An ensemble approach for attribution of hydrologic prediction uncertainty

Andrew W. Wood[1,2] and Dennis P. Lettenmaier[1]

Received 9 May 2008; revised 19 June 2008; accepted 24 June 2008; published 30 July 2008.

[1] Hydrologic prediction errors arise from uncertainty in initial moisture states (mainly snowpack and soil moisture), in boundary forcings (primarily future precipitation and temperature), and from model structure and parameter uncertainty. We evaluate the relative importance of initial condition and boundary forcing uncertainties using a hindcast-based framework that contrasts Ensemble Streamflow Prediction (ESP) with an approach that we term "reverse-ESP". In ESP, a hydrologic model with assumed perfect initial conditions (ICs) is forced by a forecast ensemble resampled from observed meteorological sequences; whereas reverse-ESP combines an ensemble of resampled ICs with a perfect meteorological forecast. The framework shows that in northern California, US, ICs yield streamflow prediction skill for up to 5 months during the transition between the wet and dry seasons, whereas during the reverse transition, climate forecast information is critical. In southern Colorado, IC knowledge outweighs climate prediction skill for shorter periods due to a more uniform precipitation regime. **Citation:** Wood, A. W., and D. P. Lettenmaier (2008), An ensemble approach for attribution of hydrologic prediction uncertainty, *Geophys. Res. Lett.*, *35*, L14401, doi:10.1029/2008GL034648.

1. Introduction

[2] In regions such as the western US where winter snowpack accumulation and melt dominates the surface hydrologic cycle, seasonal streamflow volumes can be forecasted with useful accuracy at lead times of up to six months. Such accuracy is due primarily to the systematic (physically-constrained) evolution of initial land surface moisture states from snow that accumulates in the winter wet season and melts in spring and summer. This phenomenon has been exploited for well over 50 years in streamflow forecast schemes that use regression or other methods to relate spring snowpack to the subsequent spring and summer runoff [*Pagano et al.*, 2004]. In contrast, land surface initial conditions (ICs) for forecasts made following the relatively dry summer period do not include a snowpack that influences future hydrologic fluxes. As a result, forecasts depend to a greater extent on soil moisture conditions and predictability in future meteorological inputs to the hydrologic system [*Maurer and Lettenmaier*, 2003].

[3] Understanding sources of forecast skill has long been a topic of interest in the study of geophysical systems, and examples of relevant theory and research are prevalent, for instance, in the atmospheric sciences [e.g., *Gustaffson et al.*, 1998]. *Lorenz* [1975] terms the two major sources of predictability in such systems as the "first" (related to ICs) and "second" (related to future boundary forcings) kinds. *Collins and Allen* [2002] present a framework for comparing the magnitudes of each type of predictability and conversely the potential for errors in each source to diminish forecast skill. The framework contrasts the forecast variance arising from a forecast ensemble based on small perturbations to the initial atmospheric states, and the forecast variance arising from an ensemble of boundary forcings, to the internal, climatological variance of the atmospheric system. The same perturbation concept can provide insights into the sources of predictability for hydrologic systems. Indeed, a major goal for hydrologic science identified by *National Research Council* [1999] is to improve our understanding of uncertainty in hydrologic systems, including as a technique the use of hydrologic perturbation experiments.

[4] The response of the atmosphere to perturbations differs from that of the land surface. Perturbations in atmospheric states often grow chaotically toward divergent outcomes [e.g., *Shuka*, 1981], whereas on the land surface, negative feedbacks often outweigh positive, reinforcing feedbacks, and consequently tend to dampen anomalies. For example, soil drying by evapotranspiration and percolation decreases the rate of further drying by these mechanisms due to reduced moisture availability. Wet soil moisture has an opposite effect: it increases the rate of drainage from the soil, decreases the proportion of precipitation that infiltrates into the soil column, and can increase evaporation in non-energy limited situations, all of which accelerate the restoration of soil moisture to normal levels. Positive feedbacks on surface moisture also exist. Land-atmosphere coupling [e.g., *Koster and Suarez*, 2004], for example, can involve the augmentation of boundary layer moisture by evaporation, leading to increased precipitation and further increases to soil moisture and evaporation. These feedbacks are likely to be secondary in magnitude to the negative internal feedbacks of the land surface system, but the comparison has not been quantified. The balance of these land surface feedbacks toward the negative implies that absent a strongly anomalous meteorological forecast, current drought and flood conditions will evolve toward climatological averages.

