

Quantifying Streamflow Forecast Skill Elasticity to Initial Condition and Climate Prediction Skill

ANDREW W. WOOD, TOM HOPSON, AND ANDY NEWMAN

Research Applications Laboratory, NCAR, Boulder, Colorado

LEVI BREKKE

Research and Development Office, U.S. Bureau of Reclamation, Denver, Colorado

JEFF ARNOLD

Institute for Water Resources, U.S. Army Corps of Engineers, Seattle, Washington

MARTYN CLARK

Research Applications Laboratory, NCAR, Boulder, Colorado

(Manuscript received 24 October 2014, in final form 21 October 2015)

ABSTRACT

Water resources management decisions commonly depend on monthly to seasonal streamflow forecasts, among other kinds of information. The skill of such predictions derives from the ability to estimate a watershed's initial moisture and energy conditions and to forecast future weather and climate. These sources of predictability are investigated in an idealized (i.e., perfect model) experiment using calibrated hydrologic simulation models for 424 watersheds that span the continental United States. Prior work in this area also followed an ensemble-based strategy for attributing streamflow forecast uncertainty, but focused only on two end points representing zero and perfect information about future forcings and initial conditions. This study extends the prior approach to characterize the influence of varying levels of uncertainty in each area on streamflow prediction uncertainty. The sensitivities enable the calculation of flow forecast skill elasticities (i.e., derivatives) relative to skill in either predictability source, which are used to characterize the regional, seasonal, and predictand variations in flow forecast skill dependencies. The resulting analysis provides insights on the relative benefits of investments toward improving watershed monitoring (through modeling and measurement) versus improved climate forecasting. Among other key findings, the results suggest that climate forecast skill improvements can be amplified in streamflow prediction skill, which means that climate forecasts may have greater benefit for monthly-to-seasonal flow forecasting than is apparent from climate forecast skill considerations alone. The results also underscore the importance of advancing hydrologic modeling, expanding watershed observations, and leveraging data assimilation, all of which help capture initial hydrologic conditions that are often the dominant influence on hydrologic predictions.

1. Introduction

Every day, streamflow forecasts are used to support decisions by reservoir operators and water managers who strive to balance a range of competing objectives. The need for better streamflow forecasts—from minutes

to seasons—is perennially raised in studies related to water management (e.g., Raff et al. 2013), and large investments are being made in research, science, and technology that are relevant to streamflow prediction, for example, in weather and climate forecasting, in land surface modeling, and in meteorological monitoring, among other areas. Yet, there is evidence that operational streamflow forecast quality has not substantially improved in the last decade or more (Welles et al. 2007; Pagano et al. 2004); thus, there is a pressing need to understand where forecasting investments have the

Corresponding author address: Andrew Wood, Research Applications Laboratory, NCAR, P.O. Box 3000, Boulder, CO 80307-3000.

E-mail: andywood@ucar.edu

DOI: 10.1175/JHM-D-14-0213.1

greatest potential to benefit specific streamflow predictions made to support water management.

Streamflow fluctuations are driven both by runoff discharging from the watershed's moisture stores, including soil moisture (SM), groundwater, snowpack, and the channel network itself, and by meteorological forcings to the contributing watershed. Streamflow dynamics are thus determined by two major contributing factors: 1) watershed initial moisture conditions and 2) future weather for short-range forecasts or future climate for monthly-to-seasonal (M2S) forecasts. In operational practice, M2S streamflow prediction skill attributable to climate forecasts is low, while skill attributable to initial conditions varies from low to very high (Wood and Schaake 2008). The largest predictability at seasonal scales arises in basins with large winter snow accumulation. The process through which snowmelt raises soil moisture, generates runoff, and routes runoff through a stream network to produce streamflow is relatively slow, providing useful forecast accuracy at lead times of up to six months (Wood et al. 2005; Harrison and Bales 2015). The smallest seasonal streamflow predictability is found at the end of a climatologically dry period and preceding a wetter one, such that initial hydrologic conditions (IHCs) provide little contribution to future flows relative to future climate inputs. In this case, nearly all forecast skill at M2S lead times derives from the skill of seasonal climate forecasts (SCFs).

The predictability arising from IHCs has been exploited for nearly a century in seasonal streamflow forecast practice (Pagano et al. 2014), either through statistical methods (Garen 1992; Pagano et al. 2004) or dynamical (model based) systems (e.g., Hydrologic Research Laboratory Staff 1972). Operational practice has made fewer strides to incorporate potential SCF predictability. Promising efforts in the latter area are the new NWS Hydrologic Ensemble Forecast Service, which allows for SCF incorporation from dynamical models (DeMargne et al. 2014), and the Australian Bureau of Meteorology seasonal forecasts based on a statistical combination of IHCs and SCFs through empirical and dynamical models (Wang et al. 2012).

In contrast, the land surface hydrology research community over the last two decades has been investigating the potential for SCFs to advance seasonal flow prediction and has more recently tried to quantify the relative influence of each predictability source. Wood et al. (2005) verified ensemble hindcasts from a single hydrology model to assess the marginal benefits of SCFs relative to IHCs for western U.S. streamflow forecast locations. More recently, several research efforts using SCFs from the National Multimodel

Ensemble (Kirtman et al. 2014) have confirmed the potential for improved runoff and soil moisture prediction through incorporating SCFs into hydrologic prediction approaches, focusing on the United States (Mo and Lettenmaier 2014) and Europe (Thober et al. 2015). Related research has demonstrated the influence of both sources of skill on regional- to continental-scale streamflow predictability, using one or multiple land surface models (e.g., Maurer and Lettenmaier 2003; Berg and Mulroy 2006; Mahanama and Koster 2003; Mahanama et al. 2008; Koster et al. 2010; Mahanama et al. 2012).

Recent efforts in this vein have employed an uncertainty attribution framework that contrasts the streamflow forecast variance arising from an ensemble of IHCs with the forecast variance arising from an ensemble of boundary forcings. Wood and Lettenmaier (2003, hereafter WL03) and Wood and Lettenmaier (2008, hereafter WL08) investigated seasonal streamflow forecast predictability and used climatological variance as a basis for perturbing IHCs and SCFs. Climatological variance was estimated using multidecade-length historical simulation forcings and model states. The IHC ensemble contains model moisture (e.g., soil, snow, canopy, and routed flow) variables on the day of year of the forecast initialization, from all years, and the SCF ensemble contains historical weather sequences from the initialization day of year forward through the forecast period, from all years. This idealized framework contrasted forecast uncertainty from two hindcasting "end point" experiments: one in which "perfect" IHCs are combined with climatological SCFs, and one in which climatological IHCs are paired with "perfect" SCFs. The former is identical to the traditional NWS Ensemble Streamflow Prediction (ESP; Day 1985) forecasting approach, while the latter was termed "reverse ESP." Though useful for illustrating variations in the behavior of different watersheds, WL08 noted that the reverse-ESP component is "patently artificial," because realistic IHC uncertainty is typically much less than climatological uncertainty. Operational forecasters can readily distinguish dry watershed conditions from wet ones, particularly where accumulated watershed moisture is visible in the form of snow.

The WL08 predictability construct has since been applied in a number of predictability studies. Shukla and Lettenmaier (2011) replicated WL08 using the relative error metric from WL03 for 48 hydrologic units spanning the contiguous United States (CONUS) and later extended the analysis to a global domain in Shukla et al. (2013). The work generally corroborated the earlier findings in demonstrating that IHC contributions to seasonal hydrologic predictability are highest in the arid

and snowmelt-dominated hydroclimates (which can be seasonally periodic). Li et al. (2009) applied the construct to the southeastern United States, and in addition explored changes to IHC influence when climate model forecasts are used instead of the ESP climatology. Paiva et al. (2012) focused the technique on the Amazon River basin, concluding that the importance of IHCs indicated potential value for data assimilation. Singla et al. (2012) used the framework (although with different terminology) to study soil moisture and flow predictability in France, determining that the value of SCFs varied regionally, depending on the influence of climate on IHCs. Most recently, Staudinger and Seibert (2014) extended the construct to include model parameter uncertainty in a study of 21 Swiss catchments and found persistence of IHCs ranging from 2 months to at least a year.

It is important to note that these studies have all embraced the perfect-model assumption; that is, in assessing hydrologic predictability, the forecasts are compared to model simulations that are driven with meteorological observations, but not to observed streamflow. The latter choice would recognize a third source (i.e., neither SCF nor IHC) of uncertainty in the assessment: model error (due to structure, physics, and/or parameter uncertainty) that plays a hidden role in perfect-model studies. If model persistence in streamflow simulations is stronger than observed, for instance, IHC influence on predictability could be overestimated. Perfect-model predictability estimates for the same rivers in separate studies can differ when different modeling approaches are used. Indeed, Yossef et al. (2013) highlight differing predictability estimates from those reported in Shukla et al. (2013) for several regions and hypothesize that these arise from process differences such as the inclusion of routing (in the Amazon) and snow modeling (St. Lawrence). Van Dijk et al. (2013) found that model prediction skill can be related to model skill in retrospective simulation, suggesting that the use of calibrated models in the perfect-model framework likely reduces the model error component and potentially errors in quantifying predictability. But calibration cannot eliminate model errors, and because model performance is often judged via one or a few metrics of streamflow, which is a process-integrating variable, the influence of the remaining model error on idealized predictability results is very difficult to quantify.

Limitations notwithstanding, this study extends the WL08 approach toward understanding hydrologic predictability by considering intermediate fractions of uncertainty for IHCs and SCFs that span the range from perfect knowledge to climatological uncertainty. This extension enables the calculation of forecast skill

elasticities—gradients in forecast skill relative to predictor skill—a metric that has not been previously formulated. The practical goal of this approach is to provide realistic insights into the potential benefits of efforts to improve streamflow forecasting, for different types of M2S streamflow forecasts, to support water resources program managers who may already have insight into their costs.

2. Models, data, and methods

We apply a variational ensemble streamflow prediction assessment (VESPA) approach through the generation of retrospective predictions (hindcasts) using a CONUS-wide watershed dataset, supporting demonstration, and analysis for a broad range of hydroclimatic settings. The following subsections describe the study's models, data, methods, and metrics before offering conclusions.

a. A CONUS-wide watershed modeling dataset

We use a CONUS-wide collection of lumped Sacramento Soil Moisture Accounting (SAC-SMA) and SNOW-17 watershed models, forcings, and observed streamflow that is described in Newman et al. (2015). The watersheds from this “large sample” modeling collection are a subset of the USGS HCDN-2009 (Lins 2012). Because these basins have minimal human influence, they are almost exclusively smaller, headwater-type basins. Daily forcings for the period 1980–2010 were generated for each basin by areally averaging the daily, 1-km gridded Daymet (Thornton et al. 2012) meteorological dataset over the watershed drainage areas, generating the input forcings—precipitation and temperature—required to drive the daily simulation models used in this study. Potential evapotranspiration (PET), another required forcing, was derived from Daymet solar radiation via the Priestley–Taylor method (Priestley and Taylor 1972). Daily streamflow data were compiled from the USGS National Water Information System.

