2020 Verification workshop presentation notes

9 Nov 2020 – Session 1: Opening + Error tracking (Discussions with presenters)

Barbara Casati (BC): A 20-year Journey of Forecast Verification Research

Discussion

- Are there any examples of verification of impact-based forecasts?
  - BC: That idea is more of concern for the High-Impact Weather (HIW) project, where they are looking into different approaches. It also is related to the value chain project being undertaken by HIW.
  - Participant: It is still early days for that kind of work; we are not there yet
- Our verification issues are similar to those faced by other communities (e.g., medicine). How similar are our methods? Are there things we can learn from these other groups?
  - BC: There are very many communities (in atmospheric and related sciences) doing similar work, and we are starting to connect with them (e.g., sea ice verification via the Polar Prediction Project). Other communities are less advanced and we are trying to reach out to them (e.g., hydrology, agriculture).
    - We don't know how or if they are implementing advances to the methods we are using.
    - We need a dialog between communities, and have been approached by some other groups.
    - There are many commonalities with the medical community.
- What about observation uncertainty – how are we dealing with it?
  - BC:
    - Can we learn from the DA community?
    - Linear error vs. squared error – when looking at linear error, observation error is less important
- BC: Between representativeness and observation error, representativeness has the biggest impact

Linus Magnusson (LM): Understanding medium-range forecast errors from a synoptic-dynamic perspective

Discussion

- A sophisticated center like ECMWF has many great tools. What is a good place for other groups to start?
  - LM: A good place to start would be to make use of data such as the TIGGE archive to learn about errors. It's important to see if the error evolution is captured appropriately.
- Is blocking onset the main cause of these busts?
  - LM: It's not a coincidence that blocking error often creates these kinds of large errors
- Ensemble members are not always consistent; how do you select good and bad members?

- o LM: Explore and tune the methods used to generate the members. Pragmatic approach – choose the best performing 5 members as "good" and the worst performing 5 members as "bad"
- Comment: These are very interesting techniques. Are they automated at ECMWF? Or do you look at each bust and try to identify issues?
  - o LM: So far it is an *ad hoc* approach that involves manual analysis. It is likely that artificial intelligence and machine learning techniques could help to make the process more automatic

Shannon Shields (SS): Analysis of Regional Sector Low-Skill Events in Recent Operational GFS Forecasts (NWS)

Discussion

- How is modeling changing as a result of these kinds of analyses?
  - o SS: Looking at these cases, relating low-skill events to recent events and outcomes, leads to understanding of causes of these events and helps improve understanding of model errors. The process also helps identify cases for sensitivity studies that lead to greater understanding.

General discussion and comments (from chat and raised hands)

- Is SEEPS being used for any other variable beside precipitation?
  - o B. Casati: Local climatology is used, so it could be a good option for other variables. SEEPS also might be a good option for understanding model errors
- A question for Shannon… Has NCEP's work been published?
  - o Shannon: A co-author is working on a publication on this as well as use of criteria defined by Rodwell in application to GEFS
- The methods described have generally been developed and applied for medium range forecasts. Has anything similar been done for shorter time scales (e.g., for precip, wind errors)?
  - o No one had a response to this specific question, but it was suggested that it is being investigated in the context of longer-range forecasts like S2S. The problem is more complex due to more processes being important at shorter time scales.
- Should large error information be used to select observation locations?
  - o Response: The presenters have some experience tracking errors back to specific observations as an aid in understanding large or important model errors
- Looking at results across a season it is clear that the number of bad cases/obs correlates well with overall scores. Thus, busts contribute a lot to overall scores and are an important aspect to look at!
- Looking at stratospheric impacts could also be of interest.
- A participant commented that it is important for verification results and information to be accessible to all users. This is an important issue for SERA as well.