

Verification meeting Lisbon

Report

18-20 September 2013

Participants: Alex Deckmyn (AD), Neva Pristov (NP), João Rio (JR), Piet Termonia (PT), Christophe Zingerle (CZ)

Review of the plan(s)

In order to distinguish between the different objectives for the ALADIN consortium of the tools it is summarized in the following table:

	Compute scores on the fly	Monitoring of the applications in the countries	Validation of new cycles	Science verification	Verify fields or pointwise
ALADIN Performance Monitoring Tool in Ljubljana (APMT)	yes	yes	no	no	pointwise (station data)
HARP	yes	yes (through APMT)	no	yes	both
HIRLAM verification tool: The HARMONIE system	no	no	yes	yes	pointwise

The idea is that HARP will be more and more integrated in the APMT.

2. Review of the action on identifying our end user.

Synthetic outcome of the enquiry about the “end users” of the ALADIN countries carried out by JR:

- Most of the variables identified in the forms are the usual ones. Some of the less frequent are: PBL height, MOCON, TKE, cloud water and ice, visibility, convection index, probability of occurrence of thunderstorms, forest fire index; biometeorological index;
- From the sample of answers, the verification performed at the several institutes can be considered to be mainly classical (forecast point vs observation point). Even though not explicit, some countries are expected to have implemented object-oriented or fuzzy methods (partly in HARP yet) to address the double-penalty problem inherent to the validation of high resolution forecasts of precipitation or cloud cover;
- CHMI and ZAMG do verification based on catchment areas. Austria and Slovakia appear to be

- the only ones to make use operational of SAL (object-oriented verification method);
- All or most of the countries supply forecasts for the following sectors: aviation, renewable energies, energy management, public/private companies (construction, transports) and civil protection;
- In the remaining sectors there are apparently some differences: (1) some institutes have products for specific clients - public and / or private (e.g. at ZAMG), while (2) others supply only general information (e.g. CHMI, Romania), directly from NWP or via their weather center (e.g. Portugal, Romania);
- Some examples of decisions taken by clients, based on forecast products, are: (1) hydrological warnings based on water level thresholds, (2) concentration of pollutants; (3) airport and sea/harbor/port operations, (4) type and amount of energy production either for consumption or trading, (5) winter road/rail maintenance, (6) security of outdoor events, both in land and sea, (7) estimate of visitors at selected locations and (8) irrigation and protection against severe weather in agriculture.

Although we can identify a group of sectors, the needs of every client are still too different to go into that detail.

Observation: from the discussion we had, we found it useful to make the split in different stakeholders,

- **valorization of R&D** over a longer past period, is of interest for our policy makes. Ideally one should monitor a score over one or more decades. An example is the weighted NWP index in Austria. However, to do this for the whole consortium, one should (a) have a reference system, and (b) an index might not be relevant for all the countries. So it was felt that it would be best to limit this activities to building up a *portfolio* of cases (example the Portuguese case, the CE flooding, Austria case). These should not be scientific studies but rather descriptions that should be understandable for non meteorologists.
- The **end users** that use the data in applications. The problems we identified for this are
 1. that essentially we lack always relevant data ourselves [e.g. for wind-energy applications, ideally what counts is the power output as a function of the wind to estimate how much value our forecasts are producing. Unfortunately it seems that few, if any, companies are even willing to provide such data. As an alternative we limit ourselves to EV (Economic Value) scores where the cost and the loss of the yes/no decisions does do not need to be specified. In this sense it might be useful to make an estimate of how much of our end users' activities can be reduced to a yes/no decision. It seems at least a good idea to extend HARP as much as possible with such “yes/no” scores like for instance EV.
 2. That, on the other hand, often the end users do not actually know what they need for verification either.

Conclusions:

- How much of the applications a decision taking (the event occurred or not?). It is difficult to derive this from the outcome of the enquiry.
- Extend HARP to validate more than stations data (e.g. Energy fluxes): a discussion was led if we extend HARP to include other variables. However the only extension that could be made is to make it flexible to read mast data, or more general, vertical data. It was estimated that this should not be very complicated. Given the number of specific variables (see list above) this

should then facilitate the development of specific verification in the countries.

A few illustrative cases were discussed (that may be included in an ALADIN portfolio):

- Plots of Joao: 13/9 in Portugal precipitation. ECMWF 0 mm.
- Plots of Christoph and Theresa Gorgas: SAL vs. Contingency table statistics; SAL is more appropriate for the high resolutions.

We discussed where to focus on: long term scores (subject to available manpower, Christoph Wittman already implemented it in R could be included in Ljubljana tool) or monthly Report (idea proposed during the meeting in 27-29/03/2012, Brussels). We used APMT to produce some plots for the monthly report for Austria, Czech Republic and Hungary. From this we conclude that drawing useful qualitative conclusions would be feasible. So the following approach to proceed was proposed:

- The monthly report gets priority. Action work for Jadwiga Woyciechowska to produce them.
- We will also produce some reports for a few countries in hind sight (Cz, 2007 summer/winter + 2012 same months, Austria, Hu)
- We will not perform comparisons between the different countries: plots will only contain one country at a time.
- They will only be produced for the 00-UTC runs.
- All stations will be used to compute the scores for one country (i.e. all stations per country vs. limited the model(s) of the country), then create PDF files, store them to a file systems and send them by E-mail to the LTMs without cross country exchange (i.e. the LTMs will only receive the reports for their own country). Synthetic qualitative conclusions will be drawn by CZ and PT for PAC/GA meetings.
- Add maps with mean bias and rmse, as a second document. Not necessary on the fly, but practically can be done to avoid that it is too slow.

The Shiny tool to interactively look at the verification was demonstrated by AD. Some remarks:

- It could be used as central server with DB (long term) or for local data.
- Currently we have nothing specific in mind, but the aim is to prepare a package to be available for use.
- It would be one piece of HARP, but must be more packaged.
- Currently a research tool. It is expected to evolve naturally. We will see in the consortium what the reactions are to the presentation of it.

Actions:

- PT with input of CZ prepare a document of GA, the precise content will depend on the progress of the stay in Ljubljana, see next point.
- Jadwiga Woyciechowska's stay in Si: finalize the tool to produce the monthly reports.
- Adapt the rolling plan with the outcome of this meeting.