Combining an EKF soil analysis with a 3D-Var upper-air assimilation in a limited-area NWP model: Combining an EKF with 3D-Var in a Limited-Area NWP Model

Article *in* Quarterly Journal of the Royal Meteorological Society · August 2017

DOI: 10.1002/qj.3141

CITATION
1

READS
114

6 authors, including:

Annelies Duerinckx
KU Leuven
10 PUBLICATIONS **40** CITATIONS
SEE PROFILE

Rafiq Hamdi
Royal Meteorological Institute of Belgium
77 PUBLICATIONS **1,251** CITATIONS
SEE PROFILE

Alex Deckmyn
Royal Meteorological Institute of Belgium
34 PUBLICATIONS **261** CITATIONS
SEE PROFILE

Arab Djebbar
Université Catholique de Louvain - UCLouvain
3 PUBLICATIONS **1** CITATION
SEE PROFILE

Combining an EKF soil analysis with a 3D-Var upper-air assimilation in a limited area NWP-model

Annelies Duerinckx (1,2), Rafiq Hamdi (1,2), Alex Deckmyn (1,2), Arab Djebbar (4), Jean-François Mahfouf (3), Piet Termonia (1,2)

(1) Royal Meteorological Institute, Ringlaan 3, B-1180 Brussels, Belgium.
(2) Department of Physics and Astronomy, Ghent University, Ghent, Belgium.
(3) GAME, CNRM, Météo-France, CNRS, Toulouse, France.
(4) Office national de la météorologie. Dar el Beida, Algeria.

In recent years, the Extended Kalman Filter (EKF) has been gaining more attention in the surface Data Assimilation (DA) community and has already replaced the older Optimal Interpolation (OI) scheme for the vertical component of the land surface DA system in a number of meteorological institutes. An EKF has been developed within the stand-alone land surface modeling platform SURFace Externalisée (SURFEX) for the initialisation of soil temperature and soil water content based on screen-level temperature and relative humidity. In this paper we present a new combination of the EKF with a basic (using conventional observations only) three dimensional variational (3D-Var) upper-air assimilation for the limited area model ALARO coupled to SURFEX. This new combination is compared to an Open Loop experiment where all initial conditions are interpolated from an analysis of the global Numerical Weather Prediction model Action de Recherche Petite Echelle Grande Echelle (ARPEGE) and to an experiment where the surface is initialised using the EKF, while the upper-air initial conditions are interpolated from the ARPEGE analysis. The aim of this paper is to examine whether the EKF surface assimilation coupled or not with a basic 3D-Var upper-air assimilation has an added value compared to the Open Loop, in which the more advanced upper-air data assimilation of ARPEGE with more observations used is interpolated onto the limited area model grid. All set-ups are verified during a one-year period 2013

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/qj.3141

against soil measurements, screen-level observations, radiosoundings and merged radar-rain-gauge precipitation observations. Results indicate that the EKF surface assimilation has positive effects on humidity scores and is able to produce similar or improved scores compared to the open loop. While the upper-air 3D-Var DA system of ALARO still needs improvements, the potential benefits of the combination of upper-air and surface assimilation are demonstrated through soil moisture and screen-level relative humidity verifications.

Keywords: *Data Assimilation, Soil Analysis, Extended Kalman Filter, 3 dimensional variational assimilation, limited area model, numerical weather prediction*

1. Introduction

The improvement of land surface schemes and surface assimilation techniques during the last decades has led to considerable improvements in the short-range forecasts of lower atmospheric fields in Numerical Weather Prediction (NWP) (Giard and Bazile 2000). While Optimal Interpolation is the more commonly used technique for the horizontal component of the operational land surface data assimilation system (eg. Giard and Bazile 2000; Mahfouf et al. 2000; Bélair et al. 2003; Rodriguez et al. 2003; Drusch 2007), in recent years the EKF has been gaining more attention for the vertical component and has for example replaced the old OI scheme for soil moisture analysis in the Integrated Forecast System (IFS) of the European Centre for Medium-Range Weather Forecasts (ECMWF) (de Rosnay et al. 2012). It is also used operationally at the German Weather Service (Hess 2001). An EKF has been developed for the land surface scheme SURFEX (Masson et al. 2013) by Mahfouf et al. (2009). In SURFEX, both OI and the EKF may assimilate observations of screen-level temperature and relative humidity to correct errors in soil moisture content and soil temperature. Screen-level observations contain information on the state of the soil due to the

coupling between the surface and the atmosphere governed by the sensible and latent heat fluxes (Douville et al. 2000; Draper et al. 2011). The advantage of the EKF over OI is that it has a more generic formulation of the gain coefficients and thus can be extended towards new observation types (Mahfouf et al. 2009). Therefore the EKF offers many possibilities for further improvements and extensions. For example, it has already been extended to include Advanced Microwave Scanning Radiometer - Earth observing system sensor (AMSR-E) soil moisture retrievals (Draper et al. 2009), radar precipitation information (Mahfouf and Bliznak 2011) and Advanced Scatterometer (ASCAT) soil moisture (de Rosnay et al. 2012).

However, the land surface data assimilation approach with both EKF and OI has a number of challenges. The most important one is that the coupling between the surface and the atmosphere is not always that strong (Draper et al., 2009). This means that the screen-level departures, which form the basis for the soil corrections, are not always linked to errors in the soil. In these cases the soil is erroneously corrected to compensate for errors that are elsewhere in the model. This can lead to unrealistic values for the soil parameters as well as an increase in forecast errors in the next couple of days to even weeks (Hess et al. 2008). According to Hess et al. (2008), the most problematic situation occurs when the model predicts a high radiative impact with a strong coupling between the surface and the atmosphere, while in reality the conditions are cloudy with a weak coupling between the surface and the atmosphere. In this case the model will claim that the surface is responsible for the departures, while in reality the departures are due to a different cause, such as an erroneous cloud cover.

The relatively high impact of these erroneous corrections stems from the evapotranspiration - precipitation feedback (Christensen et al. 1997; Drusch 2007; Hess et al. 2008; Masson et al.

2013). These feedback processes can be local as well as non local (Betts and Viterbo 2005) and influence the weather on timescales from hours to months (Seuffert et al. 2002; Ferranti and Viterbo 2006; Beljaars et al. 1996). For example, once the surface is too dry, the evapotranspiration will be underestimated. This leads to less humidity in the lower atmosphere and a reduced possibility of convection and cloud formation. Less clouds will in turn cause too much radiative impact at the surface and will lead to further drying out of the surface and less convective precipitation. To recover from this feedback process either a heavy precipitation event is necessary, restoring the soil moisture to a sufficient level for evapotranspiration, or the surface assimilation should detect and correct the erroneous state of the soil. The latter is not always straightforward.

The problem stems from the fact that the upper air and land surface data assimilation systems are separated. To overcome this problem there are two possibilities: the first one is to use threshold values (Giard and Bazile, 2000) to control the coupling between the surface and the atmosphere used in the assimilation process. The second possibility is to improve the state of the atmosphere and in doing so, to try to avoid these cases of erroneous predictions of cloud cover and precipitation. However, this will not solve the scientific problem of the two separated assimilation systems. It will merely try to avoid the problem from occuring in the first place. This is where upper-air assimilation comes into the picture.

In global models the aim of upper-air assimilation is to avoid model drifting and make sure that the larger scale systems are well represented by the model state. Limited area models (LAMs) can then run on a finer resolution as a dynamical adaptation of the global run, using an interpolated analysis of the global model as initial conditions. For a LAM the lateral boundary conditions (LBCs), coming from the global model, take care of the larger scales

during the forecast. Instead of using an interpolated global analysis as initial conditions, one can also use a dedicated data assimilation cycle to create the initial conditions for the LAM. The purpose of the upper-air data assimilation for LAMs is not so much to improve the larger scales but to improve the humidity, divergence and vertical velocities of the LAM to get better cloud cover and precipitation forecasts (Fischer et al. 2005). With the model resolutions increasing steadily, the need for high-density observations like Global Navigation Satelite System (GNSS) and radar data is growing (Fischer et al. 2005; Fillion et al. 2010; Seity et al. 2011; Brousseau et al. 2014). Developing a dedicated upper-air assimilation cycle for LAMs can be seen as a first step towards the assimilation of high-density observations, which are expected to further improve the upper-air humidity, cloud cover and precipitation and in doing so are also beneficial for the surface assimilation.

Since both surface and upper-air assimilations have a positive impact on the forecast scores and potentially on each other, the combination of surface assimilation and upper-air assimilation has been the subject of many studies (Tudor et al. 2013; Stanesic 2011; Randriamampianina and Storto 2008; Schneider et al. 2009) and is used operationally in many meteorological institutes[1]. Stanesic (2011) describes the combination of a 3D-Var assimilation for the upper-air in combination with OI for the surface. Their verification shows a positive impact for the upper-air fields as well as the near surface variables, especially for screen-level temperature and relative humidity and upper-air humidity. Randriamampianina and Storto (2008) report improvements on the short forecast range, while in Schneider et al. (2009) the results are more mixed.

De Rosnay et al. (2012) demonstrate how the EKF surface assimilation is used operationally

[1] eg. at Météo France, ECMWF, Met. Hungary, ZAMG Austria, Met. Slovenia, Met. Norway, Met. Sweden

at the ECMWF in combination with a four dimensional variational (4D-Var) upper-air assimilation, replacing the previous OI soil moisture analysis in the global IFS . For soil temperature the OI scheme is still used. Since the ECMWF does not yet have an externalised version of their land surface model HTESSEL (Hydrology Tiles ECMWF Scheme for Surface Exchange over Land; Van den Hurk et al. 2000; Balsamo et al. 2009), computationally intensive, fully coupled forecasts are used to calculate the Jacobian of the EKF in finite differences. For computational efficiency reasons, the surface analysis at the ECMWF is run in parallel with the upper-air 4D-Var analysis.

In this paper we will present a combination of the EKF with a basic 3D-Var upper-air assimilation for the LAM ALARO[2] coupled to SURFEX and study the impact of the different parts of this assimilation system. We call the 3D-Var upper-air assimilation a basic assimilation because we will only use conventional observations and not satellite observations. A first aim is to examine whether combining the EKF surface assimilation with a basic 3D-Var upper-air assimilation has an added value compared to the Open Loop, in which the more advanced 4D-Var upper-air data assimilation of the global NWP-model ARPEGE (Bubnová et al. 1995) is interpolated. All set-ups will be tested for a one-year period in order to compare the results for all seasons.

2. Methodology

2.1 The atmospheric model ALARO

The ALARO model is a LAM based on the Aire Limité Adaptation Dynamique Dévelopment International (ALADIN) model (Bubnová et al. 1995), that is further extended with an additional physics parametrization package designed to run at convection-permitting

[2] [2]* ALARO stands for Aladin-AROme.
* The ALADIN acronym stands for Aire Limitée Adaptation Dynamique Développement International.
* AROME stands for Application of Research to Operations at Mesoscale

resolution (the so-called grey-zone scales). This additional physics parametrization includes the 3MT (modular multi-scale micro-physics and transport) scheme of Gerard and Geleyn (2005) and Gerard (2007) that aims to improve the convection and cloud parametrization. The ALARO model has been validated up to a spatial resolution of 4km for both NWP (Gerard et al. 2009; De Meutter et al. 2015) and climate simulations (Hamdi et al. 2012; Hamdi et al. 2014; De Troch et al. 2013; Giot et al. 2016).

ALARO shares its code with the ALADIN model, the LAM version of the ARPEGE-IFS system (Bubnová et al. 1995). The physics-dynamics interface of the ALARO model uses a flux-conservative formulation described in Catry et al. (2007). The ALARO model is running operationally at the Royal Meteorological Institute (RMI) of Belgium as well as in a number of other countries of the ALADIN and HIRLAM consortia.

2.2 The land surface model SURFEX

In this study ALARO is coupled to the external land surface modelling platform SURFEX (Masson et al. 2013). SURFEX has been externalised from the mesoscale model meso-NH (Lafore et al. 1998) and is coupled to ALARO using the approach of Polcher et al. (1998) and Best et al. (2004). SURFEX can be run both in offline and coupled mode. In offline mode, the land surface model is driven by the atmospheric forcing and no feedback is possible from the surface to the atmosphere. In coupled mode on the other hand, SURFEX and the atmospheric model exchange fluxes and forcing data at every timestep. The offline mode is useful when no feedback to the atmosphere is required, for example for studying the sensitivity to land cover/land use change as in Hamdi et al. (2009), Hamdi et al. (2011) or as a computationally cheaper way for calculating the Jacobian of the Extended Kalman Filter developped for SURFEX (Mahfouf et al. 2009; Duerinckx et al. 2015). SURFEX has a modular structure so

that it can easily be extended with new parametrizations. Each gridbox consists of a number of different tiles (sea, lakes, nature and town). The fluxes calculated by SURFEX are averaged per gridbox according to the fraction of each of the tiles. Nature tiles can be build up from 12 patches of different vegetation types. The nature tiles use the ISBA scheme (Interaction between the Soil, Biosphere and Atmosphere), developed by Noilhan and Planton (1989), Noilhan and Mahfouf (1996). For sea and ocean tiles two options exist: either one can use the simple formulation with constant Sea Surface Temperature (SST) with Charnock's approach for the roughness length or one can use the one-dimensional ocean mixing layer model of Lebeaupin (2007). For the lake tiles the FLAKE model (Mironov et al. 2010) and for town tiles the TEB scheme (Town Energy Balance, Masson 2000) are available.

