Comparison of Sequential and Variational Data Assimilation

Rodolfo Alvarado-Montero (1), Dirk Schwanenberg (2), Albrecht Weerts (3)

(1) Institute of Hydraulic Engineering and Water Management, WaWi, University of Duisburg-Essen, Germany,
(2) KISTERS AG, Germany
(3) Deltares, The Netherlands

rodolfo.alvarado-montero@uni-due.de

In the framework of the HEPEX-DA Inter-Comparison experiment
Abstract EGU2017-19012

State of the art

- sequential techniques such as the Ensemble Kalman Filter (EnKF) require no additional features within the modeling process,
- variational techniques rely on optimization algorithms to minimize a pre-defined objective function. This function can be formulated as a trade-off between the amount of noise introduced into the system and the mismatch between simulated and observed variables,
- sequential techniques have been commonly applied to hydrological processes, variational techniques have been seldom used,
- lack of thoroughly comparisons in hydrological applications.

Methodology

We use a dedicated implementation of the HBV model[1], as documented in Schwanenberg et al[4]. Let us consider the following formulation[2]:

$$x_k = M_k\left(x_{k-1}, \theta, \mu_k\right) + \eta_k$$
$$z_k = H_k\left(x_k, \theta\right) + \varepsilon_k$$

x is vector of true system states, *M* represents the forward model, θ is vector of parameters, *μ* input vector, *z* the observation vector, *H* the observation operator, *η* the model error with covariance *Q*, *ε* the observation error with covariance *R*, *k* the time index

And by adding (bounded) perturbations to the model input vector :

$$u_k^i = u_k + \xi_k^i \qquad \xi_k^i \approx N\left(0, U_k\right) \qquad U_k \text{ is the covariance of } u_k$$

The EnKF setup becomes:

$$x_k^{-,i} = M_k\left(x_{k-1}^{+,i}, \theta, u_k^i, \delta_k\right)$$
$$x_k^{+,i} = x_k^{-,i} + K_k d_k^i$$
$$d_k^i = \left(z_k + \varepsilon_k^i\right) - H_k\left(x_k^{-,i}, \theta\right)$$

where *K* is the Kalman gain computed from the covariance matrices of state predictions and observations

$$K_k = E\left[X_k^- Z_k^-\right] \cdot \left(E\left[Z_k^-\left(Z_k^-\right)^T\right] + R_k\right)^{-1}$$

We formulate the variational data assimilation[3] for a forecast time $T_k=0$ over an assimilation period $k=[-N+1,0]$ of $N\geq 1$ time steps, by an optimization problem subject to model *M*, according to:

$$J = \min_{\xi_k, \eta_k} \sum_{k=-N+1}^{0} \left(w_\xi \xi_k^{\,2} + w_\eta \eta_k^{\,2} + w_\varepsilon \varepsilon_k^{\,2}\right)$$

w are weighting factors that reflect the trade-off between noise introduced to the model and the mismatch of simulated and observed variables, ξ_k is the noise added to the model input vector

In this formulation the model error becomes an optimization variable, together with the perturbation to the model input vector. Notice that the EnKF solves a very similar problem:

$$J = \min_{x_k}\left(\frac{\eta^2}{Q_k} + \frac{\varepsilon^2}{R_k}\right)$$

Case studies

Basin	Area (km2)	Elevation (mASL)	Land use (type)	Avg. discharge (m3/s)	Model NSE (calibration)
Main (Germany)	4254	220 – 1034	57% for., 38% grass, 5% built-up	44	0.926
Mistassibi (Canada)	8743	269 – 603	mostly forest	201	0.856

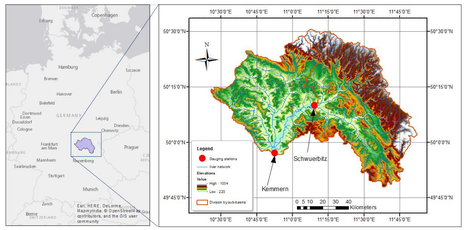

Fig.1. Location of Upper Main Basin in Germany, with Schuwerbitz and Kemmern gauging stations

