# Validation of the externalized OI scheme in SURFEX

Report of stay, 05-17 July 2009, Toulouse

László Kullmann

## 1 Introduction

Recently the OI assimilation method was implemented into SURFEX. To perform the surface analysis 2 steps has to be run: The first step is to run a "conventional" CANARI to read the observations, to interpolate the guess field to observation space, to calculate the T2m and RH2m analysis and to interpolate back to model grid. This step is done using FA files. The second step is to run the surface analysis. This is done using offline SURFEX. The surface assimilation method was externalized from ALADIN and a main driver was written which reads the prognostic surface guess fields from FM (LFI) files, the 2m analysis fields from the previous step from FA files and some atmospheric fields which are needed by the OI method.

During my stay this externalized OI method was studied and validated.

## 2 Validation

A one week assimilation cycle was performed. ALADIN was run with SURFEX and the externalized OI scheme was used. The results were compared to the operational ALADIN-France forecast.

Since there are still differences in the forecast of ALADIN with and without SURFEX the analysis increments at 2m are different resulting in different surface analysis fields (Fig. 1). However the structure is very similar. If we compare the Ts values between the reference and the experiment before and after the analysis we can see that the difference is much less after the analysis. This means that the OI works well.

If we look at the soil moisture content analysis increment at 18 UTC we can see that the magnitude is close to zero in case of the experiment but this is not the case for the reference. At 12 UTC however the increments have at least the same range of magnitude. (Fig. 2) The problem comes from the time specifications in the SURFEX file (FM), it is not the same as the analysis time. There are 2 sources for the difference. One is related to the use of DFI, since SURFEX is called only in the forward integration. The other problem is that at the beginning of integration there is a call to SURFEX at 0 time step to provide the fluxes for the atmospheric model. This step should be only a diagnostic step, prognostic variables should not be evolve but currently there is no possibility to run SURFEX with $\Delta t = 0$. To verify if the problem of the too small analysis increment of water content is really due to the time problem we have rerun the analysis by giving explicitly the analysis time instead of reading it from the file. Fig. 3 shows that the magnitude are now the same.

To solve the above mentioned problem there might be 2 options. One would be to read the analysis time from the FA file instead of the LFI one (this is what we coded currently). However it might be dangerous to use two files with different times. The proper solution would be to solve the problem of calling SURFEX at the 0 time step.

As we can see from Fig. 3, although the problem of the too small magnitude vanishes but the analysis increments are quite different. Moreover if we look at the difference of soil water content field between the two experiments we can see that after the analysis the difference is bigger than it was for the guess (Fig. 4). The reason is that while the surface and deep soil temperature increment depend only on the screen level increment, the soil moisture depends also on other atmospheric fields like precipitation, cloudiness, evaporation and these fields are different in the two forecast (reference and experiment).

# 3 Taking into account mixed tiles

In the previous section we used SURFEX to be similar to the operational surface scheme which means that over a grid box only one tile can be present: nature or sea. The offline OI surface analysis was developed only for this case. However if we use the ECOCLIMAP database to determine the physiography we will have grid-boxes with more than one tiles.

We have rewritten the driver (`oi_main`) routine to be able to handle also grid-points with mixed tiles. In the original version the program goes through every grid point and determines the tile type according to the land/sea mask and makes the specific calculation for nature or for sea according to that. In the new version the driver (`oi_main`) selects all the grid points having a given tile (by the fraction of tile) and makes the calculation according to that tile type.

Unfortunately the forecast can not be compared with the operational or the previously studied experiment since in this case the soil inertia for bare ground and vegetation has a different value. (If we use ECOCLIMAP we can not give explicitly the value of the soil inertia over vegetation, it is determined by the model.) This means that the surface and the 2m temperature forecast will differ significantly. However even in this case we can see that the analysis is working well since after the analysis the difference between the two experiment (with one tile on a grid-box and with more tiles) decrease. As an example see Fig. 5 where surface temperature is present. We can see that the guess is much colder for experiment 8319 (using ECOCLIMAP to determine physiography and using larger $C_V$ value) but the analysis increments at 2m are higher for this experiment.

We have observed that in case of more than one tile the SST analysis is quite large near the coast. This problem is due to the fact that the atmospheric model determines the land/sea mask differently than SURFEX. Those points which have some small fraction of sea in SURFEX are taken into account as land point by the atmospheric model and the CANARI analysis is calculated according to the nature type. But we think that this error will not influence significantly the forecast.

