Deltares

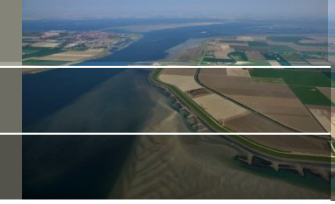
Enabling Delta Life

# Comparison of Kalman Filter Type and Variational Data Assimilation Approaches for Operational Hydrology

Dirk Schwanenberg (Deltares, University of Duisburg-Essen),
Rodolfo Alvarado Montero (University of Duisburg-Essen),
Albrecht Weerts (Deltares, Wageningen University)

2016 HEPEX Workshop
Ensemble for better hydrological forecasts
6-8 June 2016, Québec, Canada

# Introduction



Data Assimilation (DA) in hydrology:

- DA improves the system state of hydrological models at forecast time as the initial condition of subsequent forecasts
- Available techniques:
  - Sequential versus simultaneous techniques
  - Deterministic versus probabilistic approaches
  - Black box approaches such as the Ensemble Kalman Filter (EnKF) versus variational methods (4DVar) which need adjoint models and various hybrids
- Overview in Liu et al., 2012, but still a lack of comparisons in application to hydrological forecasting models

## Introduction (2)

Joint efforts to make DA broadly available for operational applications and to compare different approaches:

- H-SAF sub-project by University of Duisburg-Essen and German Federal Institute of Hydrology to develop novel variational DA methods (4DVar) and compare them against Kalman Filter style approaches in OpenDA, visiting scientist activity of Deltares
- General OpenDA development at Deltares in collaboration with TU Delft, VORtech, DHI and TNO
- Many specific OpenDA applications in operational hydrology at Deltares and Wageningen University

# Introduction (3)

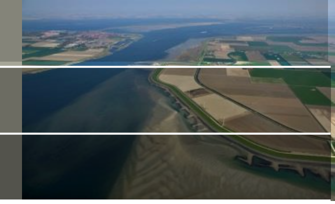

OpenDA (http://www.openda.org/):

- Toolbox for data-assimilation and calibration of arbitrary numerical models
- Ensemble KF (EnKF), Asynchronous Ensemble KF (AsEnKF), Ensemble SquareRoot KF (EnSR), Particle Filter, 3DVar

RTC-Tools (http://oss.deltares.nl/web/RTC-Tools):

- Model library including HBV implementation and routing schemes providing first-order sensitivities, i.e. adjoint models, for variational DA
- Moving Horizon Estimation (aka 4DVar)

# Short Review of DA Methodology



Hydrological model according to:

$$\underset{(true)\,state}{x^k} = \underbrace{f\left(x^{k-1}, u^k\right)}_{\text{prediction model}} + \underset{\text{process noise}}{\Delta x^k}$$

$$\underset{\text{observation}}{y^k} = \underset{\text{observation model}}{g\left(x^k\right)} + \underset{\text{observation noise}}{\Delta y^k}$$

Common assumption: noise is a zero mean multivariate normal distribution $\Delta x^k \approx N\left(0, Q^k\right), \Delta y^k \approx N\left(0, R^k\right)$

with covariance matrices $Q^k, R^k$

Deltares

# Short Review of DA Methodology (2)

## Kalman Filter

### Prediction step:

predicted state estimate $x_{sim}^{k} = \underbrace{F^{k} x^{k-1} + B^{k} u^{k}}_{\text{prediction model}}$

