Questions, Comments and Answers following the presentation

A demonstrator for an automatic operational system for the optical turbulence forecast for ESO sites (Cerro Paranal and Cerro Armazones)

Elena Masciadri

Hainaut:

1. *What additional data could the observatory provide to better constrain the forecast? Temperature sensors over the ESO land? more seeing monitors?*
2. *Assuming infinite computer power, would making the forecast "our" shrink from ~ 18h down to ~ 6h lead to better forecast?*

1. It is very important to know the homogeneity of the turbulence and atmospheric condition. Another station for atmospheric parameters located in a different position from the present one above the plateau can certainly be useful to quantify the homogeneity of these parameters above the plateau. Same thing for the DIMM. At my opinion, all DIMMs should be placed on a tower having the same height of the UT – primary (~30m). At present the old and new DIMM are located at 6m m above ground (old DIMM) and 7m above ground (new DIMM).
   This should permit to better identify a relation between DIMM and UT. The procedure used so far to pass from one to the other presents controversial elements. A more suitable height of the tower would certainly help in achieving a correct quantitative estimate of the seeing at my opinion.
2. Logic says that the larger is the delay time of the initialization data, the higher is the decreasing of the equality of the performance. We already quantified that this effect is, in our configuration not so relevant. With that I mean that, looking at differences observed between the best case (model initialized with analyses) and the operational configuration (model initialized with the forecasts), the gain does not justify such a cost to improve things. This does not mean that this is useless. Simply not among the main priorities now (at my opinion).

Tristram: *What are the plans to make the model predictions available to science operations on Paranal in order to optimize the scheduling of observations? Are there any plans how this will be implemented in practice, e.g. in the Astronomical Site Monitor?*

There is a plan we presented in the MOSE feasibility study. The idea is to set up a system similar to what we did for LBTO (ALTA Center) but dedicated to the spec./necessities of the VLT. The system is therefore an independent system from the ASM. The access for resident astronomers for the nightly use should be done by the web (as they usually do with the ASM). This is an important element for our project. We have to manage model outputs inside a consistent environment that we manage and control (of course the content of the web site is established in agreement with ESO). The access of the “next night” is protected (i.e. accessible only by ESO and developer team) because of policy. The access to “previous nights” is open.

Milli: *You are interesed in getting as many measurements as possible, for instance coming from the active optics of the UT. Is it to validate the output of the prediction or can we envisage updating the modelling at the beginning of the night with the first measurements to get more accurate predictions for the second half of the night?*

Measurements from the active optics of the UT (or other methods) should be useful in this phase for the model validation. In perspective, they might be used to improve the prediction in post-processing phase.
I think that the most important thing is, now, to have as many as possible "independent" measurements of the optical turbulence done with instruments conceived for OT measurements but also with AO systems. It is important however that our community invests more efforts towards the problem of the absolute estimate of the turbulence in measurements.

Rantakyro: You presented the confidence in the predictions averaged over the whole night. I could imagine some models on average being accurate but in shorter intervals being less accurate. How well the model tracks short time variability of the seeing ? In other words, assuming that a predicted seeing that is higher than the measured value is a "hit" for a single time-stamp, how long on average is sequence of "hits" with say 90% confidence?

There was probably a misunderstanding. We did not present the confidence of the prediction averaged over the whole nights. Each couple of points (meas., sim) is processed taking the temporal evolution of measurements and simulations all along the night doing a moving average of 1 h and taking a successive temporal sampling of 20 minutes. We analysed therefore observed and simulated values on a short time scale. We chose this approach in agreement with the ESO MOSE Board after we had tested also other approaches including the same analysis done without moving average, solution that was discarded because it was clear the high frequencies did not provided useful information.

In slides 36 and 37 we can appreciate, in a qualitative way, the trends and the dispersion between simulations and measurements from one side and between measurements from different instruments from the other side. In slide 38 the quantitative estimates are reported.

For what concerns the calculation of the PODs (second part of the question), when we calculate, for example, the probability that the seeing is weaker than the fixed threshold (e.g. the first tertile) this means that we count the number of hits from the simulations that fall in the correspondent range of values ratio the total number of time that measurements fall in the correspondent range of values. Each couple of (meas., sim) values is also calculated on a short time scale.

I highlight that we perform validation using several different strategies/methods. The principal ones are reported in the papers (slide 46), other in the reports. We also compared our method to others methods (Masciadri et al. 2017) and it appears evident an evident advantage. Further tests are planned but we are waiting for other measurements from the GS.

There is a rich literature on this kind of procedures and it is important to focus on the most useful in our context and those that can provide a useful information with respect to the measurements we can access.

As a final consideration, if one looks at slides 36 and 37 showing the temporal evolution of J observed with a DIMM and a Generalized SCIDAR we can observe that it is not trivial to define what you define short time variability because there is an intrinsic variability between different instruments. The example done in the question is therefore much more complex in the reality. I think that at this stage it is a priority to focus first on the measurements that we can take as a reference and evaluate their reliability and to do in a way to access to several independent and simultaneous measurements.

This was our most important limitation so far.

Roth: From the presentation of a series of nights with the comparison of measurements vs model prediction, can you comment on where the forecast went well, and where it failed, and what would be needed to improve the model?

The question is referring to the slide in which I show the temporal evolution of the total amount of turbulence as measured by the DIMM, the generalized SCIDAR (GS) and as forecasted by the Astro-Meso-Nh model during all the 20 nights of the PAR2007 site testing campaign (slides 36 and 37).

In all these cases the model behaviour is substantially good in the sense that the trend of the turbulence is well reconstructed and the dispersion between model and measurements is comparable to dispersion between measurements from different instruments. In this example, there are no important failures of the model. I define 'failure of the model' when the difference between the model output is much larger than the dispersion between measurements. From a quantitative point of view (slide 38) we can see that the most challenging POD for the seeing (e.g. POD1) tells us that in more than 80% of cases the model reconstruct a seeing weaker than the first tertile when it is also observed and this is more or less the same uncertainly we have if we compare measurements from independent instruments.

The causes of the model failures are of different nature such as (1) initial conditions that, for some reason fail and (2) the model parametrizations of physical phenomena that take place at sub-grid scale that are coded in the model. On the other side, it is important to consider also the errors of (uncertainties) of measurements to carry on a fair comparison.

Point (1) plays an important role. It should be therefore useful to provide the measurements of the atmospheric parameters done by the Astronomical Observatories in the Chilean region might be included in the assimilation data of the ECMWF to improve the quality of the initialisation and forcing data provided by the General Circulation Model. This is under discussion.