4.8

ASPECTS OF THE EXTENDED AND ENSEMBLE KALMAN FILTERS FOR LAND DATA ASSIMILATION IN THE NASA SEASONAL-TO-INTERANNUAL PREDICTION PROJECT

Rolf H. Reichle*, R. D. Koster, J. P. Walker, M. M. Rienecker, and P. R. Houser
NASA GSFC, Greenbelt, MD

1. INTRODUCTION

Successful climate prediction at seasonal-to-interannual time scales may depend on the optimal initialization of the land surface states, in particular soil moisture (Koster and Suarez 2001). Such optimal initialization can be achieved by assimilating soil moisture observations into the land model prior to the forecast. We assess the performance of the Extended Kalman filter (EKF) and the Ensemble Kalman filter (EnKF) for soil moisture estimation when used with the Catchment Land Surface Model (CLSM) of the NASA Seasonal-to-Interannual Prediction Project.

2. KALMAN FILTERING

Since for seasonal forecasts we are only interested in the estimates at the start time of the prediction, sequential assimilation methods like Kalman filters are ideally suited to the task. The major differences between the EKF and the EnKF concern (i) the approximation for nonlinearities of the hydrologic model and the measurement process, (ii) the treatment of horizontal correlations in the model or measurement errors, (iii) the range of model errors that can be represented, (iv) the ease of implementation, and (v) computational efficiency.

The EKF approximates the error covariance propagation by linearizing the model. However, the computational demand resulting from the error covariance integration is prohibitive unless further approximations are made (Gelb 1974). In this study we use the EKF implementation of Walker and Houser (2001) in which all correlations between different catchments are neglected. The EnKF, on the other hand, nonlinearly propagates an ensemble of model trajectories from which sample forecast error covariances are derived at the update time (Evensen 1994, Reichle et al. 2001). Its main approximation is the size of the ensemble. To simplify the comparison with the EKF, we also neglect horizontal error correlations in the EnKF.

3. TWIN EXPERIMENT

We conduct a twin experiment over the south-eastern United States by assimilating synthetic observations of near-surface soil moisture once every three days into the CLSM. The twin experiment starts with a model integration that serves as the “true” solution and is meant to represent nature. Next, we integrate the model again over the same time period but with an intentionally poor initial condition as well as different forcing data and model parameters. Collectively, these “wrong” inputs and parameters represent our imperfect knowledge of the true land processes. The resulting fields constitute our best guess prior to assimilating the remote sensing data and will be referred to as the “prior” solution. The synthetic observations used in the assimilation are derived from the true fields by adding random measurement noise. In particular, we generate synthetic observations of the surface soil moisture with an error of 2 % (volumetric) once every three days for all catchments.

4. RESULTS AND DISCUSSION

Figure 1 shows the time average (root-mean-square) actual errors of the moisture content variables from Feb to Dec 1987. The actual errors are the differences between the true soil moisture (from the control experiment) and its EKF or EnKF estimate. Obviously, the errors are higher for the surface moisture content than for the root zone and profile moisture contents. This is because the surface moisture content varies on time scales of a day or less, while we assimilate observations only once every three days. Inbetween observation times, errors in the model time scales and in the forcing (notably in precipitation) degrade the surface estimates. The situation is different for the root zone and profile moisture contents. These lower layers exhibit greater memory and variations in their moisture content occur over longer time scales. Consequently, short-term errors in the forcing do not significantly impact the root zone and profile estimates.

Overall, we find that the EKF and the EnKF are able to derive satisfactory estimates of soil moisture. In the case of the EnKF, just four ensemble members prove sufficient (Table 1). The EKF and the EnKF (with four ensemble members) show comparable performance for comparable computational effort. For five or more ensemble members, the EnKF outperforms the EKF, albeit at greater computational expense. This is attributed to the EnKF’s flexibility in representing non-additive model errors such as errors in certain forcing variables or errors in model parameters.

For both the EKF and the EnKF we find that the actual estimation errors are typically larger than filter-derived forecast and analysis error variances. The numerical differentiation scheme used in the EKF requires frequent checks in order to avoid that the error covariances diverge or lose their positive definiteness. Although these checks interrupt the integration of the error covariances and information from earlier updates is partially lost, they are not a major source of error.

* *Corresponding author address:* R. Reichle, NASA GSFC, Code 974, Greenbelt, MD 20771, email: reichle@janus.gsfc.nasa.gov

It is straightforward to extend the EnKF to account for catchment-to-catchment error correlations. Such correlations could for example arise from large-scale errors in the forcing or from unmodeled lateral fluxes such as river or groundwater flow. Moreover, satellite data are likely to exhibit horizontal error correlations. If such error correlations do in fact exist, the EnKF will be able to spread information laterally, in particular from observed to unobserved catchments. The same is not computationally feasible for the EKF. The importance of horizontal error correlations is a topic of active research.

