Developments in non-linear Kalman Ensemble and Particle Filtering techniques for hydrological data assimilation

M. Khaki[a], M. Kuhn[a], J. Awange[a], E. Forootan[b]

[a] *Western Australian Centre for Geodesy and the Institute for Geoscience Research, Curtin University, Perth, Australia.*
[b] *Institute of Geodesy and Geoinformation, Bonn University, Nussallee 17, D-53115, Bonn, Germany*
Email: Mehdi.Khaki@Postgrad.curtin.edu.au

Abstract: Quantifying large-scale (basin/global) water storage changes is essential to understand the Earth's hydrological water cycle. Hydrological models have usually been used to simulate variations in storage compartments resulting from changes in water fluxes (i.e., precipitation, evapotranspiration and runoff) considering physical or conceptual frameworks. Models however represent limited skills in accurately simulating the storage compartments that could be the result of e.g., the uncertainty of forcing parameters, model structure, etc. In this regards, data assimilation provides a great chance to combine observational data with a prior forecast state to improve both the accuracy of model parameters and to improve the estimation of model states at the same time.

Various methods exist that can be used to perform data assimilation into hydrological models. The two frequently used particle-based algorithms suitable for non-linear systems are the Ensemble Kalman Filtering (EnKF) and the Particle Filtering (PF). Despite efficiency and simplicity (especially in EnKF), both methods indicate some drawbacks. To implement EnKF, one should use the sample covariance of observations and model state variables to update a priori estimates of the state variables. The sample covariance can be suboptimal as a result of small ensemble size, model errors, model nonlinearity, and other factors. Small ensemble can also lead to the development of correlations between state components that are at a significant distance from one another where there is no physical relation. On the other hand, the efficiency of PF might be limited due to the fact that after several iterations, most of the estimated weights acts on very few particles, which often causes divergence of the PF-based assimilation.

In this study, a theoretical comparison between latest methods used in the data assimilation framework, to overcome the mentioned problems, is performed. For this, we first introduce the Local Ensemble Kalman Filter (LEnKF) utilizing covariance localization to remove long range spurious correlations and increase the effective ensemble size. Furthermore, to decrease the effect of particle concentration in PF, a re-sampling technique is considered. This technique uses the basic idea of re-sampling to draw independent particles from a specific probability density function (pdf), and then assign them uniform weights. Specifically, we will assess two strategies for re-sampling; one is the sequential importance re-sampling (SIR), and the other is based on the kernel density estimation, which leads to the so-called Regularized Particle Filter (RPF) technique. To evaluate and compare the mentioned data assimilation methods and study their performance, numerical experiments are generated based on the strongly nonlinear (40-dimensional system) Lorenz-96 model. Fundamental issues regarding the assimilation of water storage observations into the spatially extended chaotic system (as is reflected in Lorenz-96) are addressed.