KF: linear
EKF: linearized
EnKF: sampled

predicted covariance estimate $P^{k|k-1} = F^{k} P^{k-1|k-1} F^{k,T} + Q^{k}$

### Update step:

EnKF: no explicit computation, repr. by state ensemble

observed residual $\Delta y^{k} = y^{k} - \underset{\text{observation model}}{H^{k} x^{k}}$

residual covariance $S^{k} = H^{k} P^{k|k-1} H^{k,T} + R^{k}$

Kalman gain $K^{k} = P^{k|k-1} H^{k,T} S^{k,-1}$

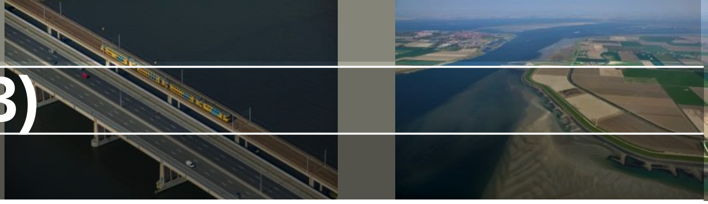
updated state estimate $x^{k} = x_{sim}^{k} + K^{k} \Delta y^{k}$

updated covariance estimate $P^{k|k} = \left(I - K^{k} H^{k}\right) P^{k-1|k}$

Deltares

# Short Review of DA Methodology (3)

## 4Dvar aka Moving Horizon Estimation



Reformulation of the DA problem by cost function and minimization by optimization algorithm:

$$\left(y^k - H^k x^k\right)^T R^{k,-1} \left(y^k - H^k x^k\right)$$

$$= \Delta y^{k,T} R^{k,-1} \Delta y^k$$

$$\approx \frac{1}{\sigma_{obs}^2} \left\| \Delta y^k \right\|$$

may handle arbitrary PDF or norms

simultaneous DA over several time steps

$$\min_{\Delta x, \Delta y} \sum_{k=-N+1}^{0} w_{obs} \left\| y_{obs}^k - y_{sim}^k(\Delta x, \Delta y) \right\| + w_{\Delta x} \left\| \Delta x^k \right\| + w_{\Delta y} \left\| \Delta y^k \right\|$$

$$\text{subject to} \quad \Delta x_L \le \Delta x^k \le \Delta x_U, \ \Delta y_L \le \Delta y^k \le \Delta y_U \quad \text{etc.}$$

optional constraints

Deltares

# Comparison of 4DVar and Ensemble KF

4DVar:

+ simultaneous technique over several time steps
+ suitable for reanalysis
- requires first-order sensitivities, i.e. adjoint code, and preferably a smooth model
- deterministic approach

Ensemble KF:

+ applicable on black-box models, simple to implement
+ probabilistic approach
- sequential technique, has issues with time lags

see also Skachko et al. (2014): Comparison of the ensemble Kalman filter and 4D-Var assimilation methods using a stratospheric tracer transport model, Geosci. Model Dev., 7, 1451–1465, 2014

