An extended Kalman filter for soil analysis in SURFEX

Annelies Duerinckx[1,2], Rafiq Hamdi[1,2], Jean-François Mahfouf[3] and Piet Termonia[1,2]

annelies.duerinckx@meteo.be

[1]Royal Meteorological Institute, Belgium

[2]Department of Physics and Astronomy, Ghent University, Belgium;

[3]GAME, CNRM, Méteo France, CNRS, Toulouse, France

Introduction

An externalised surface scheme like SURFEX [1] allows computationally cheap offline runs. This is a major advantage for surface assimilation techniques such as the extended Kalman filter (EKF), where the offline runs allow a cheaper numerical estimation of the observation operator Jacobian. In the recent past an EKF has been developed within SURFEX for the initialisation of soil water content and soil temperature based on screen-level temperature and relative humidity observations [2]. The surface serves as a lower boundary condition for the NWP model. Therefore a well initialised surface can considerably improve the boundary layer forecasts [3].

The advantage over the EKF compared to the commonly used Optimum Interpolation (OI) assimilation scheme is that it allows an easier integration of new observation types and uses dynamical gain coefficients based on the Jacobian of the model observation operator. In this poster we compare the Jacobian calculated with offline SURFEX runs and with runs coupled to the atmospheric model ALARO. We document a case of non-linearities that can hamper the linearity assumption of the observation operator Jacobian and cause spurious 2Δt oscillations in small parts of the domain. We propose a filter to remove the oscillations and show that this filter works accordingly.

1. Set-up

Belgian Operational 4 km ALARO domain:

- 181x181 grid points, 4 km horizontal resolution **(fig. 2)**

Spectral limited area model ALARO (cy36t1)

- ALARO = ALADIN + ALARO-0 physics
- Initial and boundary conditions from ARPEGE
- 46 vertical levels
- Timestep 180s

External land surface scheme SURFEX (v5):

- 2-layer ISBA-scheme
- Exchanges fluxes and forcing with ALARO every timestep **(fig. 1)**
- Initialisation with Extended Kalman Filter [2]

Figure 1: schematic overview of the model set-up, used for the forecasts. The surface is initialised by means of an Extended Kalman Filter (EKF).

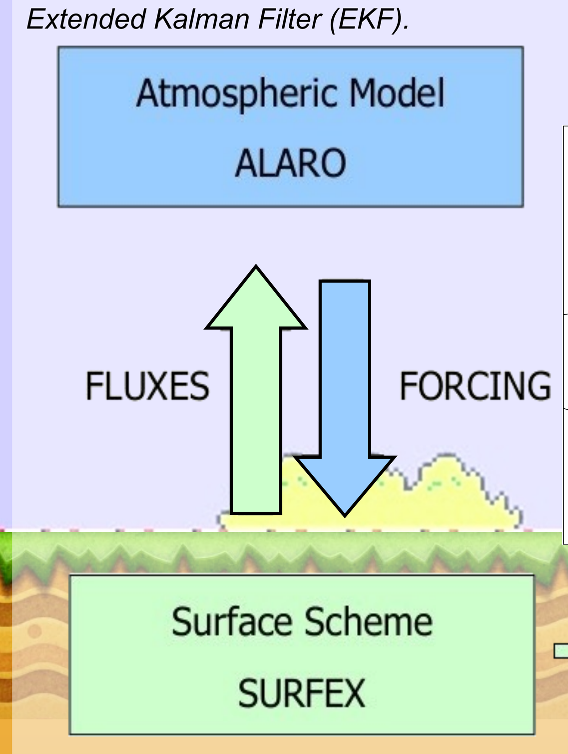


Figure 2: The operational 4 km ALARO Belgium domain.

