Data Assimilator (DA) for Hydrology Laboratory's Research Distributed Hydrologic Model (HL-RDHM)

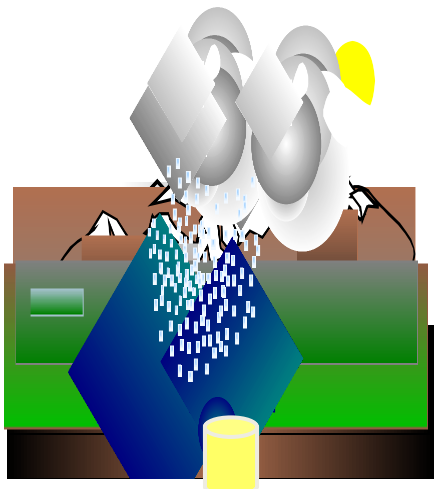
Haksu Lee[1,2], Dong-Jun Seo[1,2], Victor Koren[1]

NOAA/NWS/Office of Hydrologic Development[1]
University Corporation for Atmospheric Research[2]

Acknowledgments

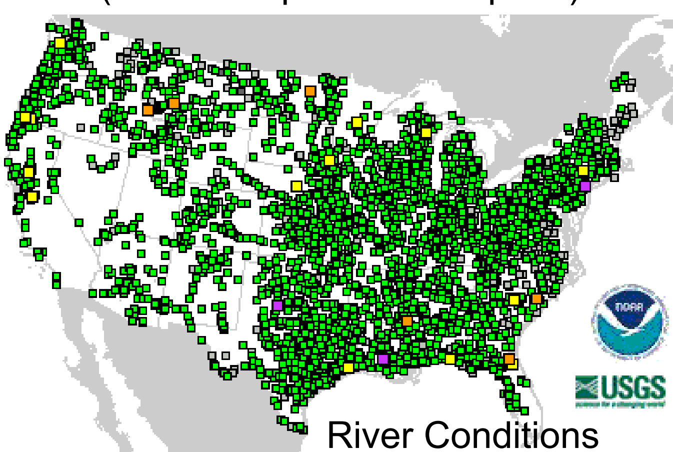
- NOAA Climate Program Office/Climate Prediction Program for the Americas (CPPA)
  - Core Project: Pedro Restrepo, John Schaake, Ken Mitchell
  - "Incorporating knowledge of observational uncertainties in streamflow forecasting applications in the Western US": Andrew Slater, Martyn Clark
- NASA
  - "Improving NOAA/NWS River Forecast Center Decision Support with NASA Satellite and Land Information System Products": Pedro Restrepo, Ashutosh Limaye, Christa Peters-Lidard
- AHPS/XEFS
- Hydrologic Ensemble Prediction Group - Limin Wu, Julie Demargne, James Brown, Satish Regonda, Yuqiong Liu
- Hydrology Group - Ziya Zhang
- Hydraulics Group - Fekadu Moreda

Predicting Floods to Droughts In Your Neighborhood

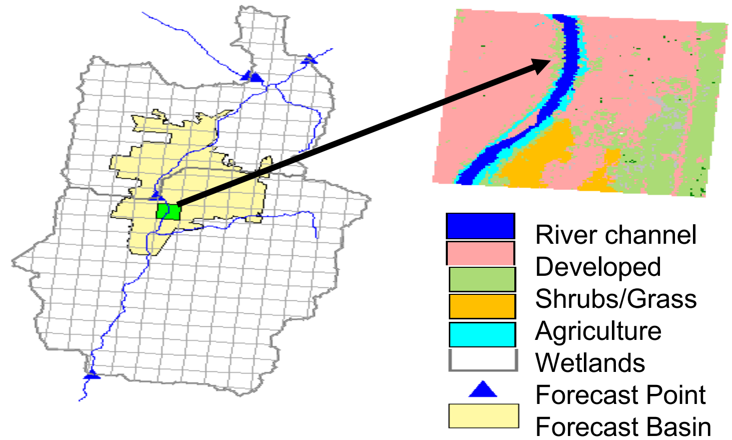


River Services

(600 miles per forecast point)

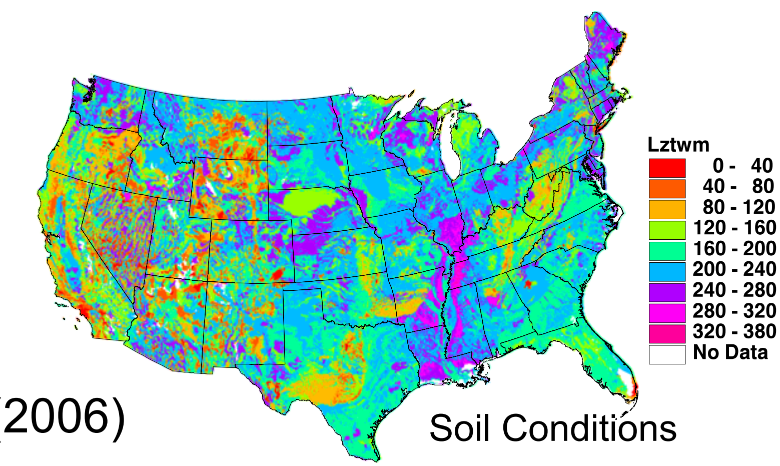


River Conditions



Water Resource Services

(6 square mile forecast basins)



Soil Conditions

From Carter (2006)

Objective of the project

- Develop a prototype data assimilator (DA) for distributed hydrologic models in HL-RDHM for more accurate, high-resolution analysis and prediction of streamflow and soil moisture
  - by reducing uncertainty in the model initial conditions (i.e. model soil moisture)

Outline of the presentation

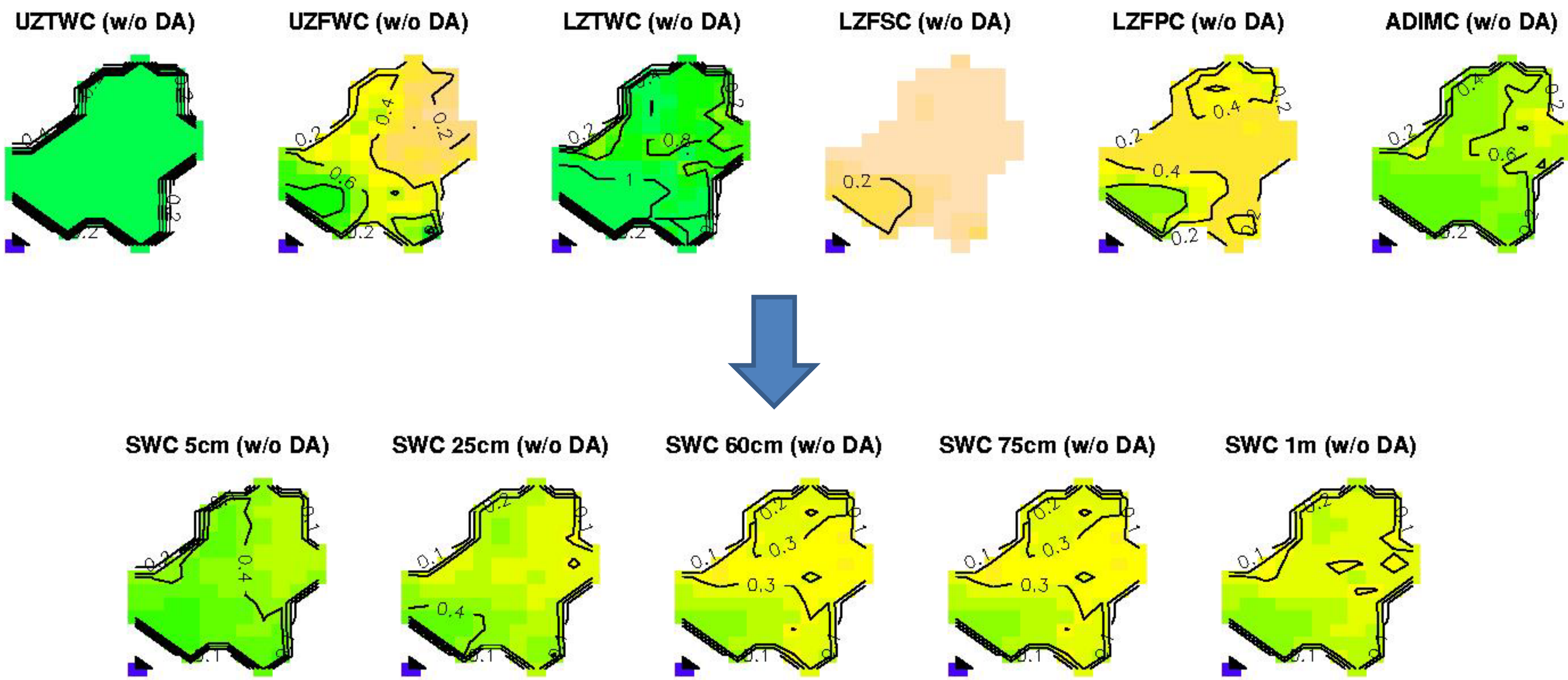
- Models used
- Technique used
  - What is 4DVAR?
  - How does 4DVAR work?
- Questions investigated
- Approach
  - Synthetic experiments
  - Real-world experiment
- Conclusions
- Next steps

Models used

- Hydrology Laboratory – Research Distributed Hydrologic Model (HL-RDHM, Koren et al. 2004)
  - Gridded (~4x4 km^2) soil moisture accounting models (SAC, API)
  - Gridded snow ablation model (SNOW-17)
  - Kinematic-wave routing

- The prototype DA assimilates the following data into gridded SAC-kinematic wave routing models (Seo et al. 2003b, Lee et al. a,b):
  - Streamflow (at outlet and interior locations)
  - Gridded precipitation
  - Potential evaporation (PE)
  - In-situ soil moisture

SAC-HT allows translation of SAC states to soil moisture, and hence assimilation of soil moisture data into SAC



WTTO2

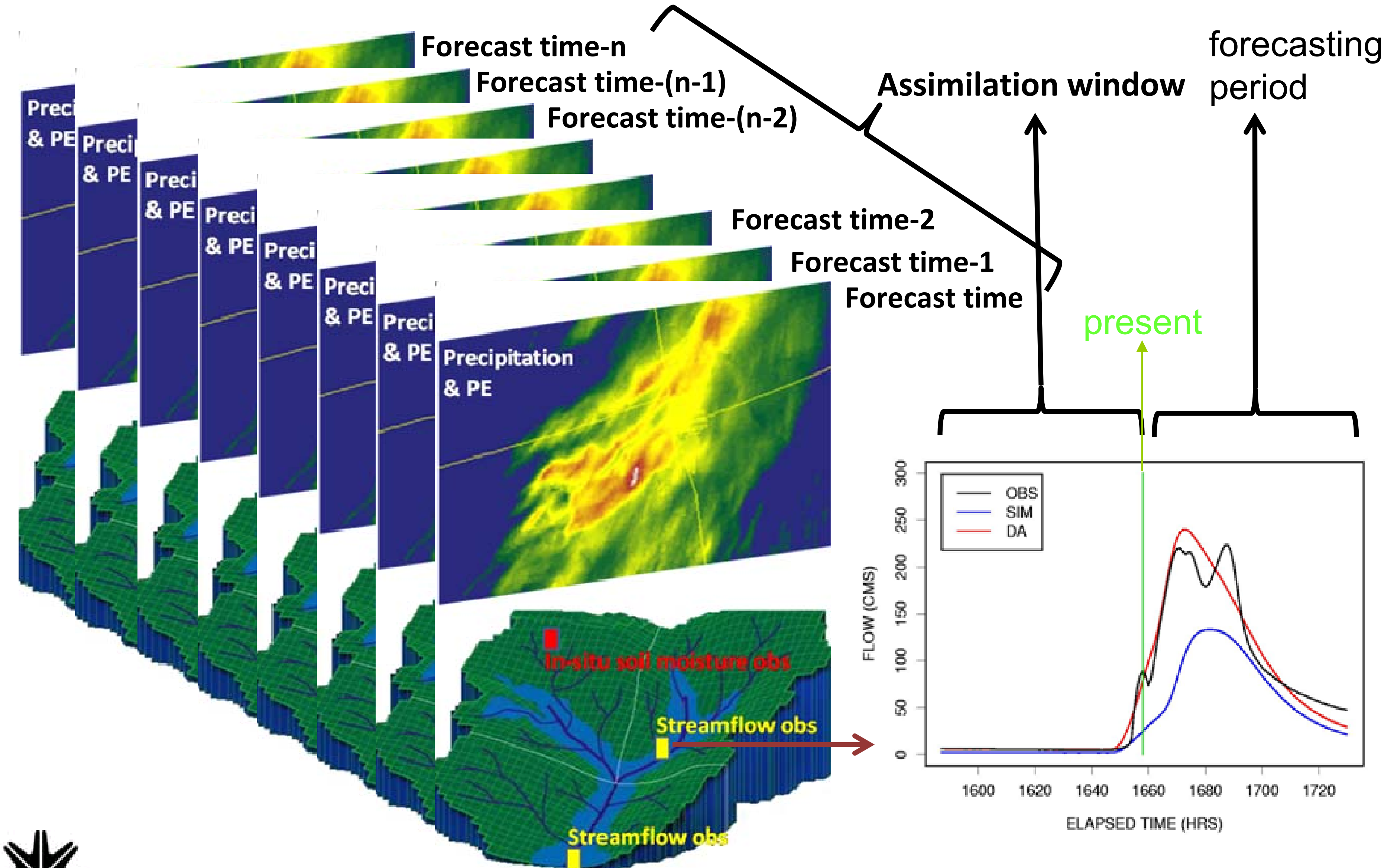
Technique used

- 4-dimensional variational assimilation, or 4DVAR
  - Arguably the most advanced data assimilation (DA) technique used in ***operational*** weather forecasting today
  - More amenable to **forecaster control** than ensemble Kalman filter/smoother (Evensen 1994, Evensen and van Leeuwen 2000)
  - Amenable to **ensemble DA** via maximum likelihood ensemble filter (MLEF) (Zupanski 2005)

What does 4DVAR do?

- Given all available data, the model(s) and the prescribed uncertainties for them, adjust the selected variables (e.g. the model states) such that the model results best fit the data
  - Under user-prescribed criterion (usually minimization of mean square errors)
  - Necessarily model-dynamically consistent
  - Not unlike what a human forecaster may do
  - As in any curve fitting, subject to over-fitting (too large a degree of freedom) and under-fitting (too small a degree of freedom)

What does 4DVAR do? (cont.)



How does 4DVAR work?

Adjust model states, and observed precipitation and PE so that the model-simulated flow is sufficiently close to the observed

$$
\begin{aligned}
\textit{Minimize} \quad J_k = & \frac{1}{2}[\mathbf{Z}_q - \mathbf{H}_{qq}(\mathbf{X}_{s,k-l}, \mathbf{X}_p, \mathbf{X}_e)]^T \mathbf{R}_{qq}^{-1} \\
& [\mathbf{Z}_q - \mathbf{H}_{qq}(\mathbf{X}_{s,k-l}, \mathbf{X}_p, \mathbf{X}_e)] \\
& + \frac{1}{2}[\mathbf{Z}_\theta - \mathbf{H}_{\theta\theta}(\mathbf{X}_{s,k-l}, \mathbf{X}_p, \mathbf{X}_e)]^T \mathbf{R}_{\theta\theta}^{-1} \\
& [\mathbf{Z}_\theta - \mathbf{H}_{\theta\theta}(\mathbf{X}_{s,k-l}, \mathbf{X}_p, \mathbf{X}_e)] \\
& + \frac{1}{2}[\mathbf{Z}_p - \mathbf{H}_{pp}\mathbf{X}_p]^T \mathbf{R}_{pp}^{-1}[\mathbf{Z}_p - \mathbf{H}_{pp}\mathbf{X}_p] \\
& + \frac{1}{2}[\mathbf{Z}_e - \mathbf{H}_{ee}\mathbf{X}_e]^T \mathbf{R}_{ee}^{-1}[\mathbf{Z}_e - \mathbf{H}_{ee}\mathbf{X}_e] \\
& + \frac{1}{2}[\mathbf{Z}_b - \mathbf{H}_{bb}\mathbf{X}_{s,k-l}]^T \mathbf{R}_{bb}^{-1}[\mathbf{Z}_b - \mathbf{H}_{bb}\mathbf{X}_{s,k-l}]
\end{aligned}
$$

How good is my streamflow data?

How good is my soil moisture data?

How good is my precipitation data?

How good is my potential evaporation (PE) data?

What do I know about the initial soil moisture states?

How good is my model?

