LACE/Météo France stay report

Estimating the Added Value of Land SAF Albedo Assimilation in ALADIN

November 2008

Jure Cedilnik*

1. Introduction

The products of Land SAF project [Geiger, 2008] are operationally available for quite some time now, yet up to now there was no real attempt to use these products in context of the ALADIN/ARPEGE/AROME NWP system.

During this stay I evaluated the use of Land SAF albedo product as a source of observation for assimilation in ALADIN NWP model. The assimilation was performed with Kalman filter to minimize the error covariance of analyzed field.

2. Experiments' design

The observation vector for Kalman filter analysis of albedo has three components: two climatological albedoes (one for bare soil and one for vegetation) and the Land SAF total bi-hemispherical albedo product, which is derived on a daily basis from MSG observations [Geiger, 2008]. For a detailed description of Kalman filter analysis of albedo, (see Carrer, MF internal note).

The initial version of analysis software uses ecoclimap data for climatology (both albedoes and fraction of vegetation). Though there is an obvious inconsistency between the two climatologies it was argued that this is not a big deficiency and since it would take only little effort to run an experiment with such data, it was decided to give it a try. This can be also considered as a sensitivity study – to see what the magnitude of impact that the albedo parameter has on model forecast. However, to at least try to isolate the impact of Land SAF alone, another experiment suite was introduced – one that would take into account the modification of ALADIN albedoes only with ecoclimap database values. The description and names of the experiments are shown in **Table 1** in the Appendix part of this report.

At this point it was also decided that later on, the system will be redesigned in such a way that it will use native ALADIN climatological albedoes and ALADIN vegetation fraction and furthermore, that there will be only one interpolation (.e.g from SEVIRI to ALADIN LCC grid) – no intermediate step with regular latlon grid of ecoclimap. The reason for this redesign is also that the architecture of initial assimilation code was not suitable for operational purposes. A simple diagram of dataflow that this scheme would use for the application in operational NWP is presented in **Figure 1**.

The chosen period for testing was from 1st of February to end of July 2007 (6 months altogether). A run was performed every day at 00 for up to +54 hours. The domain used was the

* Jure Cedilnik (jure.cedilnik@gov.si)
Slovenian Meteorological Service, Environmental Agency of Slovenia, Vojkova 1b, SI-1000 Ljubljana, Slovenia

current operational domain of Slovenian meteorological service with resolution of around 9.5 km and the size of 256x270 points, see for example **Figure 2**. The analyzed albedo values from one day before were used to modify only the initial condition file. Coupling files were not touched and were exactly the same with all the experiments – this was to speed-up the computation and economize the computational cost.

OLIVE Swapp environment was used for experiments' control and perform standard diagnostics (computation of statistical scores as it is done in COMPASS department). For more technical details about the set-up of the experiments see the appendix of this report.

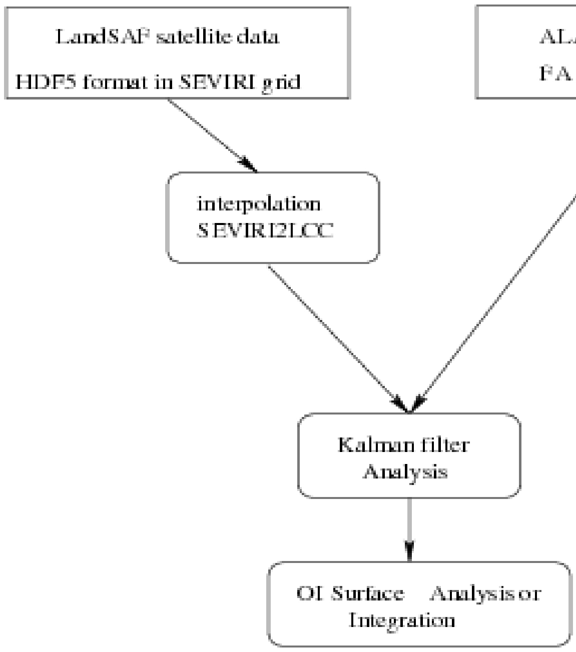

Figure 1: Dataflow of the redesigned experiment: the oval boxes are simple (one task only) programs and the process is controlled by a script. Such a design is quite suitable for a potential use in operational environment.

3. Results – using ecoclimap climatology for albedo analysis

During the first phase three experiments were performed altogether: reference run, a run with analyzed albedoes (analyzed through use of ecoclimap climatology) and a supplementary reference (a run with ecoclimap albedoes). **Figure 2, left** shows a quite significant impact of new (satellite and ecoclimap) data on the 12 hour forecast of 2m temperature. The differences seem to be more pronounced in flat terrain regions and basins – see for example the Po valley or the Rhone delta region. The highest differences are in favour of higher temperatures e.g. lower albedo values. Probably the reason for this is the difference in the albedo of vegetation – the growing period starts earlier than usual.

Figure 3 shows the same impact in terms of statistical scores. Scores were computed on a slightly smaller domain (see **Figure 10** in the Appendix). It seems that the greatest impact on temperature scores is during daytime, which corresponds to maximum solar activity and peak in shortwave radiation. However the favorable temperature bias is kept even during the night. Another interesting thing is that the difference between ALADIN with ecoclimap albedoes and reference ALADIN (Figure3, bottom) is negligible and this can serve as a circumstantial evidence that the most of improvement is a consequence of satellite measurements alone.

