Uncertainty assessment of streamflow simulation on ungauged catchments using a distributed model and the Kalman filter based MISP algorithm

Authors:

Juan Camilo Múnera¹, Félix Francés¹, Ezio Todini²

(¹) Technical University of Valencia (Spain)
Research Institute of Water Engineering and Environment
Research Group of Hydrological and Environmental Modeling
http://lluvia.dihma.upv.es

(²) University of Bologna (Italy)
Earth Sciences and Environment Department
http://www.geomin.unibo.it/
Introduction

- Important role of Predictive Uncertainty (PU) on flood forecasting and decision making:
  - Flow simulations and forecasting made from models are not error free
  - Growing interest in assessing uncertainty of predictions/simulations
  - Severe economic and social consequences derived from flood emergencies

- Developed methods and tools to assess flood forecasting uncertainty at gauging stations:
  - Hydrologic Uncertainty Processor (Krzysztofowicz, 1999)
  - Bayesian Model Averaging (Raftery, 1993; Raftery et al, 2003; 2005)
  - Model Conditional Processor (Todini, 2008)
  - And others statistical approaches…

- How to estimate PU at ungauged sites?
Introduction

- **Our proposal**: Combining simulations of the hydrological distributed model **TETIS** (Vélez et al, 2001; Francés et al, 2007) with the **MISP** technique (*Mutually Interactive State-Parameter Estimation*) (Todini, 1978)
  - Distributed models: advantage of simulating flows at any point throughout the spatial domain
  - **MISP** is a Kalman filter based algorithm, which:
    - Performs the state-parameter estimation of a discrete time dynamic system
    - Makes alternative use of two interacting filters in parallel, both with minimum variance

- **Filtering Scheme**:
  - Model predictions at an ungauged site are assumed as imperfect observations (as random variables)
  - Incorporating observations and simulations at a gauging station as on-line instrumental variables correlated with flow at the ungauged site
  - Log transformation of all input data to improve Gaussian assumptions of the K.F.
  - Inclusion of a cross covariance term in the covariance matrix of the measurement error (assumption of spatially correlated errors)
Hydrological model TETIS

- Developed at Technical University of Valencia since 1994
- Spatially distributed in regular cells
  - Reproduces spatial variability of hydrological processes
  - Reduces spatial scale effects with regard to lumped models
  - Allows to exploit all available physical and environmental spatial information

- Separate runoff modeling on slopes and channels
  - Each tank drains into the topographical downstream corresponding tank.
  - Drainage area thresholds defined for each type of tank

- Nonlinear channel routing scheme (geomorphologic kinematic wave).
Hydrological model TETIS

- Robust and parsimonious model:
  - Adequate simulation of initial state (includes balance at all times)
  - 6 storage tanks (state variables)
  - 5 external outflows (3 horizontal responses)

- Potential problem in distributed models:
  - Calibration of high number of parameters in each cell from an outflow hydrograph.
  - Proposed solution: Split Effective Parameters Structure (Frances et al, 2007):
    - Phase I: Parameter estimation from all physical and environmental information at each cell
    - Phase II: Global correction factors for each parameter map (effective parameters)
Production parameters at each cell (v.7)

- Vegetation cover index: $\lambda(t)$  
  - Crop factor
- Maximum static capacity: $H_u$  
  - Initial abstractions + upper soil capillary capacity
- Overland flow velocity: $v$  
  - Hill slope stationary velocity
- Interflow velocity: $k_{ss}$  
  - Horizontal macropore upper soil permeability
- Base flow velocity: $k_b$  
  - Upper aquifer permeability
- Infiltration capacity: $k_s$  
  - Vertical upper soil permeability
- Percolation capacity: $k_p$  
  - Vertical deep soil permeability
- Underground losses capacity: $k_{pp}$  
  - Lower aquifer permeability
Calibration Process

- Parameter estimation normally through comparison between simulated and observed values of some state variables
  - Traditionally: Discharge at the basin outlet

