July, 2005

ANALYSIS AND USE OF INPUT UNCERTAINTIES IN WATERSHED MODELING

HWANG, Yeonsang

University of Colorado, 1333 Grandview Ave., Boulder, CO80309

CLARK, Martyn P.

CIRES, Center for Science and Technology Policy Research, Boulder, CO80309

RAJAGOPALAN, Balaji

University of Colorado, Dept. of Civil, Environmental and Architectural Engg.,

Campus Box 428, Boulder, CO80309

Submitted to,
Water Resources Research.

Abstract

Among the other sources of uncertainty in hydrologic modeling, uncertainty in model inputs were quantified. The authors demonstrated how information on input uncertainty can be used in operational streamflow forecasting system. The impact of input interpolation error was tested by generating ensemble streamflows. A distributed hydrologic model (PRMS) was used in two hydrologically different basins (Animas basin at Durango, Colorado and Alapaha basin at Statenville, Georgia) to generate streamflow.

Different impacts of input uncertainty in the two basins are well shown in the generated ensemble flows - more varied ensemble flow in Alapaha basin and tighter ensemble flow range in Animas basin. The wider spread of ensemble flow in the Alapaha basin was caused by poor estimation performance of interpolation model in this rainfall-dominated basin. This ensemble streamflow generation framework was also applied for example forecasts that could improve traditional ESP method.

1. Introduction

The impact of different precipitation interpolation schemes on hydrologic modeling shown in the previous works [Hwang et al., 2005a; Hwang et al., 2005b] shows quite interesting features. First, interpolation methods based on long-term mean precipitation gradients do not capture the local character of precipitation patterns, and produce less accurate streamflow simulations, especially in rainfall-dominated basins. Second, the model calibration process can identify optimal parameter sets that compensate for differences in the statistical characteristics of various input data streams.

The result from the previous works [Hwang et al., 2005a, Hwang et al., 2005b] shows that the simulated streamflow of hydrologic model is sensitive to the different precipitation schemes. Now, we develop methods to quantify errors in the interpolation model itself, and demonstrate how information on input uncertainty can be used in operational streamflow forecasting systems.

2. Input uncertainties in watershed modeling

In watershed modeling, there are three sources of uncertainty; uncertainty in model structure, uncertainty in model inputs, and uncertainty in model parameters.

Errors on the measurement of precipitation are either instrumental or human-related, and some of them can be successfully corrected using data control techniques [Reek et al, 1992]. However, streamflow simulations are unlikely to improve until precipitation inputs improve. If better measurements are not available and hard to decompose the errors from various sources, we can at least examine the impacts of precipitation uncertainty and make the best use of exiting precipitation data [Michaud and Sorooshian, 1994a].

Input uncertainties are one of the major sources of streamflow simulation uncertainty. Early studies focused on issues of precipitation gauging density and the selection of proper gauging station. In the work of Dawdy and Bergmann [1969], annual streamflow volumes and peak volumes are shown to be sensitive to the different areal averaging technique and selection of gauging stations in a irregular measurement network. Krajewski et al. [1991] used a Monte Carlo approach to stochastically generate precipitation and showed that a distributed model is better to estimate both total streamflow and peak stream flow. Michaud and Sorooshian [1994b] showed that a decrease in temporal resolution was found to have significant impact on streamflow simulation accuracy. Shah et al., 1996 showed that streamflow is more sensitive to different precipitation aggregation in situations when soils are dry. With the deployment of radar precipitation measurement network, many authors have studied the impact of pixel aggregation of the radar measurement. In the work of Pessoa et al. [1992], both the spatial and temporal aggregation impacts on the hydrograph were found to be more important for convective storms. This finding was confirmed in the work of Winchell et al. [1998], and infiltration-excess runoff was found to be more sensitive than saturation-excess runoff to highly variable precipitation input. Radar input uncertainties in distributed modeling have been one of the main focus of previous investigators, and extensive Monte Carlo approaches have been utilized to analyze the interaction between input and parameter uncertainties [Carpenter et al., 2001; Blazkova and Beven, 2002; Carpenter and Georgaakos, 2004]. Although recent works are mostly focused on the uncertainty issues with high-resolution precipitation product coupled with highly distributed runoff models, it is worth looking at more insight on the input uncertainty issues on gauge-based precipitation estimates. Practically, watershed modeling and forecasting still largely rely on the gauged precipitation and streamflow records.

Uncertainties included in precipitation interpolation process can be examined and utilized by generating ensembles that include possible input errors. In the previous works [Hwang et al., 2005a; Hwang et al., 2005b], a proper precipitation interpolation scheme (two-step process) was suggested based on the comparison of a suite of different interpolation schemes, and the impact of different interpolation schemes on the simulated streamflows were examined using a distributed watershed model (PRMS). It was shown that the different interpolation schemes lead to streamflow differences in both high and low flow seasons. It was also suggested that the analysis on the interaction between the input and parameter uncertainties could be used to minimize overall simulation uncertainties.

The remaining question on the uncertainty issue now is focused on the possible error in the estimation model. This analysis implicitly includes the errors in measurement because there is no reason to believe that the rain gauge measurements are true. This analysis will also utilize input uncertainty for the assessment of streamflow uncertainty and estimation of daily time scale streamflow forecast.

