Laboratory of Hydrology and Water Management
Gabrielle.DeLannoy@UGent.be

Paper number H23E-1554
AGU - 12 December 2006

Ensemble Kalman filtering of soil moisture observations with model bias correction

Gabriëlle De Lannoy[1], Rolf Reichle[2], Paul Houser[3], Valentijn Pauwels[1], Niko Verhoest[1]

1, Laboratory of Hydrology and Water Management, Ghent University, Coupure links 653, B-9000 Ghent, Belgium
2, Global Modeling and Assimilation Office (Code 610.1), NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
3, Goddard Earth Sciences and Technology Center, University of Maryland, Baltimore County, Baltimore, MD 21250, USA
4, George Mason University & Center for Research on Environment and Water, 4041 Powder Mill Road, Suite 302, Calverton, MD 20705-3106, USA

1. Introduction

Land surface models are usually biased in at least a subset of the simulated variables even after calibration. Bias estimation may therefore be needed for data assimilation. Here, in situ soil moisture observations in a small agricultural field were merged with Community Land Model (CLM2.0) simulations using different algorithms for state and bias estimation with and without bias correction feedback.

2. State and bias estimation

If the error in the model formulation is biased (non-zero mean), then the truth and the model forecasts may be expressed as follows:

$$\mathbf{b}_i = \mathbf{g}_{i,i-1}(\mathbf{b}_{i-1}, \mathbf{x}_{i-1})$$
$$\mathbf{x}_i = \mathbf{f}_{i,i-1}(\mathbf{x}_{i-1}, \mathbf{u}_i) + \mathbf{w}_{i-1} + \mathbf{b}_i$$
$$\hat{\tilde{\mathbf{x}}}_i^- = \mathbf{f}_{i,i-1}(\hat{\tilde{\mathbf{x}}}_{i-1}, \mathbf{u}_i)$$

In this case the model $\mathbf{f}_{i,i-1}$ generates a biased a priori state estimate $\hat{\tilde{\mathbf{x}}}_i^-$, i.e. soil moisture and temperature at 10 soil layers and vegetation water and temperature. The meteorological forcings are given by $\mathbf{u}_i$. The random noise is $\mathbf{w}_{i-1}$. The bias vector is given by $\mathbf{b}_i$ and propagated by the bias model $\mathbf{g}_{i,i-1}$. For a linear model, the mean a priori estimation error of the forecast is now equal to the bias, that is $E[\mathbf{x}_i - \hat{\tilde{\mathbf{x}}}_i^-] = \mathbf{b}_i$.
For this poster, the bias was simply propagated as:

$$\mathbf{b}_i = \mathbf{b}_{i-1}$$

Through assimilation of observations the model results can be updated. The observations are related to the state by the operator $\mathbf{H}_i$ and are assumed to be prone to random error $\mathbf{v}_i$:

$$\mathbf{y}_i = \mathbf{H}_i \mathbf{x}_i + \mathbf{v}_i$$

Friedland (1969) proposed a method to estimate both the state and the bias through Kalman filtering by applying 2 filters:

1. a regular Kalman filter for the update of the biased state;
each ensemble of model forecasts was updated to get a bias-blind analysis estimate (1)
with the Kalman gain for state estimation (2)

$$\hat{\tilde{\mathbf{x}}}_{j,i} = \hat{\tilde{\mathbf{x}}}_{j,i}^- + \tilde{\mathbf{K}}_{x,i}[\mathbf{y}_{j,i} - \mathbf{H}_i \hat{\tilde{\mathbf{x}}}_{j,i}^-] \quad (1)$$
$$\tilde{\mathbf{K}}_{x,i} = \tilde{\mathbf{P}}_{x,i}^- \mathbf{H}_i^T [\mathbf{H}_i \tilde{\mathbf{P}}_{x,i}^- \mathbf{H}_i^T + \mathbf{R}_i]^{-1} \quad (2)$$