The SNOW-17 and SAC-SMA model parameters were calibrated to match simulated and observed daily streamflow for all basins using the shuffled complex evolution global optimization routine (Duan et al. 1992, 1993), with the single objective of minimizing daily root-mean-square error (RMSE). The forcing development and model development approach are described in Newman et al. (2015). From the 671 basins in Newman et al. (2015), we selected 424 basins (Fig. 1) by choosing up to 30 basins in each of 18 HCDN regions that had the highest Nash–Sutcliffe efficiency (NSE) scores in validation and excluding any basins with an NSE lower than

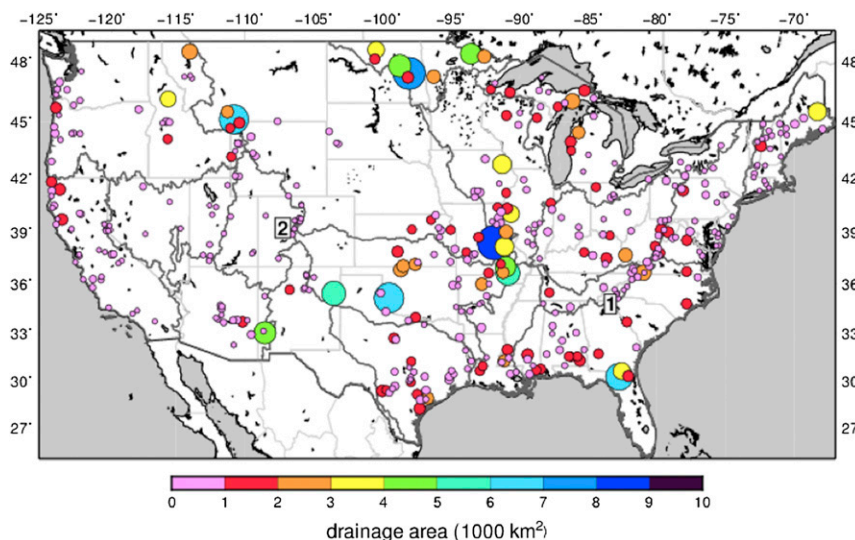


FIG. 1. The 424 study locations, with the color and relative size of the plotting symbol indicating drainage area, and HCDN region. The two locations used to illustrate detailed watershed results are denoted by numbered squares: 1 is the Chattooga River and 2 is the Crystal River.

0.5. Some regions had fewer than 30 sites: the Souris–Red–Rainy and Rio Grande regions, for example, had 8 and 7, respectively.

Two locations were selected to illustrate streamflow predictability contrasts. The Chattooga River near Clayton, Georgia (USGS 002177000), is a rain-driven watershed with a contributing drainage area of 530 km² and a mean basin elevation of 760 m MSL. The mean channel slope is 7.1 m km⁻¹, and river length is 73 km. Less than 2% of contributing drainage area is covered by storage and approximately 96% is forested. Annual precipitation is 1750 mm, with rainfall intensities reaching 133 mm day⁻¹ on average once every 2 years. January minimum temperature averages -1°C and snowfall is rare. The Chattooga River upstream flows southward along the South Carolina–Georgia border from its northernmost headwaters in North Carolina, joining the Tallulah River downstream of the gauge to form the major inflows to Lake Tugalo, held back by the Tugalo Dam.

The Crystal River above Avalanche Creek near Redstone, Colorado (USGS 009081600), is a snowmelt-driven basin of 428 km², flowing westward from headwaters in the Rocky Mountains, west of the Continental Divide, before turning north to join the Roaring Fork River, a tributary to the upper Colorado River. Relative to the Chattooga River, it is steeper (channel slope is 26.7 m km⁻¹), higher (mean elevation is 3109 m MSL), shorter (length is 49.6 km), and drier, with an annual precipitation of 830 mm and rainfall intensities of 36 mm day⁻¹ expected on average once each 2 years.

The drainage area is 60% forested, containing 2% storage and lakes, and an average January minimum temperature of -17°C leads precipitation to fall as snow equivalent (SWE) values on the order of 500 mm at the nearest snow-observing stations, the McClure Pass and North Lost Trail SNOTEL sites.

The observed precipitation and simulated monthly hydrologic water balances of the two study basins are shown in Fig. 2. The Chattooga River drainage receives a relatively uniform precipitation input throughout the year and negligible amounts fall as snow. The seasonal cycle of evapotranspiration, strongest in the summer, creates a moderate seasonal cycle in soil moisture and runoff, leading to the lowest streamflows occurring in summer. The Crystal River drainage receives moderately higher precipitation in the winter and spring than in the summer, but precipitation falls as snow in the months of October–April, on average. Note that the peak simulated basin areal average SWE in this figure is not expected to match the observed SWE at the nearest (or any) SNOTEL station location, due not only to their different spatial representation but to a myriad of terrain-, forcing-, and model-related factors. April sees a mixture of rain, snowfall, and snowmelt, spurring rises in soil moisture and runoff as SWE declines. High evaporation in summer months reduces SM and consequently runoff, which continues to decline through the fall and into the next winter snow accumulation period. The Chattooga and Crystal River models were well calibrated, exhibiting daily streamflow calibration

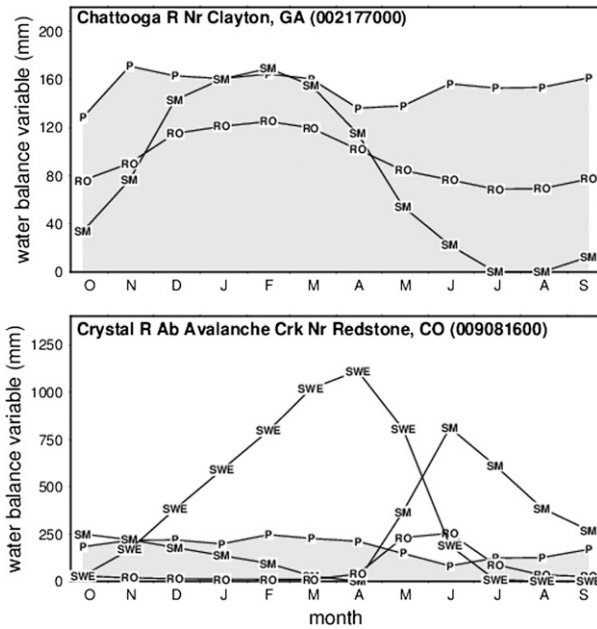


FIG. 2. Mean observed precipitation P and simulated water balance variables—active SM, SWE, and runoff (RO)—for the two study basins. For model SM, we subtract the lowest mean monthly value of the year so that the plotted values show only the active range of variation (the subtracted min values are 26 cm in August and 78 cm in April for the Chattooga, where SWE is zero, and Crystal River locations, respectively).

(validation) NSE values of 0.90 (0.78) and 0.90 (0.86), respectively.

b. Variational ensemble streamflow prediction assessment approach and application

We investigate combinations of different fractions of uncertainty for IHCs and SCFs, spanning the range from perfect knowledge to climatological uncertainty. The VESPA approach is illustrated in Fig. 3. In contrast to the ESP and reverse-ESP techniques (Figs. 3a,b) that combine a single perfect IHC or SCF with a climatological ensemble for the alternate source of predictability, VESPA uses all members of the ensemble of IHCs in turn to initialize an ensemble of SCFs. The IHC and SCF ensemble variances are scaled to between zero and the climatological variance taken from a continuous retrospective simulation dataset (Fig. 3c). In this study, the retrospective simulation dataset spanning 1981–2010 provided $N_i = 30$ initialization dates on a given day of year and $N_f = 30$ meteorological forecast members.

To form an ensemble of IHCs for a selected initialization day of year (e.g., 1 April) and historical year i , the simulated IHC moisture state vector \mathbf{M}_i is linearly blended with \mathbf{M}_j from each of the historical years j in $1 - N_i$. Each blended moisture vector creates an IHC

ensemble member. The blending function is a simple weighted average, using a weight w_{IHC} that varies from zero to one. For a given weight and hindcast initialization year i , each blended forecast initialization vector \mathbf{M}_{ij} (i.e., for initialization year i and blending year j) is

$$\mathbf{M}_{ij} = \mathbf{M}_i(1 - w_{\text{IHC}}) + \mathbf{M}_j w_{\text{IHC}}$$

for each year j in $1 - N_i$.

The state vector \mathbf{M} for the SNOW-17 and SAC-SMA models contains seven primary carryover moisture variables: the soil water components—upper- and lower-zone tension and free-water contents, the lower-zone primary water content, as well as the additional impervious area (ADIMP) and SWE. Channel input from SAC-SMA to the routing model (the unit hydrograph method) was also blended with this linear weighting approach, including inputs beginning several days prior to the forecast date, back to the time of concentration of the watershed. The moisture vector components are of course specific to the models being used and could in some models include energy-related variables. Through oversight, several temperature-related variables that could have been included in the initial snow state vector here (such as antecedent temperature index) were omitted. Subsequent assessment (not shown) indicates that their exclusion has negligible impact on the results, in part because of the time-averaged nature of the predictands being assessed.

The climate forecast ensemble members are constructed via a similar blend approach, but with an added step to account for the forecast being a time series rather than a single vector associated with a single date. In each retrospective forecast, the year of historical forcings (daily precipitation and temperature) beginning on the forecast initialization date represents the perfect meteorological and climate forecast \mathbf{C}_i for the initialization year i . Each member of the SCF ensemble \mathbf{C}_{ij} is created by blending \mathbf{C}_i with the forcings \mathbf{C}_j from the matching calendar period in each ensemble forecast year j in turn, according to

$$\mathbf{C}_{ij} = \mathbf{C}_i(1 - w_{\text{SCF}}) + \mathbf{C}_j w_{\text{SCF}}$$

for each year j in $1 - N_f$.