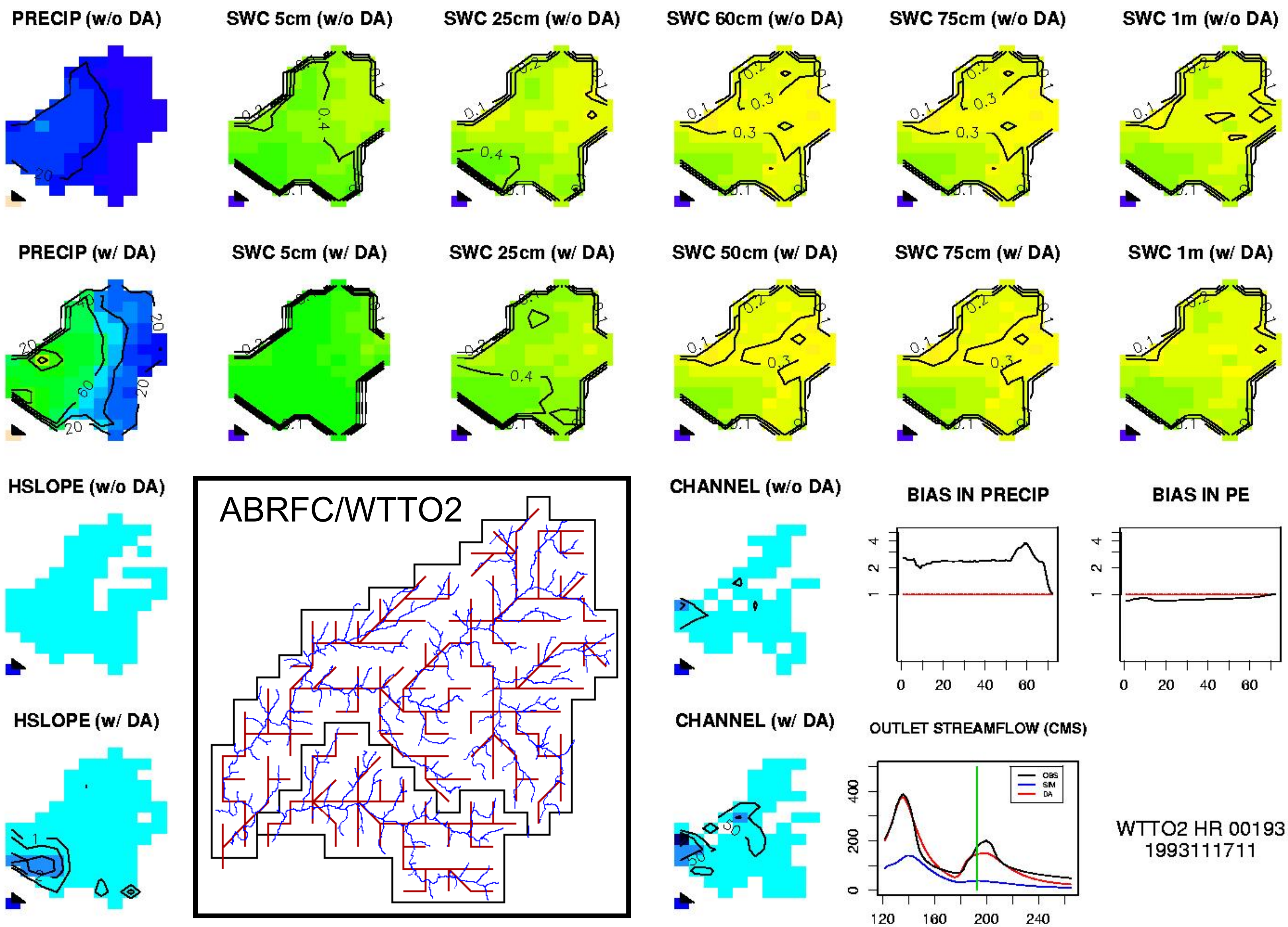
$$\textit{subject to} \quad \mathbf{X}_{s,j} = \mathbf{F}(\mathbf{X}_{s,j-1}, \mathbf{X}_{p,j}, \mathbf{X}_{e,j}), \quad j = k-l+1, \ldots, k$$

Whatever adjustments I am making must abide by the model dynamics

$$X_{s,i}^{\min} \le X_{s,k,i} \le X_{s,i}^{\max}, \quad j = k-l, \ldots, k; \quad i = 1, \ldots, 6$$

The adjustments must be within physically realistic bounds

4DVAR


PRECIP (w/o DA)
SWC 5cm (w/o DA)
SWC 25cm (w/o DA)
SWC 60cm (w/o DA)
SWC 75cm (w/o DA)
SWC 1m (w/o DA)
PRECIP (w/ DA)
SWC 5cm (w/ DA)
SWC 25cm (w/ DA)
SWC 50cm (w/ DA)
SWC 75cm (w/ DA)
SWC 1m (w/ DA)
HSLOPE (w/o DA)
ABRFC/WTTO2
CHANNEL (w/o DA)
BIAS IN PRECIP
BIAS IN PE
HSLOPE (w/ DA)
CHANNEL (w/ DA)
OUTLET STREAMFLOW (CMS)
OBS
SIM
DA
WTTO2 HR 00193
1993111711

Questions investigated

- What is the value of assimilating **streamflow** (outlet, interior) data for improved accuracy in monitoring (analysis) and prediction of streamflow and soil moisture?
  - According to uncertainty in the initial soil moisture conditions

- What is the value of assimilating in-situ **soil moisture** data?

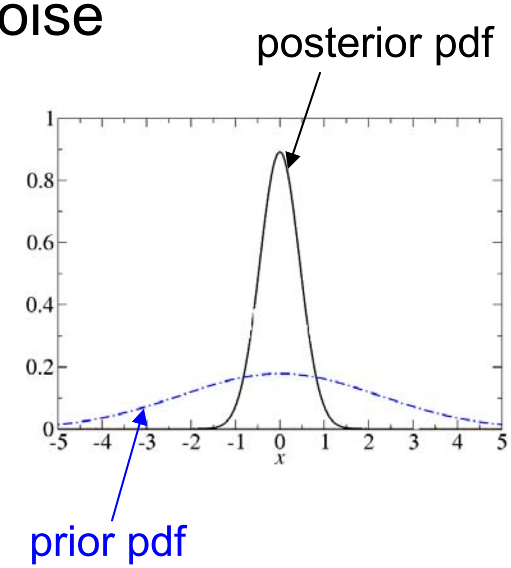
- What is the value of assimilating gridded **precipitation** data

Approach

- Carry out synthetic and real-world experiments
- Why synthetic experiments?
  - In reality, truth is unknown and many uncertainties complicate understanding and interpretation
- In synthetic experiments:
  - Truth is known
    - ✍Easier to evaluate DA performance
  - Can separate different uncertainty sources
    - ✍Initial condition uncertainty (ICU)
    - ✍Precipitation uncertainty (PU)
    - ✍Other Observational uncertainty
    - ✍Model structural and parametric uncertainty
  - More likely to gain and advance understanding on hydrologic DA with distributed models

Synthetic experiments

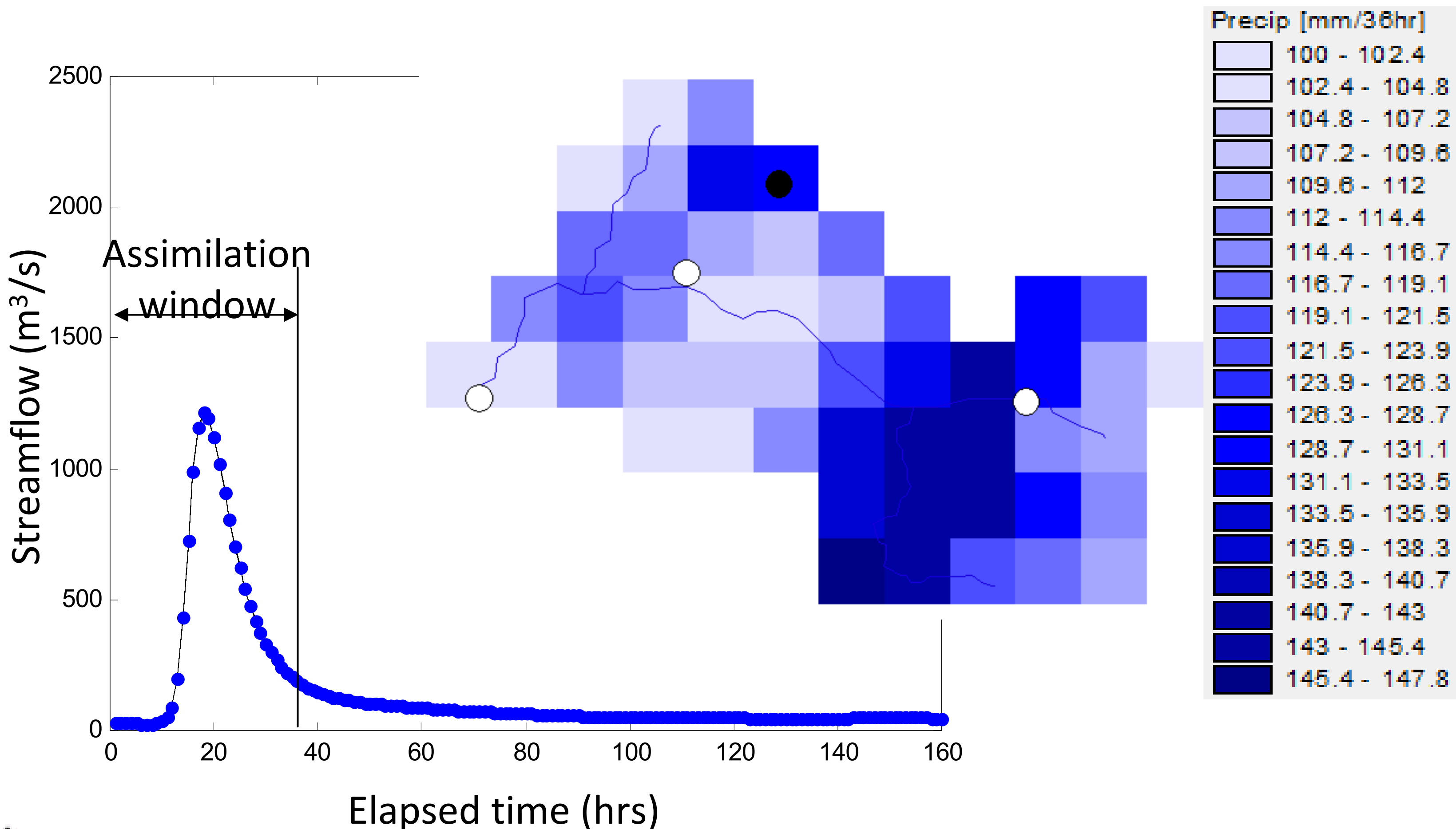
- Methodology
  - Assume "true" initial soil moisture states (IC), streamflow (Q) and soil moisture (S) observations, and observed precipitation (P)
  - Perturb with low, medium and high levels of noise
    - IC, Q, S (Experiment 1, Lee et al.[a])
    - IC, Q, S, P (Experiment 2, Lee et al.[b])
  - Assimilate the observations via 4DVAR
  - Repeat above 2 steps to generate ensembles
  - Assess the quality of posterior ensembles
  - Monte-Carlo type of 4DVAR



[a,b]in preparation

Study basin - Eldon (795 km^2)

Case studied
2000/ 6/ 22/ 18Z - 2000/ 6/ 24/ 6Z



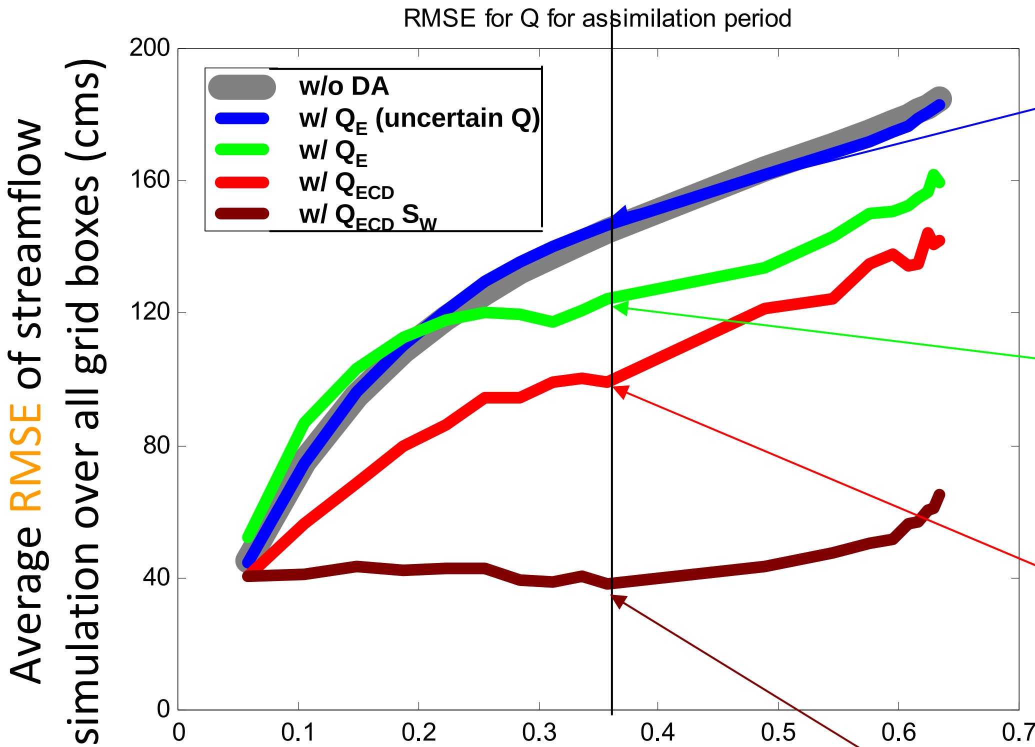
Synthetic Experiment I:

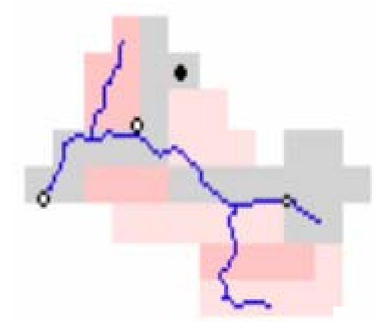
Sensitivity of DA to initial condition uncertainty (ICU) and observational uncertainties

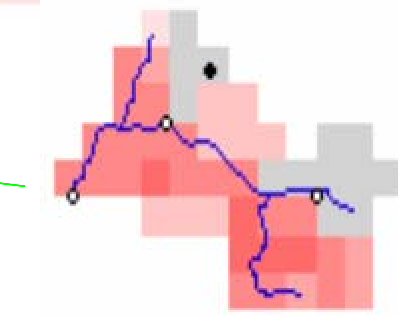
Synthetic Experiment I: Results

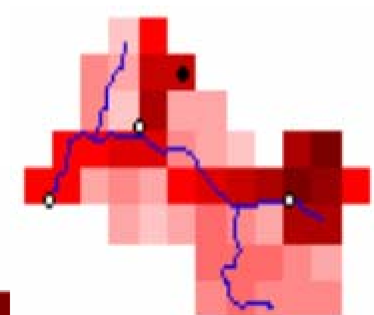
SS(Skill Score)
$=1-MSE_{DA}/MSE$

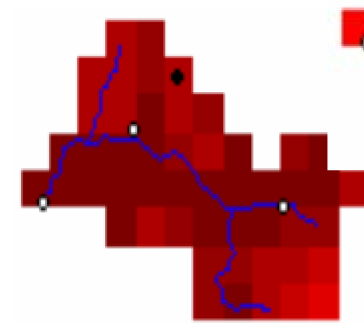
SS=1 perfect
SS>0 skillful
SS≤0 no skill



