Extended versus Ensemble Kalman Filtering for Land Data Assimilation

ROLF H. REICHLE

Goddard Earth Sciences and Technology Center, University of Maryland, Baltimore County, Baltimore, and Hydrological Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, Maryland

JEFFREY P. WALKER

The University of Melbourne, Parkville, Victoria, Australia

RANDAL D. KOSTER AND PAUL R. HOUSER

Hydrological Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, Maryland

(Manuscript received 1 October 2001, in final form 5 April 2002)

ABSTRACT

The performance of the extended Kalman filter (EKF) and the ensemble Kalman filter (EnKF) are assessed for soil moisture estimation. In a twin experiment for the southeastern United States synthetic observations of near-surface soil moisture are assimilated once every 3 days, neglecting horizontal error correlations and treating catchments independently. Both filters provide satisfactory estimates of soil moisture. The average actual estimation error in volumetric moisture content of the soil profile is 2.2% for the EKF and 2.2% (or 2.1%; or 2.0%) for the EnKF with 4 (or 10; or 500) ensemble members. Expected error covariances of both filters generally differ from actual estimation errors. Nevertheless, nonlinearities in soil processes are treated adequately by both filters. In the application presented herein the EKF and the EnKF with four ensemble members are equally accurate at comparable computational cost. Because of its flexibility and its performance in this study, the EnKF is a promising approach for soil moisture initialization problems.

1. Introduction

Climate prediction at seasonal-to-interannual timescales depends on accurate initialization of the slowly varying components of the earth's system, most notably sea surface temperature (SST) and soil moisture. While tropical SST is often the dominant source of predictability, its influence appears to be mostly limited to the Tropics (Koster et al. 2000b). Skill in the prediction of summertime continental precipitation and temperature anomalies in the extratropics may instead depend on the initialization of soil moisture and other land surface states. Since soil moisture controls the partitioning of the latent and sensible heat fluxes to the atmosphere, it can influence precipitation recycling.

The initialization of the land surface states for a seasonal climate forecast can be accomplished by assimilating soil moisture observations into the land model up to the start time of the prediction. With assimilation we attempt to combine the information from the observations and the model in an optimum way. 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 well-known extended Kalman filter (EKF) can be used for nonlinear applications, but the computational demand resulting from the error covariance integration limits the size of the problem (Gelb 1974). For this reason, the EKF has been used mostly for problems that focus on the estimation of the vertical soil moisture profile (Katul et al. 1993; Entekhabi et al. 1994). More recently, Walker and Houser (2001) have applied the EKF to soil moisture estimation across the North American continent by neglecting all horizontal error correlations and treating surface hydrological units (catchments) independently. This yields an effectively low-dimensional filter.

The ensemble Kalman filter (EnKF) is an alternative to the EKF (Evensen 1994). The EnKF circumvents the expensive integration of the state error covariance matrix by propagating an ensemble of states from which the required covariance information is obtained at the time of the update. Reichle et al. (2002) applied the EnKF to soil moisture estimation and found that it performed well against a variational assimilation method. Since the variational approach generally requires the adjoint of the hydrologic model, which is not usually available and is difficult to derive, the obvious choices for advanced land assimilation algorithms are the EKF and the EnKF. There are many variants of the EKF and the EnKF that have been used in meteorology and oceanography, notably reduced-rank square root algorithms (Verlaan and Heemink 1997), particle filters (Pham 2001), methods that use pairs of ensembles (Houtekamer and Mitchell 1998), and hybrid approaches that combine ensembles with reduced-rank approaches (Heemink et al. 2001; Lermusiaux and Robinson 1999) or with variational methods (Hamill and Snyder 2000). In this paper, we focus on the relative merits of using the traditional EKF and EnKF for soil moisture assimilation.