For the set-up of this paper the two-layer version of ISBA with one vegetation patch is used for the nature tiles. The four prognostic variables are surface and deep soil temperature (T_s and T_2) and surface and deep soil water content (W_g and W_2). Surface assimilation is only performed on the nature tiles. In offline mode, for the atmospheric forcing the variables from the first atmospheric model layer at approximately 17m height as in Mahfouf et al. (2009) are applied.

2.3 The land data assimilation system

For the surface we use the EKF to assimilate the screen-level observations of temperature (T_{2m}) and relative humidity (RH_{2m}) to correct soil moisture (W_g and W_2) and soil temperature (T_s and T_2). The EKF as formulated by Mahfouf et al. (2009) is used here. The equation for the model state analysis is:

$$\vec{x}_{\mathrm{a}}^{\,t} = \vec{x}_{\mathrm{b}}^{\,t} + \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}[\vec{y}^{\,t} - \mathsf{H}(\vec{x}_{\mathrm{b}}^{\,t_0})] \qquad (1)$$

where subscripts a,b represent the analysis and the background, respectively. The analysis model state $\vec{x}_a$ is defined as the sum of the background model state $\vec{x}_b$ and an increment based on the observation departure $[\vec{y}^t - \mathsf{H}(\vec{x}_b^{t_0})]$ and the Kalman gain matrix $\mathbf{BH}^T(\mathbf{HBH}^T + \mathbf{R})^{-1}$, $\mathbf{B}$ the background error covariance matrix, $\mathbf{R}$ the observation error covariance matrix and $\vec{y}$ the observation vector. The observation operator H projects the model state onto the observation space. In the formulation of Mahfouf et al. (2009), H contains a forward model propagation from time $t_0 = t - \Delta t$ to time t (the observation time), where Δt is the assimilation window, and the conversion of the model state into an observation equivalent. Although theoretically the increment should thus be added at the beginning of the assimilation interval, in our case the increment is added at the end of the assimilation interval, i.e. at the observation time, using the assumption that the increment is not significantly changed by the forward model propagation. $\mathbf{H}$ is the Jacobian matrix of the observation operator. It is calculated with a finite differences approach in the following way:

$$\mathbf{H^t} = \frac{\delta \mathsf{H}^t(x)}{\delta x^{t_0}} = \frac{\mathsf{H}_i^t(x^{t_0} + \delta x_j) - \mathsf{H}_i^t(x^{t_0})}{\delta x_j} \qquad (2)$$

For the calculation of $\mathsf{H}_i^t(x^{t_0} + \delta x_j)$, a run with a perturbed surface field δx_j is performed for each of the control variables x_j. This can be done with SURFEX in offline mode (the surface scheme is decoupled from the atmospheric model) or in coupled mode (the surface scheme is coupled to an atmospheric model) as described in Duerinckx et al. (2015). To filter out artificial oscillatory trajectories in the screen-level variables triggered in the situation of stable atmosphere, this study uses the temporal filter proposed in Duerinckx et al. (2015) for the calculation of the Jacobians:

$$x_{\text{t, filtered}} = 0.5 \times w \times x_{t-1} + (1-w)x_t + 0.5 \times w \times x_{t+1} \qquad (3)$$

with x corresponding to the T_{2m} or RH_{2m} value to be filtered, t representing the timestep,

$w = 0.5$ the weight attributed to the different parts of the filter and $x_{t,\,\text{filtered}}$ the predicted observation at time t for which the oscillation has been removed.

This study uses a simplified version of the EKF in which the background covariance matrix $\mathbf{B}$ does not evolve with time. This is in accordance with the work of Mahfouf et al. (2009) and Draper et al. (2009).

The EKF implemented in SURFEX performs a pointwise analysis, requiring the observations to be mapped onto the model grid beforehand. In this study, the observations were mapped onto the model grid using a horizontal optimal interpolation scheme implemented as , "Code d'Analyse Nécessaire à ARPEGE pour ses Rejets et son Initialisation" (CANARI; Taillefer 2002). This is in accordance with what has been done in Mahfouf et al. (2009). To limit the influence of the background as much as possible during this interpolation step, it is performed with observation error variances (**R**) that are ten times smaller compared to the values used for the EKF assimilation.

2.4 The upper-air assimilation system

For the upper-air assimilation system the incremental 3 dimensional variational assimilation method is used that has been developed for the ALADIN model. The ALADIN-3D-Var system is closely related to that of its global counterpart, the ARPEGE/IFS system. It uses the same incremental formulation, observation operators, minimization technique and data flow as the 3D-Var global assimilation system in ARPEGE/IFS which is described in Courtier et al. (1998). A more detailed description of the ALADIN-3D-Var system can be found in Fischer et al. (2005). The control variables for the ALADIN 3D-Var system are vorticity, divergence, temperature, surface pressure and specific humidity. Scale-dependent statistical

regressions separate the balanced and unbalanced components of the control vector (Parrish et al. 1997; Derber and Bouttier 1999). The properties of these statistical balances are described in Berre (2000), Sadiki et al. (2000), Montmerle et al. (2006). The ALADIN 3D-Var code uses the M1QN3 quasi-Newton minimization method described in Gilbert and Lemarechal (1989). For the formulation of the $\mathbf{B}$-matrix, the ensemble method is used (Berre et al. 2006; Pereira et al. 2006). This formulation was obtained as the bi-Fourier counterpart of the global, spherical harmonic formulation described in Berre (2000). The statistics are computed in spectral space assuming fields of spectral coefficients to be homogeneous and isotropic. The horizontal correlations are vertically varying and the cross-covariances are calculated using the multiple regression scheme of Berre (2000).

3 Experimental setup

Analogously to what has been done in de Rosnay et al. (2012), the EKF surface analysis is combined with an independent upper-air assimilation run. In contrast to their work, the LAM ALARO is used instead of a global model and the upper-air analysis is done with a basic 3D-Var (Fischer et al. 2005) instead of the computationally much more expensive 4D-Var. The ALARO model is coupled to the two-layer version of ISBA within SURFEX and the calculation of the EKF Jacobians of the observation operator is done using the offline, filtered approach of Duerinckx et al. (2015).

The values of the EKF parameters are the same as in Mahfouf et al. (2009). The $\mathbf{R}$-matrix is a diagonal matrix with elements set to 1K for 2m temperature and 10% for 2m relative humidity. The $\mathbf{B}$-matrix is also a diagonal matrix, with 2K for the background errors of the superficial and deep soil temperature and $0.1\times(w_{fc}-w_{wilt})$ for the superficial and deep soil moisture content, where w_{fc} is the field capacity and w_{wilt} the wilting point. A static $\mathbf{B}$-

matrix was used for the EKF, like in the set-up of de Rosnay et al. (2012), Mahfouf et al. (2009) and Duerinckx et al. (2015). Following the previous study of Duerinckx et al. (2015), in which various perturbation sizes were compared, the size of the relative perturbations were set to 10^{-7} for Wg and W2 and to 10^{-7} for Ts and T2 (see Eq. 4).

$$X_{prt} = X_i + prt * X_i \quad (4)$$

With X_{prt} the perturbed value, X_i the original value and *prt* the relative perturbation size.

The 3D-Var+EKF run is the new combination under evaluation here. It is compared to runs without local assimilation (Open Loop) and runs with only local upper-air or local surface assimilation (Table 1). When no local assimilation is applied, the initial conditions are interpolated from an ARPEGE analysis. The horizontal resolution of the ARPEGE analysis ranges from 10 km over Europe to 65 km over the Antipodes. The incremental 4D-Var data assimilation system (Desroziers et al., 1999) consists of two outer loops, the first one being at a coarser resolution (T107) than the second one (T323). The control variables are vorticity and unbalanced variables are divergence, temperature, surface pressure and humidity. The background error variances are derived from a data assimilation ensemble and are updated at every cycle (Berre and Desroziers, 2010). Data assimilations are performed every 6h, followed by 6h forecasts. Medium-range forecasts are performed every day at 0000 UTC up to 102h ahead. Conventional observations and much more satellite observations are assimilated in the operational suite (more details can be found in Chambon et al, 2014). The surface assimilation is based on OI (Giard and Bazile, 2000) at 6h windows interval using 2m temperature and relative humidity observations.

The Open Loop (OL) run uses the interpolated ARPEGE analyses, both for the surface as well as the atmosphere. The 3D-Var+OL run uses a local ALARO 3D-Var analysis for the

atmosphere combined with an 'Open Loop' (i.e. interpolated from an ARPEGE analysis) setup for the surface. The EKF run is a simulation with a local SURFEX EKF surface assimilation but an interpolated ARPEGE analysis for the atmosphere. Finally, the 3D-Var+EKF run uses only local assimilation, i.e. the ALARO 3D-Var set-up for the atmosphere and the SURFEX EKF for the surface.

All experiments were run during one year, 2013, with a 6-hour assimilation cycle and use the operational ALADIN-Belgium domain which has a 4km horizontal resolution (181 × 181 grid points) and 46 vertical levels. For the EKF surface assimilation screen-level relative humidity and temperature observations were used from SYNOP and TEMP reports in the Observation Pre-processing database for LACE (OPLACE database), Bölöni et al. (2009). Figure 1 shows the spatial distribution of the observations for the 5th of May 2013 at 12 UTC.

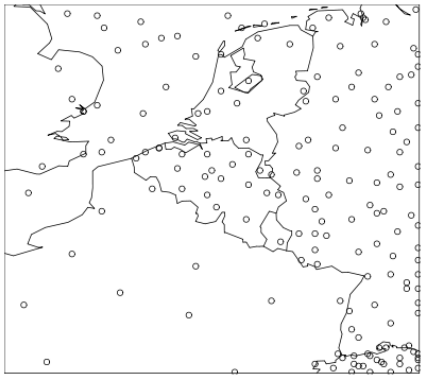

Figure 1 Locations of the screen-level observations in OPLACE used in the surface assimilation for the 5th of May 2013.

The observations used for the upper-air assimilation also come from the OPLACE database. The basic 3D-Var analysis only uses conventional observations (wind profiler, radiosonde, aircraft data and 10m wind surface synoptic observations). Since the focus is mainly on the soil analysis and the added value of combining it with an upper-air analysis, no satellite data or high-density observations such as GNSS or radar data were used. The use of such observations takes a lot of preparation and monitoring and is out of the scope of this paper. Therefore, this set-up can be seen as a preparatory step for the assimilation of high-density observations in the future.

The $\mathbf{B}$-matrix of the 3D-Var assimilation system was created using the analysis-ensemble method (Berre et al. 2006; Pereira et al. 2006; Fisher 2003; Pereira and Berre 2006) from an ensemble of ALARO 6h forecasts. The initial and lateral boundary conditions of the ensemble

are provided by a 6-member ensemble of global perturbed ARPEGE assimilation cycles. In the ARPEGE ensemble, the observations are perturbed by adding random realizations drawn from the specified observation-error covariance matrix (Fischer et al. 2005). The perturbed ARPEGE/ALARO experiments were carried out over a period of one year (2011). A forecast for each of the 6 ensemble members was launched for every 18th day (i.e. on 01/01/2011, 18/01/2011, 05/02/2011, 23/02/2011, etc.) at 00,06,12 and 18 UTC. This results in a total of 20 forecast days spread out over the year with 4 forecast-cycles per day for each of the 6 members, providing in total 480 differences. The intermittency period of 18 days between each day used for the B-matrix makes sure the B-matrix is constructed from ensemble difference for different seasons and weather situations. The ensemble method was used because it provides error correlation spectra whose shapes are more realistic than those of the NMC and Monte Carlo methods (Fischer et al. 2005).

4 Results and discussion

In this section, the different runs will be compared to each other focussing on the soil moisture and soil temperature increments, the verification against soil, screen-level and sounding observations, and the verification of SAL precipitation scores (Wernli et al. 2008).

