Comparison of precipitation forecasts by two operational NWP models:

The global model ARPEGE and the smaller mesh LAM ALADIN

Marielle Amodei and Joël Stein

DPrevi/COMPAS Météo-France

Objective: The 24 hours accumulated precipitation forecasted by ARPEGE and ALADIN is verified and compared over France according to the methodology recommended by the WGNE (Working Group on Numerical Experiment) in order to :

✓ Define a set of relevant scores to verify the precipitation forecasted by the French operational high-resolution LAM, ALADIN today, AROME in the future.

✓ Evaluate differences between **2** models and determine if they are statistically significant.

Observations and models used:

✓ The temporal period contains 9 trimesters from December 2004 until February 2006.

✓The climatological state network (Figure 1) contains almost 4000 observations per day. These observations, spaced by around 10 km, are averaged on the 0.2°verification grid to provi de the reference data.

✓ The persistence of the precipitation of the previous day provides the trivial forecast used to compute skill scores.

✓The spatial resolution of the uniform verification grid is 0.2°x 0.2°.

✓ The ALADIN forecasts, available over a 0.1°, are aver aged over 4 points and the ARPEGE ones are interpolated from a 0.25° grid.

✓ The global ARPEGE model provides lateral boundary conditions to ALADIN. The physical packages are similar for the 2 models. Since the 25th of July 2005, the ALADIN model has its own 3DVAR assimilation and assimilates the high resolution SEVIRI radiances from Meteosat 8.

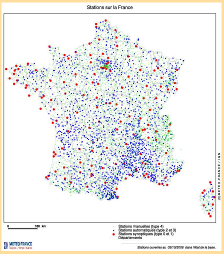


Figure 1: Climatological state network : manual stations (green), automatic stations (blue) and synoptic stations (red)

Deterministic approach:

Verification is performed in a deterministic framework. Scores are computed over every trimester from the contingency tables for the different thresholds 0.2, 1, 2, 5, 10, 20 mm/day over France. A bootstrap resampling technique is used to compute a box plot for the different scores but also to test the significance of the difference between both models according to *Hamil (1999).*

Autumn 2004 and 2005 are comparable according to the rainfall histogram and surround the 3DVAR beginning (Figure 2).

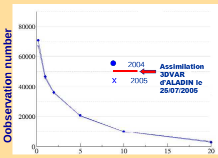


Figure 2: Number of the precipitation observations for Autumn 2004 (closed circle) and 2005 (cross)

During Autumn 2005 (cross on Figure 3), ALADIN exhibits significantly higher HSS than ARPEGE for all thresholds excepted the 20 mm/day. During Autumn 2004 (closed circle on Figure 3), the differences were smaller and in favour of ARPEGE. This evolution can be explained by the assimilation benefit for the ALADIN forecasts.

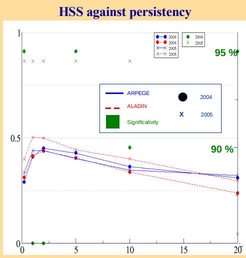


Figure 3: HSS for the Autumn 2004 (closed circle) and 2005 (cross). Green symbols indicate when the model difference is significant at 90%, 95% or not

The temporal evolution of the HSS (Figure 4) shows the significant superiority of ALADIN for the rain detection (threshold 0.2 mm/day). For stronger precipitation (20 mm/day), this is the contrary. The improvement occurring after JJA_2005 for the ALADIN model is permanent for both thresholds and is due to 3DVAR assimilation introduced in ALADIN on the 25/07/2005. Nevertheless, it only leads to a reduction of the gap between both models for strong precipitation.

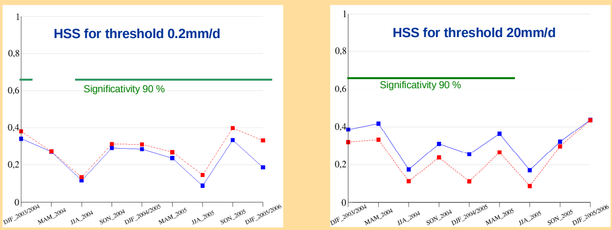


Figure 4: HSS evolution along the whole period for the 2 models ARPEGE (blue) and ALADIN (red) and for extreme thresholds values 0.2 mm/days (left), 20 mm/day (right) and for 2 NCS 60 km (top) and 242 km (bottom). Green line indicates when the model difference is significant at 90%.

Fuzzy approach:

In order to avoid the "double penalty" which affects the high resolution models, a statistical post-processing is performed to allow a probabilistic treatment of the forecasts *(Robert 2004, Theis et al. 2005).* A probability to exceed a given threshold, is computed in a square of nxn grid points for the forecasts and the persistence. The verification is performed by computing the Brier Skill Score (BSS) and Rank Probability Skill Score (RPSS). Both are estimated according to the persistence. 2 BSS are computed:

BSS-SO single observation – neighbourhood forecast gives us a local verification

•BSS-NO neighbourhood observations- neighbourhood forecast gives us a regional verification.

The neighbourhood characteristic size (NCS) vary from 60 km to 242 km. Both BSS are plotted for two seasons (DJF 2006 and JJA 2004) (Figure 5). By comparing both figures, we conclude that winter precipitation is better forecasted than summer one because of the more convective nature of the summer rainfalls.