This operation is applied to the monthly averages of forecasts because the combination of forcings time series cannot be a simple weighted average of the daily forcings, which would create unrealistic daily weather patterns. To create realistic daily patterns matching the blended monthly values, the daily values of each month in \mathbf{C}_i are shifted (for temperature) or scaled (for precipitation) so that their monthly averages equal the

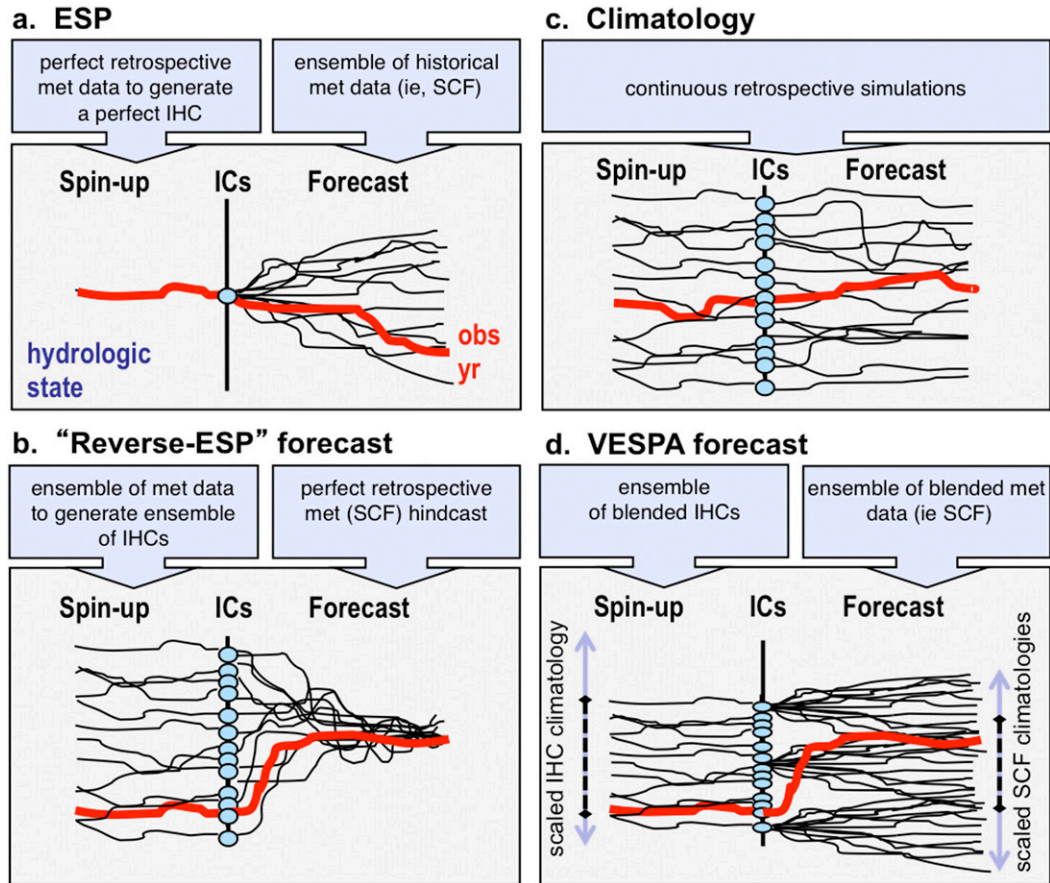


FIG. 3. Schematic illustrating hydrologic simulation in the VESPA framework. The end points of (a) ESP and (b) reverse-ESP concept pair perfect and climatological knowledge in IHCs and SCFs, where (c) climatology for IHCs and climate forcings is defined from continuous retrospective hydrologic simulations. (d) The VESPA approach combines intermediate blends (depicted by the black dashed lines) of climatological IHC and SCF variance (ranges depicted by purple arrows), exploring combinations of variance between zero (perfect knowledge) and full climatology (no knowledge). For clarity, only three combinations of the SCF ensemble with IHC members are drawn.

weighted monthly averages \mathbf{C}_{ij} . If the weight w_{SCF} equals zero, each member \mathbf{C}_{ij} represents a perfect climate forecast in this idealized framework. If w_{SCF} equals one, \mathbf{C}_{ij} has the daily pattern of \mathbf{C}_i with the monthly climate of \mathbf{C}_j . Note that the meteorological ensemble members created via this blending procedure are more similar to each other than traditional ESP ensemble members because their daily sequences arise from the \mathbf{C}_i trace alone. Although the broad variation of climate-scale amplitudes in the ensemble still leads to substantial ensemble spread, an exploration of alternative blending procedures, or the use of a weather generator in this context, may lead to improved replication of the traditional ESP, if desired.

In each case, the weights indicate the amount of uncertainty present in each predictability source: zero signifies zero uncertainty, or perfect knowledge, whereas

unity signifies complete uncertainty, as estimated from the historical period climatology. The square of the weight equals the fraction of climatological variance present in the scaled uncertainty. ESP can be represented by $w_{IHC} = 0$ and $w_{SCF} = 1$, and reverse ESP can be represented by $w_{IHC} = 1$ and $w_{SCF} = 0$. In this study, we assess the impact of nine different weight values ($\mathbf{w} = 0, 0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 0.95, 1.0$) for each predictor, leading to 81 different combinations of IHC and SCF uncertainty, or 81 hindcasts per watershed location per initialization month. Given \mathbf{w} , the approximate percentage of climatological variance explained in the predictability source $\mathbf{v} = 100(1 - \mathbf{w}^2) = (100, 99.8, 99, 94, 75, 44, 19, 10, 0)$. We treat \mathbf{v} as a measure of predictability source skill.

The hindcast period, which is a basis for estimating climatological uncertainty, is the 30-yr retrospective

period 1981–2010. All of the nonzero weight combinations lead to a total ensemble comprising 30×30 or nominally 900 members, although the ensemble members in which the verifying year (the blend year) matched the forecast year or the initial condition year were later removed during skill assessment. Twelve forecast initialization dates are used: the first day of each calendar month, for the retrospective forecast period of 30 years. The number of 30-yr-long watershed model simulations required for the study totaled approximately 370 million, not including many additional simulations performed during method development. The runs were executed on the NCAR–Wyoming supercomputer.

c. Evaluation metrics and forecast skill elasticities

We evaluate three predictands—forecasts of mean streamflow for 1-, 3-, and 6-month periods with zero lead time. The study objective is to discriminate IHC and SCF influences on prediction skill and uncertainty; thus, the forecasts are verified against simulated streamflow forced by meteorological observations. Consequently, the forecast errors depend solely on the SCF and IHC errors and do not include modeling error. For the (largely) illustration purposes of this paper, we apply one metric to define forecast skill, the squared correlation (coefficient of determination) of forecast ensemble medians with observations r^2 , which estimates the fraction of climatological variance explained by the forecast. For readability in the text below, we describe r^2 as a percentage rather than a fraction.

The assessment of streamflow forecast skill resulting from multiple combinations of IHC and SCF skill allows the calculation of gradients in forecast skill relative to IHC and SCF skill, which represent streamflow forecast skill elasticities. Elasticity E is defined in this study as the unit change in flow forecast skill for a unit change in predictability source skill. Elasticity is different for each predictability source, and our findings suggest that it is not constant across the predictor skill domain: E varies depending on the IHC and SCF skill of the system for which it is calculated. For example, the elasticity of flow forecast skill with respect to SCF or IHC skill can be different in a system with substantial uncertainty in both IHCs and SCFs than in a system with near-zero uncertainty in each predictor. We calculate E from the linear gradients in flow forecast skill between adjacent IHC and SCF variance weight combination points. Such gradients can also be calculated differently (e.g., by averaging over a larger region of the forecast skill surface, or by fitting an analytical surface—or in one dimension, line—to the hindcasting results and then calculating derivatives analytically).

In this paper, we calculate elasticities for variance weight combinations that approximately represent the current state of M2S forecasting practice, that is, the ESP approach. In essence, we ask the following question: Given our current practice, how do improvements in skill for either predictability source impact streamflow forecast skill? From the low error ranges in the streamflow simulations of calibrated models, we infer here that the model-based estimates of watershed moisture (IHCs) also have relatively low uncertainty from a year-over-year standpoint, whereas operational climate forecasts used in ESP have nearly climatological uncertainty. Consequently, for flow forecast skill elasticity to IHC skill, we calculate the elasticities across the predictor skill interval v_{IHC} between 94% and 75% skill. We opt to average this calculation over two SCF skill levels v_{SCF} of 10% and 0%, although one could also use only one v_{SCF} level or average over more levels. For flow forecast skill elasticity to SCF skill, we calculate the elasticities across the v_{SCF} interval representing 0%–10% skill, and we then average results at two IHC skill (v_{IHC}) levels, 94% and 75%. Using the notation $[v_{\text{IHC}}, v_{\text{SCF}}]$ to denote a variance weight combination point, an example of the calculation of streamflow forecast skill elasticities around this ESP point for IHCs and SCFs, respectively, is the following:

$$E_{\text{IHC}}[\text{ESP}] = \left\{ \frac{S(F[94, 10]) - S(F[75, 10])}{94\% - 75\%} + \frac{S(F[94, 0]) - S(F[75, 0])}{94\% - 75\%} \right\} / 2$$

and

$$E_{\text{SCF}}[\text{ESP}] = \left\{ \frac{S(F[94, 10]) - S(F[94, 0])}{10\% - 0\%} + \frac{S(F[75, 10]) - S(F[75, 0])}{10\% - 0\%} \right\} / 2.$$

In the numerators, $S(F[\cdot]) - S(F[\cdot])$ represents the skill deltas of streamflow forecasts, which in this example are given by their measured r^2 values. In the denominators, the skill deltas for the moisture (IHC) and climate (SCF) predictors are calculated from the explained variance values for the predictor intervals being evaluated, which is equivalent to the predictors' prescribed r^2 (from v). If a different skill metric were used, $S(\cdot)$ expressions would also appear also in the denominators of the equations above to indicate the skill deltas for the moisture (IHC) and climate (SCF) predictors. Elasticities are calculated separately for each predictand (1-, 3-, and 6-month flows) and for each forecast initialization date. An elasticity of zero