4.1 Increments

Figure 2 gives an overview of the increments of the four prognostic variables for the EKF run. For each of the four prognostic soil variables, the increments are cumulated per season and then rescaled to K/day for T_s and T_2 and mm/month for W_g and W_2. Figure 3 shows the increment distributions per season for the four prognostic variables for the EKF

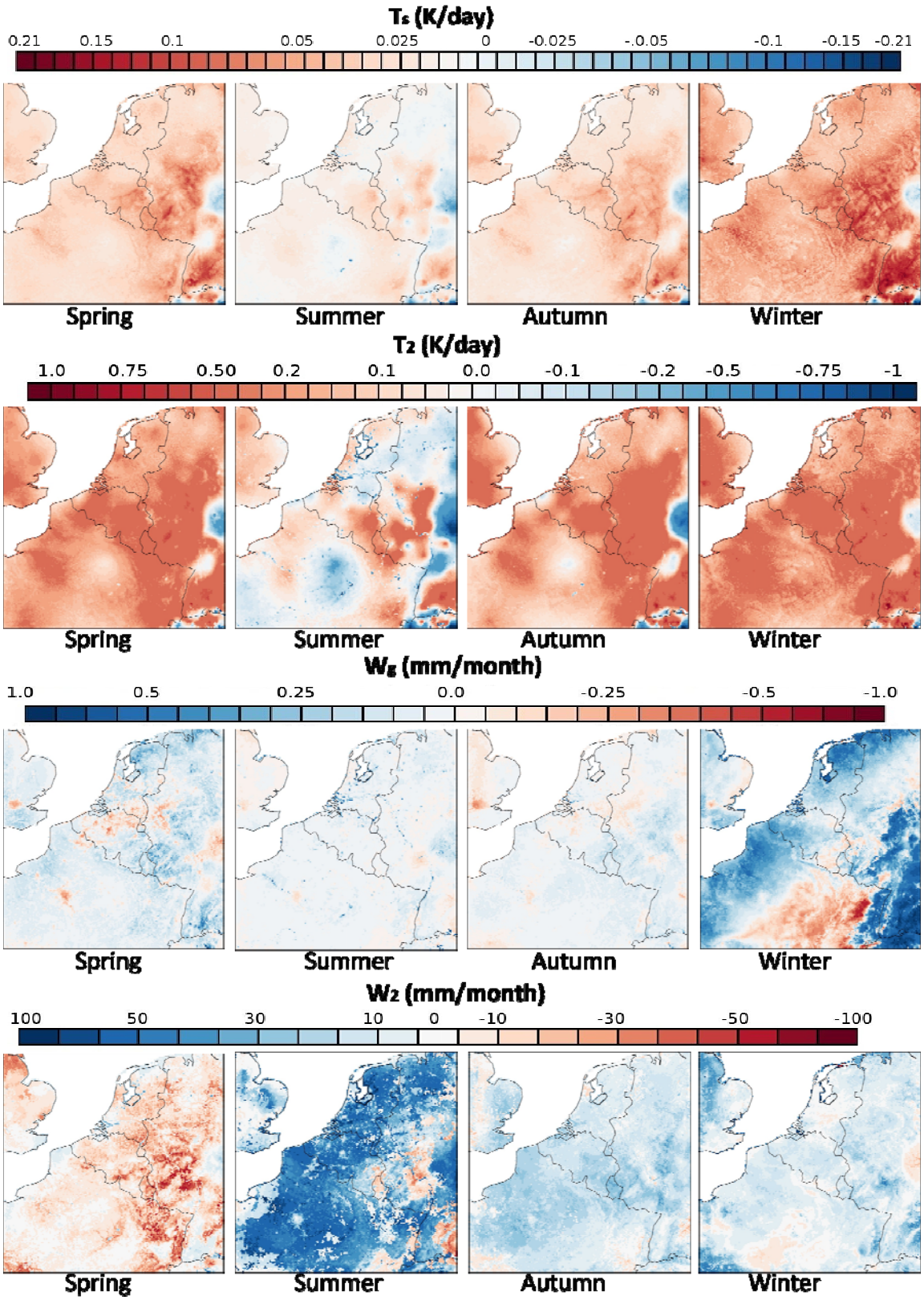

Figure 2 Increments for surface temperature (first row), deep soil temperature (second row), superficial soil moisture content (third row) and deep soil moisture content (fourth row) per season for the EKF-run

.

(white) and 3D-Var+EKF (black) with the overlap in distributions between the two indicated in grey. Figure 2 shows that the average increments for superficial soil temperature (T_s) are positive for all seasons, except for summer (June, July, August). This also shows in the distributions of figure 3, where all T_s distributions are slightly right-skewed except the one during summer. The T_s increments are strongly linked to the

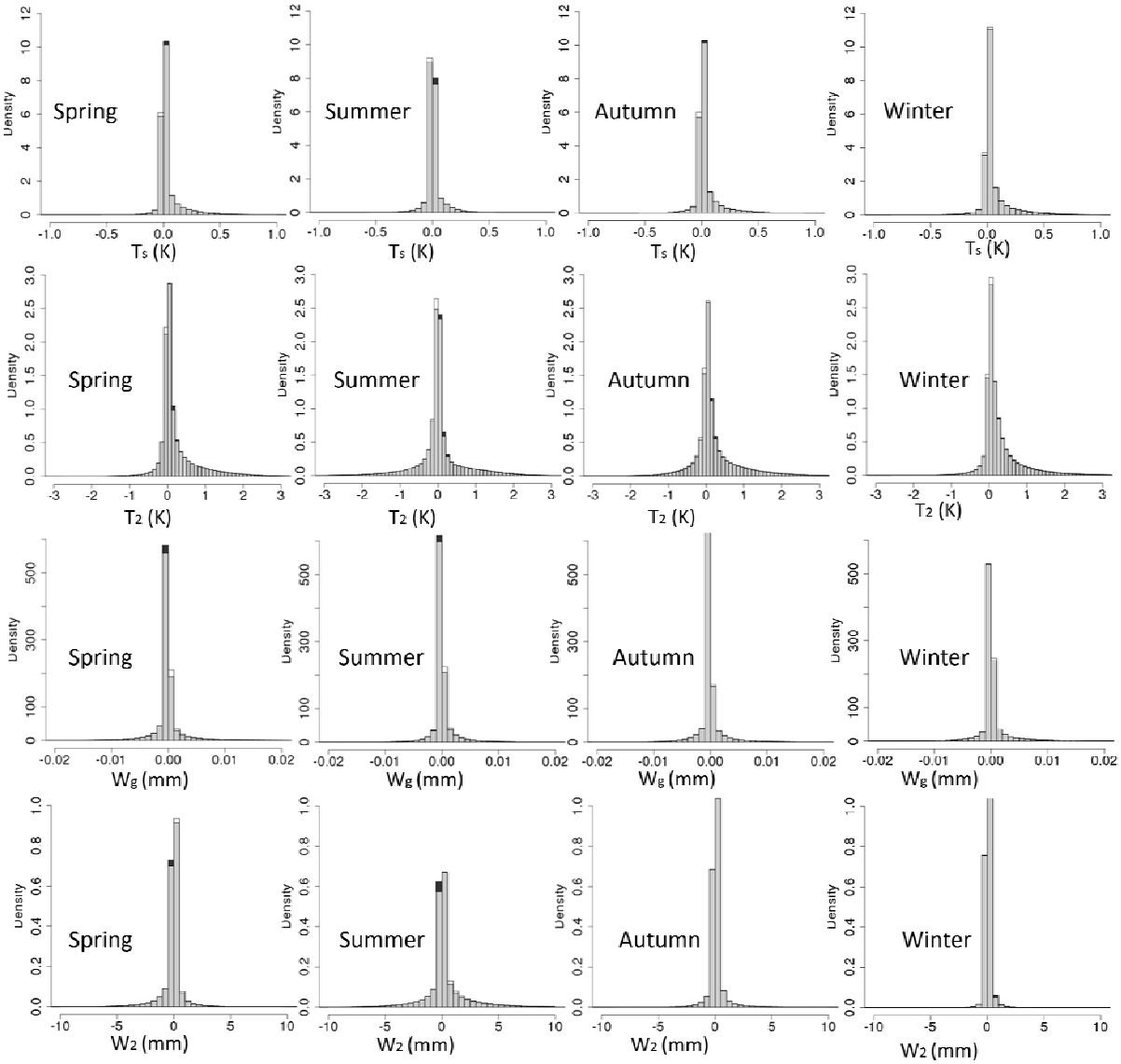

Figure 3 Soil increment histograms for surface temperature (first row), deep soil temperature (second row), superficial soil moisture content (third row) and deep soil moisture content (fourth row) per season. EKF increment distributions are indicated in white, 3D-Var+EKF increment distributions are indicated in black and the grey areas indicate the overlap between EKF and 3D-Var+EKF increment distributions. The y-axis represents the density.

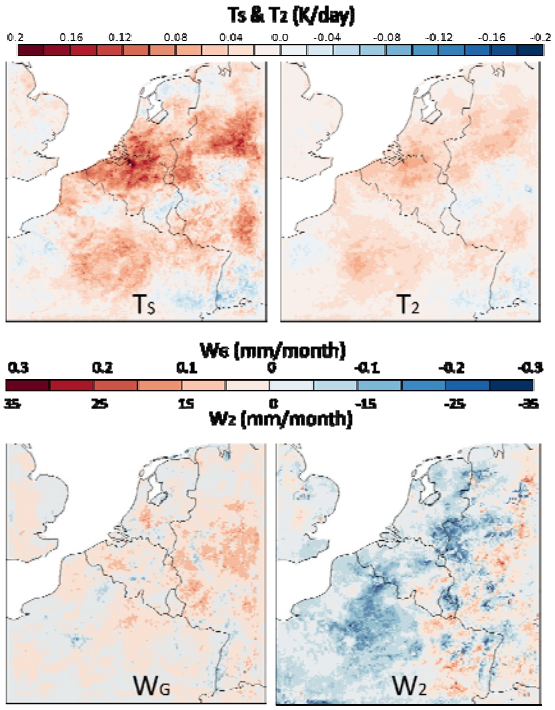

Figure 4 Difference in increments between 3D-Var+EKF and EKF for surface temperature (first row left), deep soil temperature (first row right), superficial soil moisture content (second row left) and deep soil moisture content (second row right) averaged over spring

screen-level temperature (T_{2m}) departures. This indicates that in general the model is too cold compared to the observations, except during summer. While the assimilation runs are able to correct a part of this trend, they are not able to remove the errors completely. It is likely that the soil is not the only factor causing the cold trend. The bars in figure 3 are mainly colored grey, indicating that the 3D-Var+EKF and EKF run have largely a very similar increment distribution. The small black (right) and white (left) tops in the middle of the histograms indicate that the 3D-Var+EKF run has some more small positive increments compared to the EKF run. This is confirmed by figure 4, that shows the differences in increments between 3D-Var+EKF and EKF (i.e. 3D-Var+EKF − EKF) during spring (March, April, May). For T_s the difference is mainly positive, especially in the northern half of the domain. This means that 3D-Var+EKF has higher positive increments than the EKF run, which indicates larger screen-level departures for the 3D-Var+EKF run.

The EKF T_S increments are approximately five times smaller than the T_2 increments (cfr. Figure 2). This is caused by the Jacobians that are very small for T_S (Duerinckx et al. 2015). The Jacobians measure the sensitivity of the screen-level variables to changes in the soil. They are calculated over a 6-hour time interval, measuring the effect of a change in the soil to

the screen-level variables 6 hours later. In the case of T_S this sensitivity is very low due to the very short memory of the superficial soil layer. The larger T_2 increments originate in the Jacobian values that in this case reflect the long memory for T_2. Just like for T_S, the T_2 average increments are positive except during summer. This is also reflected in the slightly right-skewed distributions during spring and winter shown in figure 3. With regards to the difference between the EKF run and the 3D-Var+EKF run, we can see that the increments of 3D-Var+EKF are somewhat higher and more positive than those of EKF (indicated by the black tops of the bars in the right part of the histograms), especially during spring and summer. The T_2 differences during spring between 3D-Var+EKF and EKF have the same spatial distributions as for T_S (cfr. Figure 4) but are smaller for T_2.

For W_g there is a slight moistening trend throughout the year, indicated by the mainly negative average increments in all four seasons shown in figure 2. This trend is the same for W_2 except during spring, where there is a drying trend. This indicates that in spring the model is too moist while especially in summer it is too dry. In autumn (September, October, November) and winter (December, January, February) the W_2 increments are smaller than in spring and summer, which is also clearly shown by the higher peak and smaller tails of the W_2 distributions for autumn and winter shown in figure 3. The screen-level scores discussed in section 4.3 indicate that the model is too moist during spring and too dry during summer. In autumn and winter the dry trend is much smaller than in summer. The W_2 increments are especially successful in removing the errors in spring and to a lesser extend in autumn and winter. Even though in summer the soil moisture increments are the largest, they are not able to correct the dry trend that can be seen at night in the screen-level scores. Once again, it is likely that the soil is not the only factor causing the dry trend in this case. One possible

explanation could be the errors in the precipitation forecasts that are highest during summer, as will be shown in section 4.5. Another possibility is an underestimated cloud cover causing too much evapotranspiration and drying out the soil too much. For W_g and W_2 the bars with black tops (i.e. there are more 3D-Var+EKF increments in this interval than EKF increments) in figure 3 are in the middle to the left of the white tops (i.e. there are more EKF increments in this interval than 3D-Var+EKF increments). Hence the increments of 3D-Var+EKF are slightly more negative (i.e. more drying) than those of the EKF in this case. This is confirmed by figure 4 which shows the spatial distribution of the increment differences during spring. For W_g the increment differences are very small and in part of the domain slightly positive. For W_2 the increment differences are larger and mainly negative. This means that especially for W_2 the 3D-Var+EKF has stronger negative increments than the EKF, which is probably caused by the higher screen-level departures. Note however that the increments of W_2 are also larger than those of W_g, resulting in larger increment differences for W_2 than for W_g.

