Subjective evaluation of different versions of the ALADIN/HU model

Helga Tóth, Hungarian Meteorological Service, *toth.h@met.hu*

1. Introduction

At the beginning of 2004 a new project started at the Hungarian Meteorological Service with the aim to subjectively evaluate our operational and quasi-operational model versions in an operational manner. This project was motivated by different reasons. On the one hand the experience gained by the subjective evaluation can be a valuable complement to the widely used objective verification scores. On the other hand we can judge and control the opinions of the forecasters about the models in a very simple way. On top of that we can have an opportunity to compare the subjective marks with the objective scores. From the common evaluation of all verification information we are able to make conclusions and decisions about development directions and the choice of operational model versions.

2. Method

During the subjective evaluation we compared different models over the Hungarian territory: the ALADIN/HU dynamical adaptation at horizontal resolutions of 12 km (former LACE resolution) and 6.5 km, the ALADIN/HU forecast based on the 3D-VAR+CANARI assimilation cycle at 12 km resolution, and the ECMWF model. The models are compared to each other and to the surface and radiosonde observations, radar and satellite measurements. At the end of the process we classify subjectively the forecast quality of the different predicted meteorological parameters. The verified variables are as follow s: precipitation, 2m temperature, total cloudiness and 10m wind. A 5-grade classification was created : from "5" mark for excellent forecasts to "1" for completely wrong predictions. The categorization is rather subjective, however some common criteria were considered, e.g. for temperature the mark is 5 if the spatial and temporal mean difference between the forecasts and measurements is within a 2 K interval, or it is 1 if the defined difference is larger than 6 K.

We were evaluating the forecasts based on the previous day integration (0-30 hours integration in case of ALADIN models and 12-42 hours for the ECMWF model), but from 1st of July 2004 the runs from two days before are considered (0-48 hours integration) and the time period is divided into two parts (0-24 hours and 24-48 hours for ALADIN models and 12-36 hours and 36-60 for ECMWF model). 5 persons in the NWP group are in charge of the subjective evaluation in weekly periods. During the first weeks we got some help from a forecaster expert regarding the evaluation and the interpretation of the model results.

3. First results

Up to now we have roughly half a year of experience about the subjective evaluation and hereafter 2 plots are showing the basic results for the period February-May 2004. The figure about the total average of marks (left) shows that the most reliable model is the ECMWF one and the two kinds of dynamical adaptation produce almost the same but a little bit less quality of forecasts. The 3D-VAR predictions are a bit worse than the others but the difference is not so significant.

We can check the variables individually as well from the other figure (right panel). It can be seen that the best predicted element is the wind, and the precipitation is also rather well represented. The cloudiness and the 2m temperature values are not forecasted too successfully especially in the case of 3D-VAR based forecasts. The largest difference (more than 0.2 mark in average) can be found between the forecasts of ECMWF and ALADIN/HU models for cloudiness. This discrepancy is coming from the problem of the ALADIN/HU cloudiness parametrization. A lot of times partly covered sky was predicted by the model, however there were no cloudiness at all in the reality (and in these cases the ECMWF model provided very good forecasts). On the top of that this kind of forecast is not sufficiently informative when the cloud cover is going to change.

Surprisingly the 2m temperature forecasts based on ALADIN/HU 3D-VAR system showed very weak quality. This was interesting because the 2m temperature is an analyzed variable of the data assimilation scheme, so the guess is corrected by the observations, therefore the 3D-VAR

analysis provides usually the best initial state for the model. Nevertheless starting from a good initial state the forecast becomes worse than the other models (probably some balance properties in the initial conditions are not kept). In the future we have to find the reason of this deterioration and correct it as far as possible.

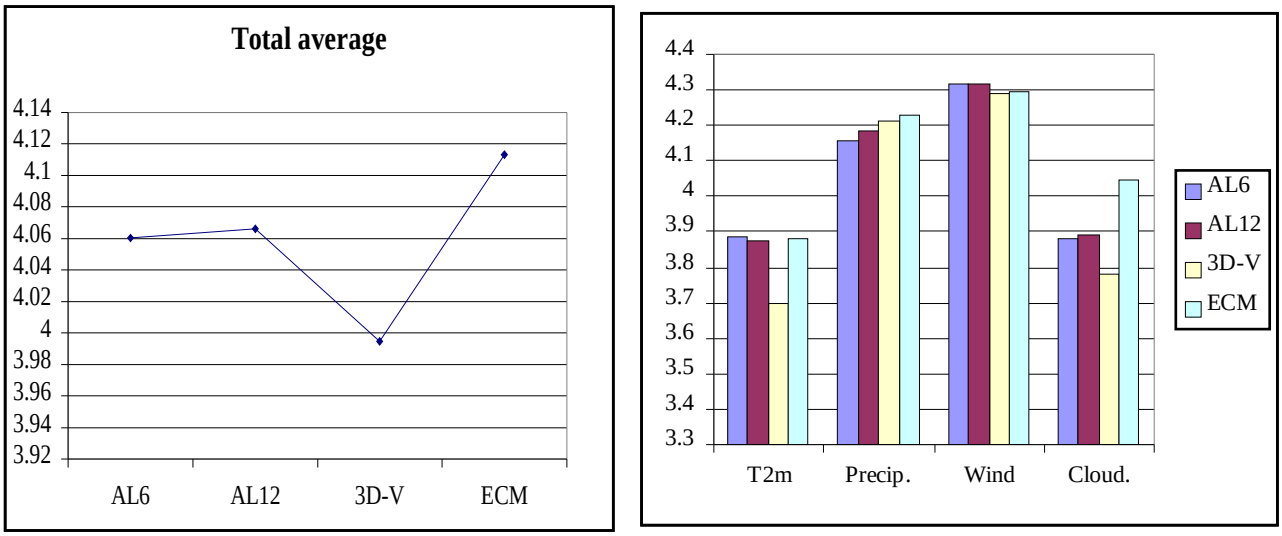

Figure 1. Total average marks (left) and averages for different parameters (right)

CONTENTS