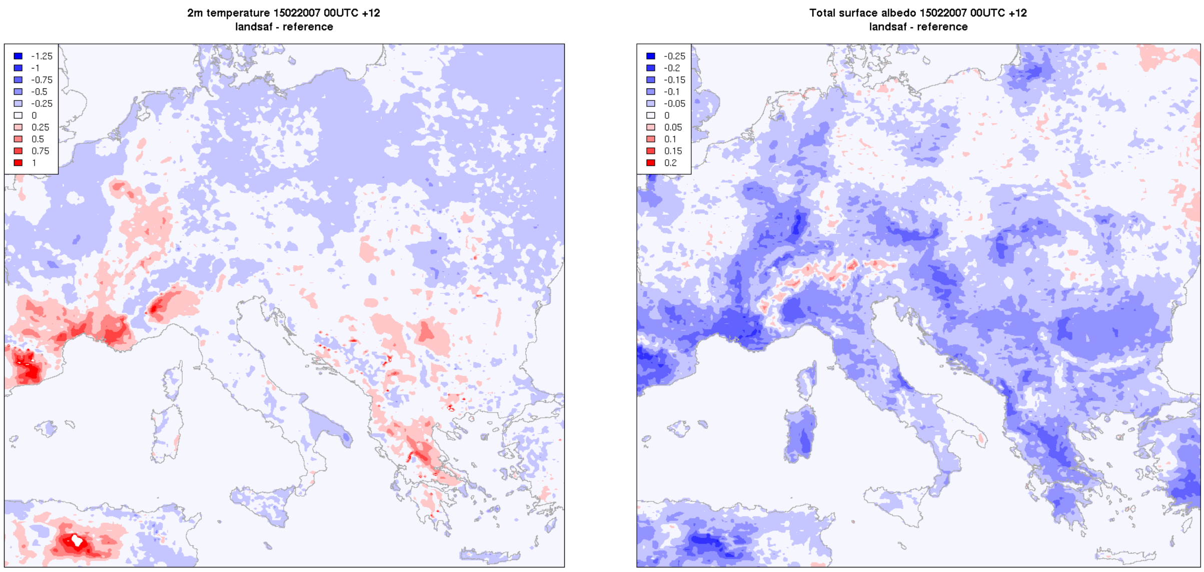

Figure2: 2m temperature after 12 hours of integration (left) and total albedo (right) difference between ALADIN run with analysed albedo (using ecoclimap climatology) and reference run. The date is 15th of February 2007.