4.2 Soil Verification

To verify the soil moisture and soil temperature analysis, measurements of soil water content and soil temperature from the Fluxnet database (Fluxnet 2015) were used at 3 locations for superficial soil moisture content (W_g, measurement depth is indicated in table 2) and 6 locations for superficial soil temperature (T_s, measurement depth is indicated in table 2) . The locations are listed in table 2 and figure 5. Table 2 also shows the vegetation type as it is mentioned on the Fluxnet website (Fluxnet2015) and the saturation and residual soil moisture content as they are used in SURFEX ('model') and in the observations. The saturation and residual soil moisture content of SURFEX are based on the percentages of clay and sand,

which stem from the FAO database (FAO 2006) used in ECOCLIMAPII (Faroux et al. 2013). No measurements for W_2 and T_2 were included, since they were not available for the year 2013 in any of the observation locations.

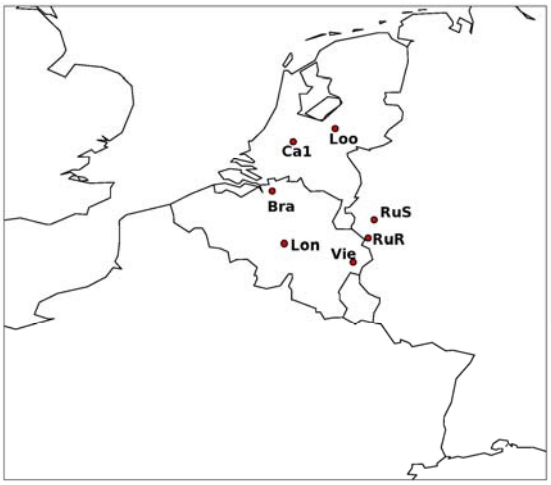

Figure 5 Locations of stations for the soil observations (cfr. Table 2).

Albergel et al. (2012) describe the soil moisture validation strategy of the ECMWF. They state that while in situ measurements of soil moisture are an important source of information for validations, these measurements also contain considerable errors and their uncertainty is often unknown. Therefore they propose temporal correlation, bias and root mean square difference (RMSD) as main validation metrics. They prefer RMSD over root mean square error (RMSE) to emphasise that in situ data may contain instrumental and representativeness errors and are by no means considered to be the 'true' soil state. Conceptually, soil moisture can vary between two extremes: the residual soil moisture (W_{res} i.e. the water that is not extractable by plants or drainage because of the molecular bounding to soil particles) and the saturation point (W_{sat}), over which water begins to flow over the surface without even entering the soil. These values often do not agree between the model parameters and the observed values. For this reason the soil moisture measurements and model values are both linearly rescaled to the $[0,1]$ interval to match the same range before calculating the bias and

RMSD.

Tables 3 and 4 show the correlation, bias and RMSD for T_s and W_g of the initial conditions averaged per season for the runs OL, EKF and 3DVar+EKF. The 3D-Var+OL run was left out of the analysis because it uses the ARPEGE interpolated OI-analysis as initial conditions for the surface, which are exactly the same ones as the ones used in the OL run. Initial conditions for the soil variables of the 3D-Var+OL run are therefore identical to those of the OL run. The correlation calculated here is a temporal correlation calculated per season between the observation timeseries and the timeseries of the nearest gridpoint to the observation location. For both variables correlations are lowest during the summer in the observation locations. A possible explanation for this is the fact that the model is generally too dry during summer (cfr. Section 4.1 on the increments) and there are not so much precipitation events. Even though in summer the soil moisture increments add water to the reservoir, they are not able to correct the dry trend that can be seen at night in the screen-level observations (cfr. Section 4.3). It is likely that the soil is not the only factor causing the dry trend. Still, the surface assimilation tries to correct this error through the soil and this causes the soil state to diverge from the observations.

In summer, the W_g correlation of 3D-Var+EKF is higher than that of OL and EKF. It seems that the 3D-Var+EKF run is better able to handle the too dry model state during the summer. During spring, autumn and winter, the W_g correlation of the OL run is slightly higher than that of the other two runs. The highest W_g correlations are found during autumn in the observation locations where the scores for all the runs lie very close together. A possible explanation for this is that W_g is closely linked to precipitation events and the precipitation is best represented

in the model during autumn (cfr. SAL precipitation verification in section 4.5). The W_g bias in the observation locations is negative in winter and spring, slightly positive in summer and close to zero in autumn. In the observation locations the model is thus too dry in winter and spring, while being well represented in summer and autumn. This corresponds to the W_g increments, discussed in section 4.1 that were in general positive during spring and winter, while being close to zero in summer and autumn. The values for the W_g bias lie very close together in all three runs, except during spring when the differences are slightly larger. During spring the EKF and 3D-Var+EKF run have a smaller bias than the OL. The W_g RMSD scores are lowest during autumn due to the better precipitation representation in the model and highest during winter when the bias scores are also slightly higher than for the other seasons. In autumn the RMSD scores of the three runs are almost equal; in winter the RMSD scores are lower for OL than for the other two runs and in spring the EKF and 3D-Var+EKF runs outperform the OL in RMSD scores. For summer the 3D-Var+EKF has the lowest RMSD score. It seems that the 3D-Var+EKF run is better able to handle the too dry model state of the summer.

For T_s OL has a slightly higher correlation in summer, autumn and winter compared to EKF and 3D-Var+EKF. In spring the correlations of all runs are the same for T_s. The bias for T_s is always negative in the observation locations except during summer, where it turns positive. This confirms that the model values for T_s in the observation locations are in general too cold except during summer when the model values are too warm. The yearly temperature cycle of the model thus seems to be larger than in the observations. In autumn and winter OL has the lowest T_s bias scores. In spring the OL already has a positive T_s bias while EKF and 3D-Var+EKF still have a negative bias which is also closer to zero in absolute values than that of

OL. In summer the EKF has the lowest T_s bias. Except for the 3D-Var+EKF run in summer, the EKF and 3-DVar+EKF run always have a colder T_s bias than OL. This corresponds to what can be seen in the bias of the screen-level temperature (cfr. section 4.3). The T_s RMSD values are highest during summer, due to the mainly positive errors in summer. The T_s RMSD values are lowest for OL in all four seasons. This corresponds to the RMSE scores of the screen-level temperature which are also lowest for OL in all seasons (cfr. Section 4.3).

4.3 Screen-level verification

To verify the effect of the assimilation on the forecast scores, the RMSE and bias are calculated for screen-level temperature (T_{2m}) and relative humidity (RH_{2m}) for 11 synoptic stations spread over Belgium (cfr. table 5 and figure 6). The observations used in the verification are independent from the ones used in the assimilation process.

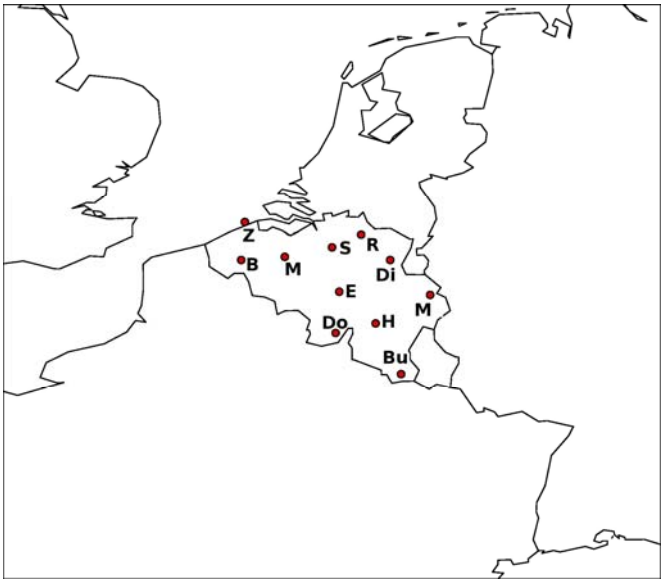

Figure 6 Location of stations for the screen-level observations used for verification (cfr. Table 5).

Figure 7 and table 6 show the RMSE and bias for RH_{2m} for each season for the OL, 3dVar+OL, EKF and 3dVar+EKF runs averaged over the 11 observation stations in Belgium. Figure 8 and table 7 show the average RMSE and bias for T_{2m} for each season for the same runs. The Open Loop (black solid line) being the reference run .

In **winter** the RMSE scores for RH_{2m} all lie very close together. Since there is only a limited amount of incoming solar radiation at the surface, the coupling between the surface and the atmosphere is rather low and the surface assimilation does not have a large impact. Still, the EKF and even more the 3D-Var+EKF manage to give an improvement in the RH_{2m} scores for the observation locations during the first twelve hours of the forecast. The 3D-Var+EKF run almost completely removes the RH_{2m} bias in the first six to twelve hours. Compared to spring and summer, the RH_{2m} bias in autumn and winter is much smaller for all four runs.

During spring, summer and autumn the EKF and 3D-Var+EKF runs seem to perform equally well or even better than the OL run during the first 6 to 12 hours of forecast for RH_{2m}. During **spring** the EKF run is able to remove the negative RH_{2m} bias completely for the first 12 hours in the observation locations, although the OL run is still better for the remaining 12-48 hours of the forecast. The 3D-Var+EKF run has a positive RH_{2m} bias for the whole forecast range during spring, while the OL has a negative bias during the first 12 hours. The EKF and 3D-VAR+OL run have a bias close to zero during the first 12 hours in spring and have a positive bias from 12h onwards.

During **summer** the EKF run is able to obtain similar RH_{2m} scores as the OL run in these 11 locations. The RMSE scores for 3D-VAR+EKF are slightly higher than those of OL and EKF from forecast range 6h up to 48h. There is a dry RH_{2m} bias for all the runs in Summer.

During **autumn**, the RH_{2m} scores already lie closer together but now the EKF and 3D-Var+EKF runs are able to improve the RMSE and BIAS of RH_{2m} for the whole forecast range up to 48 hours compared to OL. The RH_{2m} bias in Autumn is close to zero for the whole forecast range. The RMSE of 3D-Var+OL is slightly higher than that of OL during the first six hours of the forecast, it is equally low or even slightly lower during the remaining forecast ranges.

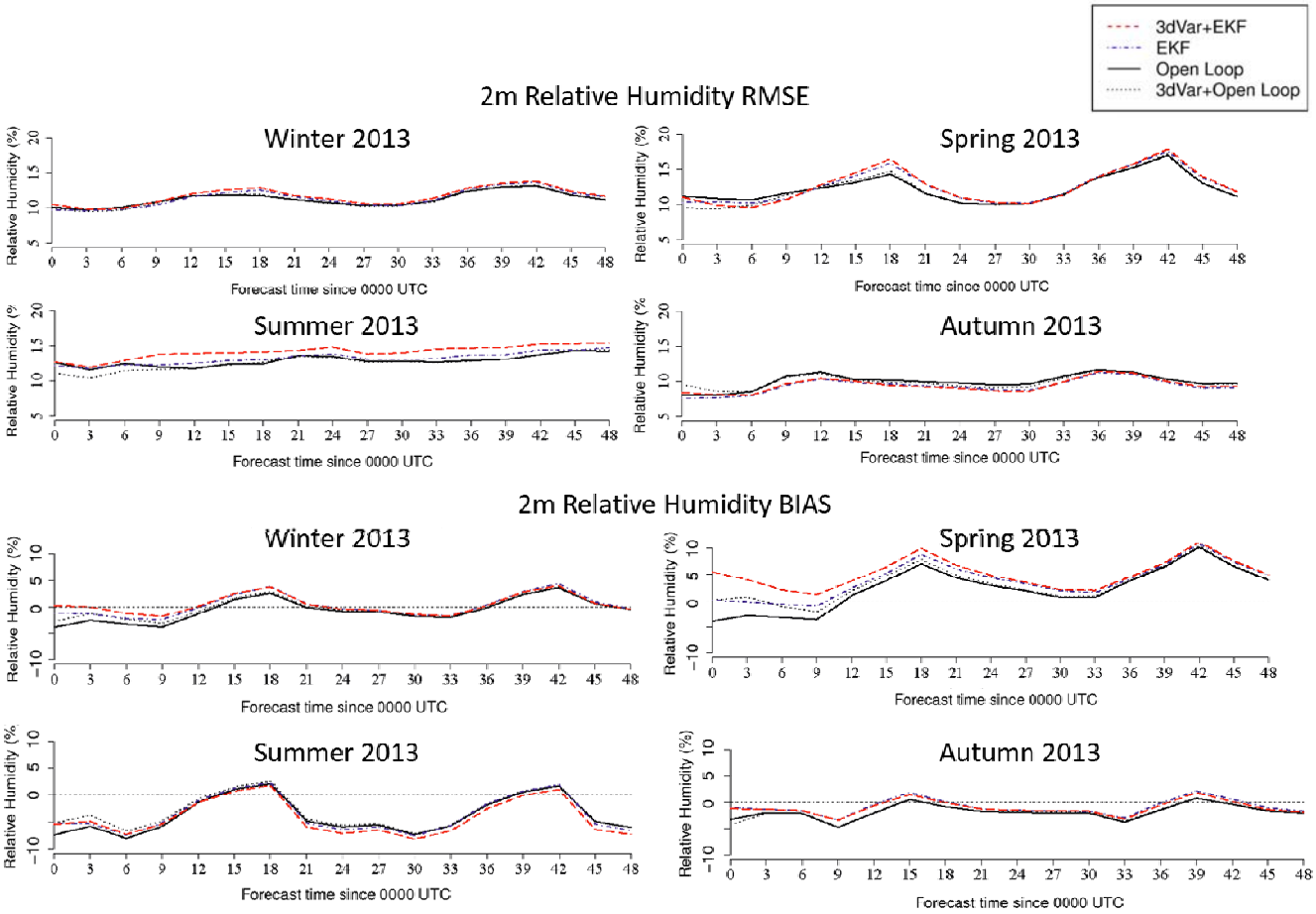

Figure 7 Screen-level relative humidity forecast scores in terms of RMSE and BIAS with a leading time up to 48 hours for OL, EKF, 3D-VAR+OL and 3D-Var+EKF averaged over 11 stations in Belgium (displayed in figure 6) calculated for spring, summer, autumn and winter.