- Model performance assessment:
  - Graphic comparison
  - Statistical Objective Functions:
    - Balance Error
      \[ \text{BE} = \frac{\sum (S_i - Q_i)}{\sum Q_i} \times 100 \]
    - Root mean square Error
      \[ \text{RMSE} = \sqrt{\frac{1}{n} \sum (Q_i - S_i)^2} \]
    - Efficiency Index (Nash-Sutcliffe)
      \[ \text{NSE} = 1 - \frac{\sum (Q_i - S_i)^2}{\sum (Q_i - Q_m)^2} \]

- Automatic calibration of correction factors: SCE-UA (Duan et al., 1992; Sorooshian et al, 1993)
Mutually interactive parameter estimation (MISP)

- The first filter performs the minimum variance state estimation, given the parameters set:

\[ x_t = \Phi_{t/t-1} x_{t-1} + \Gamma_{t-1} w_{t-1} \]
\[ z_t = H_t x_t + v_t \]

where:
- \( x_t \): State vector (nx1)
- \( z_t \): Measurement vector (mx1)
- \( \Phi \): State transition matrix (nxn)
- \( \Gamma, H \): Compatibility matrices
- \( v_t, w_t \): Normal independent processes

- The second one performs the minimum variance parameter estimation, given the state estimate, both at the present and previous time steps.

\[ \theta_t = \theta_{t-1} + \Gamma_{t-1}^* w_{t-1}^* \]
\[ z_t^* = H_t^* \theta_t + v_t^* \]

where:
- \( \theta_t \): Parameter vector (px1)
Mutually interactive parameter estimation (MISP)

- Equations required to accomplish a calculation cycle:

  **State update:**
  \[
  v_t = z_t - H_t \hat{x}_{t|t-1} - \bar{v}_{t-1}
  \]
  \[
  K_t = P_{t|t-1} H_t^T (H_t P_{t|t-1} H_t^T + R_{t|t})^{-1}
  \]
  \[
  \hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t v_t
  \]
  \[
  P_{t|t} = (I - K_t H_t) P_{t|t-1}
  \]

  **Time update:**
  \[
  \hat{x}_{t+1|t} = \Phi_{t+1|t} \hat{x}_{t|t} + \Gamma_{t+1} \bar{W}_t
  \]
  \[
  P_{t+1|t} = \Phi_{t+1|t} P_{t|t} \Phi_{t+1|t}^T + \Gamma_{t+1} Q_{t|t} \Gamma_{t+1}^T
  \]

  **Parameters update:**
  \[
  v_t^* = H_t K_t v_t
  \]
  \[
  R_t^* = H_t K_t C_{t|t} K_t^T H_t^T
  \]
  \[
  C_t^* = H_t^* P_{t|t-1}^* H_t^{*T} + R_t^*
  \]
  \[
  K_t^* = P_{t|t-1}^* H_t^{*T} (C_t^*)^{-1}
  \]
  \[
  \hat{\theta}_{t+1|t} = \hat{\theta}_{t|t-1} + K_t v_t^* + \Gamma_{t+1} \bar{W}^*
  \]
  \[
  P_{t+1|t}^* = P_{t|t-1}^* - K_t C_t^* K_t^{*T} + \Gamma_{t+1} Q^* \Gamma_{t+1}^*
  \]
Mutually interactive parameter estimation (MISP)

- Equations required to accomplish a calculation cycle:

**State update:**
- $v_t = z_t - H_t \hat{x}_{t|t-1} - \tilde{v}_{t-1}$
- $K_t = P_{t|t-1} H_t^T \left( H_t P_{t|t-1} H_t^T + R_{t|t} \right)^{-1}$
- $\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t v_t$
- $P_{t|t} = (I - K_t H_t) P_{t|t-1}$