# Example: Snow Accumulation and Melt

## Assimilation of observed flow only

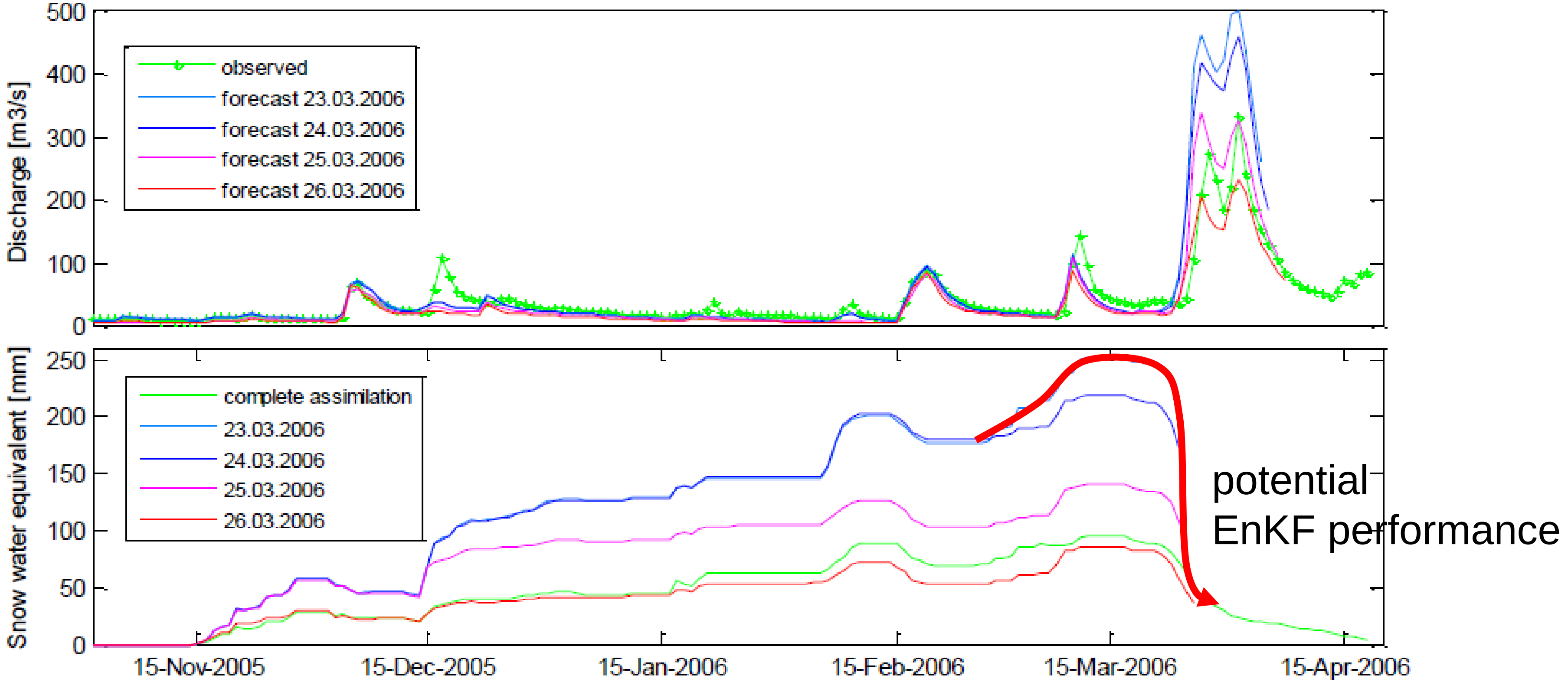



Figure 2. Lead time performance of the Main1 model for a flood event in March 2006 using observed discharge in the assimilation: i) above: comparison of observed and (assimilated) forecasted streamflow of forecast times between March 23-26, ii) below: comparison of “perfect” and assimilated snow water equivalent for the same forecasts.