3. Experimental setup

3.1 Basin

The impact of spatial interpolation methods on rainfall-runoff models (PRMS) was tested in Animas River at Durango, Colorado (Animas) and Alapaha River at Statenville, Georgia (Alapaha). Details of the basin information are given in Table 1. The two basins (Figure 1) are climatologically and hydrologically different. The streamflow of Animas basin is snowmelt driven with occasional rain-on-snow events during winter, and has large relief with elevation ranging from 680 m to 3700 m. The Alapaha basin is a lower elevation watershed, and the streamflow is mainly dominated by rainfall events. Among the various characteristics of watersheds,

topography has significant impact on the spatial variability of climate variables, especially in the Animas basin.

Mean monthly precipitation and observed stream flows from all stations in the two basins are shown in Figure 2. In Animas basin, the winter season is wetter, and the summer season is drier. The precipitation trend is quite different in Alapaha basin (right column of Figure 2). July is the wettest month but the yearly dry-wet season separation is not as clear as that of Animas basin. Further investigation on monthly flow of Animas basin (bottom left graph of Figure 2) indicates that high flows of the spring season are snowmelt runoff because the peak-flow month (June) is the driest month in this basin. Although the precipitation in Alapaha basin is quite varied through out the year and even reveals higher precipitation amount during the summer season (June, July, and August), summer streamflow volume is low due to the higher evapotranspiration rate during the summer season.

3.2 Data

The data set used in this study is compiled from the Cooperative (COOP) network managed by the National Climate Data Center (NCDC), and the SNOwpack TELemetry (SNOTEL) network managed by the Natural Resources Conservation Service (NRCS). Both NCDC (COOP) and NRCS (SNOTEL) perform some degree of data quality screening. In order to guard against erroneous values in the data set, however, additional data quality control was performed in the work of Clark and Slater [2005]. This followed the procedures of Reek et al. [1992], Kunkel et al. [1998], and Serreze et al. [1998], and Serreze et al. [1999], which include checks for a) extreme values; b) internal consistency among variables (e.g., maximum temperature less than minimum temperature); c) constant temperature (e.g., 5 or more days with the same temperature are suspect); d) excessive diurnal temperature

range; e) invalid relations between precipitation, snowfall, and snow depth; and f) unusual step changes or spikes in temperature time series. Areal consistency checks were also implemented following the method described in the work of Gandin [1988]. By comparing each data value against the interpolated value from the surrounding stations, the data values lie outside the tolerance are flagged.

Ensemble precipitation members are generated through the method developed by Clark and Slater [2005, accepted]. The ensemble generation process consists of three main steps ; (1) estimate the probability of precipitation occurrence (POP) at each grid cell (2km by 2km in this study); (2) estimate precipitation amounts (PCP) at each grid cell, conditional on precipitation occurrence; and (3) compute the cross-validated error (E) in station PCP estimates, and interpolate this error to each grid cell. Together POP, PCP, and E define the cumulative probability distribution function (c.d.f.) of precipitation at each grid cell – ensemble grids of precipitation are synthesized by sampling from the gridded c.d.f.s.

POP and PCP estimation on the grid cell are calculated using locally weighted regression scheme [Loader, 1992; Rajagopalan and Lall, 1998]. In this approach, station locations (e.g., latitude, longitude, elevation) are used as explanatory variables to predict spatial variability of precipitation.

Grids of correlated random numbers that preserve the spatial and temporal structure of the observed precipitation are used to generate precipitation ensembles. Details on the multivariate correlated random number generation technique used in this process are described in the work of Clark and Slater [2005].

In order to examine the sole impact of precipitation error on the watersheds, maximum and minimum temperature interpolation was done by multiple linear regression scheme, and the errors of these temperature interpolation were not considered.

The generated precipitation ensemble sets have 2 km by 2 km grid resolution. However, the center of each HRU is not located on the center of a grid cell. Since PRMS uses the averaged precipitation value on each HRU, it would be more reasonable to consider all the precipitation values of all the grids cells located on each HRU. For this purpose, mean of the precipitation values from the 4 nearest grid cells are used for the precipitation value of each HRU. Figure 2 shows the topography and sample maps of generated daily precipitation ensembles (#1, #2, and #3) in 2km by 2km grid in Alapaha basin. This figure is showing the precipitation event that contributed the highest streamflow events in the time series (1975~1995) used in this study. February 5, 1986 was the highest streamflow days in Alapaha. These generated precipitation ensembles are used to get precipitation values on each HRU in PRMS model. Mean precipitation maps (not shown) show that ensemble precipitation reproduces the elevation effect properly in the Animas basin (higher mean precipitation amount in higher part of the basin). The Alapaha basin is relatively flat and elevation does not exert a strong control on spatial precipitation patterns.

