WG3 : Observation acquisition and observation operators

Draft version 01 (11 September 2007) JFMD
Draft version 02 (1 October 2007) NG
Draft version 03 (22 November 2007) JFM

I. Introduction

Various working groups have been set-up during the SURFEX Workshop in Toulouse (December 2006) in order to coordinate and accelerate HIRLAM/ALADIN activities on land surface modelling and data assimilation

The goal of the present working group (WG3) was to :

1. examine the observation operators used in the present data assimilation systems
2. define a proposal for common observation operator formulations
3. define needs and priorities for obtaining relevant surface observations

First, we recall the common HIRLAM/ALADIN framework in terms of modelling constraints:

- geographical domain : Europe
- model horizontal resolution : 2 to 10 km
- temporal window : 1 to 3 days (forecast range)

to examine what are the data that could be useful in such context.

The definition of observation operators implies to know :

- what are the variables to analyse
- what are the observations that can be assimilated

Required surface variables to be analysed :

- *soil temperature and water content* (mostly deep reservoirs) – relevance of analysis for surface variables (no memory ? bias removal ? use of satellite radiances over land ? how representative is analysis increment for different surface tiles?)
- *snow* (water content, fractional cover, snow temperature?)
- sea surface temperature (SST)
- lake temperature
- sea ice extent (ice albedo (new ice or multi-year ice)? ice temperature? snow on ice?)
- physiographic data : albedo, vegetation cover, LAI
- forcing data sets : surface precipitation, radiative fluxes
-

Only variables in italic are currently prognostic variables, the other quantities are prescribed at the beginning of each forecast (can be considered as evolving slowly during the forecast duration). Within a 5 year perspective, more surface variables may become classified as prognostic variables, for example, the lake temperature.

The current observations are mostly:

- screen-level temperature and humidity, snow depth from the surface network
- satellite and in-situ data for ice extent and SST (derived products from EUMETSAT OSI/SAF are also utilized)

– climatologies for other quantities (e.g. ECOCLIMAP at Météo-France)

II. Description of TASK 1

The surface observation operators can be split in two components :

1 2D spatial interpolation schemes
2 local conversion schemes that link model variables to observations

Depending upon the surface variables that need to be analyzed either the first component is used or both of them. The spatial interpolation schemes are based on simple techniques such as optimum interpolation (CANARI/DIAGPACK at Météo-France, SPAN/MESAN at SMHI – the same basic software is used in HIRLAM) but more recently the 3D-Var ALADIN system has been adapted at Météo-France to produce frequent analyses of variables at screen-level (2m and 10 m). This system called VARPACK uses all observations of the 3D-Var system (surface data, RS, satellite radiances). The statistics have been tuned in order to give more weight to the observations than in the actual atmospheric 3D-Var. Additional levels are also included in the vertical grid in order to avoid vertical interpolations.

The spatial interpolation schemes are used to produce gridded fields of :

temperature and humidity at 2 meters
wind components at 10 m (?)
snow water equivalent or snow depth
sea surface temperature
sea-ice extent
??

The principle of the schemes is similar, but significant differences come from :

– the number of observations (availability see Task 3)
– the specification of background and observation error statistics
– the method for solving the analysis problem (OI vs 3D-Var; box technique vs point technique)
– the quality control on observations (representativeness issues of model land use w.r.t. weather station)

Such specifications depend upon many aspects of the analysis system. The correlation lengths depend on the density of the network as well as on the model resolution used to produce the background field. The specification of the error variances also depend upon the further use of surface analyses (e.g. for diagnostics or for input to the second component of the surface analysis (soil analysis)). These aspects are important when we are to address the second task of this working group.

It is foreseen that such interpolation tools will remain an important component of land surface assimilation and could be used for spatialisation of new surface variables (e.g. Lake surface temperature, Leaf area index).

The second component is already present in the spatial interpolation tools:

– vertical interpolation between the surface and the lowest model level to obtain variables at screen-level.
– Averaging over several surface tiles to obtain the background value in observation points
– conversions of snow depth in snow water equivalent (using model density ?)