[5] The importance of land surface moisture to hydrologic prediction is well established [*Day*, 1985; *Maurer and Lettenmaier*, 2003; *Berg and Mulroy*, 2006; *Mahanama et al.*, 2008]. Advances in seasonal climate prediction over the last two decades have motivated interest in improving hydrologic forecasts via the incorporation of climate predictions into operational hydrologic prediction approaches. The seasonally and geographically varying influence of land surface initial conditions relative to climate forcings on

[1]Department of Civil and Environmental Engineering, University of Washington, Seattle, Washington, USA.
[2]Now at 3TIER, Inc., Seattle, Washington, USA.

0094-8276/08/2008GL034648

hydrologic predictability dictate limits, however, to the potential value of improvements in climate forecast skill. We demonstrate an ensemble-based approach for exploring these limits and characterizing the relative importance of the two sources of predictability in hydrologic systems. We examine two hydroclimatic settings, one that has strongly seasonal precipitation typical of much of the western U.S., and one that has a more uniform seasonal precipitation regime.

2. ESP and "Reverse-ESP"

[6] To evaluate the relative importance of hydrologic initial conditions and climate forecast error as sources of seasonal runoff forecast uncertainty, we use a framework that contrasts Ensemble Streamflow Prediction (ESP) [*Day*, 1985] as developed at the U.S. National Weather Service (NWS) with an approach that we term "reverse-ESP" (*rev*ESP). Traditional ESP uses a hydrologic model driven by observed meteorological forcings up to the time of forecast to estimate what are (unrealistically) considered error-free land surface initial moisture conditions, and then produces ensemble forecasts by running the model into the future using model forcings (primarily precipitation and temperature) resampled from historical meteorology. Thus ESP represents forecast uncertainty due to boundary forcing uncertainties only, a shortcoming addressed by *Wood and Schaake* [2008] among others. Boundary forcing uncertainty also may be underestimated if a short historical record is used to supply future forcings. ESP is applied operationally by the NWS River Forecast Centers and other water management-focused groups in both the public and private sectors. Climate predictions can be readily incorporated into seasonal-lead ESP by adjusting historical forcings to reflect an alternative climate forecast [e.g., *Perica et al.*, 2000]. In practice, uncertainties in simulated ICs for ESP are reduced by adjusting model spin-up forcings or state variables to bring model outputs into closer agreement with observations [*Seo et al.*, 2003].

[7] The *rev*ESP approach reverses the ESP construct by driving the model with resampled meteorological ensembles during the spinup period (up to the date of forecast) to create an ensemble of ICs that are each paired with observed (assumed perfect) meteorology in the future period. Whereas ESP derives its skill from ICs and the ensemble spread comes from boundary forcing uncertainty, *rev*ESP skill comes from boundary forcings and the ensemble spread from IC uncertainty. Just as ESP's future forcings reflect climatological variations, climatological variations are represented in *rev*ESP's IC ensemble, which consists of hydrologic states for the same day of the year from a historical period. These two constructs are illustrated in Figure 1, along with a depiction of climatology to generate both the ICs and future conditions as a "forecast". Climatology lacks knowledge of either ICs or future forcings, and serves as a reference forecast for assessment of predictability ESP and *rev*ESP.