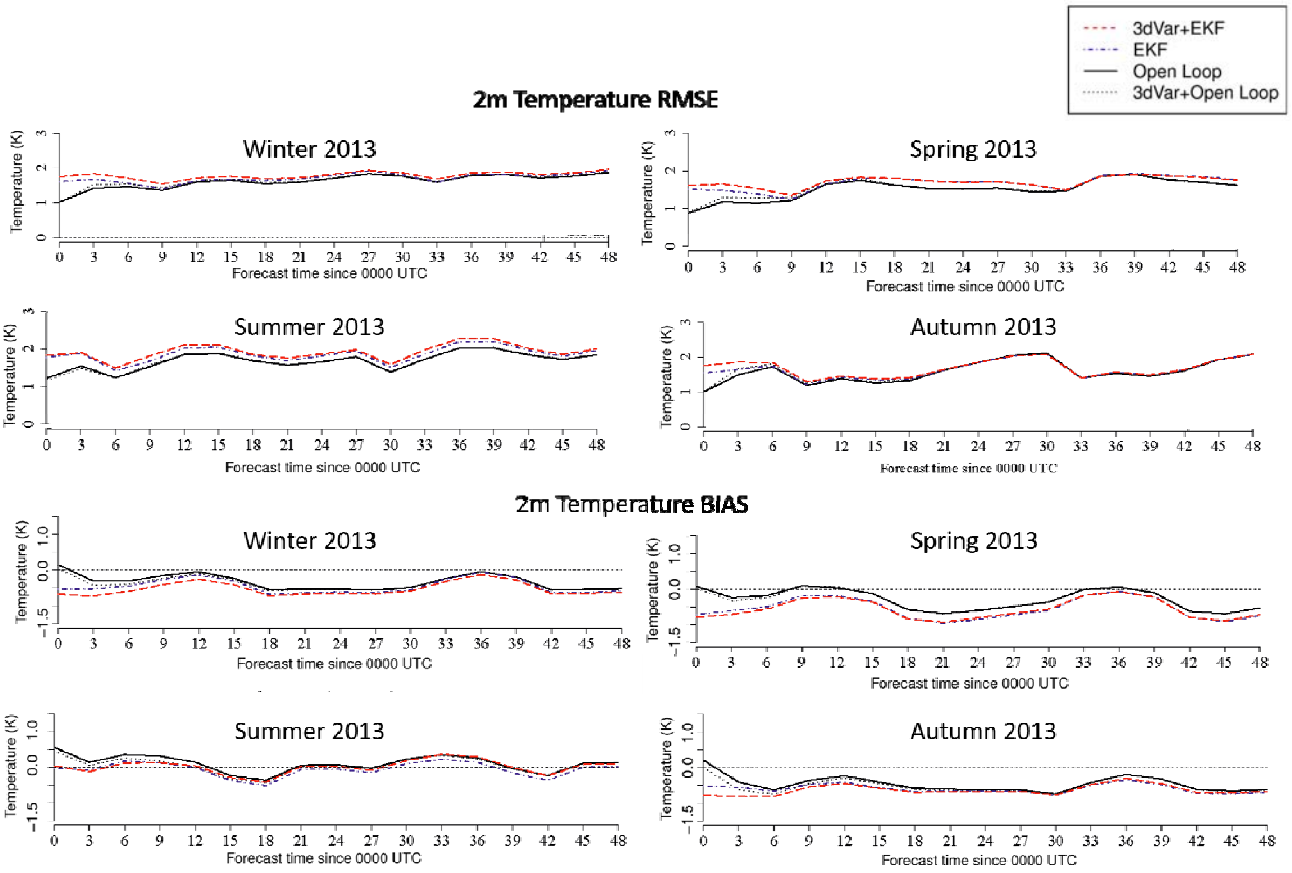

Figure 8 Screen-level Temperature forecast scores in terms of RMSE and BIAS with a leading time up to 48 hours for OL, EKF, 3D-Var+EKF and 3D-Var+OL averaged over 11 stations in Belgium (displayed in figure 6) calculated for spring, summer, autumn and winter.

For T_{2m} the surface assimilation runs seem to have more difficulties in improving the scores. Except during summer, there is a persistent cold bias in all four runs for the observation locations. Especially during spring the cold bias of EKF and 3D-Var+EKF is worse than that of OL and 3D-VAR+OL, while during autumn and winter the biases of the four runs are very similar. The difficulties of the local surface assimilation during spring could potentially be explained by the very cold spring that year, which was marked by 11 snow days in Uccle. In cold and cloudy situations the coupling between the surface and the atmosphere is weak, so the surface assimilation will not be able to correct much for errors in the soil. During the snow days the coupling is even weaker and the layer of snow acts as a barrier between the temperature of the soil and that of the air. The ARPEGE surface analysis seems to handle this situation above Belgium better. This difference might be explained by differences in parametrizations and physiography between ARPEGE and ALARO coupled to SURFEX. In summer the cold temperature bias is removed in the observation locations and all four runs have a very small bias. EKF and 3D-Var+EKF have a bias close to zero during summer for the first 12 hours in the observation locations, while the bias of OL and 3D-Var+OL is slightly higher.

4.4 Upper-air Verification

For the upper-air verification sounding measurements were used from the Wyoming Weather web database (Oolman 2015) for 4 different locations (see table 8 and figure 9). These sounding observations were not used in the 3D-Var assimilation process. Sounding scores for temperature (fig. 10) and specific humidity (fig. 11) were calculated for each of the four seasons and averaged over these 4 locations. From figures 10 and 11 it is clear that our local 3D-Var system is not able to get an equally good upper-air analysis as ARPEGE for the observation locations.

Figure 9 Locations of stations for the sounding observations (cfr. Table 8).

The temperature RMSE scores for OL and EKF are consistently lower than those of 3dVar+OL and 3dVar+EKF. This is probably due to the limited observation types that are used in our 3D-Var system and adding satellite observations could reduce this RMSE (Randriamampianina and Storto 2008). With respect to the bias, the differences between the runs with and without local 3D-Var are not as pronounced in the observation locations. During spring the 3D-Var+EKF run has a smaller temperature bias than the EKF run at a height between 700hPa and 300hPa. Also for the lowest model level the 3D-Var runs have a

smaller temperature bias than the ARPEGE runs in spring. At the lowest model levels, during summer and autumn the smallest bias is simulated by the 3D-Var+OL run, while in winter 3D-Var+EKF performs best.

Also for specific humidity the RMSE of the ARPEGE interpolated runs is lower than that of the 3D-Var runs, especially for the lower layers (> 700hPa). While the specific humidity bias on average varies around zero for all upper-air model levels, during winter there is a dry bias for all runs in the lower atmosphere. In the lowest part of the atmosphere (> 900hPa) differences between the runs arise, indicating the influence of the surface assimilation for this part of the atmosphere.

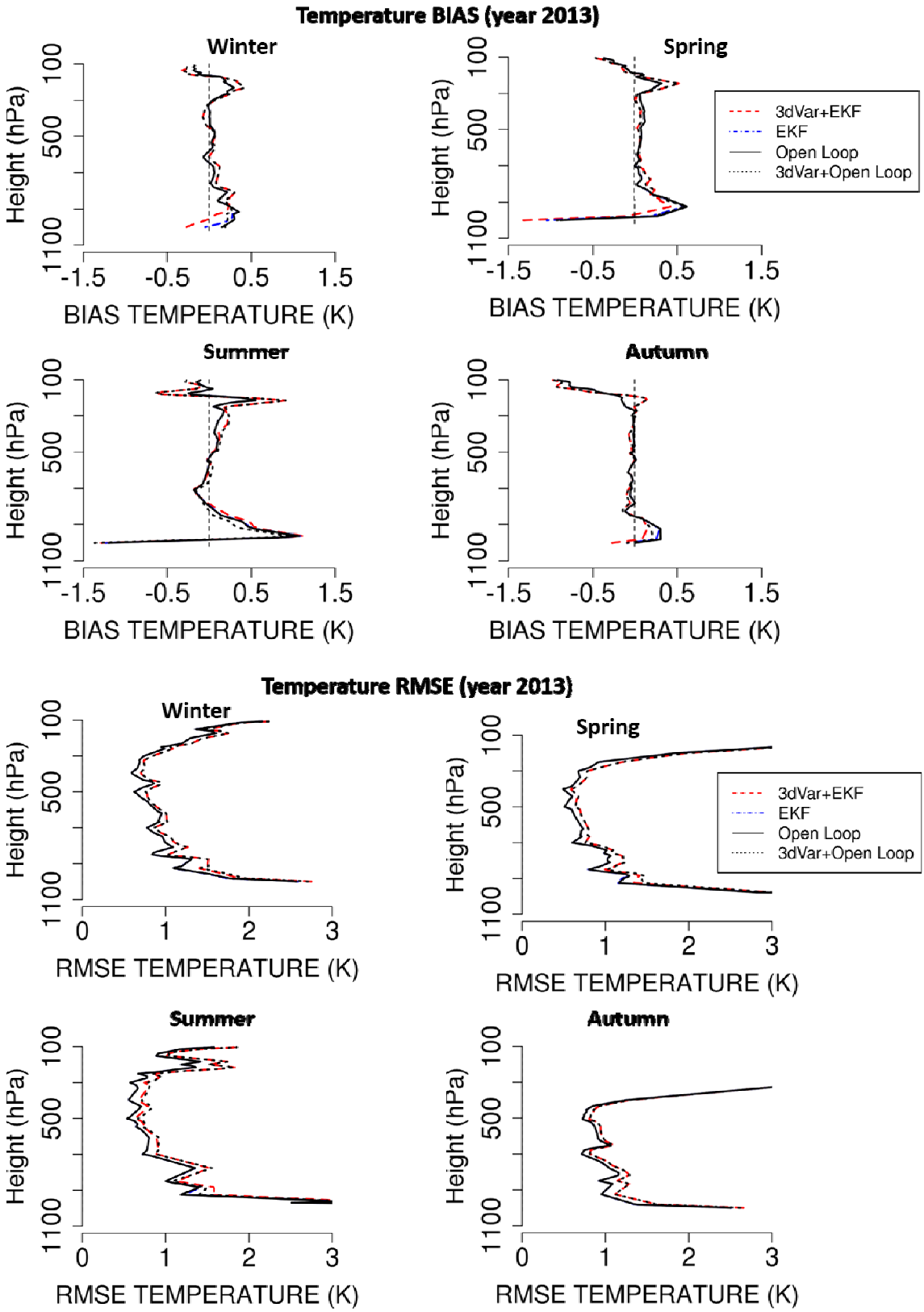

Figure 10 Temperature forecast scores against radio soundings during the four seasons of the year 2013, averaged over the four sounding stations displayed in table 8