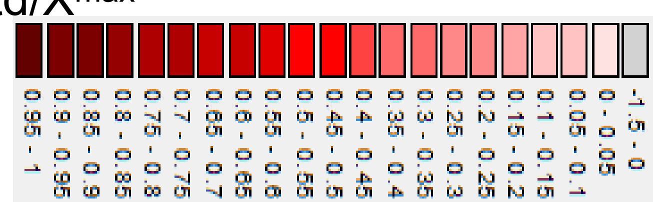




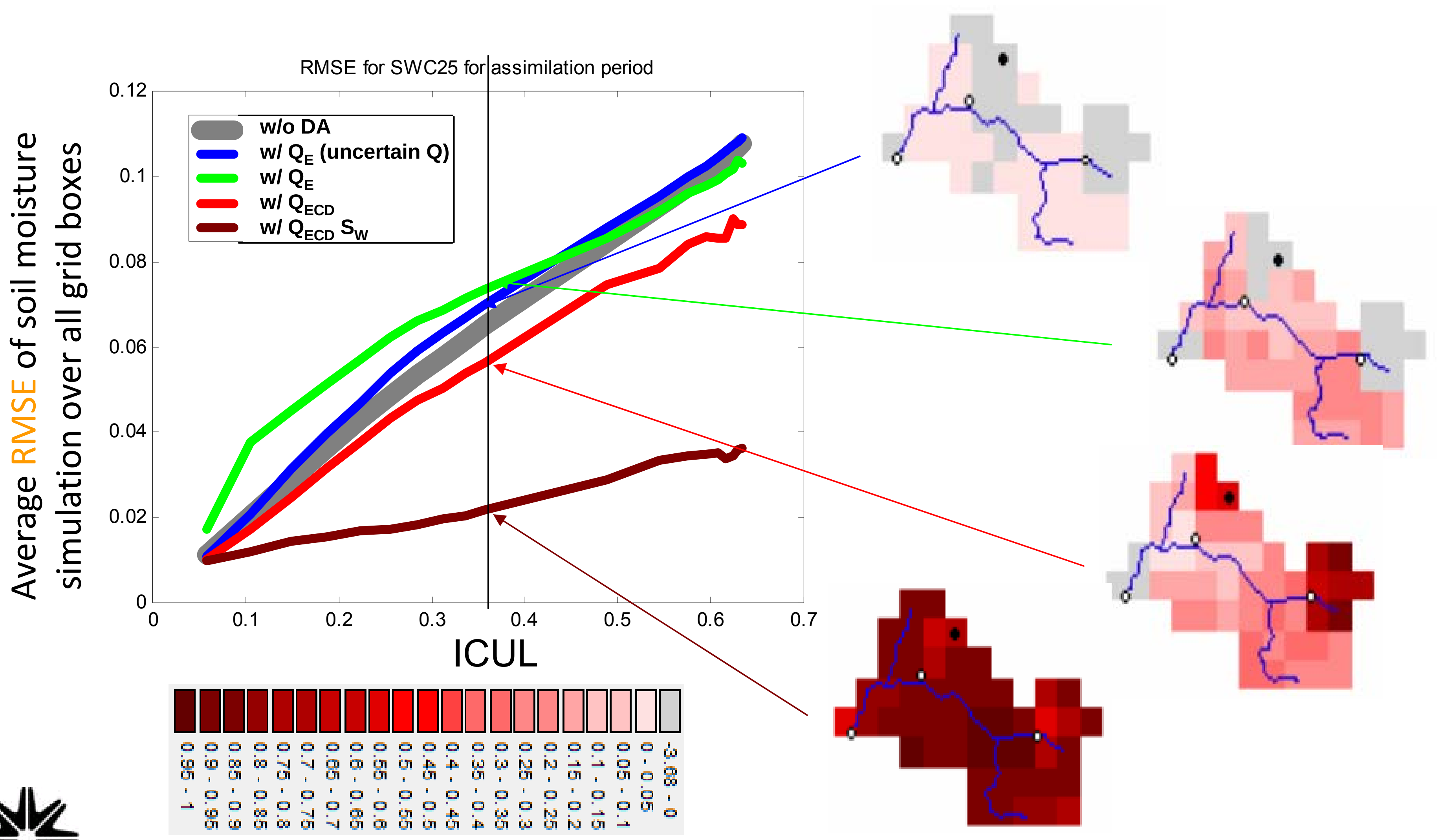




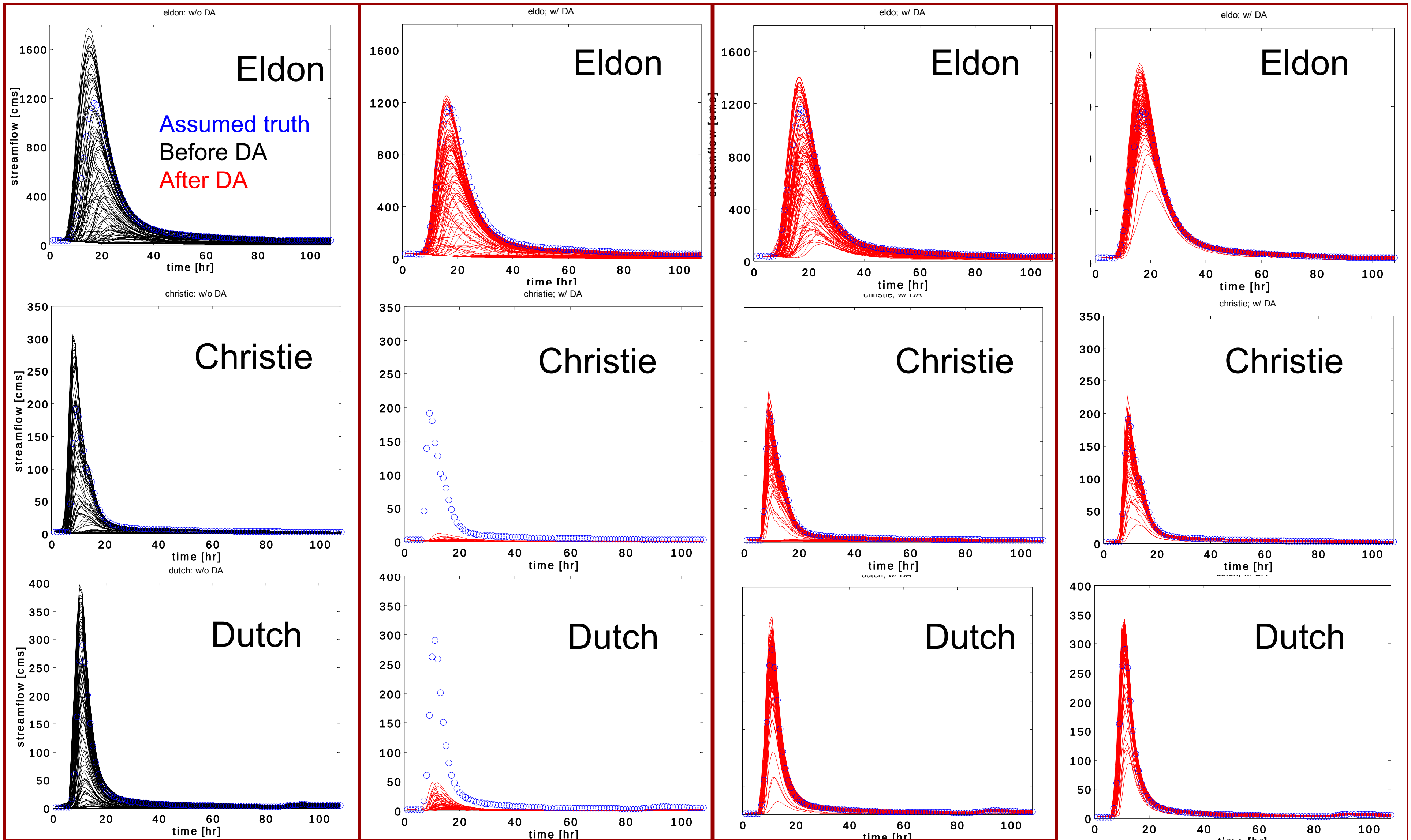
ICUL (initial condition uncertainty level):
$=std/\overline{X^{max}}$



Synthetic Experiment I: Results (cont.)



Streamflow results for assimilating accurate streamflow & soil moisture obs under uncertain IC



w/o DA | w/ Q_E | w/ Q_{ECD} | w/ Q_{ECD} S_W

Synthetic Experiment II:
Sensitivity of DA to precipitation uncertainty (PU)

Precipitation Uncertainty Model

$P_k(u)=B_k\, O_k\,(u) + \sigma\, Z_k(u)$

$\ln B_k = a_1 \ln B_{k-1} + W_k$ (Smith and Krajewski 1991)
where
$P_k(u)$: perturbed rainfall at location u at hour k (mm)
$O_k(u)$: reference rainfall at hour k (mm)
B_k: mean field bias at hour k
$Z_k(u)$: spatially-correlated standard normal random noise

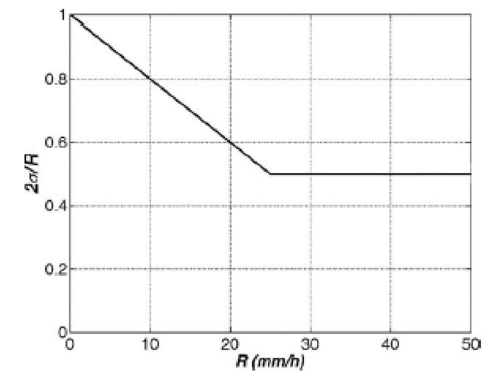
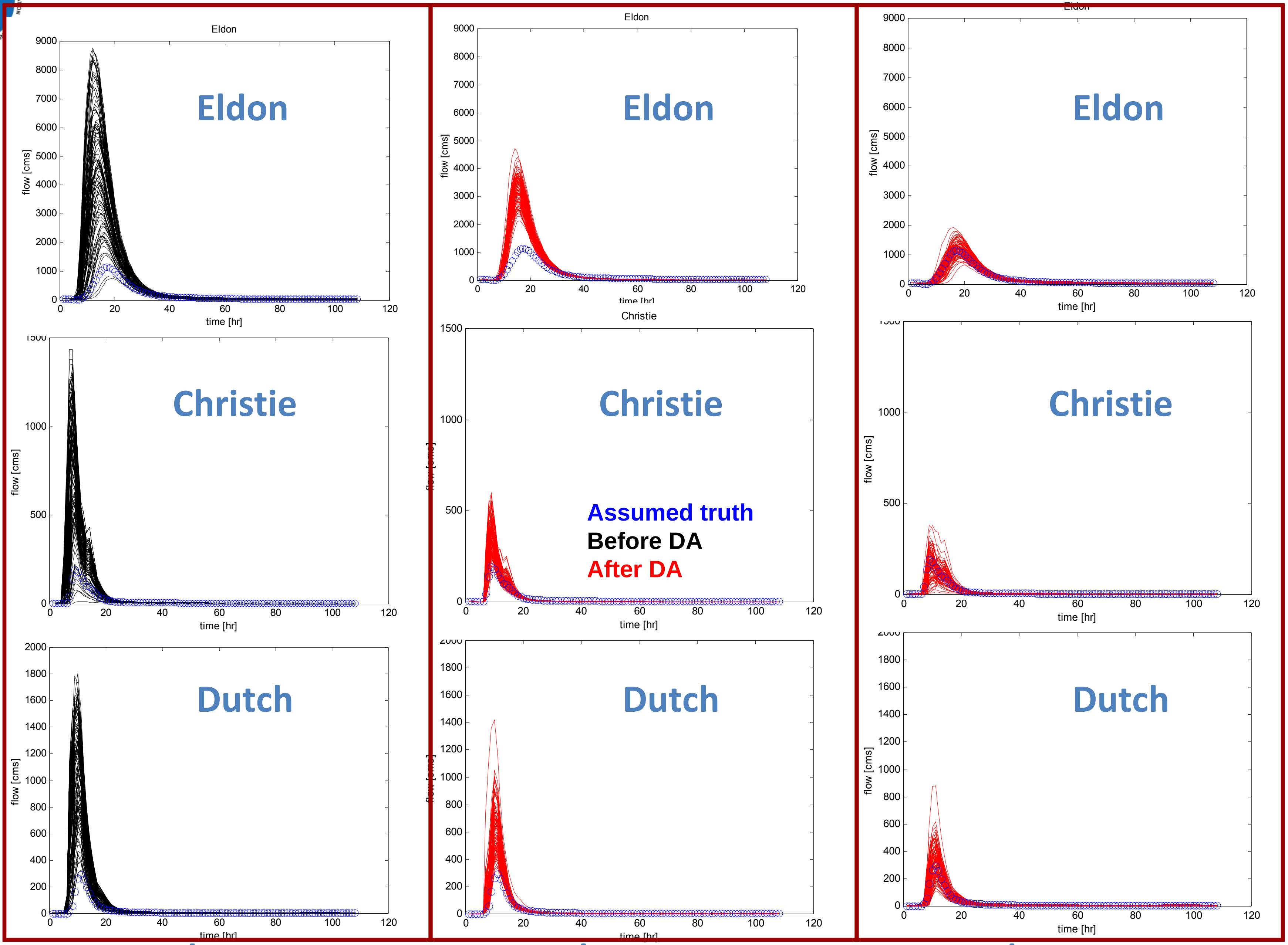


Figure 3 Standard deviation of radar-rainfall pixel error as a function of pixel rainfall value.

$2\sigma/O_k(u)=1-0.02\, O_k(u)$ if $O_k(u) \le 25.4$ (mm)
$2\sigma/O_k(u)=0.5$ if $O_k(u) > 25.4$ (mm)
(Carpenter and Georgakakos 2006)

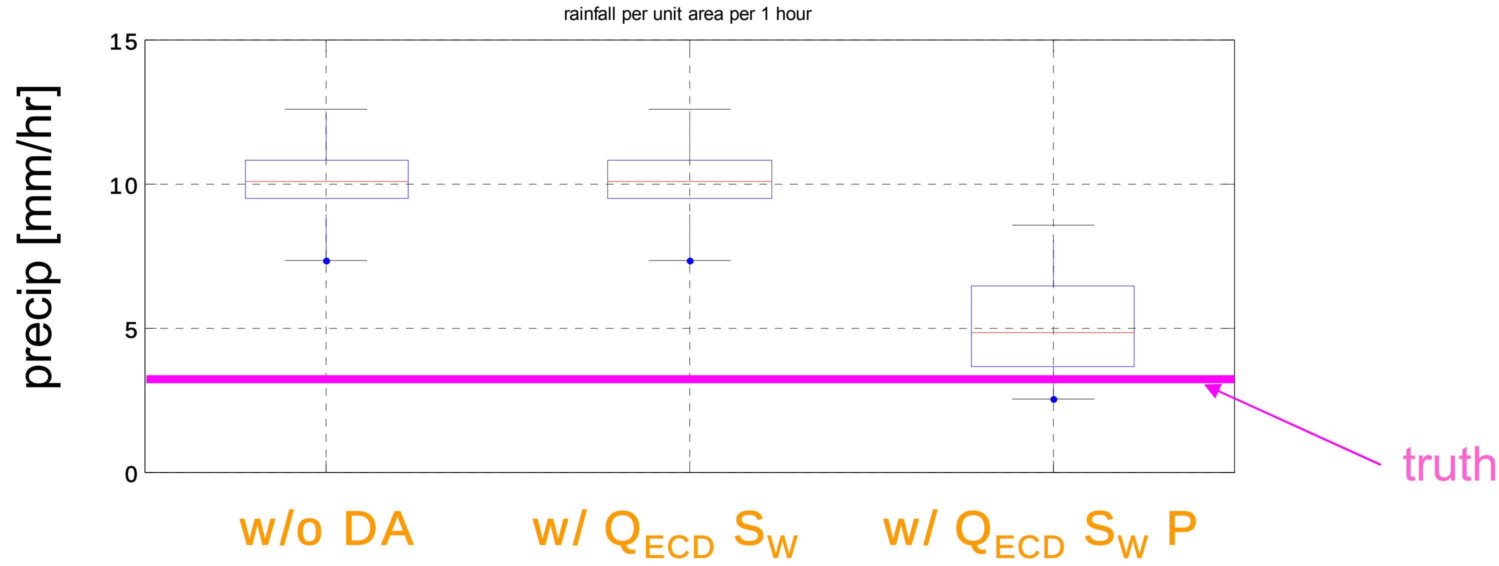
where
σ: rainfall amount-dependent standard deviation of the noise

