Short Time Augmented Extended Kalman Filter for Soil Analysis. A feasibility study

Carrassi A., Hamdi R., Termonia P., Vannitsem S.

Reykjavik, 15-19th April 2013

IRM KMI

- **Recently, Carrassi and Vannitsem (2011, QJRMS) introduced an alternative formulation of the EKF where the uncertain model parameters are estimated along with the system state variables.**

- **The algorithm, Short Time Augmented Extended Kalman Filter (STAEKF), uses a deterministic formulation for the model error dynamics (Nicolis, 2003, JAS).**

- **The same formulation has been used for the treatement of the error arising from the unresolved scales (Carrassi and Vannitsem, 2011, IJBC) and in the context of variational assimilation (Carrassi and Vannitsem, 2010, MWR).**

- **We undertake here a set of numerical twin experiments designed to test the STAEKF in estimating three land surface parameters: LAI, the albedo, and the minimum stomatal resistance RSmin.**

Assimilation of 2m temperature and relative humidity using an offline version of ISBA.

- **The two-layers version of the land surface model ISBA.**

The model equations read:

$$\frac{\partial T_s}{\partial t} = C_T(R_n - H - LE) - \frac{2\pi}{\tau}(T_s - T_2)$$
$$\frac{\partial T_2}{\partial t} = \frac{1}{\tau}(T_s - T_2)$$
$$\frac{\partial w_g}{\partial t} = \frac{C_1}{\rho_w d_1}(P_g - E_g) - \frac{C_2}{\tau}(w_g - w_{geq})$$
$$\frac{\partial w_2}{\partial t} = \frac{1}{\rho_w d_2}(P_g - E_g - E_{tr}) - \frac{C_3}{d_2 \tau} max[0., (w_2 - w_{fc})]$$

- **The model is available within a surface externalized platform (SLDAS, Mahfouf 2007).**

- **The state vector,** $X = (T_s, T_2, W_g, W_2)$ **and the equation can be formally written as a dynamical system,** $\frac{dx}{dt} = g(X, \lambda)$

- **The vector λ is taken to represent the set of model parameters.**

The forecast model, is augmented with P model parameters:

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

$\mathbf{z} = (\mathbf{x}, \lambda)$ 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.

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.

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

$$\mathbf{C} = \begin{pmatrix} \mathbf{M} & \frac{\partial g}{\partial \lambda}|_{\lambda^a}\mathcal{T} \\ 0 & \mathbf{I}_P \end{pmatrix} \quad (3)$$

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 $\frac{\partial g}{\partial \lambda}|_{\lambda^a}\mathcal{T}$ 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.

An augmented observation operator is introduced, $\mathcal{H}_z = [\mathcal{H} \quad 0]$ with $\mathcal{H}$ as for the standard EKF. Its linearization, $\mathbf{H}_z$ is now a $M \times (I+P)$ matrix in which the last P columns contain zeros. The augmented state and covariance update complete the algorithm and are equivalent to those of the EKF except that they refer now to the augmented system, and the gain matrix has dimension $(I+P) \times M$ (see Carrassi and Vannitsem, 2011a, for details).

Experimental Set-up

- **Observation system simulation experiments (OSSE).**

- **The forcing consist of 1-hourly air temperature, specific humidity, atmospheric pressure, incoming global radiation, incoming long-wave radiation, precipitation rate and wind speed relative to the ten summers in the decade 1990-1999 extract from ECMWF Re-analysis ERA40.**

- **ISBA is run in one offline single column mode for a 90 day period.**

- **The simulated observations are T2m and RH2m at 00, 06, 12 and 18 UTC.**

- **The initial Pf (B) and Pm (Q) required by the EKF, are taken from Mahfouf (2007).**
  diag(R)=(1,10-2) diag(Pf)=(1,1,10-2,10-2) diag(Pm)=(25.10-2,25.10-2,4.10-4,4.10-4)

- **Parametric errors are introduced by perturbing either alternatively or simultaneously, the Leaf ea Index, LAI , the albedo, and the minimum stomatal resistance.**

IRM KMI

Experimental Set-up

- **For each summer in the period 1990 - 1999, a reference trajectory is generated by integrating the model with LAI = 1 m2 /m2, albedo = 0.2, and Rsmin=94 s/m.**

- **Around each of these trajectories, Gaussian samples of 100 initial conditions and uncertain parameters are used to initialize the assimilation cycles.**

- **The initial conditions are sampled from a distribution with standard deviation:**

