Summer 2014 Peak Performance Newsletter

Performance Branch, NWS Office of Climate, Water, and Weather Services, Silver Spring, MD

This Issue: August 2014

Expanded Level of Quality Control Coming to Verification 1
Service Assessment Program: NWS Director Approves Public Release of CO Floods Service Assessment 3
Fly...with Ointment: Introduction to Selecting Performance Measures 4
Did You Know? 5
Inappropriate GPRA Goals 7
Status of Service Assessment Actions 8
Contact Information 9

Expanded Level of Quality Control Coming to Verification

By Brent MacAloney, Performance Branch, NWS Headquarters, Silver Spring, Maryland

Those of you who have been visiting the Performance Management website for more than five or ten years are probably well aware that the Performance Branch would regularly release three or four new programs or major enhancements per year. Due to the staffing issues in recent years that impacted everyone in the NWS, the Performance Branch was forced into survival mode, dedicating most of its resources to maintaining what was currently running. Unfortunately though, not all programs could be maintained; some had to be discontinued or put in hibernation mode. The plan was to re-develop those in hibernation mode (once the staffing situation became more stable) into faster, more accurate, and easier to use programs. A major part of this development effort had to deal with better ways to quality control the data.

The original design of the verification systems often made it difficult for the Performance Branch staff members to omit or deal with questionable data from the verification dataset once it was in the system. For example, if a buoy went adrift and no one noticed for a few weeks, it was likely that the observations used to verify the forecast would be derived from a location unrepresentative of the location for which the forecast was made. In another example, construction occurring in the area of an ASOS location kicks-up dust and erroneously reduces visibilities, which then negatively impacts visibility verification. In both of these examples, depending upon how long the issue continued, a local office's verification score could be significantly affected.

To help the Performance Branch staff members better handle requests to deal with questionable

Continued on next page...

Expanded Level of Quality Control Coming to Verification – Continued from Page 1

data, a series of three tools were recently developed. These tools are scheduled to become operational over the next few months. The Marine Station Management Tool, River Station Management Tool, and Aviation/Public Station Management Tool were all developed with the ability to easily:

- Add new observation points to the verification program
- Modify those stations with erroneous or changing attributes
- Remove questionable data from the verification database
- Permanently remove observation points from inclusion in the verification routine which are no longer supported, maintained, or operational

These tools will work hand-in-hand with the various verification programs to allow for better overall administration capabilities. Using the case from above regarding the construction near an ASOS location, the Performance Branch staff would be able to access the Aviation/Public Station Management Tool and make a note in the data quality logs that the visibility data for the 2 week construction period is erroneous. The next time the TAF verification matching process ran, the program would see the note about the erroneous data and automatically remove it permanently from use in the verification database.

Over the next year or so, as these features are fully implemented into the verification systems, you should begin to notice a significant reduction in the amount of time needed to take action on any data quality requests.

The estimated timeline for release and implementation of this data quality tool is:

- Marine Wind/Wave Verification Stats on Demand – Fall 2014
- River Forecast Verification Stats on Demand – Spring / Summer 2015
- Point Forecast Matrix (PFM) Verification Stats on Demand – Summer /Fall 2015
- Aviation (TAF) Verification Stats on Demand – TBD

We hope that these efforts to better quality control the observational data used for verification will result in more accurate, meaningful, and useful statistics for managing your office's performance.◆

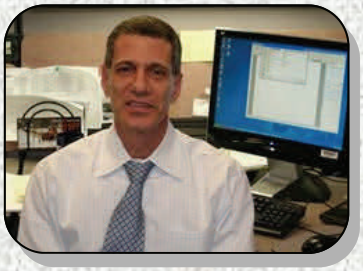
Forecaster Missing From Your Office Roster or Improperly Listed as Being at Your Office on the Performance Management Website?

It is important that everyone's profile is up to date with their current office location so that everyone can fully utilize tools like TAF Stats-on-Demand, NOEES, and our mailing lists.

New Employee—Have a new employee at the office with no account on the Performance Management website? They can register for an account here: http://goo.gl/zsR9gm

Employee Transfer—Users can always update their profile's location information to their current location by logging into the Performance Management website (http://goo.gl/Br2h9Q), clicking the "Account" button, and selecting the "Update your profile information" link.

Retired Employee—Email NWS.Verification@noaa.gov and we will remove this employee, freeing up their forecaster ID number for others to use.



Service Assessment Program

By Sal Romano, Performance Branch, NWS Headquarters

NWS Director Approves Public Release of Colorado Floods Service Assessment

The Record Front Range and Eastern Colorado Floods of September 11-17, 2013 Service Assessment report has been reviewed and signed by the National Weather Service (NWS) Director.