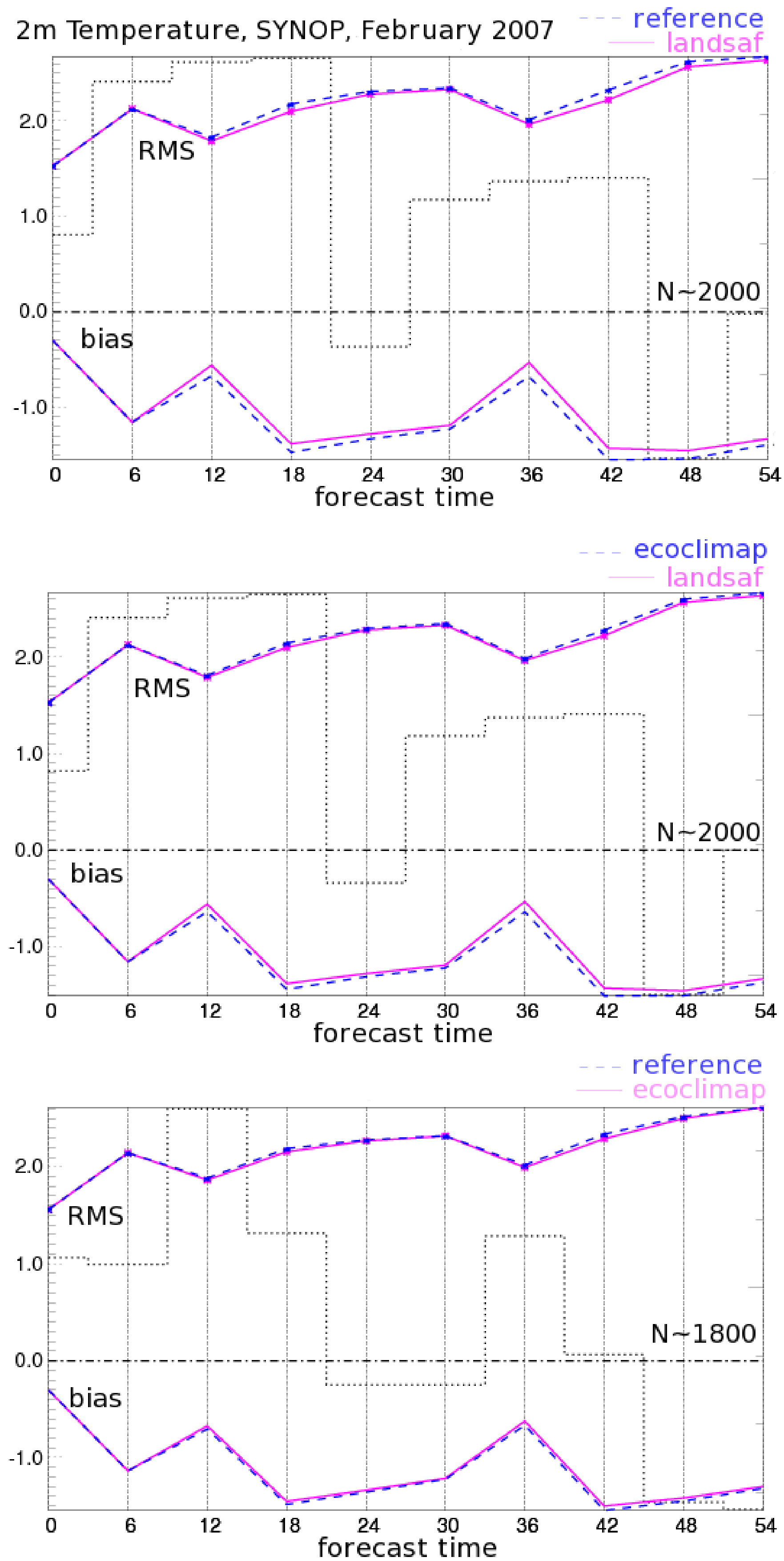

Figure 3: Statistical scores (bias and RMSE) for 2m SYNOP temperature for February 2007 for three pairs of experiments: a run with analyzed albedo ("landsaf") compared to ALADIN "reference" (top), analyzed albedo run ("landsaf") compared to ALADIN with albedo values from ecoclimap ("ecoclimap") (middle) and ALADIN with ecoclimap ("ecoclimap") values of albedo compared to a "reference" ALADIN run (bottom).

However, scores for early summer season are less pronounced. As it can be seen from **Figure 4**, the impact is rather neutral. When looking at a particular case (**Figure 5, left**) one observes a rather strangely looking noisy pattern for the difference of the two 2m temperature fields (albedo analyzed and reference run). The emergence of this pattern is peculiar especially due to the fact that the new albedo field difference (**Figure 5, right**) (which is the only thing we modify) does not exhibit any similar pattern. A very likely explanation for this is that very small differences in albedo can cause small differences in temperature which differently feed the convection and turbulence schemes which are in turn producing different cloud cover and even small differences in precipitation amount. And due to these differences in radiative forcing and latent heat release the final temperature fields can be quite different with great fluctuation from one grid point to another. See **Figure 6, left** for the difference in the cloud cover in one case. Further evidence that the noise is caused by convection is that at forecast time +6 (in the morning) the noisy pattern is practically non-existing (**Figure 6, right**).

Notice also that the average difference for February is much higher than in June.

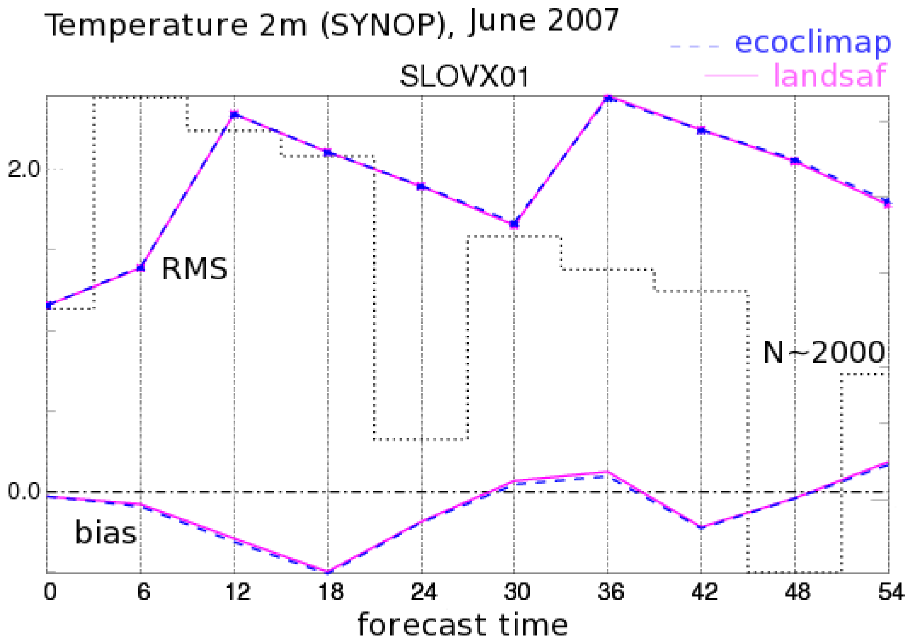

Figure 4: Same as **Figure 3**, but for June 2007 and only comparison of experiment with analyzed albedo ("landsaf") and the run with ecoclimap ("ecoclimap") albedo values in ALADIN.