Deltares

# Example: Snow Accumulation and Melt (2)

## Assimilation of observed flow and “perfect” SWE

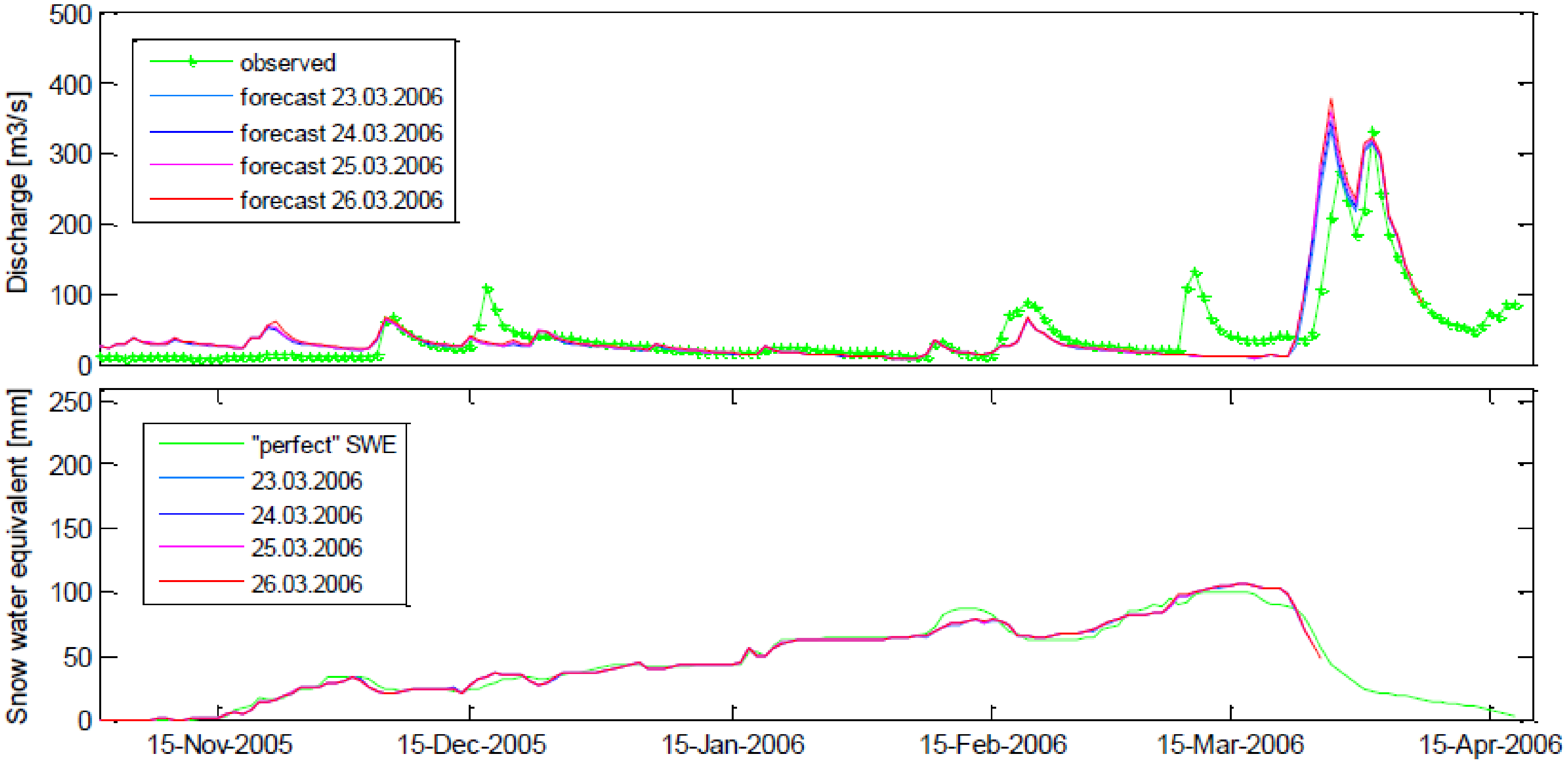
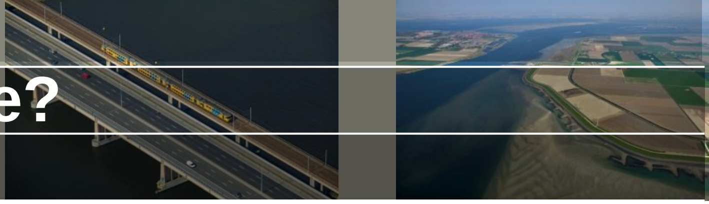


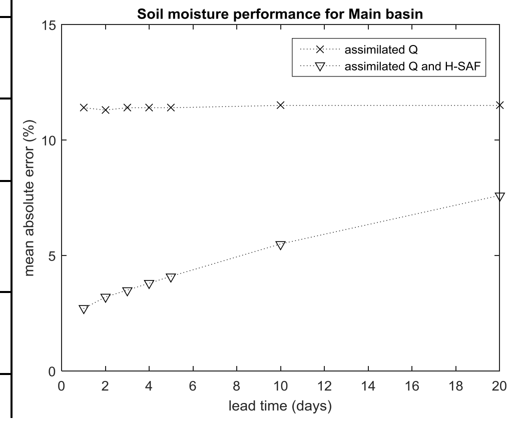

Figure 3. Lead time performance of the Main1 model for a flood event in March 2006 using observed snow water equivalent and discharge in the assimilation: i) above: comparison of observed and (assimilated) forecasted streamflow of forecast times between March 23-26, ii) below: comparison of “perfect” and assimilated snow water equivalent for the same forecasts.

