

Kevin Smith (WFO Paducah) Interviews Chuck Kluepfel on TAF Verification - Past and Future

Kevin Smith at the Paducah, Kentucky Weather Forecast Office (WFO) asked us some very thoughtful, probing questions about Terminal Aerodrome Forecasts (TAF) verification. They take us back to when we started building the system over five years ago and also take us forward to where we might want to be 5 or 10 years into the future. Here are Kevin's questions and my responses:

Question 1: Although most NWS field offices do not have inexpensive access to five minute Automated Surface Observing System (ASOS) data (usually requires telephone charges for remote sites), how did you decide whether to use 5- or 1-minute data for TAF verification?

Answer: We are not yet able do true 5-minute verification of the TAFs, but we have employed a "poor person's" version of it. We only store the hourly and special Aviation Routine Weather Reports (METARs) that are transmitted to the world. We then look for the latest observation every 5 minutes and use it for verification. For example, if an ASOS didn't transmit any specials during a given hour, we would end up verifying the same hourly observation twelve times for that hour. However, for precipitation and thunderstorm verification, we also store and use the METAR begin- and end-time remarks, which adds some critical temporal resolution to precipitation and thunderstorm events when the weather changes, but special criteria are not met. Having said all of this, please keep in mind, most people are interested in categorical verification of ceilings and visibilities. The categories were designed with operational considerations in mind and, in most cases, are similar (but not identical) to the METAR special criteria. We have never considered using 1-minute data because we didn't think we would get the bang for the buck from 1-minute data. The management of 5-minute data has been challenging enough to the system. If you were in the NWS a couple years ago and remember how long it used to take to get your requests for TAF verification reports answered, you know what I'm talking about. Through more efficient programming and faster servers, the delivery time of your longer requests has been reduced from many hours to minutes.

Currently, the acquisition of true 5-minute ASOS data would require us to dial into over 500 ASOS sites daily. Such a technological feat would drastically increase the number of "points of failure" for data acquisition and, in the end, could cost us as much data as we would gain. However, the times are changing, and past limitations need not continue into the future. I just learned that the Meteorological Assimilation Data Ingest System (MADIS) system in Boulder, Colorado, is now receiving 5-minute ASOS data centrally from individual ASOS units in the Northeast US and the western Gulf coast area. Consequently, the possibility for a TAF verification enhancement that

incorporates true 5-minute data in the future is not out of the question. We will seriously consider switching to using true 5-minute data once MADIS gets access to *all* METARs in the entire country. I don't know if this would open up any legal issues since most of the aviation community only receives the transmitted hourly and special observations in real time, but we will certainly look into the matter. We may also need to add some memory and horsepower to the system to keep it running efficiently.

Question 2: Does the NWS receive Surface Weather Observation Stations (AWOS) and Automated Weather Information Systems (AWIS), or other FAA-supported observation system on a 5-minute basis as well for verification? I know at our own office, we forecast for one TAF site that does not use an ASOS-based observation setup.

Answer: We are not aware of any 5-minute data available from AWOS sites. Nonetheless, we use AWOS data for TAF verification. Most AWOS units only provide a new observation every 20 minutes, regardless of whether or not special criteria were met. No specials are issued between each 20-minute report. Just as with the ASOS data, we take the latest observation every 5 minutes and use it. Therefore, AWOS data are inferior to ASOS for verification, but they are better than nothing.

Question 3: Although there are many caveats and cautions in using digitized radar, satellite, or other supplementary observing systems within the 5 mile radius of a verifying TAF location, are there any plans in the future to incorporate this data into the TAF verification process, given some of the technical and statistical limitations of a single point ASOS site?

Answer: We've not made any attempts in this area. However, the thunderstorms included in the ASOS observations come directly from thunderstorm detection equipment. Terminals that do not have thunderstorm detection equipment contain the remark, *Thunderstorm information not available* (TSNO) and are not verified for thunderstorms. Such terminals are identified by the TSNO remark near the end of the observation. It is important to remember that we are only tasked with verifying TAFs at the terminal for which they are written. Technically, this means within a 5 mile radius of the center of the airfield. We do not verify any of the vicinity (VC) forecasts (5 to 10 statute miles from the center of the field) or utilize any of the VC observation remarks. They get stored in our database, but we do not use them for verification.

Question 4: Given the 5-minute verification profile used for prevailing and/or Temporary/Probability (TEMPO/PROBxx) groups, are there any plans to allow

the use of 15 minute resolution start/end times in TEMPO/PROBxx groups? The use of these higher resolutions start times in prevailing forecasts have always helped to reduce extraneous time periods in which a weather element occurrence may be unjustified.