In summary we can say that the EnKF is more robust and offers more flexibility in covariance modeling (including horizontal error correlations). This leads to its superior performance in this study and makes the EnKF a promising approach for soil moisture initialization of seasonal climate forecasts.

Acknowledgments

This research was sponsored by the NASA Seasonal-to-Interannual Prediction Project. Thanks to Max Suarez and Sarith Mahanama for their help.

References

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *J. Geophys. Res.*, **99**, 10 143–10 162.

Gelb, A., Ed., 1974: *Applied Optimal Estimation*, The M.I.T. Press, Cambridge, MA.

Koster, R. D. and M. J. Suarez, 2001: Soil moisture memory in climate models, *J. Hydrometeorol.*, in press.

Reichle, R., D. B. McLaughlin, and D. Entekhabi, 2001: Hydrologic data assimilation with the Ensemble Kalman filter, *Mon. Weather Rev.*, in press.

Walker, J. P. and P. R. Houser, 2001: A methodology for initializing soil moisture in a global climate model: Assimilation of near-surface soil moisture observations, *J. Geophys. Res.*, **106**, 11 761–11 774.

		prior	EKF	EnKF				
Ensemble size *N*	[–]	n/a	n/a	4	5	10	100	500
Surface m.c.	[%]	6.09	3.32	3.35	3.33	3.26	3.21	3.21
Root zone m.c.	[%]	5.29	1.72	1.71	1.70	1.59	1.49	1.50
Profile m.c.	[%]	5.59	1.64	1.62	1.60	1.49	1.39	1.40

Table 1: Actual errors of the moisture content (m.c.) in volumetric percent (root-mean-square average over all catchments from Feb to Dec 1987).

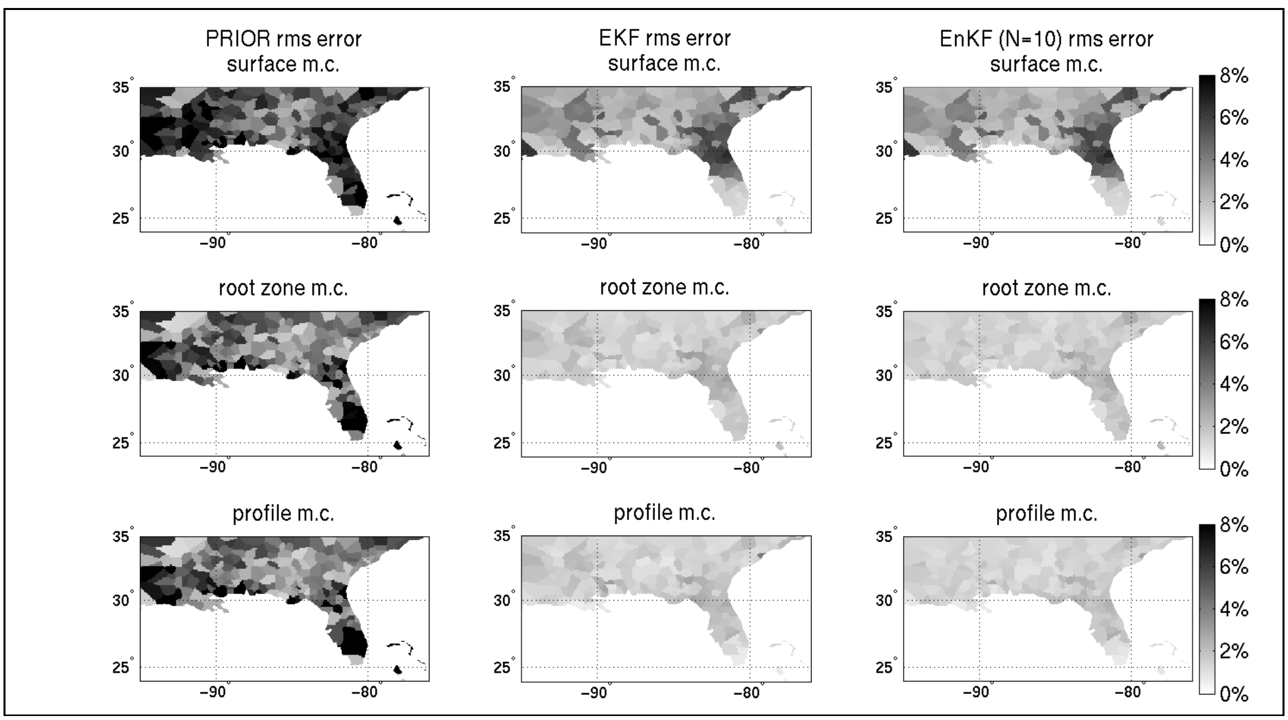

Figure 1: Time-average error of the moisture content (m.c.) prior to the assimilation (first column), for the EKF (second column), and for the EnKF with *N*=10 ensemble members (last column). The first, second, and third rows show the errors for the surface, root zone, and profile soil moisture content, respectively. The average is from Feb to Dec 1987 in the root-mean-square (rms) sense. Units are volumetric moisture percent.