Different climatological characteristics of the two basins are expected to impact generated ensemble precipitation. Generated ensemble sets can represent the possible input uncertainty (e.g., larger ensemble spread upon storm-dominated climatology) that is often not being reproduced by the mean estimation of the interpolation models. The comparison of mean and variance statistics between observed precipitation and generated ensembles is shown in Figure 4. On each station or grid cell, mean and variance in January and July were calculated through the entire time series. The length of the boxes indicates the inter-quartile range of the measures at each location in the basin and the whiskers show the 5^{th} and 95^{th} percentile range; the horizontal line in the box is the median of the estimates. A

larger box length indicates increased variability in the estimates. Mean and variance of ensemble sets show general agreement with the observations. However, July statistics of the ensemble sets number 1 and 36 in Alapaha basin are more varied than that of observed statistics shown in the left side of the set of three boxplots. This is because of the storm-dominated climate during the summer season in this basin. This implies that larger input uncertainties should be considered in watershed modeling. This result is well agreeable with the generally lower performance of most interpolation models shown previously in the work of Hwang et al. [2005a]. Note that deterministic approaches to precipitation estimation result in lower variability than what occurs at station locations.

The spatial correlation in precipitation fields is one of the critical features that needs to be properly preserved. In order to check this, the lagged distance correlation of the observed and generated 100 ensembles is shown in Figure 5. In this plot, 10 distance categories are defined first. Then, the correlation between the entire station pairs are calculated and plotted against the defined distance categories, and only the median values are plotted in the figure as a single line for each precipitation set. The thick gray lines represent generated 100 ensemble members, and a solid black line represents the observed precipitation. Left columns are for Alapaha basin (graphs (a) and (c)), and right columns are for Animas basin (graphs (b), and (d)). As shown in the graphs, the median line of the observation is located inside the ensemble plots, and the ensemble members are well spread around the observation except for July in Animas basin. Note that the deterministic methods for precipitation estimation presented in previous works [Hwang et al., 2005a, Hwang et al., 2005b] over-estimated the spatial correlations.

3.3 Watershed Model

Runoff Modeling System (PRMS) is a physically based distributed hydrologic model based on modular-design modeling concept developed and used by USGS [Leavesley et al., 1983; Leavesley et al., 1996]. The response of the watershed upon input forcing is conceptually reproduced in the different modules to simulate streamflow at the basin outlet. The distributed modeling capability in daily mode PRMS simulation is implemented using Hydrologic Response Unit (HRU) concept while storm mode (e.g., at hourly time steps) simulation is performed using plane and channel segments that allow surface runoff routing. In the daily simulations that are used in this study, the hydrologic response on each HRU is calculated separately, and then combined to produce basin wide response. Each HRU is assumed to be homogeneous in hydrologic and geophysical characteristics (aspect, slope, soil type, and precipitation distribution).

The watershed system is conceptually simplified for modeling purpose [Leavesley et al., 1983]. The precipitation exceeding the interception storage from the vegetation canopy is the net input to the watershed. Accumulated snow pack melts based on the maximum, minimum temperature, and solar radiation of the basin. Once the impervious storage capacity is filled with precipitation or melted snow pack, surface runoff occurs. Basically, the amount of surface runoff is controlled by antecedent soil moisture and given precipitation amount. The water body that doesn't contribute to the surface runoff and evapotranspiration will go further down to the subsurface reservoir system divided into three layers based on the speed of response. The top-level storage is called soil zone storage and it holds water body in root zone. Subsurface storage, second layer, controls the water that is available for rapid movement to streams. The third layer, ground water storage, contains slow responding water in the basin and subject to a routing scheme that determines the response time and the amount of the water that goes to groundwater

sink. Basically, these three layers are designed to simulate basin's streamflow response characteristics.

PRMS has been used for many watershed management practices such as water balance study, streamflow prediction, and climate change impact studies (e.g., Leavesley et al. [2002], Salvetti et al. [2002], etc.).

3.4 Model Calibration

Model parameter calibration work was done with the general-purpose global optimization strategy, Shuffle Complex Evolution (SCE-UA) Algorithm [Duan et al., 1993]. This method is based on the multi-start simplex optimization procedure, but includes parameter shuffling concepts to improve the parameter optimization performance for complex watershed models. The core of this method is shuffling initial parameter population at periodic stages in order to share information for the next evolution. Sorooshian et al. [1993], demonstrated that SCE-UA to be very efficient and effective for calibrating the National Weather Service SMA-NWSRFS flood forecast model.

Four major PRMS model states were identified in the work of Hay et al. [2005] as; (1) solar radiation, (2) potential evapotranspiration (PET), (3) annual water balance, (4) daily streamflow components. Among these model states, only the fourth state was calibrated in this study. The model parameters were calibrated against a single objective function that is defined as a sum of two different objective functions; peak streamflow and daily mean flow;

$$OF = OF_{peak} + OF_{daily} \tag{1}$$

$$OF_{daily} = \sum_{i=1}^{n}\left[\frac{abs(Q_{ei} - Q_{oi})}{Q_{oi}}\right] \tag{2}$$

$$\mathrm{OF}_{\mathrm{peak}} = \sum_{i=1}^{n'} \left[\frac{abs(Q_{ei} - Q_{oi})}{Q_{oi}} \right] \tag{3}$$

where Q_o is the observed daily flow, Q_e is the estimate daily flow, n is the number of days, and n` is the number of days of upper 10^{th} quartile flows.

Before the parameter calibration procedure performed in this study, pre-calibration was done against three other objective functions (solar radiation, evapotranspiration, water balance) in the previous work of Clark and Hay [2004]. Only limited manual calibration on those parameters were done for this study.