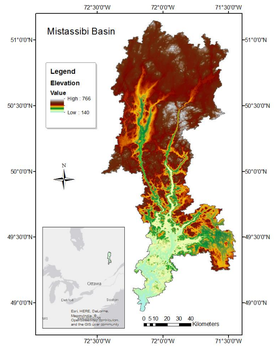

Fig.2. Location of Mistassibi basin in Canada

Experiment Setups

We assimilate only discharge data

Sequential EnKF DA	Variational DA
Warm-up period of 1 year	Assimilation window of 180 days
Perturbations to model inputs (*μ*) - Precipitation: normal distr. N(0, ±15%), tail limits at ± 30% - Temperature: normal distr. N(0,±0.5°C) , tail limits at ± 1.0°C Updates the complete matrix of model states	Noise to model inputs (*μ*) - Precipitation: bounded to ± 30% - Temperature: bounded to ± 1.0°C Noise to model states (*η*) - Soil moisture: bounded to ± 1.0mm - Upper zone: bounded to ± 1.0mm - Lower zone: bounded to ± 1.0mm
Tests using 50, 100 and 200 members	Deterministic method
Observation error from 2.5% to 5%	Observation mismatch weights from 0.001 to 100, model noise weights kept constant at 1.0

Results for a single event

Comparison of flood event in May 2013 in Main basin (similar performance)

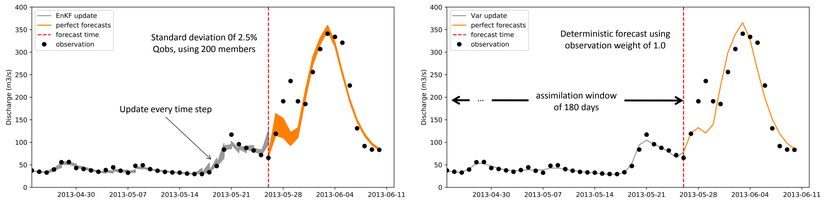

Fig.3. Update and forecast runs for a single event in late May 2013 at Kemmern gauging station (Main basin) using a perfect forecast; left: using EnKF, right: with variational data assimilation

Results for hindcasting experiment

Results at Main:

- Variational DA shows a better mean Continuous Ranked Probability Score (CRPS) performance particularly for short lead time up to 3 days.
- Better Mean Absolute Error (MAE) performance compared to EnKF
- Improvements of MAE up to 10% for 15 days lead times for the highest 25-percentile events

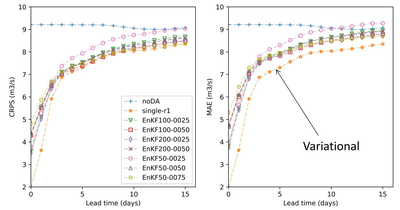

Fig.2. Comparison of performance for EnKF and Var at Main basin

Table 1: Improvement of variaitional data assimilation with respect to EnKF200-0025

Lead time (days)	Improvement all events (%)	Improvement 25-perc. (%)
0	47%	51%
3	8%	7%
6	6%	7%
9	5%	9%
12	5%	10%
15	5%	10%

Results at Mistassibi:

- Best CRPS for EnKF corresponds to observation uncertainty of 3%-5%
- Best CRPS for Variational is reached with observation weights between 0.01 to 1.0 at noise weights of 1.0
- Improvement of MAE from 51% to 8% for thes first 8 days of lead time.
- Performance of both DA for the 25-percentile events diminishes after 8 days. EnKF however, remains closer to results without DA

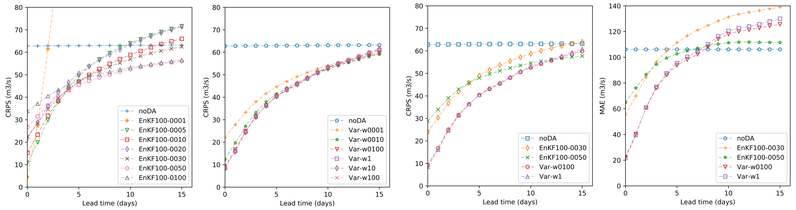

Fig.4. Performance indicators for: a) different setups of EnKF, b) different setups of Var, c) best MAE, d) MAE 25th percentile

References

1. Lindström, G., Johansson, B., Persson, M., Gardelin, M., & Bergström, S. (1997). Development and test of the distributed HBV-96 hydrological model. *Journal of Hydrology*, *201*(1–4), 272–288.
2. Liu, Y., & Gupta, H. V. (2007). Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water Resour. Res.*, *43*(7), W07401-.
3. Montero, R. A., Schwanenberg, D., Krahe, P., Lisniak, D., Sensoy, A., Sorman, A. A., & Akkol, B. (2016). Moving horizon estimation for assimilating H-SAF remote sensing data into the HBV hydrological model. *Advances in Water Resources*, *92*, 248–257.
4. Schwanenberg D, B Bernhard. RTC-Tools Reference Manual. (2013).

Acknowledgements

This project was funded by the German Federal Institute of Hydrology (BfG), in the context of the EUMETSAT funded project Hydrological Satellite Application Facility (H-SAF). We acknowledge the contribution from Marie-Amelie Boucher (Université du Quebec à Chicoutimi, Canada) for supplying the data for the Mistassibi catchment.