

Recommended Guidelines for testbeds and proving grounds

5/19/11

Drafted by:

P. Davidson, J. Tuell, L. Uccellini (NWS); J. Gaynor, S. Koch, R. Pierce, M. Ralph (OAR)

I. Vision: Orderly transition of research capabilities to operational implementation through developmental testbeds and pre-deployment testing and operational readiness/suitability evaluation in operational proving grounds.

II. Purpose: Based on the Vision, recommend roles and responsibilities, function and governance for developmental testbeds and operational proving ground. The recommendations include testbeds and proving grounds documenting, in Charters and/or Terms of Reference, their scope, operations (roles and responsibilities, and function/execution) and governance. The guidelines are not intended to direct revision of existing Charters/Terms of Reference that are consistent with the recommendations.

III. Background

A. Definitions and objectives for NOAA testbeds and proving grounds

1. Testbeds

a. Definition and Purpose: A NOAA testbed is a working relationship for developmental testing, in a quasi-operational framework among researchers and operational scientists/experts (such as measurement specialists, forecasters, IT specialists) including partners in academia, the private sector and government agencies, aimed at solving operational problems or enhancing operations, in the context of user needs. A successful testbed involves physical assets as well as substantial commitments and partnerships.

b. What is tested: Advances to be considered include candidates for more effective observing systems, better use of data in forecasts, improved forecast models, and applications for improved services and information with demonstrated economic/public safety benefits.

c. Objectives:

Testbeds (TB) accelerate the translation of R&D findings into better operations, services, and decision-making. Outcomes from a testbed are capabilities that have been shown to work with operational systems and could include more effective observing systems, better use of data in forecasts, improved forecast models, and applications for improved services and information with demonstrated economic/public safety benefits. Successfully demonstrated testbed capabilities are ready for advanced pre-deployment testing, in a full simulation of real-time operational conditions, leading to “go/no go” deployment decisions.

2. Operations and Services Proving Ground:

a. Definition and Purpose: Operations and services proving grounds (proving grounds, or PG) are a framework for NOAA/NWS to conduct testing of advanced operations, services and science and technology capabilities that address the needs of both internal and external users. Successful testing demonstrates readiness to implement into operations.

b. What is tested: Capabilities to be tested in operational proving grounds have already passed developmental testing. Such capabilities include advanced observing systems, better use of

data in forecasts, improved forecast models, and applications for improved services and information with demonstrated economic/public safety benefits.

c. Objectives: Testing in real-time, in an operations-like setting to demonstrate achievement of performance metrics, including testing any workflow changes, needed for implementing in operations as well as end-to-end delivery of services. Performance metrics are defined for each candidate capability, in categories of: objective performance (e.g. accuracy/skill); subjective evaluations of utility (e.g. user feedback on balance positive); and production/engineering readiness (e.g. systems and communications reliability/security/backup, data retention). Performance criteria for objective and subjective evaluations by users internal to NWS include expected impacts to workflow and workload, except when advanced capabilities have no impact on workflow/workload; e.g. in the case of improvements to numeric quality of current operational guidance and tools. Successful pre-deployment testing is necessary for approval to implement into operations.

B. Guiding Principles for Recommendations

1. Guidelines are consistent with NOAA policies and plans.

The recommended roles and responsibilities, functions, and governance are consistent with, but do not supersede the following overarching NOAA plans and policies: NOAA and NWS Strategic Plans, NOAA Research Plans, S&T and Services Roadmaps; NOAA transition policy (NAO216-105), and related activities of the LOTM; NWS Operations and Services Improvement Policy (NWS directive 10-103) and NWS' Research and Innovation Transition Team.

NOAA research/laboratory reviews have called for additional clarity and emphasis on testing procedures as conducted in current testbeds and proving grounds, which have been established independently to address specific service improvements.

2. Guidelines adopt and promote best practices.

Recommended guidelines for roles and responsibilities, function and governance recognize the common goals and objectives in these facilities in adopting and promoting best practices.

Based on reviews of current testbeds and proving grounds (TB/PG) and existing charters and terms of reference for NOAA testbeds, the following best practices for operating procedures are incorporated into recommendations:

- a. Procedures for initiating, conducting, and successfully completing testing are clearly communicated, including entry criteria, risk reduction/adjustments and restarts, and metrics for success.
- b. TB/PG personnel assist in risk-reduction
- c. Real-time collaboration of R&D and operations partners is facilitated
- d. Testing includes appropriate participation of field operations personnel (e.g. forecasters)

3. Governance is transparent and follows consistent, but not identical procedures.

Common practices among testbeds and proving grounds that promote transparency include

executive steering committees, local execution/management, and appropriate stakeholder involvement. Executive committees have oversight for review of incoming proposed tests and results of testing activity. Local management oversees testing operations, to ensure rigorous testing and evaluation procedures are followed. Appropriate stakeholder involvement (from NOAA operational and research organizations, external) is encouraged.

