Soil moisture data assimilation:
Error modeling, adaptive filtering, and the contribution of soil moisture retrievals to land data assimilation products

R. Reichle\textsuperscript{1,2}, W. Crow\textsuperscript{3}, R. Koster\textsuperscript{1,2}, C. Keppenne\textsuperscript{2}, S. Mahanama\textsuperscript{1,2}, and H. Sharif\textsuperscript{4}

Rolf.Reichle@nasa.gov

1 – Goddard Earth Sciences and Technology Center, UMBC
2 – Global Modeling and Assimilation Office, NASA-GSFC
3 – Hydrology and Remote Sensing Lab, USDA-ARS
4 – Civil Engineering Dept., University of Texas, San Antonio
Outline

• **Motivation**
  • Soil moisture data assimilation

• **Part 1** *(doi:10.1029/2007WR006357)*
  • Impact of input error parameters on soil moisture estimates
  • Adaptive filtering

• **Part 2** *(doi:10.1029/2007GL031986)*
  • Contribution of soil moisture retrievals to land assimilation products

http://userpages.umbc.edu/~reichle/
Large-scale soil moisture is needed, for example, for water cycle studies and for initializing weather/climate models. It is available from:

- **AMSR-E surface soil moisture**
  - *Upper 1cm, ~50km, ~daily.*

- **Catchment land surface model**
  - Forced w/ observed meteorology. *Complete space-time coverage, incl. root zone.*

Weights based on respective uncertainties.

**Assimilation**

- Soil moisture retrievals (subject to error)
- Model soil moisture (subject to error)

*“Optimal” soil moisture*

*a.k.a. “Level 4 product”*
Global assimilation of AMSR-E soil moisture retrievals

Assimilate AMSR-E surface soil moisture (2002-06) into NASA Catchment model

Validate with USDA SCAN stations (only 23 of 103 suitable for validation)

<table>
<thead>
<tr>
<th></th>
<th>Anomaly time series correlation coeff. with in situ data [-] (with 95% confidence interval)</th>
<th>Confidence levels: Improvement of assimilation over</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Satellite</td>
</tr>
<tr>
<td>Surface soil moisture</td>
<td>23</td>
<td>.38±.02</td>
</tr>
<tr>
<td>Root zone soil moisture</td>
<td>22</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Assimilation product agrees better with ground data than satellite or model alone. Modest increase may be close to maximum possible with *imperfect* in situ data.

Reichle et al., *JGR*, 2007
Outline

- **Motivation**
  - Soil moisture data assimilation

- **Part 1** (doi:10.1029/2007WR006357)
  - Impact of input error parameters on soil moisture estimates
  - Adaptive filtering

  - Contribution of soil moisture retrievals to land assimilation products

http://userpages.umbc.edu/~reichle/
Input error parameters $Q$ and $R$

Weights based on respective uncertainties.

Soil moisture retrievals (subject to error) → Assimilation → “Optimal” soil moisture → Model soil moisture (subject to error)
Input error parameters Q and R

Weights themselves are subject to error!!!
Wrong weights may lead to poor estimates.

Model soil moisture (subject to error)

Soil moisture retrievals (subject to error)

“Optimal” soil moisture

Retrieval error covariance R (subject to error)

Model error covariance Q (subject to error)
Synthetic assimilation experiment

Investigate impact of wrong model and obs. error inputs on assimilation estimates:

"True" precip., radiation, ...

"True" land model

"True" soil moisture

Repeat for many different sets of model and retrieval error cov’s.

Precip., radiation, … (subject to error)

Land model (subject to error)

Model soil moisture (subject to error)

Retrieval error covariance R (subject to error)

Model error covariance Q (subject to error)

Soil moisture retrievals (subject to error)

Assimilation (EnKF)

“Optimal” soil moisture

compare

Reichle et al., doi:10.1029/2007WR006357
Red-Arkansas river basin

Red-Arkansas river basin (308 catchments)
NASA Catchment land surface model (identical twin experiment)

Annual Precipitation (mm)

West: Dry with sparse vegetation
East: Wet with dense vegetation

Sharif et al., JHM, 2007
Impact of Q and R on assimilation estimates

RMSE of assimilation estimates v. truth for:

Surface soil moisture $m^3/m^3$

sqrt(R_true) = 0.05, OL = 0.035

Each “+” symbol represents one 19-year assim. experiment over the Red-Arkansas with a unique combination of input model and observation error parameters.

