Ensemble-based Streamflow Data Assimilation for an Operational Distributed Hydrologic Model

Minxue He[1,2], Haksu Lee[1,3], and Yuqiong Liu[1,2]

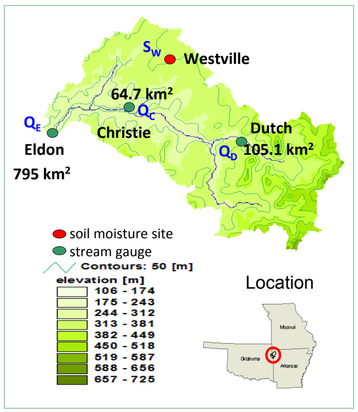
[1]NOAA, National Weather Service, Office of Hydrologic Development
[2]Riverside Technology, inc., Fort Collins, Colorado 80525, USA [3]University Corporation for Atmospheric Research

H13A-0936

Introduction

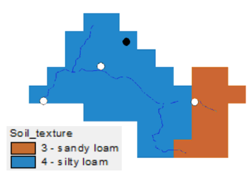
- Applying advanced data assimilation techniques to distributed modeling holds great potential for improving operational streamflow forecasting, particularly in light of increasing computing power and the availability of remotely sensed data with high spatial and temporal resolutions.
- The objectives of this work are to
  1) Integrate an ensemble-based data assimilation framework (i.e. Ensemble Kalman Filter (EnKF)) within the National Weather Service (NWS) Hydrology Lab's Research Distributed Hydrologic Model (HL-RDHM);
  2) Assess the performance of this ensemble framework in estimating flows as compared to observed flows, flows derived from stand alone HL-RDHM and from a deterministic DA technique (i.e. four dimensional variational data assimilation (4DVAR)).

Study Area



Left: Eldon basin (ELDO2)
- elevation map (50m resolution)
- main channel network
- soil moisture gauge
- streamflow gauges

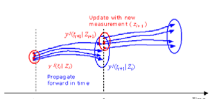
Below: ELDO2 soil map



Dataset: hourly gridded precipitation & hourly outlet flow (2000/6/21-2005/9/30)

Methodology

I. EnKF



Step 1: Propagation

$$y_j(t) = A[\alpha_j, u_j(t), t]; \quad y_j(t=t_0) = y_{0j}$$

Step 2: Updating

$$y_j^+ = y_j^- + K[z + \varepsilon_z - M_j(y_j^-)]$$

$$K = C_{yz}[C_{zz} + C_v]^{-1}$$

$$j = 1, 2, \ldots, N$$

Z: Flow measurement

y: Model states

N: Ensemble size

II. 4DVAR

Step 1: Create adjoint model (from Tapenade (http://tapenade.inria.fr:8080/tapendade/index.jsp))

Step 2: Minimize

$$J_k = \frac{1}{2}[\mathbf{Z}_q - \mathbf{H}_{qq}(\mathbf{X}_{s,k-l}, \mathbf{X}_p, \mathbf{X}_e)]^T \mathbf{R}_{qq}^{-1}[\mathbf{Z}_q - \mathbf{H}_{qq}(\mathbf{X}_{s,k-l}, \mathbf{X}_p, \mathbf{X}_e)] \quad Z_q\text{: flow measurement}$$

$$+ \frac{1}{2}[\mathbf{Z}_\theta - \mathbf{H}_{\theta\theta}(\mathbf{X}_{s,k-l}, \mathbf{X}_p, \mathbf{X}_e)]^T \mathbf{R}_{\theta\theta}^{-1}[\mathbf{Z}_\theta - \mathbf{H}_{\theta\theta}(\mathbf{X}_{s,k-l}, \mathbf{X}_p, \mathbf{X}_e)] \quad Z_\theta\text{: soil moisture meas.}$$

$$+ \frac{1}{2}[\mathbf{Z}_p - \mathbf{H}_{pp}\mathbf{X}_p]^T \mathbf{R}_{pp}^{-1}[\mathbf{Z}_p - \mathbf{H}_{pp}\mathbf{X}_p] \quad Z_p\text{: precipitation data}$$

$$+ \frac{1}{2}[\mathbf{Z}_e - \mathbf{H}_{ee}\mathbf{X}_e]^T \mathbf{R}_{ee}^{-1}[\mathbf{Z}_e - \mathbf{H}_{ee}\mathbf{X}_e] \quad Z_e\text{: potential evaporation (PE) data}$$