BSS-NO outperforms BSS-LO for all thresholds and NCS. BSS-NO grows continuously with NCS to reach a maximum of 0.9 in winter while BSS-LO is quasi-constant. However in winter the BSS-LO slowly decreases for the largest neighbourhood. BSS-LO peaks for NCS= 130-140 km, indicating that information at greater distances is not pertinent for the local forecast.

Strong BSS-NO values in winter show the ability of the ALADIN model to predict regional precipitations. Note that the relative bad performances for the 0.2 mm/d is probably due to measurement threshold of certain raingaujes. During summer, the probabilistic scores reduce the double penalty influence for the convective cases more frequent in this season for the strong threshold 20 mm/d. Therefore, BSS-NO increases the most quickly for the 20 mm/d among all these thresholds.

For these reasons, BSS-NO is more relevant in the framework of precipitation verification by operational model.

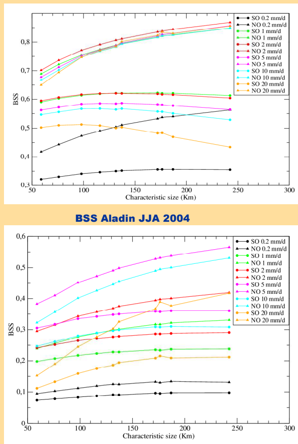


Figure 5: BSS_SO (closed circle) and BSS_NO (triangle) for several NCS, Winter 2006 (top) and summer 2004 (bottom). The thresholds (0.2 to 20 mm/d) are differentiated by colours.

The comparison between the two models is shown for the BSS-NO along the whole period, for extreme thresholds values (0.2 mm/d, 20 mm/d) and 2 NCS (60 km, 242 km) (Figure 6).

The largest neighbourhood always achieves the largest forecast improvement whatever the trimester. The BSS-NO_242km is 10% larger than the BSS-NO_60km.

The fuzzy approach shows an improvement of the ALADIN model for the rain detection (0.2 mm/d) during the two last trimesters after the ALADIN assimilation introduction. This is coherent with the classical HSS scores. But for the 20 mm threshold, the ALADIN improvement after JJA 2005 is greater for BSS than for HSS.

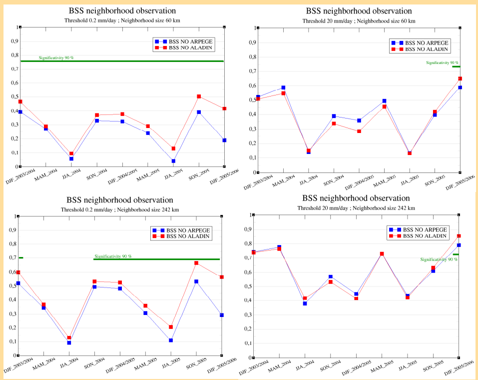
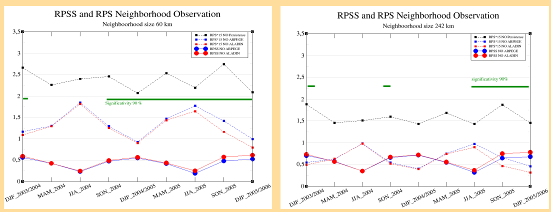


Figure 6: BSS_NO evolution along the whole period for extreme thresholds values 0.2 mm/days (left), 20 mm/day (right) and for 2 NCS 60 km (top) and 242 km (bottom). Green line indicates when the model difference is significant at 90%.

This is related to a reduction of the double penalty importance allowed by the fuzzy method. Nevertheless, the impact is not as important as for very high resolution models because the resolution ratio between ARPEGE and ALADIN is only 2.3 and they share the same physical package.

A more synthetic information is given by the Rank Probability Skill Score (RPSS) (Figure 7). RPSS-SO and RPSS-NO are computed in the same way as the BSS. Reference is provided by the persistence. RPS of the two models and of the persistence have been multiplied by a factor 15 and plotted together with the RPSS. These curves confirm, that the RPSS variations are mainly due to the RPS variations and the ALADIN improvement is clear after the introduction of its proper assimilation. We also notice the improvement in term of skill scores when a larger verification box is used.

Figure 7: RPSS_NO (full line) and RPS_NO (dot line) evolution along the whole period for 2 NCS 60 km (top) and 242 km (bottom). Green line indicates when the model difference is significant at 90%. RPS_NO of the 2 models and of the persistence have been multiplied by a factor 15.



Conclusion

•The probabilistic approach has been used to evaluate the differences between the global ARPEGE model and the LAM ALADIN. BSS-NO against persistency is a good candidate to measure the impact of higher resolution forecasts. It partly reduces the influence of the double penalty. Statistical significance can be added to this treatment.

•This method allowed us to show the ALADIN improvement for the precipitation forecast brought by the inclusion of a proper assimilation scheme using higher resolution data.

Future

• The daily raingaujes will be replaced by the ANTILOPE hourly analysis mixing raingaujes with radar data.

•The AROME prototype of the future high resolution model will be verified according to this methodology.

METEO FRANCE
Toujours un temps d'avance