For the rest of the ensemble stream flow simulations, model calibration was done using an extended ensemble set combined by 15 ensemble sets. Ten years of each ensemble (1983-1992 for Animas basin, 1974-1983 for Alapaha basin) is attached to make 150 years of input time series. This procedure is designed to filter out the possible noise included in the single ensemble through this extended time series, and give more reliable parameter set. The rest of the stream flow simulations were done with this calibrated parameter set. Since the ensemble members have the same statistical property, the parameter sets should be transferable.

4. Results

4.1 Impact on ensemble streamflow generation

PRMS was run in batch mode to produce 100 ensemble stream flow simulations in both Alapaha and Animas basins. The period of record for all simulations is from January 1, 1975 to December 31, 1995. The entire flow simulation process is shown in Figure 6. This flow chart includes the entire calibration, initialization, and ensemble forecast procedure explained in this paper. Figure 7 and 8 show the daily ensemble streamflow simulation and observed stream flow values for the later half of the entire time series (from 1986 to 1995). Solid

circles represent observed streamflow values on each day. The distribution of daily ensemble simulations is showing in gray shades. Outer shade (lightest gray color) shows the range of 95th and 5th quartile of the generated ensemble streamflows on each day. The next shade (next darker gray color) shows the range of 75th and 25th quartile. The narrowest boundary (darkest gray color in the middle) shows the 55th and 45th quartile range. The entire ensemble range, however, is much broader than the shown shaded boundary, and covers most of the observed stream flow values.

One can quickly notice that the ensemble range in the Animas basin is tighter than in the Alapaha basin. This is expected because daily variations in streamflow in the Animas basin are dominated by snowmelt, which is more closely tied to temperature than precipitation (note that temperature is estimated deterministically). Uncertainties in precipitation inputs in the Animas basin are related to uncertainties in total winter accumulation, rather than uncertainties in individual precipitation events. We will demonstrate later that uncertainties in precipitation inputs constitute a small component of total model uncertainty, whereas in the Alapaha basin uncertainties in precipitation inputs are much more important.

It is therefore not surprising that there are instances in the Animas basin when the ensemble streamflow is not well matched with observations (see water years 1991 and 1995).

Annual and monthly flow statistics are shown in Figure 4.8 and 4.9 for the two basins. Gray lines represent simulated ensemble streamflows and the black solid line represents the observed streamflow. Monthly flow graph in Alapaha basin (upper left corner graph of Figure 4.8) shows slightly lower ensemble flow during the peak season and higher flow during low flow season. In Animas basin, the monthly flow graph in Animas basin shows the impact of spring streamflow mismatch shown in Figure 4.7. This lower mean flow during the spring season is

compensated with the higher flow in the fall season. However, in general, both of the two figures show the relatively more varied annual and monthly mean flows in Alapaha basin, which is symptomatic of the strong relationship between uncertainties in precipitation estimates and total model uncertainty in the Alapaha basin.

Although the ensemble flow generation is not a streamflow forecasting process, some of the forecast verification metrics are helpful to understand the behavior of the ensemble streamflow upon input uncertainty. The rank histogram is a useful tool to allow quick examination of ensemble quality [Hamill, 2001]. In an ideal case, the rank histogram should have a flat shape (i.e., a uniform distribution)

The rank histogram is computed as follows. Assume that there are *n* ensemble flow points on a time step and one point of observed flow value. Based on rank of the observation among the n+1 points, plotting position of the observed value is calculated in each time step. The plotting position of this observation x_o in the vector $\mathbf{X}=(x_1, x_2, \ldots, x_n, x_o)$ is calculated as;

$$pp = \frac{rank(x_o)}{n+2} \tag{2}$$

where *pp* is the plotting position and n is the total number of ensemble members on each day. Then, the frequency of this calculated position *pp* on each pre-defined bin is keep added along the time series. The left column of the Figure 10 shows this rank histogram in Alapaha (upper left graph) and Animas (lower left graph) basin.

It is obvious that Animas basin has U shape rank histogram. This means that the ensemble spread is too small. Since we have only accounted for uncertainties in precipitation inputs, the U-shaped ranked histogram suggests that other sources of modeling uncertainty are important in the Animas basin. These additional sources of

uncertainty include uncertainty in temperature inputs, uncertainty in model parameters, and weaknesses in model structure.

In the contrary, Alapaha basin has quite flat shape of rank histogram. This means that the input uncertainty ensemble streamflow captures the observed flow very well. This also implies that the input uncertainty in Alapaha basin is 'large' with respect to other sources of uncertainty. The difficulties and poor performance in this rainfall-dominated basin was already shown in the work of Hwang et al. [2005a].

The reliability diagram is another method to analyze probabilistic model simulations. The reliability diagram of the generated daily ensemble streamflows is shown in Figure 4.11 (right column). The reliability diagram plots the observed frequency against the forecast probability, where the range of forecast probabilities is divided into the pre-defined number of bins. The perfect reliability can be found when a plotted reliability curve overlaps the 1:1 line, which means, forecasted flows shows exactly the same probability on each streamflow category.

The probability of low frequency ensemble streamflows matches quite well with the observed streamflow. In Animas basin, frequency of observation is also quite well reproduced in generated ensemble set. In general, it is quite noticeable that the generated ensembles in both basins reproduced the proper frequency of the observed streamflow quite well. It is quite interesting that the input error based ensembles are not designed as a forecasting tool but only based on the possible input uncertainty of an interpolation model. The last section of this paper will test the possibility of this input uncertainty as a streamflow forecast tool.