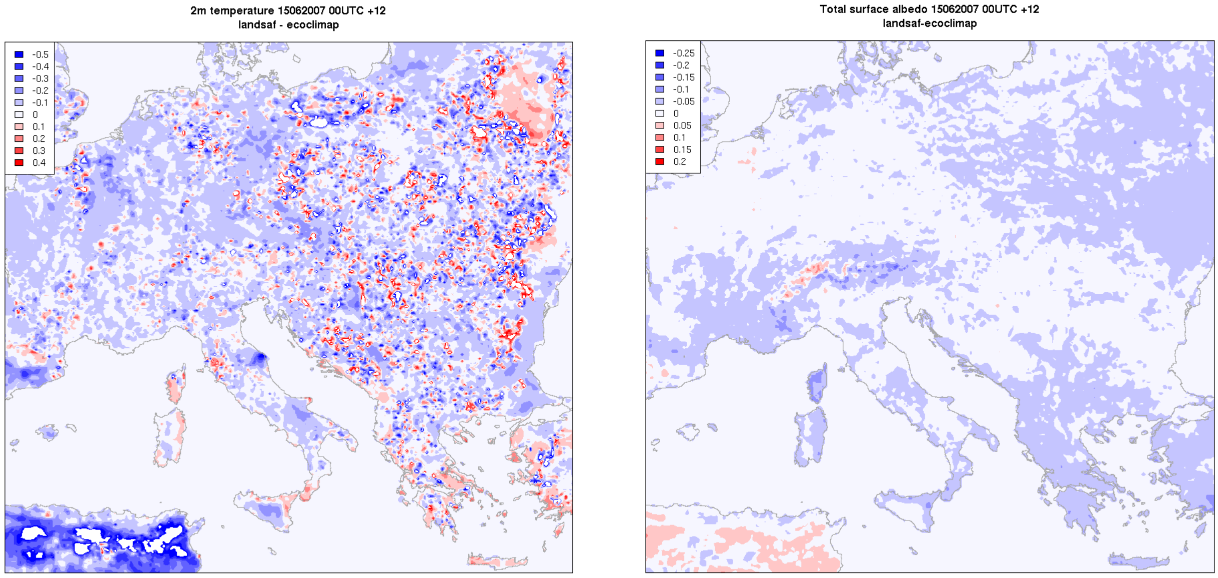

Figure 5: Same as **Figure 2** but for 15th of June 2007.

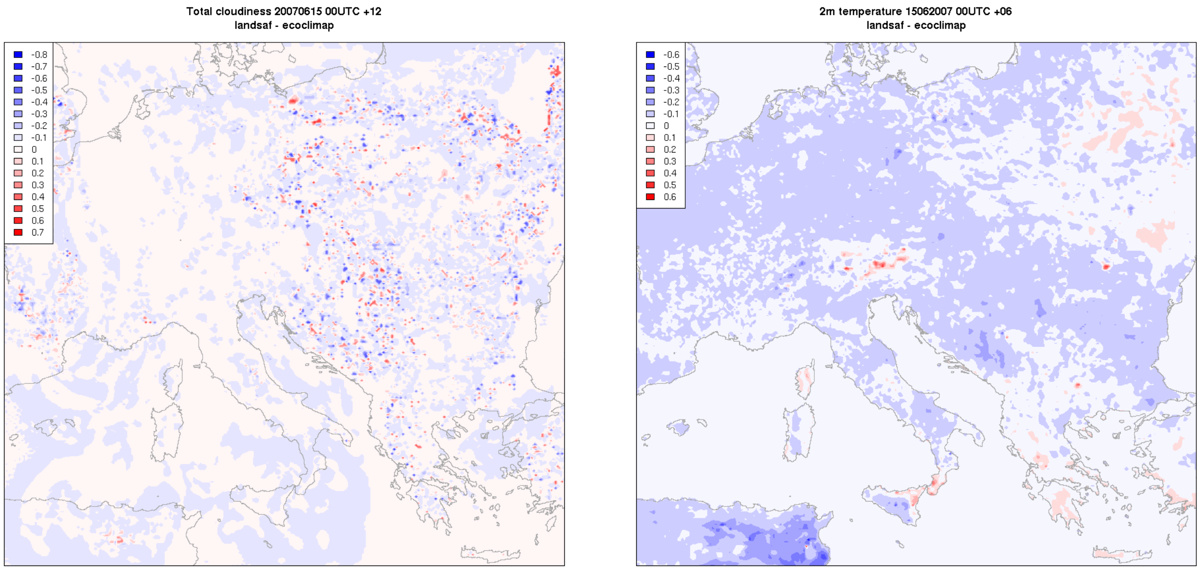

Figure 6: The difference in total cloud cover at noon (+12 hrs) (left) and early morning 2m temperature (at +6 hrs) for an early summer day (15th of June 2007) for the run with analyzed albedoes ("landsaf") and the run with ecoclimap ("ecoclimap") value of albedo.

4. Results – using ALADIN climatology for albedo analysis (or consistent)

Though the results in the previous section are quite convincing it is important to notice, that the differences between the climatology of ecoclimap and the one of ALADIN exist and even might be quite significant. And it is not only the two climatological albedo values that matter, one of the largest climatological differences is the vegetation fraction (**Figure 7**). The vegetation fraction is important because it appears in the obs operator of the Kalman filter analysis.