Impact of additionally assimilating precipitation on streamflow prediction

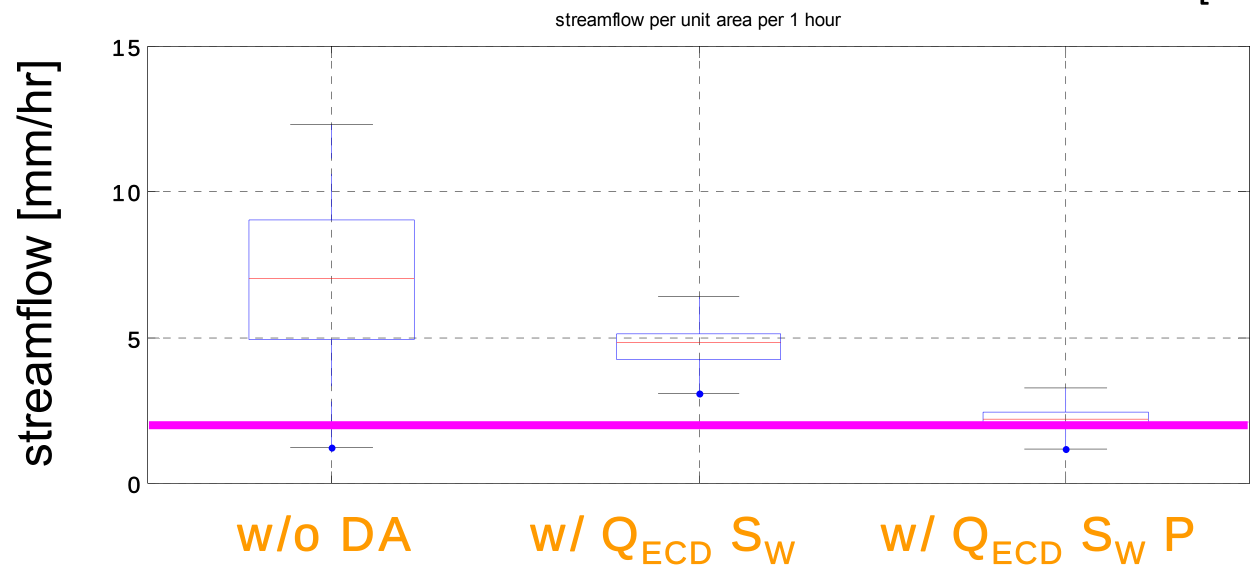


w/o DA | w/ Q_{ECD} S_W | w/ Q_{ECD} S_W P

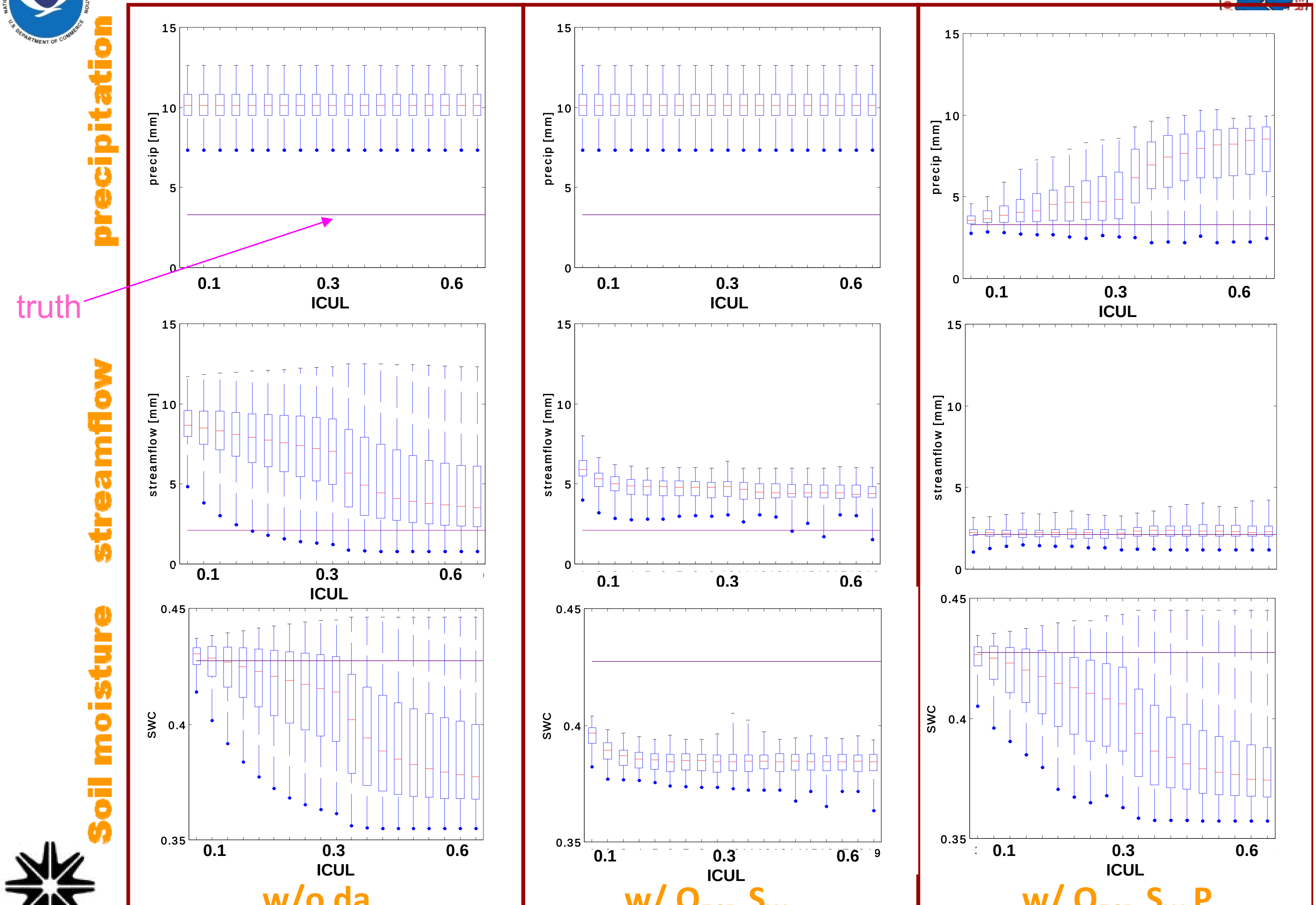
Total precipitation over the assimilation window [mm/hr] (input)



Total streamflow at the outlet over the assimilation window [mm/hr] (output)

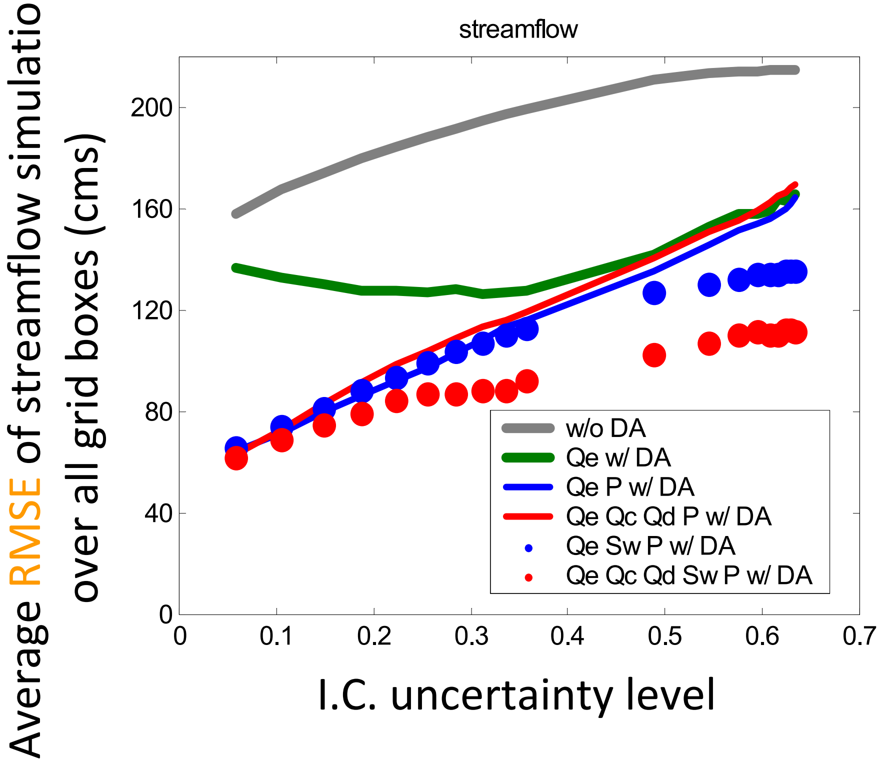


Water balance !!!

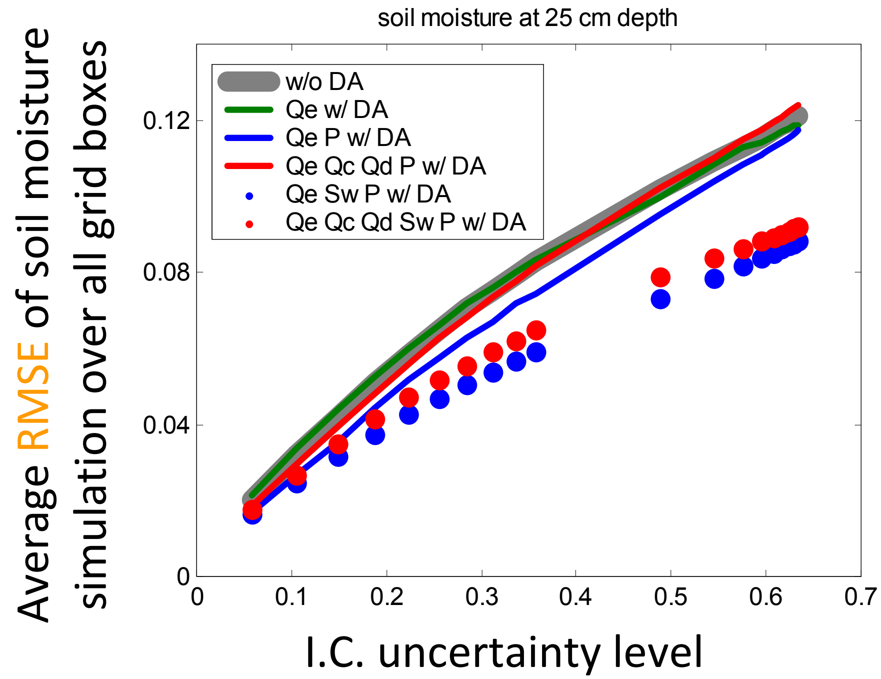


Impact of additionally assimilating precipitation to streamflow and soil moisture simulation

streamflow



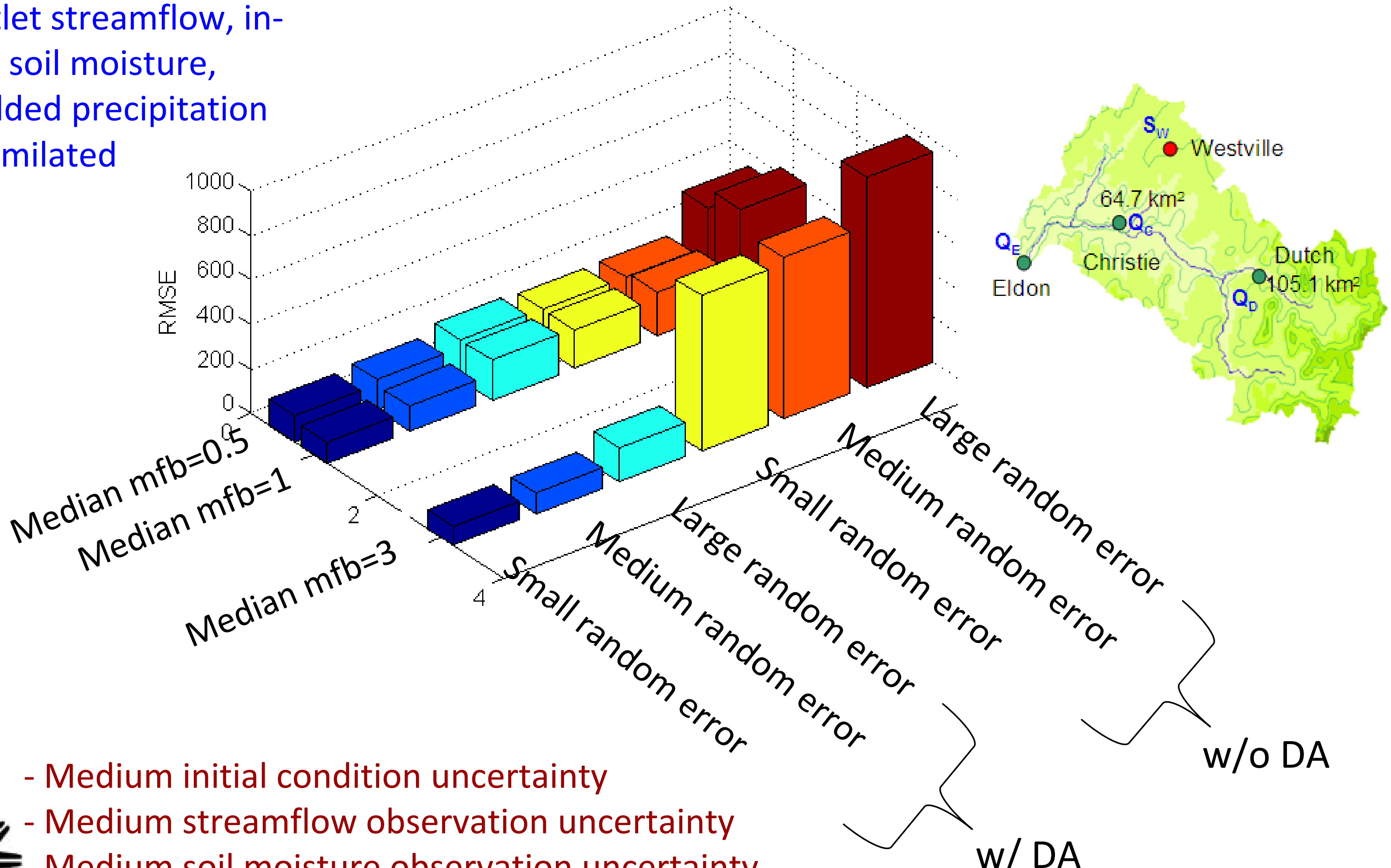
soil moisture



- Large precipitation uncertainty
- Medium streamflow observation uncertainty
- Medium soil moisture observation uncertainty

Impact of mean field bias (B'=B?α) and noise ($\sigma'=a_3\sigma$) to streamflow simulation via DA

Outlet streamflow, in-situ soil moisture, gridded precipitation assimilated



- Medium initial condition uncertainty
- Medium streamflow observation uncertainty
- Medium soil moisture observation uncertainty

Real-World Experiment

Real-World Experiment: Questions

- The models are never perfect
  - Structural errors
  - Parametric errors

- Soil moisture is seldom observed directly, and never at the model grid scale

- How to account for these uncertainties?

- How do these uncertainties impact DA?

Experiment Design

- Setup
  - Assimilation window: 36 hrs
  - Error variance for precipitation: sample variance
  - Error variance for streamflow: sensitivity analysis
  - Error variance for soil moisture: data analysis & model simulation

- Data
  - Soil moisture: 1997 – 2000 (Oklahoma Mesonet, Brock et al. 1995)
  - Streamflow: 1997 – 2000
  - Precipitation: ABRFC-produced operational multisensor QPE

Acknowledgment: We would like to thank the Oklahoma Climatological Survey for allowing the use of the Oklahoma Mesonet soil moisture data.

Uncertainties associated with in-situ soil moisture obs (OK Mesonet)

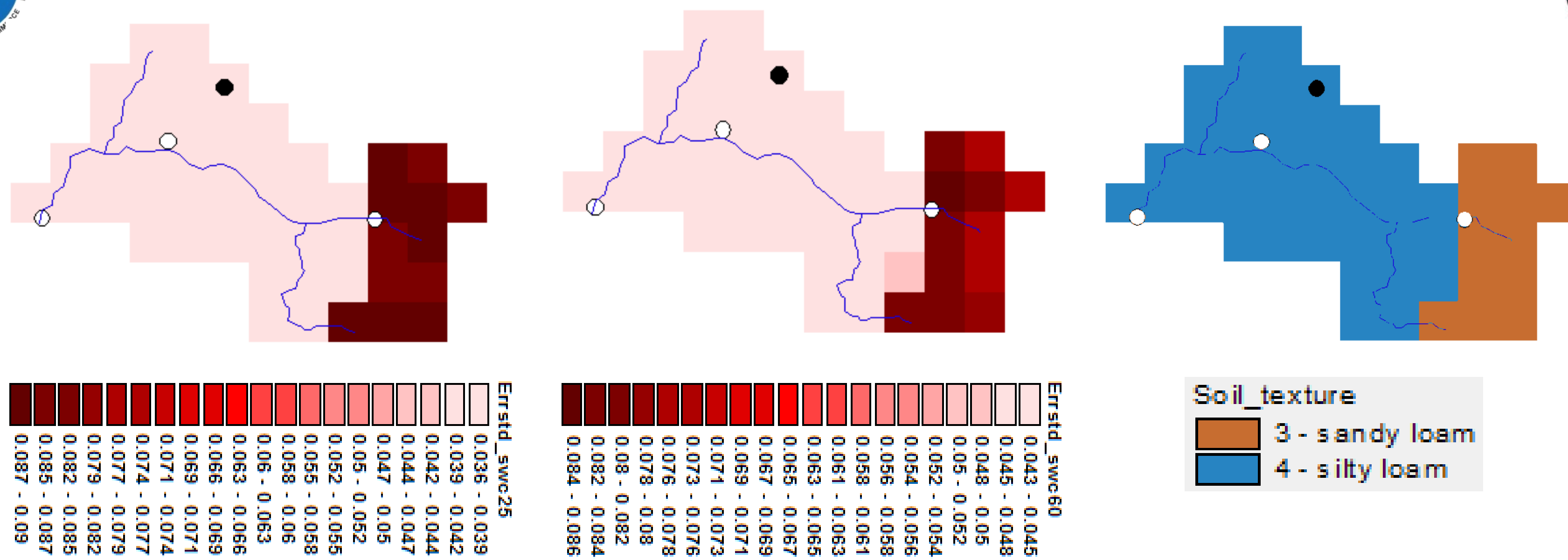
- Device error:
- Soil moisture sensor error (CSI 229-L) (e1)
- Numerical precision error (e2)
- Device limit to measure extreme values (e3)