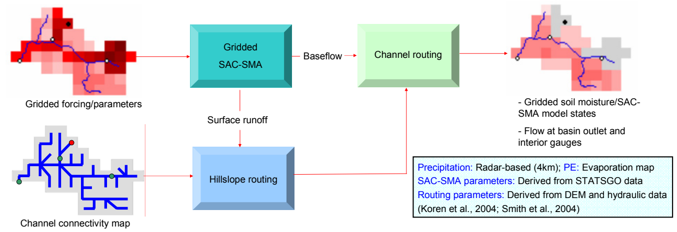
$$+ \frac{1}{2}[\mathbf{Z}_b - \mathbf{H}_{bb}\mathbf{X}_{s,k-l}]^T \mathbf{R}_{bb}^{-1}[\mathbf{Z}_b - \mathbf{H}_{bb}\mathbf{X}_{s,k-l}] \quad Z_b\text{: initial model soil moisture state}$$

$$\textit{subject to } \mathbf{X}_{s,j} = \mathbf{F}(\mathbf{X}_{s,j-1}, \mathbf{X}_{p,j}, \mathbf{X}_{e,j}),\ j = k-l+1, \ldots, k \quad \text{Abide by the model dynamics}$$

$$X_{s,i}^{\min} \le X_{s,k,i} \le X_{s,i}^{\max},\ j = k-l, \ldots, k;\ i = 1, \ldots, 6 \quad \text{Model states } (X_j) \text{ in feasible ranges}$$

Methodology (cont.)

III. Gridded SAC-SMA / Kinematic-wave Routing Models of HL-RDHM



IV. Experimental Design

a) Focus

Biggest four events of the dataset
- 1st: June 21-23, 2000 (Q_{peak} = 1549 cms)
- 2nd: April 22-26, 2004 (Q_{peak} = 1255 cms)
- 3rd: April 6-12, 2002 (Q_{peak} = 445 cms)
- 4th: Dec.15-21, 2001 (Q_{peak} = 368 cms)

b) Error models
- EnKF:
  Precipitation uncertainty: log-normal
  Flow measurement error: Gaussian
  (covariance depends on flow magnitudes)
- 4DVAR:
  Standard setting of Lee et al., 2010

c) Sensitivity test (1st event)

Baseline setting (BS):
- Precipitation uncertainty variance (C_{vp}): 0.5mm²
- Flow measurement error variance (C_{vq}): $(Q_t/60)^2$
- Ensemble size (N): 50

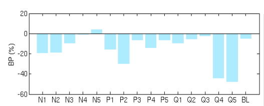
Scenarios
- N1-N5: N = 10,25,75,100,150 (C_{vp}, C_{vq} unchanged)
- P1-P5: C_{vp} = 0.1,0.25,0.36,1,1.5 (C_{vq}, N unchanged)
- Q1-Q5: Cvq = $(Qt/100)^2$, $(Qt/80)^2$, $(Qt/40)^2$, $(Qt/20)^2$, $(Qt/10)^2$ (C_{vp}, N unchanged)

d) Flow simulation

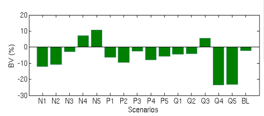
Evaluation metrics: bias in peak (BP), bias in volume (BV), timing error (TE)

Results

I. Sensitivity test



Left: Percent bias in simulated peak flow (BP) for the baseline and 15 sensitivity test scenarios. When N > 50, its impact on peak flow estimate is not significant; when uncertainty in precipitation (C_{vp}) is small and errors in flow observations (C_{vq}) are large, the filter tends to underestimate the peaks by a larger amount, indicating that proper error setting is needed to enable satisfactory filter performance.