Answer: Changing the TAF code would require interagency (DOC and FAA) and international coordination. Given our ability to forecast the kinds of things that go into TEMPO and PROB groups, it would probably be difficult to make a convincing case that a forecast precision of less than one hour for these groups would benefit flight operations and, therefore, add value to the TAF. In verification, we mainly use the "5-minute data" to look back and forth through the observation record to perform the variability test. Yes, we still do the verification every 5 minutes, but the rules for determining the operational impact forecast are very forecaster friendly when you pass the variability test for a large percentage of the TEMPO valid period. For example, if you forecast TEMPO thunderstorms, and your TEMPO passes the variability test for thunderstorms for the entire valid period of the TEMPO, you do not (and *should* not) need to receive thunderstorms for the entire valid period to get a perfect verification score for the thunderstorms in that TEMPO. A brief period of thunderstorms during the valid period will do just fine for verification. Plus, that is what is preferable. You do not want it to thunder during an entire TEMPO valid period; for such cases should have had thunderstorms in the prevailing (FM group) forecast! Not surprisingly, you risk failing the variability test if it thunders too long. For greater detail, see the new TAF training module titled, *Interpretation of TAF Verification Statistics, the Impact of TEMPO Forecasts*.

Question 5: I frequently witness ground fog rise up past an ASOS sensor. It either dissipates or becomes a ceiling in a one hour time frame. The ASOS visibility may move through each of the visibility categories from VFR to VLIFR back to VFR in 60-90 minute time period. From a verification standpoint, is there any strategy in the use of a one hour long TEMPO group that would not harm the aviation community, but would not hurt the TAF verification as well?

Answer: TEMPOs are a help to the aviation community if the variability in the TAF really does occur. You can easily monitor the quality of your TEMPO forecasts through the TEMPO verification statistics. They tell you whether and for how long you passed the variability test (the same test used for the operational impact forecast). They also tell you how much of your TEMPO forecast valid time should have been forecast as prevailing conditions with FM groups. You will also see the percentage of time that your TEMPOs hurt flight planning operations (by crying wolf) and the percentage of time that they were benign (wrong, but inconsequential toward flight planning). From the perspective of the operational impact forecast (OIF), the verification rules for TEMPOs are very generous if you pass the variability test, but that's a very

important "if." If you fail the variability test when a TEMPO is in effect, you get stuck with the most pessimistic forecast (prevailing versus TEMPO) for verification.

Question 6: I know that a "VC" is not critical for verification, but how does the NWS handle partial obscuration (BC/MI/etc...) of fog below 5/8's of a mile (FG). I believe this could be valid concern to some pilots, but does it have any positive impact on TAF verification, especially if you have a wide difference between overall flight categories of visibility VFR vs. IFR/LIFR.?

Answer: We do not do anything with partial obscurations. We use the cloud height, vertical visibility (the ceiling during *total* obscurations), and prevailing visibilities reported in the METAR. Elements forecasted or reported with a prefix of VC are not verified or used for verification because we only verify for the terminal, which is defined as within a 5 statute mile radius of the tower.

Comments:

Although we are encouraged to amend frequently, we are also encouraged not to "chase observations" and utilize flight categories when possible. I noticed in your verification profile in ceiling and visibility groups for the OIF utilizes the worst case scenario regardless of the flight category of the individual groups. I understand the reasoning, but from a verification scenario, would it not be prudent to forecast the overall flight category that has the greatest probability of occurrence for the longest duration in a FM/TEMPO/PROBxx group, even if it is well above the worst condition (assuming this condition has a high probability of occurrence, but short duration). If this were the case, we would never have to worry about being "pinched" by verification stats that show 1/4 mile visibility for 10 minutes...versus 3 mile visibility for 50 minutes over a two hour TEMPO period. However...if you are landing during a peak travel window in 1/4 mile visibility...you might think different.

Response: The "pinching" you describe doesn't seem very likely unless you use a lot of PROB40 groups, which are not allowed in the first 9 hours of the TAF and may not be used unless the PROB40 forecast includes a forecast for precipitation or thunderstorms. An operational impact forecast (OIF) associated with PROB40 forecast conditions will cause the OIF to be defined by the worse case forecast scenario (prevailing versus PROB40). This may appear punitive, and probably is. Be assured, however, that the relative use of PROB40 groups in TAFs these days is so infrequent that their impact on verification scores is slim to none.

Under all other circumstances, the OIF is defined a lot differently. If no PROB40 or TEMPO forecasts are in effect, the OIF is defined as the only thing left in a TAF—the prevailing forecast. If a TEMPO forecast is in effect, the OIF is strongly influenced by the variability test, and the worse case forecast scenario is only used to define the OIF when the TAF *failed* the variability test. If the variability test passes, the OIF is defined as the forecast (TEMPO or prevailing) categorically equal to or closest to the observation.