Colorado Flooding Assessment

The event began on September 9 as a large, slow-moving upper-level circulation that became nearly stationary over the Great Basin of the southwest United States. The broad flow around this system pulled plumes of tropical moisture northward from the Pacific Ocean off the western coast of Mexico and the western Gulf of Mexico. A frontal system became stationary along the Front Range of the Rockies while upslope easterly flow became established. Three episodes of torrential rainfall struck the Front Range from Fort Collins southward to Colorado Springs and east to Denver and Aurora, CO. The most intense events occurred on the nights of September 11-12, and September 15. Rainfall totals far exceeded existing records. In Boulder, 24-hour amounts exceeded 9 inches by the morning of September 12, nearly doubling the previous record. Event rainfall totals exceeded 17 inches in the climatologically favored upslope areas of the Front Range with a large area in eastern Colorado measuring 8 -17 inches of precipitation.

Eight people lost their lives as a direct result of the flooding. Most of these fatalities occurred during the height of flash floods on the evenings of September 11 and 12. Local authorities evacuated more than 18,000 people. There were approximately 19,000 homes and commercial buildings damaged with more than 1,500 destroyed. Authorities estimate the flooding damaged or destroyed almost 485 miles of roads and 50 bridges in the impacted counties.

Figure 1. Flash Flash flood damage in Jamestown, CO along Left Hand Creek (M. Leffingwell, Boulder Daily Camera)

The service assessment team completed on-site visits and telephone interviews. The team leaders presented the team's findings to NWS senior leadership on March 11, 2014. The service assessment team's report was reviewed and edited at various offices within NWS Headquarters, including the Office of the Assistant Administrator. The report was then released to the public on June 24, 2014.◆

Fly...with Ointment

By Beth McNulty, Performance Branch, NWS Headquarters

Introduction to Selecting Performance Measures

Performance measures don't just "happen," even if it seems that way. At some point someone had to make a decision about what to measure, and how to go about it. If you're the person responsible for developing or evaluating metrics, you get to select what to measure.

Let's assume, for this discussion, that you have a newly developed aviation product that you think will enhance a pilot's understanding of the forecast and overall situational awareness of flight hazards, local or otherwise. You need to find a metric to analyze how well the new product serves your target user. Today we'll introduce the concept of metric selection, and in later columns entertain a more in-depth analysis of metric creation.

The first question to ask is, "What do you want to measure?" Now, you may wonder why this is an important question. What you want to measure in large part stems from what you hope to learn from the metric. For instance, if you want to know the forecast accuracy in predicting changing conditions, you will need different data than if you simply wanted to know if the forecast and eventual observed conditions were reasonably similar. Once you have determined what you want to learn from the data, then you can determine how to do it. Next ask yourself, "Why?" You may find that answering "why" helps you to define the "what." For example, you need to know how many forecasts have no verifying observations (why). To determine this number you decide to compute or compare forecast valid periods with available observation time groups (what).

A third question is, "How will measuring this quantity help my forecast?" If you measure quantity "x" but do nothing else, is the measurement helping you to improve the forecast? Not likely. A corollary question is, "Will I apply the metric results to improving my forecast?" If the answer is "no," then don't collect the metric data; unless, of course, you like exercises in futility.

Suppose you design a metric to track the accuracy of the time of onset for an event, and you plan to use the results to determine what types of events you need more training on to improve your timing accuracy. One consideration is whether you are using a meteorological metric or an operational one (more on these in another article). It's important to focus the metric on the aspect of your product you hope to improve with time. Using a decision support metric to improve forecast accuracy is like trying to haul garden mulch with a hammer instead of a wheelbarrow. It doesn't work. The performance metric you develop can only help your product improve if you have focused on the appropriate aspect, and apply the results.

But there's another wrench to consider: how will you compute the metric, analyze the data, and apply the results? Usually some form of statistical analysis will be necessary to actually compute and analyze the metric results; the analysis may be

Continued on page 6...

Did You KNOW?

By Doug Young, Performance Branch,
NWS Headquarters

"*Did You Know*" is a column dedicated to making you aware of performance and evaluation-related information that you may not have known existed. This could be information within the pages of our website, techniques we are using to verify forecasts and warnings, performance trends, new research papers, training, customer feedback results, or a host of other topics of interest. We hope you find this column interesting and informative and, as always, we welcome your feedback!

