Data Assimilation in Environmental Modelling

Henrik Madsen
DHI

HOBE-TR32 Joint workshop, Sandbjerg Castle, 1-4 April 2012
Data assimilation framework

1. Correction of model forcing
2. State updating
3. Parameter estimation
4. Error forecasting

Refsgaard (1997)
Filtering framework

Model forecast

$$x_k^f = \Phi(x_{k-1}^a, u_k)$$

Measurements

$$z_k = C_k(x_k^f)$$

Update

$$x_k^a = x_k^f + G_k(z_k - C_k(x_k^f))$$

**Φ(.)** Model operator (one-step ahead predictor)

**x_k** System state

**u_k** Model forcing

**z_k** Measurements

**C_k** Mapping of state space to measurement space

**G_k** Weighting matrix

**k** Time step
Kalman filter

Model forecast

\[ x_k^f = \Phi(x_{k-1}^a, u_k) + \varepsilon_k \]

Measurements

\[ z_k = C_k(x_k^f) + \eta_k \]

Update

\[ x_k^a = x_k^f + G_k(z_k - C_k(x_k^f)) \]

- Stochastic interpretations of model and measurement equations
- Weighting matrix (Kalman gain) determined based on a least squares minimisation of the expected error of the updated state estimate
- Estimation of model prediction uncertainty
Kalman filter

Model forecast
\[ x_k^f = \Phi(x_{k-1}^a, u_k) + \epsilon_k \]

Measurements
\[ z_k = C_k(x_k^f) + \eta_k \]

Update
\[ x_k^a = x_k^f + G_k(z_k - C_k(x_k^f)) \]

Statistical properties:
- If the model is linear, and model and measurement errors are independent white noise with known covariance matrices, the Kalman filter is the best linear unbiased estimator (BLUE).
- If, in addition, model and measurement errors are Gaussian, the Kalman filter is the maximum a posteriori (MAP) or maximum likelihood (ML) estimator.
Application in environmental modelling

Challenges
• Non-linear model dynamics
• Non-Gaussian and possible biased model and measurement errors
• Unknown model and measurement error statistics
• High-dimensional systems

Solutions
• Steady-state Kalman filter
  – Kalman gain calculated "off-line"
  – Assumes constant model and measurement error statistics
• Ensemble-based Kalman filters
  – Covariance matrix represented by an ensemble of states (ensemble size << dimension of state)
  – Error propagation according to full non-linear model dynamics
  – Sampling uncertainty decreases slowly with increasing ensemble size
Joint state updating and model forcing correction

- Model forcing main error source
- Temporal correlation model (e.g. AR(1) model) of model forcing included in filtering scheme using augmented state vector formulation
- Model forcing error updated along with the model state
Twin test experiment

- False run: Model forced with erroneous boundary conditions
- Update of false model using water level measurements at two locations (from reference run)
Joint state updating and bias correction

- Bias aware Kalman filter
  - Include bias using augmented state formulation
  - Separate bias Kalman filter (Dual Kalman filter)

Drecourt et al. (2006)
Twin-test experiment: Groundwater model

Classical KF

CoIKF no feedback

Bias

Bias

www.dhigroup.com
Joint state updating and parameter estimation

- Include model parameter estimation in Kalman filter
  - Augmented state vector formulation
  - Dual Kalman filter formulation
Twin-test experiment: Groundwater model

Jørn Rasmussen: Data assimilation in hydrological models – evaluation of filter performances (Poster)
Hybrid filtering and error forecasting

- In practice model innovations are not white noise
- Hybrid filtering and error forecasting includes forecast of model innovation
Hybrid filtering and error forecasting

- Error forecast model applied to forecast innovation in measurement points -> virtual measurement
- Filtering using virtual measurements

Madsen and Skotner (2005)
Regularisation

Challenges with ensemble based Kalman filtering methods

- Computationally expensive
- Introduces sampling uncertainties in covariance and Kalman gain estimates (spurious correlations)
- Performance sensitive to description of model and measurement errors (bias in error covariance estimation)

- Include regularisation to provide more robust and computationally efficient filtering schemes
Regularisation methods

• Distance regularisation (localisation)
  – Update state only in local region around measurement
  – Reduces the effect of spurious correlation in data sparse regions

• Covariance or Kalman gain smoothing
  – Temporal smoothing of covariance or Kalman gain
  – Reduces the effect of sampling uncertainty
  – But fade out high-frequency variability

\[ K_k^{\text{smooth}} = (1 - \alpha) K_{k-1}^{\text{smooth}} + \alpha K_k, \quad 0 < \alpha < 1 \]

\( \alpha = 0 \): Steady-state Kalman filter
\( \alpha = 1 \): Normal Kalman filter

*Sørensen et al. (2004)*
Regularisation: Application example

Forecast skills
RMSE [m] as function of lead time

Assimilation stations

Validation stations
Implementation

- OpenDA-OpenMI open framework for data assimilation
- OpenDA is an open interface standard for data assimilation and offers a number of ensemble based algorithms
- OpenMI is an open source standard model interface

Marc-Etienne Ridler: Open Framework for Data Assimilation (Poster)
Concluding remarks

• A general filtering data assimilation framework
  – Allows joint updating and estimation of model state, forcing, parameters and bias
  – Can be combined with error forecasting in measurement points
• Kalman filtering with regularisation efficient for real-time applications
• Hybrid filtering and error forecasting procedures can increase forecast skills for longer lead times

• Challenges
  – Representation of model and measurement uncertainty
  – Computational efficiency
  – Multi-variate data assimilation in integrated models using in-situ and remote sensed data
Thank you for your attention

Henrik Madsen
DHI
hem@dhigroup.com