The potential of an Extended Kalman Filter for Soil Analysis in conjunction with a 3D-var system in a Limited Area NWP Model.

Extended abstract of the doctoral thesis, December 2015

Annelies Duerinckx, University Gent, Faculty of Physics and Astronomy
Promotors: Prof. dr. Piet Termonia and dr. Rafiq Hamdi

link full PhD thesis : https://weersprookjes.files.wordpress.com/2015/12/phd.pdf

1 Introduction

The surface exerts an important influence on numerical weather predictions (NWP), especially for the planetary boundary layer (PBL). A good initialisation of the surface can therefore improve the short and medium-range forecast scores considerably. The link between the surface and the atmosphere is made by fluxes that transport energy and momentum between the surface and the atmosphere. The fluxes are regulated by the soil temperature and soil moisture content, that regulate the partitioning in sensible and latent heat flux. To initialise the surface, data assimilation techniques can be used. Those techniques combine observations with model data to estimate the real state of the system, or in our case the surface. Due to a lack of direct observations of soil temperature and soil moisture content, the assimilation process uses indirect observations of screen-level temperature and relative humidity. These screen-level observations are influenced by the surface fluxes and thus contain information about the state of the soil. A commonly used data assimilation technique for the surface is Optimal Interpolation (OI). Despite its operational usage in numerous NWP-centers, it has a few important shortcomings. The OI coefficients are pre-calculated and so they do not depend on the specific location or weather situation. Moreover, these pre-calculated coefficients make it cumbersome to include new observation types, like satellite observations. Recently an EKF has been developed for surface assimilation in the surface model SURFEX that meets these shortcomings. In the EKF the coefficients are calculated in an ad hoc manner, so their values are dependent on the specific location and weather situation. Moreover, the more general and ad hoc manner of calculating these coefficients allows the EKF to be more easily extended towards new observation types.

In this research the EKF is validated for the Numerical Weather Prediction (NWP) model ALARO coupled to the surface model SURFEX. The EKF is combined with a three dimensional variational (3D-var) assimilation for the upper-air and the added value of this combination, with respect to surface or upper-air assimilation separately, is investigated. The combination of the EKF with 3D-var for a limited area model is a new one that, to our knowledge, has not been tested before. The purpose of this thesis is to find an optimal set-up for the initialisation of the operational NWP-model of the Royal Meteorological Institute (RMI) of Belgium, that is currently initialised using an interpolation of the global ARPEGE data assimilation analyses for both the surface and the atmosphere.

The research is build up in in four stages. First, the importance of the surface for the planetary boundary layer and the upper atmosphere is described and the surface model SURFEX coupled to ALARO is validated above Belgium within this context. Next, data assimilation theory is discussed, with special attention for OI and the EKF as candidate assimilation techniques for the surface and 3D-var as a technique for the upper-air. The third stage is a thorough validation of the EKF, including a search for the optimal perturbation sizes for the finite differences calculation of the Jacobian of the observation operator. A comparison is made of the offline and coupled finite differences approach of calculation the Jacobian. In a fourth and final stage, the EKF is combined with a 3D-var upper-air assimilation and this combination is compared to a number of other set-ups. The comparison is made with regards to increments and validated with observation of the soil, the screen-level temperature, screen-level relative humidity, atmospheric soundings and precipitation observations.

2 Results

The validation of SURFEX

The validation of SURFEX show that SURFEX improves the forecast scores compared to the current operational ISBA surface scheme. The results over Belgium show that the introduction of SURFEX either shows improvement for or has a neutral impact on the 2m temperature (cfr. Figure 1), 2m relative humidity and 10m wind. However, it seems that SURFEX has a tendency to produce a too high maximum temperature at a high-elevation station during winter daytime, which degrades the scores. In addition, surface radiative and energy fluxes improve compared to observations from the Cabauw tower. It was found that the use of SURFEX has a neutral impact on the precipitation scores. Overall, it can be stated that forecast performance can be improved on average when using SURFEX in ALARO.