2. a second Kalman filter to estimate the bias.
the bias (no ensembles) was updated by (3)
with a Kalman gain for bias estimation (4)

$$\hat{\mathbf{b}}_i = \hat{\mathbf{b}}_i^- + \mathbf{K}_{b,i}[\mathbf{y}_i - \mathbf{H}_i(\hat{\tilde{\mathbf{x}}}_i^- + \hat{\mathbf{b}}_i^-)] \quad (3)$$
$$\mathbf{K}_{b,i} = \mathbf{P}_{b,i}^- \mathbf{H}_i^T [\mathbf{H}_i \mathbf{P}_{b,i}^- \mathbf{H}_i^T + \mathbf{H}_i \tilde{\mathbf{P}}_{x,i}^- \mathbf{H}_i^T + \mathbf{R}_i]^{-1} \quad (4)$$

The bias estimate can be used to correct the biased-blind model state estimates for output (5), but the bias-bind state estimate is fed back into the model (without bias correction).

$$\hat{\mathbf{x}}_{j,i} = \hat{\tilde{\mathbf{x}}}_{j,i} + [\mathbf{I} - \tilde{\mathbf{K}}_{x,i}\mathbf{H}_i]\hat{\mathbf{b}}_i \quad (5)$$

All P-matrices are referring to error covariance matrices of either the state estimation error ($\tilde{\mathbf{P}}_{x,i}^-$) or the bias estimation error ($\mathbf{P}_{b,i}^-$).

Alternative methods

→ For the output states, the bias-corrected estimates (if available) are generally of most interest.
→ For the model re-initialization step, it is not clear whether feeding back bias-blind, bias-corrected, or fully biased estimates is most beneficial.
The table 1 and figure 1 illustrate several possible methods to correct the model state and/or the resulting output.

Table 1 : Methods for state and/or bias estimation

EnKF	conventional state updating without explicit bias estimation
EnBKF_0	bias estimation without state update
EnBKF_1	separate state and bias estimation of Friedland (1969)
EnBKF_2	partial feedback of bias correction on the analysis through bias-corrected innovations
EnBKF_3	complete feedback of the bias corrections on the analysis

Superscript + means that the bias estimate is added to the output state.

Method	State update	Bias update	Bias feedback
EnKF	yes	no	n/a
EnBKF_0	no	yes	no
EnBKF_1	yes	yes	no
EnBKF_2/2$^+$	yes	yes	partial
EnBKF_3/3$^+$	yes	yes	complete