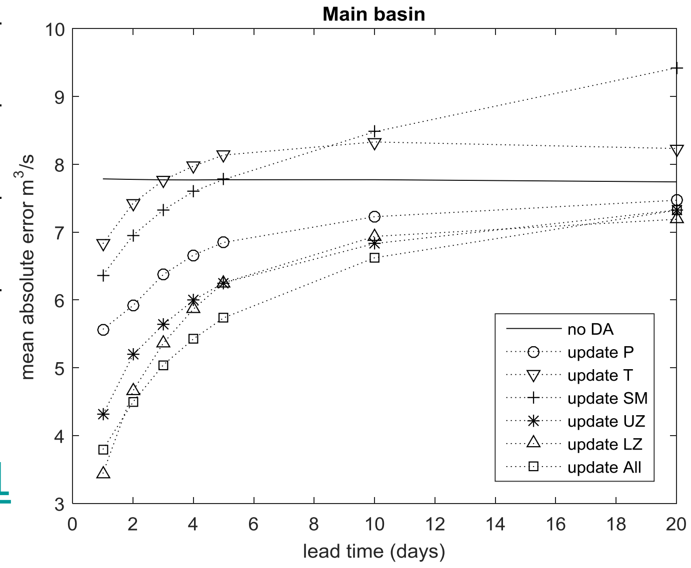
Deltares



# Which observation has added value?

| Variable | | Objective Function Term |
|---|---|---|
| Model Inputs | Precipitation ($P$) | $w_P(\Delta P^k)^2$ |
| | Temperature ($T$) | $w_T(\Delta T^k)^2$ |
| Model States | Snow Water Equivalent ($SWE = SP + WC$) | $w_{SWE}(\hat{s}^k_{SWE} - s^k_{SWE})^2$ |
| | Soil Moisture ($SM$) | $w_{SM}(\hat{s}^k_{SM} - s^k_{SM})^2 + w_{\Delta SM}(\Delta s^k_{SM})^2$ |
| | Upper Zone Storage ($UZ$) | $w_{\Delta UZ}(\Delta s^k_{UZ})^2$ |
| | Lower Zone Storage ($LZ$) | $w_{\Delta LZ}(\Delta s^k_{LZ})^2$ |
| Model Outputs | Snow Covered Area ($SCA$) | $w_Q(\hat{A}^k_{SCA} - A^k_{SCA})^2$ |
| | Discharge ($Q$) | $w_Q(\hat{Q}^k - Q^k)^2$ |



