the target performance level and, if appropriate, corresponding tolerances.
- *Identify and address operational issues.* NEON will track performance and maintenance anomalies that arise during the commissioning period and which may affect the ability to gather and publish data from each subsystem. NEON will investigate each anomaly, attempt to identify the root cause, and take appropriate action.

We explain the activities associated with each of these objectives below.

4.1.3 Activities

For each subsystem, NEON will carry out the following high-level activities to complete the science commissioning process:

- *Assess process quality metrics.* NEON scientists will evaluate metrics that measure whether the subsystem is collecting data with consistent process quality over space and time. These metrics will consider how well NEON is adhering to the *procedures* which affect data quality, rather than evaluating the quality of the resulting data. Metrics will include both a numerical target (benchmark) and a corresponding tolerance (when appropriate). Each test will consider a sample of sites (or domains) representing a range of environments. If a process quality test fails, then NEON will follow the process outlined in the *NEON Science Commissioning Anomaly Resolution Process* (RD [10]) to diagnose and resolve the issue. Section 5.2.1 presents additional detail on the planned process quality metrics.
- *Assess data quality metrics.* NEON scientists will evaluate data quality metrics that characterize the ability of the subsystem to deliver data of consistent quality over space and time. These metrics will assess the quality of the resulting data relative to adopted benchmarks, with corresponding tolerance (when applicable), across a sample of sites (or domains) representing a variety of environments. If a data quality test fails, then NEON will follow the process outlined in the *NEON Science Commissioning Anomaly Resolution Process* (RD [10]) to diagnose and resolve the issue. Section 5.2.2 further explains the data quality metrics.
- *Assess system performance metrics.* NEON scientists will evaluate system performance metrics that reflect the subsystem's ability to build and maintain quality instrument systems. These metrics, applicable only to the AIS, TIS, MDP, and AOP subsystems, will compare a site or payload's performance during the commissioning period to adopted benchmarks, with corresponding tolerance (where applicable). To ensure consistency, NEON will select a set of sites that span a wide range of environmental conditions. If a system performance test fails, then NEON will follow the process outlined in the *NEON Science Commissioning Anomaly Resolution Process* (RD [10]) to diagnose and resolve the issue. See Section 5.2.3 for more on system performance metrics.
- *Identify and address operational issues.* During the commissioning phase, NEON will identify issues related to operating and maintaining each subsystem and which may affect the ability to gather and publish data. NEON will track each issue, attempt to isolate the root cause, and take appropriate action in response. These responses may include: (1) formulating and implementing an appropriate mitigation strategy within the NEON construction phase; (2) requesting a deviation (indicating that the subsystem is not meeting a performance goal) and recommending a mitigation strategy for future implementation under the NEON operations phase; or (3) requesting a waiver (indicating that the subsystem is unable to achieve a performance target and that no mitigation activity is recommended). NEON will document the

process of identifying, investigating, resolving, tracking, and reporting these issues in *NEON Science Commissioning Anomaly Resolution Process* (RD [10]).

4.1.4 Schedule

The schedule of the NEON science commissioning process depends on the particular site subsystem or payload under consideration. Typically, commissioning begins following completion of the Initial Operational Capability Review (IOCR) for the payload or site subsystem. In certain cases, commissioning can make use of data gathered prior to IOCR; for example, observations gathered as part of site characterization activities may be usable for commissioning when the relevant protocols and field staff executing the protocols remain unchanged. In some cases, executing a test can begin immediately after IOCR once field operations staff are available to gather data. In other cases, evaluating the performance of a component may require waiting for data to accumulate over a sufficiently long timespan to enable evaluation of performance. The commissioning period is not formally limited in time, but is practically constrained by the need to hold the FOCR and FOOCR meetings prior to the completion of the NEON construction phase.

4.1.5 Inputs

The key inputs to the subsystem-level commissioning process are:

- *Subsystem Commissioning Plan.* Each NEON subsystem will have a document that describes the specific process quality, data quality, and (when applicable) system performance tests that NEON will apply to confirm that the subsystem will generate data of consistently high quality and will operate in a robust and reliable manner. Please see RD [03] through RD [08].
- *Commissioning Test Reports.* These documents will describe the detailed steps for executing each specific process quality, data quality, and system performance test.

4.1.6 Outputs

The key deliverables from the subsystem-level commissioning process are:

- *Commissioning Test Reports.* The same documents that describe the detailed steps for executing each commissioning test will also document the results obtained from the testing. Please refer to the complete description of the document storage locations and formats in RD [09].
- *Science Commissioning Archive.* Commissioning test authors will store data files associated with the commissioning tests in a file hierarchy dedicated to commissioning. Material stored in this Science Commissioning Archive (SCA) will include text files, Excel files, AOP-generated imagery, and possibly other data types. The intent of the archive is to allow NEON to reproduce the original data analysis procedures at will. Please see the complete description of the SCA in RD [09].

4.2 Final Operational Capabilities Review

4.2.1 Scope

To conclude the commissioning process for a given NEON subsystem (or component), NEON will convene a Final Operational Capabilities Review (FOCR) to evaluate the outcome of the commissioning process. The FOCR process described below will apply to the AIS, AOS, TIS, TOS, and MDP subsystems. Please refer to section 4.4 below for notes on how NEON will modify the FOCR process for AOP.

4.2.2 Objectives

The subsystem-level commissioning stage formally ends with a successful completion of the Final Operational Capability Review (FOCR). The review has two key objectives:

- *Confirm performance.* Determine whether the subsystem performance as characterized by the process quality, data quality, and system performance assessments demonstrate consistency and quality over time (short term) and space (within site/payload);
- *Confirm operation.* Affirm that NEON has resolved all operation and maintenance issues identified during the commissioning process, or has provided a mitigation plan and timeline for resolving the issues during operations which is acceptable to all stakeholders.

4.2.3 Prerequisites

All commissioning tests on the given subsystem must be completed before initiating the FOCR process.