[8] The *rev*ESP is patently artificial: in a real-time forecast setting, the future forcing is not known, and the climatological range of ICs almost certainly overestimates uncertainty about the true moisture states. ESP and *rev*ESP are here intended to provide the endpoints in a hindcast diagnosis framework. If the hindcast-based ensembles are compared

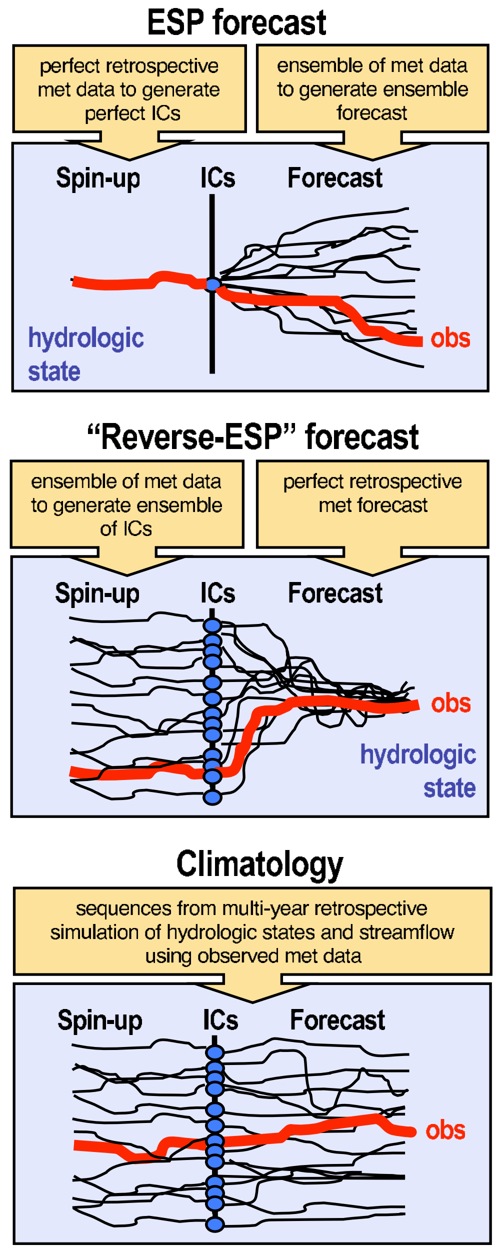

Figure 1. Schematic illustrating hydrologic state evolution during spinup and forecast using (a) ESP approach, (b) reverse ESP approach, and (c) climatology.

with observations (e.g., of streamflow), the resulting error for both *rev*ESP and ESP includes model error. The ESP and *rev*ESP hindcasts can also be evaluated against retrospective model simulations of hydrology driven by observations, in which case the forecast errors are related solely to future climate uncertainty and the predictability is solely related to knowledge of initial moisture state. We take the latter route, and use retrospective model simulations for comparison.

3. Case Study Evaluation

[9] We illustrate the ESP/*rev*ESP uncertainty attribution framework at two locations in the western U.S. shown in

