Workshop on SURFEX data assimilation (5-6 March 2012, Toulouse)
Minutes of the discussions (written by Jean-François Mahfouf)

Attendees : J.-C. Calvet (GMME), E. Martin (GMME), E. Brun (GMGEC), A. Duerinckx (RMI), R. Hamdi (RMI), S. Lafont (GMME), M. Parrens (GMME), D. Carrer (GMME), J.-L. Roujean (GMME), M. Dumont (CEN), S. Morin (CEN), S. Jebali (GMAP), F. Taillefer (GMAP), V. Pedinotti (GMME), F. Bouyssel (GMAP), C. Soci (GMAP), A. Barbu (GMAP), E. Kourzeneva (FMI-RHSU), M. Homleid (MetNo), T. Aspelien (MetNo), W. Lahoz (NILU), T. Svendby (NILU), E. Bazile (GMME), J.-F. Mahfouf (GMME), R. El Khatib (GMAP), C. Fischer (GMAP), S. Faroux (GMME)

The presentations given during the first day concerned two main classes of surface analysis tools:

- A 2D optimum interpolation scheme based on CANARI (for T2m, RH2m analyses, SST analyses, SWE analyses, and to be adapted for LST and surface precipitation analyses) – Developments on CANARI OI take place at MetNo (M. Homleid) and at Météo-France (C. Soci).
- A number of 1D soil/vegetation assimilation schemes based on OI, EKF, SEKF, STAEKF, EnKF

They are available on the following web site:
http://www.cnrm.meteo.fr/aladin/spip.php?article239

I provide hereafter a summary of discussion that took place during the second day of the workshop about possible evolutions of the SURFEX data assimilation systems from scientific and technical points of view.

I proposed the acronym **SODA** (SURFEX Offline Data Assimilation) to design the surface analysis tools developed around SURFEX.

We discussed a number of current SURFEX developments that could be of interest for surface data assimilation:

- The multi-layer soil version (ISBA-DIF): Developments will be undertaken very soon by M. Parrens (The Jacobians should be examined and compared to those ones from the ISBA 2L scheme: not discussed during the meeting)
- The double energy balance version (ISBA-MEB): Important for snow/vegetation interactions and also for the ISBA-A-gs version (for an improved description of the radiative transfer within the canopy that will correct a number of known weaknesses) – to be used at GMME/VEGEO in 2013.

We discussed various issues related to the efficiency of the PGD, PREP and OFFLINE applications (relevant but not specific to SODA):

- OFFLINE: a MPI driver will be written (***Action: S. Faroux and T. Aspelien***). An action has started on the suppression of « global variables » in SURFEX (CERFACS) in order to allow OPEN-MP capabilities without explicit « THREAD PRIVATE » instructions in the code.
- PGD and PREP applications are not parallelized. There is a need for an efficient tool to interpolate SURFEX surface fields from one grid to another. SURFEX file (either a more efficient version of PREP or a modified version of FULL-POS that could accommodate for the « tile approach ») (***Action: R. El Khatib or P. Marguinaud ? We should also explore with C. Fischer and P. Termonia the possibility of having this action undertaken by an ALADIN***)

The current EnKF has been developed by NILU with SURFEX V4.8. T. Svendby and W. Lahoz agreed to upgrade their system to the most recent SURFEX version: V7.2 and to make it available in a next SURFEX version release (***Action: S. Faroux and T. Svendby, in coordination with A. Barbu for compatibility with EKF***)

It has been agreed to provide a first version of the EKF in the next SURFEX version (to be delivered in October 2012 for the preparation of CY39, which will be the last cycle for a parallel suite during the first semester of 2013 before the change of supercomputer at Météo-France). This version will be based on the one used by A. Barbu (SURFEX V6++) for the GEOLAND2 project. It will include both the assimilation of screen-level observations, LAI and superficial soil moisture content and also the option with « 12 patches » (necessary with ISBA-A-gs ; also useful for model comparison with point scale observations). It was recognized that the so-called « OI_MAIN » assimilation should be merged with the « EKF_MAIN » under a common driver. Moreover, the analysis should be split in various subroutines according to the surface tile (nature, town, water, sea) (***Action: T. Aspelien***). The observation files should be split in with one file per observation type (***Action: A. Barbu***).