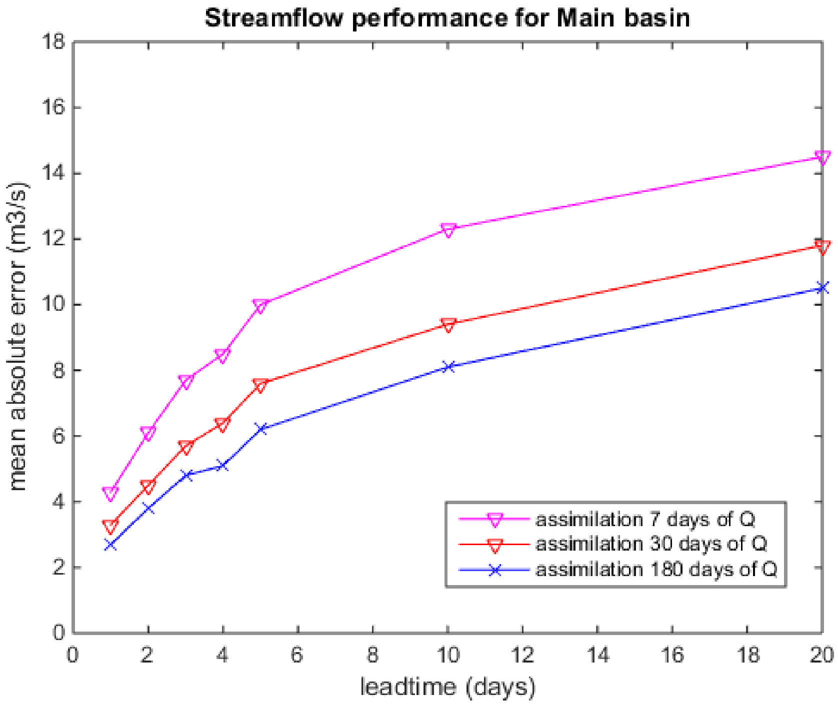

http://dx.doi.org/10.1016/j.advwatres.2016.04.011

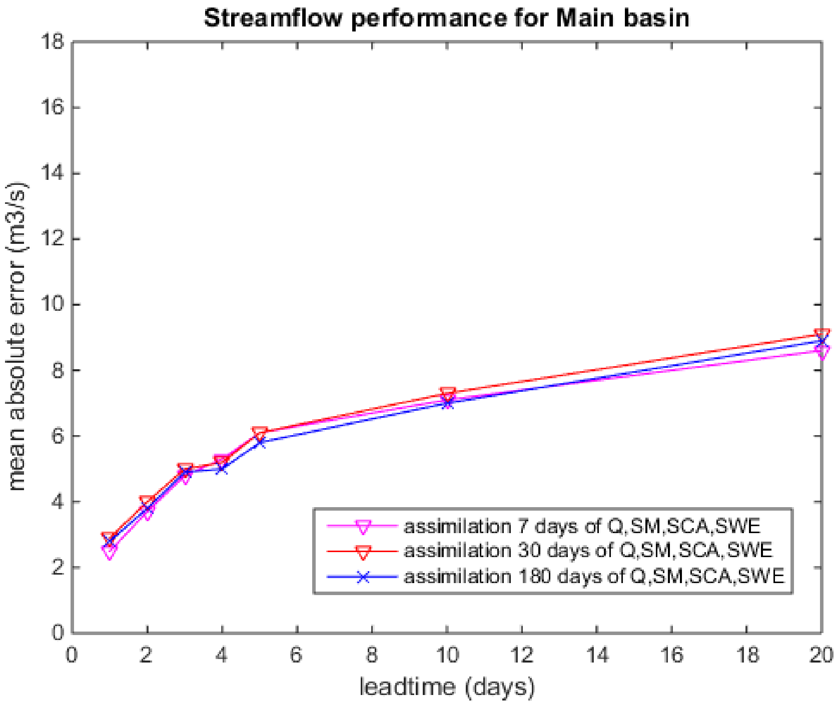
Deltares

# What optimization window to use?

Assimilation of streamflow only

Assimilation of streamflow, SCA, SWE and SM



# Example: Snow Accumulation and Melt (3)
## Discussion / Conclusions

Benefit of the simultaneous approach for processes such as snow accumulation and melt (with long memory):

- EnKF corrects when detects (challenge to update with sufficient skill to keep the snow pack meaningful from physical perspective)
- MHE/4DVar reconstructs the snow pack within the estimation window leading to
  - Snow pack with more physical meaning
  - Historical reanalysis
  - Robust against missing and noisy data
- Difference between approaches becomes less relevant if snow products from observations or remote sensing get used for continuous updates



Deltares

# Representation of Uncertainty in Initial State

Important benefit of EnKF (over MHE/4DVar):

- Sampled distribution of states represent the model uncertainty in the initial condition and the subsequent forecast

Our current research looks into approaches to make the MHE/4DVar approach more probabilistic:

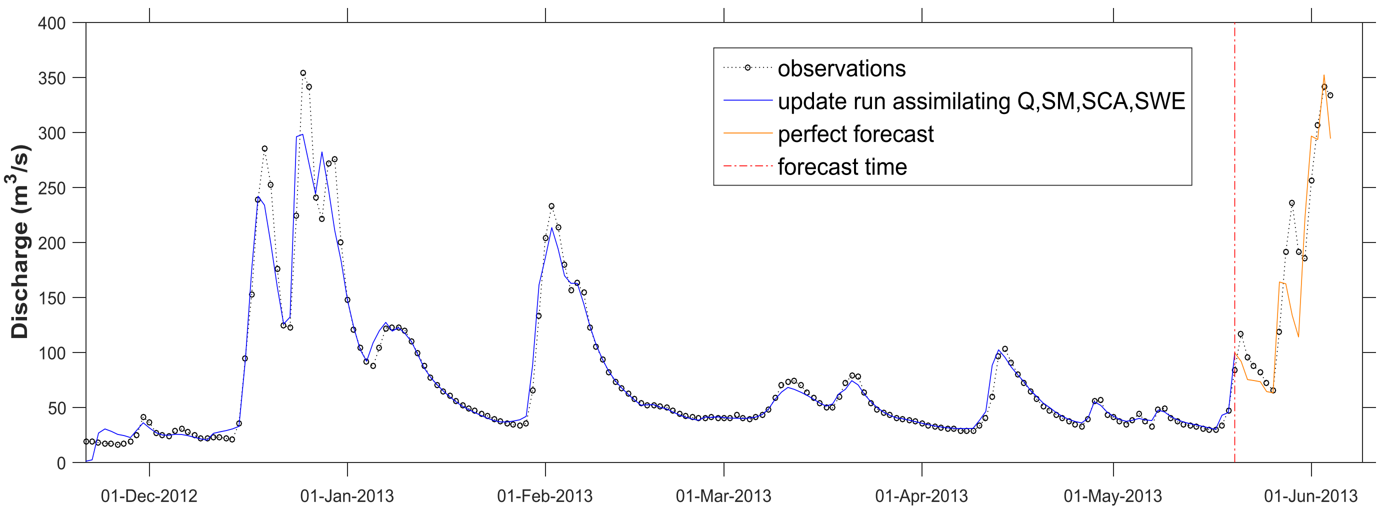
- Introduction of a model pool instead of a single model
- Variation of model structure and model parameters

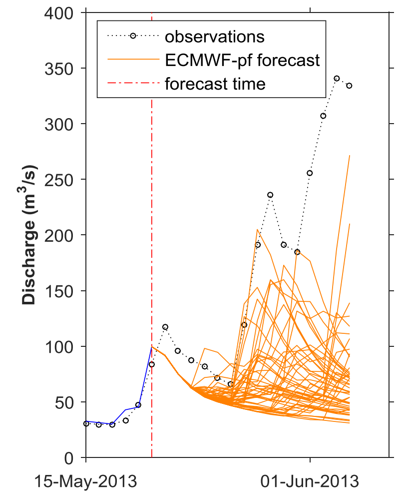
see poster: Rodolfo Alvarado Montero, Dirk Schwanenberg, Peter Krahe, VARIATIONAL DATA ASSIMILATION BY MOVING HORIZON ESTIMATION AND A PROBABILISTIC POOL OF CONCEPTUAL HYDROLOGICAL MODELS