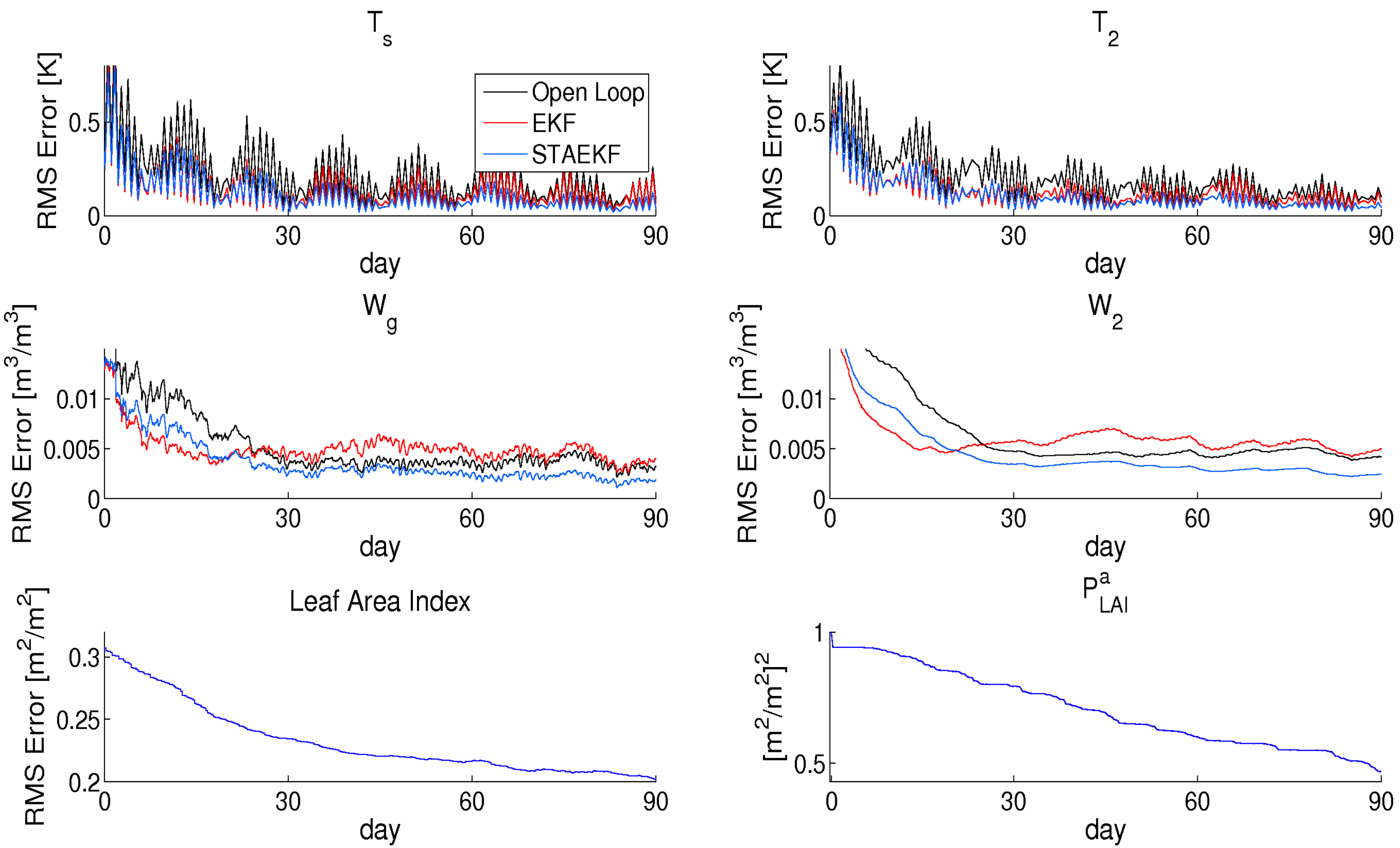
  (σTs , σT2 , σwg , σw2) = (5, 5, 1, 1)

