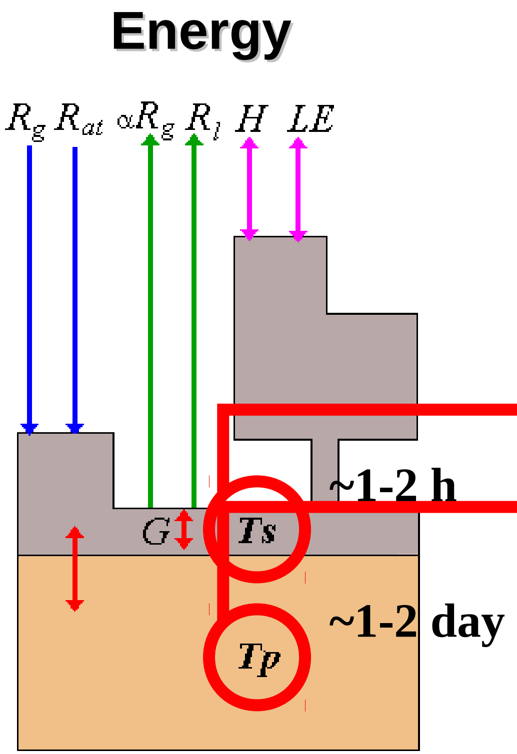
Surface data assimilation at RMI

Rafiq Hamdi, Annelies Duerinckx

The **ISBA-2L** scheme evolves 4 prognostic variables. (Giard and Bazile 2000)



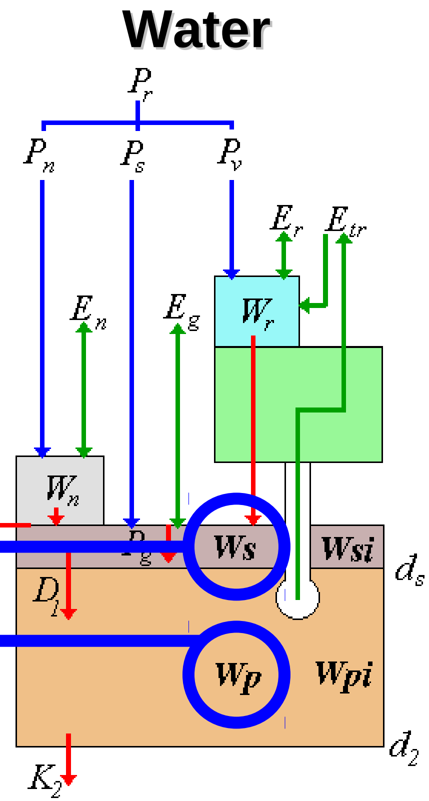
Analysis {
- surface temperature
- mean soil temperature
- superficial soil water content
- total soil water content

$$\frac{\partial T_s}{\partial t} = C_T(R_n - H - LE) - \frac{2\pi}{\tau}(T_s - T_2)$$

$$\frac{\partial T_2}{\partial t} = \frac{1}{\tau}(T_s - T_2)$$

$$\frac{\partial w_g}{\partial t} = \frac{C_1}{\rho_w d_1}(P_g - E_g) - \frac{C_2}{\tau}(w_g - w_{geq})$$

$$\frac{\partial w_2}{\partial t} = \frac{1}{\rho_w d_2}(P_g - E_g - E_{tr}) - \frac{C_3}{\tau}\max\left[0., (w_2 - w_{fc})\right]$$



In this model the ground is conceptually split into two layers, a relatively thin layer near the top with uniform temperature, and a deep soil layer also of uniform but different temperature. The net result is that flux from the deep soil layer tends to restore the top layer, opposing any radiative forcing from the atmosphere. This method is an alternative to a multilayer soil model, which is computationally more expensive.

1) **Optimum Interpolation of T_{2m} and RH_{2m} using 2m observations interpolated at the model grid-point by a 2m analysis (2-D CANARI OI)**

$$\Delta T_{2m} = T_{2m}{}^{a} - T_{2m}{}^{b} \qquad \Delta RH_{2m} = RH_{2m}{}^{a} - RH_{2m}{}^{b}$$

2) **Correction of 4 surface parameters (T_s, T_p, W_s, W_p) using 2m increments between analysed and forecasted values.**

$$T_p{}^{a} - T_p{}^{b} = \Delta T_{2m} / 2\pi \qquad T_s{}^{a} - T_s{}^{b} = \Delta T_{2m}$$

$$W_s{}^{a} - W_s{}^{b} = \alpha_{WsT} \Delta T_{2m} + \alpha_{WsRH} \Delta RH_{2m}$$

$$W_p{}^{a} - W_p{}^{b} = \alpha_{WpT} \Delta T_{2m} + \alpha_{WpRH} \Delta RH_{2m}$$

OI coefficients

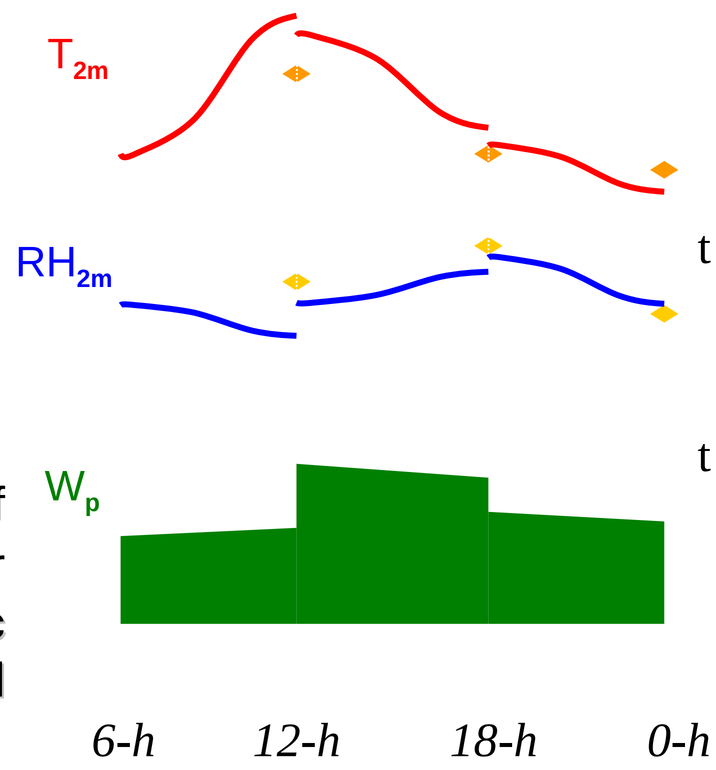
$$\alpha_{Ws/pT} = \frac{\sigma^b_{Ws/p}}{\Phi \sigma^b_{T2m}} \left\{ \left[1 + \left(\frac{\sigma^a_{RH2m}}{\sigma^b_{RH2m}} \right)^2 \right] \rho_{T2m,Ws/p} - \rho_{T2m,RH2m} \rho_{RH2m,Ws/p} \right\}$$