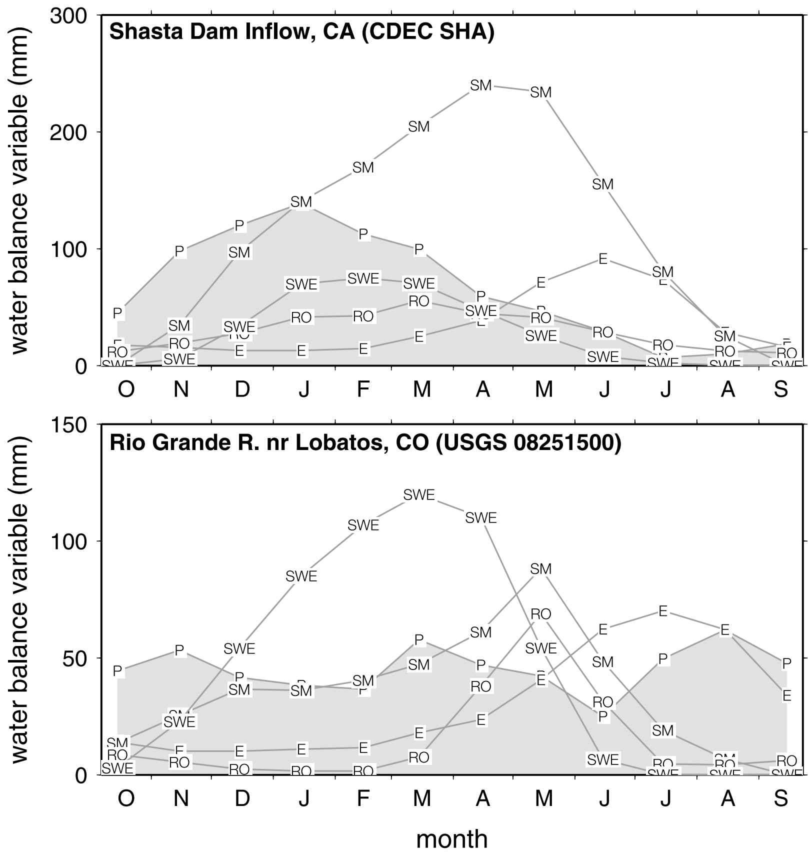

Figure 2. Mean observed precipitation (P) and simulated water balance variables – soil moisture (SM), snow water equivalent (SWE), runoff (RO) and evaporation (E)—for the two study basins. Model SM is reduced by the lowest mean monthly value so that the plotted values shown only the active range.

Figure S1: the upper Sacramento River drainage to Shasta Reservoir in northern California, and the Rio Grande River drainage upstream of Lobatos, Colorado (basin average elevations of 1,475 and 2,870 meters, and drainage areas of 16,430 and 11,930 km^2, respectively).[1] The Variable Infiltration Capacity hydrologic model [*Liang et al.*, 1994] was implemented at a daily time step as described by *Wood et al.* [2005], to which the reader is referred for details of model calibration, data sources and the generation of the retrospective forecast dataset that is used in the analysis presented here. In brief, the hindcast data set provides four ESP and *rev*ESP forecast ensembles per year, for the period 1979–1999, with initialization dates on the 25th day of January, April, July and October, each of lead approximately 6 months. Each forecast ensemble contained 21 members drawn from the period 1979–1999.

[10] The observed precipitation and simulated monthly hydrologic water balances of the two study basins are shown in Figure 2. The Shasta Reservoir drainage receives precipitation predominantly in the winter months, most of which is stored in the soil column and as snow water equivalent (SWE), and is depleted as runoff and evaporation during the spring and summer. The lowest runoff occurs in the late summer and early fall (note, runoff is approximately equivalent to streamflow because channel routing lags are minimal at a monthly time step for small basins). The Rio Grande River drainage exhibits a similar snow accumulation and melt cycle driven by winter precipitation, but also receives summer precipitation that produces a minor runoff response before the low streamflow period in the winter. The four initialization dates for the forecasts capture markedly different hydrologic conditions and future climate expectations.

[11] To quantify predictability for the forecasts, we use a ratio of variances framework [*Collins and Allen*, 2002] that compares the mean squared error (MSE) from each of the two ensemble forecast approaches with the MSE resulting when climatology is used as an ensemble prediction, which is equivalent to the climatological variance. To specify the error and predictability metrics, let $\mathbf{h_{sm}(t)}$ be the hydrologic forecast value (e.g., streamflow, soil moisture) at lead time $\mathbf{t}$ from initial hydrologic state $\mathbf{s}$ and boundary meteorological forcing $\mathbf{m}$. ICs are the hydrologic states for the same day of year from a period of $\mathbf{S}$ years. Meteorological forcings are in this case the daily temperature and precipitation model input sequences from a period of $\mathbf{M}$ historical years, where the start of each sequence begins on the day of year of the ICs.

[12] Omitting the functional dependence on time (both lead time and initialization day of the year) from subsequent formulations, the mean squared error of ESP, the future meteorological condition ensemble forecast, is

$$E[ESP] = \frac{1}{S}\sum_{s=1}^{S}\left[\frac{1}{M}\sum_{m=1}^{M}(h_{sm}-h_{ss})^2\right]; \qquad (1)$$

[1]Auxiliary materials are available in the HTML. doi:10.1029/2008GL034648.