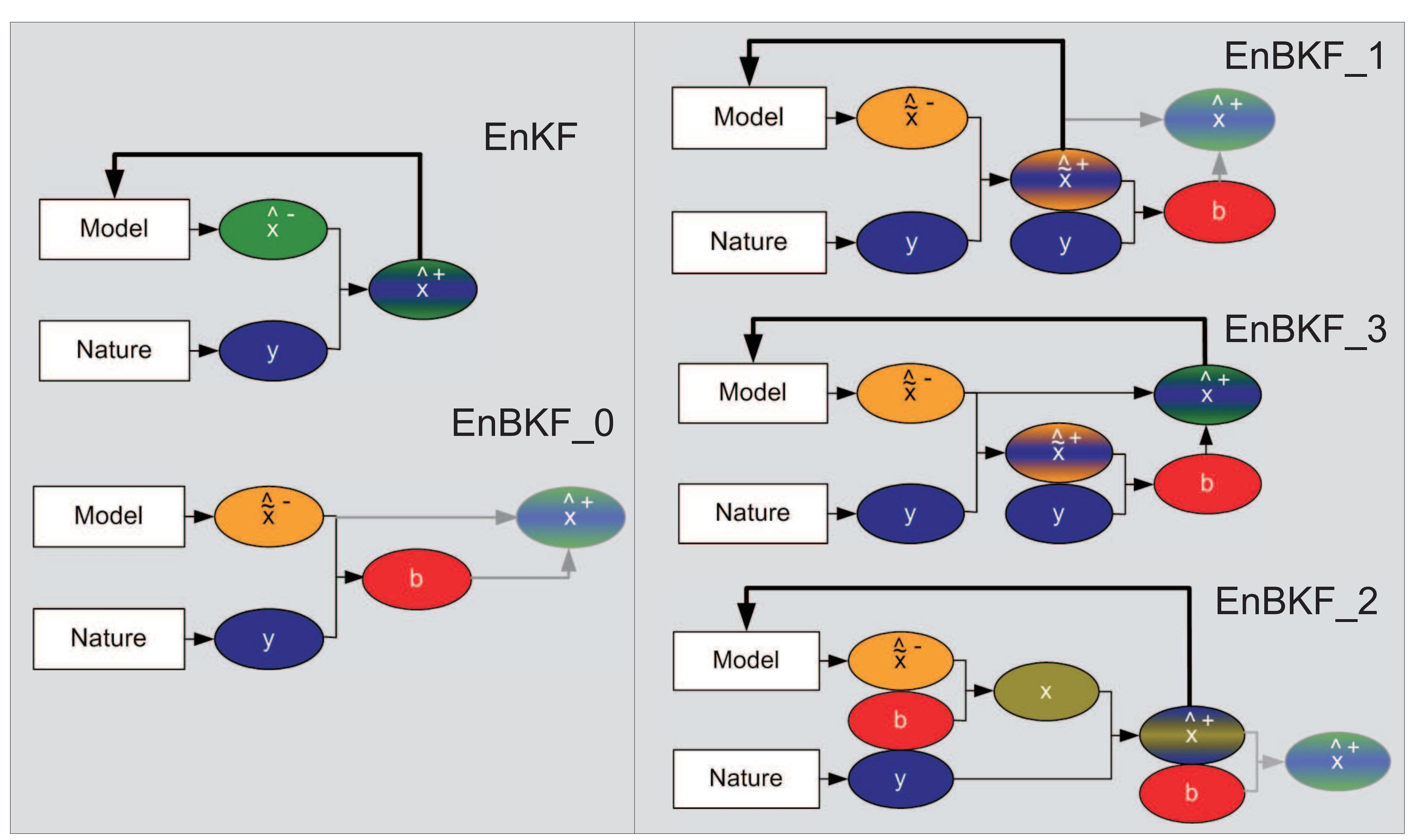

Figure 1 : Methods for state and/or bias estimation

3. Data and Model

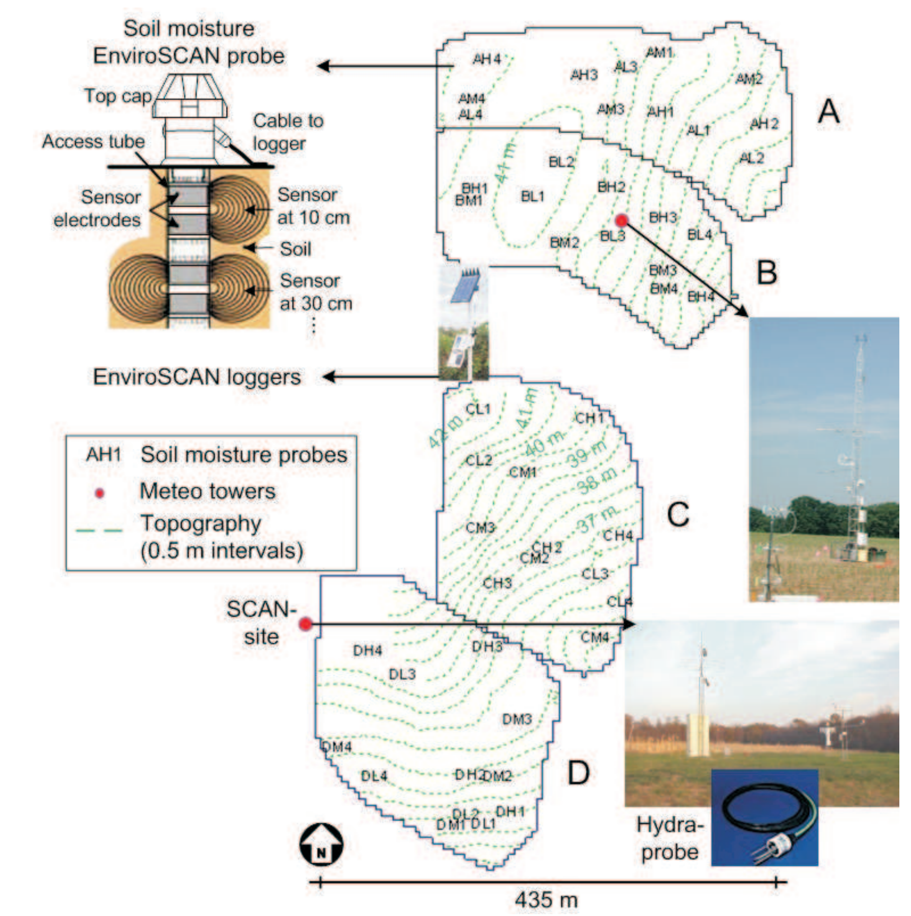

Figure 2 : OPE3 field with location of soil moisture probes and meteorological towers.