The analysis over inland water is quite important especially if no lake model is used since the temperature is kept constant during the forecast. If no assimilation is run the initial lake surface temperature is determined from the deep soil temperature of the coupling model. The reason for using the deep soil temperature instead of the surface temperature is that the fluctuation of water surface temperature is much less than the surface temperature over land.

In case when assimilation is run we can not use the above mentioned method since there is no deep soil temperature analysis in CANARI when we run the model with SURFEX. If the fraction of lake on a grid-box is large it may happen that this grid-box is seen in the atmospheric file as sea point. In this case we could use the CANARI analysis (like we use for sea) but the problem is that the analysis will be an interpolation from the closest sea points which may be very far from the lake point, so the analyzed value will be far from the reality.

Our first idea was to calculate the water temperature analysis similar to what is done for the deep soil temperature over nature tile: $T_{s,water}^{anal} = T_{s,water}^{guess} + (T_{2m}^{obs} - T_{2m}^{guess})/(2\pi)$ The problem is that the forecast (guess) value in this case is in fact the analysis value at the previous analysis time (no time evolution of the water surface temperature). The conclusion was that it would be better to take the deep soil temperature from the nearest grid-box which has land cover fraction. This solution is not yet worked out.

# 4 Figures

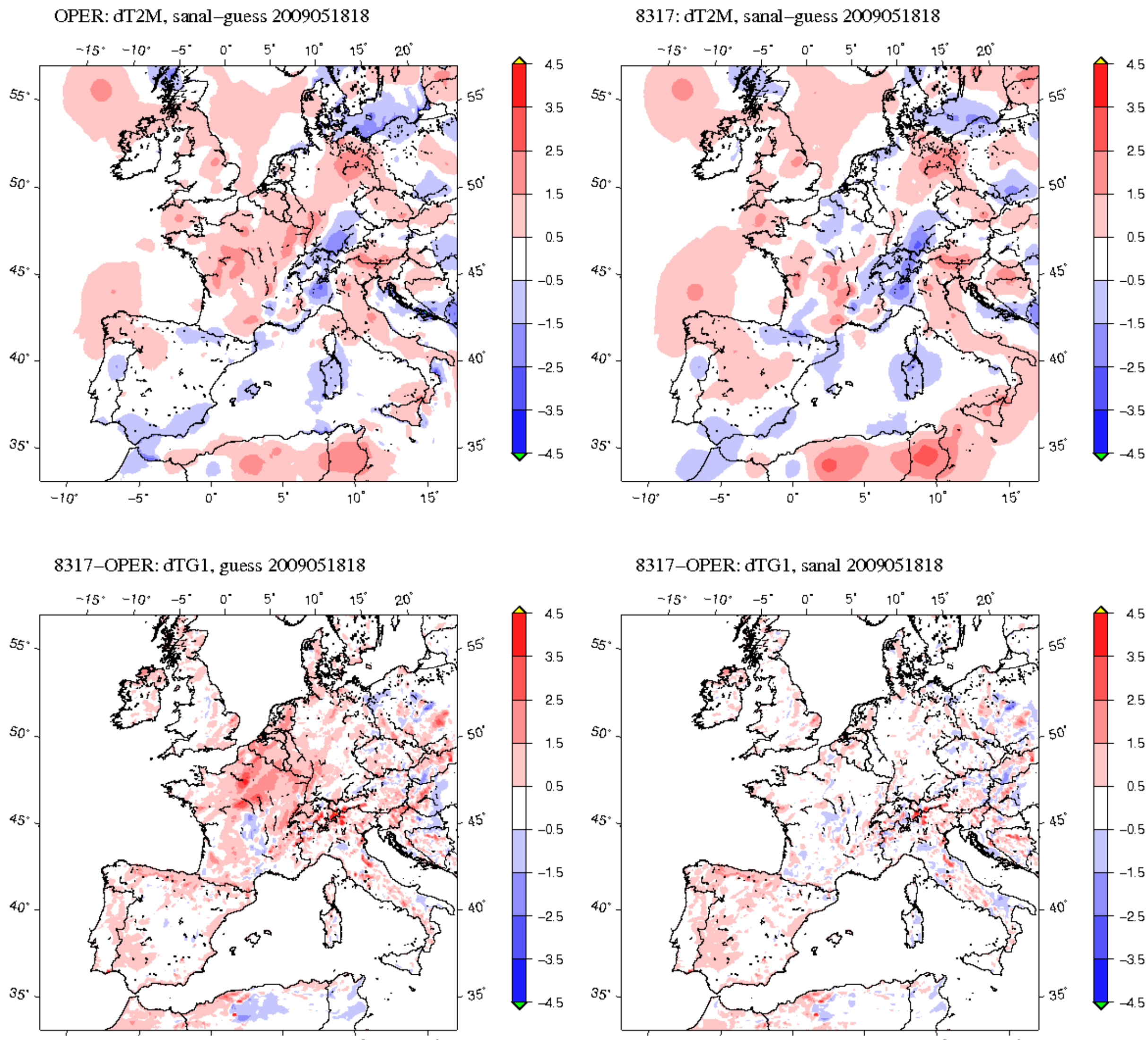


Figure 1: Comparison of the T2m analysis increments for the reference run (top left) and for the experiment (top right) after 18 UTC. The surface temperature difference between the reference and experiment before the analysis (bottom left) and after the analysis (bottom right).

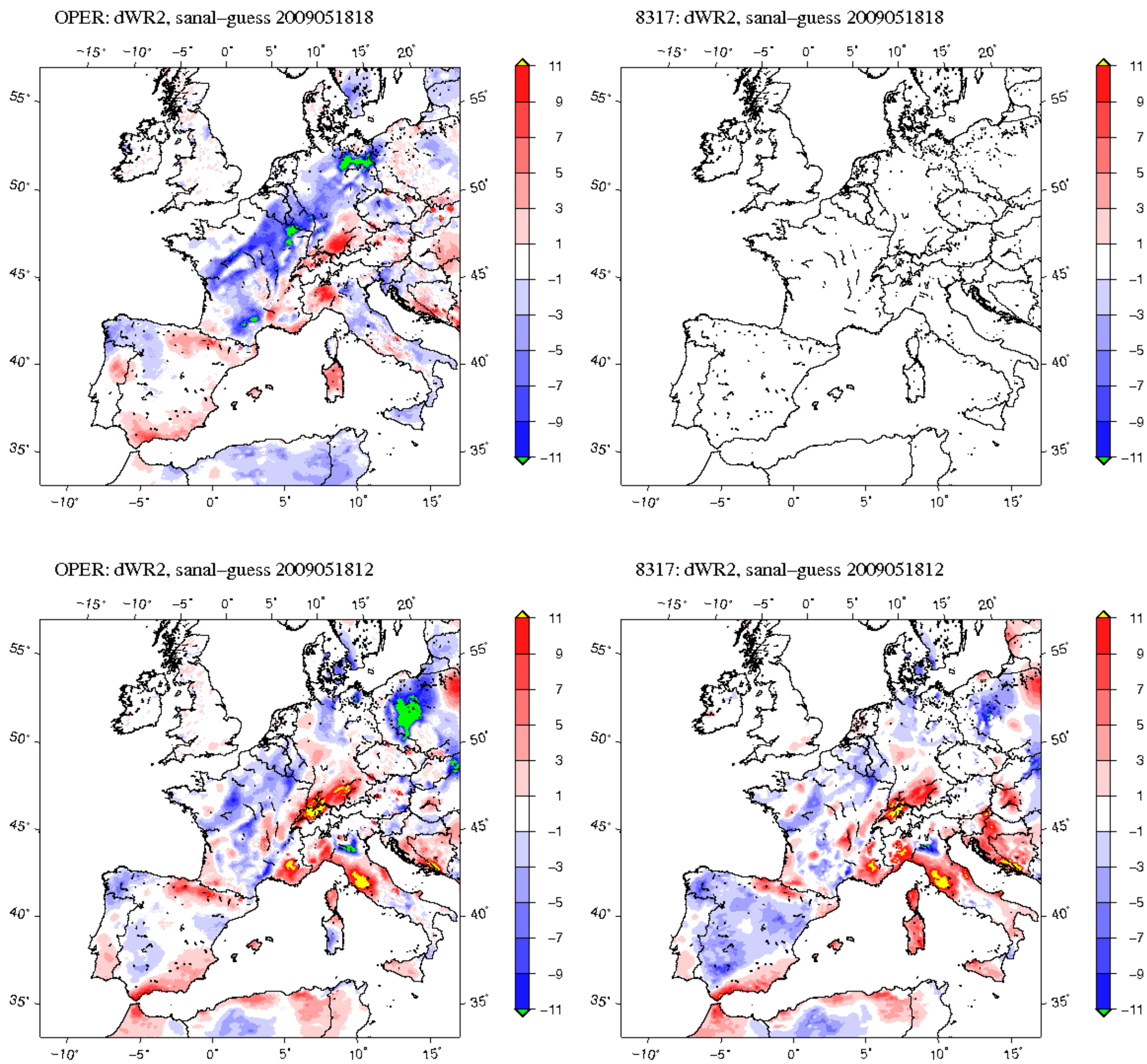


Figure 2: Analysis increment of the total soil moisture at 18 UTC for reference (top left) and for experiment (top right). The same is presented at the bottom but for 12 UTC.