$$\alpha_{Ws/pRH} = \frac{\sigma^b_{Ws/p}}{\Phi \sigma^b_{RH2m}} \left\{ \left[1 + \left(\frac{\sigma^a_{T2m}}{\sigma^b_{T2m}} \right)^2 \right] \rho_{RM2m,Ws/p} - \rho_{T2m,RH2m} \rho_{T2m,Ws/p} \right\}$$

$$\Phi = \left[1 + \left(\frac{\sigma^a_{T2m}}{\sigma^b_{T2m}} \right)^2 \right] \left[1 + \left(\frac{\sigma^a_{RH2m}}{\sigma^b_{RH2m}} \right)^2 \right] - \rho^2_{T2m,RH2m}$$

- Very strong dependency of these background error statistics to **physiographic** properties and **meteorological conditions.**

Sequential analysis (every 6h)



EKF for soil analysis

Background error covariance matrix

Observation error covariance matrix

$$\mathbf{x}_t^a = \mathbf{x}_t^b + \underbrace{\mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}}_{\text{Kalman gain (weight)}}\underbrace{[\mathbf{y}_t^o - \mathcal{H}(\mathbf{x}_o^b)]}_{\text{Departure (error)}}$$

$\mathbf{y}_t^o$ Observations (T2m, RH2m)

$\mathbf{x}_o^b$ Model variables (Wg, W2, Ts, T2)

$\mathcal{H}(\mathbf{x}_o^b)$ Model counterpart of Observations (T2m, RH2m)

	Parameter	Value
$R-matrix$	T_{2m}	1 K
	RH_{2m}	10 %
$B-matrix$	T_s	2 K
	T_2	2 K
	W_g	$0.1(w_{fc} - W_{wilt})$
	W_2	$0.1(w_{fc} - W_{wilt})$

EKF: The Jacobian

$$\mathbf{x}_t^a = \mathbf{x}_t^b + \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}[\mathbf{y}_t^o - \mathcal{H}(\mathbf{x}_o^b)]$$

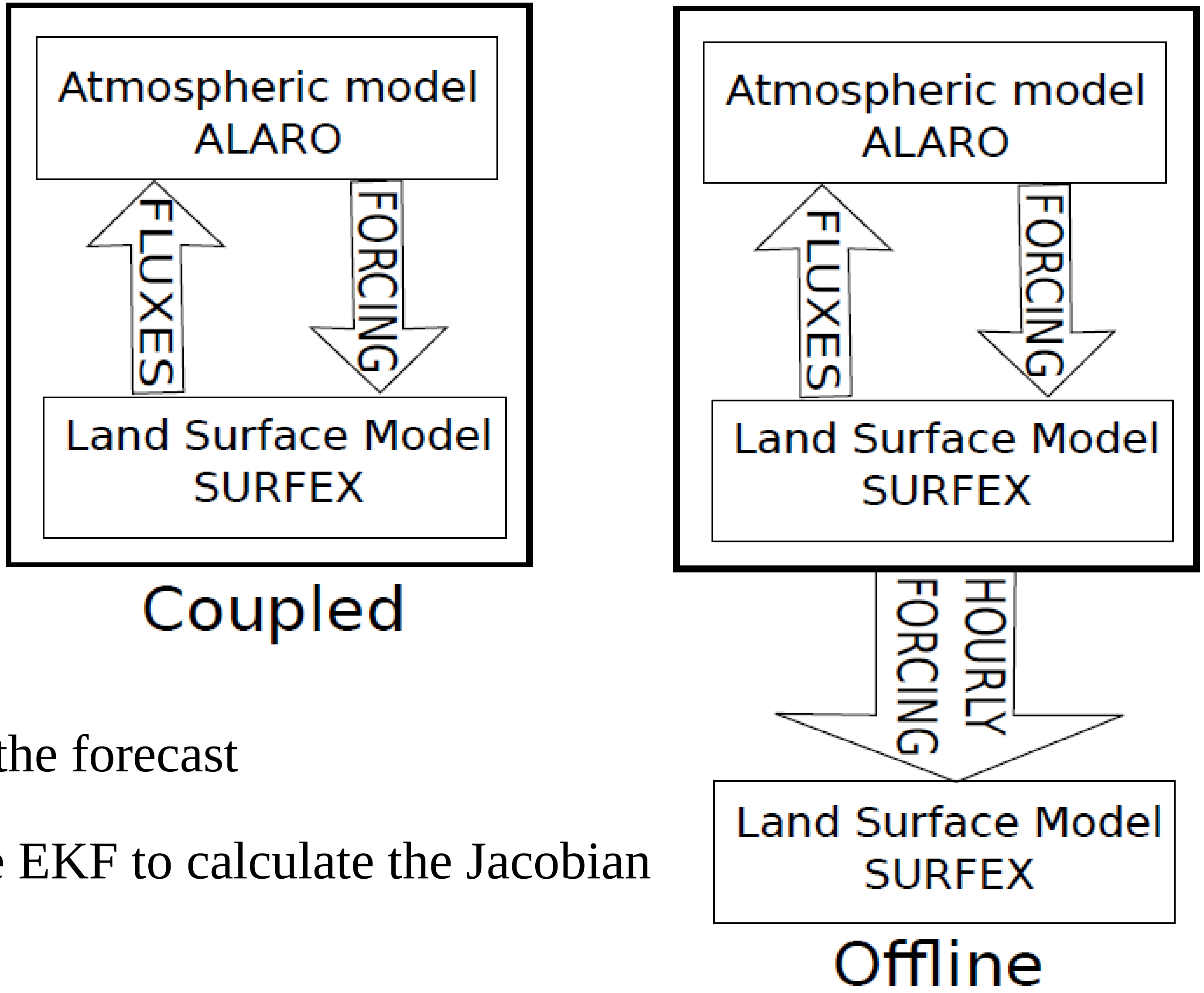
- $\mathcal{H}$: observation operator

  includes a model propagation
- **H**: Jacobian of the observation operator

  Calculated with finite differences

$$H_{i,j} = \frac{\delta y_{i,t}}{\delta x_{j,t0}} = \frac{y_i(x + \delta x_j) - y_i(x)}{\delta x_j}$$