**Parameters update:**
- $v_t^* = H_t K_t v_t$
- $R_t^* = H_t K_t C_{t|t}^* K_t^T H_t^T$
- $C_t^* = H_t^* P_{t|t-1}^* H_t^{*T} + R_{t|t}^*$
- $K_t^* = P_{t|t-1}^* H_t^{*T} \left( C_t^* \right)^{-1}$
- $\hat{\theta}_{t+1|t} = \hat{\theta}_{t|t-1} + K_t^* v_t + \Gamma_{t+1} \bar{w}^*$
- $P_{t+1|t} = P_{t|t-1}^* - K_t^* C_{t|t}^* K_t^{*T} + \Gamma_{t+1} Q_{t+1} \Gamma_{t+1}^{*T}$

- The estimation of the state vector provided by the first filter ($\hat{x}_{t|t}$) is used as observations, in order to estimate the parameter vector $\theta$.
- Optimality is reached after a series of runs through the historical data, as model residuals become less and less correlated.
- At the same time, it is possible to estimate the unknown noise statistics: $\bar{w}$, $\bar{v}$, $Q$, $R$. 

EGU Leonardo Conference Series on the Hydrological Cycle
FLOODS IN 3D: PROCESSES, PATTERNS, PREDICTIONS
Bratislava, Slovakia, 23-25 November 2011
MISP implementation

- System equations (matrix form):

\[
x_t = \Phi_{t/t-1} x_{t-1} + \Gamma_{t-1} w_{t-1}
\]

\[
\begin{bmatrix}
Q_2(t) \\
Q_2(t-1) \\
Q_1(t) \\
Q_1(t-1) \\
Q_{sl}(t) \\
Q_{sl}(t-1)
\end{bmatrix} =
\begin{bmatrix}
\alpha_1 & \alpha_2 & \beta_1 & \beta_2 & \chi_1 & \chi_2 \\
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \beta_1^2 & \beta_2^2 & \chi_1^2 & \chi_2^2 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \chi_1^3 & \chi_2^3 \\
0 & 0 & 0 & 0 & 1 & 0
\end{bmatrix}
\begin{bmatrix}
Q_2(t-1) \\
Q_2(t-2) \\
Q_1(t-1) \\
Q_1(t-2) \\
Q_{sl}(t-1) \\
Q_{sl}(t-2)
\end{bmatrix} +
\begin{bmatrix}
w_{2(t-1)} \\
w_{1(t-1)} \\
w_{sl(t-1)}
\end{bmatrix}
\]

- Parameter equation:

\[
z_t = H_t x_t + v_t
\]

\[
\begin{bmatrix}
z_{1(t)} \\
z_{2(t)} \\
z_{3(t)}
\end{bmatrix} =
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0
\end{bmatrix}
\begin{bmatrix}
Q_2(t) \\
Q_2(t-1) \\
Q_1(t) \\
Q_1(t-1) \\
Q_{sl}(t) \\
Q_{sl}(t-1)
\end{bmatrix} +
\begin{bmatrix}
v_1(t) \\
v_2(t) \\
v_3(t)
\end{bmatrix}
\]

\[
\theta_t = \theta_{t-1} + \Gamma_{t-1}^* w_{t-1}^*
\]

\[
\theta_t = \begin{bmatrix}
\alpha_1 \\
\alpha_2 \\
\vdots \\
\chi_1 \\
\chi_2
\end{bmatrix}
\]

\[
z_t^* = H_t^* \theta_t + v_t^*
\]
Case study: simulation at Baron Fork

- **Distributed Model Intercomparison Project (DMIP2), NOAA/NWS.**
  - Series of experiments to guide NOAA/NWS research into advanced hydrologic models for river and water resources forecasting.