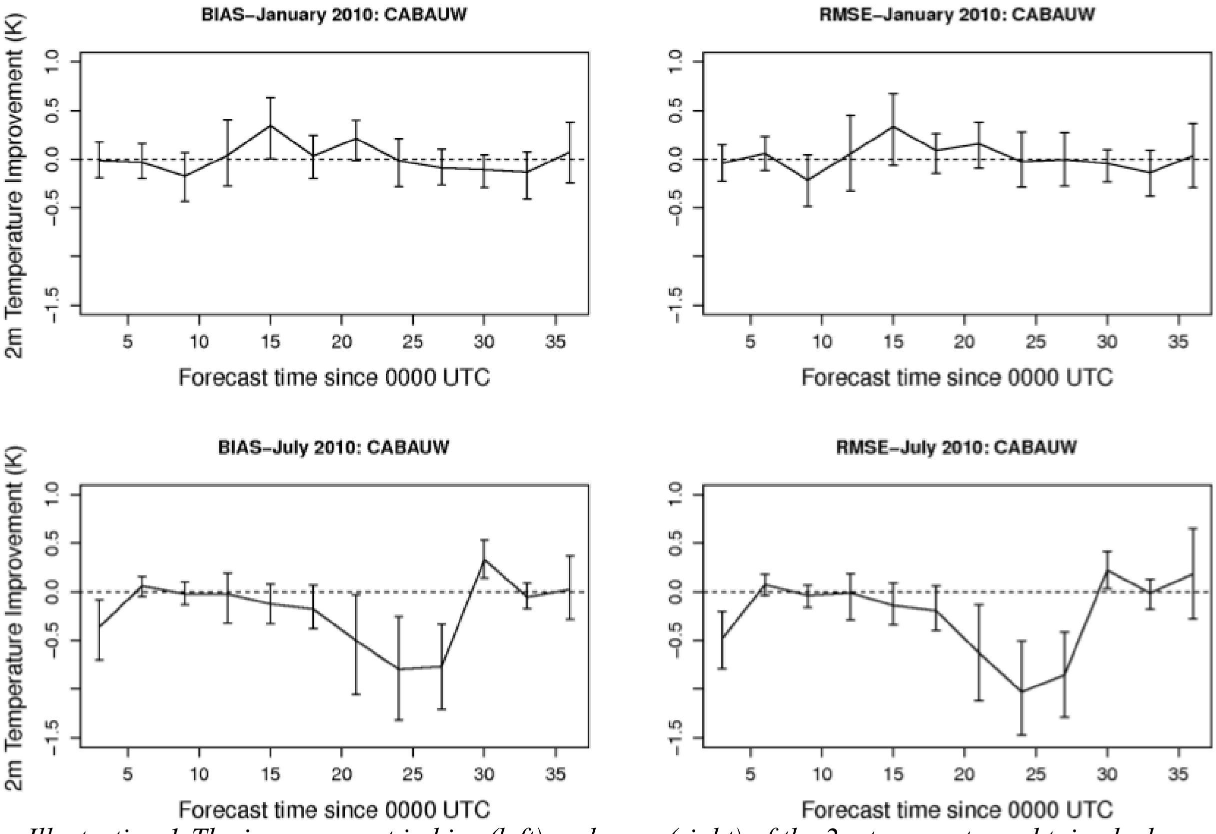

Illustration 1:The improvement in bias (left) and rmse (right) of the 2m temperature obtained when using SURFEX for January (top) and July (bottom). The 95 % confidence intervals for |biasOPER + SFX| − |biasOPER| and rmseOPER + SFX − rmseOPER were calculated with the bootstrap method.

The Jacobian of the observation operator of the EKF

The study of the observation operator Jacobian of the EKF shows that the offline and coupled approach have similar spatial patterns and values. Still, the offline approach has a few advantages over the coupled one. Firstly, the offline approach allows for smaller perturbation sizes due to which there is a better validity of the linearity assumption of the finite differences approach (see figure 2). Moreover, the offline approach is computationally much cheaper, allowing it to be used in an operational setting. A case of spurious 2-delta-t oscillations is documented. The oscillations arise in the late afternoon in Summer when a stable boundary layer sets in. It can be linked to an oscillation in the Richardson number (see figure 3) that is reflected in a number of surface related variables, amongst which 2m temperature (see figure 4) and 2m relative humidity, used to calculate the Jacobians. Test ware made with different perturbation sizes, time steps, values of the critical Richardson number and SURFEX coupling methods (implicit vs. explicit), but none of the different settings were able to remove or reduce the oscillation. Although the oscillation disappears again after a while and does not have a detrimental effect on the forecast scores, it introduces considerable noise in the Jacobian of the EKF (see figure 5) and thus in the increments. For this reason a filter was proposed to deal with these oscillations and it is shown that the filter works accordingly (see figure 6). Results show that the coupled, filtered approach gives the best forecast scores. Still our preference goes out to the offline, filtered approach that also improves the non-filtered EKF but is computationally much cheaper and thus more feasible for operational usage.