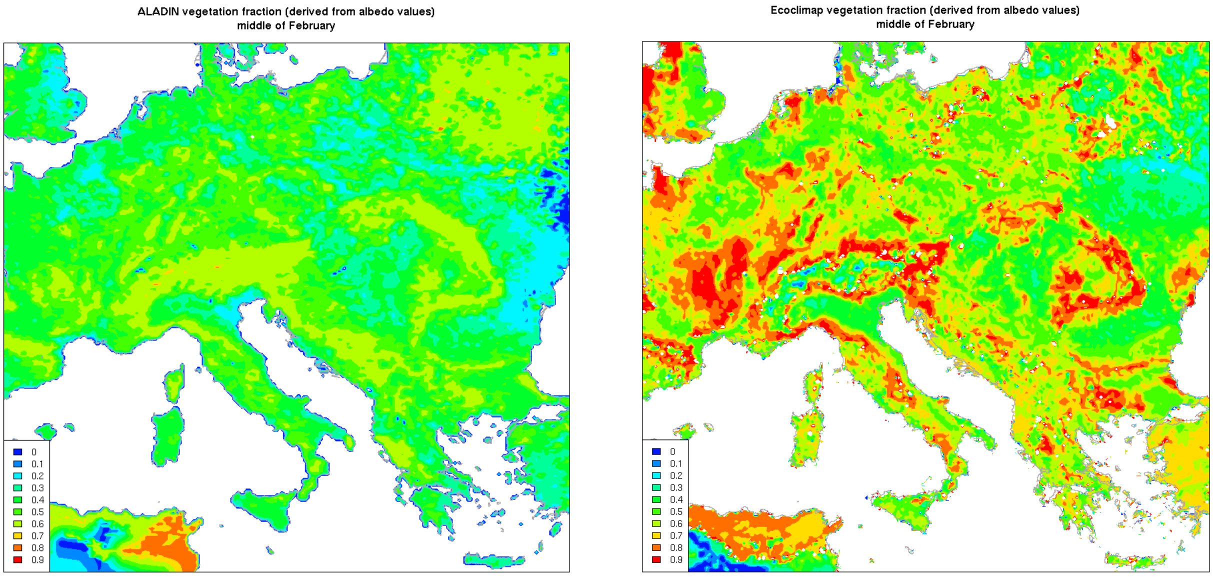

Figure 7: The vegetation fraction in ALADIN (left) and in ecoclimap (right) for February. Both are derived from the vegetation, bare soil and total albedo values and both are scaled to ALADIN grid and the legend of the plot is the same.

In spite of these large differences, it can be seen that the consistent experiments' results (**Figure 8**) seem very similar to the results in the previous section. And the same goes for the statistical scores (**Figure 9**).

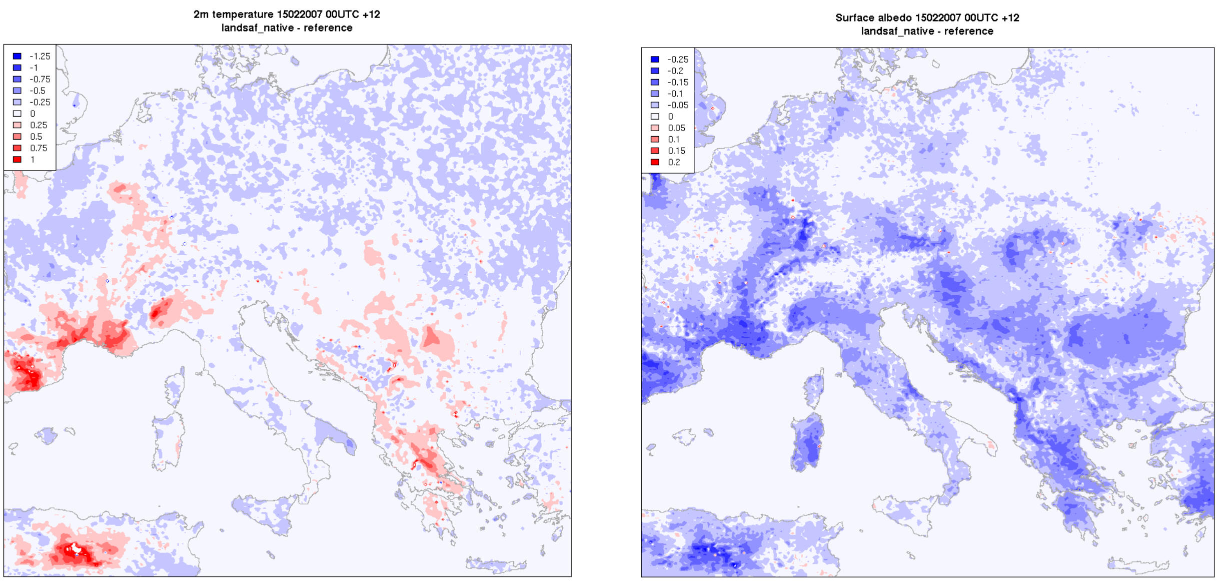

Figure 8: Same as **Figure 2,** but for LandSAF assimilation in native ALADIN LCC grid.