EKF: The Jacobian

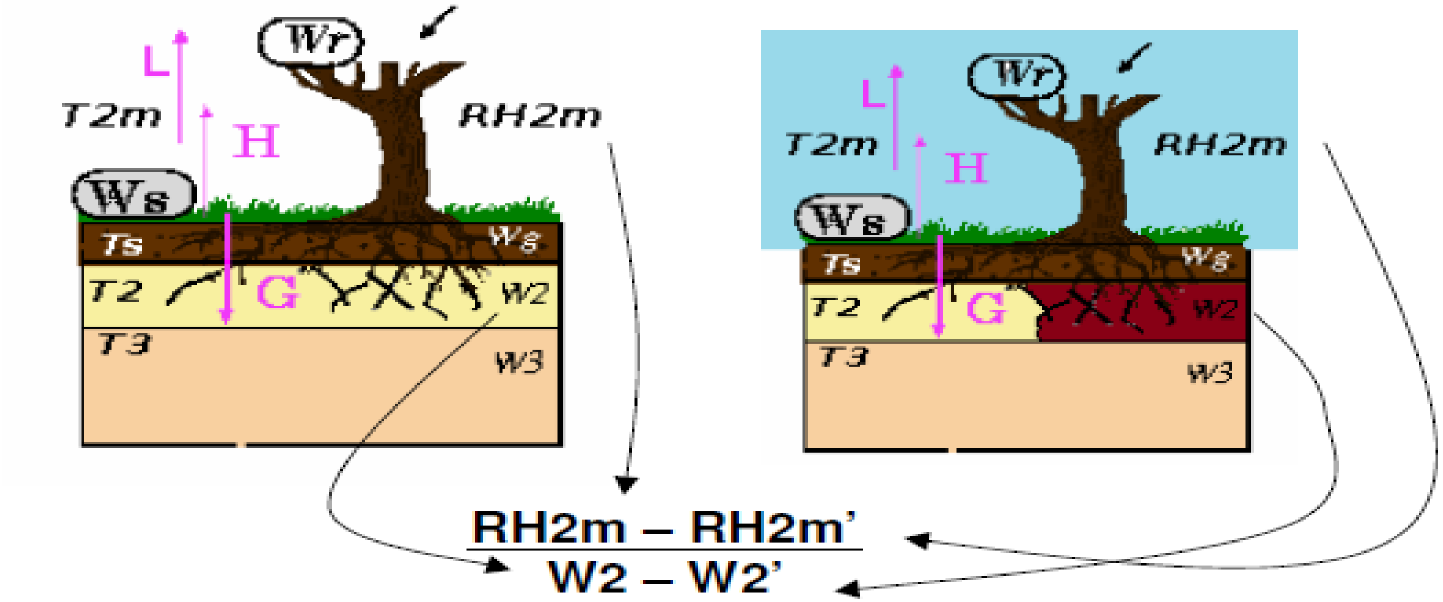


Coupled : used for the forecast

Offline : used in the EKF to calculate the Jacobian

EKF: The Jacobian

- **dRH2m/dW2**: how big is the change in RH2m if we introduce a small change in W2?