NWS Customer Satisfaction Survey Archive

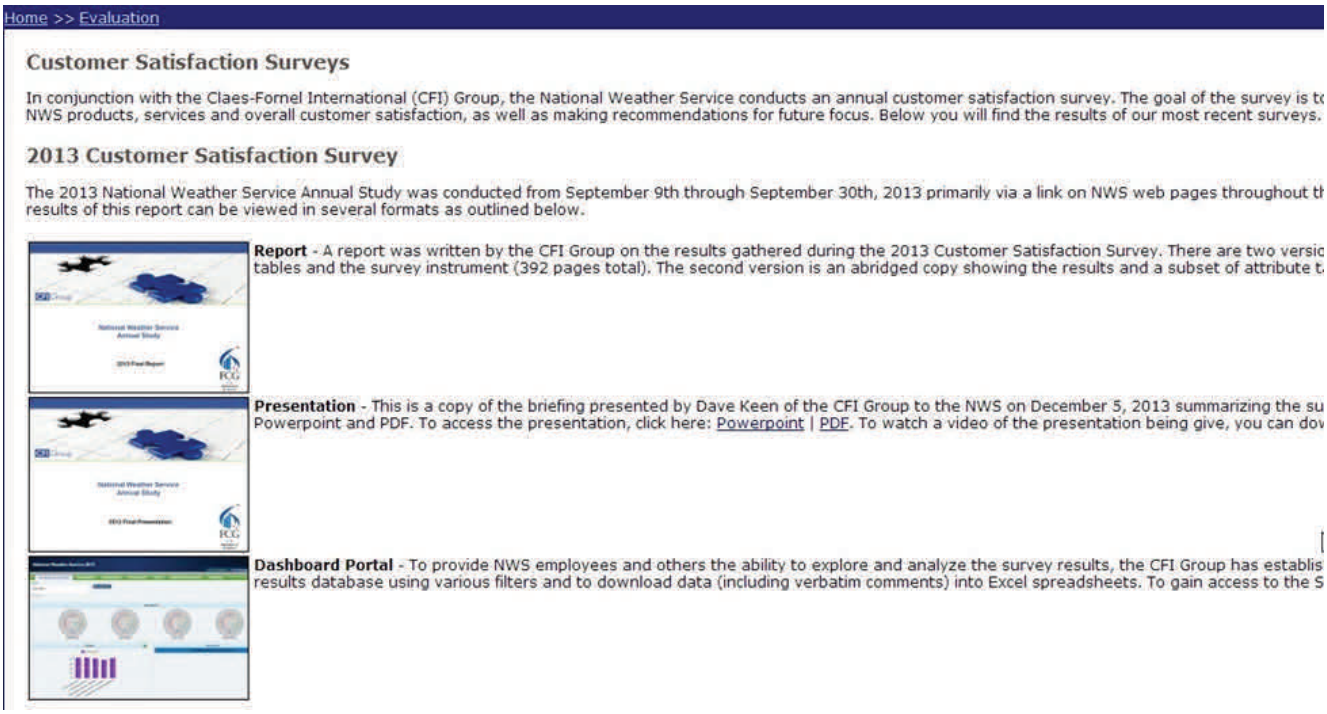
Did you know that the Performance Management website archives NWS Customer Satisfaction Surveys dating back to 2003?

To find previous customer satisfaction surveys, select "Evaluation" from the left-hand menu and click on "Customer Satisfaction" from the pop-out menu (Figure 2).

Figure 2. Selecting the "Service Assessments" page from the Performance Management home page.

You will now be on the Customer Satisfaction page (Figure 3). This page lists past customer satisfaction surveys. The 2010 through 2013 surveys are listed in reverse chronological order with links to abridged and full reports, presentations briefed by CFI representatives, and the Dashboard Portals for each year. In the case of the 2013 survey, the recorded briefing has been captured as well. For your convenience, the presentations are made available in both Powerpoint and Adobe PDF. Continued on next page…

Did You Know? – Continued from page 5


Home >> Evaluation

Customer Satisfaction Surveys

In conjunction with the Claes-Fornel International (CFI) Group, the National Weather Service conducts an annual customer satisfaction survey. The goal of the survey is to NWS products, services and overall customer satisfaction, as well as making recommendations for future focus. Below you will find the results of our most recent surveys.

2013 Customer Satisfaction Survey

The 2013 National Weather Service Annual Study was conducted from September 9th through September 30th, 2013 primarily via a link on NWS web pages throughout th results of this report can be viewed in several formats as outlined below.