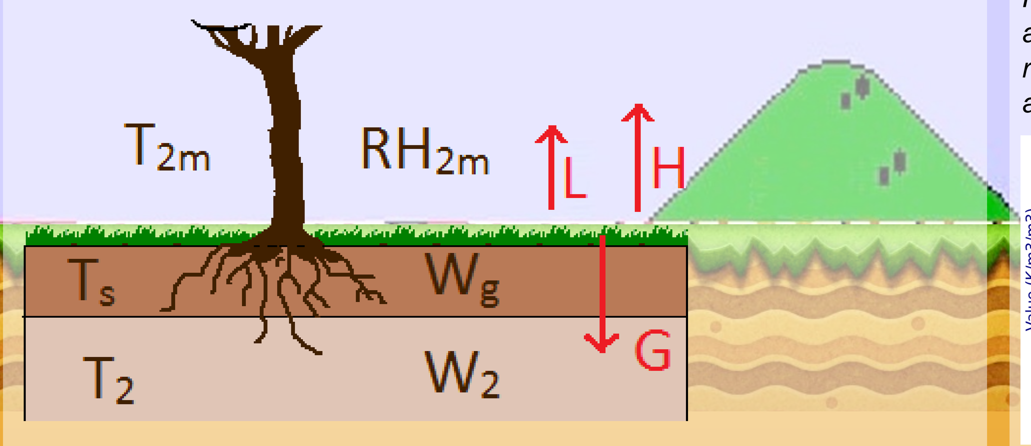
2. The extended Kalman filter

The aim of the EKF is to assimilate observations of screen level temperature (T_{2m}) and screen level relative humidity (RH_{2m}) to initialise superficial and root zone soil moisture content (W_g and W_2) and surface and deep soil temperature (T_s and T_2) **(fig. 3)**. It has been tested here using the same set-up and covariance values as in [2]. The equation of the EKF is the following:

$$x_a^t = x_b^t + \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}[y_o^t - \mathcal{H}(x_b^{t_0})]$$

- The EKF is applied point-wise, i.e. for each model grid point separately
- The observations are interpolated to the model grid before being assimilated
- The background error covariance matrix (**B**) is kept constant. The observation operator $\mathcal{H}$ contains a model state evolution from time t_0 to time t. Therefore, the **B** matrix is implicitly evolved by the linearised observation operator **H**.

Figure 3: Schematic overview of the most important aspects of the surface scheme.



3. The observation operator Jacobian

H is the Jacobian of the observation operator $\mathcal{H}$. This Jacobian allows the EKF to create dynamical coefficients that depend on the specific conditions of each grid point. The numerical computation of the Jacobian uses a finite differences approach in the following way:

$$\mathbf{H} = \frac{\delta y^t}{\delta x^{t_0}} = \frac{y_i^t(x^{t_0} + \delta x_j) - y_i^t(x^{t_0})}{\delta x_j}$$

The calculation of the Jacobian requires one additional perturbed model run for each of the four prognostic variables. There are two possibilities for calculating these runs **(fig. 4)**:

- **Offline approach**: the surface scheme is decoupled from the atmospheric scheme. A previous coupled run provides the hourly forcing for the offline run.
- **Coupled approach:** The surface scheme is coupled to the atmospheric scheme and they exchange fluxes and forcing every timestep.

A comparison of the perturbation sizes **(fig. 5)** shows that the offline approach allows smaller perturbations so that the linearity assumption for the calculation of the Jacobians with finite differences is better approximated.

Figure 4: Schematic overview of the coupled and offline set-ups, used for the perturbed runs of the EKF.

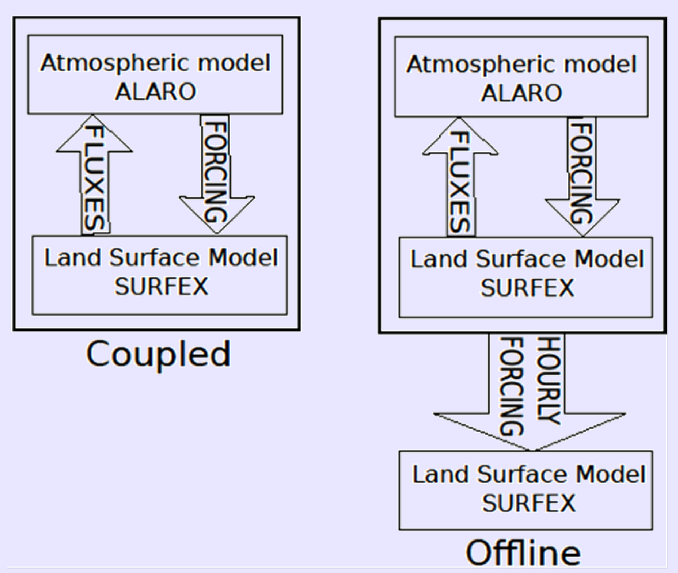
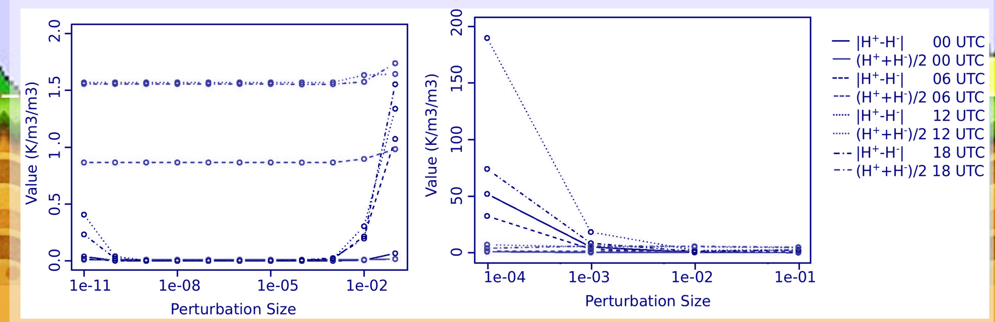