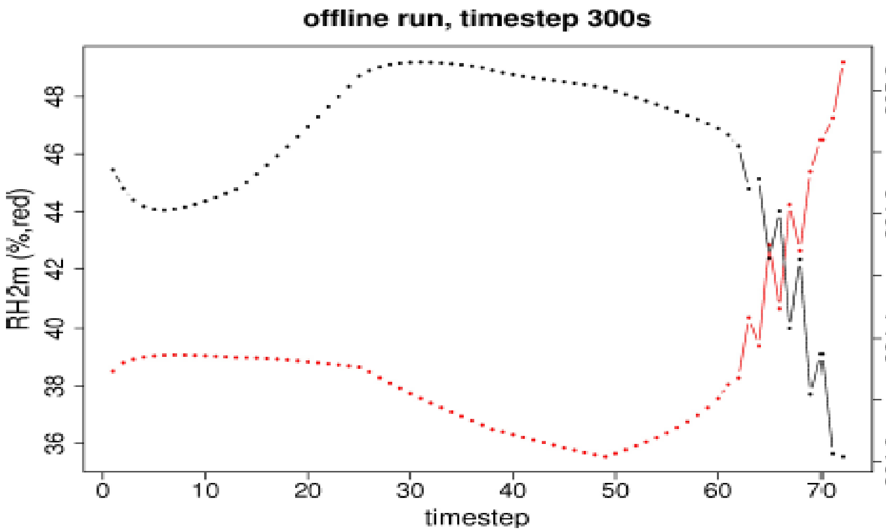


Calculation of the Jacobian requires 4 perturbed runs

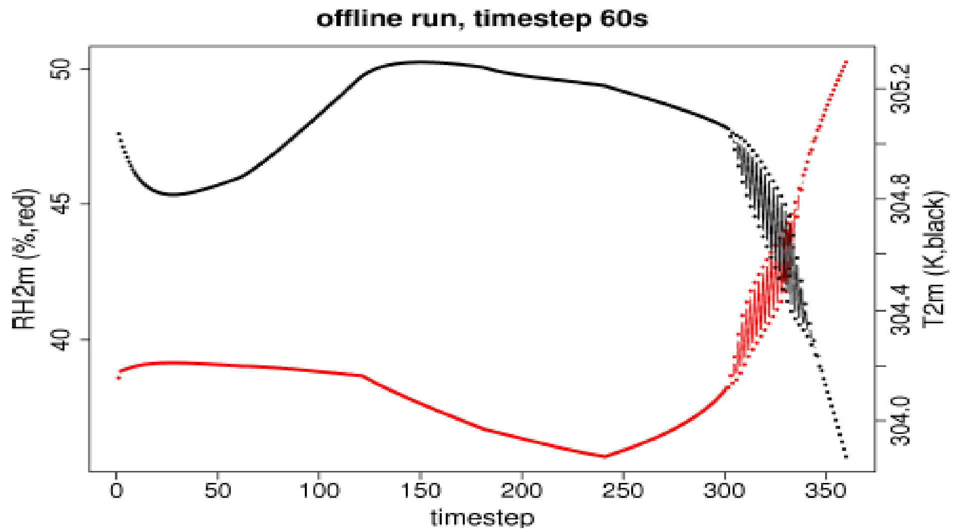
_ In offline or coupled mode

_ offline mode is computationally cheaper and allows smaller perturbation sizes

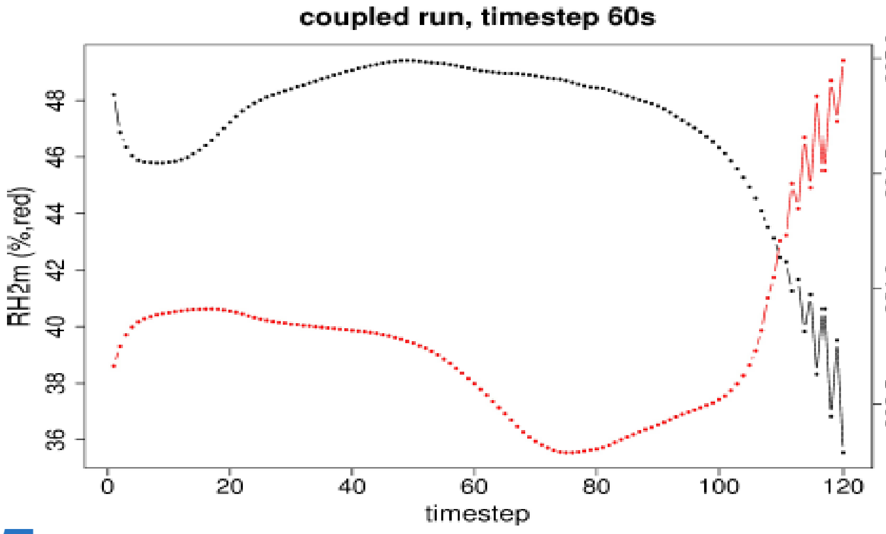
Oscillations

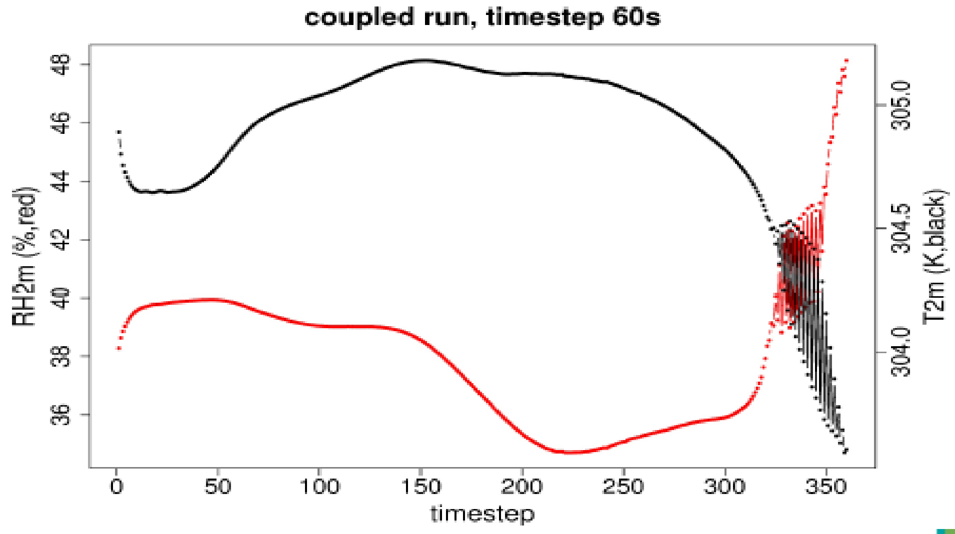


(a)



(b)





Oscillations

- In coupled and offline runs
- Between 12UTC and 18UTC when RI>0, when stable boundary layer starts to form
- In all surface variables related to fluxes
- $2\Delta t$ oscillations that remain when timestep is decreased from 300s to 60s : artificial !
- No fibrillations, but caused by small inconsistencies in turbulent fluxes between the surface and the atmosphere

	Offline		Coupled	
	End	Total	End	Total
$\frac{\delta RH}{\delta W_{g2}}$	4.8 %	24 %	11 %	53 %
$\frac{\delta T}{\delta W_{g2}}$	5.2 %	21 %	13 %	55 %
$\frac{\delta RH}{\delta T_{g2}}$	2.4 %	21 %	11 %	66 %
$\frac{\delta T}{\delta T_{g2}}$	3.6 %	10 %	11 %	57 %

Geosci. Model Dev., 8, 845–863, 2015
www.geosci-model-dev.net/8/845/2015/
doi:10.5194/gmd-8-845-2015

Geoscientific Model Development
Open Access

Study of the Jacobian of an extended Kalman filter for soil analysis in SURFEXv5

A. Duerinckx[1,2], R. Hamdi[1,2], J.-F. Mahfouf[3], and P. Termonia[1,2]

[1]Department of Physics and Astronomy, Ghent University, Ghent, Belgium
[2]Royal Meteorological Institute, Ringlaan 3, 1180 Brussels, Belgium
[3]GAME, CNRM, Météo-France, CNRS, Toulouse, France

Correspondence to: A. Duerinckx (annelies.duerinckx@meteo.be)

Received: 5 September 2014 – Published in Geosci. Model Dev. Discuss.: 28 October 2014
Revised: 9 February 2015 – Accepted: 3 March 2015 – Published: 27 March 2015

Abstract. An externalised surface scheme like SURFEX allows computationally cheap offline runs. This is a major advantage for surface assimilation techniques such as the extended Kalman filter (EKF), where the offline runs allow a cheaper numerical estimation of the observation operator Jacobian. In the recent past an EKF has been developed within SURFEX for the initialisation of soil water content and soil temperature based on screen-level temperature and relative humidity observations. In this paper we make a comparison of the Jacobian calculated with offline SURFEX runs and with runs coupled to the atmospheric ALARO model. Comparisons are made with respect to spatial structure and average value of the Jacobian, gain values and increments. We determine the optimal perturbation size of the Jacobian for the offline and coupled approaches and compare the linearity of the Jacobian for these cases. Results show that the offline Jacobian approach gives similar results to the coupled approach and that it allows for smaller perturbation sizes that better approximate this linearity assumption. We document a new case of non-linearities that can hamper this linearity assumption and cause spurious $2\Delta t$ oscillations in small parts of the domain for the coupled as well as offline runs. While these oscillations do not have a detrimental effect on the model run, they can introduce some noise in the Jacobian at the affected locations. The oscillations influence both the surface fluxes and the screen-level variables. The oscillations occur in the late afternoon in summer when a stable boundary layer starts to form near the surface. We propose a filter to remove the oscillations and show that this filter works accordingly.