The vertical interpolation of variables at 2 meters can be done in three ways in SURFEX:

- Use of the analytic formulation proposed by Geleyn (1988) that requires the knowledge (specification) of surface variables (temperature and specific humidity) that may be difficult to obtain when considering a tile approach for the surface (already in ISBA without tiles, the various evaporation flux components make the definition of surface specific humidity rather complex).
- Use of the Monin-Obukhov similarity profiles that depend only on surface fluxes (various stability functions can be chosen, particularly in stable regime) and on one level in the surface boundary layer
- Use of a multi-layer modeling of the surface boundary layer (Canopy model, Masson, 2008)

A difficulty of the last two SURFEX formulations is the derivation of the linearized versions (tangent-linear an adjoint).

In HIRLAM (Rodriguez et al., 2003), the analytical formulation proposed by Geleyn (1988) is applied. The application of the Geleyn vertical interpolation is preceded by an adjustment of the atmospheric profile to the station height, averaging of surface variables over the land tiles and adjustment of the surface temperature to preserve the lapse rate in the surface layer. The calculation of an effective local roughness length is following Mason (1988). The averaging of surface variables over land tiles has been the origin of difficult problems in HIRLAM, in particular in connection with introduction of a more advanced surface scheme, including also a canopy temperature (What is the 2 meter temperature in a forest?). The averaging over land surface tiles assumes implicitly a local quasi-equilibrium for each surface tile, an assumption that is questionable and that may cause errors in the assimilation phase.

Various problems concern possible inconsistencies between hypotheses for the fluxes computations and the screen-level diagnostics, the specification of surface characteristics (e.g. use lawn values for surface roughness length ?), and the link with atmospheric analyses (using screen-level observations) that require similar observation operators (activities of WG1).

The tools used to provide locally the correspondence between soil variables (temperature, moisture content) and observations are of two types:

- the current optimal interpolation (Mahfouf, 1991; Giard and Bazile, 2000) that provides a linear relationship between analysis increments produced by the 2D spatial interpolation scheme and soil corrections, with calibrated (empirical) coefficients (used both foe ARPEGE/ALADIN and HIRLAM)
- a new simplified EKF scheme developed within SURFEX that computes dynamical OI coefficients (with the advantage of being portable to any model coupled with SURFEX and that can usenew observation types). Such development will make the soil analysis independent of the definition of the land surface scheme (e.g. carbon version of ISBA, multi-layer version of ISBA)

III Description of TASK 2

The definition of common observation operators will have to encompass four aspects:

- the 2D spatial interpolation tool
- the vertical interpolation scheme in the surface boundary layer
- the soil analysis scheme (the use of SURFEX should help, moving from a simplified EKF to an EnKF)
- the microwave emission model (the development of a common model at ECMWF within the NWPSAF should help as well as its inclusion within SURFEX). This new component has to be considered in the near future.

What would be the advantages and constraints of using the same tools for surface assimilation ?

IV Description of TASK 3

We should seek for currently used observational data set and prepare those that could be used in a not so distant future.

- Availability of screen-level observations from national high density networks (e.g. are RADOME data over France distributed in real-time world wide ?)
- Performing the soil analysis within an offline version of SURFEX means that precipitation and radiative fluxes could be used as forcing data. LANDSAF fluxes derived from MSG are are available over Europe. What about precipitation data from national climatological networks, and from radars ?
- Soil moisture data should be available from ASCAT MetOp operationally to NWP centres
- SMOS brightness temperatures will be produced in real-time by ESA and transmitted to ECMWF (availability mid-2009)
- What is the availability of other products (temporal frequency, spatial resolution):
  - lake temperatures (from which instruments ?) – national in-situ lake data
  - surface albedo and LAI (LANDSAF, MODIS) : future analysed products combining satellite data and ECOCLIMAP (ongoing activities at Météo-France CNRM/GMME/MATIS) ?
  - Snow albedo, fractional snow cover (would it be already useful ?)
  - Sea-ice extent (SSM/I mostly) and SST, also EUMETSAT OSI/SAF data are available

Note: The problem of inconsistencies between the surface assimilation and the atmospheric assimilation has been experienced with HIRLAM. For example, the information contained in a surface temperature assimilation increment may be lost, if the same source of information is not used in the atmospheric assimilation. On the other hand, the introduction of detailed surface observations in the atmospheric assimilation may be pre-mature due to too simple assumptions about background errors. This issue of inconsistencies between the surface and atmospheric assimilation modules has been discussed also by WG1, and it was proposed to recommend introduction of the concept of **an external assimilation solver** that could handle constraints originating from both the atmospheric assimilation and from the surface assimilation in order to achieve consistency. In the definition of the joint observation operators, this possibility needs to be considered.