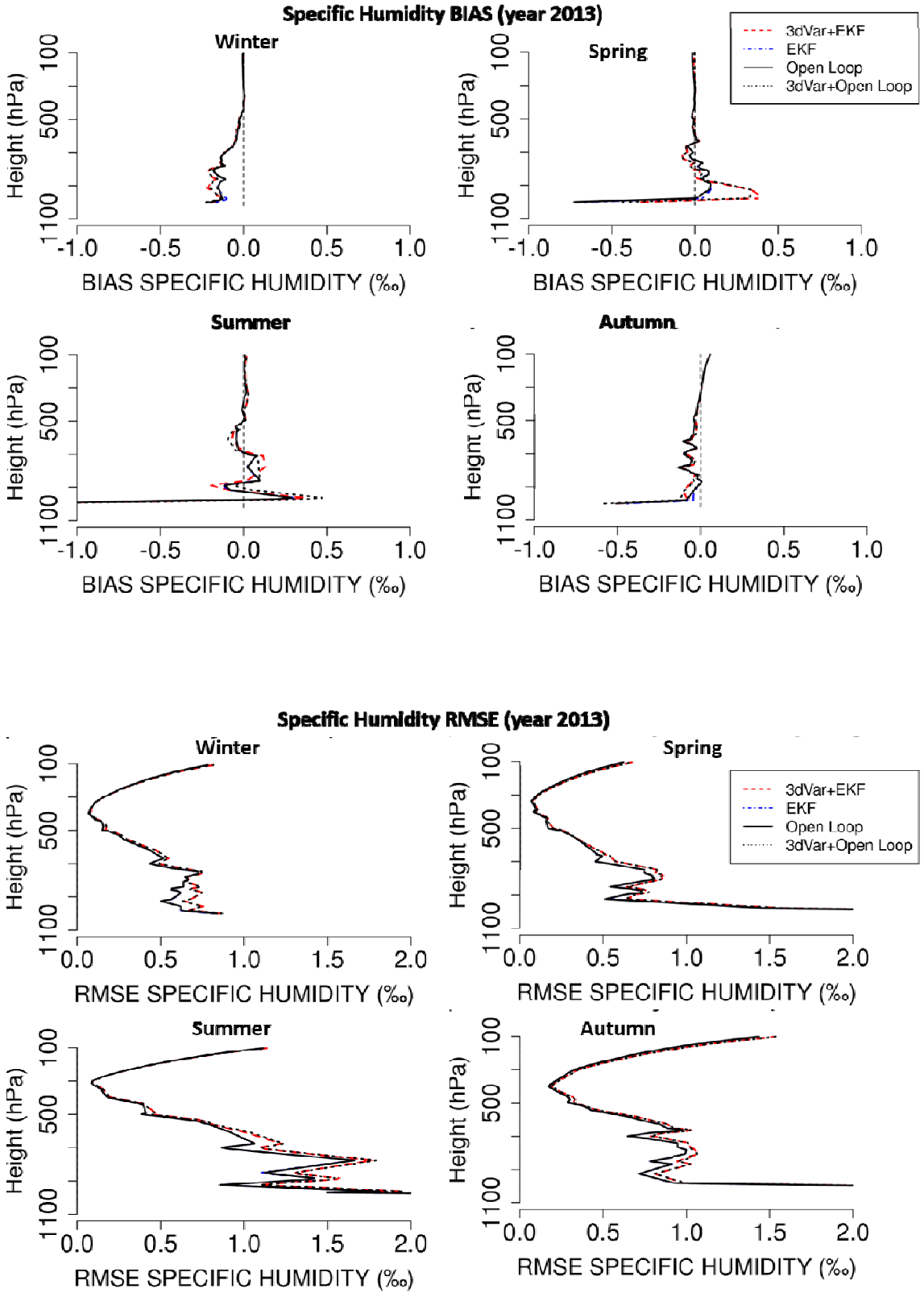

Figure 11 Specific Humidity forecast scores against radio soundings during the four seasons of the year 2013, averaged over the four sounding stations displayed in table 8

4.5 Precipitation Verification

The precipitation forecasts are verified against quantitative precipitation estimates with a radar-gauge merging method following Goudenhoofdt and Delobbe (2009), using the SAL (structure, amplitude and location) method of Wernli et al. (2008). The structure component characterizes the size and shape of the precipitation objects. Its values range from -2 (predicted precipitation objects too small or too peaked) to 2 (predicted precipitation objects too large or too flat). For a value of $S = 0$, the model has the correct structure. The amplitude component is related to the total precipitation amount. It varies between -2 (an under-predicted total precipitation amount) and 2 (an over-predicted total precipitation amount), with a value of 0 corresponding to a perfect forecast amplitude. Finally, the location component quantifies whether the predicted precipitation objects are situated at the correct location. It ranges from 0 (predicted precipitation objects at correct position) to 2 (predicted precipitation objects at incorrect position).

Table 9 shows the SAL-scores for the different experiments, averaged per season. For all three components all values are positive, indicating that in general the precipitation forecasts are too large and intense and are somewhat mislocated. Except during autumn, the best (lowest) structure (S) and amplitude (A) scores are for the EKF. The ARPEGE 4D-Var analysis assimilates satellite data and probably has better humidity profiles due to this, while the local 3D-Var set-up assimilates only conventional observations and no satellite or high-density observations. Compared to the OL, with ARPEGE downscaling for both the surface and the upper-air, the EKF has a small positive effect on the structure and amplitude. With respect to the location, during winter and autumn the EKF outperforms the other runs. In spring the OL has the best location score, although all scores lie very close together. During summer the 3D-Var+OL run has the best location score.

Compared to the other seasons, the structure (S) and length (L) values are highest in summer. This indicates that during summer the predicted precipitation objects are too large and at a wrong location. The model has some problems predicting the small scale features of the precipitation. This is probably due to the fact that summer is typically characterized by more localized convective precipitation than the other seasons. The amplitude (A) scores are in general much lower during summer than those during spring, so the amount of precipitation is well captured by the model during summer.

In autumn the scores of all components are low compared to spring and summer. The precipitation is only slightly too intense. While September 2013 was a relatively dry month, the largest amount of precipitation in autumn fell during October and November and consisted mostly of frontal precipitation. The model seems to handle this precipitation very well. With regards to the amplitude (A) in autumn, all runs with 3D-Var for the upper-air have somewhat lower scores than the runs with an interpolated ARPEGE upper-air analysis, so 3D-Var seems to be able to better represent the upper-air humidity in this season. The SAL scores for winter are somewhat higher than in autumn but still relatively low and all larger than zero. During winter the EKF has the best scores for all three components.

Overall the EKF, with an interpolated ARPEGE upper-air analysis but a local surface assimilation, outperforms the other runs for all seasons except for location during summer and spring and amplitude during autumn. Summer is characterised by the highest SAL scores, probably due to the convective nature of precipitation in summer that is harder to predict precisely.

5 Conclusions and perspectives

In this paper the offline filtered EKF presented in Duerinckx et al. (2015) for soil analysis was combined with a 3D-Var assimilation for the upper-air analysis for the LAM ALARO

coupled to SURFEX. This 3D-Var+EKF set-up for ALARO was tested for the first time and was compared to a number of other possible initialisation methods for the surface and the atmosphere. A first aim was to examine whether the EKF surface assimilation combined or not with a basic 3D-Var upper-air assimilation has an added value compared to the Open Loop, in which the more advanced upper-air data assimilation of ARPEGE is interpolated. A second aim was to investigate whether the results of this comparison depend on the upper-air data assimilation method. All initialisation methods have been tested for a one-year period in order to compare the results for all seasons.

To achieve these goals, the different experiments were compared in terms of soil moisture and soil temperature increments and a verification was done against soil measurements, independent screen-level and radiosounding observations and merged radar-rain-gauge precipitation observations.

The analysis of increments indicates that the model screen-level values are in general too cold except during summer. The surface assimilation runs (EKF and 3D-VAR+EKF) are able to reduce part of this cold trend, without completely removing it. It is likely that the soil is not the only factor causing the cold trend. With regards to soil moisture content, there is a drying trend in the increments of the EKF and 3D-Var+EKF run during spring and a moistening trend during the other seasons. By examining the screen-level forecast scores, the increments are especially successful in removing the wet trend during spring and to a lesser extend in autumn and winter. Differences between the runs with and without a local 3D-Var upper-air assimilation indicate that the runs with 3D-Var have larger screen-level departures that result in larger soil increments.

For the verification of the soil variables, measurements from the Fluxnet database were used. The soil verification was done according to the soil moisture validation strategy of the ECMWF (Albergel et al. 2012). The biases for soil moisture and soil temperature indicate

that in the observations locations the soil is too moist and cold during winter, while in summer the soil is too warm but with only a very small moisture bias. During summer, autumn and winter the soil moisture biases for the four runs all lie very close together. In Spring the bias of 3D-VAR+EKF is lower than the other biases. In autumn and winter the soil temperature bias is lowest for the OL, while in spring the 3D-Var+EKF and EKF runs outperform OL. In summer the EKF run has the lowest soil temperature bias. Since in situ measurements of soil moisture and temperature contain considerable errors and often have an unknown uncertainty, the correlation of these measurements with the model values is an important additional verification score. Correlations for all variables were lowest during summer. For soil moisture the correlations during summer are very low, with the highest correlation for the 3D-VAR+EKF run. The model is generally too dry during summer and there are not so many precipitation events. The surface assimilation tries to correct this bias through the soil, while this is probably not the only cause of the bias. This causes the soil state to diverge from the observations. It seems that the 3D-Var+EKF run is better able to handle the too dry model state of the Summer. This is also reflected in the lower RMSD scores for 3D-VAR+EKF during summer.

The screen-level verification forecast scores indicate the benefits of the EKF surface assimilation for the relative humidity scores, especially during the first six to twelve hours. The EKF runs are able to gain similar or even better relative humidity scores compared to the Open Loop run. For screen-level temperature, the EKF runs seem to have more difficulties in improving the Open Loop scores and the model suffers from a persistent cold trend during all seasons except for summer. When looking at the 3D-Var+OL run, it is clear that the 3D-Var upper-air assimilation is able to improve the relative humidity scores during the first 12 hours of the forecast compared to the Open Loop. During winter the combination of surface and upper-air assimilation of the 3D-Var+EKF run outperforms the other runs. For summer on the

other hand, the EKF run performs better than the 3D-Var+EKF run.

To verify the upper-air, radiosounding observations were used from 4 different locations. The scores indicate that the current 3D-Var system is not able to get an upper-air analysis as accurate as the ARPEGE interpolated atmosphere. This is probably due to the more sophisticated 4D-Var technique used for ARPEGE and the limited number of observation types used in our 3D-Var set-up. Adding satellite observations could reduce this weakness. For the lowest model levels (pressure > 900hPa) the influence of the different surface-assimilation set-ups is visible in the forecast scores of the 0h forecast range against radiosoundings.

Similarly, the precipitation verification shows the need to add observation types for the upper-air analysis. The best precipitation scores from 24h accumulations are in general for the runs with the interpolated ARPEGE upper-air analysis and not the 3D-Var runs. Only in autumn the 3D-Var runs (EKF and 3D-VAR+EKF) outperform the ARPEGE interpolated runs. In spring the precipitation forecasts are too large and intense. In summer the precipitation scores are the worst, with objects that are generally too large and at the wrong location, while the amount of precipitation is well captured. In autumn and winter the precipitation scores are very good. The model appears to be able to capture frontal precipitation systems very well but has more difficulties with the small-scale features of the convective precipitation in summer. For the SAL precipitation scores the combination of interpolated ARPEGE upper-air analysis with the EKF surface analysis (the 'EKF' run) seems to be the most beneficial. This run shows the best (lowest) scores for structure and amplitude in all seasons except for autumn. Also with respect to location, this EKF run outperforms the other runs in winter and autumn. In winter the EKF run outperforms the other runs for all three SAL-components.

Overall it can be concluded that the EKF surface assimilation has a positive effect on

the soil moisture and screen-level humidity scores, especially during spring and summer and it is able to produce similar or improved scores compared to the Open Loop set-up. Also for the SAL precipitation scores the EKF surface assimilation is beneficial, especially during winter. For temperature the benefits of the surface assimilation are less pronounced in the screen-level temperature scores, but it still manages to get similar scores as the Open Loop in most cases.

The 3D-Var assimilation mainly suffers from a lack of observations resulting in higher RMSE values against radiosounding and larger precipitation errors. Still, the potential benefits of the combination of upper-air and surface assimilation is shown in the soil moisture and screen-level relative humidity verification. When the 3D-Var set-up is extended to include satellite, GNSS and radar data, this should benefit all verification scores and improvements compared to the Open Loop set-up are very likely.

From all the above results it appears beneficial to invest in an EKF soil analysis, especially to improve the model humidity scores. Moreover it should be investigated whether the extension of the basic 3D-Var setup with satellite, GNSS and radar data can help to resolve the issues with the upper-air noticed in this paper.

Acknowledgements

The authors thank the reviewers for their useful comments and suggestions, which greatly helped to improve this manuscript. The authors also wish to thank Julie Berckmans for her advice and insightful discussions.

This work used soil moisture and soil temperature observation data acquired and shared by the FLUXNET community, including these networks: AmeriFlux, AfriFlux, AsiaFlux, CarboAfrica, CarboEuropeIP, CarboItaly, CarboMont, ChinaFlux, Fluxnet-Canada, GreenGrass, ICOS, KoFlux, LBA, NECC, OzFlux-TERN, TCOS-Siberia, and USCCC.

References

Albergel C, De Rosnay P, Balsamo G, Munoz-Sabater J, Boussetta S, Isaksen L. 2012. ECMWF soil moisture validation activities. *ECMWF Newsletter* **133**: 23–29.

Balsamo G, Viterbo P, Beljaars A, van den Hurk B, Hirschi M, Betts A, Scipal K. 2009. A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the integrated forecast system. *J. Hydromeor.* **10**: 623–643.

Bélair S, Crevier LP, Mailhot J, Bilodeau B, Delage Y. 2003. Operational implementation of the ISBA land surface scheme in the canadian regional weather forecast model. part i: Warm season results. *J. Hydrometeorol.* **4**: 352–370.

Beljaars A, Viterbo P, Miller M, Betts A. 1996. The anomalous sensitivity over USA during july 1993: sensitivity to land surface parameterization. *Mon.Wea.Rev.* **124**: 362–383.