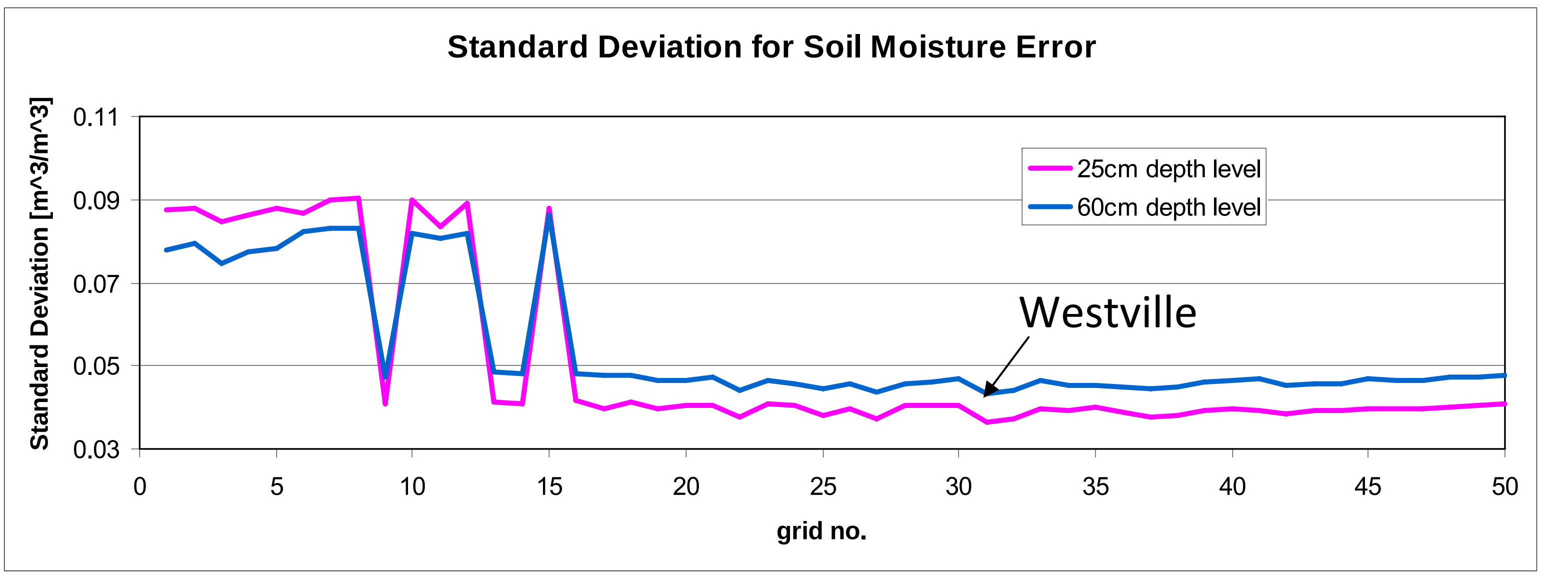
- Scaling (e4)
  - pt to HRAP scale error estimated by cdf matching technique
  - bias correction is done by cdf matching

- Spatial variability (e5)

- Overall error variances (=e1+e2+e3+e4+e5)

Estimated standard deviation for soil moisture error





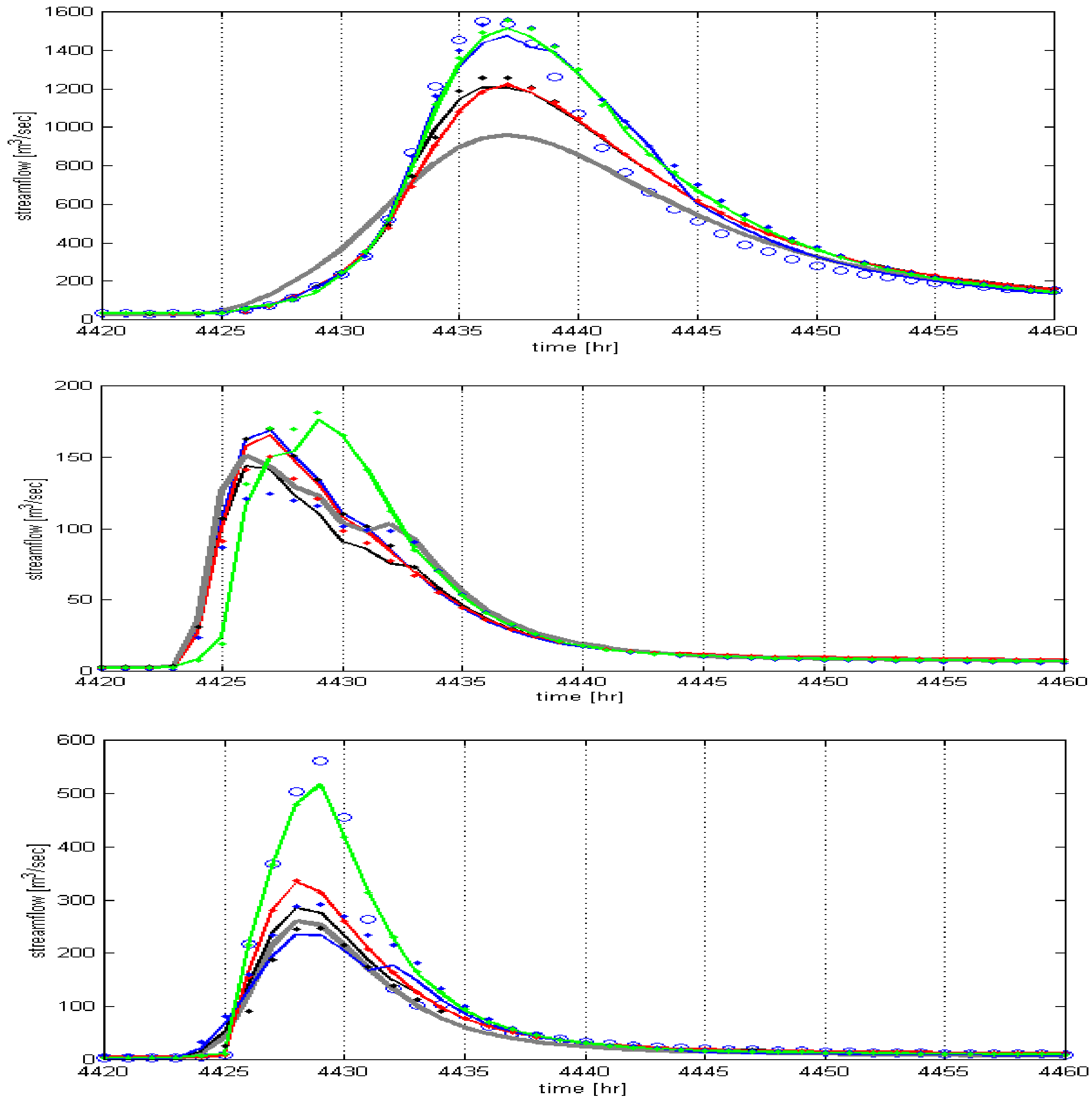
≤ 0.05 m^3/m^3 (Walker and Houser, 2004) is useful for data assimilation

Streamflow time series for the biggest event in yr 2000: lead time = 0 hr

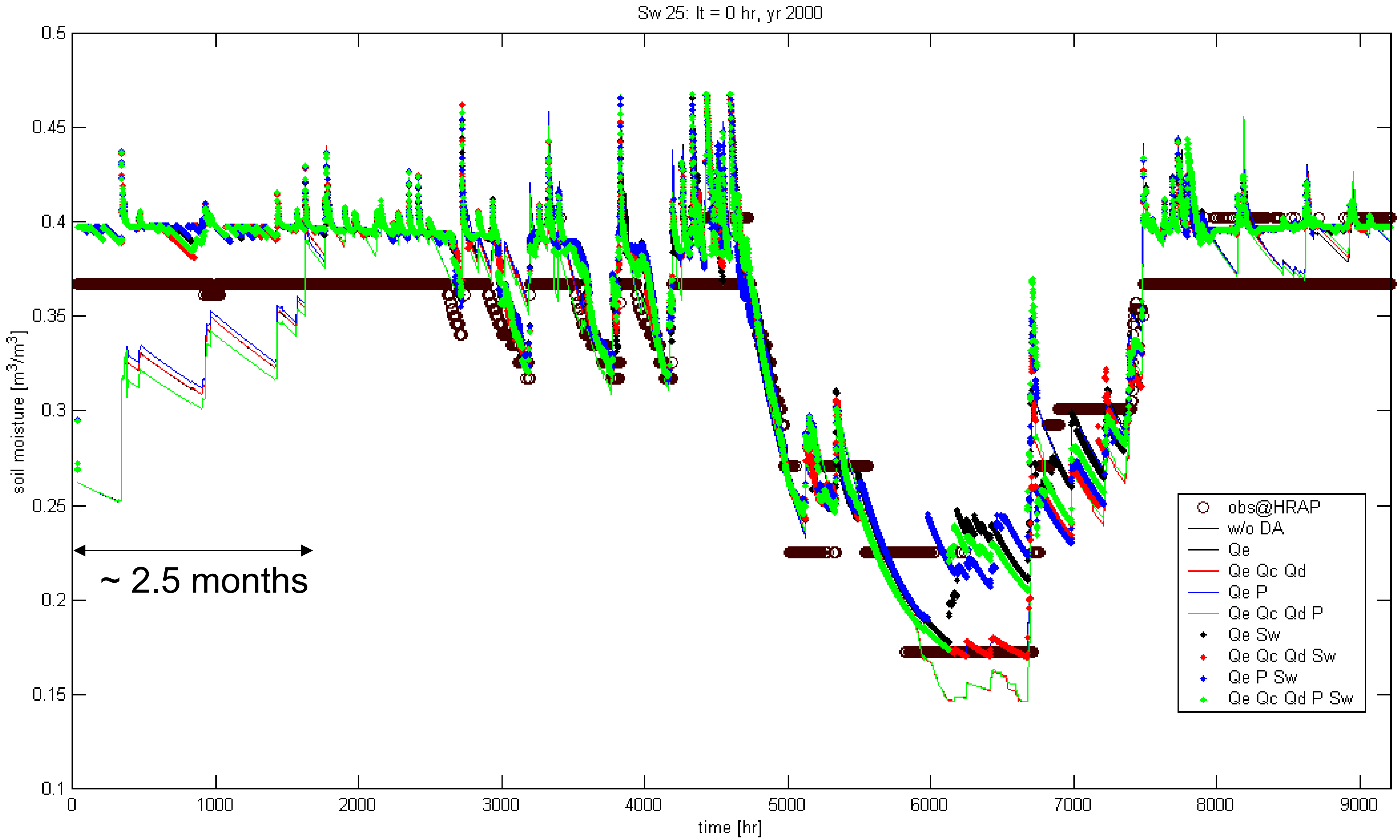
Eldon

Christie

Dutch

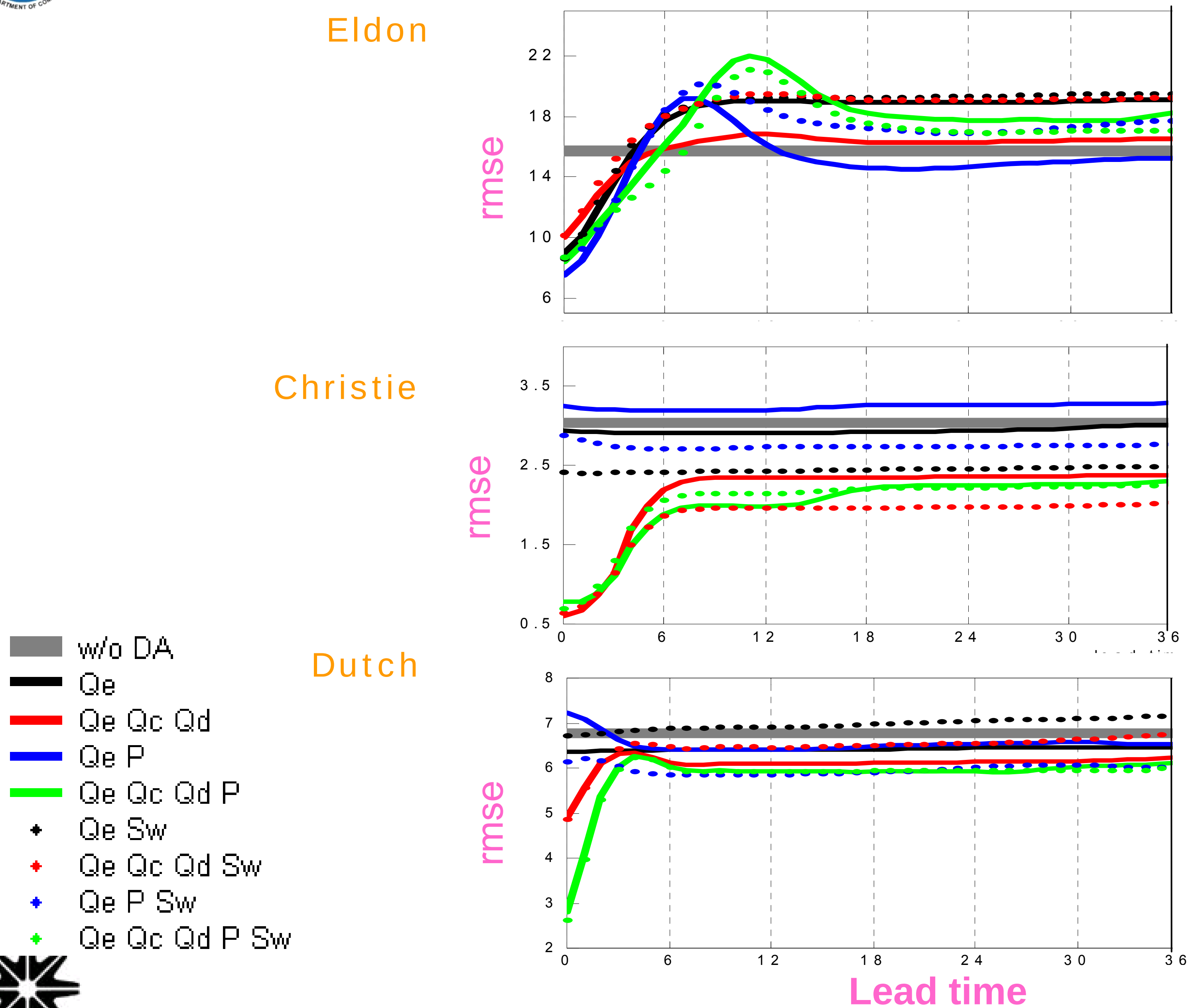


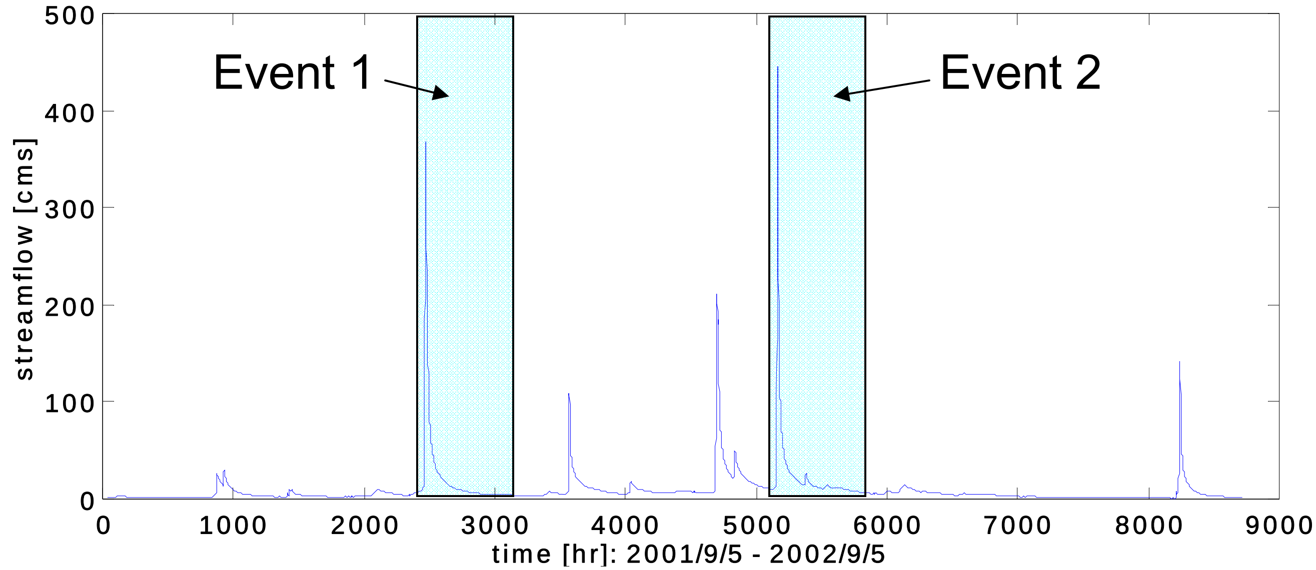
Hourly Soil Moisture at Westville for yr 2000