4.2.4 Process

The NEON Systems Engineering (SE) group will organize and facilitate the review and will convene a review panel including the following key stakeholders at a minimum:

- NEON Program Management (PMO)
- NEON Director of Science (or delegate)
- NEON Lead Systems Engineer (or delegate)
- NEON Commissioning Scientist
- Representatives from the NEON science team responsible for the subsystem or payload
- Representative from NEON Operations staff

All stakeholders need to participate in the process and be prepared to discuss the status of the deliverables. The review panel will assess the outcome according to the FOCR criteria defined below and make its recommendation to the NEON Director of Science, the NEON Lead Systems Engineer, and the NEON PMO. The NEON Lead Systems Engineer will make the final assessment of pass/fail based on input from all stakeholders and final agreement from the NEON Project Management Office and NEON Director of Science.

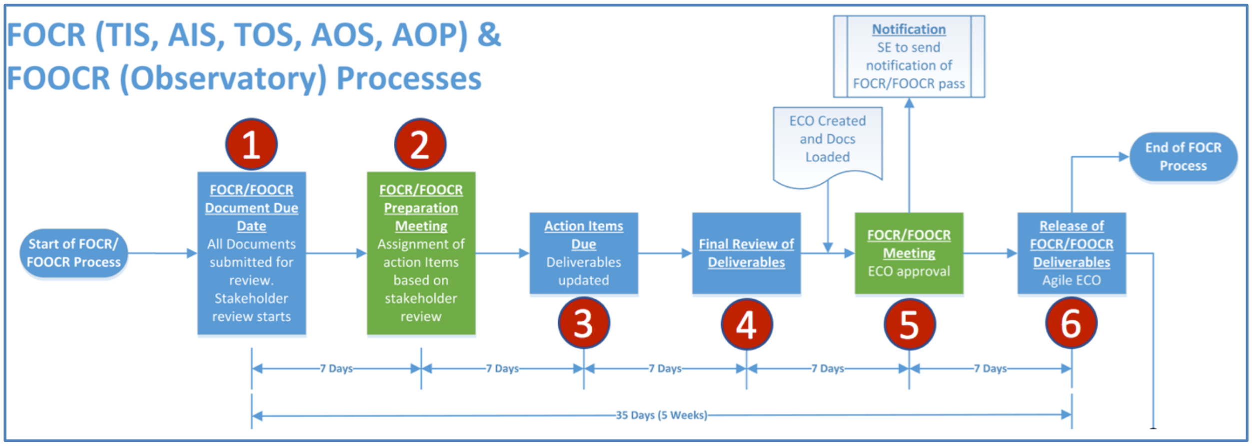

Figure 2: FOCR Process Flow Diagram

Figure 2 illustrates the FOCR review and approval process flow, which is shared with the FOOCR review and approval process up to the release of FOCR/FOOCR deliverables. These key steps in the process are indicated by number in the figure:

1. The FOCR/FOOCR process flow starts with the submission of all documents and deliverables at the FOCR/FOOCR Document Due Date. Once documents and deliverables are submitted, the FOCR/FOOCR review period begins and reviewers are asked to record their questions, issues and action items in a common area.
2. One week following the document due date, the FOCR/FOOCR Preparation Meeting will be held to review all questions, issues and action items submitted as part of the document review. At this review, all action items are officially assigned with a due date one week later. Action items that cannot be resolved within one week will be considered for waivers and/or deviations.
3. All action items are due one week after the FOCR/FOOCR Preparation Meeting. If action items remain unresolved (e.g., a complex anomaly may require more investigation to compose an acceptable mitigation plan) then the approval process will halt until the item is resolved.
4. Once action items are addressed, the reviewers who originally submitted action items will be allowed one week to perform a final review of deliverables that verifies that their action items are completed. Once this final review is completed, the updated documents and deliverables will be loaded into an enterprise change order (ECO) in the configuration management (CM) system.
5. The FOCR/FOOCR Meeting occurs three weeks after the FOCR/FOOCR Preparation Meeting. Given that that all action items should have been completed and reviewed prior to this time, the FOCR/FOOCR Meeting will ideally consist of ECO stakeholders approving documents in the CM system.
6. Following the FOCR/FOOCR Meeting, the documents and deliverables will be officially released.

The FOCR process ends here with the release of documentation. The FOOCR process continues, as described in section 4.3.3.

4.2.5 Criteria

The NEON FOCR panel shall apply the following criteria in the Subsystem Final Operational Capability Reviews.

1. Does the performance of each NEON subsystem, as characterized by the process quality, data quality, and system performance assessments, demonstrate consistency and quality performance relative to the adopted benchmarks?
2. Has NEON resolved all operation and maintenance issues identified during the commissioning process by fixing the issue, providing a mitigation plan and timeline for resolving the issue during operations which is acceptable to all stakeholders, or requesting a waiver for non-compliance?

4.2.6 Inputs

Specific deliverables NEON will require to complete the FOCR include:

- *Commissioning Test Reports* recording the outcomes of each commissioning test.
- *Commissioning Summary Report* providing a high-level review of the commissioning test outcomes and the issues that arose in commissioning.

4.2.7 Outputs

The deliverables NEON will generate at completion of the FOCR include:

- *FOCR Report* recording the results of the FOCR process. The SE group will prepare this report and provide it to the NEON Program Manager for review and approval.

4.3 Final Observatory Operational Capabilities Review

4.3.1 Scope

NEON will hold the Final Observatory Operational Capabilities Review (FOOCR) as the final step in the Observatory construction process. The FOOCR will occur after all subsystem FOCRs are finished and will confirm that all NEON project deliverables are complete, that NEON has generated all of the agreed-upon L1 data products, and that these data products are science-grade data (as defined in the *NEON Acceptance Plan*, AD [02]) that are accessible through the NEON portal or otherwise made available to all key stakeholders. The FOOCR will assess the following subsystems and capabilities:

- AIS

- AOP
- AOS
- MDP
- TIS
- TOS
- All minimum viable data products to which NEON has committed delivery

4.3.2 Objectives

The observatory-level commissioning process (and the entire NEON construction phase) formally concludes with successful completion of the Final Observatory Operational Capability Review (FOOCR). The objectives of this review are to:

- *Confirm deliverables.* Verify that NEON has completed all planned construction deliverables.
- *Confirm data products.* Establish that NEON is able to provide all intended L1 data products from the data-generating subsystems, either through the NEON data portal or by an acceptable alternative means of distribution.

4.3.3 Prerequisites

All NEON generating subsystems must have transitioned to operations through the FOCR process described above before the FOOCR process can begin.