Report - A report was written by the CFI Group on the results gathered during the 2013 Customer Satisfaction Survey. There are two versic tables and the survey instrument (392 pages total). The second version is an abridged copy showing the results and a subset of attribute t

Presentation - This is a copy of the briefing presented by Dave Keen of the CFI Group to the NWS on December 5, 2013 summarizing the su Powerpoint and PDF. To access the presentation, click here: Powerpoint | PDF. To watch a video of the presentation being give, you can dov

Dashboard Portal - To provide NWS employees and others the ability to explore and analyze the survey results, the CFI Group has establis results database using various filters and to download data (including verbatim comments) into Excel spreadsheets. To gain access to the S

Figure 3. Active and Previous Service Assessments web page displaying all active service assessment teams and previous service assessments.

Prior to 2010, customer satisfaction surveys were conducted on an ad hoc basis for different user groups and services, depending upon funding and need. These surveys can be found by scrolling to the bottom of the page. The surveys range in years from 2003 through 2010 and are listed in reverse chronological order within each program area. Both the reports and presentations are available in Adobe PDF.

A wealth of information, including verbatim user comments, is available within these past surveys and via the dashboard portals. The user feedback is fertile ground for research projects—some of it may have never been examined.

If you have any questions about customer satisfaction surveys, please contact Sal Romano (Salvatore.Romano@noaa.gov) or Doug Young (Douglas.Young@noaa.gov).◆

Fly,,, with Ointment – Continued from page 4

simple (i.e., average or mean, or percentages) or complex (i.e., everything else).

Finally, even though you anticipate using the metric to improve forecasts, what if there turns out to be a better way to measure performance? Do you continue to compute the first metric "just because it's there" (like Mt. Everest)? Is it necessary to duplicate your efforts in measuring performance if you have a better metric, just to maintain an established baseline?

Next: Specific discussion of representative classes of performance metrics.◆

Inappropriate GPRA Goals

By Chuck Kluepfel, Performance Branch, NWS Headquarters, Silver Spring, Maryland

The NWS Headquarters Performance Branch recently received a question from a field office about why NWS-wide (national) GPRA goals are often applied to individual NWS regions and sometimes to individual field offices.

Our answer is that, while this is often done throughout the NWS, the Performance Branch does not support, nor encourage this practice because some forecast elements have huge differences in forecast performance from one part of the country to the other. Therefore, we encourage each region and each field office to develop separate goals for each GPRA element. These goals should be challenging, yet achievable. When developing these goals, several years of data should be used to establish meaningful baselines.

Example 1: For fiscal year (FY) 13, Alaska region had a marine wind speed percent correct value equal to 55. In Southern Region, that score was 85, and the NWS-wide score was 76 (the NWS-wide goal that year was 74). It is evident from just from one year of data that a huge variation in forecast difficulty exists between these two regions. One year of data does not provide much of a baseline for developing any goal so here are six consecutive years of Alaska Region scores, starting with FY08: 56, 57, 55, 53, 52, and 55. Here are the Southern Region scores for the same six years: 82, 84, 85, 86, 85, and 85. The NWS-wide goal of 74 should not be applied to either of these separate regions. Alaska isn't likely to reach 74 anytime soon, and 74 is way too low of a goal to use for Southern Region.

Similarly, climate can also vary substantially across a single NWS region, so it is also unwise to apply a regional goal to one of its local field offices.

Example 2: In FY13, the tornado warning false alarm ratio (FAR) for a given field office was 57. That office resides in a NWS region that had a FAR equal to 53 that year, and the NWS-wide FAR was 74 (the NWS-wide goal was 72). The NWS-wide goal was reasonable when applied to the NWS as a whole, but 72 would not be a challenging goal for that field office or the region within which it resides.

Concluding Remark: Sometimes extreme weather variations from one year to another can produce wild fluctuations in the scores from one year to another. Returning to example 2, the respective tornado warning FARs for that field office from FY08 thru FY13 were 71, 58, 81, 74, 78, and 57. This high inter-annual variability illustrates how any goal needs to be based upon several years of data, and all interpretations of performance metrics need to be done in terms of the weather actually experienced during the year being interpreted. The NWS typically takes a beating on tornado warning verification statistics during "dry years" with relatively few organized tornado outbreaks. A similar thing happens with regard to flood warnings and the forecasting of instrument flight rule (IFR) conditions. If forecasters receive relatively few opportunities to accurately predict adverse weather, their scores tend to suffer. It is difficult to "shine" when you get so few opportunities.◆

Status of Service Assessment Actions

Summary – August 2014

- **Currently, there are 455 total actions *from Open Events*.**
- **128 actions remain open.**
- **327 actions are closed.**
- **There were no active service assessments as of the publication of this newsletter**