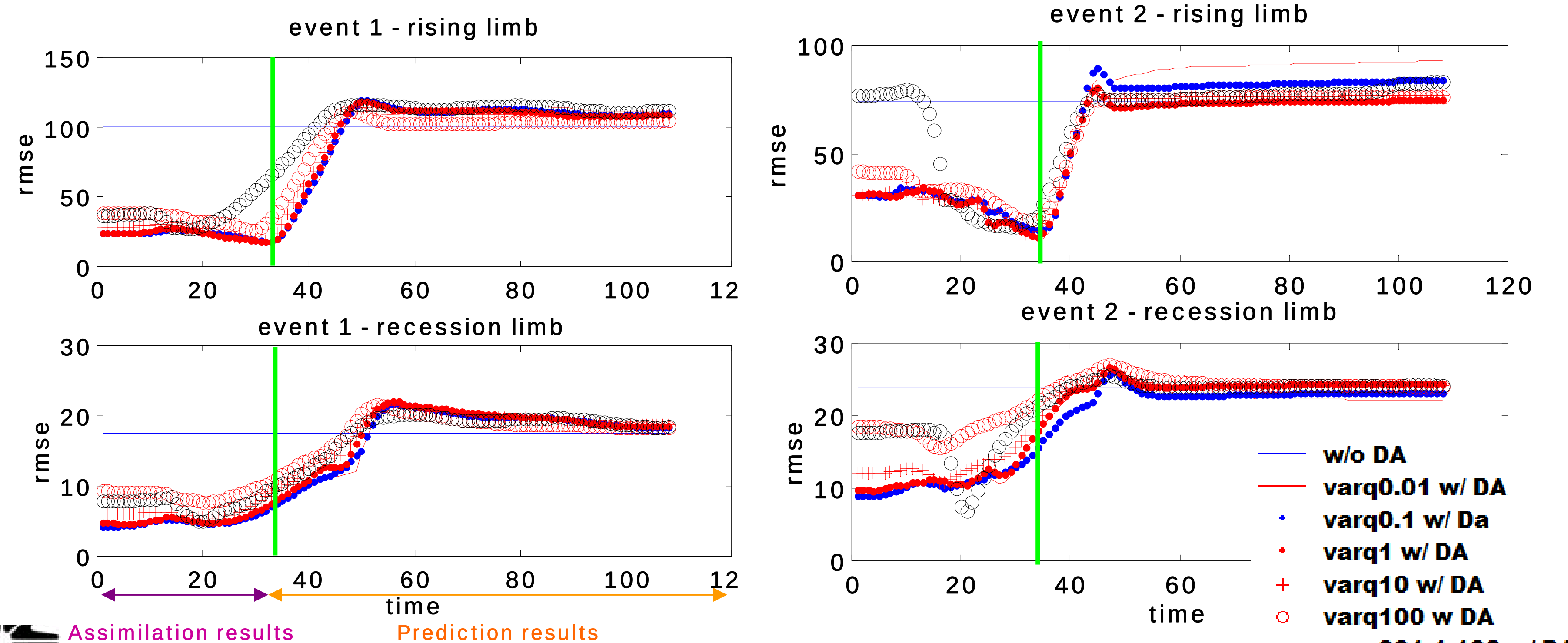
S_W at 25cm

rmse vs. lead time for streamflow for yr 2000





Event 1
Event 2
streamflow [cms]
time [hr]: 2001/9/5 - 2002/9/5
0
100
200
300
400
500
1000
2000
3000
4000
5000
6000
7000
8000
9000



event 1 - rising limb
event 2 - rising limb
event 1 - recession limb
event 2 - recession limb
rmse
time
Assimilation results
Prediction results
w/o DA
varq0.01 w/ DA
varq0.1 w/ Da
varq1 w/ DA
varq10 w/ DA
varq100 w DA
varq001 1 100 w/ DA

Conclusions

- Assimilating streamflow, in-situ soil moisture and QPE data in real time has large potential value for high-resolution analysis and prediction of streamflow and soil moisture
- However, its potency is sensitive to the uncertainty in the initial soil moisture conditions, the quality of observations and the goodness of the models (and their parameters) used
- It is seen that:
  - If the initial conditions are highly uncertain, soil moisture observations have larger positive impact than streamflow observations
  - If the initial conditions are less uncertain, accurate streamflow observations have larger positive impact than soil moisture observations

Conclusions (cont.)

- Assimilating QPE, in addition to streamflow and soil moisture observations, improves water balance calculations
  – If precipitation uncertainty is not properly accounted for in DA, streamflow balance may be improved, but only at the expense of deteriorated soil moisture balance
- If there are large uncertainties in QPE and in the initial conditions, assimilating soil moisture observations has large positive impact on analysis and prediction of soil moisture and streamflow
- Assimilating streamflow observations at both the outlet and interior locations generally improves streamflow prediction at those locations
- Assimilating soil moisture observations have large positive impact on model soil moisture states on cold starts

Upshot of all this

- A prototype DA has been developed that is capable of assimilating streamflow, in-situ soil moisture and gridded QPE into SAC and kinematic wave routing models of HL-RDHM
- Results thus far are encouraging, and points out salient observational, scientific and practical issues to be addressed
- Gained much understanding on how the major sources of uncertainty impact the performance of DA and what the next steps are toward improving operational worthiness
- The immediate next step is to simplify the current prototype to avoid “overfitting” and reduce computational burden (ongoing – should also help forecaster control of the DA), and to evaluate performance for multiple basins (ongoing)
- The new prototype to be considered for integration with HL-RDHM in the CHPS/FEWS/XEFS environment

Next Steps

- Simplify the current prototype
  - Avoid overfitting, reduce amount of computation
- Further assess model errors and their impact
- Better understand in-situ soil moisture measurement (HMT/Robert Zamora)
- Assimilate satellite-derived soil moisture data (w/ NCEP/EMC)
  - Into SAC-HT via LIS
  - Assimilate satellite-aided model soil moisture fields into the prototype DA
- Develop 4DVAR into ensemble DA using, e.g., maximum likelihood ensemble filter

Thank you

Q&A, discussion

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NOAA
U.S. DEPARTMENT OF COMMERCE

NATIONAL WEATHER SERVICE

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NOAA
U.S. DEPARTMENT OF COMMERCE

NATIONAL WEATHER SERVICE

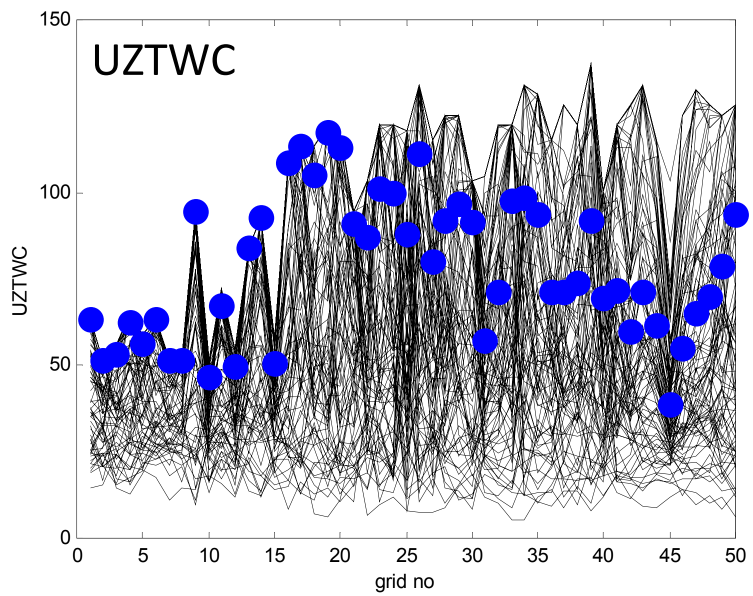
Appendix

Uncertainty model for initial SAC states

$X_k(0{:}k) = X^{max}\,[\exp(\eta_k)-1]+X^{true}$

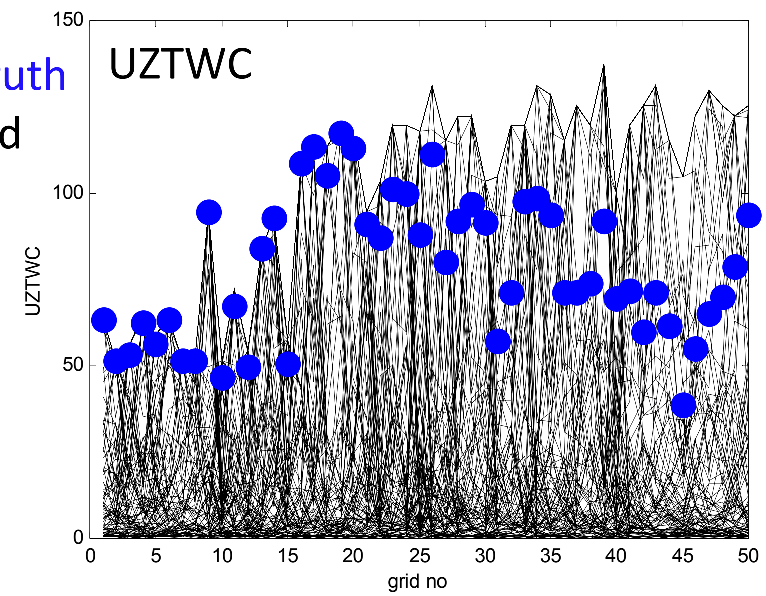
where $\eta_k = -0.5\ \ln[1+ (a_{IC}X^{max})^{\wedge}2] + \varepsilon_k\ [\ln(1+(a_{IC}X^{max})^{\wedge}2)]^{1/2}$,

ε_k ~ k-th spatially correlated N(0,1) random deviate



Assumed truth

Perturbed



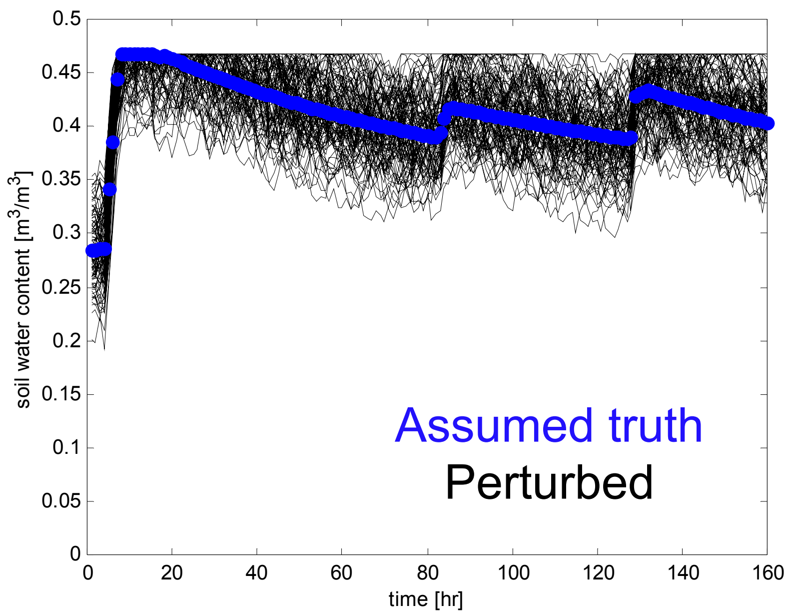
Uncertain IC
(a_{IC}=0.01)

Highly uncertain IC
(a_{IC}=0.1)

Uncertainty model for in-situ soil moisture obs

$Z_S(t:k) = Z_S(t) + a_S\, w(t:k)$

Where w(t:k) is k-th temporally correlated N(0,1) random deviate



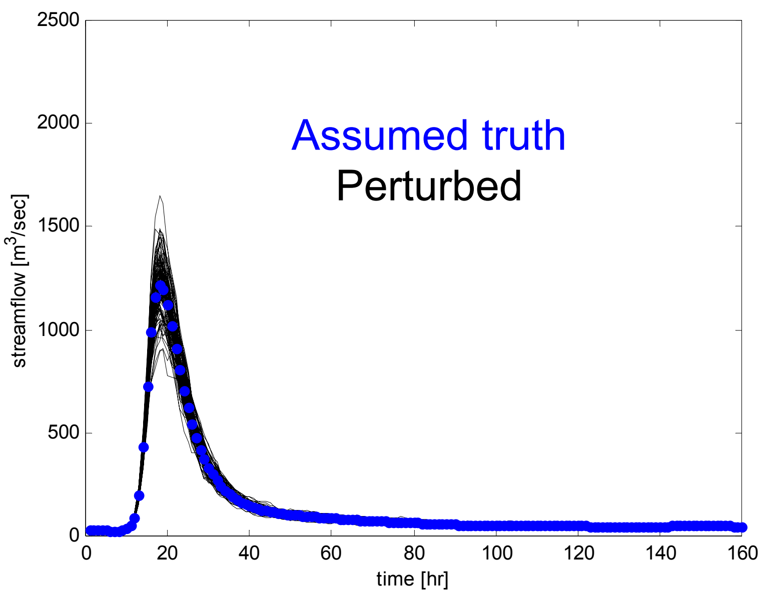
generated soil moisture obs (a_S=0.03)

≤ 0.05 m^3/m^3 (Walker and Houser, 2004) is useful for data assimilation

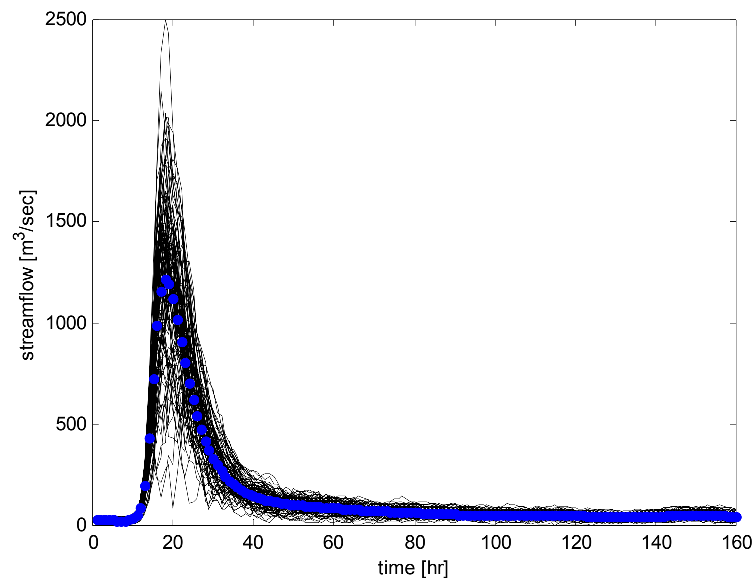
Uncertainty model for streamflow obs

$Q(t{:}k) = Q(t) + a_q\, Q(t)\, w(t{:}k)$

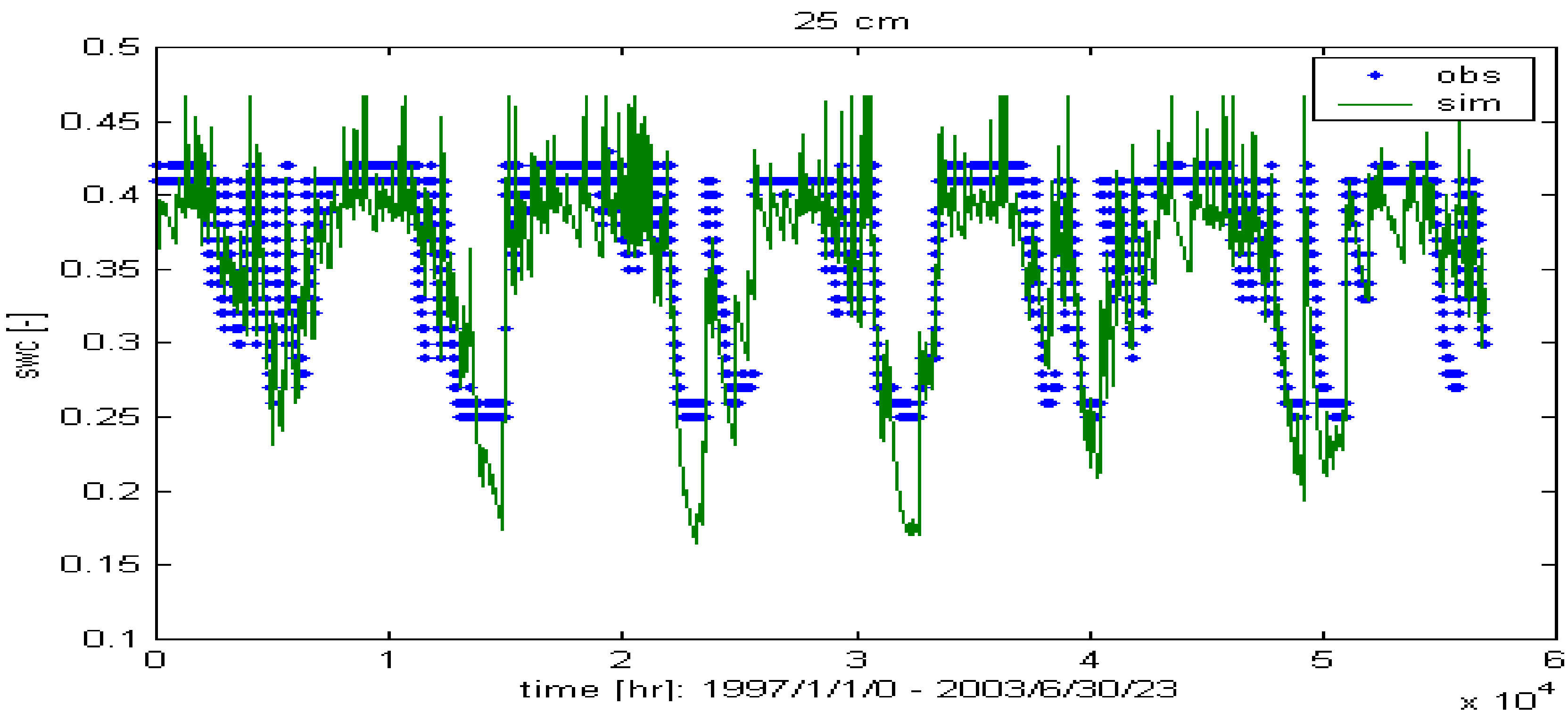
where w(t) is k-th temporally correlated N(0,1) random deviate