Q = model error (including errors in precip, radiation, and soil moisture tendencies)

P = P(Q) = soil moisture error variance

Reichle et al., doi:10.1029/2007WR006357
Impact of Q and R on assimilation estimates

- "True" input error covariances yield minimum estimation errors.
- Wrong model and obs. error covariance inputs degrade assimilation estimates.
- In most cases, assimilation still better than open loop (OL).

Reichle et al., doi:10.1029/2007WR006357
Impact of Q and R on assimilation estimates

Root zone soil moisture $m^3/m^3$ vs. surface soil moisture $m^3/m^3$ for:
- $\sqrt{R_{true}} = 0.05$, $OL = 0.035$
- $\sqrt{R_{true}} = 0.05$, $OL = 0.020$

- Root zone more sensitive than surface soil moisture.

Reichle et al., doi:10.1029/2007WR006357
Impact of Q and R on assimilation estimates (fluxes)

RMSE of assimilation estimates v. truth for:

- Sensible heat flux $W/m^2$
  \[ \text{sqrt}(R_{true})=0.05, \text{OL}=21.749 \]

- Latent heat flux $W/m^2$
  \[ \text{sqrt}(R_{true})=0.05, \text{OL}=26.932 \]

- Runoff $mm/d$
  \[ \text{sqrt}(R_{true})=0.05, \text{OL}=0.305 \]

- Fluxes more sensitive to wrong error parameters than soil moisture.
- Sensible/latent heat more sensitive to model error cov than obs error cov (probably related to ensemble propagation).

Reichle et al., doi:10.1029/2007WR006357
Motivation

Soil moisture data assimilation


Impact of input error parameters on soil moisture estimates

Adaptive filtering


Contribution of soil moisture retrievals to land assimilation products

http://userpages.umbc.edu/~reichle/
Diagnostics of filter performance and adaptive filtering

Find true Q, R by enumeration?
- RMSE plots require “truth” (not usually available).
- Too expensive computationally.

Use diagnostics that are available within the assimilation system.

Filter update: \( x^+ = x^- + K (y - x^-) \)

K \( = P (P + R)^{-1} \) = Kalman gain

Diagnostic: \( E[(y - x^-) (y - x^-)^T] = P + R \)

innovations \( \equiv \) obs – model prediction
(internal diagnostic)

state err cov + obs err cov
(controlled by inputs)

Example: Average “obs. minus model prediction”
distance is much larger than assumed input uncertainties
Diagnostics of filter performance and adaptive filtering

Find true Q, R by enumeration?
- RMSE plots require “truth” (not usually available).
- Too expensive computationally.

Use diagnostics that are available within the assimilation system.

Filter update: \[ x^+ = x^- + K(y - x^-) \]

K = \[ P (P + R)^{-1} \] = Kalman gain

Diagnostic: \[ E[(y - x^-)(y - x^-)^T] = P + R \]

innovations ≡ obs – model prediction

diagnostic

state err cov + obs err cov

(controlled by inputs)

Contours: misfit between diagnostic and what it “should” be.
Adaptive filter: Nudge input error parameters (Q, R) during assimilation to minimize misfit.

Reichle et al., doi:10.1029/2007WR006357
Diagnostics of filter performance and adaptive filtering

Find true Q, R by enumeration?
- RMSE plots require “truth” (not usually available).
- Too expensive computationally.

Use diagnostics that are available within the assimilation system.

Filter update:  \( x^+ = x^- + K(y - x^-) \)

\[ K = P(P + R)^{-1} = \text{Kalman gain} \]

Diagnostics:
\[ E[(y - x^-) (y - x^-)^T] = P + R \]

Innovations \( \equiv \) obs – model prediction (diagnostic)

State err cov + obs err cov (controlled by inputs)

Contours: misfit between diagnostic and what it “should” be.

Adaptive filter: Nudge input error parameters (Q, R) during assimilation to minimize misfit.

Diagnostic 1: \[ E[(y - x^+) (y - x^-)^T] = R \]

Diagnostic 2: \[ E[(x^+ - x^-) (y - x^-)^T] = P(Q) \]

Reichle et al., doi:10.1029/2007WR006357
Adaptive algorithm

1. EnKF propagation and update

Propagate model:
\[ x_{t,i}^- = f(x_{t-1,i}^+, q_{t,i}) \]

Forecast error cov.:
\[ P_t = E\{(x_{t,i}^- - E(x_{t,i}^-))^2\} \]
Kalman gain:
\[ K_t = P_t H_t^T (H_t P_t H_t^T + R_t)^{-1} \]
Analysis update:
\[ x_{t,i}^+ = x_{t,i}^- + K_t (y_{t,i} - H_t x_{t,i}^-) \]