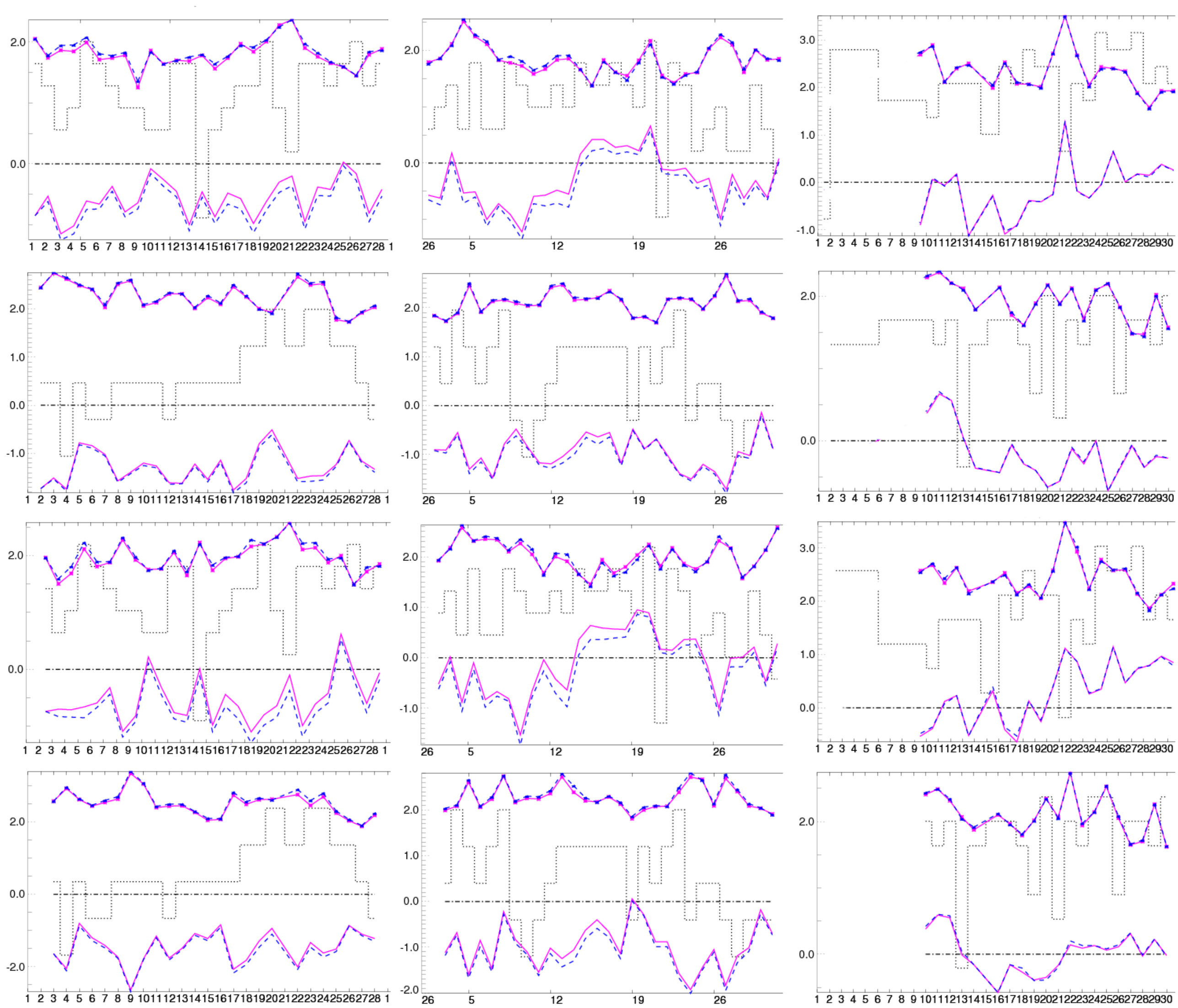

Figure 9: Time series of 2m temperature scores for February (left column), March (middle column) and June (right column). The first row is a 12 hour forecast, second row 24 hour, third one 36 hour and last row is for 48 hour forecast. Lower two lines on each graph are biases and upper ones are RMSE. The full magenta line is the assimilation run and the blue broken line is the reference.

Furthermore it can be seen from **Figure 9**, that the model has a significant cold bias in winter time and satellite data is reducing this bias. However this bias correction is more pronounced during day-time: at forecast times +12 and +36 (rows 1 and 3 in left column of **Figure 9**). The overall impact (in terms of statistical scores) of satellite data analysis is practically negligible in June (right column).

An interesting phenomena can be observed in timeseries of scores in March. It appears that the influence of satellite data is particularly strong from 10th to 20th of March when the gap between the two runs is quite large. During those days the run with albedo assimilation is “overcorrecting” the bias. This process is yet to be investigated, but the first guess is that it must have something to do with snow-melting: **Figure 10** shows how the snow cover decreased during those days of March.