Berre L, Stefanescu S, Belo Pereira M. 2006. The representation of the analysis effect in three error simulation techniques. *Tellus* **58A**: 196–209.

Berre L. 2000. Estimation of synoptic and mesoscale forecast error covariances in a limited area model. *Mon. Weather Rev.* **128**: 644–667.

Berre L, Desroziers G. 2010. Filtering of background error variances and correlations by local spatial averaging. Mon. Weather Rev. 138: 3693–3720, doi: 10.1175/2012MWR3111.1.

Best M, Beljaars A, Polcher J, Viterbo P. 2004. A proposed structure for coupling tiled surfaces with the planetary boundary layer. *J. of Hydromet.* **5**: 1271–1278.

Betts A, Viterbo P. 2005. Land-surface, boundary layer and cloud field coupling over the southwestern amazon in ERA-40. *J. Geophys. res.* **110**: 15pp.

Bölöni G, Szabó L, Trojáková A. 2009. OPLACE: common observation preprocessing for lace. Technical report, Hungarian Meteorological Service, URL

http://www.rclace.eu/File/Data_Assimilation/2008/oplace.pdf .

Brousseau P, Desroziers G, Bouttier F, Chapnik B. 2014. A posteriori diagnostics of the impact of observations on the AROME-france convective-scale data assimilation system. *Quarterly Journal of the Royal Meteorological Society* **140**(680): 982–994.

Bubnová R, Hello G, Bénard P, Geleyn JF. 1995. Integration of the fully elastic equations cast in the hydrostatic pressure terrain-following coordinate in the framework of the ARPEGE/ALADIN NWP system. *Mon. Weather Rev.* **123**: 515–535.

Catry B, Geleyn J, Tudor M, Bénard P, Trojáková A. 2007. Flux-conservative thermodynamic equations in a mass-weighted framework. *Tellus A* **59**(1), URL http://www.tellusa.net/index.php/tellusa/article/view/14856 .

Chambon, P., Meunier, L.-F., Guillaume, F., Piriou, J.-M., Roca, R. and Mahfouf, J.-F. (2015), Investigating the impact of the water-vapour sounding observations from SAPHIR on board Megha-Tropiques for the ARPEGE global model. Q.J.R. Meteorol. Soc., 141: 1769–1779. doi:10.1002/qj.2478

Christensen H, Machenhauer B, Jones R, Schar C, Ruti P, Castro M, Visconti G. 1997. Validation of present-day regional climate simulations over Europe: LAM simulations with observed boundary conditions. *Climate Dynamics* **13**: 489–506.

Courtier P, Andersson E, Heckley W, Pailleux J, Vasiljevic D, Hamrud M, Hollingsworth A, Rabier F, , Fisher M. 1998. The ECMWF implementation of three dimensional variational assimilation (3D-Var). i: Formulation. *Q. J. R. Meteorol. Soc.* **124**: 1783–1807.

Deardorff J. 1977. Parameterization of ground surface moisture content for use in atmospheric prediction models. *J. Appl. Meteorol.* **16**: 1182–1185.

Deardorff J. 1978. Efficient prediction of ground surface temperature and moisture, with inclusion of a layer of vegetations. *J. Geophys. Res.* **83**: 1889–1903.

Derber J, Bouttier F. 1999. A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus* **51A**: 195–221.

De Meutter P, Gerard L, Smet G, Hamid K, Hamdi R, Degrauwe D, Termonia P. 2015. Predicting small-scale, short-lived downbursts: case study with the NWP limited-area ALARO model for the Pukkelpop thunderstorm. *Mon. Wea. Rev.* **143**: 742–756, doi:10.1175/MWR-D-14-00290.1 .

Desroziers G, Pouponneau B, Thépaut JN, Janisková M, Veersé F. 1999. Four-dimensional variational analyses of FASTEX situations using special observations. Q.J.R. Meteorol. Soc. 125: 3339–3358, doi: 10.1002/qj.49712556111.

de Rosnay P, Drusch M, Vasiljevic D, Balsamo G, Albergel C, Isaksen L. 2012. A simplified extended kalman filter for the global operational soil moisture analysis at ECMWF. *Q.J.R. Meteorol. Soc.* **139**: 1199–1213.

De Troch R, Hamdi R, Van de Vyver H, Geleyn JF, Termonia P. 2013. Multiscale performance of the ALARO-0 model for simulating extreme summer precipitation climatology in belgium. *J. Climate* **26**: 8895–8915.

Douville H, Viterbo P, Mahfouf JF, Beljaars A. 2000. Evaluation of the optimum interpolation and nudging techniques for soil moisture analysis using fife data. *Monthly Weather Review* **128**: 1733–1756.

Draper CS, Mahfouf JF, Walker JP. 2009. An EKF assimilation of AMSR-E soil moisture into the ISBA land surface scheme. *Journal of Geophysical Research: Atmospheres* **114**(D20104).

Draper CS, Mahfouf JF, Walker JP. 2011. Root zone soil moisture from the assimilation of screen-level variables and remotely sensed soil moisture. *Journal of Geophysical Research: Atmospheres* **116**(D02127).

Drusch M. 2007. Initializing numerical weather prediction models with satellite-

derived surface soil moisture: Data assimilation experiments with ECMWF's integrated forecast system and the tmi soil moisture data set. *Journal of Geophysical Research: Atmospheres* **112**(D03102).

Duerinckx A, Hamdi R, Mahfouf JF, Termonia P. 2015. Study of the Jacobian of an extended kalman filter for soil analysis in surfexv5. *Geoscientific Model Development* **8**(5): 1–19.

FAO. 2006. World reference base for soil resources 2006, a framework for international classification, correlation and communication, world soil resources report no. 103.

Faroux S, Kaptué Tchuenté A, Roujean JL, Masson V, Martin E, Le Mogne P. 2013. Ecoclimap-ii/europe: A twofold database of ecosystems and surface parameters at 1km resolution based on satellite information for use in land surface, meteorological and climate models. *Geosci. Model Dev.* **6**: 563–582.

Ferranti L, Viterbo P. 2006. The European summer of 2003: Sensitivity to soil water initial conditions. *J. Climate* **19**: 3659–3680.

Fillion L, Tanguay M, Lapalme E, Denis B, Desgagne M, Lee V, Ek N, Liu Z, Lajoie M, Caron JF, Pagé C. 2010. The Canadian regional data assimilation and forecasting system. *Weather and Forecasting* **25**: 1645–1669.

Fischer C, Montmerle T, Berre L, Auger L, Stefanescu S. 2005. An overview of the variational assimilation in the ALADIN/France numerical weather-prediction system. *Q.J.R.Meteorol.Soc.* **131**: 3477–3492.

Fisher M. 2003. Background error covariance modelling. *Proc. ECMWF Seminar on Recent Developments in Data Assimilation for Atmosphere and Ocean* : 45–63.

Fluxnet. 2015. Available online [http://fluxnet.ornl.gov from ORNL DAAC, Oak Ridge, Tennessee, U.S.A.

Gerard L, Geleyn JF. 2005. Evolution of a subgrid deep convection parametrization in a limited-area model with increasing resolution. *Quarterly Journal of the Royal Meteorological Society* **131**(610): 2293–2312.

Gerard L, Piriou JM, Brožková R, Geleyn JF, D B. 2009. Cloud and precipitation parameterization in a meso-gamma-scale operational weather prediction model. *Mon. Wea. Rev.* **137**: 3960–3977.

Gerard L. 2007. An integrated package for subgrid convection, clouds and precipitation compatible with meso-gamma scales. *Quarterly Journal of the Royal Meteorological Society* **133**(624): 711–730.

Giard D, Bazile E. 2000. Implementation of a new assimilation scheme for soil and surface variables in a global NWP model. *Monthly Weather Review* **128**: 997–1015.

Gilbert J, Lemarechal C. 1989. Some numerical experiments with variable-storage quasi-newton algorithms. *Mathematical Programming* **45**: 407–435.

Giot O, Termonia P, Degrauwe D, De Troch R, Caluwaerts S, Smet G, Berckmans J, Deckmyn A, De Cruz L, De Meutter P, Duerinkcx A, Gerard L, Hamdi R, Van den Bergh J, Van Ginderachter M, Van Schaeybroeck B. 2016. Validation of the ALARO-0 model within the euro-cordex framework. *Geosci. Model Dev.*, **9**: 1143-1152

Goudenhoofdt E, Delobbe L. 2009. Evaluation of radar-gauge merging methods for quantitative precipitation estimates. *Hydrol. Earth syst. Sci.* **13**: 195–203.

Hamdi R, Deckmyn A, Termonia P, Demarée GR, Baguis P, Vanhuysse S, Wolff E. 2009. Effects of historical urbanization in the Brussels capital region on surface air temperature time series: A model study. *J. Appl. Meteorol. Clim.* **48**: 2181–2196.

Hamdi R, Degrauwe D, Duerinckx A, Cedilnik J, Costa V, Dalkilic T, Essaouini K, Jerczynki M, Kocaman F, Kullmann L, Mahfouf JF, Meier F, Sassi M, Schneider S, Váňa F, Termonia P. 2014. Evaluating the performance of SURFEXv5 as a new land surface scheme

for the ALADIN cy36 and ALARO-0 models. *Geoscientific Model Development* **7**(1): 23–39, URL http://www.geosci-model-dev.net/7/23/2014/ .

Hamdi R, Termonia P, , Baguis P. 2011. Effects of urbanization and climate change on surface runoff of the Brussels capital region: a case study using an urban soil-vegetation-atmosphere-transfer model. *Int. J. Climatol* **31**: 1959–1975.

Hamdi R, Van de Vyver H, , Termonia P. 2012. New cloud and microphysics parameterisation for use in high-resolution dynamical downscaling: application for summer extreme temperature over Belgium. *Int. J. Climatol.* **32**: 2051–2065.

Hess H. 2001. Assimilation of screen-level observations by variational soil moisture analysis. *Meteorol. Atmos. Phys.* **77**: 145–154.

Hess R, Lange M, Wergen W. 2008. Evaluation of the variational soil moisture assimilation scheme at Deutscher Wetterdienst. *Q.J.R. Meteorol. Soc.* **134**: 1499–1512.

Lafore JP, Stein J, Asencio N, Bougeault P, Ducrocq V, Duron J, Fischer C, Héreil P, Mascart P, Masson V, Pinty JP, Redelsperger JL, Richard E, Vilà Guereau de Arellano J. 1998. The meso-NH atmospheric simulation system. part I: adiabatic formulation and control simulations. *Annales Geophysicae* **19**: 90–109.

Lebeaupin C. 2007. Etude du couplage ocean-atmosphere associe aux episodes de pluies intenses en region mediterranennee. PhD thesis, University Paul Sabatier, Toulouse, France.

Mahfouf JF, Bergaoui K, Draper C, Bouyssel F, Taillefer F, Taseva L. 2009. A comparison of two off-line soil analysis schemes for assimilation of screen level observations. *Journal of Geophysical Research: Atmospheres* **114**(D8): n/a–n/a.

Mahfouf JF, Bliznak V. 2011. Combined assimilation of screen-level observations and radar-derived precipitation for soil moisture analysis. *Q.J.R. Meteorol. Soc.* **137**: 709–722.

Mahfouf JF, Douville H, Viterbo P, Beljaars A, Saarinen S. 2000. A revised land

surface analysis scheme in the integrated forecasting system. *ECMWF Newsl.* **88**: 8–13.

Masson V, Le Moigne P, Martin E, Faroux S, Alias A, Alkama R, Belamari S, Barbu A, Boone A, Bouyssel F, Brousseau P, Brun E, Calvet JC, Carrer D, Decharme B, Delire C, Donier S, Essaouini K, Gibelin AL, Giordani H, Habets F, Jidane M, Kerdraon G, Kourzeneva E, Lafaysse M, Lafont S, Lebeaupin Brossier C, Lemonsu C, Mahfouf JF, Marguinaud P, Mokhtari M, Morin S, Pigeon G, Salgado R, Seity Y, Taillefer T, Tanguy G, Tulet P, Vincendon B, Vionnet V, Voldoire A. 2013. The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes. *Geosci. Model Dev. Discuss.* **5**: 3771–3851.

Masson V. 2000. A physically-based scheme for the urban energy budget in atmospheric models. *Bound.-Layer Meteor.* **94**: 357–397.

Mironov D, Heise E, Kourzeneva E, Ritter B, Schneider N, Terzhevik A. 2010. Implementation of the lake parameterisation scheme FLAKE into the numerical weather prediction model COSMO. *Boreal Env. Res.* **15**: 218–230.

Mokhatari M, Gomes L, Tulet P, Rezoug T. 2012. Importance of the surface size distribution of erodible material: an improvement on the dust entrainment and deposition (DEAD) model. *Geosci. Model Dev.* : 581–598.

Montmerle T, Lafore JP, Berre L, Fischer C. 2006. Limited-area model error statistics over western Africa: Comparisons with midlatitude results. *Q. J. R. Meteorol. Soc.* **132**: 213–230.

Moors E. 2015. Loobos website. http://www.climatexchange.nl/sites/loobos/index.htm . Accessed: 2015-09-10.