4.2 Expansion of ensemble streamflow simulation to forecast

Most operational streamflow forecasting systems ignore hydrologic model error (e.g., Day, 1985). This means that probabilistic risk assessments may be unreliable. Therefore, we evaluate the use of the precipitation input ensembles within a forecasting framework as a first step to integrating model error into forecast applications.

One easy forecasting example is a traditional forecasting method (**E**nsemble **S**treamflow **P**rediction, ESP). In traditional ESP (Day, 1985), the historical data set is repeated for the forecasting period to generate multiple ensemble stream flow. Although the forecasting process is designed to consider uncertain nature of climate variables during the forecasting period, everything is entirely deterministic up to the starting point of the forecast. The range of the generated ensembles are limited mainly because of the two reasons; 1) forecasted ensembles are only generated using historical climatologic events, 2) historical observations and the estimations on the modeling locations are assumed to be true (no uncertainty). The input uncertainty analysis shown in the previous sections are now applied to the ensemble stream flow generation to see the possible improvements in streamflow forecast.

4.2.1 Forecast setup

During the ensemble stream flow generation performed in this section, the same model parameters are used as the previous sections using SCE-UA method. For the traditional ESP, streamflow simulation until the forecast starting point is fully deterministic since the runoff model is calibrated and validated using historical observations. Input climate variables can be available from external products such as weather generator or downscaled output from larger scale climate models for stream flow forecast. In this study, this climatologic forecast is simulated basically by sampling the historical data on the same day of the past years. From the 20 years of

historical data, 20 possible ensemble inputs (resample from ensemble #1) are used for the forecast simulations. For the forecast including input uncertainty, each of the 100 input uncertainty ensemble is used to generate 20 individually re-sampled historical trails for both Alapaha and Animas basins. This procedure makes 2000 ensemble flows compare to the 20 ensembles of traditional ESP.

In Alapaha basin, ensemble forecast is done from January 1, 1996 since the ensemble input trails are generated up to 1995 in the previous section. In Animas basin, two forecast starting points are selected in order to see the different predictability based on the knowledge of snow pack during the snow-accumulating season (September to March). The selected two forecast starting points are April 1, 1995 and September 1, 1994.

4.2.2 Ensemble forecast comparison

Figure 11 shows the trace of calibrated daily streamflow during 1995 and the forecasted streamflow ensembles in 1996 in Alapaha basin. The upper graph shows the traditional ESP results and the bottom graph shows the revised ensemble including input uncertainty. Solid black circles in the upper and lower graphs are showing the observation. Since the streamflow simulation up to the starting point of the forecast is fully deterministic, 20 ensemble traces during 1995 are collapsed into one thick solid line rather than the gray shades such as shown in 1996.

For the ESP simulation, the observed streamflows are generally found within the 5^{th} and 95^{th} quartile range shown by light gray shade. Enhanced ensemble runs show similar but a bit wider range than traditional ESP, and the observations are also within the 5^{th} and 95^{th} quartile range. However, the observed values are much lower during the high flow season (January, February, and March), and the peak timing is quite off the ensemble range. This is expected as we cannot forecast daily variations

in streamflow several months into the future; due to the chaotic nature of environmental systems and the inherent limits on predictability, we can only hope to provide reliable probabilistic estimates of flow attributes on seasonal time scales (e.g., total seasonal volume, the number of days flow is above/below a given threshold, etc.). Consistent with the results in Figure 11, 1996 was a dry year (not shown).

In Animas basin, estimates of the water stored in the winter snowpack provides capabilities for skillful multi-season forecasts of seasonal flow attributes. So, the forecast predictability of the two different forecast starting points tested. September 1st was chosen for the forecast on pre-snow season (less predictability, graph not shown), and April 1st was chosen for the post-snow season (more predictability) forecast. The forecast starting on April shows a much tighter ensemble spread than the September forecast. Ensemble ranges are more effectively capturing the observation, and the 75th quartile ranges of both ESP and enhanced ensembles are expanded near observations. One of the obvious findings in these graphs is forecasts initialized in April better capture the higher flow in the spring melt period.

The spread characteristics of the enhanced ensemble forecast versus traditional ESP are also illustrated between the top and bottom streamflow graphs in the Figure 12. Note that the enhanced ESP forecasts include information on uncertainties in precipitation inputs, and have larger spread than the standard ESP forecasts. While extensive verification is not possible with these example cases, explicitly accounting for additional sources of uncertainty in the forecasting process should improve the reliability of probabilistic risk assessments.

These simple boxplots give some valuable information in terms of the decision-making in water resources management. One can easily convert these

boxplots into probability distribution of streamflows. A boxplot of input uncertainty forecast (wider whiskers than traditional ESP plots) should be transferred into a probability distribution that has longer tail. This means that the risk (or exceeding probability) estimated by the input uncertainty forecast is higher than the probability estimated by traditional ESP forecast.

