An Adaptive Ensemble Kalman Filter for Land Data Assimilation

Rolf Reichle and Christian Keppenne, in collaboration with Wade Crow (USDA)

Project Goals: The input error parameters required by data assimilation systems are a source of uncertainty in the ensuing analysis. Errors in their specification may be alleviated with an adaptive filtering approach as we developed for the GEOS-5 land data assimilation system.

Project Description: Data assimilation products are sensitive to input observation and model error variances. With very poor input error parameters, assimilation analyses may be worse than model estimates without assimilation. To examine this issue, a suite of experiments was conducted with the GEOS-5 Land Data Assimilation System (G5-LDAS) using synthetic surface soil moisture observations. Each assimilation experiment has a unique set of input error parameters that leads to a unique pair of scalars: the (space and time) average forecast error variance (P_0) and the input observation error variance (R_0) for surface soil moisture. We can thus plot 2D surfaces of filter performance as a function of $\sqrt{P_0}$ and $\sqrt{R_0}$. The figure to the left shows one such surface using the RMSE of surface soil moisture estimates from a non-adaptive assimilation over the entire Red-Arkansas domain as the performance measure.

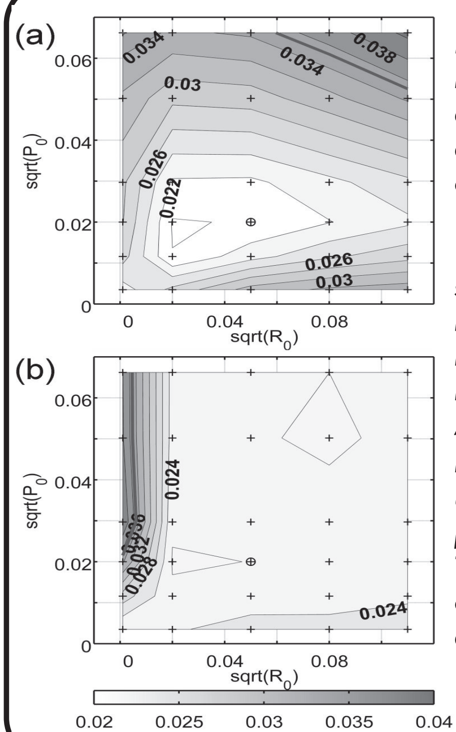


Figure 1: MSE of surface soil moisture for the (a) non-adaptive and (b) adaptive assimilation system as a function of input error parameters: (ordinate) forecast and (abscissa) observation error standard deviation. Units are m^3m^{-3}. Each + indicates the result of a 19-year assimilation integration over the Red-Arkansas domain. The ⊕ indicates the experiment that uses the true input error parameters for the assimilation. Thick gray lines indicate RMSE of the open loop (no assimilation) integration.

The estimation error in surface soil moisture is smallest near the experiment that uses the true error inputs. The minimum estimation error is around 0.02 m^3m^{-3}, down from the open loop (no assimilation) value of 0.035 m^3m^{-3}. The estimation error increases as the input errors deviate from their true values. The figure also indicates where the estimation error surface intersects the open loop error. For grossly overestimated error variances, the assimilation estimates of surface soil moisture are in fact worse than the open loop estimates. Ultimately, the success of the assimilation (measured through independent validation) suggests whether the selected input error parameters are acceptable.

Results: Adaptive filtering methods can assist with the estimation of input error parameters. The central idea behind adaptive filtering is that internal diagnostics of the assimilation system should be consistent with the values that are expected from input parameters provided to the data assimilation system. The most commonly used diagnostics for adaptive filtering are based on the observation-minus-forecast residuals or innovations (computed here as $v_t \equiv E\{y_{t,i} - H_t x_{t,i}^-\}$, where $E\{\cdot\}$ is the ensemble mean operator). For a linear system operating under optimal conditions, the lagged innovations covariance is