Innovations:
\[ v_t = E\{y_{t,i} - H_t x_{t,i}^-\} \]
Analysis departures:
\[ w_t = E\{y_{t,i} - H_t x_{t,i}^+\} \]
Analysis increments:
\[ u_t = E\{H_t (x_{t,i}^+ - x_{t,i}^-)\} \]

2. Moving average of filter diagnostics

\[ \text{a.) } MA[u v^T]_t = (1-\gamma) MA[u v^T]_{t-1} + \gamma u_t v_t^T \]
\[ \text{b.) } f_Q = \beta \frac{MA[u v^T]_t}{MA[HPH]^T_t} \]
\[ \text{c.) } \alpha_{Q,t} = \alpha_{Q,t-1} \max(\min(f_Q, \alpha_{\text{max}}), \alpha_{\text{min}}) \]
\[ \text{d.) } \alpha_{R,t} = \alpha_{R,t-1} \max(\min(f_R, \alpha_{\text{max}}), \alpha_{\text{min}}) \]
\[ \text{e.) } Q_{t+1} = \alpha_{Q,t} Q_0 \] (and generate \( q_{t,i} \))
\[ R_{t+1} = \alpha_{R,t} R_0 \]

3. Adaptive scaling coefficients

- Adapted Dee et al. for land
- Cheap
- Need parameters

Reichle et al., doi:10.1029/2007WR006357
Adaptive scaling factors generally converge to true values (thick lines).
Convergence is slow (order of years).
Spatial variability (thin lines) much greater for \( \alpha_Q \) than for \( \alpha_R \).

\[ \sqrt{R_0} = 0.02 \]
\[ \sqrt{R_0} = 0.08 \]

Reichle et al., doi:10.1029/2007WR006357
Adaptive v. non-adaptive EnKF (soil moisture)

- Adaptive filter: Map experiment onto contour plot based on initial guess of $R$, $P(Q)$.
- Adaptive filter yields improved assimilation estimates for initially wrong model and observation error inputs (except for $R_0=0$).

Reichle et al., doi:10.1029/2007WR006357
Adaptive v. non-adaptive EnKF (fluxes)

Contours: RMSE of assim. est. v. truth

- Adaptive filter generally yields improved flux estimates.
- Degradation when R is severely underestimated.
  → Simply choose large R at the start and let the filter adapt it.

Reichle et al., doi:10.1029/2007WR006357
Adaptive v. non-adaptive EnKF (filter diagnostics)

<table>
<thead>
<tr>
<th>Error in estimate of obs error std $\sqrt{R}$ m³/m³</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-adaptive</td>
</tr>
</tbody>
</table>

Log10 of innov. misfit

- Adaptive filter (by design) improves innovations stats.

Error in estimate of analysis error std “$\sqrt{P^+}$” m³/m³

- Adaptive filter retrieves obs error std (except for $R_0=0$).

- On balance, adaptive filter improves estimate of error bars on assimilation product (surface soil moisture).
Wrong model and observation error inputs degrade assimilation estimates. 
Degradation quantified with synthetic experiment over Red-Arkansas river basin.

**Adaptive EnKF:**
+ Generally improves assimilation estimates.
+ Cheap.

**Future applications:**
Use for AMSR-E soil moisture assimilation.
Estimates of AMSR-E obs. error variance (not provided by official NASA product).
Outline

• Motivation
  • Soil moisture data assimilation

  • Impact of input error parameters on soil moisture estimates
  • Adaptive filtering

  • Contribution of soil moisture retrievals to land assimilation products

http://userpages.umbc.edu/~reichle/
**Problem statement**

Design problem for future satellite missions (eg. NASA Soil Moisture Active Passive “SMAP” mission)

*How uncertain can retrievals be and still add useful information in the assimilation system?*

<table>
<thead>
<tr>
<th></th>
<th>Anomaly time series correlation coeff. with in situ data [-] (with 95% confidence interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
</tr>
<tr>
<td>Surface soil moisture</td>
<td>23</td>
</tr>
</tbody>
</table>

*Example:* If target skill=0.5 and model skill=0.43, need retrieval skill $\geq 0.38$.

*Goal:* Contour plot based on many such triplets of numbers.
Soil moisture retrieval
“Observing System Simulation Experiment” (OSSE):
Can we achieve a retrieval accuracy of ~0.04 m³/m³ (“4%”) in absolute soil moisture with realistic errors in brightness temperatures and retrieval parameters?
Soil moisture assimilation OSSE: Design

1.) Add data assimilation.