5. Summary and Discussion

Errors in a precipitation interpolation model could be one of the important sources of input uncertainty in watershed modeling while different interpolation schemes also will impact the streamflow simulation as shown in the work of Hwang et al. [2005b]. In this paper, the impact of input interpolation error was tested by generating ensemble stream flows, and the possible forecast application of the generated ensemble are presented. The results found in this paper are as follows,

- Generated input ensembles show different characteristics on the two basins.
- The interpolation models' lower performance in rainfall-dominated basin (Alapaha) shown in the work of Hwang et al. [2005a] is represented in the statistical feature of the generated ensembles.
- Lower interpolation performance (reflected in higher ensemble spread) implies that the simulated streamflow might be less accurate.
- Differences in input uncertainty in the two basins are well shown in the generated ensemble flows (more varied ensemble flow in Alapaha basin and tighter ensemble flow range in Animas basin).
- Input uncertainty is tested as a possible ensemble forecast tool to enhance traditional ESP in Alapaha and Animas basin. Generated ensemble flows

range captures the observation quite well, and possible wet-dry year forecast can improve the skill (correct skew in forecast).

Input uncertainty in hydrologic modeling is one of the difficult factors of modeling uncertainty problems. A way of considering input uncertainty in hydrologic modeling (albeit it's still a partial input uncertainty) is shown in this paper, and the possible extension to forecasting is also shown. However, the interaction between the uncertainties in stream flow simulation and forecast such as parameter uncertainty, model structure uncertainty, and input uncertainty obviously important.

Although only one forecasting example was shown in this paper using the historical data, this approach can be easily applied for further forecasting practices. Replacing the ensemble sets with random selection, conditioned selection upon climate forecast, or even the output from GCM can be easily used to improve the forecasting skill. The application of climate or weather model to streamflow prediction can be found in many literatures (e.g., Cameron et al., [2000], Jasper et al., [2002], etc). For example, with the framework suggested in this paper, the input precipitation ensembles given for the 1995 streamflow forecast (starting on April 1, 1995) could be selected only from the years that showed high streamflows in the history. This selected sampling can be done using climate signal or any type of long-range weather forecast. For the short-term streamflow forecast, numerical weather prediction output could be downscaled (e.g., Gangopadhyay et al. [2005]) to generate input ensembles for the streamflow forecast suggested in this paper.

Acknowledgments

Partial support of this work by NOAA GAPP program (Award NA16GP2806) and the NOAA RISA Program (Award NA17RJ1229) is thankfully acknowledged. The authors also wish to thank George Leavesley, Lauren Hay, and Steve Markstrom at USGS for providing valuable comments and data that greatly enhanced the quality of the entire analysis.

Reference

Blazkova, S. and K. Beven (2002), Flood frequency estimation by continuous simulation for a catchment treated as ungauged (with uncertainty), *Water Resources Research*, *38*(8), Art. No. 1139.

Cameron, D., K. Beven, and P. Naden (2000), Flood frequency estimation by continuous simulation under climate change (with uncertainty), *Hydrology and Earth System Sciences, 4(*3), 393-405.

Carpenter, T.M., K.P. Georgakakos, and J.A. Sperfslagea (2001), On the parametric and NEXRAD-radar sensitivities of a distributed hydrologic model suitable for operational use, *Journal of Hydrology*, *253*(1-4), 169-193.

Carpenter, T.M. and K.P. Georgakakos (2004), Impacts of parametric and radar rainfall uncertainty on the ensemble streamflow simulations of a distributed hydrologic model, *Journal of Hydrology*, *298*(1-4), 202-221.

Clark, M.P. and A.G. Slater (2005, accepted), Probabilistic quantitative precipitation estimation in complex terrain. Paper submitted to the *Journal of Hydrometeorology*.

Dawdy, D.R. and Bergmann J.M. (1969), Effect of rainfall variability on streamflow simulation, *Water Resource Research*, *5*(5), 958-966.

Duan, Q., V.K. Gupta, and S. Sorooshian (1993), A shuffled complex evolution approach for effective and efficient global minimization, *Journal of Optimization Theory and Applications, 76*(3), 501-521.

Faures, J.M., D.C. Goodrich, D.A. Woolhiser, And S. Sorooshian (1995), Impact of small-scale spatial rainfall variability on runoff modeling, *Journal of Hydrology*, *173* (1-4), 309-326.

Gandin, L.S. (1988), Complex quality-control of meteorological observations, *Monthly Weather Review*, *116*(5), 1137-1156.

Hamill, T.M. (2001), Interpretation of Rank Histograms for Verifying Ensemble Forecasts, *Monthly Weather Review*, 129, 550-560.

Hay, L. E., M.P. Clark, R.L. Wilby, W.J. Gutowski, G.H. Leavesley, Z. Pan, R.W. Arritt, and E.S. Takle (2002), Use of Regional Climate Model Output for Hydrologic Simulations, *Journal of Hydrometeorology*, vol.3, 571-590.

Hwang, Y., M.P. Clark, B. Rajagopalan, S. Gangopadhyay, and L.E. Hay (2005a), Spatial interpolation schemes for daily precipitation, submitted to *Journal of Hydrometeorology*.

Hwang, Y., M.P. Clark, B. Rajagopalan, G. Leavesley, and L.E. Hay (2005b), Impact of different spatial interpolation methods on runoff simulation, submitted to *Advances in Water Resources.*

Jasper, K., J. Gurtz, and L. Herbert (2002), Advanced flood forecasting in Alpine watersheds by coupling meteorological observations and forecasts with a distributed hydrological model, *Journal of Hydrology*, *267*(1-2), 40-52.