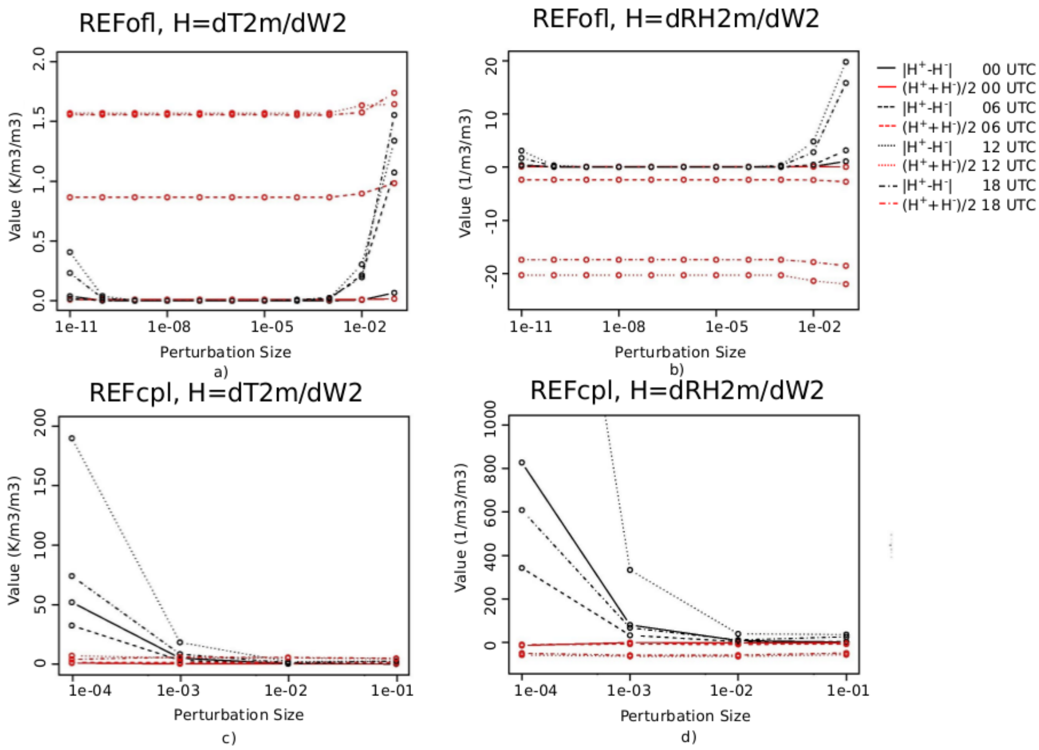

Illustration 2: Comparison of the optimal perturbation size for the offline (top) and coupled(bottom) approach. |H+ − H− | (black) and H+ + H− /2 (red) for different perturbation sizes on 2 July 2010 at 00:00, 06:00, 12:00 and 18:00 UTC averaged over the whole domain with H = δT2m /δWg2 (left) and H = δRH2m /δWg2 (right).

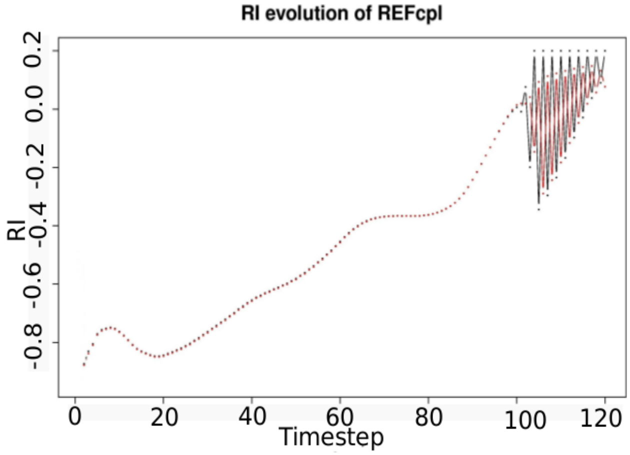

Illustration 3: Evolution of the Richardson number (RI) during a 6h coupled run for 2 July 2010 from 12:00 until 18:00 UTC as it