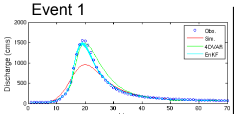
* Timing error (TE) only observed for scenarios N1 and Q5 (1 hour), suggesting that small N and large C_{vq} leads to TE

Above: Percent bias of total flow volume (BV). With increasing N, the bias changes gradually from negative to positive values; the sensitivity to errors in precipitation and flow observations shows a similar pattern to that of BP.

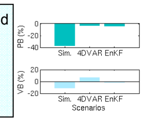
* The EnKF is configured with the setting of Baseline scenario and applied in flow simulation.

II. Flow simulation

Event 1

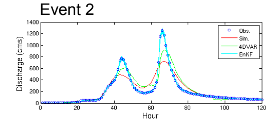


Left: HL-RDHM, 4DVAR, and EnKF-simulated flows of Event 1. Both 4DVAR and EnKF outperform the original model. EnKF under-catches the peak. Right: Bias in peak and volume. EnKF performs the best.

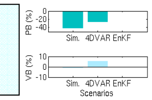


Results (cont.)

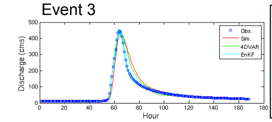
Event 2



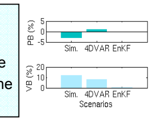
Left: Flow simulations for Event 2. EnKF simulations catch the peak and variation patterns well. Right: EnKF performs the best.



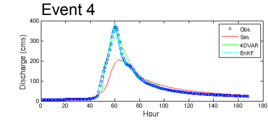
Event 3



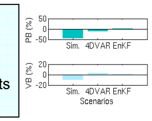
Left: Flow simulations for Event 3. Both 4DVAR and the model perform well, yet not as well as the EnKF. Right: EnKF results have the least bias in peak and volume.



Event 4



Left: Flow simulations for Event 4. 4DVAR under-catches the peak. EnKF mimics the peak and flow variations best. Right: EnKF results again have the least bias.



Right: TE of model-simulated, 4DVAR and EnKF derived flow for four events. EnKF estimates have no TE for the first three events. For Event 4, EnKF and 4DVAR estimates have comparable TE.

	Event 1	Event 2	Event 3	Event 4
Sim.	1	1	1	3
4DVAR	1	1	0	1
EnKF	0	0	0	-1

Summary

- Increasing ensemble size for EnKF does not necessarily lead to flow estimates with less bias in both peak and volume
- Inaccurate representation of errors in precipitation and flow observations in EnKF might lead to flow estimates with insufficient ensemble spread, as well as underestimated flow peak and volume
- TE tends to be mostly caused by insufficient ensemble size and over-estimated flow observation errors
- EnKF outperforms both the model and the 4DVAR in terms of estimating peak flows and overall variations in flow, even for the flow event with multiple peaks (i.e. Event 2)
- EnKF flow estimates are generally of lower TE in comparison to model simulation and 4DVAR results

Ongoing / Future Work

- Investigating the capability of the proposed ensemble filter in improving flow forecasting (and hence stage forecasting)
- Evaluating the performance of the filter in providing gridded soil moisture estimates and flow estimates at interior gauges
- Applying the filter in long-term simulations (rather than event-based)
- To assimilate in-situ soil moisture data at various depths
- To assimilate flow observations at interior gauges

References

Koren, V., S. Reed, M. Smith, Z. Zhang, and D.-J. Seo, 2004. Hydrology laboratory research modeling system (HL-RMS) of the US national weather service, *Journal of Hydrology*, 291, 297-318.

Lee, H., D.-J. Seo, and V. Koren, 2010. Assimilation of streamflow and in-situ soil moisture data into operational distributed hydrologic models: Effects of uncertainty in the data and initial soil moisture conditions, *Advances in Water Resources* (in review).

Smith, B., D.-J. Seo, V. Koren, S. Reed, Z. Zhang, Q. Duan, F. Moreda, and S. Cong, 2004. The distributed intercomparison project (DMIP): motivation and experimental design, *Journal of Hydrology*, 298, 4 -26.

UCAR
The University Corporation for Atmospheric Research

Contact: Kevin.He@noaa.gov

Riverside Technology, inc.
Water Resources Engineering and Consulting