Krajewski, W.F., V. Lakshmi, K.P. Georgakakos, and S.C. Jain (1991), A Monte Carlo study of rainfall sampling effect on a distributed catchment model, *Water Resources Research*, *27*(1), 119-128.

Kunkel, K.E., K. Andsager, G. Conner, W.L. Decker, H.J. Hillaker, P.N. Knox, F.V. Nurnberger, J.C. Rogers, K. Scheeringa, W.M. Wendland, J. Zandlo, and J.R. Angel (1998), An expanded digital daily database for climatic resources applications in the midwestern United States, *Bulletin of the American Meteorological Society*, *79*(7), 1357-1366.

Landman, W.A., S.J. Mason, P.D. Tyson, and W.J. Tennant (2001), Statistical downscaling of GCM simulations to streamflow, *Journal of Hydrology, 252*(1-4), 221-236.

Leavesley, G.H., Lichty, R.W., Troutman, B.M., and Saindon, L.G. (1983), Precipitation-Runoff Modeling System: User's Manual: U.S. Geological Survey Water-Resources Investigations 83-4238, 207 p.

Leavesley, G.H., P.J. Restrepo, S.L. Markstrom, M. Dixon, and L.G. Stannard (1996), The modular modeling system - MMS: User's manual, *Open File Report 96-151*, U.S. Geological Survey.

Leavesley, G.H., S.L. Markstrom, P.J. Restrepo, and R.J. Viger (2002), A modular approach to addressing model design, scale, and parameter estimation issues in distributed hydrological modeling, *Hydrological Processes, 16*(2), 173 – 187.

Loader, C. (1992), *Local Regression & Likelihood*, Springer-Verlag, New York

Michaud, J.D. and S. Sorooshian (1994a), Comparison of simple versus complex distributed runoff models on a midsized semiarid watershed, *Water Resources Research, 30*(3), 593-605.

Michaud, J.D. and S. Sorooshian (1994b), Effect of rainfall-sampling errors on simulations of desert flash floods, *Water Resources Research 30*(10), 2765-2775.

Pessoa, M.L., R.L. Bras, E.R. Williams (1993), Use of weather radar for flood forecasting in the Sieve River Basin; a sensitivity analysis, *J. of Applied Meteorology*, *32*, 462-475

Rajagopalan, B. and U. Lall (1998), Locally weighted polynomial estimation of spatial precipitation, *Journal of Geographic Information and Decision Analysis*, *2*(2), 44-51.

Reek, T., S. R. Doty, and T. W. Owen, 1992, A deterministic approach to the validation of historical daily temperature and precipitation data from the cooperative network, *Bulletin American Meteorological Society*, *73*(6), 753-762.

Serreze, M.C., M.P. Clark, R.L. Armstrong, D.A. McGinnis, and R.S. Pulwarty (1999), Characteristics of the western United States snowpack from snowpack telemetry (SNOTEL) data, *Water Resources Research*, 35(7), 2145-2160.

Seyfried M.S., B.P. Wilcox (1995), Scale and the nature of spatial variability - field examples having implications for hydrologic modeling, *Water Resources Research*, *31*(1), 173-184.

Shah, S.M.S., P.E. O'Connel, and J.R.M. Hosking (1996), Modeling the effect of spatial variability in rainfall on catchment response: 2. Experiments with distributed and lumped models, *J. Hydrol*., *175*, 89-111.

Singh, V.P. (1997), Effect of spatial and temporal variability inrainfall and watershed characteristicson stream flow hydrograph, *Hydrol. Process*, *11*, 1649-1669.

Winchell, M., H.V. Gupta, and S. Sorooshian (1998), On the simulation of infiltration- and saturation-excess runoff using radar-based rainfall estimates: Effects of algorithm uncertainty and pixel aggregation, *Water Resources Research*, *34*(10), 2655-2670.

Tables

Table 1. Basin information

	Basin	
	Animas	Alapaha
Data Period	1970.1~1995.12	1974.1~1995.12
# of HRUs	121	180
# of stations	37	28
Elevation Range (m)	2000-3700	40-125
Drainage Area (km^2)	1792	3626
Network Density (station/km^2)	0.021	0.008