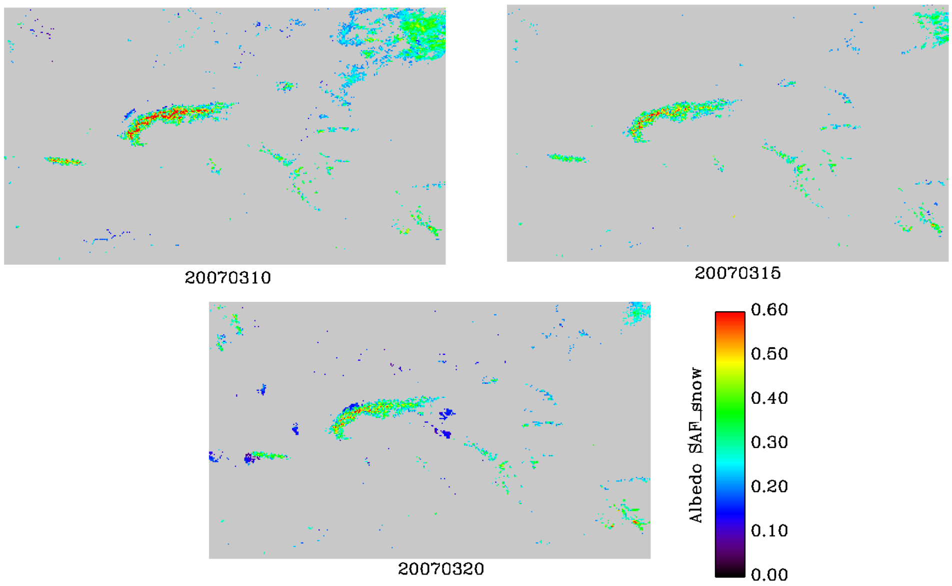

Figure 10: The albedo of snow for 10th (upper left), 15th (upper right) and 20th (bottom) of March. The domain is latlon over central Europe. Perhaps the only geographical marker is the comma shaped pattern which is the Alpine ridge.

Even when considering the non-favorable evolution of scores during the above mentioned period in spring, the average statistics for the entire month looks very similar to what was observed with first "non-consistent" results (**Figure 11**).

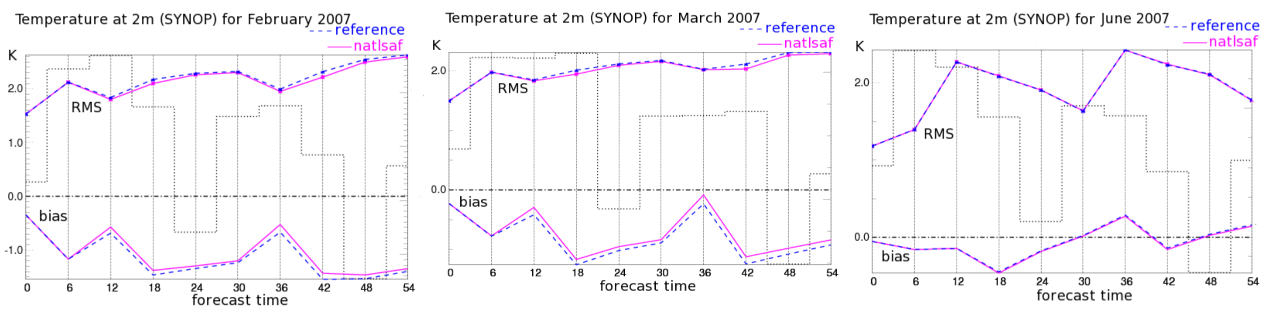

Figure 11: Average monthly scores for February (left), March (middle) and June (right) 2007 for 2m temperature for native grid Land SAF assimilation ("natlsaf") and reference.

5. Conclusions and further work

The use of Land SAF albedo product as the input for surface assimilation in ALADIN clearly has a positive impact in winter time and (so far it seems) neutral impact in summer time. Some situations need to be further examined, particularly time evolution of temperature and albedo for some grid points. Another scope of evaluation would be by using fluxes at certain gridpoints. There is also a need to do some tuning: it seems that the uncertainty values of the albedo product are pretty low and hence much more weight is put to the satellite.

The work presented here clearly has the potential for operational use. The code and the scripts are available and only minor modifications would be necessary for operational use.

Furthermore, it is meaningful to say that one would want to include snow cover analysis and snow albedo analysis in the same manner; particularly the snow cover is one of the points where improvement would be most urgent (ALADIN is currently using climatological snow cover).

Another possibility is the exploitation of other Land SAF products which are not necessarily at the same level of sophistication as the albedo (maybe even in experimental phase) – vegetation fraction and LAI, where the former would probably need to be analyzed before albedo (the obs operator for albedo is a function of vegetation fraction).

Appendix: Short technical documentation

A.1 Original Kalman filter software

The initial version of Kalman filter software was written by Dominique Carrer, it used ecoclimap data as the climatology (note that ecoclimap uses tiles – the vegetation area of each grid point is split into 12 (in this case) tiles).

This software is actually a set of tools:

- a look-up table generator (to be used for interpolation from SEVIRI Land SAF grid to latlon grid of ecoclimap),
- an interpolation part (a tool that actually performs interpolation by making use of SEVIRI input file and look-up tables)
- and finally the analysis part (which read all the data and performs Kalman filtering, giving binary files in the output).