4.3.4 Process

The NEON Systems Engineering (SE) group will organize and facilitate the FOOCR and will convene a review panel including the following key stakeholders at a minimum:

- NEON Director of Science (or delegate)
- NEON Lead Systems Engineer
- NEON Commissioning Scientist
- Representatives from NEON science team who support the AOP, AOS, AIS, TOS, and TIS subsystems
- Representatives from NEON science team who support the development of NEON data products
- Representative from NEON Operations team

All stakeholders need to participate and be prepared to discuss the status of the deliverables. The review panel will assess the outcome according to the formal FOOCR criteria defined below. The Lead Systems Engineer will prepare a formal FOOCR report and provide it to the NEON Program Manager and the NSF for review and approval. The FOOCR report approval will include a formal sign-off process where approval is required by the NEON Director of Science, the NEON PMO, and the NSF.

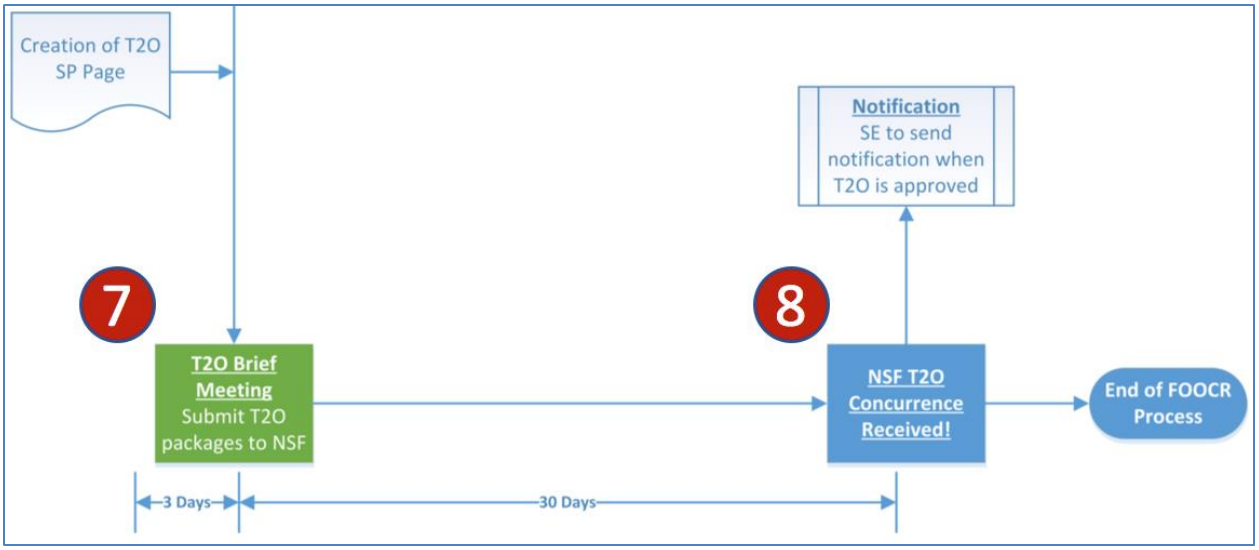

Figure 3: FOOCR Process Flow Diagram (continuing from Figure 2)

The FOOCR review and approval process flow follows the same flow as described for FOCR (Figure 2); however, the FOOCR requires a transition to operations (T2O) following the release of FOOCR deliverables. The following steps extend the FOCR process and correspond to numbered items in Figure 3:

7. Following the release of FOOCR deliverables, the NEON Systems Engineering team will hold a meeting with NEON Program Management to review a T2O briefing. At the end of that review, NEON will submit all T2O deliverables to NSF for a 30-day review period after which the NSF will approve if they concur with transition to operations.
8. Once NEON receives T2O concurrence from NSF, the remaining subsystems under construction will transition to operations.

4.3.5 Criteria

The FOOCR panel will apply the following criteria in evaluating the Observatory's readiness to enter full operation:

- Has NEON completed all Observatory deliverables?
- Are all of the NEON data-generating subsystems delivering L1 data products?
- Is NEON publishing all required L1 data products either on the NEON Data Portal or available through a suitable alternative distribution channel?
- Has NEON adequately addressed all issues that arose during commissioning and initial operations, by:
  - resolving the issue through the construction warranty process as defined in the *NEON Configuration Management Plan* (AD [03]),

  - recommending mitigation plans to occur later during the Operational phase,
  - obtaining approval for a deviation for temporary non-compliance with a corresponding mitigation plan, or
  - obtaining approval for a waiver for permanent non-compliance?

4.3.6 Inputs

Specific deliverables NEON will require to complete the FOOCR include:

- *AOP Payload 1 FOCR Report*
- *AOP Payload 2 FOCR Report*
- *AOP Payload 3 FOCR Report*
- *AIS FOCR Report*
- *AOS FOCR Report*
- *TIS FOCR Report*
- *TOS FOCR Report*
- *MDP FOCR Report*
- *NEON FOOCR Deliverables Status Report*
- *NEON FOOCR Data Products Status Report*
- *NEON FOOCR Issue Status Report*

4.3.7 Outputs

Specific deliverables NEON will generate as part of FOOCR include:

- *FOOCR report* detailing the outcome of the FOOCR.

4.4 AOP Commissioning

The NEON AOP subsystem consists of three separate payloads with differing instrument configurations and capabilities, and which will come online at different times. Given the distinct characteristics of the AOP subsystem relative to the other data-generating subsystems, the commissioning process for AOP will deviate slightly from the process employed for the remaining NEON subsystems. Specific differences include the following:

- Instead of a single FOCR applicable to all AOP payloads, NEON will hold a separate FOCR for each AOP payload. The objectives of each AOP FOCR will remain as described in 4.1.2.
- Rather than applying the same commissioning criteria to each AOP payload, NEON will modify the suite of tests to eliminate redundant tests (i.e., tests that concern the performance of staff rather than the performance of instrumentation) for payloads 2 and 3, and to eliminate or modify tests that would not apply to the instruments of payloads 2 and 3 because of their distinct capabilities.

5 SCIENCE COMMISSIONING ASSESSMENTS

Both the site- and payload-level commissioning process and the observatory-level commissioning process involve objective assessments of NEON's performance. Below we describe the types of assessments NEON will employ, their key attributes, and the key inputs to these assessments.

