Research to Operations Case Study: The Forecast Evaluation Tool
-- Holly Hartmann, CLIMAS, University of Arizona

Early work in the CLIMAS project showed that stakeholders across many sectors faced several barriers in using seasonal climate forecasts. Key barriers included poor understanding of how to correctly interpret the products, poor ability to place the forecasts in the context of historical or recent conditions, and poor understanding of the skill of the forecasts, or even how to assess forecast quality (Hartmann et al., 2002a). At the time, the Climate Prediction Center (CPC) provided only a single verification score for the entire nation (Hartmann et al., 2002b) and users had to print and archive hardcopy maps before they were replaced on the CPC website by new outlooks. Further, climate observations were reported in different units and time steps by the National Climatic Data Center (NCDC).

The Forecast Evaluation Tool (FET) was designed to remove these barriers for users. The FET provides a tutorial on interpreting the outlooks and a searchable archive of the original climate outlook maps. The FET also includes a track showing recent seasonal climate observations and the official 30-year seasonal climatology, using the same terminology as the climate outlooks. Finally, the FET allows users to quantitatively assess forecast performance customized to focus on the specific seasons, forecast lead times, variables, and verification metrics that matter most to their decisions. The opportunity to consider increasingly sophisticated performance metrics helps build user capacity for working with probabilistic forecasts, and users can 'drill down' into their results to see the forecast and observation data used to calculate their chosen scores.

The FET was developed through extensive interactions with users, from discussions testing the understandability and relevance of mockups, to usability assessment workshops where observers documented web interface issues by watching individual users as they attempted to perform directed analyses. Some components of the FET were even designed by users. Since 2001, the FET has gone through three major revisions, with substantial software code rewrites to implement additional functionalities and incorporate advancing software technologies. Even minor revisions, e.g., to accommodate changes in web browsers, are thoroughly tested to ensure they work across the major operating systems (Linux, Windows, Apple) and multiple browsers (Explorer, Firefox, Mozilla, Opera, etc.).

The FET was developed without commitment to a specific 'business model' for sustained operations and maintenance. A NASA program that contributed funding toward FET development, and the RISA program manager at the time (H. Hill) encouraged consideration of privatization options (e.g., subscription access through the University of Arizona, transfer to a private company such as the Weather Channel). However, a market assessment and discussions with stakeholders revealed resistance to privatization and a preference for equitable access through a public agency.

Under the NOAA Climate Transitions Program (NCTP), we had several highly rated but unfunded proposals to transfer the FET to NCDC. The NCDC was seen as having sufficient computer resources (physical and human) to make a transfer practical, a reliable record of archiving data (forecasts are data about the state of science and forecasting), being independent from any sensitivity about computed

forecast skill scores, and being institutionally able to encompass a variety of seasonal outlooks over time (e.g., international climate outlooks from the International Research Institute for Climate and Society, water supply outlooks from the NWS River Forecast Centers). With the termination of the NCTP before we could devise a successful proposal, we were left with no formal mechanism for transfer of the FET to NOAA operations.

In 2004, the Climate Services Division (CSD), within the Office of Climate, Water, and Weather Services (OCWWS), provided a small contract to explore transferring the tool to their office. Our assessment found that NWS policies prohibited dynamic initiation of computations by users. Citing security concerns, the NWS Chief Information Officer (CIO) required all possible computations and graphics to be generated beforehand, allowing users only to display selected results. When the transfer to the NWS proved impractical, CSD provided another small contract (~$7500/year after university indirect costs) to maintain and expand the FET. Because that amount isn't sufficient to sustain FET operations and expansion, we continue to support the FET through CLIMAS funding as well.

In 2007, we began collaborating with the CPC Operations Branch (OB) on a Climate Test Bed project aimed at (1) developing dynamic forecast products that users can build to match their cognitive frameworks and decision needs, (2) connecting more CPC forecasts to the FET, and (3) implementing field-testing of communication effectiveness of new forecast products. It did not include transfer of the FET to CPC. In the process of working on the first objective, we were able to get NWS CIO approval for user-initiated dynamic computations. And in the process of working on the second objective, CPC-OB decided they wanted to transfer the FET to CPC, to jumpstart internal verification enhancement and build web applications programming capacity.

Subsequently, we shifted our CTB project to focus on the transfer of the FET to CPC. We worked with the NCEP Web Operations Center (WOC) and CPC to develop a collaborative software development framework and process, which includes software version control tools and task tracking (or bug reporting) tools. This allows software code to reside at CPC, while also allowing external groups to modify code, enabling our group to support CPC staff as they learned the FET, and providing a clear pathway for transition of ongoing research into operations. Our success in transferring collaborative software development to CPC has been transformative for CPC, changing how they handle all their software development. For this outcome alone, CPC recently rated our CTB project as 2.97, with 3 being the maximum rating.

We transferred the FET code onto a CPC server and got it running reliably as an internal prototype, but work stopped when CPC-OB decided to develop a new Forecast Verification Tool independent from FET code. In simple terms, the CPC has said they are too limited in time, skill, and experience to entrain such a large web application, and prefer to develop a limited application from scratch. CPC's decision was unilateral rather than collaborative, and we are still struggling to appreciate their rationale.

On the surface, CPC's decision seems illogical, but CPC-OB has given a variety of reasons for their choice. An early comment was that they philosophically do not want to have code 'dumped' on them; they are bothered that other groups get to have 'the fun' of the research and development while they are left to

'take care of everything'. (CPC-OB made similar comments in reference to transfer of the local 3-month temperature outlook [L3MTO] product and web application from the CSD.) CPC-OB has begun to emphasize that they also do research and development, not just operations.

The size of the FET code was seen by CPC-OB as too large, based on the naïve metric of number of lines of code. CPC-OB also considered the FET as 'old school' because it doesn't make use of software application frameworks, which can make coding easier but pose many risks for overall software project success. Further, the FET code was managed by less user-friendly version control and task tracking tools than the ones ultimately selected by the CPC-OB, although the FET is well-documented and could easily be integrated into another management system.

The CPC-OB requires that operational software run without modification or recompiling for months to years, although the WOC has no such requirements and approved routine recompiling of FET code. Self-updating can be enhanced in the FET, but large web applications benefit from frequent recompiling to accommodate updates in operating systems and browsers. We think CPC will be confronted with this situation as their tool expands in functionality and browsers evolve, but are avoiding the issue by limiting the functionality and interface choices in their verification tool.

Our perspective is that CPC has little interest in entraining externally developed software applications, and limited capacity to do so with sufficient proficiency to correct bugs or extend functionalities. CPC staffing restrictions preclude hiring computer scientists, and programming must be done by meteorologists and other physical scientists that lack the education and experience of software engineers. While CPC employees are dedicated and committed to learning new technologies, they simply don't have the required experience to entrain large applications developed elsewhere or design large applications with intuitive interface components preferred by users. The L3MTO transition experience has been similar to the FET, with CPC supporting only routine computations and a private contractor still maintaining the web application.

We continue to collaborate with the CPC on software projects, but only for new code and new products, not for transfer of existing technology. We are using the ongoing collaboration to more clearly understand CPC, NCEP, and NOAA software standards and develop CPC capacity for advanced applications.