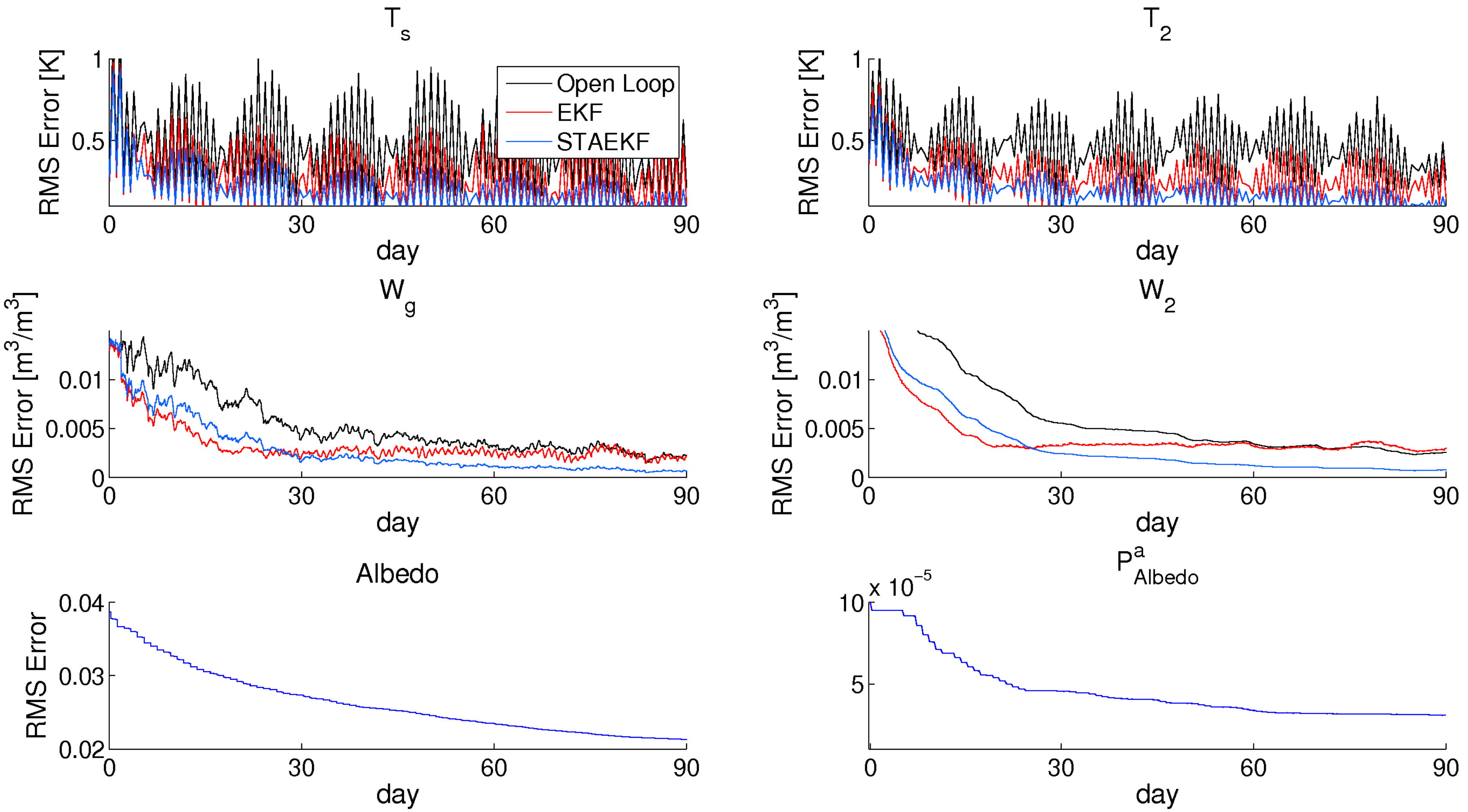
- **LAI, albedo, and RSmin are sampled with standard deviations:**