5.1 Attributes

Each science commissioning assessment is based on a criterion associated with some measureable aspect of NEON's performance. Each criterion must be *verifiable*, meaning that it includes a stated performance level. To make these quantitative assessments, each should include a specific numerical target. In some cases, this target represents a tolerance based on the "true" value of some quantity as derived from an independent analysis; for example, temperature measured by NEON instrumentation could be measured to an accuracy of 0.1 C based on comparison with a calibrated thermometer. In other cases, the performance level is implicit (e.g., 100% success or 0% failure) and the stated performance level represents either a minimum success rate or maximum rate of failures.

Commissioning criteria ideally correspond to the established NEON technical requirements for the relevant subsystem (as recorded in the database of requirements). In practice, requirements either don't exist for aspects of certain NEON components or are not written in a verifiable form amenable for application in commissioning tests. In such cases, NEON science staff members will compose suitable verifiable criteria for use in commissioning tests based on their assessment of performance levels needed to produce data of suitable quality.

5.2 Types

NEON will employ three types of assessments in the site/payload and observatory commissioning levels: process quality metrics, data quality metrics, and system performance metrics. Below we define the distinctions between them.

5.2.1 Process quality metrics

Process quality metrics measure whether NEON is collecting data with consistent process quality over space and time. These metrics evaluate how well NEON is adhering to the *procedures* which affect collection of good data, rather than evaluating the quality of the resulting data themselves.

Table 1: Example process quality metrics

Subsystem	Metric
AOP	AOP shall conduct at least 95% of acquisitions during the period of peak greenness.
AIS	AIS shall measure ground and surface water stage such that the NEON cyber infrastructure receives 95% of the expected L0 measurements over a 60-day period

	excluding downtime for scheduled maintenance, calibration/validation, and extreme weather events.
AOS	AOS shall collect at least 95% of scheduled water chemistry samples per site per bout (minimum of 6 bouts) and process all samples (filtered) within 6 hours.
TIS	TIS shall measure 2D wind speed and direction such that each tower-based 2D anemometer transmits 95% of the expected measurements over a 60-day period following IOCR that excludes downtime for scheduled maintenance and calibration.
TOS	NEON shall gather phenology observations at a frequency appropriate to the growth form and phenophase as specified in the plant phenology protocol. Tolerance shall be ±1 day for sampling intervals < 1 week and shall be ±2 days for sampling intervals ≥ 1 week. The minimum number of bouts to be used for assessment of sampling frequency shall be 10 bouts that include a minimum of one phenophase transition.
MDP	MDP shall measure CO_2 concentration such that the 30-minute L1 averaged measurements over a 30-day period show that an acceptable fraction of those data are valid (i.e., Q_F=0). This period shall exclude downtime for scheduled maintenance and calibration.

5.2.2 Data quality metrics

Data quality metrics characterize the ability of NEON to deliver data of consistent quality over space and time. These metrics will assess the quality of the resulting data relative to adopted benchmarks, with corresponding tolerance (when applicable).

Table 2: Example data quality metrics

Subsystem	Metric
AOP	AOP shall measure horizontal locations of LiDAR reference points to an root-mean-squared accuracy of 58 cm (using wide divergence mode at 1000 m above ground level).
AIS	AIS shall confirm that the difference between the L1 5-minute averaged pH data product and the lab-analyzed pH measurement of the sample acquired within the same 5-minute window is less than 0.2 pH units RMS (root mean square).
AOS	AOS shall measure surface water alkalinity to an accuracy of 5% (or 10% for low conductivity water) as determined by replicate analysis of alkalinity samples over at least four bouts at each station.
TIS	TIS shall confirm that the root mean square error (RMSE) of 30 minute soil temperature averages over a 3 month window for each sensor depth shall be at least 10% greater than the RMSE value for the sensor immediately below it in the soil temperature profile.
TOS	TOS shall measure the relative proportion of the stable isotopes for C and N in soil biogeochemical samples to an accuracy of 2 standard deviations for δ13C (0.3 ‰) and δ15N (0.4 ‰).

5.2.3 System performance metrics

System performance metrics reflect NEON's ability to build and maintain quality instrument systems. These metrics will compare a site or payload's performance during the commissioning period to adopted benchmarks, with corresponding tolerance (where applicable). NEON will select metrics that characterize the performance of the corresponding subsystem over time and over a variety of environmental conditions. System performance metrics are applicable only to the AIS, TIS, and AOP subsystems, not to AOS or TOS.

Table 3: Example system performance metrics

Subsystem	Metric
AOP	AOP shall demonstrate that for sites flown during the commissioning flight campaign, the AOP survey has covered at least 80% of the area in the Priority 1 Flight Box.
AIS	The wireless network for the NEON AIS subsystem will operate at least 99% of the time when measured over a period of 60 consecutive days.
TIS	TIS shall confirm that the rate of aspiration of each aspirated temperature sensor at the site shall remain within the nominal range at least 95% of the time for 60 consecutive days during the commissioning period.
MDP	The pitch/yaw/roll orientation of the anemometer prior to removal and reinstallation shall be within its specified range for at least two trials (separated by two weeks or more) during the commissioning period. The elevation (pitch) shall be initial angle ±2° (i.e., level) and the azimuthal orientation (yaw/roll) shall be initial angle ±2°.

5.3 Key inputs

The two key inputs to the commissioning assessments are measurement requirements and data products.

5.3.1 Measurement requirements

The measurement requirements are formally-defined stipulations indicating:

1. what quantities NEON will measure, and
2. with what accuracy to make these measurements.

These properties derive from higher-level requirements dictating the scientific questions that NEON intends to address. When measurement requirements exist, we can typically adapt them to serve as criteria for commissioning assessments. We generally must make certain modifications to the formal measurement requirements to render them suitable for use as commissioning assessments; for example, placing explicit limits on the number of samples we will test, defining the specific time frame applicable to the assessment, or adding information on what techniques we will use to confirm the quality of the data.

5.3.2 Data Sources

Different types of commissioning assessments involve distinct types of data sources, as described below.