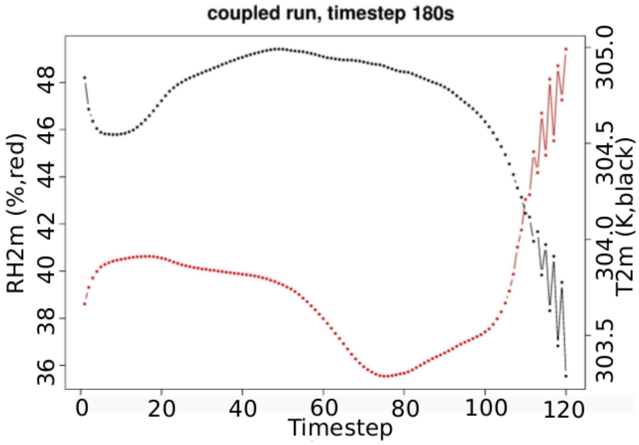

Illustration 4: Evolution of T2m (black) and RH2m (red) during a 6h coupled SURFEX run for 2 July 2010 from 12:00 to 18:00 UTC

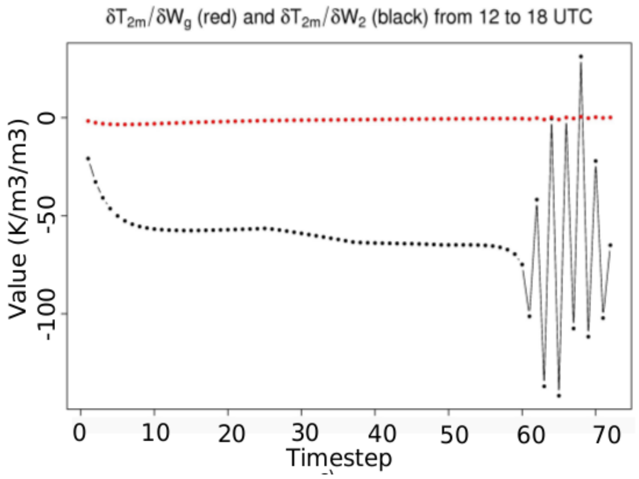

Illustration 5: δT2m /δWg1(red) and δT2m /δWg2 (black) from 12:00 to 18:00 UTC for an offline SURFEX run on 2July 2010.

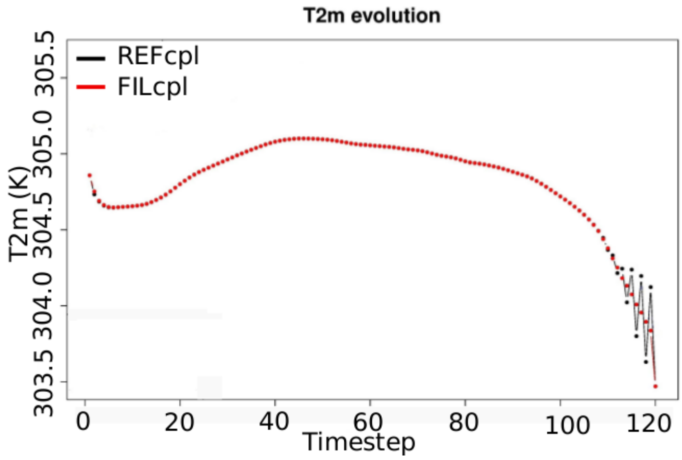

Illustration 6: Evolution of T2m for the coupled reference run (REF, black) and the filtered run (FIL, red).

Validation of the combination of the EKF with a 3D-var assimilation for the upper-air

In the final stage the EKF is combined with a 3D-var upper-air assimilation and this set-up is compared to a number of other initialisation set-ups. Experiments were performed for 1 year and for eight different set-ups. The goal of the verification is to get similar scores as the Open Loop, that uses the interpolated ARPEGE analysis as initial conditions for the surface and the atmosphere as it is done in the correct operational set-up at the RMI. Results show that the planetary boundary layer in the model is in general too cold and too wet, except during summer. The surface assimilation is capable of partly eliminating this bias. The importance of the surface assimilation is confirmed by the much larger bias and root mean square error of the free run , in which the surface is not reinitialised after each assimilation cycle but is allowed to run freely during the whole year. The combination of surface and upper-air assimilation provides better scores for soil moisture content and screen-level humidity (see figure 7) compared to the Open Loop, especially during the first twelve hours of the forecast. Comparisons of the model values with atmospheric soundings and precipitation observations show that the 3D-var assimilation experiments are not able to reproduce the scores of the Open Loop for the upper layers of the atmosphere. This is probably due to a lack of observations, since only conventional observations are used in the assimilation so far. Only during Autumn, the 3D-var assimilation is able to improve the Open Loop scores for precipitation. The scores also show the positive effect of surface assimilation on the precipitation forecasts.