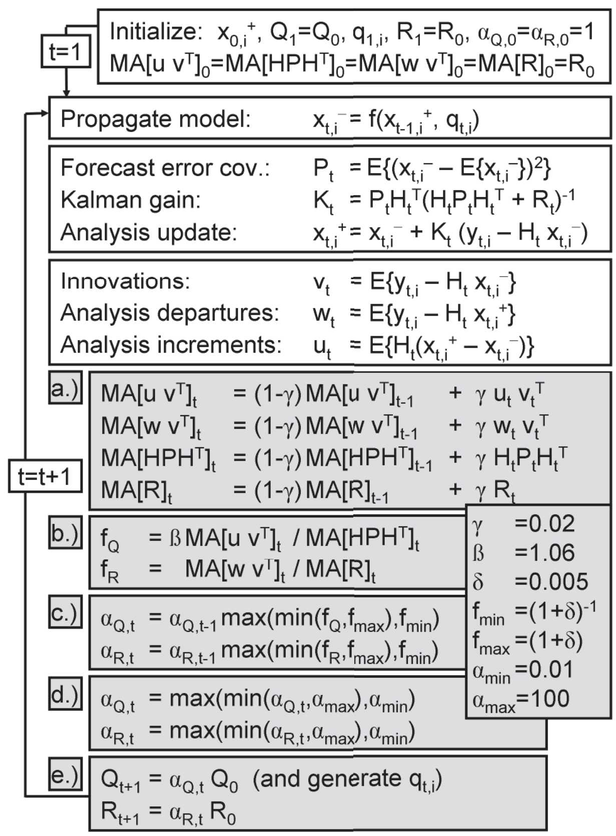
(1) $E[v_t\, v_{t-k}^T] = \delta_{k,0}\, (H_t\, P_t\, H_t^T + R_t)$

where $E[\cdot]$ is the expectation operator and $\delta_{k,0}$ is the Kronecker delta. Equation (1) implies that the innovations sequence is uncorrelated in time and that its covariance is equal to the sum of the forecast error covariance $H_t\, P_t\, H_t^T$ (in observation space) and the observation error covariance R_t. Now recall that the forecast error covariance P depends on the model error covariance Q. If the innovations show less spread than expected, the input error covariances (Q and/or R) are too large, and vice versa. Such information can be used for adaptive tuning of Q and/or R.

Alternative diagnostics are based on the analysis departures $w_t \equiv E\{y_{t,i} - H_t\, x_{t,i}^+\}$ and the (observation space) analysis increments $u_t \equiv E\{H_t(x_{t,i}^+ - x_{t,i}^-)\}$. For linear systems under optimal conditions we have

(2) $E[u_t\, v_t^T] = H_t\, P_t\, H_t^T$

(3) $E[w_t\, v_t^T] = R_t$



Equations (2) and (3) suggest a simple way of estimating Q and R separately by tuning the input error parameters so that the output diagnostics on the left-hand-side of (2) and (3) match the right-hand-side error covariances. This approach is employed in the G5-LDAS as summarized in the flow diagram to the right.

An example of the benefits of the adaptive module is given in Figure 1. The adaptive estimation of input error parameters leads to improved estimates of surface soil moisture regardless of *initial* error estimates, except for the case of severe underestimation of the input observation error variance. The poor performance in this special case is due to technicalities in the implementation and can easily be avoided in applications.

The adaptive filtering module will be used for estimation of retrieval error parameters of NASA satellite products (such as surface soil moisture from AMSR-E) and for the development of a SMAP Level 4 soil moisture assimilation product.

Publications:

Reichle, R. H., W. T. Crow, and C. L. Keppenne, 2008: An adaptive ensemble Kalman filter for soil moisture data assimilation, *Wat. Resour. Res.*, **44**, W03423, doi:10.1029/2007WR006357.

Crow, W. T. and R. H. Reichle, 2008: Comparison of adaptive filtering techniques for land surface data assimilation, *Wat. Resourc. Res.*, **44**, W08423, doi:10.1029/2008WR006883.

Reichle, R. H., and co-authors, 2008: Recent Advances in Land Data Assimilation at the NASA Global Modeling and Assimilation Office, *In: Data Assimilation for Atmospheric Oceanic and Hydrologic Applications*, Springer Verlag DOI: 10.1007/978-3-540-71056-1, Chapter 21.