1. *Process quality assessments* typically measure aspects of the data collection procedures, rather than formal NEON data products. For example, they may consider the number of times a procedure was run, the conditions under which a procedure was run, the timing of procedures, the quantity of data received, or other characteristics that do not correspond to a formal data product but affect the quality of the resulting data product. Process quality assessments may involve additional work by scientists or field staff to capture data which is not typically recorded under normal operations.
2. *Data quality assessments* confirm that the data products NEON is providing to the community are robust and reliable. Assessments of data quality should involve specific data products that NEON will publish, as opposed to intermediate data products intended only for internal use.
3. *System performance assessments* typically measure characteristics associated with reliability, repairs, and maintenance. They may measure the frequency of failures on a given system, the uptime of the system, the ability of a system to remain in proper working order, or the number of times that NEON had to repeat data collection due to breakdowns. The information NEON uses in these assessments is typically drawn from maintenance records or logs that monitor system performance, rather than from formal data products.

5.4 Scope

The scope of a given commissioning assessment is constrained on the basis of costs (i.e., effort and expense required to collect and analyze the measurements) and benefits (i.e., addressing an important aspect of system or subsystem performance). Costs include:

1. Effort by science staff to compose and document the data collection and analysis procedure;
2. Effort by the field staff to collect the data, perform in-house analysis, or process samples for external analysis;
3. Cost of equipment and supplies needed to execute a test;
4. Costs of contracts to have outside experts or laboratories verify results;
5. Effort by science staff to develop data analysis procedures and routines;
6. Effort by science staff to analyze and document results.

We do not consider potential expenses involved with investigating and resolving issues that arise in the investigation; such effort is covered under the warranty provisions of the NEON construction agreements (see AD[03]).

Potential commissioning assessments range widely in scope, from simple to highly involved. In practice, we have aimed to limit the costs associated with each assessment by designing analysis procedures that involve a minimum of effort from NEON Field Operations staff to collect and process measurements, a minimum of external contracts, and can be completed in a limited time by science staff. Further consideration of resources appears in each subsystem's commissioning plan (RD [03-08]).

6 DATA QUALITY VERIFICATION METHODS

The goal of each data quality test is to confirm that the corresponding NEON subsystem is generating data of acceptable quality. Various techniques can be used to measure data quality. Some of these techniques enable checks on data accuracy by comparing a given NEON measurement to an independent measure of "ground truth". Other methods offer value by establishing that a measurement is consistent without addressing whether the value is "correct". The most appropriate technique to apply in a given data-quality test depends on numerous factors, including the nature of the data source (instrumental or sampling-based), timeliness, available resources, availability of external data, etc. The following sections describe the various techniques we have considered for tests applicable to instrument-based systems and sampling-based systems.

6.1 Field Sampling

NEON considered the following data quality verification methods as applicable to field sampling cases. This list is ordered by the approximate desirability of the methods, with the methods listed first being those most able to assess the accuracy of the results, and the more limited methods last.

6.1.1 Independent lab analysis

Description	Multiple samples acquired by the same team in the same place are analyzed at different labs and the resulting measurements compared.
Property tested	Accuracy of laboratory analysis.
Advantages	• Straightforward to perform • Resistant to natural variation • Good test of accuracy
Disadvantages	• Expense of laboratory testing
Example	Split a single soil sample into multiple parts and send to different labs for analysis.
Comments	Well suited to analysis of abiotic components. Not necessarily well suited to biotic samples.

6.1.2 Redundant sampling

Description	Multiple samples acquired by the same team in the same place are analyzed in the identical fashion at the same lab and the resulting measurements compared.
Property tested	Consistency of sampling and lab analysis.
Advantages	• Straightforward to perform
Disadvantages	• Does not address accuracy of results • Possibly expensive

	• Potentially subject to bias from natural variation (although this can be minimized by carefully homogenizing the sample)
Example	We perform identical soil composition analysis on multiple samples acquired from the same plot.
Comments	None.

6.1.3 Independent sampling

Description	Samples acquired in the same place by different field teams are subjected to identical analysis.
Property tested	Consistency of field sampling process.
Advantages	• Focuses on field sampling
Disadvantages	• Logistically difficult • Does not address accuracy of results • Potentially subject to bias from natural variation • Could result in disturbance of plot
Example	We perform identical soil composition analysis on samples acquired by different teams from the same plot.
Comments	Well suited to properties that are consistent from day to day: tree height and diameter, plant species prevalence, etc.

6.1.4 Field analysis of training samples

Description	Subject matter experts create "training samples" for field classifications and send these to our field teams periodically to have them classified, then compare results to "truth".
Property tested	Accuracy of field classifications
Advantages	• Straightforward to perform • Established "truth" • Immune from natural variation • Re-usable for multiple sites over years • Low cost to perform; no travel required
Disadvantages	• Requires effort to set up
Example	We create a collection of ground beetles that have been classified by experts and send it to our field teams for classification. We compare the field results to "truth" to determine the number of correct and incorrect classifications.
Comments	May require domain-specific samples; for example, different sites will have different species of ground beetles.

6.1.5 Lab analysis of training samples

Description	Subject matter experts create "training samples" for lab analysis and send these to our labs periodically to have them analyzed, then compare results to "truth".
Property tested	Accuracy of lab analysis

Advantages	• Straightforward to perform • Potentially re-usable • Establishes "truth"
Disadvantages	• Requires effort to set up • Possibly expensive
Example	We create soil samples with known composition and send them to labs for analysis. We compare the lab results to "truth" to determine the accuracy of the results.
Comments	NEON's Calibration/Validation group is also planning such testing.

6.1.6 Expert verification of field results

Description	Field classifications are confirmed through similar analysis by experts on site (e.g., contractor) or at NEON HQ.
Property tested	Accuracy of field classifications
Advantages	• Straightforward to perform
Disadvantages	• Potentially time-consuming • Potentially expensive
Example	We take samples of ground beetles collected and classified in the field and have them independently classified by beetle experts.
Comments	None

6.1.7 Laboratory verification of field results

Description	Field classifications are confirmed through independent analysis.
Property tested	Accuracy of field classifications
Advantages	• Straightforward to perform • Possibly more authoritative than expert classification
Disadvantages	• Potentially expensive
Example	We take samples of ground beetles collected and classified in the field and have them independently classified in a DNA lab.
Comments	None

6.1.8 Data triangulation

Description	Field classifications are confirmed by comparison with similar sampling performed by non-NEON teams.
Property tested	Entire process: field sampling protocols, lab analysis, ATBDs
Advantages	• Most comprehensive test of our methods
Disadvantages	• Relies on external data of unknown quality • Subject to natural variation • Availability limited to certain sites and data products

Example	We compare the relative abundance of various ground beetle species in NEON samples to the results obtained by an outside group sampling at the same location and at the same time.
Comments	None

6.2 Instrumental Sampling

NEON considered the following data quality verification methods as applicable to instrumental sampling cases. This list is ordered by the approximate desirability of the methods, with the methods listed first being those most directly able to assess the accuracy of the results, and the more indirect methods last.