Project: Optimizing Production Inputs for Economic and Environmental Enhancement (OPE[3])
- BARC-ARS (USDA)
- Agricultural corn field
- Greenbelt, MD, USA
- 21 ha, 4 sub watersheds, drains in a first order stream

Soil moisture:
- 36 working capacitance probes during May 2001- May 2002
- H-probes: sensors at 10, 30 and 80 cm depth
- L- and M-probes: sensors at 10, 30, 50, 120, 150, 180 cm depth
- L-probes: additional sensor at 80 cm

Land surface model:
- Community Land Model v2.0
- Simulates water and heat fluxes for individual soil profiles
- Monte Carlo optimization of initial states and parameters
- Uncertainty simulated by ensembles (patches per grid cell)
- Ensemble mean = control run

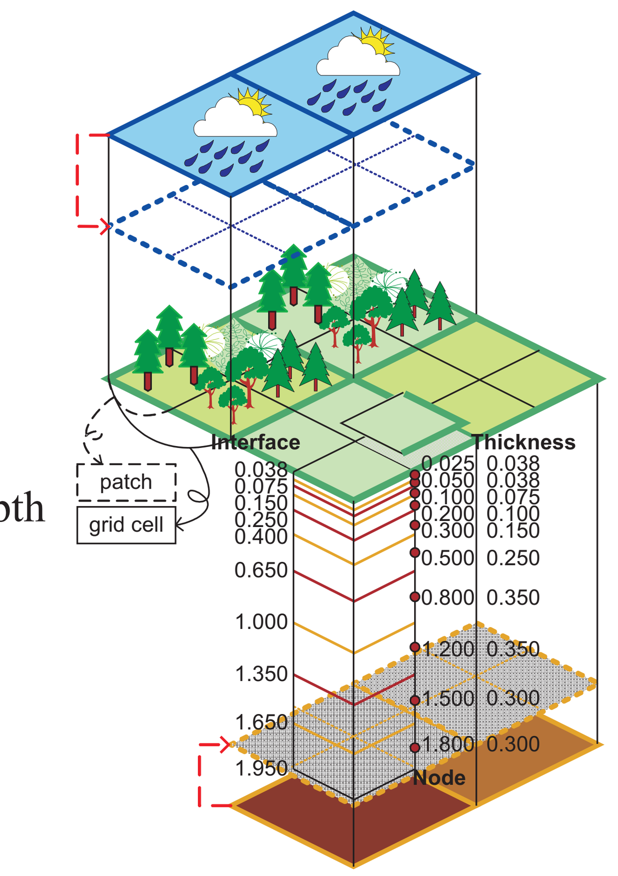

Figure 3 : Structure for CLM2.0 simulations.