The major differences between the EKF and the EnKF are (i) the approximation of nonlinearities of the hydrologic model and the measurement process (the EKF uses a linearized equation for the error covariance propagation while the EnKF nonlinearly propagates a finite ensemble of model trajectories), (ii) the range of model errors that can be represented (the EnKF can account for a wider range of model errors), (iii) the ease of implementation (the EKF requires derivatives of the nonlinear hydrologic model, evaluated numerically or from a tangent-linear model), (iv) computational efficiency (it must be determined how many ensemble members are needed in the EnKF to match the performance of the EKF), and (v) the treatment of horizontal correlations in the model or measurement errors (the EKF cannot account for horizontal error correlations in large systems for computational reasons). Insights into many important issues can be gained from low-dimensional versions of both filters.

Although approximate nonlinear filters such as the EKF and the EnKF have been found to work well in some applications, their value in a particular nonlinear problem cannot be assessed a priori but must be determined by simulations (Jazwinski 1970). We investigate the above differences in the context of soil moisture initialization for seasonal prediction using synthetic data in a twin experiment. Since all uncertain inputs are known by design, such experiments are well suited for a first assessment of algorithm performance. Tests with actual observations will be conducted in future studies. For retrospective analysis, surface soil moisture can be retrieved from the Scanning Multifrequency Microwave Radiometer (SMMR) for the period 1979–87 (Owe et al. 2001). These retrievals are derived from the 6.6-GHz (C band) and 37-GHz channels. Similar retrievals should soon be available from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E). In the future, passive 1.4-GHz (L band) sensors should also become available (Kerr et al. 2001).

5. Summary and conclusions

In this paper, we compare two promising data assimilation methods for soil moisture initialization in seasonal climate prediction. The extended Kalman filter (EKF) and the ensemble Kalman filter (EnKF) were used to assimilate synthetic surface soil moisture observations into the Catchment Model, with model error parameters calibrated against actual estimation errors. The best results are obtained for both filters when the model error in the root zone excess is large compared to the model errors in the surface excess and the catchment deficit. Using the calibrated filter parameters we find that the EKF and the EnKF produce satisfactory estimates of soil moisture.

The EKF and the EnKF (with four ensemble members) show comparable performance for comparable computational effort. For 10 or more ensemble members, the EnKF outperforms the EKF. This is ascribed to the EnKF's flexibility in representing nonadditive model errors. The actual estimation errors of the EnKF converge quickly with increasing ensemble size, even though the filter-derived (expected) error covariances are noisy for small ensembles. The numerical differentiation scheme used in the EKF requires frequent checks in order to avoid divergent error covariances or loss of 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.

The normalized innovations are found to be inconsistent with a standard normal distribution. This is because our representation of model errors cannot fully account for the effects of uncertainties in the forcing and imperfectly known model parameters that we use in our twin experiment. Nonlinearities in the land model generate skewness in the distribution of ensemble states. But this skewness information is only very approximately used in the EnKF update and is not available in the EKF. Fortunately, the nonlinearities are not a dominant source of error, because the local linearization strategy of the EKF is for the most part successful and because the nature of the soil moisture bounds limits the actual estimation errors.

Catchment-to-catchment error correlations could 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. The present paper compares the EKF and EnKF under the assumption that horizontal error correlations can be neglected. The importance of such correlations is a topic of active research. If horizontal error correlations turn out to be important, information can be spread laterally, in particular from observed to unobserved catchments. When horizontal error correlations are taken into account in the EnKF, small error correlations associated with observations that are far apart must be filtered out (Mitchell and Houtekamer 2000). For computational reasons, the EKF must be approximated using a rank-reduction technique such as the reduced-rank square root method (Verlaan and Heemink 1997).

Before soil moisture assimilation can become a routine tool for seasonal climate prediction, many more questions will need to be addressed. Important areas of research include the investigation of multivariate assimilation using more Catchment Model prognostic variables as states, the direct assimilation of radiances as opposed to soil moisture retrievals, and the assimilation of other types of remote sensing data such as soil temperatures or vegetation parameters. Finally, soil moisture estimates from the assimilation must then be shown to improve the accuracy of seasonal climate forecasts. 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 slightly superior performance in our study and makes the EnKF a promising approach for soil moisture initialization of seasonal climate forecasts.