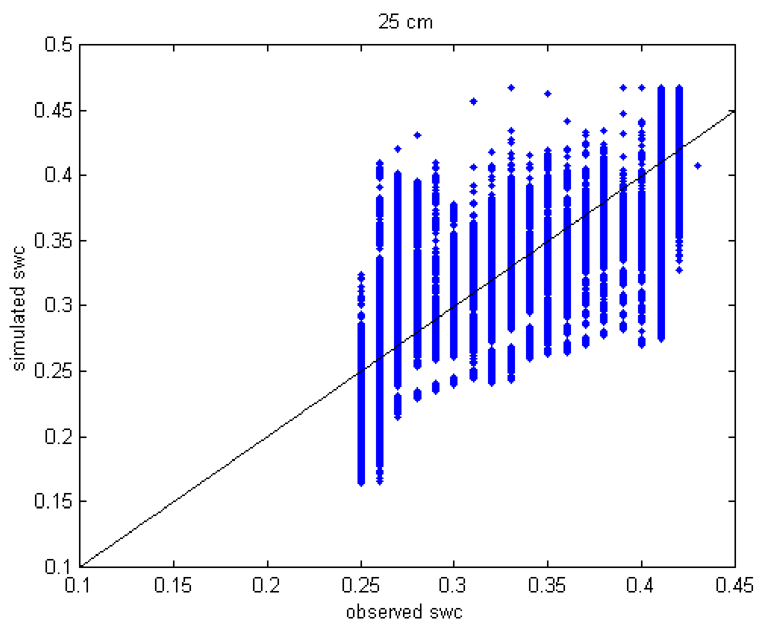
Accurate streamflow obs (aq=0.1)



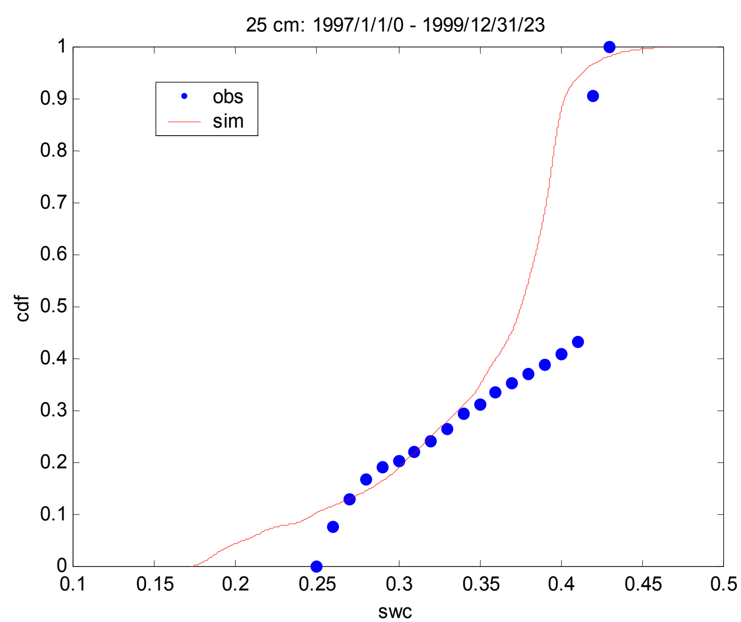
Uncertain streamflow obs (aq=0.3)



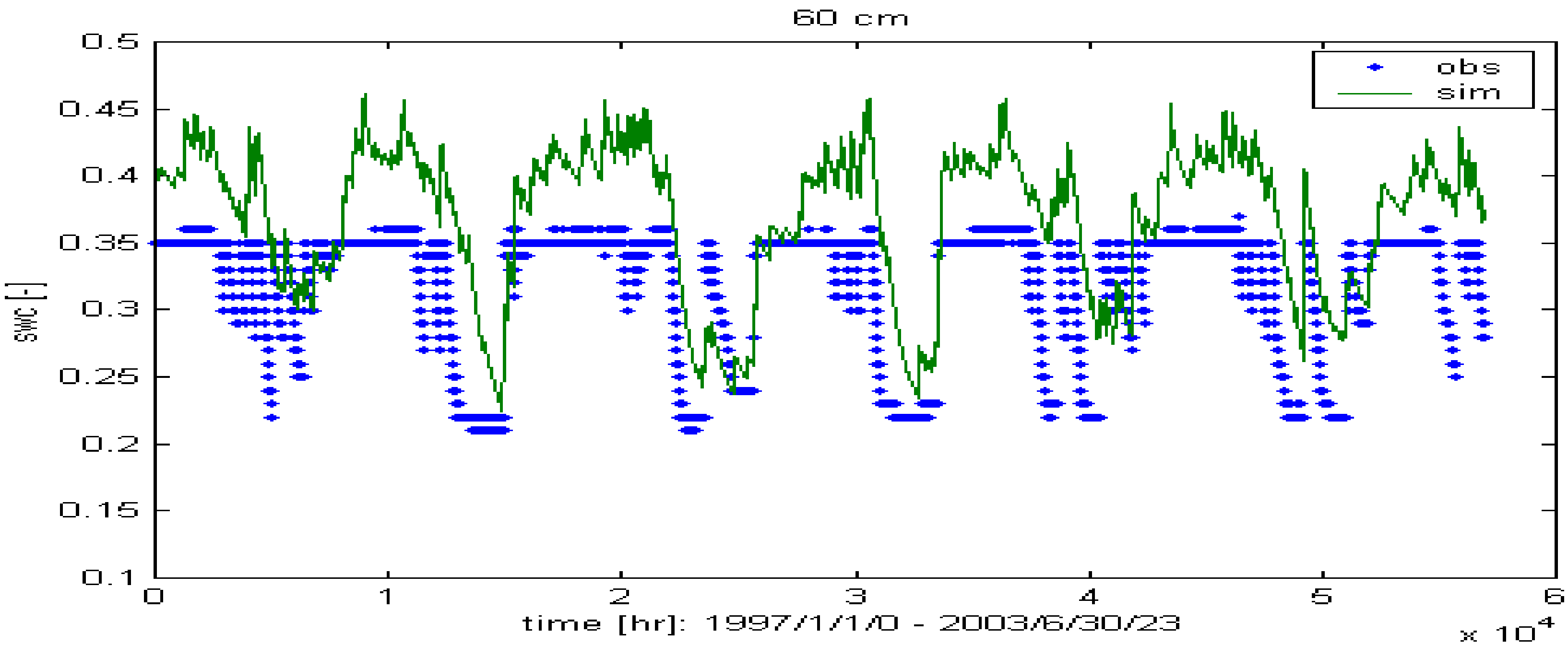
25 cm
obs
sim
SWC [-]
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0
1
2
3
4
5
6
time [hr]: 1997/1/1/0 - 2003/6/30/23
x 10^4

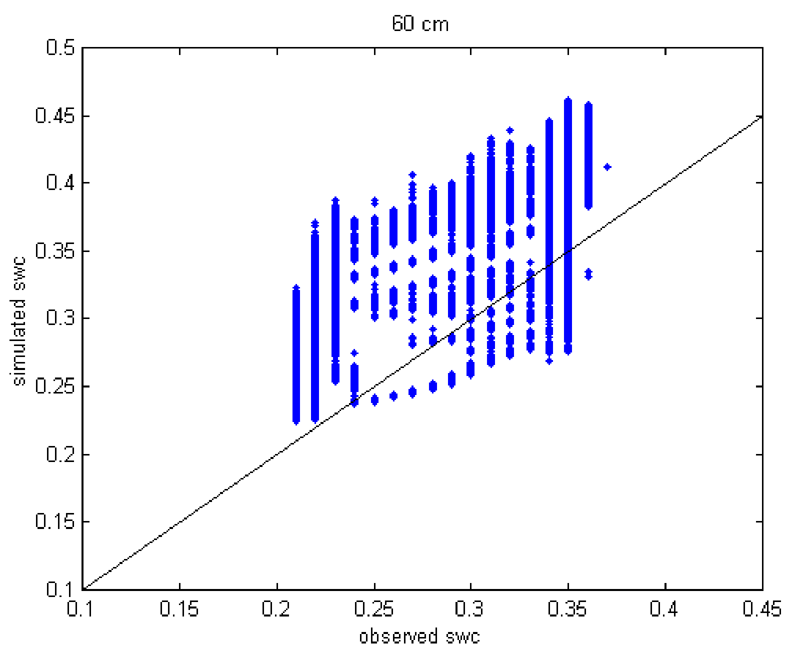


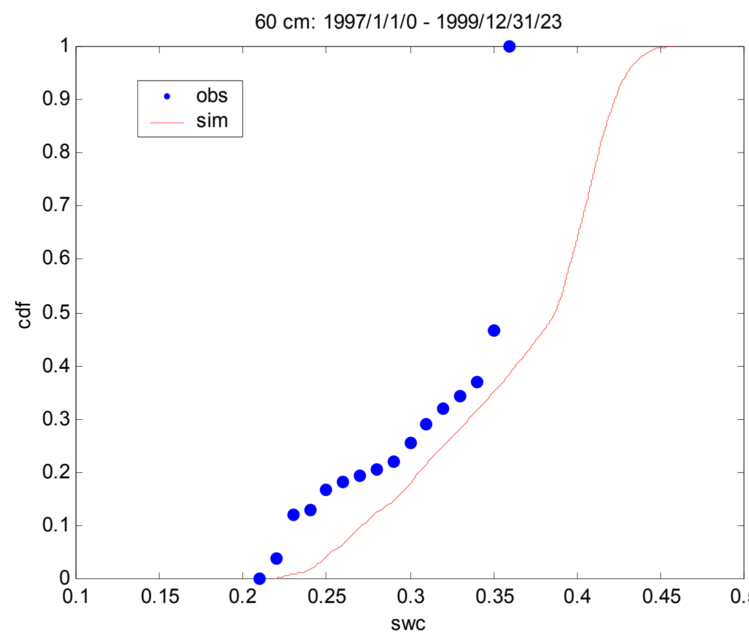
25 cm
simulated swc
observed swc
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5



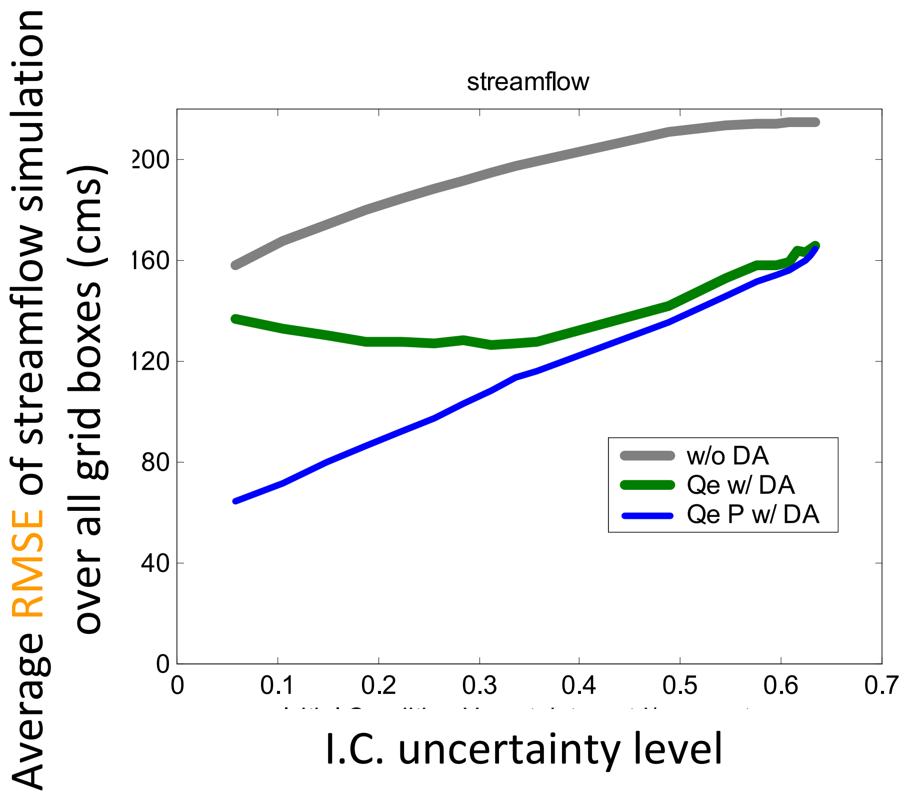
25 cm: 1997/1/1/0 - 1999/12/31/23
obs
sim
cdf
swc
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0.15
0.25
0.35
0.45

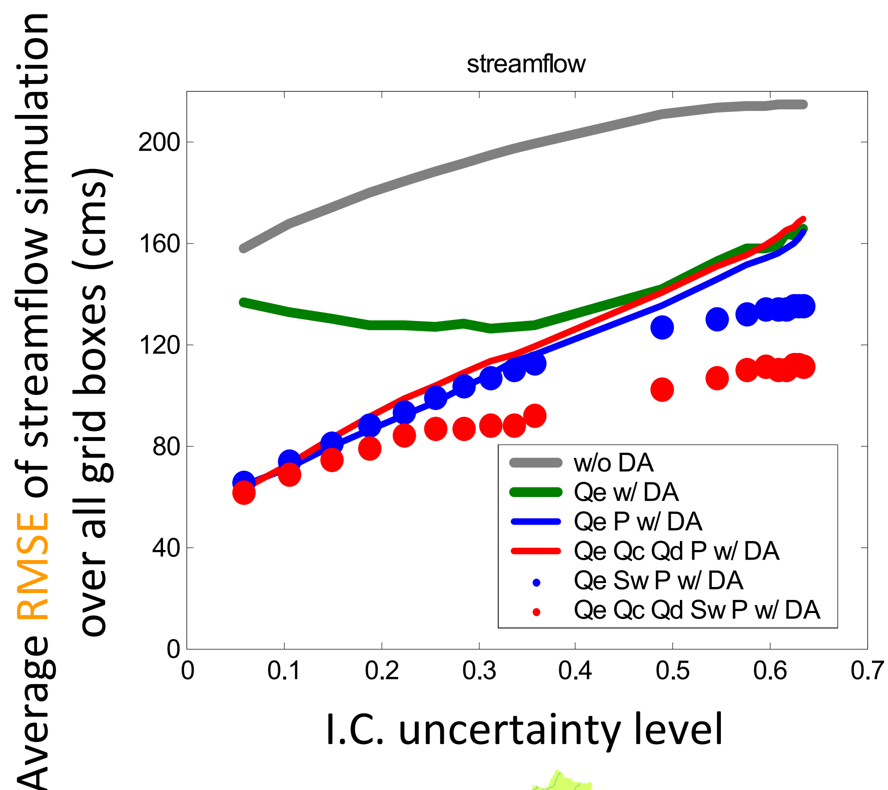

60 cm
obs
sim
swc [-]
time [hr]: 1997/1/1/0 - 2003/6/30/23
x 10^4


60 cm
simulated swc
observed swc


60 cm: 1997/1/1/0 - 1999/12/31/23
obs
sim
cdf
swc

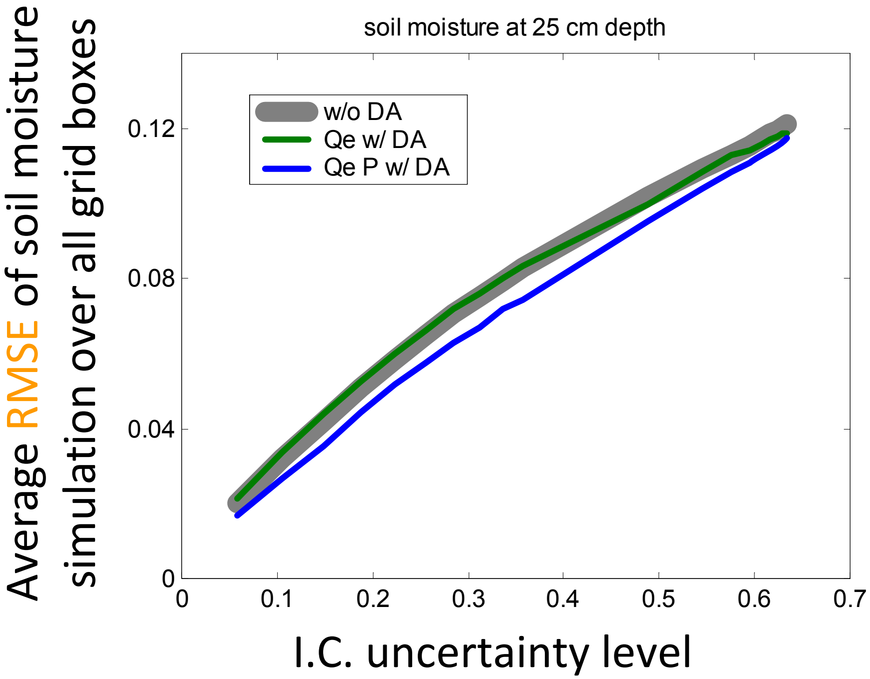
Impact of additionally assimilating precipitation to streamflow simulation