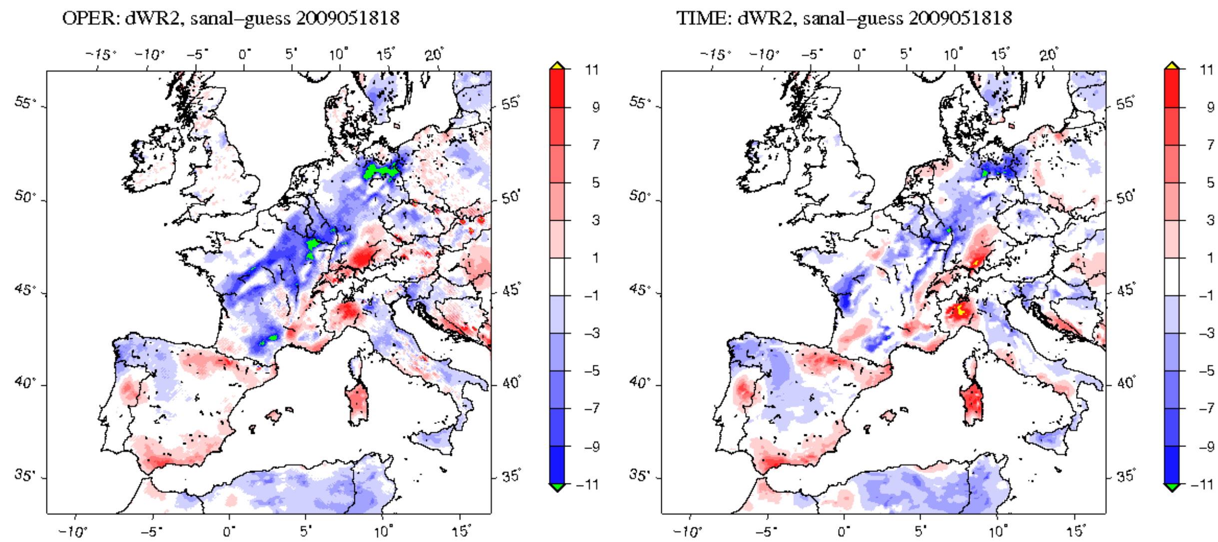


Figure 3: Same as the top row of Fig. 2 (analysis at 18 UTC) but using the correct time.

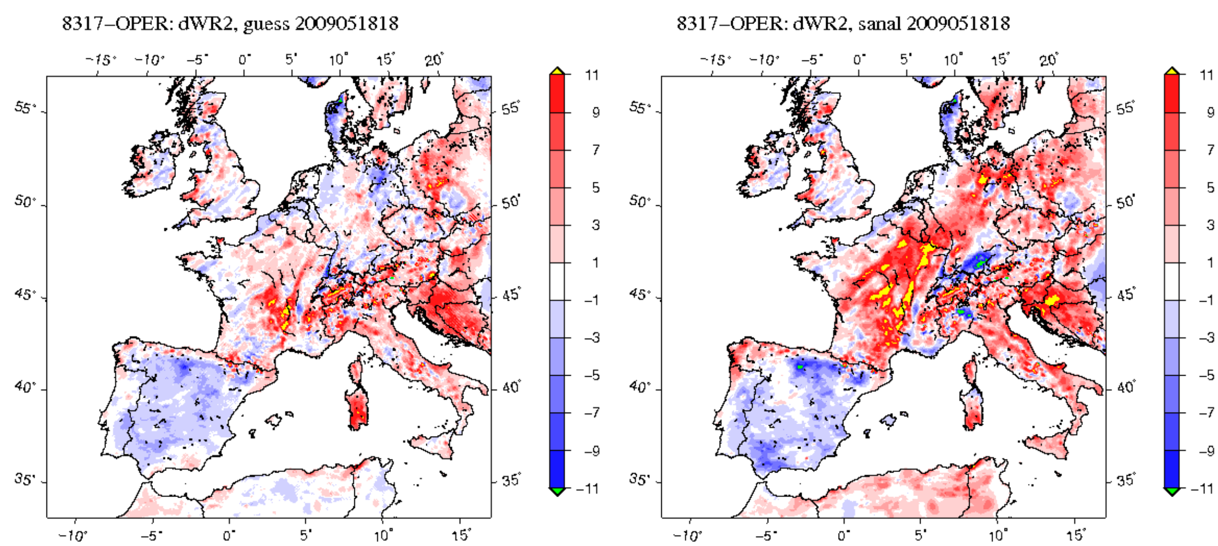


Figure 4: Difference of soil water content between reference and experiments before analysis (top left) and after the analysis (top right).