Once a version of the EKF system is available in SURFEX (version 7.3) we agreed to add the following capabilities in the following versions:

- Lake data assimilation (currently done outside SURFEX by E. Kourzeneva)
- The EnKFs developed at NILU (also PF ?)
- The STAEKF developed at RMI (for parameter adjustment)
- An adaptive EKF (for a dynamical estimation of the bias correction)

The interest of Météo-France to include the SURFEX analyses (SODA) in CANARI has been expressed (code parallelization, reduced I/O, interface with ODB, maintenance and evolutions through the ARPEGE/IFS environment). In particular with the availability of SURFEX files in FA format, the OI soil analysis can be called from CANARI and will be available in CY38T1 (***Contacts: P. Marguinaud and F. Taillefer***). The possibility of including the EKF within CANARI should be explored (***Action: F. Taillefer and R. El Khatib***). C. Fischer made use aware of the OOPS project (recoding of the ARPEGE/IFS with C++ routines at the control levels) that will affect CANARI in the coming years (could it make the inclusion of SODA easier ?)

It is important for a number of applications outside the NWP environment to maintain an offline version of SODA (i.e. outside CANARI), but to improve its efficiency. However, in that context, the observations cannot be managed through the ODB software (that could not be tested by the SURFEX team). It means that a fully offline version of SODA and a version coupled to CANARI will have to co-exist (therefore, parts of the code that will differ should be clearly identified and isolated to avoid complete duplications).

Even though not discussed extensively, J.-F. Mahfouf mentioned the interest for having in the future a common pre-processing tool that could read observations in native formats (BUFR, HDF5, …) and project them onto the model grid with a number of gross data quality controls. The issue of scale discrepancy between a fine scale model and coarse satellite observations has been mentioned (it can noticed that a 2D EnKF addresses this issue in a natural way but currently the NILU EnKF is 1D).

The extensive use in OI_MAIN of routines that correct for mismatches between a surface type on the native grid and the corresponding one on the target grid (looking for the nearest grid point of same nature) by HIRLAM, should be taken into account before entering the soil analysis (possible weakness of the HIRLAM interpolation software GL ? a revised PREP/FULL-POS should be able to handle such mismatches). Currently the corresponding routine prevents the code reproducibility (communication between processors). When including OI_MAIN in CANARI a solution has been found to avoid its usage (by defining SST field defined over sea and land).

An increase usage of LSA SAF (and H-SAF) products that have spatial and temporal resolution with LAMs has been encouraged by J.-F. Mahfouf both for data assimilation and improved atmospheric forcings.

The link with atmospheric ensemble techniques (assimilation and prediction systems) has been underlined by J.-F. Mahfouf: outputs from such systems could be used to characterize model errors of surface schemes.

A number of issues (or facts):

- Interest in monitoring of NWP soil analyses (done at ECMWF)
- It has been shown that ECOCLIMAP overestimates « permanent snow » cover over Norway. New fields will be provided for correction (***Action: S. Faroux and M. Homleid***)
- The possibility to overwrite in SURFEX surface fields that are computed from averaging cover values should be used more extensively to gain computing time and for including surface analyses.
- Poor current quality of the ESA GlobSnow product (too much constrained by SYNOP kriging as background) – E. Brun informed that the snow model Crocus forced by ECMWF analyses, without using snow observations or analyses, provide SWE simulations over open fields of similar performance as Globsnow. Preliminary discussions have started with Kari Luojus (FMI) in order to evaluate the potential of using Crocus simulations as a background information for future Globsnow improvements.
- S. Morin suggested to M. Homleid to perform the snow analysis with CANARI in terms of snow depth rather than in SWE (using model snow density) ***(Action : M. Homleid and F. Taillefer)***.

- S. Morin suggested to use the results from the research activities of Richard Essery (University of Edinburgh) in order to get ideas when moving from the one layer snow scheme ISBA-D95 to the multi-layer version ISBA-ES
- Current microwave emissivity models are not so useful for snow because of the importance of layering and grain type that are not adequately described in snow models.
- CEN will use Level 1b products than Level 2 products; for example, data from VIS/NIR satellite radiometers (such as MODIS) will be used to assimilate surface snow properties rather than operational snow products (except maybe for the snow cover extent). The same will apply to MW data.
- Possibility of using PALM (already used in a number of hydrological applications at CNRM): flexibility for research applications and very good support from CERFACS.

7-8 March: Discussion of the HIRLAM delegation with GMME and GMEGEC for the development of a simple sea-ice model in SURFEX (of interest for the assimilation of satellite radiances over such surfaces).