ICAP Workshop on Aerosol Forecast Verification Oxford, UK 30 September – 1 October

Rationale

- This second meeting of the International Cooperative for Aerosol Prediction (ICAP) will focus on defining common verification measures that can be used to assess model performance in general terms.

- The first meeting, held in Monterey, CA, in April 2010, centered on issues of Aerosol Observability requirements for global operational aerosol models and was extremely successful in establishing links between the ICAP members and the data providers.

- A meeting on verification was deemed necessary as there is yet to be agreed upon measures of model skill for which forecasting capabilities and data assimilation impact can be adequately assessed.

- The Numerical Weather Prediction (NWP) community has over 30 years of experience with the definition of internationally accepted verification measures to quantify model skill, while the aerosol forecasting community is still at the stage of defining suitable "scores" that can fulfil a similar function. We can learn from its long-established experience.

- Possible topics for subsequent meetings need also to be discussed. These could include aerosol predictability, model development, ensemble aerosol forecasting, and multi-model ensembles.

Why do we care about aerosol forecast verification?

- Monitor progress in forecasting in this field

- Derive a benchmark so that systems can be appropriately assessed

- Improve forecast quality through better understanding of forecast errors

- Assess the impact of observations in the model

Overarching goals and hoped-for outcomes

a) Provide baseline information on how operational centres currently verify NWP models, the climate community verifies climate models (with a focus on aerosol), and the air quality community verifies (regional) models.

b) Compare how aerosol forecast verification is currently performed at different centres, identifying common issues faced by all centres (verification measures/scores currently adopted, independent data used, etc.)

c) Provide a venue for the discussion of suitable scores based on both analysis data and independent (non-assimilated) observations.

d) Focus on issues faced in verification of OPERATIONAL forecasts

e) Plan for advancing research on relevant verification issues.

Proposed Questions

Questions for general discussion

a) What do we want to "get better at"? Or, what forecast variables do we want to focus our verification on?

b) What are appropriate aerosol forecast "scores" ? I.e., is the equivalent of the 500hPa anomaly what we want?

c) What issues are specific to aerosol forecasts and what existing measures can be borrowed from standard NWP forecast verification?

d) For verification purposes, should aerosols be looked at as "homogeneous" variables (such as temperature) or as "inhomogeneous" variables (such as precipitation)?

e) Is it useful to see aerosol and meteorological NWP variables as two separate entities?

f) What is the role of external validation outside of operational centres? How can the research community contribute?

g) Is there the need for common verification measures? Is the community ready to specify some common formats for data distribution and display?

Discussion on verification using ground-based observations

a) How well do the ground-based measurements reflect the aerosol distribution surrounding the monitoring station?

b) What is the spatial representativeness of these observations? And what does it depend on? (Topography, proximity to emission sources, lifetime of chemical species under evaluation, …)

c) Should we verify the model with gridded data or with station data?

d) Should ground-based observations used as "ground truth" or should we also include errors in our comparisons with model data?

e) Are the existing ground-based observing networks adequate to meet the verification needs of the aerosol forecasting community? What type of improvements can be made to ensure that?

f) What are the issues faced by data providers of ground-based observations in providing observations for an operational forecast verification?

Discussion on verification using satellite observations

a) What is the role of satellite data for routine verification of aerosol forecasts?

b) Is there a reliable aerosol climatology based on available satellite data?

c) How do we verify with uncertain data?

d) What is the role of research and pathfinder missions in supporting the verification activities?

e) Are the existing suite of available satellite observations adequate to meet the verification needs of the aerosol forecasting community? What type of improvements can we propose for future satellite missions?

Discussion on aerosol "scores"

a) What are appropriate aerosol forecast "scores"?

b) What should determine how we choose the "best" scores for the aerosol forecasts?

c) What issues are specific to aerosol forecasts and what existing measures can be borrowed from standard NWP forecast verification?

d) ....

Discussion on model-based verification

a) Does it make sense to verify against own analysis?

b) How much do we trust that our analysis provide a good base for forecast verification?

c) Should we also use other model references in our verification? For example, should we compare with persistence or a consensus "climatology" or the model climate?

d) Should we use reanalysis as a benchmark to check "real" improvements in forecast capabilities (i.e. improvements that are not related to predictability)?

Summary of the discussion

Discussion on aerosol "scores"

a) What are appropriate aerosol forecast "scores"?

b) What should determine how we choose the "best" scores for the aerosol forecasts?

Scores for whom? What users exist for forecasted aerosol ?