Figure 5: Comparison of the optimal perturbation size for the offline (left) and coupled (right) approach. $|H^+ - H|$ (black) and $(H^+ + H)/2$(red) for different perturbation sizes on 2 July 2010 at 00:00, 06:00, 12:00 and 18:00 UTC averaged over the whole domain with $H = \partial RH_{2m}/\partial W_2$



4. Oscillations cause noisy Jacobian

A 2Δt oscillation in the Jacobian value evolution introduces noise in the increments. The oscillation can be linked to the stability parameters and the formation of a stable boundary layer in the late afternoon:

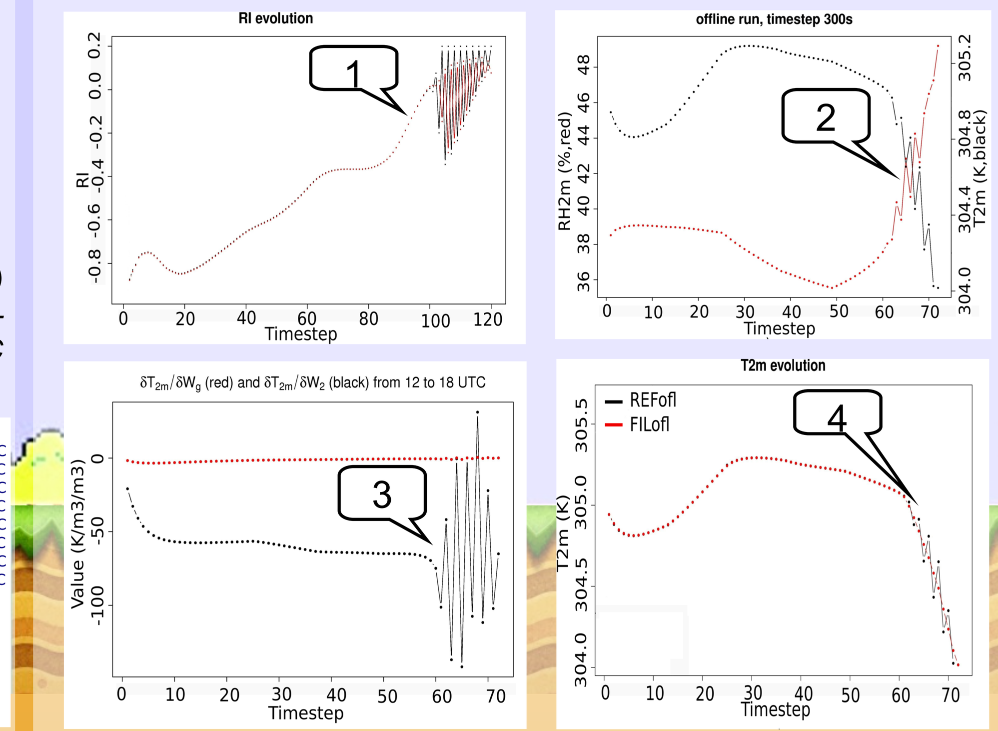
- Oscillations in the Richardson number (RI, 1) introduce oscillations in T_{2m} and RH_{2m} 2 which in turn are reflected by the Jacobian values 3.
- They are ppresent in the offline and coupled approach of the Jacobian and for different time step intervals
- They occur in a limited number of grid points. They do not have a detrimental effect on the performance of the model run, but can introduce noise locally into the Jacobian of the EKF.

- We propose **a temporal filter** for the reference and perturbed values of T_{2m} and RH_{2m} :

  $x_{filtered} = 0.5 \times w \times x_t\text{-}1 + (1\text{-}w)\ x_t + 0.5 \times w \times x_t\text{+}1$

  with x the T_{2m} or RH_{2m} value to be filtered, *t* indicating the time step and *w* the weight attributed to the different parts of the filter.
- A number of values for w have been tested and a value of 0.5, the most optimal choice for filtering the 2Δt mode, appeared to filter out the oscillation best.
- The filter is successful in removing the oscillation 4 .

Figure 6: (top, left) Evolution of the Richardson number (RI) at location A during a 6h coupled run for 2 July 2010 from 12:00 until 18:00 UTC. (top right) Evolution of T_{2m} (black) and RH_{2m} (red) during a 6h offline run for 2 July 2010 from 12:00 until 18:00 UTC at location A. (bottom, left) The evolution of the Jacobian value of $\partial T_{2m}/\partial W_g$ (red) and $\partial T_{2m}/\partial W_2$ (black) during the same offline run at location A. (bottom, right) Evolution of T_{2m} (black: unfiltered, red: filtered) at location A during the 6h offline run for 2 July 2010 from 12:00 until 18:00 UTC.