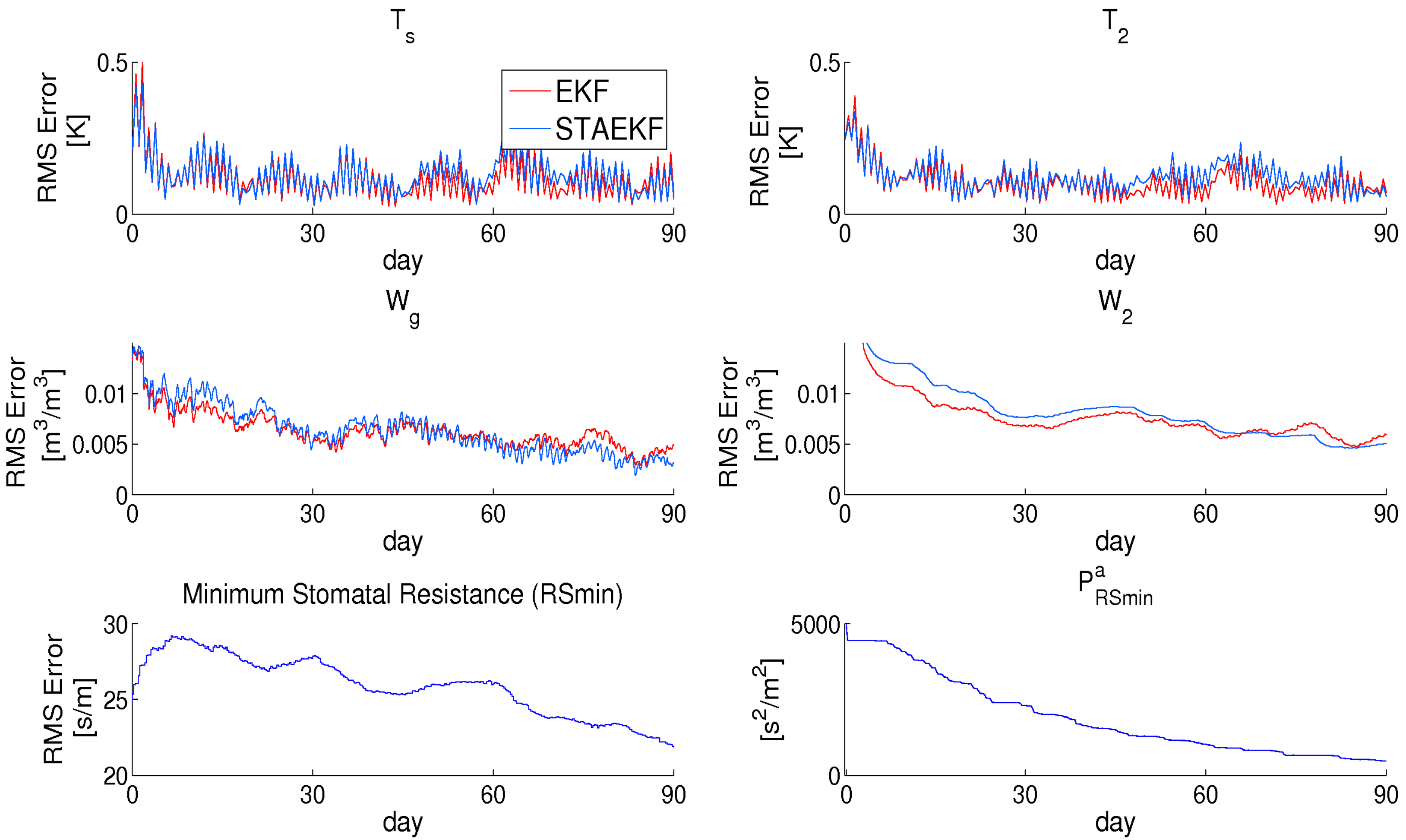
  σ LAI = 0.5 σalbedo = 0.1 σRSmin = 50

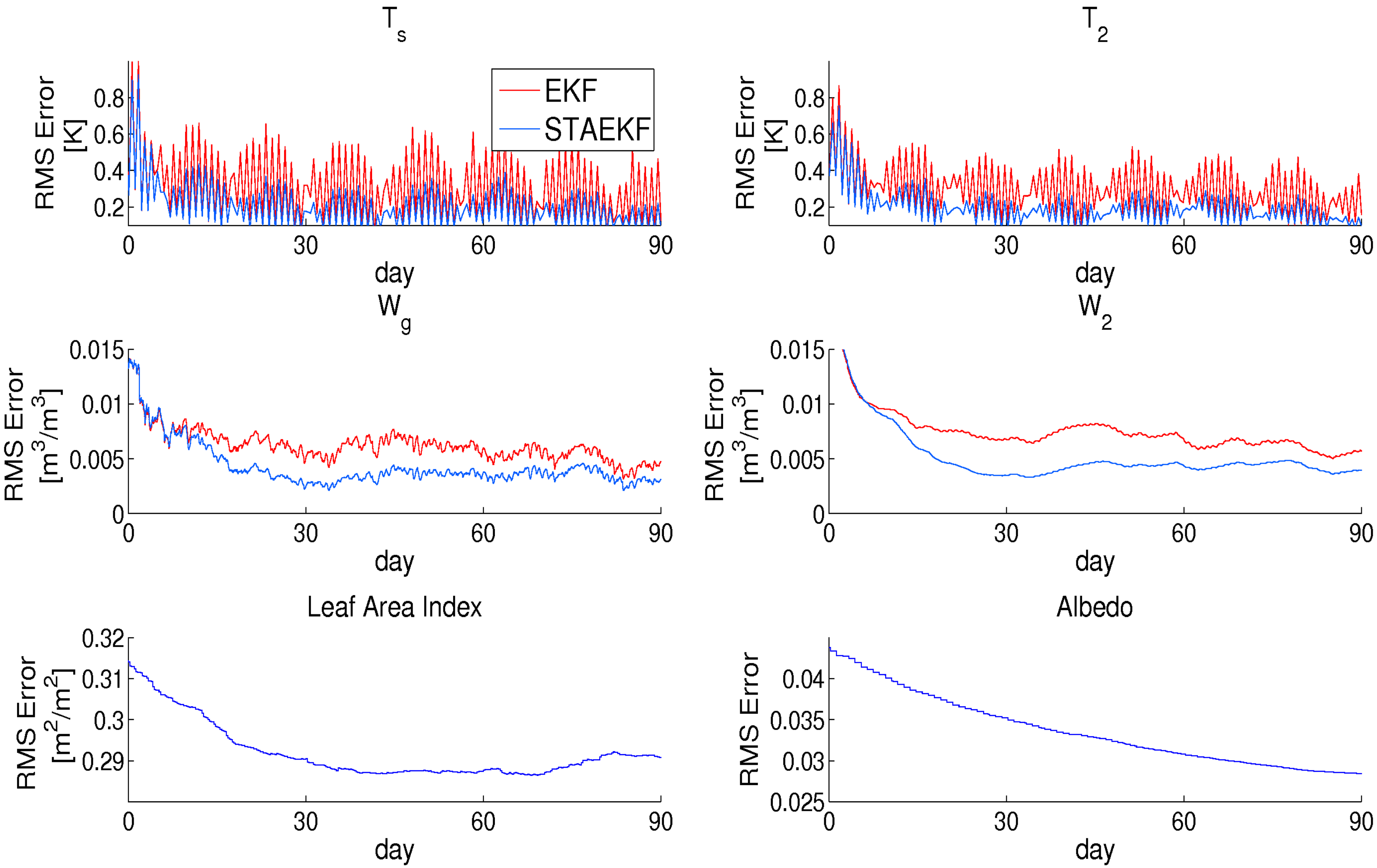
- **PLAI=1 Palbedo=10-4 PRSmin=5000 in the STAEKF, while Pax is taken as in the EKF; Px,λ is initially set to zero.**

IRM KMI


T_s
T_2
Open Loop
EKF
STAEKF
RMS Error [K]
0.5
0
30
60
90
day
W_g
W_2
RMS Error [m^3/m^3]
0.01
0.005
Leaf Area Index
RMS Error [m^2/m^2]
0.3
0.25
0.2
P^a_LAI
[m^2/m^2]^2
1


T
s
RMS Error [K]
Open Loop
EKF
STAEKF
day
T
2
W
g
RMS Error [m3/m3]
W
2
Albedo
RMS Error
P
a
Albedo
x 10−5
0
30
60
90
1
0.5
0.01
0.005
0.04
0.03
0.02
10
5


T s
RMS Error [K]
EKF
STAEKF
T 2
W g
RMS Error [m3/m3]
W 2
Minimum Stomatal Resistance (RSmin)
RMS Error [s/m]
P a RSmin
[s2/m2]
day


T_s
T_2
EKF
STAEKF
RMS Error [K]
W_g
W_2
RMS Error $[m^3/m^3]$
Leaf Area Index
RMS Error $[m^2/m^2]$
Albedo
RMS Error
day
0
30
60
90

- **Continuing the evaluation of the STAEKF with data from the Cabauw tower (work of Annelies's PhD).**

- **A scientific paper is published: Short Time Augmented Extended Kalman Filter for Soil Analysis: A feasibility study.** *Carrassi A., Hamdi R., Termonia P., and Vannitsem S., DOI: 10.1002/asl.394, Vol. 13, p. 268-274, 2012.*

- **The STAEKF is able to reduce the parameter estimation errors.**

- **The accuracy of these estimates is inherently related to the type of parameter to be estimated.**

- **The rate of error convergence in the STAEKF is related to the initial parametric error variance.**

- **Implementing the STAEKF in SURFEX and study its behaviour within ALARO.**