6.2.1 Calibration unit

Description	Measurements are checked via comparison with calibrated instruments at the same time and location
Property tested	Accuracy of measurements
Advantages	Using different instruments provides a more robust check on instrumentation May reduce effect of systematic errors
Disadvantages	Requires sending a person to the site to perform testing
Example	We compare measurements from a TIS air temperature sensor to a portable, calibrated air temperature sensor at the same site.
Comments	Possibly the most accurate independent technique.

6.2.2 Co-located instrumentation

Description	Measurements are checked via comparison with *identical, co-located* instrumentation
Property tested	Consistency of measurements
Advantages	• Provides a high-quality, direct comparison.
Disadvantages	• If the data are processed through identical pipelines then this does not provide a completely independent check on data processing algorithms. • Does not address systematic error; similar errors could impact both measurements.
Example	We compare measurements from identical temperature sensors on the same tower boom arm.
Comments	Few NEON instruments have multiple instances installed at the same location.

6.2.3 Nearby instrumentation

Description	Measurements are checked via comparison with identical, *nearby* instrumentation
Property tested	Consistency of measurements
Advantages	• Eliminates instrumentation as a source of error
Disadvantages	• If the data are processed through identical pipelines then this does not provide a completely independent check on data processing algorithms.

	• Does not address systematic error; similar errors could impact both measurements.
Example	We compare measurements from identical temperature sensors on different tower boom arms.
Comments	Few NEON instruments have duplicate instances installed at a nearby location.

6.2.4 Similar instrumentation

Description	Measurements are checked via comparison with similar, nearby instrumentation
Property tested	Accuracy of measurements
Advantages	• Using different instruments provides a more robust check on instrumentation • May reduce effect of systematic errors
Disadvantages	None
Example	We compare measurements from an AIS air temperature sensor to measurements from a TIS sensor on a nearby tower.
Comments	None

6.2.5 Remote instrumentation

Description	Measurements are compared to remote instrumentation
Property tested	Accuracy of measurements
Advantages	• Available over wide areas
Disadvantages	• Subject to natural variation • Unlikely to give great agreement
Example	We compare barometric pressure measurements to data derived from satellites
Comments	Only applicable to a few measurements.

6.2.6 Correlated measurements

Description	Measurements are compared to other data products which should correlate
Property tested	Correlation of measurements
Advantages	• Provides a good sanity check
Disadvantages	• Does not verify accuracy
Example	We compare measurements of relative humidity with the precipitation monitor
Comments	None

6.2.7 Profile trends

Description	Measurements are compared to nearby, identical measurements which need not agree
Property tested	Consistency of measurements
Advantages	• Provides a good sanity check
Disadvantages	• Does not verify accuracy

Example	We confirm that soil temperature measurements get colder as the sensors get deeper
Comments	None

6.2.8 Theoretical predictions

Description	Measurements are compared to expected values based on theory
Property tested	Accuracy of measurements
Advantages	• Good sanity check, or even an accuracy check
Disadvantages	• Typically not a precise comparison
Example	Compare measured PAR against the predicted value, given model for atmospheric transmission
Comments	None

7 COMMISSIONING PROGRAM OPTIMIZATION STRATEGIES

A comprehensive commissioning program for NEON would test every protocol and sensor at every site to confirm performance; however, such a program would require substantial time and resources to complete. It should be possible to optimize the return on investment by deploying commissioning resources strategically rather than comprehensively. We can optimize the commissioning program by selecting a subset of tests to analyze, and a subset of sites to test for each assessment. Although the resulting commissioning program will not assess all NEON components, the observatory's IOCR/T2O process will nevertheless ensure a basic level of functionality at all sites.

7.1 Assessment Selection Strategies

Rather than test every protocol and sensor, we can optimize the commissioning program by focusing on certain aspects of each NEON subsystem. The following criteria can be used to develop the list of assessments to consider:

- *Importance to NEON science.* Certain tests are more relevant to the scientific goals of NEON than others. For example, measuring the relative humidity at a lake site may be less important than measuring the amount of nitrate in the water. When possible, NEON commissioning will test protocols that are of higher scientific importance as determined by the science team.
- *Areas of concern.* Certain NEON protocols and sensors may be more susceptible to faults and errors than others, and testing these aspects of a subsystem provides a more critical test of the subsystem. For example, based on feedback from NEON's Field Operations and Engineering teams, we suspect that the fans in the aspirated air temperature sensors will be among the most error-prone operational components of the TIS subsystem. Commissioning tests should emphasize testing of such components that are believed to present a risk to reliable gathering of quality data.

- *Variety of science topics.* Each NEON subsystem will address numerous ecological "grand challenge" questions. Rather than focusing on a single topic, NEON's commissioning assessments should touch on as many different questions as possible. For example, instead of exclusively testing protocols that deal with the diversity of species, TOS assessments should also touch on land use and other ecological questions.
- *Availability of data products.* NEON commissioning is slated to occur at the same time as the NEON data products are being developed, meaning that not all data products will be available in time to use for commissioning. To avoid delaying the completion of science commissioning, we will favor testing those data products that are available for analysis earlier in the commissioning phase.
- *Feasibility of testing.* Certain aspects of NEON are more readily tested than others. For example, the NEON secondary precipitation sensors consist of small "buckets" that only generate a reading when they have accumulated a certain amount of liquid. Confirming that these sensors generate data reliably would be difficult; hence, we will instead test only the primary precipitation sensors that transmit data on a defined schedule. Similarly, NEON commissioning assessments will favor those components of the system that are most amenable to testing.