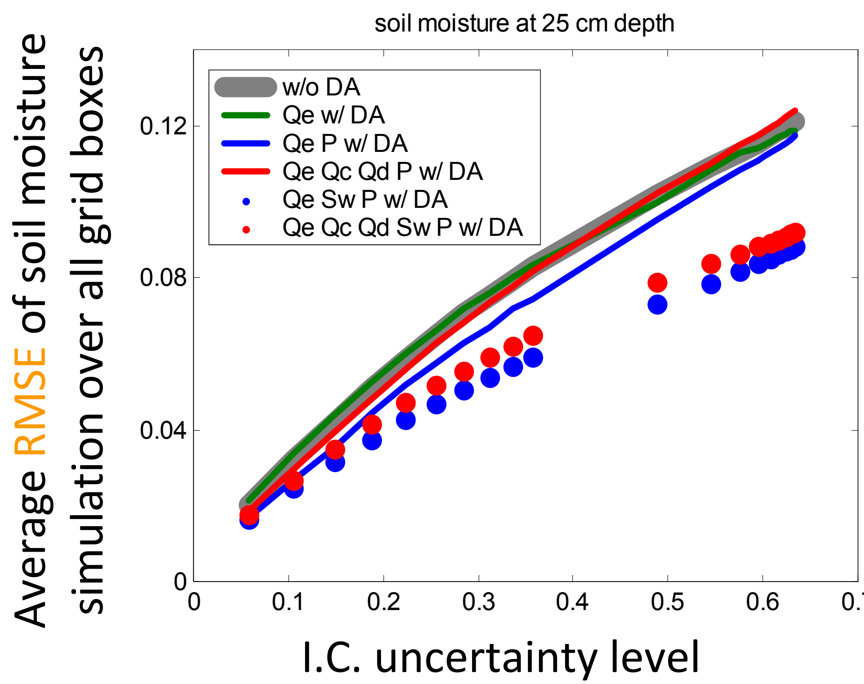




- Large precipitation uncertainty
- Medium streamflow observation uncertainty
- Medium soil moisture observation uncertainty

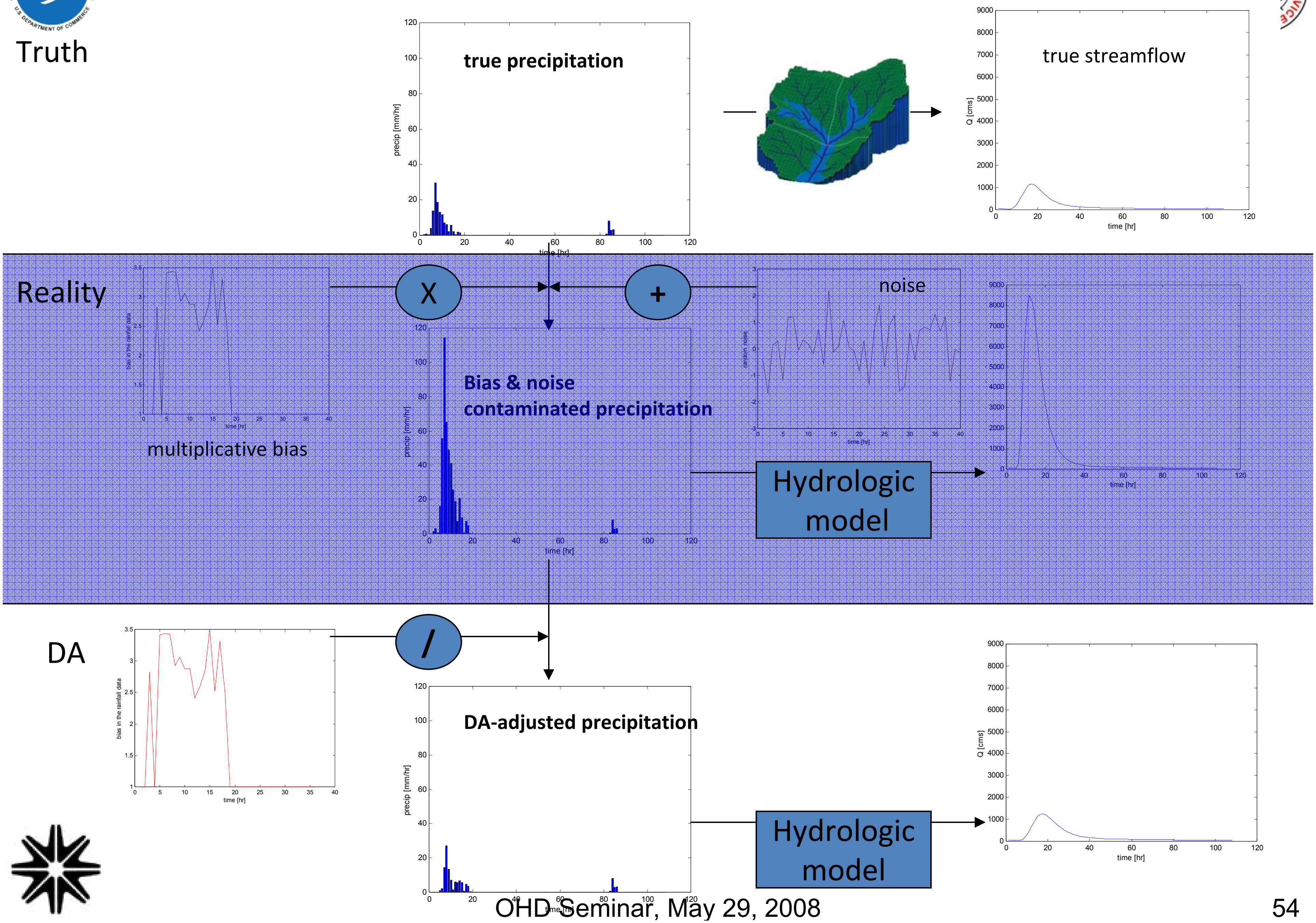
Impact of additionally assimilating precipitation to soil moisture simulation at 25-cm depth



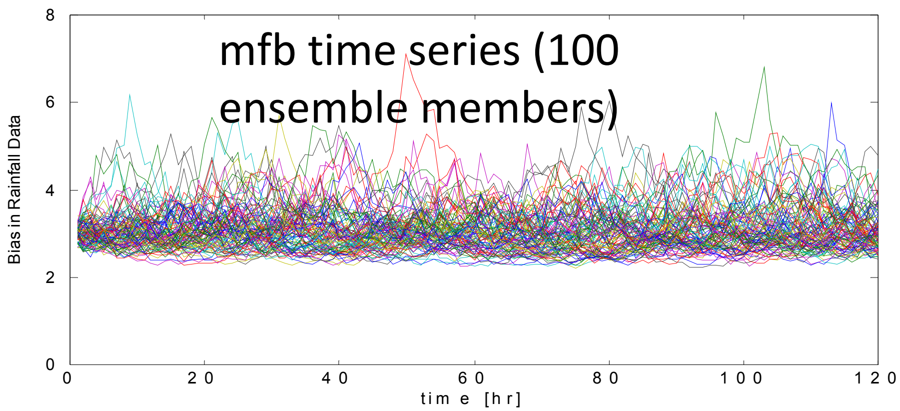


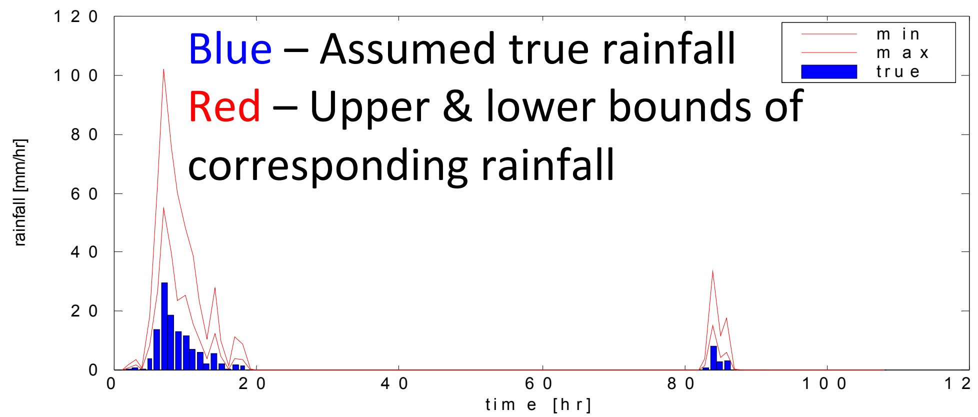
- Large precipitation uncertainty
- Medium streamflow observation uncertainty
- Medium soil moisture observation uncertainty

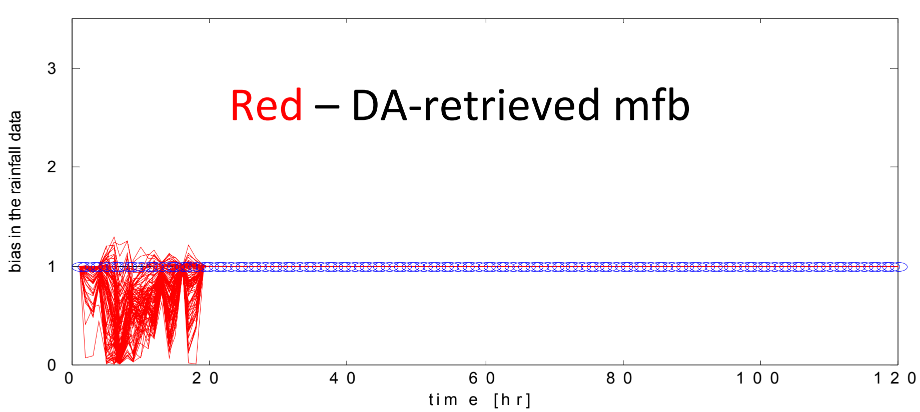
How additionally assimilating precipitation may reduce PU

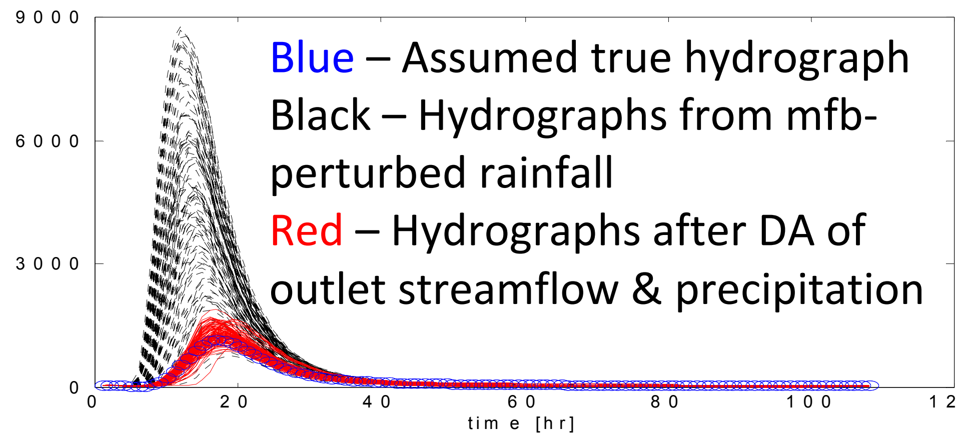


Large perturbations to mean-field bias (median=3)



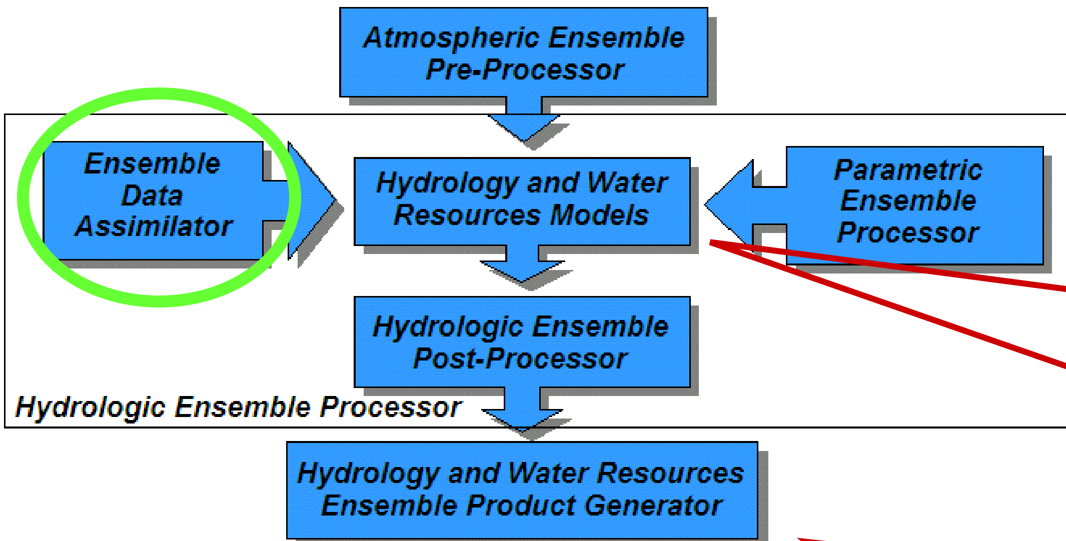


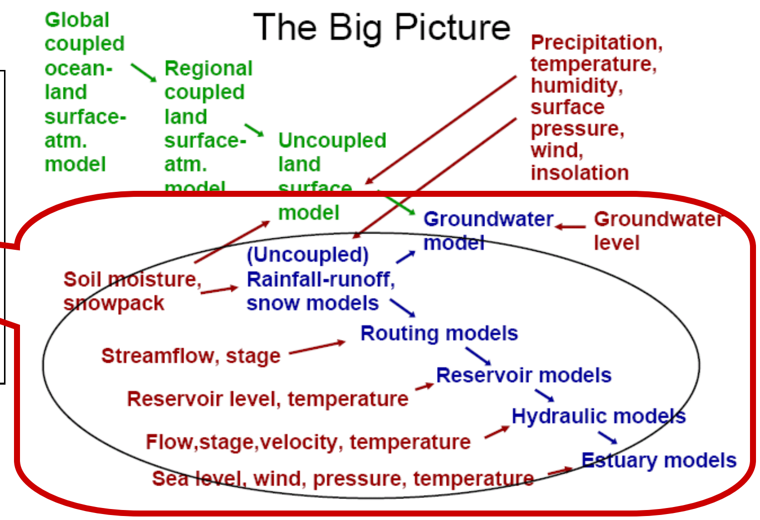




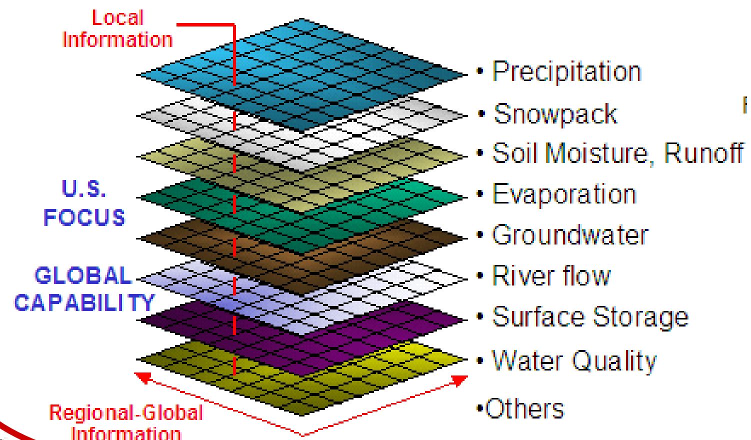
- Medium initial condition uncertainty
- Medium streamflow observation uncertainty

Vision for Ensemble & DA

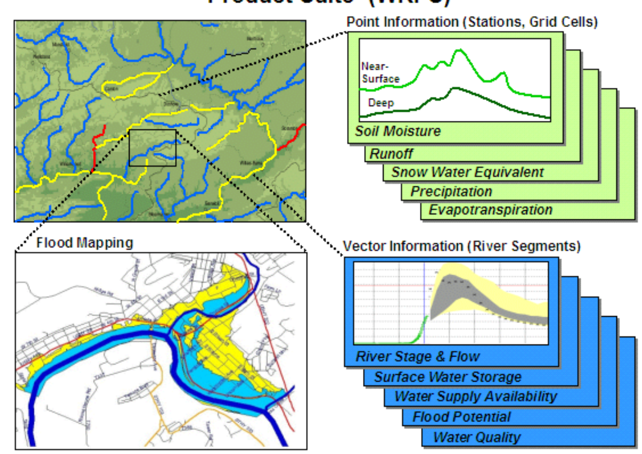




a) NWS-NDFD High-Resolution Gridded Water Resources Product Suite (WRPS)



b) NWS-NDFD High-Resolution Geospatial Water Resources Product Suite (WRPS)



Improved accuracy, Reliable uncertainty estimates, Benefit-cost effectiveness maximized