1 Introduction

Externalising surface schemes from upper-air atmospheric models has many advantages. If the interface between the different parts is defined in a flexible manner (see Best et al., 2004, for an example), then it provides the possibility to plug one scheme into different models, even targeting different applications, ranging from climate to high-impact weather. Another major advantage is that the scheme can also be used in an offline mode, allowing for cheap solutions in specific applications. An example of this is studied in the present paper: the implementation of an extended Kalman filter (EKF) for surface assimilation (Mahfouf et al., 2009), where cheap offline runs with the SURFEX external land surface model (Masson et al., 2013; Hamdi et al., 2014a) allow one to numerically estimate the observation operator Jacobian.

Surface assimilation techniques, like this EKF, can improve the boundary layer forecasts of a numerical weather prediction (NWP) model considerably (Douville et al., 2000; Hess, 2001; Drusch and Viterbo, 2007). The surface serves as a lower boundary condition for the NWP model and has an important impact on the lower atmosphere. Land surface models (LSMs) determine the partitioning of the energy into latent and sensible heat fluxes (e.g. by means of evapotranspiration processes) and these fluxes provide the main link between the surface and the atmosphere. In the past two decades LSMs have been improved considerably. Still, there are a lot of uncertainties and errors in model parameterisations, model resolution and observation measurements of soil variables. In order to provide an optimal initial surface state for an NWP forecast, the assimilation of surface observations into the land surface model is necessary. The amount and fre-

Published by Copernicus Publications on behalf of the European Geosciences Union.

Small inconsistencies between the surface and the atmosphere cause oscillations in coupled and offline SURFEX runs when RI>0

No problem for model runs, but important effect on Jacobian values of the EKF in certain locations

Solutions :

_ Use a filter to remove oscillations

_ Use a lagged atmosphere

Results :

_ Oscillations dissapear with solutions

_ Better forecast scores for lagged atmosphere

Combining the EKF soil analysis with a three dimensional variational upper-air assimilation for ALARO

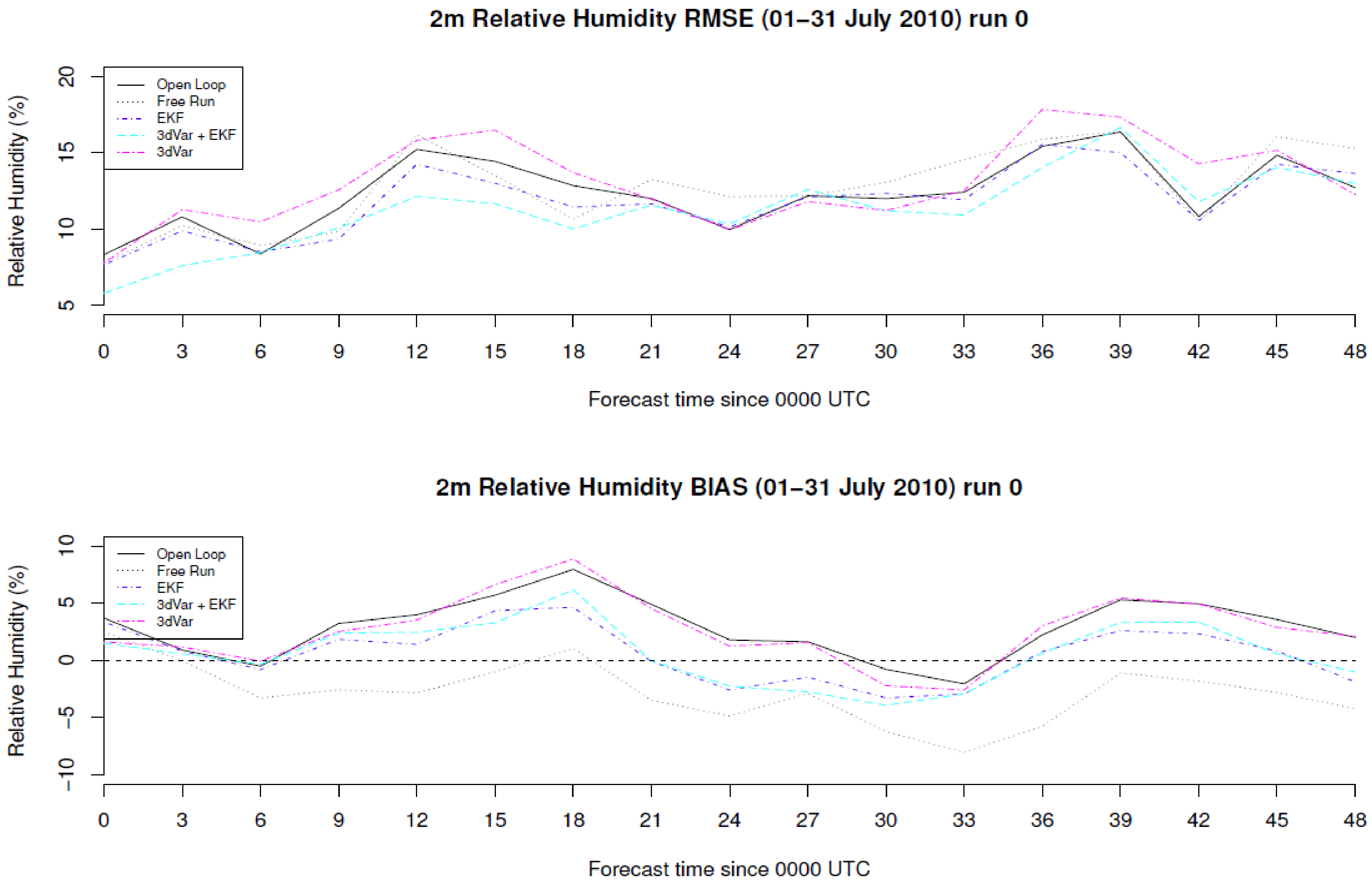


Figure 3.12: Root-mean-square error and BIAS for relative humidity at Uccle averaged over the month July 2010 for extended Kalman filter (EKF), open loop, free run, 3D-Var, 3D-Var+EKF.