Public&Industry / Surface and tropospheric concentration levels (PM alert, ground and air traffic perturbation, Industry production , solar energy sector, education)

Science / Radiative effects (weather prediction model, forcing fields and forcing estimate, visibility, satellite retrieval, ecosystem modelling)

Science / Composition and aerosol transport (Source receptor relationships; Emission trend analysis, Field campaign planning)

c) What issues are specific to aerosol forecasts and what existing measures can be borrowed from standard NWP forecast verification?

- *Is there anything specific ???*
- Specific Data? Specific dispersion phenomena? Specific users?
- Is Persistence Test useful? Any useful to use climatology for shortlived aerosols?
- Data sharing "difficult"
- Special scientists - attention: chemists included.

What issues are specific to aerosol forecasts and what existing measures can be borrowed from standard NWP forecast verification?

- *Is there anything specific [about aerosols]???*
- Specific Data? Specific dispersion phenomena? Specific users?
- Is Persistence Test useful? Any useful to use climatology for short-lived aerosols?
- Data sharing "difficult"

- Data is less and less homogeneous (not a thought through plan to where to put stations)
- Is there a willingness to exchange scores.
- Aerosols are similar to precip, trying to split the aerosol from the meteorology.
- Threat score → introduce some uncertainty.
- It would be nice to have a score that is rather smooth to show the evolution.

- AOD, event based scores
- Equitable Threat score → event based using aeronet (pm2.5, AOD)
- Continuous scores (ACC) → AOD

- Keep track of fine and coarse modes of optical depth.
- Common database of obs.
- Important parameters: AOD, visibility (critical parameter)

- Common climatology

Discussion (cont.)

…Is there anything specific [about aerosols]??? Specific Data?
Specific dispersion phenomena? Specific users?

**Action: do a "user" survey to identify the specific needs that may inform
the type of model skill that are more valuable: that would provide a motivation
on what scores to prioritize.**

Is Persistence Test useful? Any useful to use climatology for short-lived aerosols?
[I think the answer was yes to both questions.] Jeff was especially keen on having
a consensus climatology based on satellite observations that could be made available
as a reference and also to compute scores of the "anomaly-with-respect-to-climatology" type.

**Action: Build this climatology based on the years 2008 and 2009 that were deemed
interesting from the point of view of aerosol. What? Total and Fine/coarse AOD
from satellite/ground based stations/model. PM10/PM2.5: model + ground-based stations?
Fine/coarse concentrations - model. [Am I forgetting something?]**

- Data sharing is "difficult" [but not impossible].

- Data is less and less homogeneous (not a thought through plan to where to put stations).

- Is there a willingness to exchange scores. [I am not sure that was a question, but I believe that, yes, there is willingness from all centres to communicate over this].

- **Action: Use an existing "template" from one of the groups and see if things can be ported to the other centres.**

- Aerosols are similar to precip, trying to split the aerosol from the meteorology [Some people also underlined again the importance of joint verification of the aerosols and the underlying meteorology, as suggested by Doug].

- Common database of obs.

- Threat score: [useful, but need to] introduce some uncertainty.

- Continuous scores (ACC): [possible for] AOD It would be nice to have a score that is rather smooth to show the evolution [see Jean-Jacques' proposal in the section below].

- AOD, event based scores [I said that we cannot NOT verify in AOD space because that's one of the main observables]. Equitable Threat score event based using AERONET (AOD) and other observing networks (PM 2.5)

- Keep track of fine and coarse modes of optical depth [this is possible with AERONET data].
- Important parameters: AOD, visibility (critical parameter), PM

- **Action: visibility data from SYNOP - get list of "good stations" fromWalter Sessions (NRL).**

- **Action: get PM data from NCEP available hourly on ftp (contact Sarah Lu, NCEP).**

- Sharing tools?

- **Action: start from common colorbar**

- **Action: look at online tools (NCAR, for example)**

- Many issues are faced in mantaining a network meant for research and not operations, on a tight budget: find funding to buy instruments?

- **Action: ask for a minimum of 10-12 consistent stations, compatible with budget restrictions**

- Issues of representativeness of station data: study representativeness using the model?

- Verification against own analysis: useful for model development and to monitor the performance of the forecast (biased/as a function of range, etc.)

- Model inter-comparison is seen as a strength.

- **Action: start building a multi-model ensemble for dust forecasts. Common data format needs to be specified, GEMS RAQ example can be followed, contact Miha Razinger (ECMWF).**