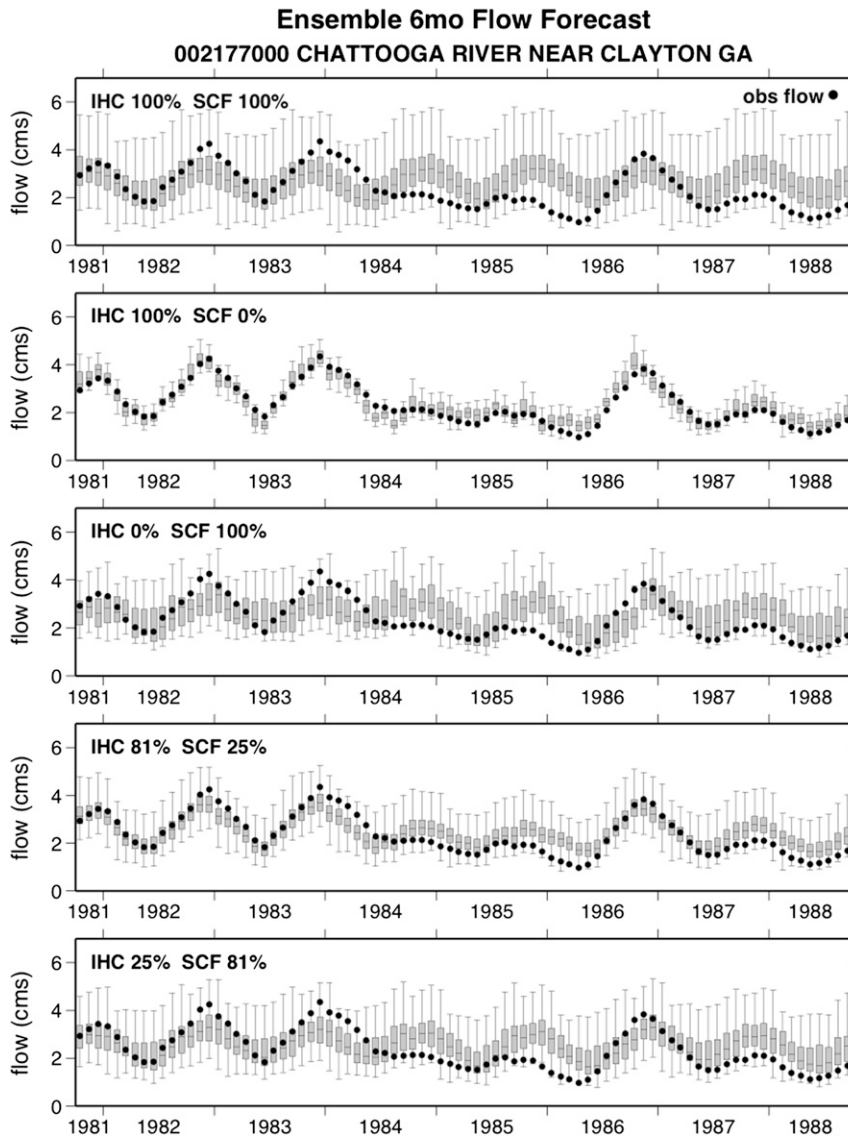


FIG. 4. Time series of 6-month ensemble flow forecasts initialized each month, compared to simulated observations, for varying levels of IHC and SCF climatological uncertainty, for the Chattooga River location. Gray box-and-whisker symbols show each hindcast distribution min, max, and quartiles. The symbols are plotted in the month of initialization; for example, the first data in each year represent a January–June hindcast and observation.

means that changes in the skill of a predictability source have no influence on the skill of the flow forecast, whereas positive elasticities indicate that improving a predictability source will improve the flow forecast skill.

3. Results

The application of the VESPA approach is presented first via a comparison of results for the two focus watersheds, followed by aggregated regional analyses.

a. Predictability variations in two watersheds

On any given day of the year, a watershed experiences season-specific land surface conditions and also expects season-specific forecast period climate conditions; thus, the climatology of a forecast can vary markedly with initialization season and forecast period. Figures 4 and 5 illustrate such variations, showing 6-month hindcasts made with different levels of uncertainty for the Chattooga and Crystal River study locations, respectively. The top plot in each figure shows the combination of

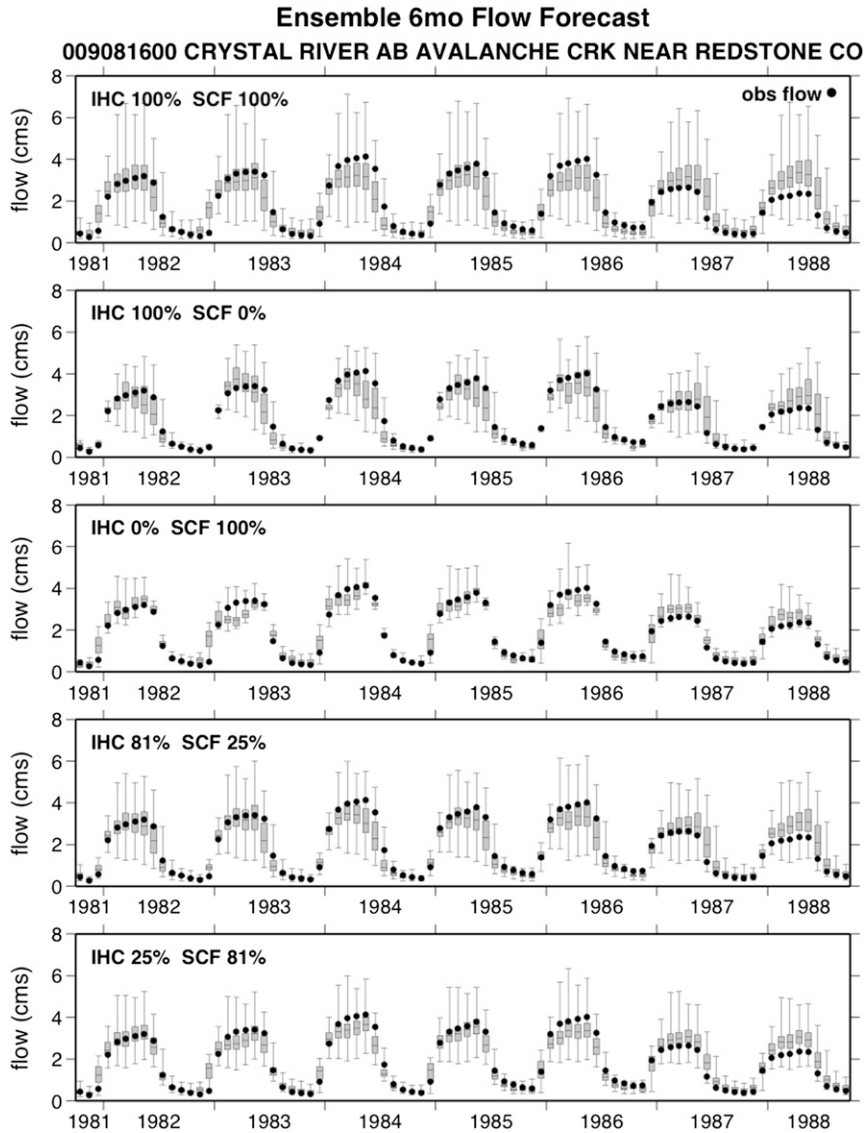


FIG. 5. As in Fig. 4, but for the Crystal River location.

100% climatological uncertainty for both IHCs and SCFs, revealing broad spread and a pattern that repeats each year because of a lack of any information specific to a given year. The next two plots show reverse-ESP and ESP forecasts, and the bottom two plots show two combinations of partial reductions in IHC and SCF uncertainty—a high reduction leading to 25% uncertainty, and a low reduction leading to 81% uncertainty.

For the Chattooga River, the elimination of SCF uncertainty in reverse ESP dramatically reduces the spread and improves the skill of the forecasts, even while the IHCs contain no information, for all forecast initialization months. In contrast, the ESP hindcasts in which IHC uncertainty is eliminated retain ample spread and

show less improvement in the hindcast correspondence with observations. These effects are also evident in the bottom two plots, where high reductions in SCF uncertainty (to 25%) are more effective at improving forecast skill than low ones with high reductions to IHC uncertainty. The Crystal River (Fig. 5) hindcasts have greater seasonal climatological variation in both mean and spread (top plot), with large uncertainty in high-flow months and narrow uncertainty in low-flow months. The reverse-ESP and ESP hindcasts show greater seasonal variations in uncertainty influences than were found for the Chattooga River. Reducing SCF uncertainty to zero most improves the hindcasts (i.e., reduces spread and median error) in fall and winter, whereas reducing IHC uncertainty most improves the hindcasts in spring and