STAEKF

The forecast model, is augmented with P model parameters:

$$\mathbf{z}^f = \begin{bmatrix} \mathbf{x}^f \\ \lambda^f \end{bmatrix} = \mathcal{F}\mathbf{z}^a = \begin{bmatrix} \mathcal{M}\mathbf{x}^a \\ \mathcal{F}^\lambda \lambda^a \end{bmatrix}, \tag{1}$$

$\mathbf{z} = (\mathrm{x}, \lambda)$ is the augmented state vector. The augmented dynamical system $\mathcal{F}$ includes the dynamical model for the system's state, $\mathcal{M}$, and a dynamical model for the parameters $\mathcal{F}^\lambda$. In the absence of additional information, a persistence model for $\mathcal{F}^\lambda$ is often assumed so that $\mathcal{F}^\lambda = \mathrm{I}$ and $\lambda^f_{t_{k+1}} = \lambda^a_{t_k}$; the same choice has been adopted here.

STAEKF

The forecast/analysis error covariance matrix, $\mathbf{P}_z^{f,a}$, for the augmented system reads:

$$\mathbf{P}_z^{f,a} = \begin{pmatrix} \mathbf{P}_x^{f,a} & \mathbf{P}_{x\lambda}^{f,a} \\ \mathbf{P}_{x\lambda}^{f,a\,T} & \mathbf{P}_\lambda^{f,a} \end{pmatrix} \quad (2)$$

where the $I \times I$ matrix $\mathbf{P}_x^{f,a}$ is the error covariance of the state estimate $\mathbf{x}^{f,a}$, $\mathbf{P}_\lambda^{f,a}$ is the $P \times P$ parametric error covariance and $\mathbf{P}_{x\lambda}^{f,a}$ the $I \times P$ error correlation matrix between the state vector, $\mathbf{x}$, and the vector of parameters λ. These correlations are essential for the estimation of the parameters. In general one does not have access to a direct measurement of the parameters, and information are only obtained through observations of the system's state.

STAEKF

The forecast error propagation in the STAEKF is given by $\mathbf{P}_z^f = \mathbf{C}\mathbf{P}_z^a\mathbf{C}^T$, with $\mathbf{C}$ being the STAEKF forward operator defined as:

The short-time truncation of the dynamics

$$\mathbf{C} = \begin{pmatrix} \mathbf{M} & \frac{\partial g}{\partial \lambda}|_{\lambda^a}\tau \\ 0 & \mathbf{I}_P \end{pmatrix} \tag{3}$$

where $\mathbf{I}_P$ is the $P \times P$ identity matrix. Equation (3) embeds the key feature of the STAEKF; the presence of the term $\frac{\partial g}{\partial \lambda}|_{\lambda^a}\tau$ allows for accounting for the contribution of the parametric error to the forecast error as well as to the error correlation between model state and parameters.

ATMOSPHERIC SCIENCE LETTERS
Atmos. Sci. Let. (2012)
Published online in Wiley Online Library
(wileyonlinelibrary.com) **DOI:** 10.1002/asl.394

RMetS
Royal Meteorological Society

Short time augmented extended Kalman filter for soil analysis: a feasibility study

Alberto Carrassi,[1,2]* Rafiq Hamdi,[1] Piet Termonia[1,3] and Stéphane Vannitsem[1]
[1] *Institut Royal Météorologique de Belgique, Bruxelles, Belgique*
[2] *Institut Catalá de Ciéncies del Clima (IC3), Barcelona, Spain*
[3] *Department of Physics and Astronomy, Ghent University, Ghent, Belgium*

*Correspondence to:
A. Carrassi, Institut Catalá de
Ciéncies del Clima, Carrer Doctor

Abstract

This paper presents a soil analysis scheme based on an extended Kalman filter (EKF), the

Twin experiment study for Albedo and LAI

Accepted: 24 May 2012

Keywords: **data assimilation; parameter estimation; soil analysis**

1. Introduction

Current soil moisture analysis schemes in most numerical weather prediction (NWP) centers are based on analyzed or observed screen-level variables. An optimal interpolation technique (OI) developed by Mahfouf (1991) on the basis of the short-range forecasts of 2 m temperature (T_{2m}) and relative humidity (RH_{2m}) is operational in a number of centers: Météo-France (Giard and Bazile, 2000), Environment Canada (Bélair *et al.*, 2003), and the European Centre for Medium-Range Weather Forecasts (Douville *et al.*, 2000) among others. However, some fundamental limitations are inherent with the use of the OI (Balsamo *et al.*, 2004; Mahfouf *et al.*, 2009), among these the fact that it does not allow for the use of nonlinear observation operator (note that the physical link between near-surface observations and soil variables is nonlinear) and that the OI coefficients are usually derived using very simple assumptions and/or model configurations.

A particularly promising approach to address the above shortcomings is the use of an extended Kalman filter (EKF). The EKF and a simplified version (SEKF), in which a static forecast error covariance matrix is adopted, have been successfully applied to the assimilation of direct and/or indirect soil observations (Draper *et al.*, 2009; Drusch *et al.*, 2009; Mahfouf *et al.*, 2009; Albergel *et al.*, 2010; Mahfouf and Bliznak, 2011).

Recently, Carrassi and Vannitsem (2011a) introduced an alternative formulation of the EKF where the uncertain model parameters are estimated along with the system state variables. The algorithm, referred to as short time augmented extended Kalman filter (STAEKF), uses a deterministic formulation for the model error dynamics (Nicolis, 2003); an approach already used for the treatment of the error arising from the unresolved scales (Carrassi and Vannitsem, 2011b) and in the context of variational assimilation (Carrassi and Vannitsem, 2010).

With the aim of contributing to the current discussion on soil data assimilation strategies, we undertake here a set of numerical 'twin' experiments with observations designed to test the STAEKF for the assimilation of screen-level observations using an off-line version of the land surface model, Interactions between surface, biosphere, and atmosphere (ISBA) (Noilhan and Mahfouf, 1996) where errors in three land surface parameters, the leaf area index (LAI), the albedo, and the minimum stomatal resistance (RS_{min}) are introduced.