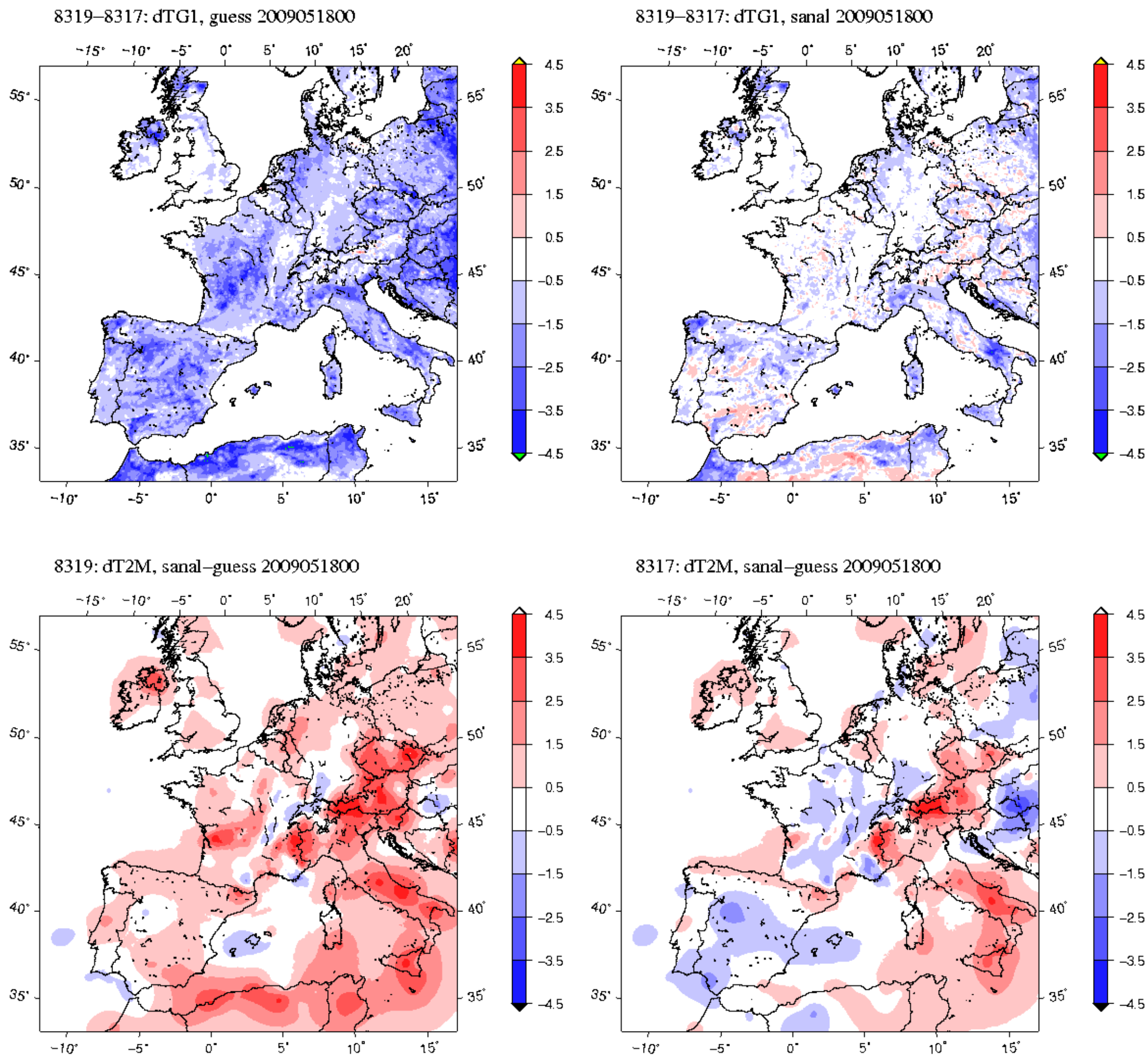


Figure 5: Comparison of surface temperature between experiment **8317** (one tile on a gridbox) and **8319** (more tiles) before analysis (top left) and after (top right). On the bottom row the 2m temperature analysis increment is presented for both models, 8319 left and 8317 right.