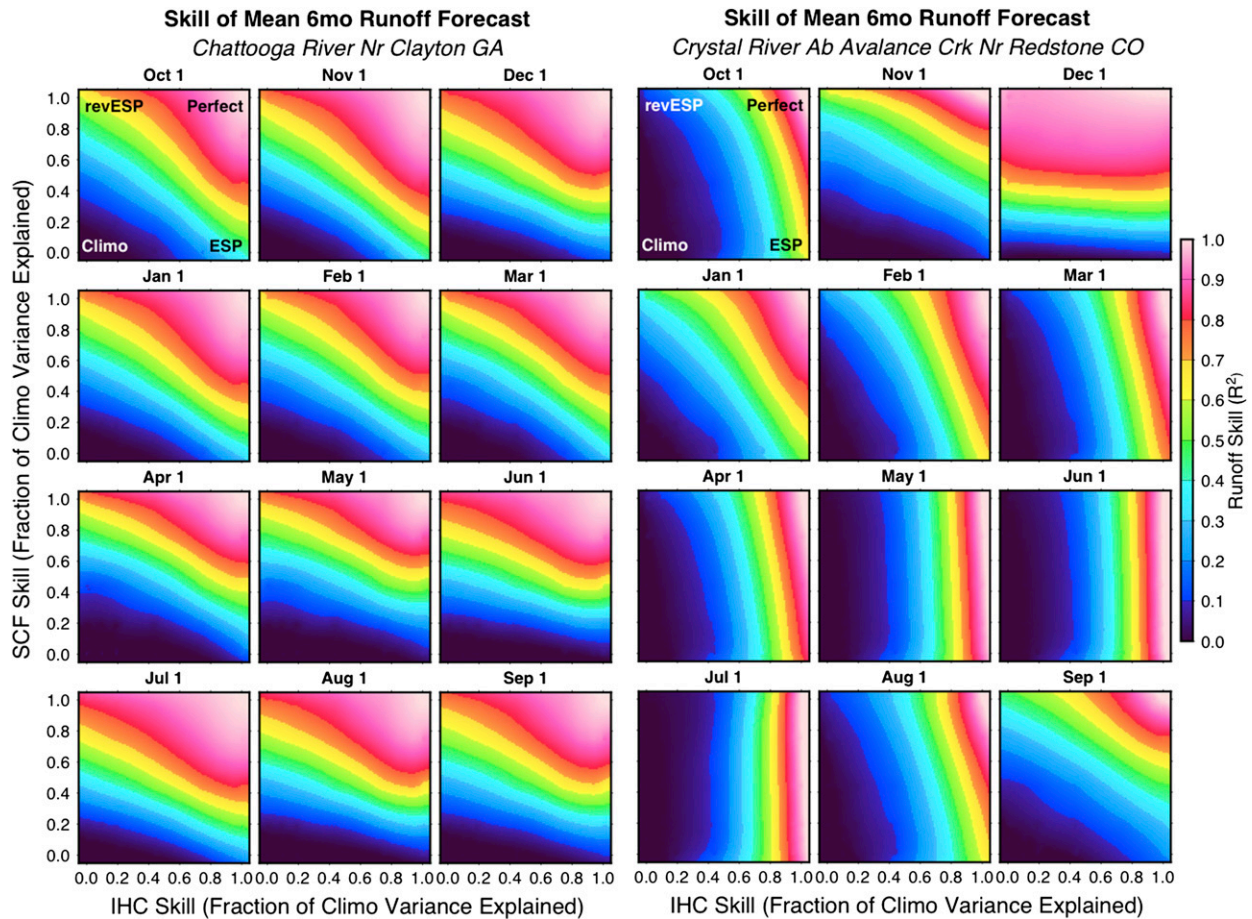


FIG. 6. The 6-month streamflow ensemble median forecast skill R^2 for two locations, for initializations on the first day of each month vs skill in the two predictability sources (fraction of climatological variance r^2). In the top left thumbnail, the points on the surface represented by ESP, climatology (Climo), perfect forecast (Perfect), and reverse ESP (revESP) are indicated.

summer. It is clear that the uncertainty of a December 6-month forecast is almost entirely controlled by SCF skill, whereas IHC skill almost entirely determines June hindcast uncertainty. The final two plots also show these effects, albeit less clearly because of the blending of uncertainty contributions.

Further insight into the performance of streamflow hindcasts illustrated in Figs. 6 and 7 can be gained by plotting the surface describing the skill of the forecast versus the skill in the two predictability sources, IHCs and SCFs, as estimated from the eighty-one 30-yr hindcasts. Figures 6 and 7 show contours of the streamflow forecast skill for the two locations for all 12 initialization dates, for the 6- and 1-month predictands, respectively. Estimates for skill of the ESP, reverse-ESP, perfect, and climatological predictions are located in the bottom right, top left, top right, and bottom left of each thumbnail figure, respectively.

In Fig. 6 (left), the rainfall-driven Chattooga River watershed shows relatively uniform predictability gradients

throughout the year, with slightly greater predictability in fall than in summer. For all forecast initialization months, the skill gradient slope indicates that SCF skill has a stronger influence on the relatively long duration predictand than IHC skill, though IHC skill also affects streamflow forecast skill. The controls on snowmelt-driven Crystal River 6-month forecast skill, in contrast, vary by month according to the hydrologic cycle depicted in Fig. 2. The water year begins with IHC skill as the stronger influence because runoff from October to March is largely driven by base flows, a function of watershed soil moisture. The climate variability of the late fall and winter begins to affect runoff during snowmelt in late March and April. As April and the following months enter the forecast period (e.g., in November and December), the skill gradients in the plots first flatten, indicating that SCFs are critical, and then gradually steepen again as the accumulating snowpack transfers predictability from the climate to the land surface. From March through August, 6-month predictability is

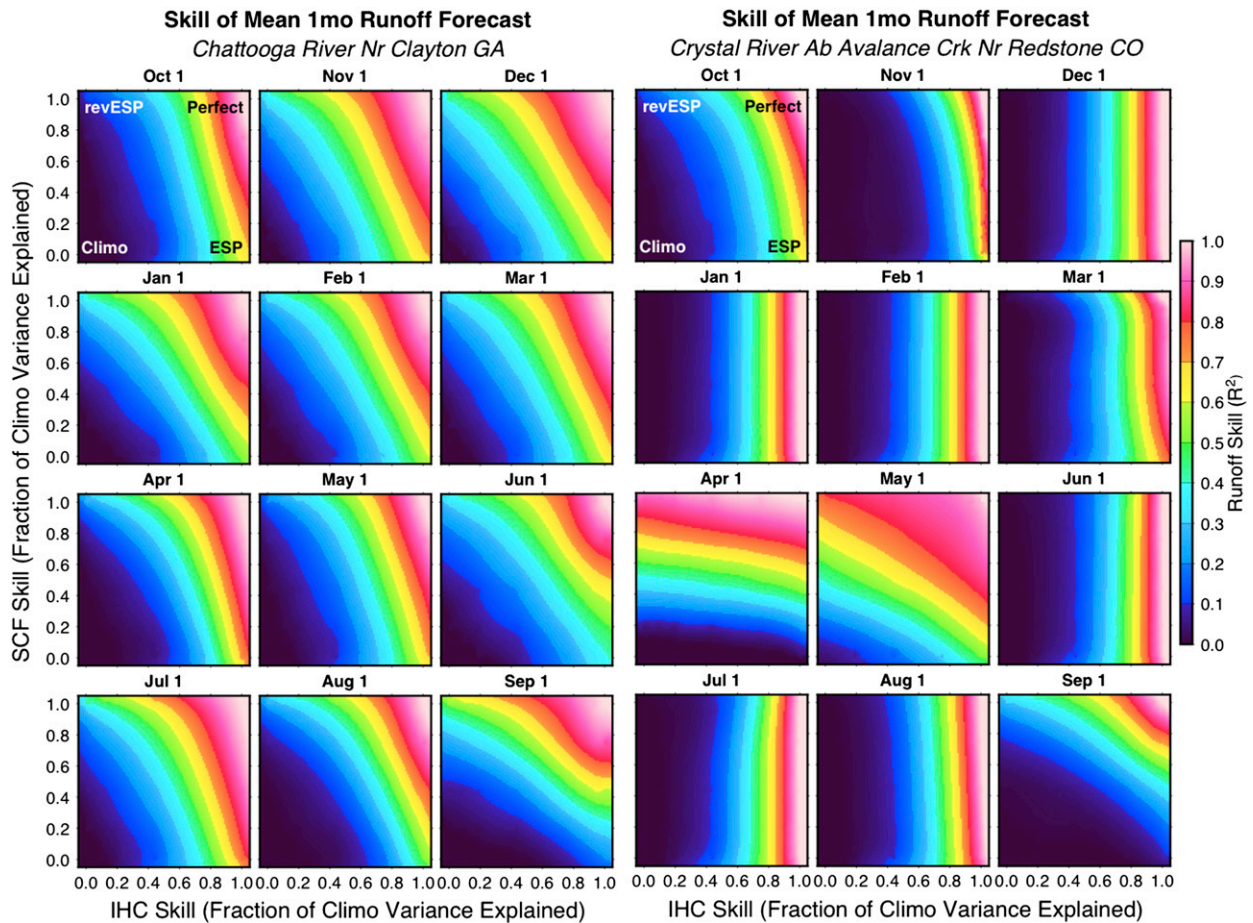


FIG. 7. As in Fig. 6, but for 1-month streamflow forecast skill.

dominated by the IHC skill and uncertainty because the majority of runoff in this period consists of summer snowmelt. In September, SCFs begin to contribute to predictability for runoff variability through fall, given the diminishing influence of snowmelt combined with direct late summer rainfall effects on runoff.

The 1-month streamflow forecast skill gradients (Fig. 7) show similar contrasts between locations, and for the Crystal River site, between forecast months. The Chattooga River gradients are steeper than for the 6-month forecasts because IHCs exert greater control over runoff. In the Crystal River watershed, SCFs are no longer influential in the fall and early winter months because the snowmelt runoff that is linked to SCF variability is not within the forecast period. IHCs dominate 1-month forecast skill in the fall and late summer. April and May show a strong SCF control (the horizontal skill gradients), indicating that climate variability within the months, the heart of the snowmelt period, can drive runoff variation directly. SCFs and IHCs have mixed influence in September and October, when declining SM

variability (from snowmelt) balances climate variability in determining runoff (see Fig. 2).

Another way to use the multiple skill combinations of the variational ensemble assessment to understand contributions to streamflow forecast skill is to examine the effects of adding skill in one predictability source, while fixing skill in the other. Figure 8 shows, for each forecast initialization month, the effects of augmenting skill in each source from a baseline of climatological uncertainty, for increments of 19%, 44%, 75%, and 100% skill (i.e., variance reduction in either IHCs or SCFs), while the other predictability source is held at climatological uncertainty. For the Chattooga River location, adding SCF skill has relatively more benefit than adding IHC skill as the prediction period lengthens from 1 to 6 months (read from top to bottom). The variations from month to month reflect sampling uncertainty ($N = 30$ for each hindcast) given the modest seasonal cycle of the watershed, although the results suggest a relatively stronger IHC (and lower SCF) influence for fall month seasonal forecasts.