Open Service Assessments

- **Colorado Flooding of September 11-17, 2013**
  Released June 24, 2014
  26 Total Actions, 3 (12%) Closed Actions
  22 (88%) Open Actions

- **May 2013 Oklahoma Tornadoes and Flash Flooding**
  Released March 21, 2014
  29 Total Actions, 6 (21%) Closed Actions
  23 (79%) Open Actions

- **Hurricane and Post-Tropical Cyclone Sandy**
  Released May 5, 2013
  25 Total Actions, 10 (40%) Closed Actions
  15 (60%) Open Actions

- **Remnants of Tropical Storm Lee and the Susquehanna River Basin Flooding of September 6-10, 2011 (Regional Service Assessment)**
  Released July 26, 2012
  11 Total Actions, 7 (64%) Closed Actions
  4 (36%) Open Actions

- **Historic Derecho of June 29, 2012**
  Released February 05, 2013
  14 Total Actions, 4 (29%) Closed Actions
  10 (71%) Open Actions

- **The Missouri/Souris River Floods of May – August 2011 (Regional Service Assessment**
  Released June 05, 2012
  29 Total Actions, 22 (76%) Closed Actions
  7 (24%) Open Actions

- **May 22, 2011 Joplin Tornado (Regional Service Assessment)**
  Released September 20, 2011
  16 Total Actions, 12 (75%) Closed Actions
  4 (25%) Open Actions

- **Hurricane Irene in August 2011**
  Released October 05, 2012
  94 Total Actions, 65 (69%) Closed Actions
  29 (31%) Open Actions

- **Spring 2011 Mississippi River Floods**
  Released April 11, 2012
  31 Total Actions, 24 (77%) Closed Actions
  7 (23%) Open Actions

- **The Historic Tornado Outbreaks of April 2011**
  Released December 19, 2011
  32 Total Actions, 29 (91%) Closed Actions
  3 (9%) Open Actions

- **Record Floods of Greater Nashville: Including Flooding in Middle Tennessee and Western Kentucky, May 1-4, 2010**
  Released January 12, 2011
  17 Total Actions, 16 (94%) Closed Actions
  1(6%) Open Action

- **South Pacific Basin Tsunami of September 29-30, 2009**
  Released June 04, 2010
  131 Total Actions, 128 (98%) Closed Actions
  3 (2%) Open Action

Most Recent Closed Events (all actions completed)

- **Washington, D.C. High-Impact, Convective Winter Weather Event of January 26, 2011**
  Released April 01, 2011
  6 Total Actions,- Closed

- **Southeast US Flooding of September 18-23, 2009**
  Released May 28, 2010
  29 Total Actions - Closed

- **Mount Redoubt Eruptions of March - April 2009 -**
  Released March 23, 2010
  17 Total Action - Closed

- **Central US Flooding of June 2008**
  Released February 03, 2010
  34 Total Actions - Closed ◆

Contributors to this Summer 2014 edition of Peak Performance include . . .

Chuck Kluepfel
Contributor
Performance Branch, NWS Headquarters
Forecast Verification
Charles.Kluepfel@noaa.gov

Sal Romano
Contributor
Performance Branch, NWS Headquarters
Service Assessment and Evaluation
Salvatore.Romano@noaa.gov

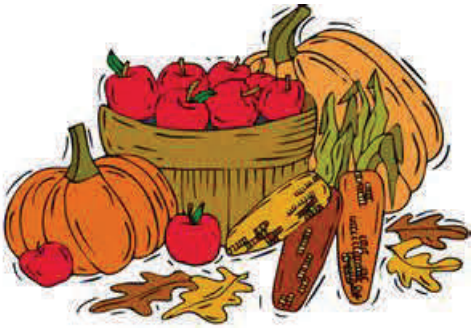
Brent MacAloney
Contributor
Performance Branch, NWS Headquarters
Warning Verification
Brent.Macaloney@noaa.gov

Freda Walters
Co-Editor and Designer
Performance Branch, NWS Headquarters
Service Assessment and Evaluation
Alfreda.Walters@noaa.gov

Beth McNulty
Performance Branch, NWS Headquarters
Aviation Performance and Verification
Beth.Mcnulty@noaa.gov

Doug Young
Contributor and Editor
Performance Branch Chief, NWS Headquarters
Douglas.Young@noaa.gov

Questions and comments on this publication should be directed to Freda Walters.



Please consider contributing to our next issue of Peak Performance: Fall 2014

Articles Due September 30, 2014