4. Result

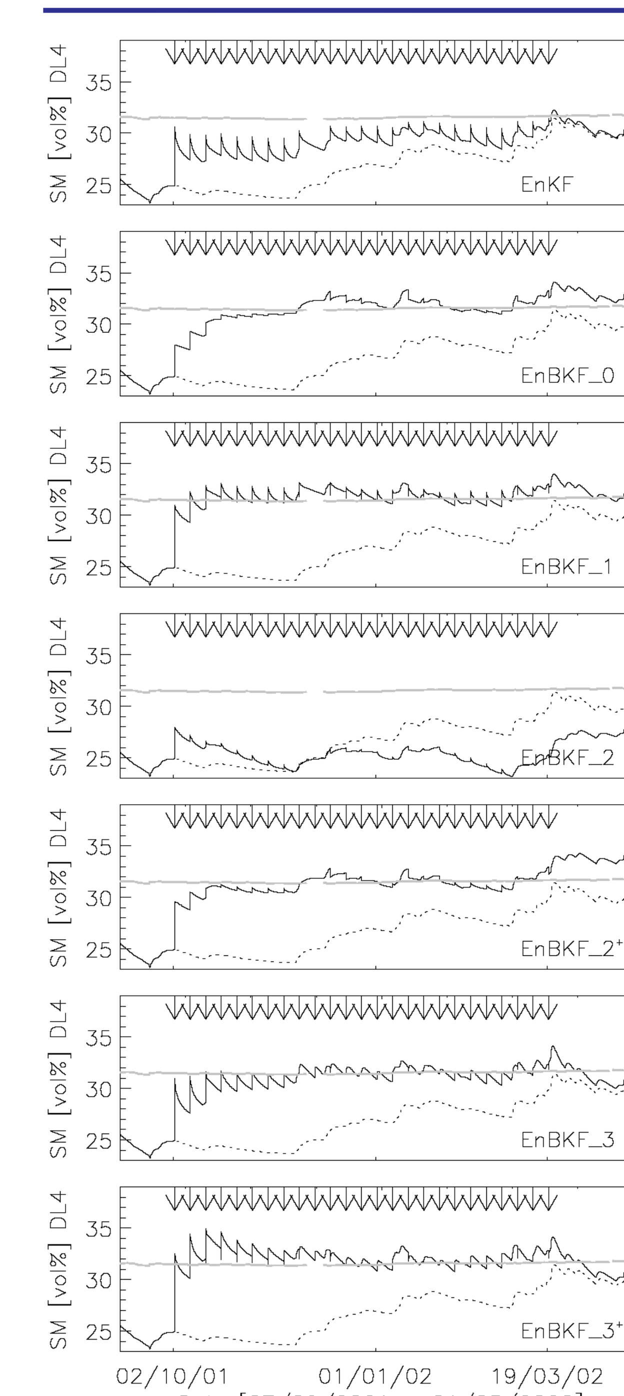

Figure 6 : Soil moisture at 80 cm for probe DI4 for different estimation algorithms. Shown are (gray) observations, (dotted) ensemble mean without filtering, and (black solid line) filtering runs. The assimilation was weekly (arrows) and only at 80 cm.