5. Forecast Scores improve

- The filter is capable of removing noisy Jacobian values where they occur **(fig. 7)**
- The filter slightly improves the RH_{2m} RMSE and BIAS scores for the station of Beitem **(fig. 8)**
- Averaged over 13 stations above Belgium, the filtered runs give a small improvement in scores for RH_{2m} and similar scores for T_{2m} **(table 1)**. In the coupled case the improvement in RH_{2m} scores due to the filtering is larger than in the offline case.

Conclusions

- We identified 2Δt oscillations during the late afternoon when a stable boundary layer sets in. The oscillations can introduce noise locally into the Jacobian of the EKF.
- We proposed and tested a temporal filter to deal with these oscillations.
- Results show that the filter is sucessful in removing the oscillation and produces a small improvement in the RH_{2m} scores for the offline as wel as the coupled approach.
- Due to limited computational resources, our preference goes to the filtered, offline approach and not the coupled approach.

Figure 7: Forecast scores (RMSE and BIAS) for RH_{2m} in Beitem (Belgium) averaged over July 2010 for the reference run without filtering (black, REFofl) and the run with filtering (red, FILofl)

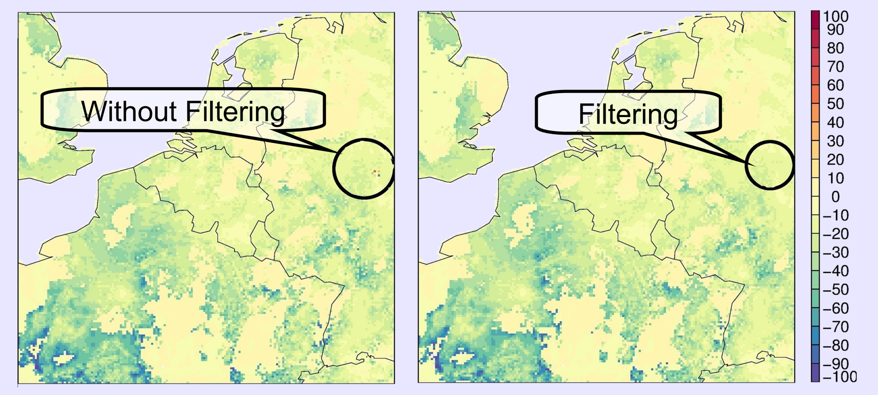


Figure 8: Map of the Jacobian value for $\partial T_{2m}/\partial W_2$ for 6 July 2010 for the reference run (top) and the filtered run (bottom)

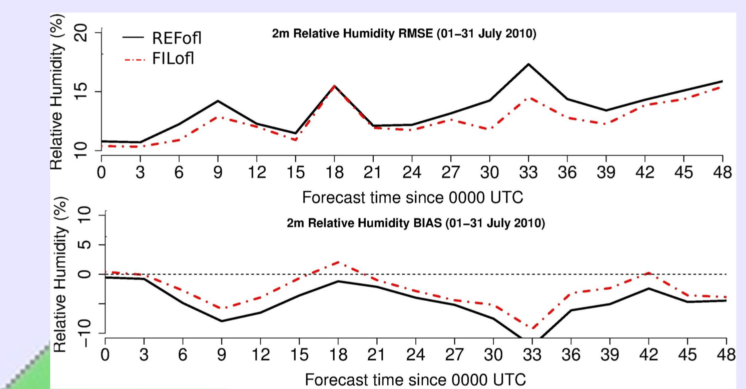


Table 1: Overview of the RMSE and BIAS scores for T_{2m} and RH_{2m} averaged over 13 stations and over July 2010.

	Offline		Coupled	
	REFofl	FILofl	REFcpl	FILcpl
T2m RMSE (K)	2.2	2.2	2.2	2.2
T2m BIAS (K)	0.1	0.1	0.1	−0.1
RH2m RMSE (%)	15.2	15.0	15.2	14.5
RH2m BIAS (%)	-4.9	−4.6	−4.5	−3.2

References:

[1] Masson et al. (2013), The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes, Geosci. Model Dev., 6, 929–960, doi:10.5194/gmd-6-929-2013.
[2] Mahfouf et al. (2009), A comparison of two off-line soil analysis schemes for assimilation of screen level observations, J. Geophys. Res.-Atmos., 114, D08105, doi:10.1029/2008JD011077.
[3] Drusch and Viterbo (2007), Assimilation of screen-level variables in ECMWF's Integrated Forecast System: a study on the impact on the forecast quality and analyzed soil moisture, Mon. Weather Ref., 135, 300-314.

This poster is based on Duerinckx, A., Hamdi, R., Mahfouf, J.-F., and Termonia, P., 2015, Study of the Jacobian of an extended Kalman filter for soil analysis in SURFEXv5, Geosci. Model Dev., 8, 1-19.