Noilhan J, Mahfouf JF. 1996. The ISBA land surface parameterisation scheme. *Global and Planetary Change* **13**: 145 – 159, URL http://www.sciencedirect.com/science/article/pii/0921818195000437 .

Noilhan J, Planton S. 1989. A simple parameterization of land surface processes for meteorological models. *Monthly Weather Review* **117**: 536–549.

Oolman L. 2015. Wyoming weather web database. http://weather.uwyo.edu/upperair/sounding.html . Accessed: 2015-09-10.

Parrish D, Derber J, RJ P, Wu WS, Z-X P. 1997. The NCEP global analysis system: recent improvements and future plans. *J. Meteorol. Soc. Japan* **75**: 359–365.

Pereira M, Berre L. 2006. The use of an ensemble approach to study the background error covariances in a global NWP model. *Montly Weather Review* **134**: 2466–2489.

Polcher J, McAvaney B, Viterbo P, Gaertner MA, Hahmann A, Mahfouf JF, Noilhan J, Phillips T, Pitman A, Schlosser C, Schulz JP, Timbal B, Verseghy D, Xue Y. 1998. A proposal for a general interface between land-surface schemes and general circulation models. *Global Planet. Change* : 263–278.

Randriamampianina R, Storto A. 2008. ALADIN-HARMONIE/Norway and its assimilation system - the implementation phase. *HIRLAM Newletter* **54**: 20–29.

Rodriguez E, Navascues B, Ayuso J, Jarvenoja S. 2003. Analysis of the surface variables and parameterisation of the surface processes in HIRLAM. part I: Approach and verification by parallel runs. *HIRLAM Tech. Rep.* **59**: 52pp.

Sadiki W, Fischer C, Geleyn JF. 2000. Mesoscale background error covariances: recent results obtained with the limited-are model ALADIN over Morocco. *Mon. Weather Rev.* **128**: 3927–3935.

Schneider S, Meier F, Yan X. 2009. Data assimilation activities at zamg (Austria).

Seity Y, Brousseau P, Malardel S, Hellop G, Bénard P, Bouttier F, Lac C, Masson V. 2011. The AROME-france convective scale operational model. *Mon. Weather Rev* **139**: 976–991.

Seuffert G, Gross P, Simmer C, Wood E. 2002. The influence of hydrologic modelling

on the predicted local weather: Two-way coupling of a mesoscale weather prediction model and a land surface hydrologic model. *J. Hydrometeorol.* **3**: 505–523.

Stanešic A. 2011. Assimilation system at DHMZ: Development and first verification results. *Croatian Meteorological Journal* **44/45**: 3–17.

Stevanescu S, Berre L, Belo Pereira M. 2006. The evolution of dispersion spectra and the evaluation of model differences in an ensemble estimation of error statistics for a limited area analysis. *Mon. Wea. Rev.* **134**: 3454–3476.

Tudor M, Ivatek-Šahdan S, Stanešic A, Horvath K, Bajic A. 2013. Forecasting weather in Croatia using ALADIN numerical weather prediction model.

Van den Hurk B, Viterbo P, Beljaars A, Betts A. 2000. Offline validation of the ERA-40 surface scheme. *ECWF Tech. Memo.* **259**: 43pp.

Wernli H, Paulat M, Hagen M, Frei C. 2008. SAL – a novel quality measure for the verification of quantitative precipitation forecasts. *Mon. Weather Rev.* **136**: 4470–4487.

WMO. 1997. The ALADIN project: Mesoscale modelling seen as a basic tool for weather forecasting and atmospheric research. *WMO bull.* **46**: 317–324.

	Initial State		LBC
	Atmosphere	**Surface**	
Open Loop (OL)	Interpolated ARPEGE 4D-Var Analysis	Interpolated ARPEGE OI Analysis	3-hourly coupling to ARPEGE forecast
EKF	Interpolated ARPEGE 4D-Var Analysis	Local SURFEX Extended Kalman Filter	
3DVar+OL	Local ALARO 3D-Var	Interpolated ARPEGE OI Analysis	
3DVar+EKF	Local ALARO 3D-Var	Local SURFEX Extended Kalman Filter	

Table 1 Overview of the initial and lateral boundary conditions used for the soil and for the atmosphere in the different experiments that were run.

Location	Lon (°E)	Lat (°N)	height (m)	Vegetation	Wres (m^3/m^3)	Wsat (m^3/m^3)	T_s depth (m)	W_g depth (m)
Brasschaat (Bra, BE)	4.52	51.31	16	broadleaved forest			0.02	/
Lonzée (Lon, BE)	4.74	50.55	165	cropland	Model: 0.16 Obs: 0.17	Model: 0.44 Obs: 0.43	/	0.05
Vielsalm (Vie, BE)	6.00	50.30	491	broadleaved forest	Model: 0.16 Obs: 0.27	Model: 0.44 Obs: 0.41	0.03	0.02
Loobos (Loo, NL)	5.74	52.17	25	needle leaf evergreen forest	Model: 0.08 Obs: 0.03	Model: 0.40 Obs: 0.43	0.05	0.05
Cabauw (Ca1, NL)	4.93	51.97	0.7	grassland polder			0.03	/
Rollesbroich (RuR, DE)	6.30	50.62	508	grassland			0.04	/
Selhausen (RuS, DE)	6.45	50.87	103	urban and cropland			0.04	/

Table 2 Overview of the Fluxnet observations used for the soil verification indicating the location of the observations (where BE is Belgium, NL is the Netherlands and DE is Germany), the longitude, the latitude, , the altitude(height), the vegetation type, the residual soil moisture level, the saturation soil moisture level and the depth of the observations. When no observations were available of the given variable at that location this is indicated with "/"). The residual and saturation soil moisture level are only indicated for the locations in which Wg observations are available.

T_s	**spring**			**summer**			**Autumn**			**winter**		
	Cor	BIAS	RMSD	Cor	BIAS	RMSD	Cor	BIAS	RMSD	Cor	BIAS	RMSD
OL	0,77	0,69	**4,34**	**0,55**	2,75	**5,28**	**0,82**	**-0,99**	**3,47**	**0,72**	**-1,60**	**3,83**
EKF	0,77	-0,28	4,54	0,52	**2,46**	5,52	0,79	-2,08	4,29	0,66	-2,64	4,68
3DVAR+EKF	0,77	**-0,19**	4,51	0,51	2,83	5,96	0,79	-2,03	4,34	0,64	-2,77	4,81

Table 3 Superficial soil temperature (Ts) correlation, Bias (K) and Root Mean Square Difference (K) calculated for season for OL, EKF and 3D-Var+EKF during the year 2013. The values are averaged over the stations of Bra, Ca, Loo, Vie, Rur, RuS (cfr. Table 2). The best scores per parameter per season are indicated in bold.

W_g	**Spring**			**summer**			**autumn**			**winter**		
	Cor	BIAS	RMSD	Cor	BIAS	RMSD	Cor	BIAS	RMSD	Cor	BIAS	RMSD
OL	**0,28**	-0,196	0,270	0,22	**0,043**	0,219	**0,63**	0,010	0,162	**0,47**	**-0,278**	**0,357**
EKF	0,23	-0,137	**0,237**	0,22	0,058	0,217	0,59	0,009	0,162	0,38	-0,290	0,383
3DVAR+EKF	0,16	**-0,122**	0,238	**0,33**	0,045	**0,207**	0,55	**0,006**	0,169	0,37	-0,297	0,393

Table 4 Superficial soil water content (Wg) correlation, Bias and Root Mean Square Difference calculated for season for OL, EKF and 3D-Var+EKF during the year 2013. The values are averaged over the stations of Loo, Lon, Vie (cfr. Table 2). The best scores per parameter per season are indicated in bold.

Name and abbreviation	Longitude (°E)	Latitude (°N)	Altitude (m)
Beitem (B)	3.12	50.91	25
Zeebrugge (Z)	3.20	51.35	8
Melle (M)	3.83	50.98	15
Sint-Katelijne-Waver (S)	4.53	51.08	10
Dourbes (Do)	4.60	50.10	233
Ernage (E)	4.69	50.58	157
Retie (R)	5.03	51.22	21
Humain (H)	5.26	50.19	296
Diepenbeek (Di)	5.45	50.92	39
Buzenol (Bu)	5.59	49.62	324
Mont Rigi (M)	6.07	50.51	673

Table 5 Stations used for the screen-level verification.

RH_{2m}

RMSE	**winter**		**spring**		**summer**		**autumn**	
	0-12	**12-24**	**0-12**	**12-24**	**0-12**	**12-24**	**0-12**	**12-24**
OL	10.5	**11.3**	11.3	**12.3**	12.0	**12.9**	9.3	10.0
3D-VAR+OL	10.3	11.4	**10.5**	12.5	**11.2**	**12.9**	9.7	9.8
EKF	**10.2**	11.8	10.8	13.4	12.1	13.3	**8.6**	9.5
3D-Var+EKF	10.6	12.1	10.8	13.7	13.0	14.3	8.9	**9.4**
BIAS	**winter**		**spring**		**summer**		**autumn**	
	0-12	**12-24**	**0-12**	**12-24**	**0-12**	**12-24**	**0-12**	**12-24**
OL	-3.0	**0.7**	-2.6	**4.5**	-5.7	-1.9	-2.8	-1.0
3D-VAR+OL	-2.1	1.0	**-0.1**	5.1	**-4.2**	**-1.4**	-3.0	-1.0
EKF	-1.5	1.6	**0.1**	6.1	-4.8	-2.0	**-1.6**	**-0.2**
3D-Var+EKF	**-0.6**	1.6	3.2	6.9	-4.9	-2.6	-1.7	-0.4

Table 6 Screen-level relative humidity forecast scores in terms of RMSE and BIAS for ranges [0-12] and]12-24] for OL, EKF, 3D-VAR+OL and 3D-VAR+EKF averaged over 11 stations in Belgium (displayed in figure 6) calculated for spring, summer, autumn and winter. Lowest values are indicated in bold.

T_{2m}								
RMSE	**winter**		**spring**		**summer**		**autumn**	
	0-12	**12-24**	**0-12**	**12-24**	**0-12**	**12-24**	**0-12**	**12-24**
OL	**1.37**	**1.62**	**1.20**	**1.60**	1.47	**1.69**	**1.35**	**1.51**
3D-VAR+OL	1.42	1.64	1.30	1.61	**1.45**	1.70	1.42	1.53
EKF	1.57	1.69	1.45	1.75	1.76	1.84	1.51	1.53
3D-Var+EKF	1.70	1.74	1.57	1.76	1.83	1.89	1.63	1.56
BIAS	**winter**		**spring**		**summer**		**autumn**	
	0-12	**12-24**	**0-12**	**12-24**	**0-12**	**12-24**	**0-12**	**12-24**
OL	**-0.13**	**-0.47**	**-0.05**	**-0.50**	0.30	**-0.12**	**-0.29**	**-0.55**
3D-VAR+OL	-0.22	-0.49	-0.08	**-0.50**	0.19	-0.17	-0.42	-0.58
EKF	-0.38	-0.56	-0.44	-0.74	**0.04**	-0.25	-0.52	-0.63
3D-Var+EKF	-0.53	-0.61	-0.51	-0.73	**0.04**	-0.16	-0.67	-0.65

Table 7 Screen-level temperature forecast scores in terms of RMSE and BIAS for ranges [0-12] and]12-24] for OL, EKF, 3D-VAR+OL and 3D-VAR+EKF averaged over 11 stations in Belgium (displayed in figure 6) calculated for spring, summer, autumn and winter. Lowest values are indicated in bold.

Name and abbreviation	Latitude (°N)	Longitude (°E)	Time of sounding (UTC)
Trappes (FR)	48.77	2.00	12:00
Beauvecchain (BE)	50.76	4.77	00:00
Essen (DE)	51.40	6.96	00:00 and 12:00
Idar-Oberstein (DE)	49.70	7.33	12:00

Table 8 Soundings used for the upper-air verification

		S	A	L
winter	OL	0.266	0.233	0.117
	EKF	**0.212**	**0.220**	**0.114**
	3D-Var+OL	0.267	0.278	0.121
	3D-Var+EKF	0.250	0.267	0.122
spring	OL	0.235	0.347	**0.099**
	EKF	**0.226**	**0.343**	0.107
	3D-Var+OL	0.326	0.449	0.114
	3D-Var+EKF	0.310	0.463	0.121
summer	OL	0.407	0.153	0.203
	EKF	**0.363**	**0.097**	0.195
	3D-Var+OL	0.427	0.232	**0.189**
	3D-Var+EKF	0.398	0.201	0.197
autumn	OL	0.125	0.130	0.092
	EKF	**0.118**	0.127	**0.090**
	3D-Var+OL	0.126	0.095	0.096
	3D-Var+EKF	0.132	**0.087**	0.094

Table 9 Mean precipitation forecast scores per season for 2013 expressed in terms of S(structure)A(amplitude)L(location) scores against radar observations for the different runs listed in table 1, calculated with 24h-accumulations from forecast ranges 6h to 30h. The scores in bold indicate the best scores for each season and parameter.