Streamflow Forecast Skill Given Single Source Skill

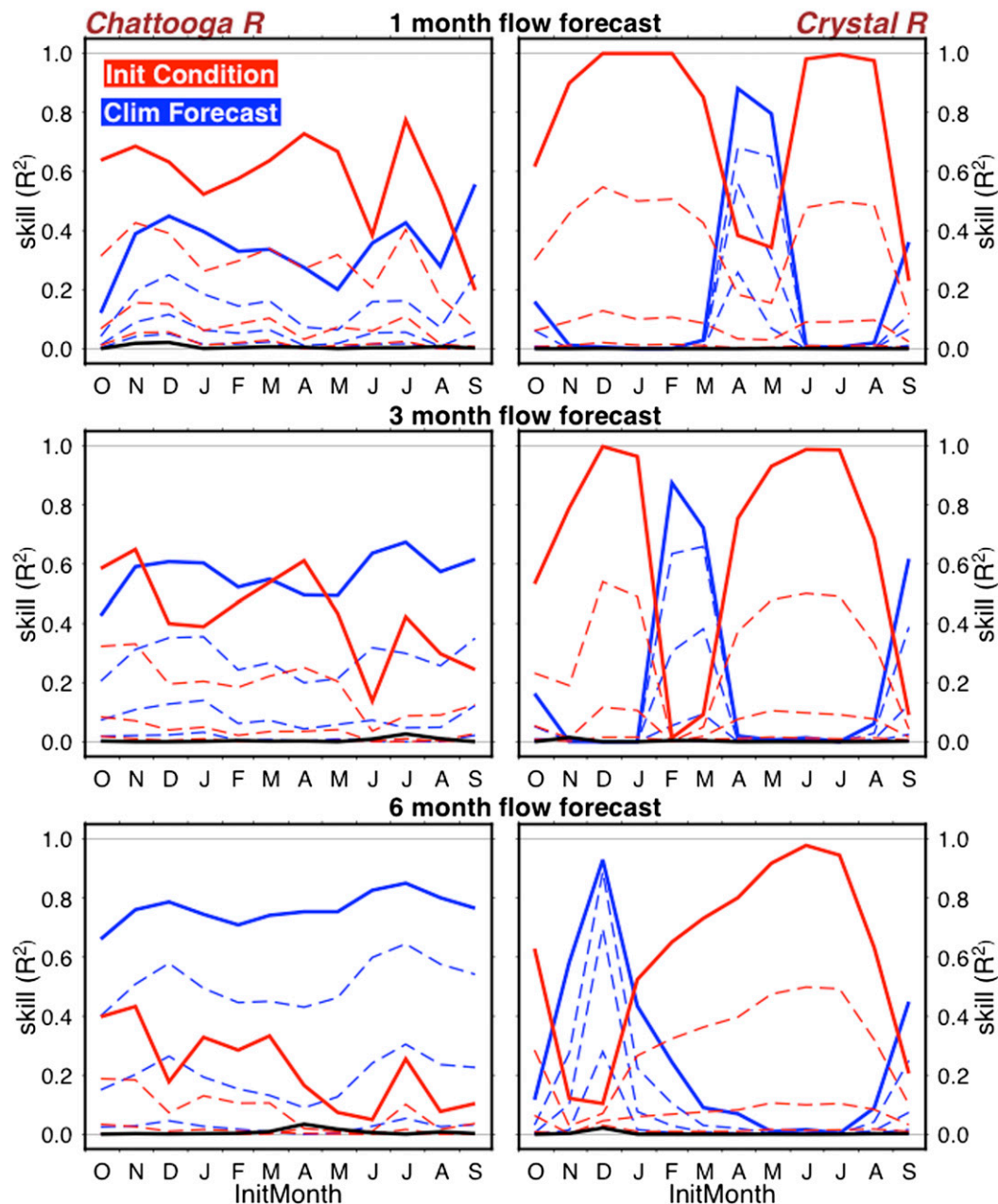


FIG. 8. Streamflow forecast skill sensitivity to skill in each predictability source, initial condition (IHC) and climate forecast (SCF), for different initialization months. The thick black line shows the climatological skill, each of the thin colored dashed lines shows the addition of 19%, 44%, and 75% skill (from lowest to highest), and the thick colored line shows the addition of 100% skill, for each predictability source separately (assuming 0% skill in the alternate source).

For the Crystal River (right), IHCs dominate predictability during most forecast initialization months, and for the 1- and 3-month forecasts, can lead to perfect skill despite a lack of information in the climate forecasts, for snow accumulation months when soil moisture controls runoff. SCFs are important for forecast periods

that include the spring months of high runoff variability (mostly March–May).

b. Regional comparisons

The results for the two focus locations gave an example of the regional difference in the relative

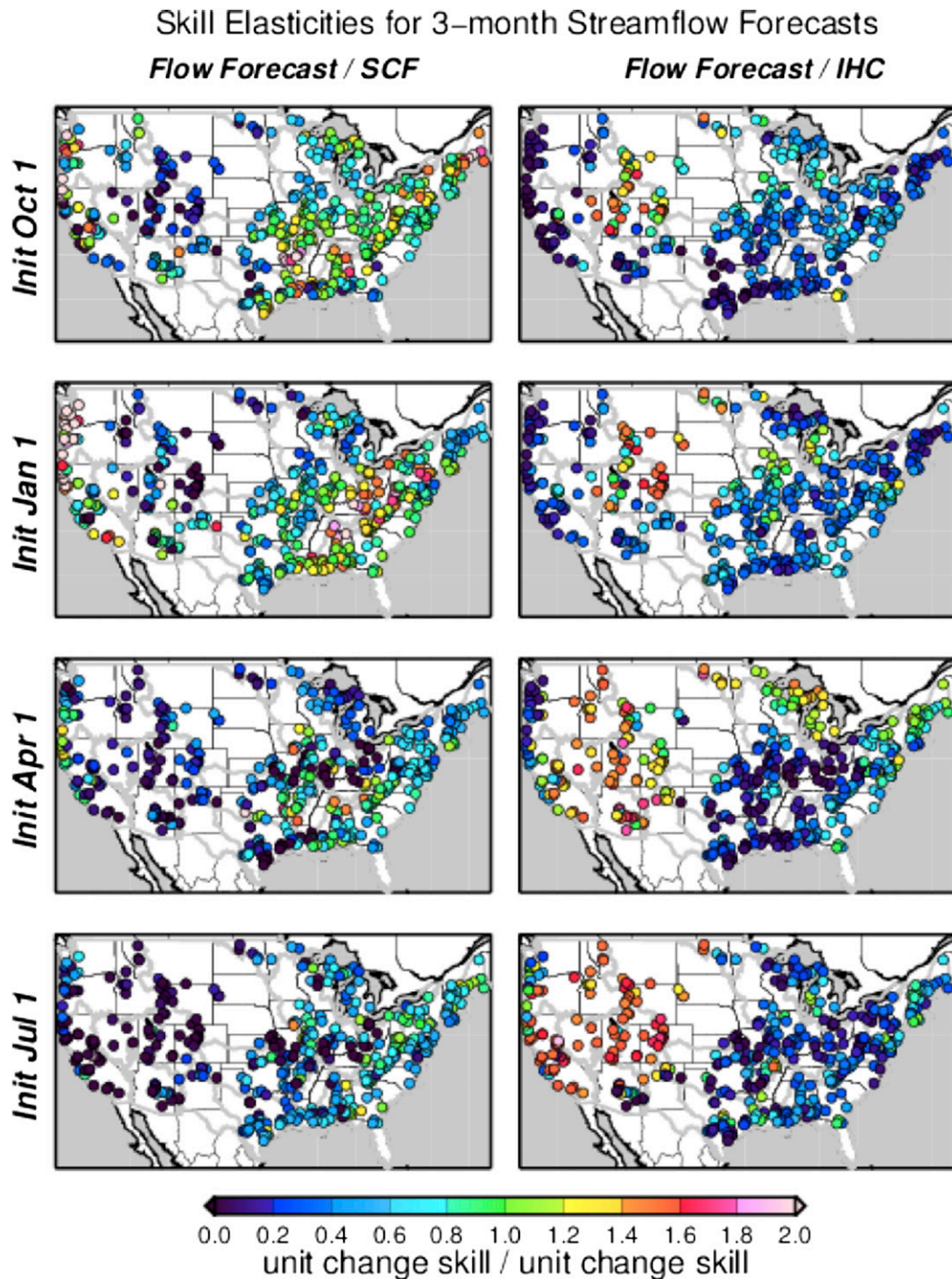


FIG. 9. Skill elasticities for 3-month streamflow forecast initialized on four dates, relative to SCF and IHC skill. Elasticities are calculated around the study's skill combination point that approximates skill levels in an ESP forecast (i.e., assuming minimal climate predictor skill and high initial condition skill; see text for details).

contributions of IHC and SCF skill to flow forecast skill. Forecast skill elasticity—the unit change in flow forecast skill per unit change in a predictor skill—provides a metric for intercomparing skill sensitivities and ultimately

for discussing the value of improvements in a predictability source for a given location, initialization date, and predictand. Figure 9 shows elasticities for all study locations for four forecast dates (1 October, 1 January,

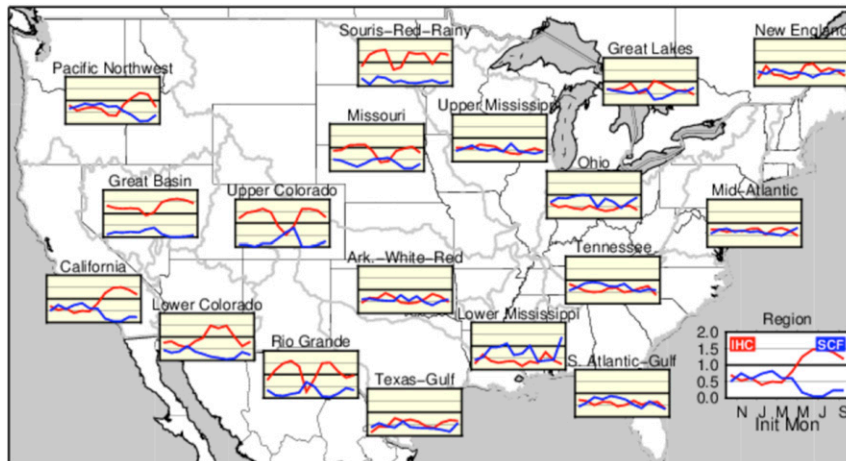


FIG. 10. Regionally averaged streamflow forecast skill elasticities for 1-month predictions on the first day of every month. Elasticities are shown for a baseline capability with a combination of IHC and SCF uncertainty levels of approximately 16% and 95%, respectively.

1 April, and 1 July) and predictand (3-month streamflow) to illustrate the regional variation in skill dependence in more detail. As noted earlier, the elasticities are calculated around the skill combination point that approximates skill levels in an ESP forecast: low IHC uncertainty and near-climatological SCF uncertainty (Wood and Schaake 2008; Werner et al. 2004).

Despite much site-by-site variation, three broad elasticity regions emerge. On the U.S. West Coast, which has a maritime climate regime, flow forecast skill is sensitive to SCF but not IHC skill in most of the wet season (fall, winter, and early spring, with rainfall-driven runoff), and the strongest sensitivity to IHC skill comes to the southern United States in spring and summer. The Intermountain West and northern Great Plains regions, which experience subfreezing winter temperatures and a prominent snow cycle, in contrast, indicate a strong role for IHC skill but not SCF skill during much of the year. Last, the eastern United States south of the Great Lakes, which has a more humid regime with rainfall-driven runoff but variable winter snow in the north, shows stronger influences of SCF than IHC skill throughout the year. This is especially evident in fall and winter, when soil moistures are relatively higher and strengthen the precipitation-runoff response. As one can surmise from the more detailed results of the two study locations, the hydrologic dynamics behind these sensitivities vary by forecast start date and predictand. They also vary (see Figs. 6, 7) depending on the “predictability point” at which the elasticities are calculated.