- **Study basins: Baron Fork river and Peacheater Creek (tributary)**
  - Complete description (Smith et al, 2004)
  - Availability of cartographic information of physical and environmental parameters
  - Concurrent time series (1995-2002) of discharges, radar precipitation (NEXRAD), temperature and ETP from Reanalysis (NCEP-NCAR)

- **Basin areas**
  - Baron Fork: 795 km²
  - Peacheater: 65 km²
TETIS model calibration

- Calibration results, Baron Fork at Eldon (oct/2001-sep/2002)

Statistical Index
- BE (%): -14.6
- RMSE (m³/s): 6.61
- NSE: 0.91
### TETIS model calibration

- Calibrated correction factors (cell parameters) for the DMIP2 case study

<table>
<thead>
<tr>
<th>Cell parameter</th>
<th>TETIS Param. decomposition</th>
<th>Calibrated values of R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static storage capacity</td>
<td>$h_u^* = R_1 \cdot h_u$</td>
<td>0.3831</td>
</tr>
<tr>
<td>Index of monthly vegetation density coverage (i=1, 2, ..., 12) for ET.</td>
<td>$\lambda_i^* = R_2 \cdot \lambda_i$</td>
<td>0.7232</td>
</tr>
<tr>
<td>Infiltration capacity at constant rate</td>
<td>$k_s^* = R_3 \cdot k_s$</td>
<td>1.2594</td>
</tr>
<tr>
<td>Direct Runoff hillslope velocity (linear reservoir)</td>
<td>$u_{OF}^* = R_4 \cdot u_{OF}$</td>
<td>2.0</td>
</tr>
<tr>
<td>Percolation capacity at constant rate</td>
<td>$k_p^* = R_5 \cdot k_p$</td>
<td>0.3329</td>
</tr>
<tr>
<td>Interflow velocity (linear reservoir)</td>
<td>$k_{if}^* = R_6 \cdot k_s$</td>
<td>30.00</td>
</tr>
<tr>
<td>Groundwater loss capacity at constant rate</td>
<td>$k_{pp}^* = R_7 \cdot k_p$</td>
<td>0.0</td>
</tr>
<tr>
<td>Baseflow velocity (linear reservoir)</td>
<td>$k_{bf}^* = R_8 \cdot k_p$</td>
<td>114.43</td>
</tr>
<tr>
<td>Streamflow velocity in the river network</td>
<td>$u_{CF}^* = R_9 \cdot u_{CF}$</td>
<td>0.2009</td>
</tr>
</tbody>
</table>
TETIS model validation

- Temporal validation results, Baron Fork at Eldon (oct/1996-sep/2001)

![Graph showing observed and simulated flows with statistical indices]

**Statistical Index**
- BE (%): -11.2
- RMSE (m³/s): 14.89
- NSE: 0.83
TETIS model validation

- Spatial validation results, (Peacheater Cr. At Christie (oct/1996-sep/2002)

Statistical Index
- BE (%): 9.3
- RMSE (m³/s): 0.88
- NSE: 0.67
Results, application of MISP

- MISP results, Peacheater Cr. (oct/1996-sep/2002)

<table>
<thead>
<tr>
<th>Date</th>
<th>Observed flow (assumed unknown)</th>
<th>Updated flow (MISP K.F.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>01/10/99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31/10/99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>01/12/99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31/12/99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>01/04/00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31/01/00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>01/03/00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>01/04/00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>01/05/00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>01/06/00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>01/07/00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>01/08/00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31/08/00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>01/09/01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>02/03/01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>02/04/01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>02/05/01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>02/06/01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>02/07/01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>02/08/01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>01/09/01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>02/10/01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>02/11/01</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Statistical Index**

- BE (%) 9.3
- RMSE (m³/s) 0.83
- NSE 0.71

EGU Leonardo Conference Series on the Hydrological Cycle
FLOODS IN 3D: PROCESSES, PATTERNS, PREDICTIONS
Bratislava, Slovakia, 23-25 November 2011
Results, application of MISP

- Peacheater Creek (15/03/2002 – 29/04/2002)

![Graph showing observed, simulated, and updated flows for Peacheater Creek and Baron Fork.](image-url)
Results, application of MISP

- Peacheater Creek, uncertainty bounds (15/03/2002 – 29/04/2002)

![Graph showing flow data and uncertainty bounds for Peacheater Creek from 15/03/2002 to 29/04/2002.](image)

- Observations and updates for observed and updated flows with lower and upper uncertainty limits.