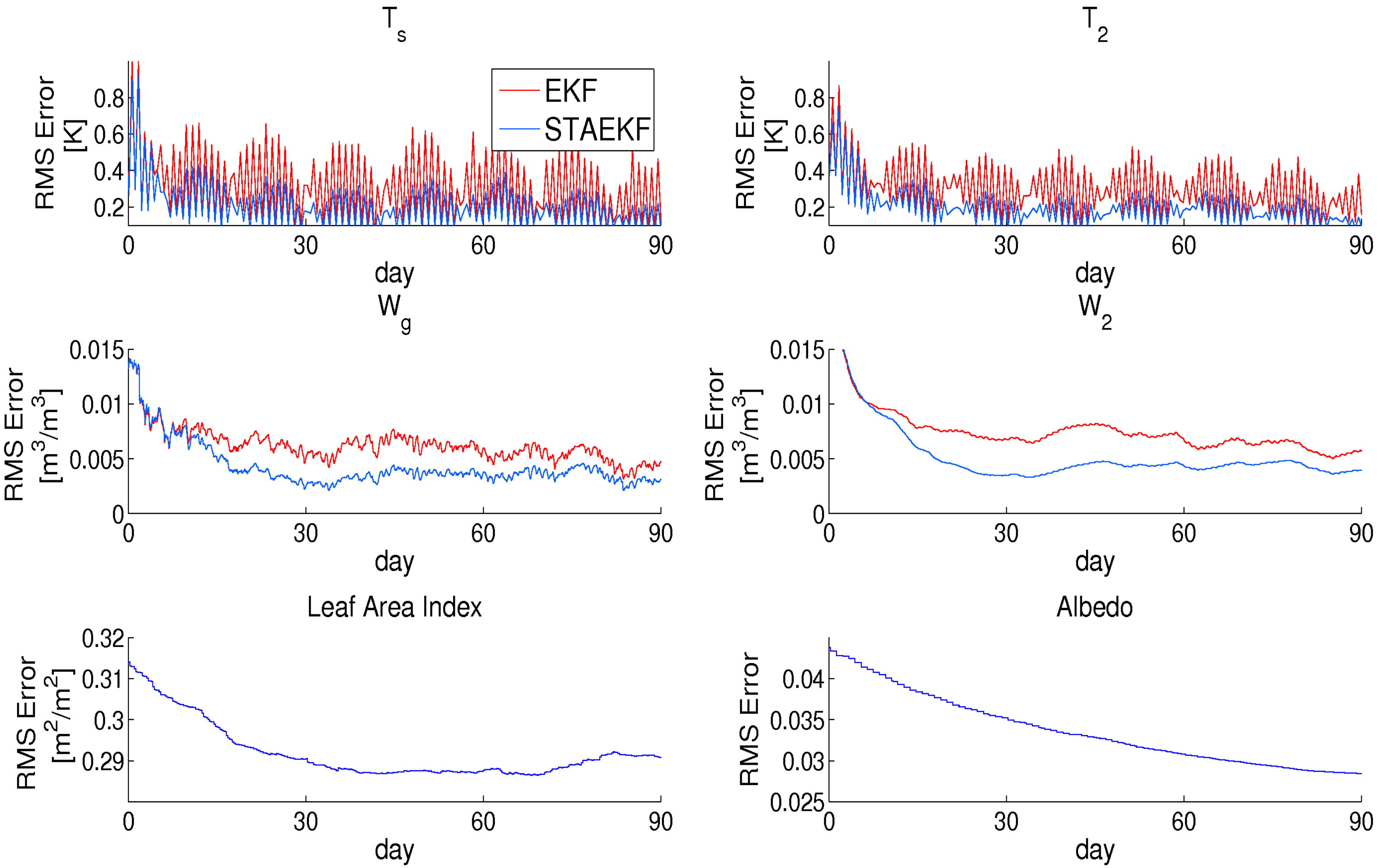
2. The land surface model

The two-layer version of the land surface model ISBA describes the evolution of soil temperature and moisture contents based on the force-restore method. The model is available within a surface externalized platform (SLDAS; Mahfouf, 2007). The equations can be formally written as a dynamical system, $\frac{d\mathbf{x}}{dt} = \mathbf{g}(\mathbf{x}, \boldsymbol{\lambda})$. The state vector, $\mathbf{x} = (T_s, T_2, w_g, w_2)$, contains the surface and deep soil temperatures T_s and T_2 and the corresponding water contents w_g and w_2. The vector $\boldsymbol{\lambda}$ is taken to represent the set of