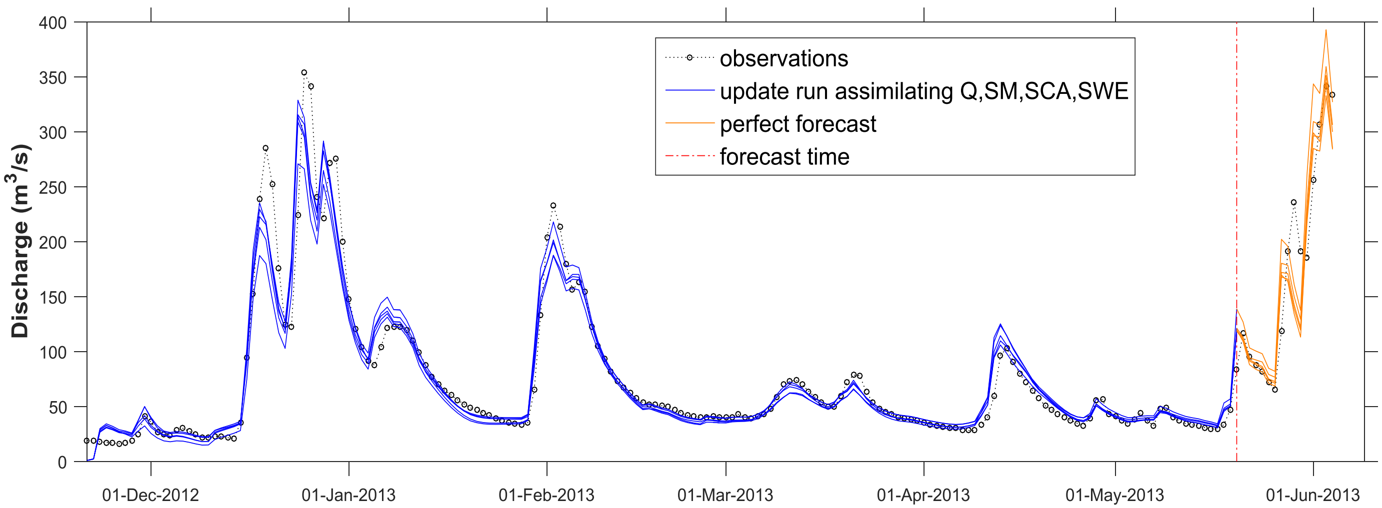


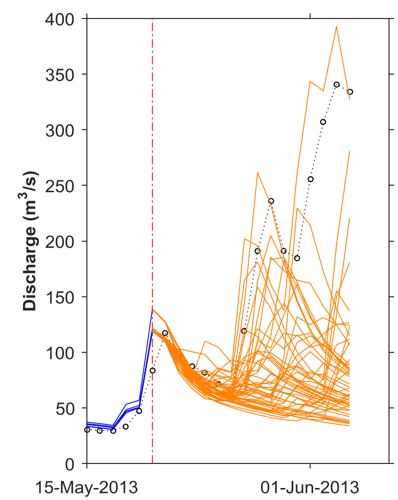
Deltares

# Representation of Uncertainty in Initial State (2)




Deltares

# Representation of Uncertainty in Initial State (3)
## Discussion / Conclusions

The use of model pools instead of a single deterministic model in variational DA seems to have some benefits to represent model uncertainty (which is not propagation out of the model).

Our current work:

- Include different model structures in the model pool (so far only parametric uncertainty)
- Comparison against EnKF, Asynchronous EnKF, Particle Filter etc.

Remark: The approach is quite challenging from a technical perspective (adjoint of model pool etc.).

# Outlook

- Broad application of various DA techniques for a representative number of river basins:
  - <u>Break-out session "Assimilate your basin"</u> on Thursday
  - Dissemination activities in the H-SAF project and at Deltares
- Further efforts to streamline freely available DA tools and make them available to the public
- Approaching scientific papers in 2016:
  - Comparison of several DA techniques in application to conceptual hydrological models
  - Probabilistic MHE/4DVar

Deltares


# Thank you for your attention!

Recent publications:


- D. Schwanenberg, R. Alvarado Montero, Total Variation Diminishing and Mass Conservative Implementation of Hydrological Flow Routing, J. Hydrology, http://dx.doi.org/10.1016/j.jhydrol.2016.05.007
- R. Alvarado Montero et al., Moving horizon estimation for assimilating H-SAF remote sensing data into the HBV hydrological model, Advances in Water Resources, Volume 92, June 2016, Pages 248-257, ISSN 0309-1708, http://dx.doi.org/10.1016/j.advwatres.2016.04.011
- F. Mainardi Fan, D. Schwanenberg, R. Alvarado Montero, A. Assis dos Reis, W. Collischonn, S. Naumann, Performance of deterministic and probabilistic hydrological forecasts for the short-term optimization of a tropical hydropower reservoir, Water Resources Management (accepted)


Deltares