It is not possible to show similar results for the remaining 32 combinations of forecast date and predictands. We instead summarize regional findings for monthly variations in skill elasticities in Figs. 10–12,

which are for 1-, 3-, and 6-month streamflow forecasts, respectively. The regional summaries in some cases include substantial subregional variation that can be seen in Fig. 9—for example, the Pacific Northwest combines dry and cold interior locations with wetter, more maritime locations—but show cross-regional variations nonetheless. We use the same predictability combination point so that the results describe, in effect, potential skill improvement benefits relative to current practice. As might be expected, the generally higher IHC elasticities in Fig. 10 compared to the generally higher SCF elasticities in Fig. 12 confirm that progressively longer forecast periods translate the streamflow prediction challenge from a mostly IHC-driven problem toward an SCF-driven one. Even for the shorter 1-month predictand in Fig. 10, though, several regions exhibit higher SCF elasticities. The Ohio, Tennessee, and lower Mississippi regions may have faster-responding soils that drain quickly and reduce the runoff persistence from soil moisture and may also experience high-intensity tropically fed storms that overwhelm runoff regulation by land surface storages.

For 3-month streamflow forecasts, eastern regions (such as the mid-Atlantic) show SCF skill rising above IHC skill in importance, while in the western United States, the hydroclimatic patterns of either wet and dry seasons, or cold and warm seasons, lead to seasonal reversals in the dominant influences on streamflow prediction skill. During the dry California and Northwest summers, for instance, SCF skill has little value for improving streamflow predictions, versus high value for forecasts leading into or beginning in the wetter fall and winter months. In the Rio Grande River basin, like the Colorado, the snow cycle leads to three distinct

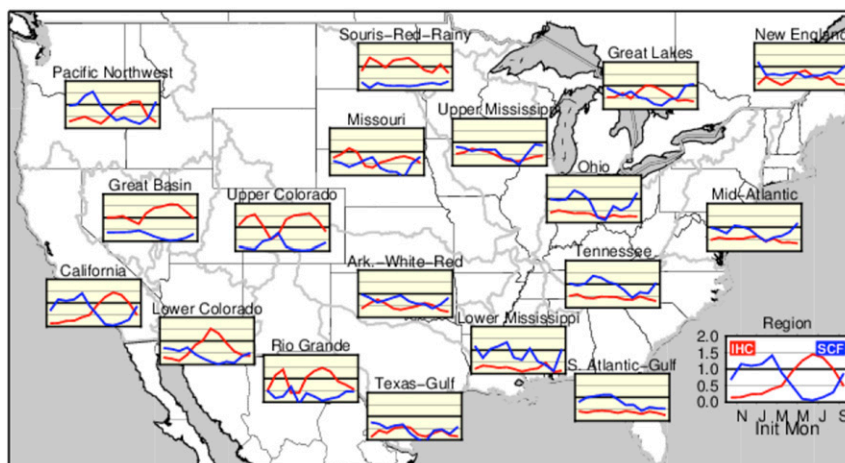


FIG. 11. As in Fig. 10, but for 3-month flow predictions.

predictability regimes: the late summer to winter high IHC predictability as runoff that is driven by soil moisture variability, the late spring snowmelt runoff regime that is driven by a mixture of climate variability, and the late spring to summer regime that is driven by snowmelt and rising soil moisture.

For 6-month streamflow predictions (Fig. 12), SCFs become the dominant streamflow predictability source in the Pacific Northwest, California (until spring), the midwestern regions, and much of the eastern United States. The Great Basin, upper Colorado, and Souris–Red–Rainy regions sustain strong spring through end-of-summer IHC influence, in part because of the snow cycle. A dip in the summer through early fall SCF predictability in the Ohio, upper Mississippi, Tennessee, and both Atlantic regions appears to be related to the depletion of soil moisture during summer and the role of precipitation in soil moisture recharge during fall.

4. Discussion and conclusions

We apply a model-based variational ensemble streamflow prediction assessment (VESPA) approach to investigate the major sources of streamflow predictability across a wide range of different CONUS streamflow locations, forecast dates, and periods. The approach is motivated by interest in quantifying the relative potential for advancing monthly to seasonal streamflow forecasting through investing in improvements to the major sources of predictability—our ability to monitor current watershed conditions and/or to upgrade and apply watershed-scale climate predictions. To this end, we define the metric forecast skill elasticity to estimate the linkage of source skill capability improvements to potential benefits for streamflow prediction

skill. For a specific forecast to support a specific reservoir operation (e.g., a 1 May 2-month runoff forecast to inform reservoir refill), forecast skill elasticities can suggest both where efforts should be devoted and also how much improvement might be expected, given the current state of the practice and realistic potential improvements in predictability sources.

Many of the skill sensitivity results in this paper will be easily recognized by those engaged in the practice of hydrologic prediction or in research into land surface predictability. For instance, the dominant contribution of IHCs (particularly SWE) to long lead predictability in the western United States is a foundation of operational seasonal forecasting, and a reason why climate forecasting, with both weaker skill and flow forecast skill elasticity in many months, has been relatively undeveloped. The link between dry or subfreezing climate periods, which restrict variation in hydrologic fluxes, and higher streamflow predictability corroborates earlier findings that linked wet climate periods with lower predictability.

The quantification of forecast skill elasticities nonetheless adds a useful dimension to these insights. A key finding is that in many western U.S. locations, the elasticities are greater than one for both IHCs and SCFs. For SCFs, this result argues that the low levels of skill for precipitation forecasting (e.g., <10% improvement over climatology) can lead to higher, potentially usable levels of skill for seasonal streamflow prediction (e.g., >10% improvement). This result contradicts the perception of low SCF utility that has been common in operational flow forecasting, inhibiting the development of SCF-based flow predictions even in regions of relatively strong climate predictability and potential management value (such as the hydropower-rich Northwest). The

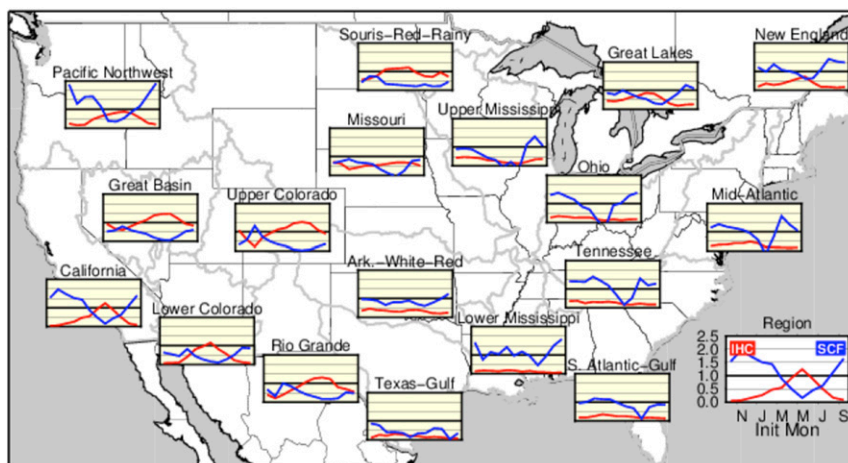


FIG. 12. As in Fig. 10, but for 6-month flow predictions.

reasons for the amplification of skill between IHCs or SCFs and streamflow warrant further investigation but may relate to the time-integrated formation of snow or soil moisture anomalies over the winter and spring, as well as to subunity runoff ratios and nonlinearities or threshold behavior in the forcing–runoff response.

The variation of elasticities by month, region, and predictand has an important implication for climate prediction evaluation and services from the water-sector perspective. The best climate forecast across a range of these dimensions (e.g., for all 3-month seasons and regions) is unlikely to be the best climate forecast for any particular season, location, prediction, and decision use. In the age of multiple climate forecast options (such as the North American Multi-Model Ensemble), users would likely benefit from determining their selection or weighting of climate forecast sources with specific predictands in mind. We note, however, that the forecast skill elasticities alone, lacking knowledge of the hydroclimatological variability and water management practice in a region, can provide a misleading depiction of the real value of the component forecast capabilities. For instance, the highest skill elasticities associated with low-variability summer base flows or subfreezing regimes arise because of high natural hydrologic persistence and predictability, which obviates the need for a streamflow forecast.

The major goal of this paper is to demonstrate VESPA as a strategy for understanding forecast skill dependencies, yet we make a number of choices that could be handled differently in other studies. We use a single skill metric (i.e., R^2), relatively simple calibrated models (SNOW-17 and SAC-SMA), a linear scaling approach to blending IHCs and SCFs, a limited number of predictands (3) and forecast dates (12), and a specific

set of skill combinations (81) that may not resolve skill gradients optimally for many purposes. We also note the sensitivities and skills presented are dependent on the quality of the model simulations, which, though calibrated, are not perfect. Given the number of watersheds assessed here, it lies beyond the scope of this paper to explore these implications, but further research in this area would be of interest.

A second objective of the paper is to provide a broad-based analysis of forecast skill sensitivities across CONUS, sufficient to provide insights into hydroclimatically driven variations in skill at the watershed level. The summary-level perspectives and insights from the two focus watersheds illustrate the detailed assessment that has been created for all 424 watersheds. This assessment supports a key finding from prior work: IHCs are so critical to many prediction objectives that continuing the substantial investment toward improving their estimation is surely warranted. Such IHC-related skill upgrades can come through improvements to the reanalysis and real-time monitoring of meteorology and hydrology (e.g., of snow and streamflow), to watershed models (via better parameter estimation and forcing preparation), and/or through data assimilation techniques that reduce IHC error by leveraging unexploited observations (Liu et al. 2012). Second, we find that for some predictands, the application of SCFs provides the only effective avenue to add new skill and thus also deserves investigation from forecasting enterprises.

As discussed earlier in the paper, the elasticity results of this paper, like the predictability findings of prior ESP/reverse-ESP analyses, are likely to depend in complex ways on the modeling and other choices made in the analysis. A clear direction for future work in this area is to move beyond perfect-model experiments

where possible. The incorporation of observations (of streamflow or other variables) will be essential to begin to understand and quantify the role of modeling uncertainty in model-based predictability studies. For now, this paper describes only part of a strategy toward full decomposition of streamflow prediction skill and uncertainty, the end goal of which is to support the design of research and development toward improving streamflow forecasting and ultimately, water management.

Acknowledgments. The research and analysis reported herein were supported by the Bureau of Reclamation and the U.S. Army Corps of Engineers. We would like to acknowledge high-performance computing support from Yellowstone (ark:/85065/d7wd3xhc) provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation. We also thank two anonymous reviewers for critical and constructive comments that have substantially enhanced the clarity and readability of the paper.