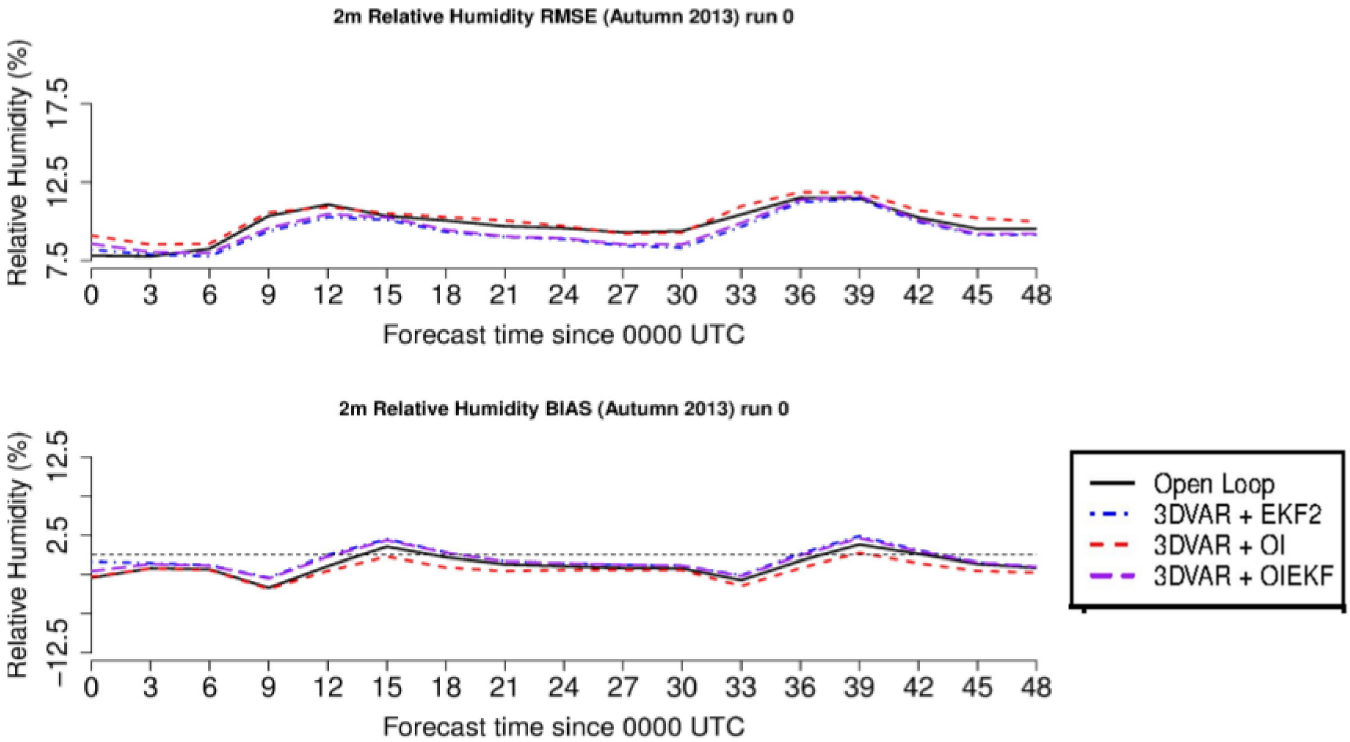

Illustration 7: creen-level relative humidity RMSE and BIAS scores for OL, 3D-Var+OI,3D-Var+EKF and 3D-Var+OIEKF averaged over 13 stations in Belgium for Autumn 2013

3 Conclusion

In general it can be concluded that the surface assimilation, and in particular the EKF, improves the surface and boundary layer humidity compared to the Open Loop. For temperature the results are more mixed, but also in this case the surface In general it can be concluded that the surface assimilation, and in particular assimilation experiments are able to achieve similar results as the Open Loop in most cases. The 3D-Var upper-air assimilation should be improved by a better B-matrix and adding satellite data, GNSS ZTD data and radar data. The advantages of the combination of surface and upper-air assimilation are clear from the improved scores for soil moisture content and relative humidity in the lower parts of the atmosphere, compared to the runs with only surface or upper-air assimilation. As a conclusion for the operational set-up of the RMI it can be said that the surface assimilation runs, and particularly the EKF, are able to get similar or improved scores compared to the current operational initialisation set-up. The 3D-var upper-air assimilation however, should first be improved by using satellite, GNSS and/or radar observations and a better B-matrix before it can be considered for operational usage.