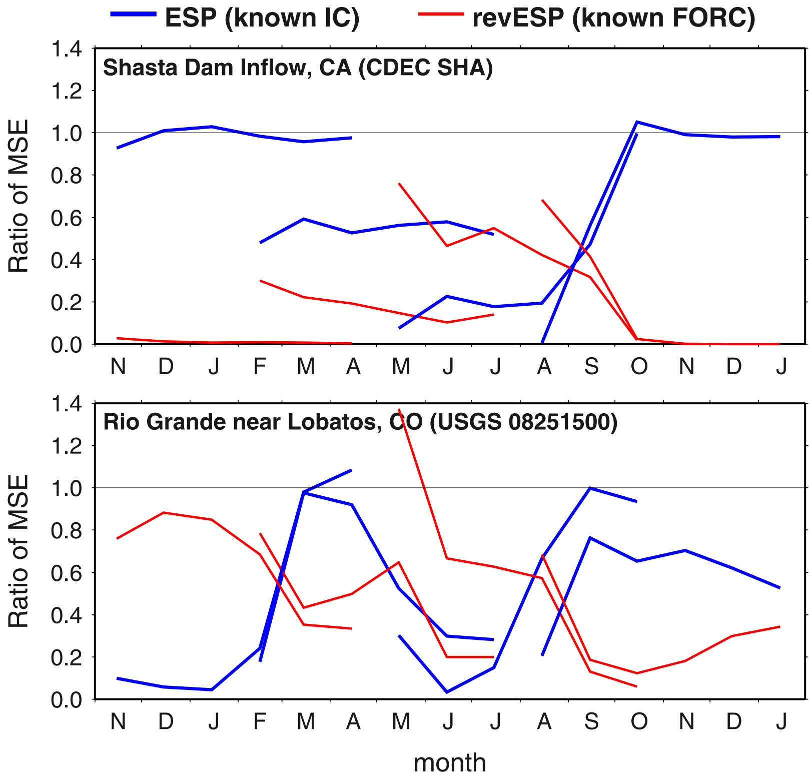

Figure 3. The ratios of mean squared error (MSE) of the ESP and *rev*ESP 6-month lead forecast ensembles to MSE from using a climatological sample as a forecast, for two basins and for four starting dates.

the forecast error of *rev*ESP, the initial condition ensemble forecast is

$$E[revESP] = \frac{1}{M}\sum_{m=1}^{M}\left[\frac{1}{S}\sum_{s=1}^{S}(h_{sm} - h_{mm})^2\right]; \quad (2)$$

and the error of the naïve climatology (Clim) forecast is

$$E[C\ \lim] = \frac{1}{S}\sum_{s=1}^{S}\left[\frac{1}{S}\sum_{t=1}^{S}(h_{tt} - h_{ss})^2\right]. \quad (3)$$

[13] When the ratio of the MSE of either forecast ensemble to the MSE of the climatology is less [greater] than 1, the forecast is more [less] skillful than a climatological forecast. Ratios greater than 1 arise from combinations of the ICs with forcing ensembles that produce a wider range of hydrologic responses than are present in the simulated climatology. The ratio approaches zero for a perfect forecast.

4. Results

[14] The ratios of MSE for ESP/Clim and *rev*ESP/Clim for streamflow forecasts from the four initiation dates and for all monthly forecast lead times are plotted for both locations in Figure 3. For the late October ESP forecast at the Shasta Reservoir location, the MSE at all lead times is close to the climatology ensemble MSE, whereas the *rev*ESP forecast (known future meteorology with climatological IC spread) leads to a nearly perfect hydrologic forecast. This result is consistent with the depleted moisture states (see SWE and soil moisture in Figure 2) at this time of year, and implies that any hydrologic forecast skill must come from knowledge of future climate.