---

EGU Leonardo Conference Series on the Hydrological Cycle
FLOODS IN 3D: PROCESSES, PATTERNS, PREDICTIONS
Bratislava, Slovakia, 23-25 November 2011
Results, application of MISP

- Peacheater Creek (06/03/1999 – 17/05/1999)

<table>
<thead>
<tr>
<th>Date</th>
<th>Flow (m³/s) Peacheater Cr.</th>
<th>Flow (m³/s) Baron Fork</th>
</tr>
</thead>
<tbody>
<tr>
<td>06/03/99</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>11/03/99</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>16/03/99</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>21/03/99</td>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td>26/03/99</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>31/03/99</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>05/04/99</td>
<td>3.5</td>
<td></td>
</tr>
<tr>
<td>10/04/99</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>15/04/99</td>
<td>4.5</td>
<td></td>
</tr>
<tr>
<td>20/04/99</td>
<td>5.0</td>
<td></td>
</tr>
<tr>
<td>25/04/99</td>
<td>5.5</td>
<td></td>
</tr>
<tr>
<td>30/04/99</td>
<td>6.0</td>
<td></td>
</tr>
<tr>
<td>05/05/99</td>
<td>6.5</td>
<td></td>
</tr>
<tr>
<td>10/05/99</td>
<td>7.0</td>
<td></td>
</tr>
<tr>
<td>15/05/99</td>
<td>7.5</td>
<td></td>
</tr>
</tbody>
</table>
Results, application of MISP

- Peacheater Creek, uncertainty bounds (06/03/1999 – 17/05/1999)

![Graph showing flow rates and uncertainty bounds for Peacheater Creek from 06/03/1999 to 17/05/1999. The graph illustrates observed and updated flows, along with lower and upper uncertainty limits. The dates are marked on the x-axis and flow rates on the y-axis. A small map inset shows the location of Peacheater Creek.]
Results, application of MISP

- Peacheater Creek (15/06/2000 – 10/07/2000)

![Graph showing observed, simulated, and updated flows for Peacheater Creek from 15/06/2000 to 10/07/2000. The graph compares observed flow (unknown), simulated flow (Tetis), updated flow (MISP K.F.), and Qobs3_BF. The dates on the x-axis are 15/06/2000, 15/06/2000, 15/06/2000, 07/06/2000, 07/06/2000, and 07/06/2000. The y-axis shows flow in m³/s for Peacheater Creek and Baron Fork.]
Results, application of MISP

- Peacheater Creek, uncertainty bounds (15/06/2000 – 10/07/2000)

![Graph showing flow rates and uncertainty bounds for Peacheater Creek from 15/06/2000 to 10/07/2000. The graph includes observed flow, updated flow with MISP, and uncertainty bounds at 5% and 95% confidence levels.](image-url)
Conclusions

- A methodology based on the Kalman filter MISP algorithm is presented, aimed to improve predictions performed with a distributed hydrological model at ungauged sites and to estimate the related predictive uncertainty.

- The observed and simulated discharges at the basin outlet are used as instrumental variables in the Kalman filter implementation.

- Importance of including a cross covariance error term in matrix $R$.

- The results of the proposed approach are considered very satisfactory. The NSE was improved from 0.67 to 0.71 for the Peacheater Creek (assumed as ungauged site).

- The Kalman filter based MISP approach allows assess the predictive uncertainty associated to the model prediction (simulation or forecasting mode).

- The uncertainty band is strongly affected by model performance, and is expected that any improvement of the model would reduce predictive uncertainties.
Thank you for your attention!

Contact:
Juan Camilo Múnera
E-mail: juamues1@doctor.upv.es
juancmunera@yahoo.es

Technical University of Valencia (Spain)
Research Institute of Water Engineering and Environment
Research Group of Hydrological and Environmental Modeling
http://lluvia.dihma.upv.es