Sequential data assimilation for subsurface flow modeling in civil engineering

Shinya Yamamoto
Shimizu Corporation
Roles of subsurface flow analysis in civil engineering

- Impact assessment of the construction in design phase
- Regular monitoring in operation and maintenance
- Safety assessment with long-term forecast for geological nuclear waste disposal
Difficulties in the subsurface flow simulation

We have carried out deterministic simulation under huge uncertainties.

- Sparse observations
- Geological heterogeneity and discontinuity
- Incomplete knowledge about hydraulic parameters and boundary conditions
Objectives

- Estimate the hydraulic parameters that can represent the current state of subsurface flow.
- Construct the reliable subsurface flow model for forecasting

Our solution: sequential data assimilation

- Efficient use of observations
- Easy implementation
- Online method
Sequential-based data assimilation methods

- Particle Filter
  - No assumption
  - Very easy implementation
  - Expensive computation cost
- Ensemble Kalman Filter (EnKF)
  - Relatively small ensemble size
  - Gaussian assumption

Versatile PF or Efficient EnKF?
Numerical experiments

Application to an actual full-scale model

Step 1: Twin experiments with synthetic data
- Evaluate the performance of PF and EnKF
- Identify the issues of the data assimilation method

Step 2: Assimilation test using real observation (in progress)
- Identify the issues of the model and observations

Kikuma underground oil storage
Saturated-unsaturated subsurface flow model equation

\[
\frac{\partial}{\partial x} \left( K \frac{\partial h}{\partial x} \right) = \left( C + \beta S_S \right) \frac{\partial h}{\partial t}
\]

Discretize in time using

\[
\frac{\partial h}{\partial t} \approx \frac{h_t - h_{t-1}}{\Delta t}
\]

\[
h_t = \left\{ 1 + \frac{\Delta t}{C + \beta S_S} \frac{\partial}{\partial x} \left( K \frac{\partial}{\partial x} \right) \right\} h_{t-1}
\]

\[h\] : Pressure head
\[K\] : Permeability
\[C\] : Specific water capacity
\[S_S\] : Specific storage
\[\beta = 1\] saturated \[/ = 0\] not saturated
Summary of subsurface flow simulation

<table>
<thead>
<tr>
<th>Solver</th>
<th>Finite Element Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimension</td>
<td>34,000 nodes</td>
</tr>
<tr>
<td></td>
<td>50,000 elements</td>
</tr>
<tr>
<td>Boundary conditions</td>
<td>• Water seal pressure</td>
</tr>
<tr>
<td></td>
<td>• Internal tank pressure (Oil level)</td>
</tr>
<tr>
<td></td>
<td>• Groundwater recharge</td>
</tr>
<tr>
<td></td>
<td>• Shoreline</td>
</tr>
<tr>
<td>Initial condition</td>
<td>Use the result of static analysis</td>
</tr>
</tbody>
</table>

Elements of tanks (overhead view)
Geology

5 Geological units
- Weathered zone
- Granite (Fresh rock)

10 Fracture zones

We estimate 15 hydraulic parameters
Observations

- Groundwater level at 14 boreholes
- Inflow and outflow volumes around 3 oil tanks (total 6 observations)
Summary of Condition of data assimilation

<table>
<thead>
<tr>
<th></th>
<th>PF (SIR)</th>
<th>EnKF (PO)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated parameter</td>
<td>15 permeability parameters $\log(K)$</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>20 observations (Pressure head and flow volume)</td>
<td></td>
</tr>
<tr>
<td>State vector</td>
<td>$x=(h_1, \ldots, h_{Nn}, Q_1, \ldots, Q_{Nn}, \log(K)<em>1, \ldots, \log(K)</em>{15})^T$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$h$ : Pressure head, $Q$ : Flow rate, $Nn$ : number of nodes</td>
<td></td>
</tr>
<tr>
<td>System noise</td>
<td>$\log(K_{t,j}) = \log(K_{t-1,j}) + v_{t,j}$</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>$v_{t,j} \sim N(0, \sigma)$</td>
<td></td>
</tr>
<tr>
<td>Observation noise</td>
<td>$N(0, \sigma^o_j)$</td>
<td></td>
</tr>
<tr>
<td>Initial distribution of $\log(K)$</td>
<td>Geological unit : $N(-4, \sigma^{ini}_j)$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fracture zone : $N(-2, \sigma^{ini}_j)$</td>
<td></td>
</tr>
</tbody>
</table>
Assimilation results of water inflows in tanks

PF

Ensemble size = 96

EnKF
Assimilation results of pressure head

PF

Ensemble size = 96

EnKF
Estimation results of log(K)

- **PF**
  - Ensemble size = 96

- **EnKF**
Degeneration of particle filter

- **Prior PDF**
- **Posterior PDF**

Probability distribution of log(K) at 60\textsuperscript{th} day (ensemble size=512)

- PF collapses completely in every time step.

Using more than thousands ensemble member is impractical!
Number of ensemble members in EnKF

Comparison of the errors of pressure head observations
Assimilation results of water inflows with EnKF (unsaturated model)

Ensemble size = 96

Flow rate [m³/day]

Time (day)

TK-1

TK-2

TK-3
Assimilation results of pressure head with EnKF (unsaturated model)

Ensemble size = 96
Estimation results of log(K) (Unsaturated model)

Ensemble size = 96
Non-linearity of unsaturated model

Prior PDF of Permeability ($\log(K)$)

Prior PDF of Pressure head

Ensemble size = 96

Ensemble size = 96
Data assimilation using real observations with EnKF (in progress)

Ensemble size = 96
Data assimilation using real observations (in progress)

Ensemble size = 96
Summary of the results

- EnKF produces good performance under the twin experiments with perfect model and observations.
- Application of particle filter is difficult at acceptable computation cost.
- The non-linearity of the unsaturated model collapses Gaussian distributions of state variables and reduces the accuracy.
- The assimilation result using real observations has biased errors. This implies that the current model has a limitation.
Future tasks

- Model evaluation and refinement
  - More accurate geological model
  - Considering the uncertainty of the groundwater recharge

- Find a solution to deal with non-gaussian distributions in EnKF
  - Gaussian mixture density

- Build the appropriate scheme for successful data assimilation in an actual operation.
  - Assimilation cycle
  - Ensemble size