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. J.*, **51**, 642–654, doi:10.1623/hysj.51.4.642.
- Day, G. N., 1985: Extended streamflow forecasting using NWSRFS. *J. Water Resour. Plann. Manage.*, **111**, 157–170, doi:10.1061/(ASCE)0733-9496(1985)111:2(157).
- Demargne, J., and Coauthors, 2014: The Science of NOAA's Operational Hydrologic Ensemble Forecast Service. *Bull. Amer. Meteor. Soc.*, **95**, 79–98, doi:10.1175/BAMS-D-12-00081.1.
- Duan, Q., S. Sorooshian, and V. K. Gupta, 1992: Effective and efficient global optimization for conceptual rainfall–runoff models. *Water Resour. Res.*, **28**, 1015–1031, doi:10.1029/91WR02985.
- , V. K. Gupta, and S. Sorooshian, 1993: A shuffled complex evolution approach for effective and efficient optimization. *J. Optim. Theory Appl.*, **76**, 501–521, doi:10.1007/BF00939380.
- Falcone, J. A., 2011: GAGES-II: Geospatial Attributes of Gages for Evaluating Streamflow. USGS, accessed XX XXXXXX XXXX. [Available online at http://water.usgs.gov/GIS/metadata/usgswrd/XML/gagesII_Sept2011.xml.]
- Garen, D. C., 1992: Improved techniques in regression-based streamflow volume forecasting. *J. Water Resour. Plann. Manage.*, **118**, 654–670, doi:10.1061/(ASCE)0733-9496(1992)118:6(654).
- Gustafsson, N., E. Kallen, and S. Thorsteinsson, 1998: Sensitivity of forecast errors to initial and lateral boundary conditions. *Tellus*, **50A**, 167–185, doi:10.1034/j.1600-0870.1998.t01-1-00002.x.
- Harrison, B., and R. Bales, 2015: Skill assessment of water supply outlooks in the Colorado River basin. *Hydrology*, **2**, 112–131, doi:10.3390/hydrology2030112.
- Hydrologic Research Laboratory Staff, 1972: National Weather Service River Forecast System forecast procedures. NOAA Tech. Memo. NWS HYDRO-14, 264 pp.
- Kirtman, B. P., and Coauthors, 2014: The North American Multimodel Ensemble: Phase-1 seasonal-to-interannual prediction; Phase-2 toward developing intraseasonal prediction. *Bull. Amer. Meteor. Soc.*, **95**, 585–601, doi:10.1175/BAMS-D-12-00050.1.
- Koster, R. D., S. P. P. Mahanama, B. Livneh, D. P. Lettenmaier, and R. H. Reichle, 2010: Skill in streamflow forecasts derived from large-scale estimates of soil moisture and snow. *Nat. Geosci.*, **3**, 613–616, doi:10.1038/ngeo944.
- Li, H., L. Luo, E. F. Wood, and J. Schaake, 2009: The role of initial conditions and forcing uncertainties in seasonal hydrologic forecasting. *J. Geophys. Res.*, **114**, D04114, doi:10.1029/2008JD010969.
- Lins, H. F., 2012: USGS Hydro-Climatic Data Network 2009 (HCDN-2009). USGS Fact Sheet 2012-3047, 4 pp. [Available online at <http://pubs.usgs.gov/fs/2012/3047/>.]
- Liu, Y., and Coauthors, 2012: Advancing data assimilation in operational hydrologic forecasting: Progresses, challenges, and emerging opportunities. *Hydrol. Earth Syst. Sci.*, **16**, 3863–3887, doi:10.5194/hess-16-3863-2012.
- Mahanama, S. P. P., and R. D. Koster, 2003: Intercomparison of soil moisture memory in two land surface models. *J. Hydrometeorol.*, **4**, 1134–1146, doi:10.1175/1525-7541(2003)004<1134:IOSMMI>2.0.CO;2.
- , —, 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.*, **31**, 1333–1343, doi:10.1016/j.advwatres.2008.06.004.
- , B. Livneh, R. D. Koster, D. P. Lettenmaier, and R. H. Reichle, 2012: Soil moisture, snow, and seasonal streamflow forecasts in the United States. *J. Hydrometeorol.*, **13**, 189–203, doi:10.1175/JHM-D-11-046.1.
- Maurer, E. P., and D. P. Lettenmaier, 2003: Predictability of seasonal runoff in the Mississippi River basin. *J. Geophys. Res.*, **108**, 8607, doi:10.1029/2002JD002555.
- Mo, K. C., and D. P. Lettenmaier, 2014: Hydrologic prediction over the conterminous United States using the National Multi-Model Ensemble. *J. Hydrometeorol.*, **15**, 1457–1472, doi:10.1175/JHM-D-13-0197.1.
- Newman, A. J., and Coauthors, 2015: Development of a large-sample watershed-scale hydrometeorological data set for the contiguous USA: Data set characteristics and assessment of regional variability in hydrologic model performance. *Hydrol. Earth Syst. Sci.*, **19**, 209–223, doi:10.5194/hess-19-209-2015.
- Pagano, T., D. Garen, and S. Sorooshian, 2004: Evaluation of official western U.S. seasonal water supply outlooks, 1922–2002. *J. Hydrometeorol.*, **5**, 896–909, doi:10.1175/1525-7541(2004)005<0896:EOOWUS>2.0.CO;2.
- , A. W. Wood, K. Werner, and R. Tama-Sweet, 2014: Western U.S. water supply forecasting: A tradition evolves. *Eos, Trans. Amer. Geophys. Union*, **95**, 28–29, doi:10.1002/2014EO030007.
- Paiva, R. C. D., W. Collischonn, M. P. Bonnet, and L. G. G. de Gonçalves, 2012: On the sources of hydrological prediction uncertainty in the Amazon. *Hydrol. Earth Syst. Sci.*, **16**, 3127–3137, doi:10.5194/hess-16-3127-2012.
- Priestley, C. H. B., and R. J. Taylor, 1972: On the assessment of surface heat flux and evaporation using large-scale parameters. *Mon. Wea. Rev.*, **100**, 81–82, doi:10.1175/1520-0493(1972)100<0081:OTAOSH>2.3.CO;2.
- Raff, D., L. Brekke, K. V. Werner, A. Wood, and K. White, 2013: Short-term water management decisions: User needs for improved climate, weather, and hydrologic information. Tech. Rep. CWTS-2013-1, 233 pp. [Available online at http://www.cawwg.us/docs/Short-Term_Water_Management_Decisions_Final_3_Jan_2013.pdf.]

- Shukla, S., and D. P. Lettenmaier, 2011: Seasonal hydrologic prediction in the United States: Understanding the role of initial hydrologic conditions and seasonal climate forecast skill. *Hydrol. Earth Syst. Sci.*, **15**, 3529–3538, doi:10.5194/hess-15-3529-2011.
- , J. Sheffield, E. F. Wood, and D. P. Lettenmaier, 2013: On the sources of global land surface hydrologic predictability. *Hydrol. Earth Syst. Sci.*, **17**, 2781–2796, doi:10.5194/hess-17-2781-2013.
- Singla, S., J.-P. Céron, E. Martin, F. Regimbeau, M. Déqué, F. Habets, and J.-P. Vidal, 2012: Predictability of soil moisture and river flows over France for the spring season. *Hydrol. Earth Syst. Sci.*, **16**, 201–216, doi:10.5194/hess-16-201-2012.
- Staudinger, M., and J. Seibert, 2014: Predictability of low flow—An assessment with simulation experiments. *J. Hydrol.*, **519**, 1383–1393, doi:10.1016/j.jhydrol.2014.08.061.
- Thober, S., R. Kumar, J. Sheffield, J. Mai, D. Schaefer, and L. Samaniego, 2015: Seasonal soil moisture drought prediction over Europe using the North American Multi-Model Ensemble (NMME). *J. Hydrometeorol.*, **16**, 2329–2344, doi:10.1175/JHM-D-15-0053.1.
- Thornton, P. E., M. M. Thornton, B. W. Mayer, N. Wilhelm, Y. Wei, and R. B. Cook, 2012: Daymet: Daily surface weather on a 1 km grid for North America, 1980–2012. ORNL DAAC, accessed 15 July 2013. [Available online at <https://daymet.ornl.gov/>.]
- van Dijk, A. I. J. M., J. L. Peña-Arancibia, E. F. Wood, J. Sheffield, and H. E. Beck, 2013: Global analysis of seasonal streamflow predictability using an ensemble prediction system and observations from 6192 small catchments worldwide. *Water Resour. Res.*, **49**, 2729–2746, doi:10.1002/wrcr.20251.
- Wang, Q. J., A. Schepen, and D. E. Robertson, 2012: Merging seasonal rainfall forecasts from multiple statistical models through Bayesian model averaging. *J. Climate*, **25**, 5524–5537, doi:10.1175/JCLI-D-11-00386.1.
- Welles, E., S. Sorooshian, G. Carter, and B. Olsen, 2007: Hydrologic verification: A call for action and collaboration. *Bull. Amer. Meteor. Soc.*, **88**, 503–511, doi:10.1175/BAMS-88-4-503.
- Werner, K., D. Brandon, M. Clark, and S. Gangopadhyay, 2004: Climate index weighting schemes for NWS ESP-based seasonal volume forecasts. *J. Hydrometeorol.*, **5**, 1076–1090, doi:10.1175/JHM-381.1.
- Wood, A. W., and D. P. Lettenmaier, 2003: Comparing hydrologic forecast uncertainty due to initial condition error versus climate forecast error. *Proc. EGS-AGU-EUG Joint Assembly*, Vol. 5, Nice, France, EGS-AGU-EUG, 1374.
- , and —, 2008: An ensemble approach for attribution of hydrologic prediction uncertainty. *Geophys. Res. Lett.*, **35**, L14401, doi:10.1029/2008GL034648.
- , and J. C. Schaake, 2008: Correcting errors in streamflow forecast ensemble mean and spread. *J. Hydrometeorol.*, **9**, 132–148, doi:10.1175/2007JHM862.1.
- , 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.
- Yosief, N. C., H. Winsemius, A. Weerts, R. van Beek, and M. F. P. Bierkens, 2013: Skill of a global seasonal streamflow forecasting system, relative roles of initial conditions and meteorological forcing. *Water Resour. Res.*, **49**, 4687–4699, doi:10.1002/wrcr.20350.