7.2 Site Selection Strategies

Rather than test every NEON site, we can optimize the commissioning program by conducting commissioning tests at a subset of sites. The optimal subset will depend on the nature of the test, as indicated below.

7.2.1 Strategies for Observational Subsystems

NEON employs Field Technicians to gather data for the AOS and TOS subsystems by executing defined protocols. Most NEON domains have multiple sites at which the Field Technicians execute the same protocol. We can assume that the primary variable governing the variability in performance from site to site within the AOS and TOS subsystems is not the properties of the site itself, but the training and capabilities of the people who execute the protocols. Given this assumption, the optimal strategy for comparing performance across a given subsystem is to test at no more than one site per domain, since the same personnel will execute a given protocol at each site within a domain. Since sites within a domain may vary in difficulty of implementing a given protocol (even with the same staff), we aim to select sites that present a range of challenges rather than simply choosing those at which the protocol is most easily executed. The strategy of testing one site per domain can reduce the scope of testing AOS subsystems by over 40% and the scope of TOS subsystems by almost 60%.

7.2.2 Strategies for Instrumented Subsystems

Each NEON TIS and AIS site will be configured with nearly identical instrumentation. The primary driver of performance will be the design and construction of individual sensors, but an important secondary consideration will be the environmental conditions to which the instrumentation will be subjected. Instead of testing instrument performance at every NEON site, we can choose to assess performance at a subset of sites that represent the range of environments over which the instrumentation will operate. This may include sites spanning a range of temperatures, humidity, wind, elevation, snowfall, and other environmental conditions that could conceivably affect the function of the instrumentation. This strategy can reduce the resources required to complete AIS and TIS testing by 50-75%.

8 DOCUMENTATION PLAN

The present document represents the top-level document in a document tree that will describe the NEON Science Commissioning process and results. Commissioning of each subsystem will require additional documentation.

8.1 Document Types

Within each group, the documents fall into three categories:

- *Subsystem Commissioning Plans* provide a general description of the commissioning assessments applicable to a given subsystem. These documents indicate the justification for selecting each test, outline the strategy for performing the test, and tabulate the estimated resources required to complete the test plan. They also list the proposed set of sites, domains, or payloads at which NEON will execute each test. The primary purpose of these documents is to provide a high-level overview of the commissioning strategy applicable to a given subsystem at a level appropriate for a review panel. NEON will compose, review, and approve a Subsystem Commissioning Plan for each subsystem prior to developing detailed test plans.
- *Commissioning Test Reports* provide detailed descriptions of the rationale for a given test, the comprehensive procedure required to complete the test, and the detailed results of executing the test. They also document the specific data files used as input and output from the tests as well as any software or other tools necessary to complete the analysis. The document will thus serve as both a rationale and instruction manual for completing the test as well as a record of the performance of each tested site or payload. NEON science staff will compose the Commissioning Test Reports for a given subsystem only after approval of the corresponding Subsystem Commissioning Plan. As NEON staff members complete the assessment for each site, domain, or payload, they will update the document to add a new section describing their findings.
- *Commissioning Summary Reports* provide a synopsis of commissioning results from a particular subsystem or AOP payload. The primary purpose of the document is to summarize the results of

all applicable tests to be reviewed at a given FOCR or FOOCR so that review panelists do not need to consult a large number of independent commissioning test reports in order to determine the test outcomes. The Commissioning Summary Report will include:

- A compliance matrix indicating the status of all of the scheduled science commissioning assessments, including process quality metrics, data quality metrics, and system performance metrics (when applicable).
- A list of all issues NEON has identified during the commissioning process on the corresponding subsystem, including issues affecting commissioning assessments, actions that NEON took to resolve the issues, and plans for resolving any outstanding issues.

8.2 Document Hierarchy

The figures below illustrate the relationships between site- and payload-level commissioning documents pertaining to the plans and results from the commissioning tests, respectively. Figure 4 shows the hierarchy of documents describing commissioning plans. The top-level NEON science commissioning plan (this document) covers generic aspects of the site- and payload-level commissioning plans. The subsystem commissioning plans describe tests specific to each respective subsystem. Each subsystem plan has a corresponding distinct set of Commissioning Test Reports that detail the plans for completing a certain test and document the results. For clarity, the figure does not show the Commissioning Test

Reports for AOS, TIS, and MDP.