[15] In contrast, for the Rio Grande River, and the same (late October) forecast date, knowledge of ICs (ESP) greatly reduces the forecast variance for up to four months, whereas knowledge of future forcings (*rev*ESP) has a much smaller contribution to forecast skill for equivalent leads; and the signals swap relative importance for leads 5 and 6 months. This result seems counter-intuitive, given that SWE and soil moisture are also relatively depleted in the basin at the forecast initialization time (Figure 2). A number of points are relevant. The late summer precipitation regime (not present in the Shasta drainage) results in substantial variations in soil moisture, which then influence late fall and winter runoff. Also, the snow accumulation season begins earlier than in the Shasta Reservoir drainage, with the result that flows during the forecast period are mostly influenced by soil moisture rather than by late fall and winter precipitation. Although late fall and winter runoff is a small fraction of spring snowmelt runoff, hence of lesser importance to water resources, significant forecast skill may be attainable from knowing ICs (e.g., ESP) for flows in the Rio Grande drainage for this period.

[16] The subsequent three forecasts in the Shasta Reservoir drainage show the increasing importance of knowing ICs and decreasing importance of future climate prediction to hydrologic prediction skill, which is a result of the progressive accumulation of moisture in the soil and snowpack and (for the late May and August forecasts) the onset of the dry summer period. By October, for both late May and August forecasts, the contribution of ICs is gone, and the effect of fall precipitation variability in determining

hydrologic response becomes pronounced, leading to nearly zero forecast variance given known future forcings (*rev*ESP). The January, April and August forecasts in the Rio Grande River, in contrast, show an alternating influence of ICs versus climate forcings. Knowledge of late January ICs (ESP) becomes less important than knowledge of climate during the transitional months from a baseflow regime to a snowmelt regime (February to April), when climate variation determines the timing of melt. For the late April forecast start date, knowledge of ICs is again very important, as the strong melt-based generation of streamflow is underway. Toward the end of the summer, as snowmelt effects wane, precipitation regains control over the hydrologic response, and consequently the *rev*ESP forecast again provides the greater reduction in forecast error.

5. Discussion and Conclusions

[17] The contrasting results of ESP and *rev*ESP across different seasons and locations indicate the large variation in the influence of errors in initial conditions and the influence of climate forecast skill on hydrologic prediction. In the Shasta Reservoir drainage, knowledge of ICs provides streamflow prediction skill for up to 5 months during the transition from the wet to the dry season, whereas during the reverse transition, climate forecast information is critical. In the Rio Grande River drainage, two regimes control hydrologic persistence. The first is a baseflow regime in winter, and the second is a melt regime in spring. During the winter/spring transition, IC knowledge is more important than climate information, as in the Shasta Reservoir inflow case. The periods during which IC predictability dominates are shorter than in California, however, because they are followed by greater temporal variability in the climate-runoff relationship caused in part the late summer precipitation. The ratios of errors for each ensemble are indicative of the fractional contribution of the unknown feature of that ensemble to forecast errors. For example, the ratio of nearly 1 for the October ESPs show that future climate uncertainty almost completely controls future flow forecast uncertainty.

[18] While these results will come as no surprise to those engaged in the practice of hydrologic prediction, they have important implications for hydrologic predictability that have not generally been recognized. Common practices such as aggregating climate forecast assessments across space and/or up to a season in the assessment of forecast skill, for instance, may obscure the forecasts' performance at the scales of spatial and temporal variability that matter for hydrologic or other end uses. The ESP/*rev*ESP construct helps determine the tradeoff between improved ICs (e.g., via land data assimilation) versus improved climate forecast accuracy.

[19] Our results suggest that the emphasis placed by hydrologic forecasters on obtaining better information about future climate versus initial conditions should vary during the year and by forecast lead time, yet commonly used ESP-type frameworks lack a mechanism for representing IC errors. Fortunately, initiatives such as the Hydrologic Ensemble Prediction Experiment (HEPEX) [*Schaake et al.*, 2007] are re-evaluating ensemble hydrologic prediction approaches to address IC and climate forecast related uncertainties. A central tenet of HEPEX is the requirement that real-time forecasts be accompanied by hindcasts produced via consistent methods. This article underscores the importance of such hindcasts to quantification and attribution of hydrologic prediction skill.