The best source of information regarding this package is: dominque.carrer@meteo.fr

A.2 Tool to intepolate and inject fields in ALADIN file

For the first part of the evaluation experiment I used a program called injectalbedo. This is a simple program for manipulation of FA files. It opens an ALADIN file, extracts coordinates of ALADIN LCC grid, reads in the albedo analysis (for bare soil and vegetation) in binary format in ecoclimap latlon grid and performs a bilinear interpolation of the values of four nearest neighbors of latlon grid to every ALADIN grid point and finally replaces the field in ALADIN file by the values obtained with interpolation from albedo analysis.

A.3 Software for assimilation in native ALADIN (LCC) grid

After the initial tests with assimilation in ecoclimap grid turned out promising, I coded a version of the software similar to 1) that performs the analysis on ALADIN LCC grid.

The programs are the following:

- SEVIRI2LCC_lookup: creates a look-up table for SEVIRI grid from ALADIN LCC grid; for every point in ALADIN, the closest satellite point is found and the two indexes of this point are written in two files (LOOKUP_FILE_LCC-SEVIRI_LAT and LOOKUP_FILE_LCC-SEVIRI_LON) in binary format
- SEVIRI2LCC_interp: the two look-up files from previous point are used to interpolate SEVIRI HDF5 data (albedo, error of albedo and quality flag) to ALADIN grid, again the output is a binary file
- KALMAN4ALBEDO: the program that does the analysis, a modified version of the third line under 1).

On top of that, another script was written that merges all the three parts in a time loop and controls the input and output. So that the programs themselves are only tools and the script is the main part.

A.4 Performed experiments and their names

All the experiments were performed on Slovenian operational ALADIN domain (covering central Europe, the Alps and central Mediterranean area, see for example Figure 1) with a recent Météo France ALADIN setup (cy32t0.oper). The names of experiments, taken from the OLIVE environment, are in Table 1.

All the results of the experiments can be found on cougar: ~mrpa669/xp/$EXP. The coupling files were produced only with the reference run, the other runs used the same coupling files to minimize the computational cost.

Table 1: Description of the experiments used and their names as used by OLIVE.

Exp. name	Description
62ZY	A reference run, from 1/2/2007 to 31/7/2007
630L	Same as reference, but with LandSAF analysed (ecoclimap climatology was used for analysis) albedo values “injected” in INIT file
631D	Same as reference, but with ecoclimap albedo values “injected” in INIT file
636I	Same as reference, but with LandSAF analysed albedo values in INIT file, where the analysis was performed in native grid and model climatology was used

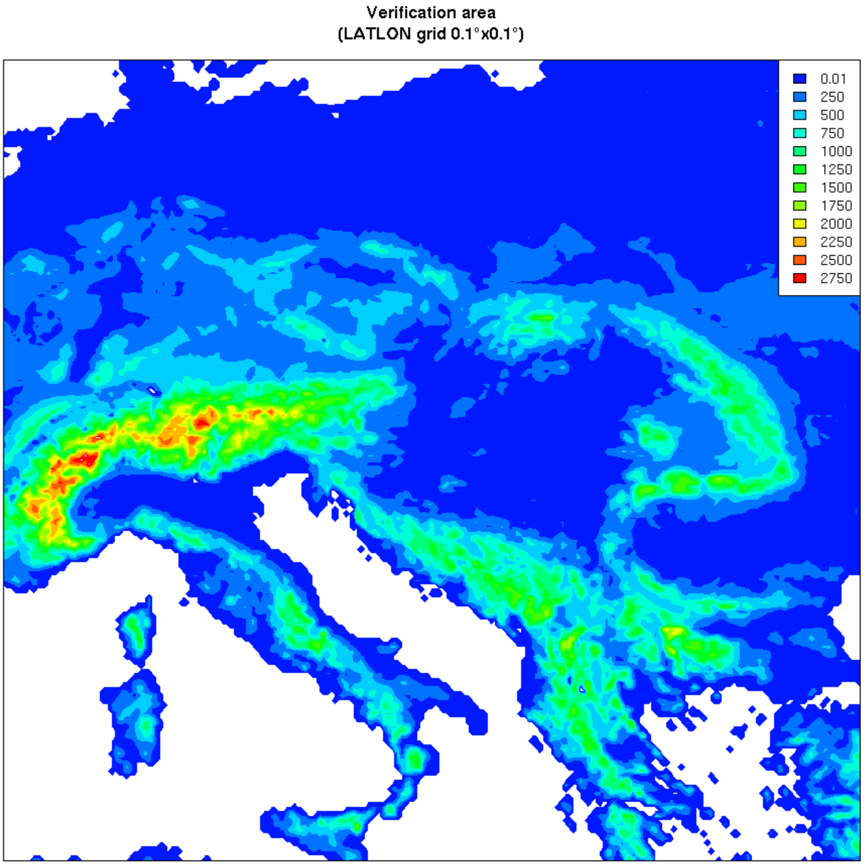

Figure 10: The verification domain size used for the computation of statistical scores. The shaded colors are height of orography [m].