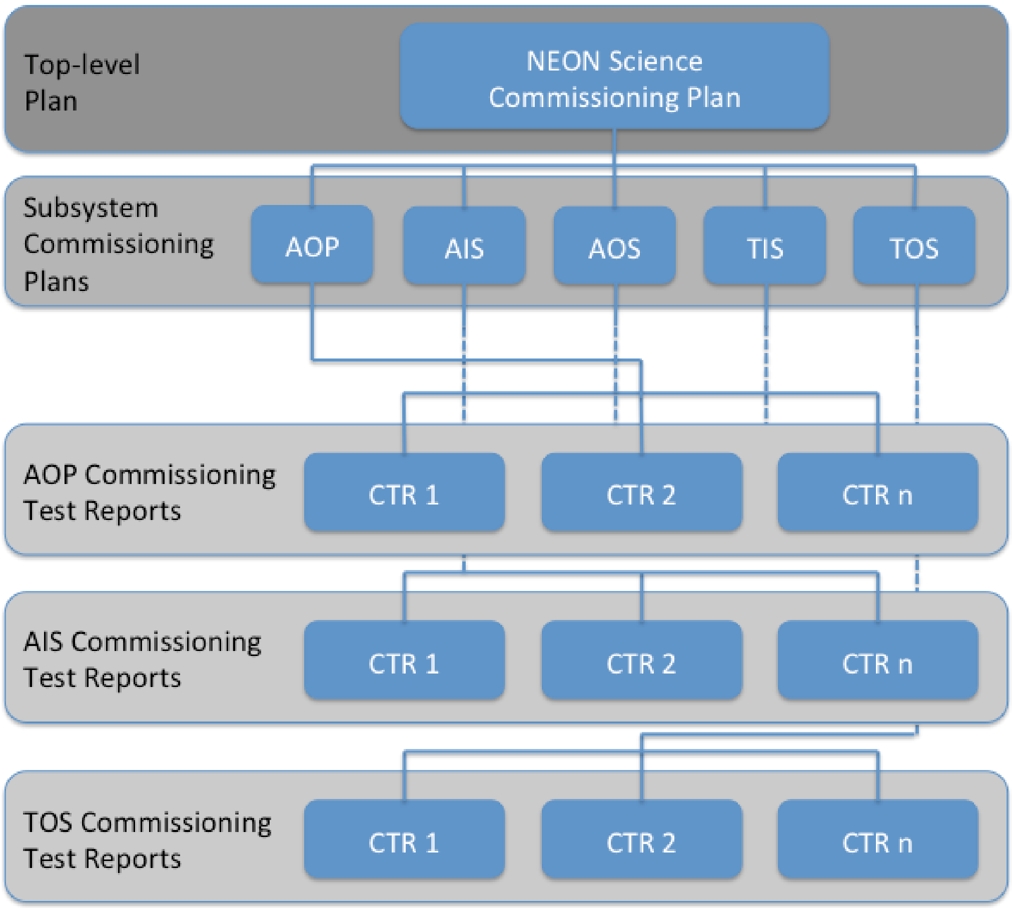

Figure 4: NEON site- and payload-level science commissioning plan hierarchy

Figure 5 illustrates the relationship between the Commissioning Review Report and the Commissioning Test Reports for each non-AOP subsystem. The Commissioning Review Report draws details from each of the Commissioning Test Reports for that subsystem.

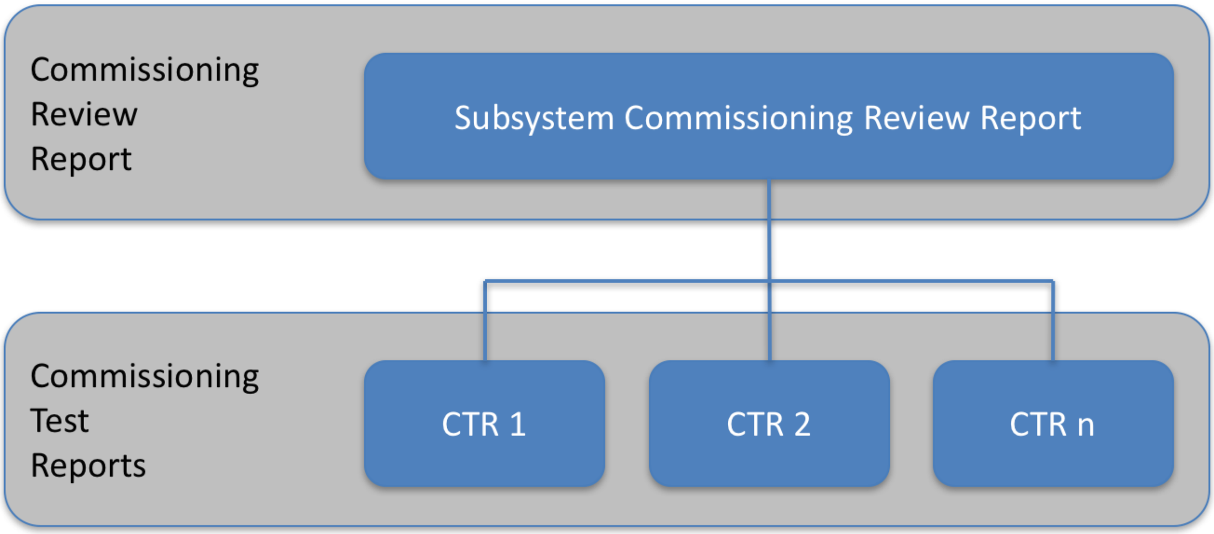

Figure 5: NEON FOCR material hierarchy (non-AOP)

Figure 6 illustrates the relationship between the Commissioning Review Report and the Commissioning Test Reports for the AOP payloads. Since each AOP Commissioning Test Report records results from all payloads associated with the given test, the relationship between the AOP Commissioning Review Report and the AOP Commissioning Test Reports is many-to-many rather than a strict hierarchy.

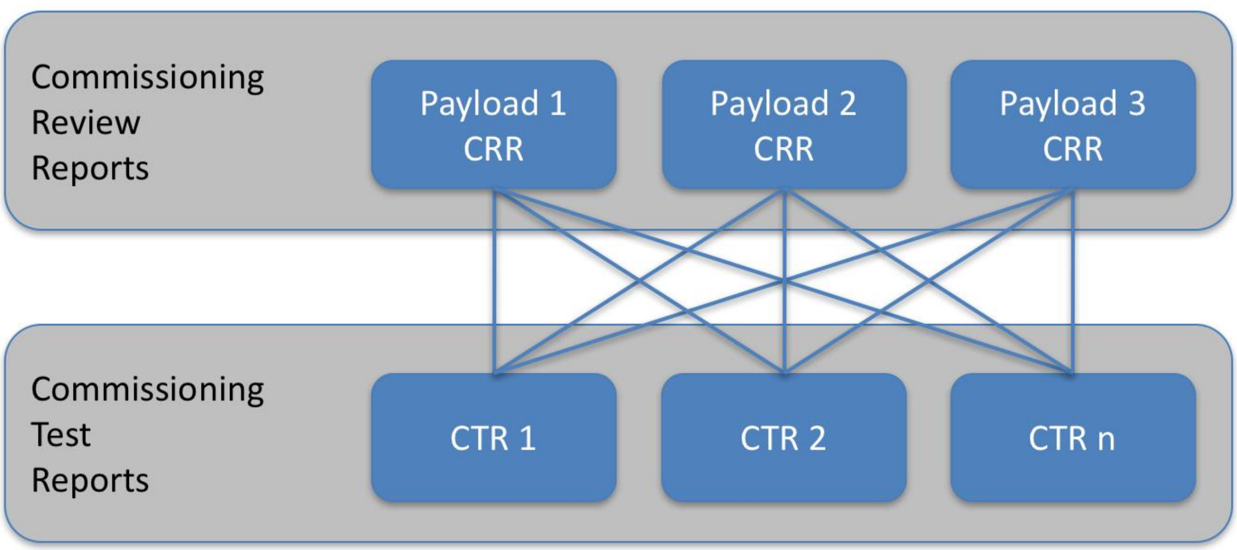

Figure 6: NEON FOCR material hierarchy (AOP)