[20] **Acknowledgments.** We acknowledge the helpful suggestions of Frank Weber, Sarith Mahanama and one anonymous reviewer. This article was supported by 3TIER, Inc. and by grants to the University of Washington under NASA Agreement NNX06AE34G and NOAA Cooperative Agreements NAOAR50AR4310015 and NA07OAR4310458.

References

Berg, A. A., and K. A. Mulroy (2006), Streamflow predictability in the Saskatchewan/Nelson River basin given macroscale estimates of the initial soil moisture status, *Hydrol. Sci.*, *51*, 642–654.

Collins, M., and M. R. Allen (2002), Assessing the relative roles of initial and boundary conditions in interannual to decadal climate predictability, *J. Clim.*, *15*, 3104–3109, doi:10.1175/1520-0442(2002)015.

Day, G. N. (1985), Extended streamflow forecasting using NWSRFS, *J. Water Resour. Plann. Manage.*, *111*, 157–170.

Gustafsson, N., E. Kallen, and S. Thorsteinsson (1998), Sensitivity of forecast errors to initial and lateral boundary conditions, *Tellus, Ser. A*, *50*, 167–185.

Koster, R. D., and M. J. Suarez (2004), Suggestions in the observational record of land–atmosphere feedback operating at seasonal time scales, *J. Hydrometeorol.*, *5*, 567–572.

Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges (1994), A simple hydrologically based model of land surface water and energy fluxes for GSMs, *J. Geophys. Res.*, *99*, 14,415–14,428.

Lorenz, E. N. (1975), Climate Predictability. The Physical Basis of Climate Modeling, *WMO GARP Publ. Ser.*, *16*, 132–136.

Mahanama, S. P. P., R. D. Koster, R. H. Reichle, and L. Zubair (2008), The role of soil moisture initialization in subseasonal and seasonal streamflow predictability: A case study in Sri Lanka, *Adv. Water Resour.*, doi:10.1016/j.advwatres.2008.06.004, in press.

Maurer, E. P., and D. P. Lettenmaier (2003), Predictability of seasonal runoff in the Mississippi River basin, *J. Geophys. Res.*, *108*(D16), 8607, doi:10.1029/2002JD002555.

National Research Council (1999), *Hydrologic Science: Priorities for the U.S. Global Change Research Program: An Initial Assessment*, Natl. Acad. Press, Washington, D. C.

Pagano, T., D. Garen, and D. Sorooshian (2004), Evaluation of official western US seasonal water supply outlooks, 1922–2002, *J. Hydrometeorol.*, *5*, 896–909.

Perica, S., J. C. Schaake, and D.-J. Seo (2000), Hydrologic application of global ensemble precipitation forecasts, paper presented at Conference on Hydrology, Am. Meteorol. Soc., Long Beach, Calif.

Schaake, J. C., T. M. Hamill, R. Buizza, and M. Clark (2007), HEPEX: The Hydrological Ensemble Prediction Experiment, *Bull. Am. Meteorol. Soc.*, *88*, 1541–1547.

Seo, D. J., V. Koren, and N. Cajina (2003), Real-time variational assimilation of hydrologic and hydrometeorological data into operational hydrologic forecasting, *J. Hydrometeorol.*, *4*, 627–641.

Shukla, J. (1981), Dynamical predictability of monthly means, *J. Atmos. Sci.*, *38*, 2547–2572.

Wood, A. W., and J. C. Schaake (2008), Correcting errors in streamflow forecast ensemble mean and spread, *J. Hydrometeorol.*, *9*, 132–148.

Wood, A. W., A. Kumar, and D. P. Lettenmaier (2005), A retrospective assessment of National Centers for Environmental Prediction climate model–based ensemble hydrologic forecasting in the western United States, *J. Geophys. Res.*, *110*, D04105, doi:10.1029/2004JD004508.

D. P. Lettenmaier, Department of Civil and Environmental Engineering, University of Washington, Seattle, WA 98195-2700, USA.

A. W. Wood, 3TIER, Inc., Seattle, WA 98121, USA. (awood@3tiergroup.com)