Precip., radiation, … (subject to error) → Land model (subject to error) → “Optimal” soil moisture

Model soil moisture → Assimilation → Soil moisture retrievals


Retrieval algorithm (subject to error) → Brightness temp. (subject to error)

Reichle et al., doi:10.1029/2007GL031986
Soil moisture assimilation OSSE: Design

1.) Soils are initialized from a land model and then evolved with a land model and radiative transfer model to produce true soil moisture and true brightness temp.
2.) Repeat for many different sets of model and retrieval error characteristics to get contour plots.

Reichle et al., doi:10.1029/2007GL031986
Soil moisture assimilation OSSE: Implementation

Sharif et al 2007 forcing (1km) → TOPLATS (1km) → “True” soil moisture, ET (1km) → H-pol. ω,τ radiative transfer model → “True” brightness temp. (1km)

Model forcing (subject to error, ~35km) → Catchment LSM (35km)

Model soil moisture, ET → Assimilation products: soil moisture, ET → Adaptive 1d EnKF w/ cdf-matching → Surf. Soil moisture retrievals (36km) → Inverse horiz.-pol. ω,τ model (subject to error) → Brightness temp. (subject to error, 36km)

Reichle et al., doi:10.1029/2007GL031986
Soil moisture assimilation OSSE: Implementation

Model scenario | M1 | M2 | M4 | M3 | ... | M8
---|---|---|---|---|---|---
Base forcing dataset | F1 | F2 | F3 | F1 | ... | F1
Forcing shift [days] | n/a | n/a | n/a | 7 | ... | 365
Rsf (skill) | 0.76 | 0.63 | 0.41 | 0.5 | ... | -0.01
Rrz (skill) | 0.78 | 0.55 | 0.46 | 0.64 | ... | 0.01
R_{ET} (skill) | 0.65 | 0.38 | 0.37 | 0.58 | ... | 0.02

8 x 12 = 96 assimilation experiments

Adaptive 1d EnKF w/ cdf-matching

Assimilation products: soil moisture, ET

Surf. Soil moisture retrievals (36km)

Inverse horiz.-pol. $\omega, \tau$ model (subject to error)

Perturbations to VWC, Tsoil, and parameters for vegetation opacity

Catchment LSM (35km)

Model soil moisture, ET

"True" brightness temp. (1km)

Brightness temp. (subject to error, 36km)

Reichle et al., doi:10.1029/2007GL031986

Aggregation errors

Model forcing (subject to error, ~35km)
Skill of soil moisture estimates

Skill is measured in terms of $R$ (=anomaly time series correlation coefficient against truth).

Contours show the skill of the assimilation product

X-axis: Skill of retrievals
Y-axis: Skill of model product

Each plus sign indicates the result of one 19-year assimilation integration over the entire Red-Arkansas domain.

Reichle et al., doi:10.1029/2007GL031986
The skill of the soil moisture (surface and root zone) assimilation product increases with the skill of the retrievals and the skill of the model. The skill of the assimilation product is more sensitive to model skill than to retrieval skill.

Reichle et al., doi:10.1029/2007GL031986
Assimilation of soil moisture retrievals adds skill (relative to model product).

Even retrievals of poor quality contribute information to the assimilation product.
Skill improvement (soil moisture)

- Assimilation of soil moisture retrievals adds skill (relative to model product).
- Even retrievals of poor quality contribute information to the assimilation product.
- Published AMSR-E and SMMR assimilation products are consistent with expected skill levels for surface soil moisture, to a lesser degree also for root zone soil moisture.

Reichle et al., doi:10.1029/2007GL031986
• Assimilation of surface soil moisture retrievals yields, on average, modest improvements in ET estimates.
• Negative ΔR related to technicalities (EnKF bias issues and adaptive filtering).
General DA-OSSE framework developed:

- Quantify the information added to land assimilation products by satellite retrievals for detailed and comprehensive error budget analyses for data assimilation products.
- **Adaptive filtering** is major component of the DA-OSSE.
- Success of DA-OSSE depends on realism of imposed model errors.

Soil moisture assimilation study for the Red-Arkansas:

- Even retrieval data sets of poor quality contribute information to the assimilation product.
- Published AMSR-E and SMMR assimilation products are consistent with expected skill levels for surface soil moisture, to a lesser degree also for root zone soil moisture.

Future applications:

- Extending the DA-OSSE to continental/global scales is straightforward but computationally demanding.
- Same applies for higher-resolution soil moisture retrievals (e.g. from active/passive MW sensor).