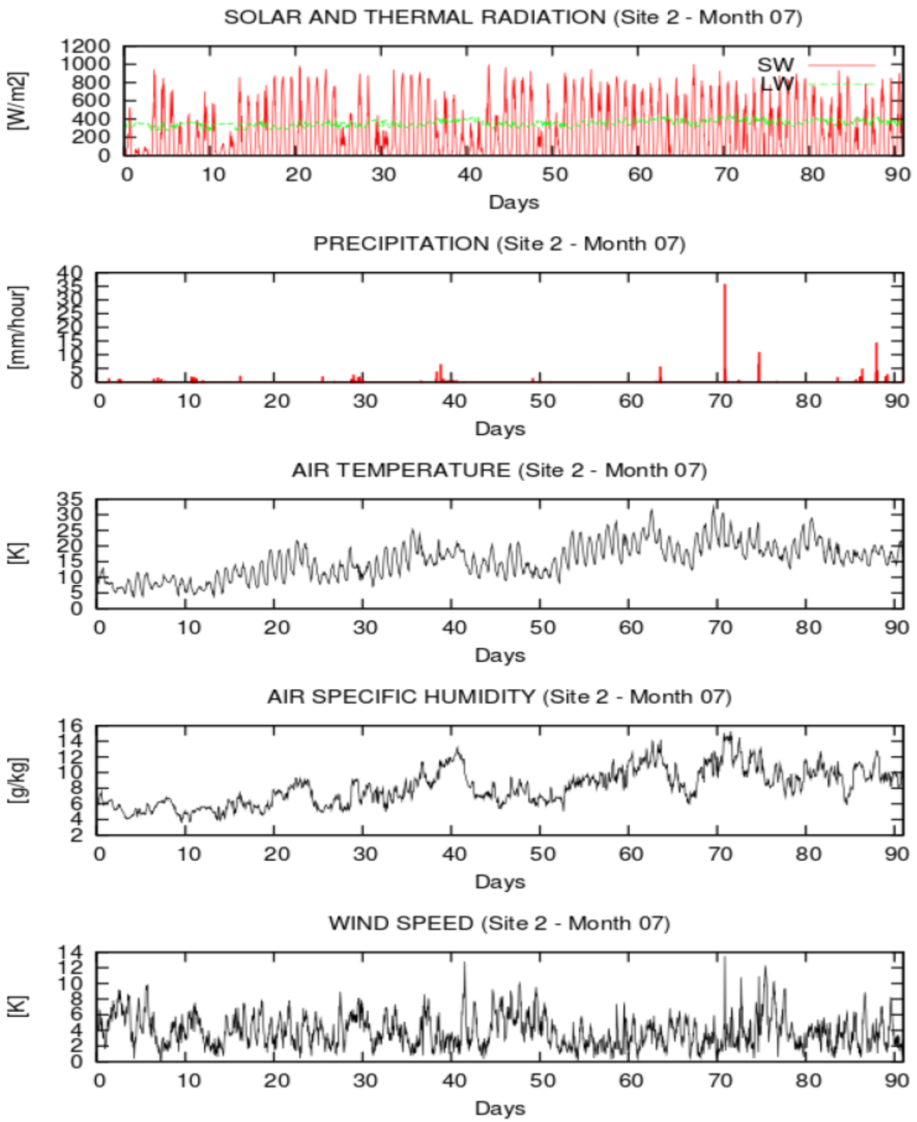
IRM depuis 100 1913

belspo.be Politique scientifique fédérale

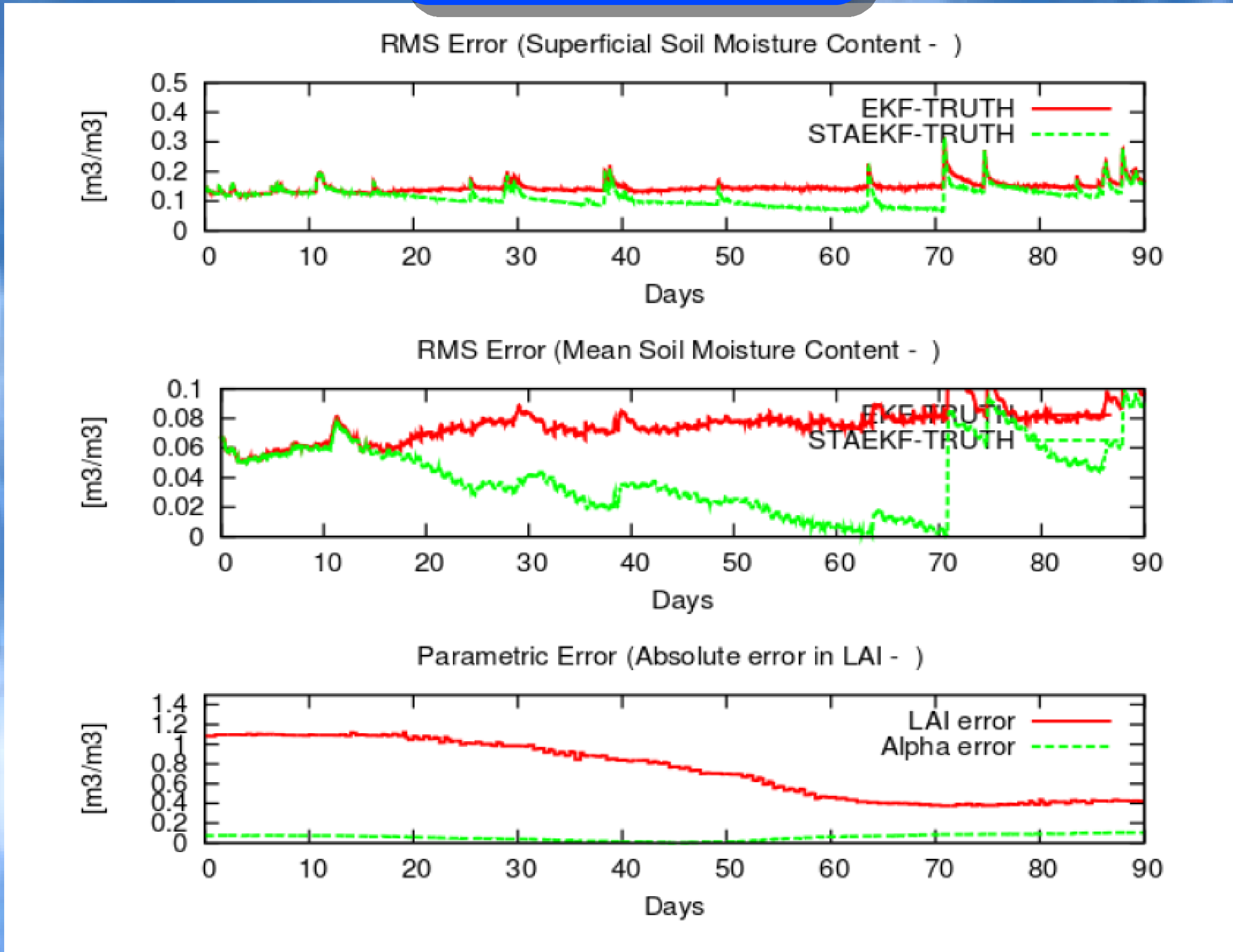
STAEKF


T_s
RMS Error [K]
EKF
STAEKF
day
T_2
W_g
RMS Error [m^3/m^3]
W_2
Leaf Area Index
RMS Error [m^2/m^2]
Albedo
RMS Error

STAEKF at Cabauw



STAEKF at Cabauw



Outlook:

1. Use of Land SAF product as forcing for SURFEX offline runs: solar downward and longwave radiation

2. Use albedo and LAI from new sensors ProbaV and combine this information with STAEKF.

3. Test STAEKF in a real enviromenet within SODA and ALARO for NWP application

4. Combine 3d-var with STAEKF for NWP application.

5. Use of STAEKF for study about the coupling between atmosphere and land surface and for seasonal forecasting.