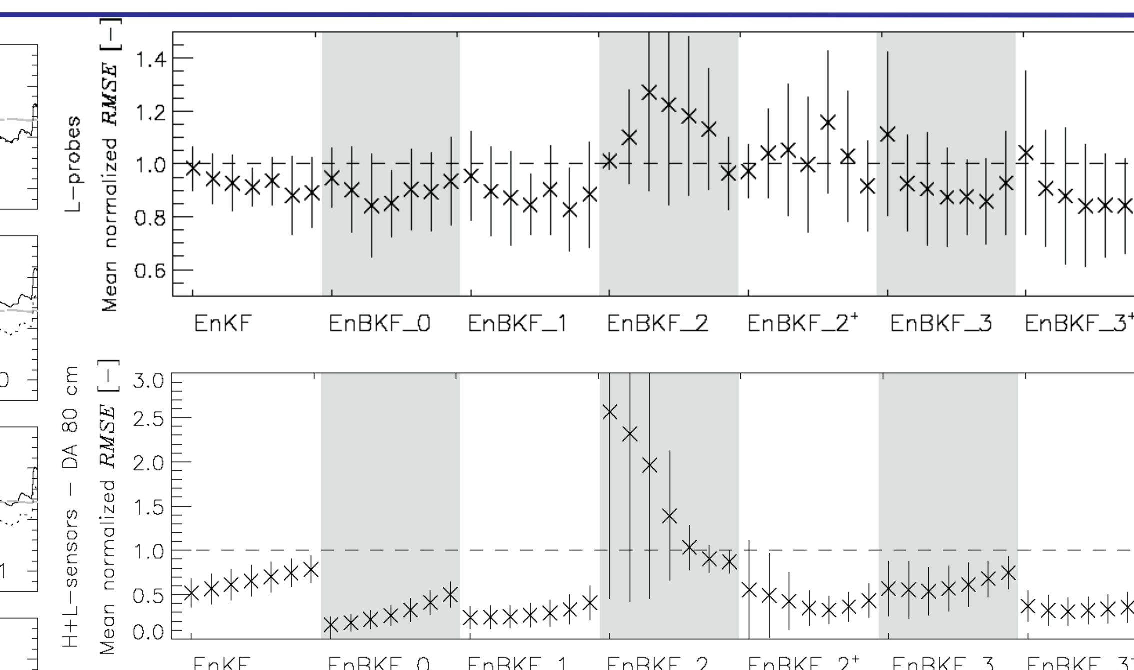

Figure 4 : (Top) Profile-integrated, area-average, normalized RMSE for different algorithms with weekly assimilation at 7 different depths (10, 30, 50, 80, 120, 150 or 180 cm) per algorithm. Normalization of RMSE is with respect to the control run. Averages were calculated over all layers and all L-probes. One spatial standard deviation is also shown.

Figure 5 : (Bottom) Normalized single layer RMSE different algorithms with 7 different assimilation frequencies per algorithm: every day, 2 days, 4 days, week, 2 weeks, 4 weeks and 8 weeks. The assimilation depth was only at 80 cm.

- The performance for the assimilation layer itself was greatly improved relative to the control run through inclusion of bias estimation (*figure 5*), even though the profile-average performance improvement was modest (*figure 4*)
→ Problem in observability of the bias system
- Assimilation of complete profile had little effect on evapotranspiration and runoff, but large impact on subsurface drainage (*table 2*)
→ Upper layers simulated already well in the control run

Table 2 : Area-average states and total fluxes (2 October 2001 – 30 April 2002) for 2-weekly assimilation of complete observed profiles, plus/minus one (spatial) standard deviation. Per column, the two listed algorithms yield the same output for all variables except for soil moisture and for the increment. EnCtrl stands for the control, i.e .the ensemble mean run without filtering.

	Units	EnCtrl EnBKF_0	EnKF EnBKF_1	EnBKF_2 EnBKF_2$^+$	EnBKF_3 EnBKF_3$^+$
Profile average soil moisture	[vol%]	16.5± 2.8 19.2± 5.1	17.9± 3.6 19.7± 5.3	16.8± 2.8 19.2± 4.8	18.2± 3.8 20.0± 5.4
Evapotranspiration	[mm]	67± 11	67± 11	67± 11	67± 11
Ground evaporation	[mm]	42± 10	42± 10	42± 10	42± 10
Canopy evaporation	[mm]	0.2± 0.7	0.1± 0.7	0.1± 0.7	0.1± 0.7
Canopy transpiration	[mm]	25± 10	25± 10	25± 10	25± 10
Surface runoff	[mm]	38± 11	39± 11	37± 11	39± 11
Subsurface drainage	[mm]	109± 32	427± 393	149± 60	571± 526
Soil moisture *RMSE*	[vol%]	6.74± 3.47 2.00± 0.70	4.98± 2.74 1.94± 0.71	6.21± 3.38 1.88± 0.70	4.82± 2.73 2.08± 0.75
Average absolute increment*	[mm/event]	n/a 65± 58	26± 29 39± 40	7± 4 51± 50	38± 39 36± 41

*For EnCtrl, EnKF, EnBKF_2, and EnBKF_3 there are 13 assimilation events. For EnBKF_0, EnBKF_1, EnBKF_2$^+$, and EnBKF_3$^+$ there are 24∗211=5,064 hourly increments (events) over the 211-day period.

5. Conclusion

Simple state updating with the conventional ensemble Kalman filter (EnKF) allows for some implicit bias correction. It is possible to estimate the soil moisture bias explicitly and derive superior soil moisture estimates with a generalized EnKF that uses a simple persistence model for the bias and assumes that the a priori bias error covariance is proportional to the a priori state error covariance. Significant improvements, however, are limited to layers for which observations are available. Therefore, it is crucial to measure the state variables of interest. The best variant for state and bias estimation depends on the nature of the model bias. In a biased model, low errors in soil moisture estimates may require large and frequent increments which in turn negatively impact the water balance and output fluxes.

Friedland, B., Treatment of bias in recursive filtering, IEEE Transactions on 13 Automatic Control , AC-14 , 359–367, 1969.