IV. Recommended Roles and Responsibilities

A. NOAA participants

1. Host facilities:

- a. Provide Charter and/or Terms of Reference outlining scope and general procedures, including infrastructure requirements, availability of staff support for testing, etc. as described under Section VI.A.
- b. Provide Notice of opportunity to participate in testing
- c. Establish and lead management team, to oversee, support and facilitate testing operations. The management team designates a representative to serve on its executive oversight committee, and reports to this oversight committee on test selection and results.
- d. Are governed (Section VI) by the local management team, with an executive oversight committee
- e. Participate in NOAA-wide coordinating Testbed and Proving Ground coordinating committee

2. Research partners (outside host facility): Provide peer-reviewed capabilities, testing support

3. Operations partners (outside host facility): Provide statement of needs/requirements, testing support

B. External stakeholders/participants

- Respond to announcements of opportunities for testing advanced S&T to support operational mission requirements
- Participate in testing and evaluation of selected test capabilities

V. Functions

A. Testbeds

1. Objective: Controlled testing of peer-reviewed capabilities to determine if they can work with operational systems

2. Operations.

a. Provide announcements of opportunity for testing cycles, with the following recommended information:

- Scope and eligibility requirements, timing and funding availability/strategies
The announcements state the scope of activities to be tested, eligibility to participate, timing, and resources support (both internal: logistical, systems for testing, funding availability and any necessary external support). Eligibility constraints include demonstration of relevance to mission/service requirements, organizations/leads that may participate (e.g. whether restricted to NOAA), and additional requirements, for example in code standardization or demonstrated peer review.

- Criteria to be demonstrated in successful tests
The announcement also documents or refers to standard criteria for testing success, in areas of objective performance, subjective utility, and engineering criteria (which may include, but are not limited to reliability, interoperability, systems size/timing/IT constraints). Any additional, customized criteria for particular capabilities are included.
- Required documentation needed for consideration/selection
Forms and expected documentation (e.g. proposals), deadline, and the form of response.

b. Recommended project prioritization/selection strategies

Peer review is conducted to evaluate and recommend appropriate tests, on the basis of relevance to mission requirements, excellence (appropriateness of approach, expected outcome, related prior success), and cost effectiveness. The executive oversight committee (or its designate) approves recommendations for testing based on scope, chances of success and resource availability.

c. Facilitate testing

Testbed personnel assist/facilitate the testing to negotiate the most effective test plan/schedule, coordinate with any internal/external facilities needed, ensure that tests are conducted according to NOAA guidelines and procedures, and recommend and/or assist with any minor revisions needed to make most effective use of the testing time.

d. Reporting plans and results

The testbed Management Team, provides an annual operating test plan, based on the prioritization of proposed tests in section b., and reports the testing results on an annual basis. It is recommended that an annual summary of results be provided, via a convenient, low-overhead medium like a web-site.

B. Proving Grounds

1. Objective: Demonstration of operational readiness and performance criteria, including impacts to operational workload and workflow

2. Operations:

Operations and Services Proving Grounds operate in similar ways to testbeds, although the capabilities eligible for testing, testing environment, and outcomes, are different. Proving grounds provide an environment for controlled real-time testing of capabilities proven to work with operational systems, as well as workflow, workload options and impacts. There are options to link to capabilities in several testbeds, and to simulate capabilities in several collaborative operational settings. Recommended operations are therefore similar and the recommendations for testbeds, in Section A, are adopted, with the following differences.

a. Provide announcement of opportunity, as done by testbeds. Eligibility of capabilities for PG testing includes having passed testbed/developmental testing, and demonstrated impact on meeting operational requirements.

b. Recommended project prioritization/selection strategies, as done by testbeds.

c. Facilitate testing, as done by testbeds. In addition, following successful testing, PG personnel participate in review/approval processes for implementing into operations, e.g. “readiness reviews”.

d. Reporting plans and results, as done by testbeds.

VI. Recommended Governance

A. Charters/Terms of Reference Documentation

It is recommended that each TB/PG develop charters or terms of reference that specify their scope, operations (roles and responsibilities, and function/execution), resourcing framework (i.e., how they will be resourced) and governance. Authorities for these charters should be at the lowest level appropriate to the executive oversight committee (see below).

B. Individual testbeds and proving grounds

TB/PGs are locally managed, by management teams responsible for conducting and supporting testing operations, and reporting the results. It is recommended that executive committees (or boards) oversee the scope of activities, selection of tests, and review of quality of results.

1. TB/PG Management team

- a. Includes research and operational partners.
- b. Oversees testing operations, to ensure rigorous testing and evaluation procedures are followed.
- c. Prepares plans, programs resources and executes budget for supporting testing.
- d. Designates a representative to serve on Executive oversight committee.
- e. Reports to oversight committee on test selection and results
- f. Communicates annual summary results for the broader community; use of web-sites is encouraged.

2. Executive committees

- a. Includes executive representatives for research and operational partners, as well as a designate from the local management team.
- b. Oversees the scope of activities, testing selection, quality of results.
- c. Advises on the scope and priority of activities in view of strategic considerations, including funding availability.
- d. Recommends strategic priorities and facilitates/communicates funding availability for them
- e. Oversees selection of proposed tests. The committee may delegate the responsibility for review and selection to the management team.
- f. Ensures quality of tests and results is acceptable. The committee may delegate the responsibility to the management team.
- g. Reviews and approves recommendations for capabilities’ successful passing. The committee may delegate the responsibility to the management team.

C. Coordinating committee

To foster coordination, leverage related activities and increase communication among TB/OG and across NOAA, a coordination committee is proposed, with representatives from each TB/OG, and Line Office Focal points. A draft Charter is provided separately.