Figures

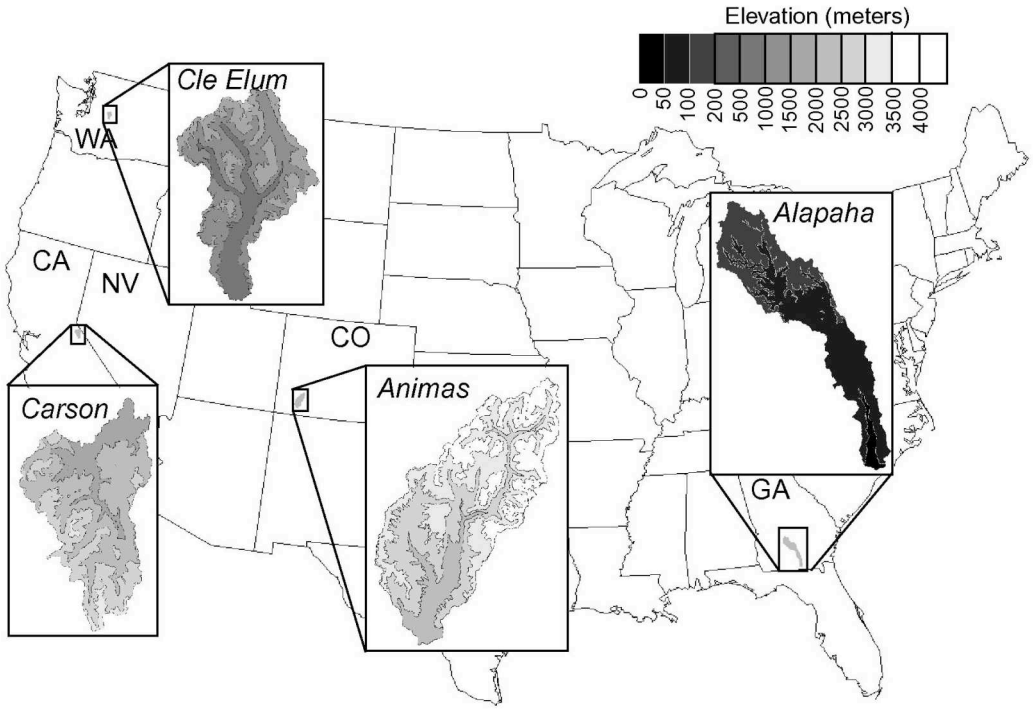

Figure 1. Map showing the study regions (detailed results from Carson and Cle Elum basins are shown here)

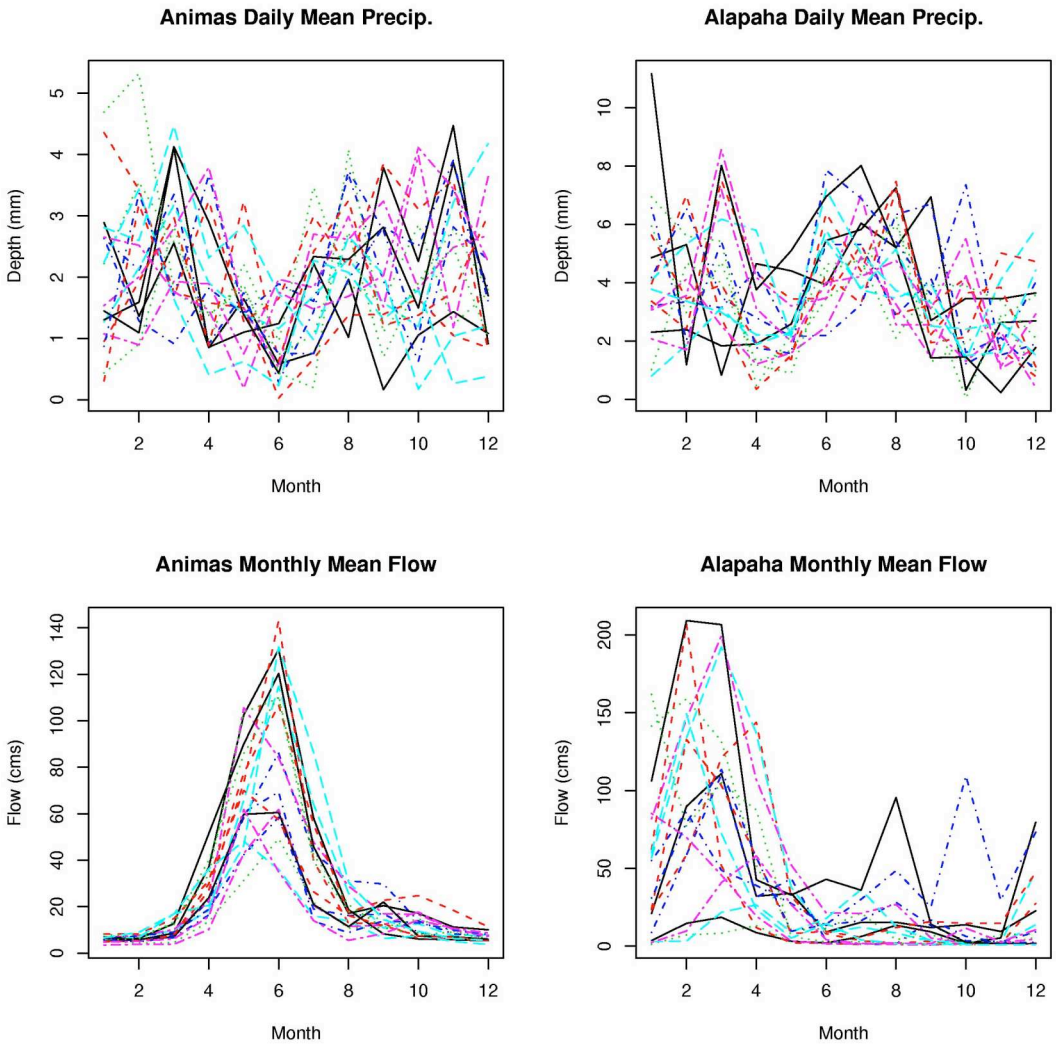

Figure 2. Monthly mean precipitation and observed stream flow in Animas and Alapaha basins. Each line stands for each year in the entire time series used in this study.

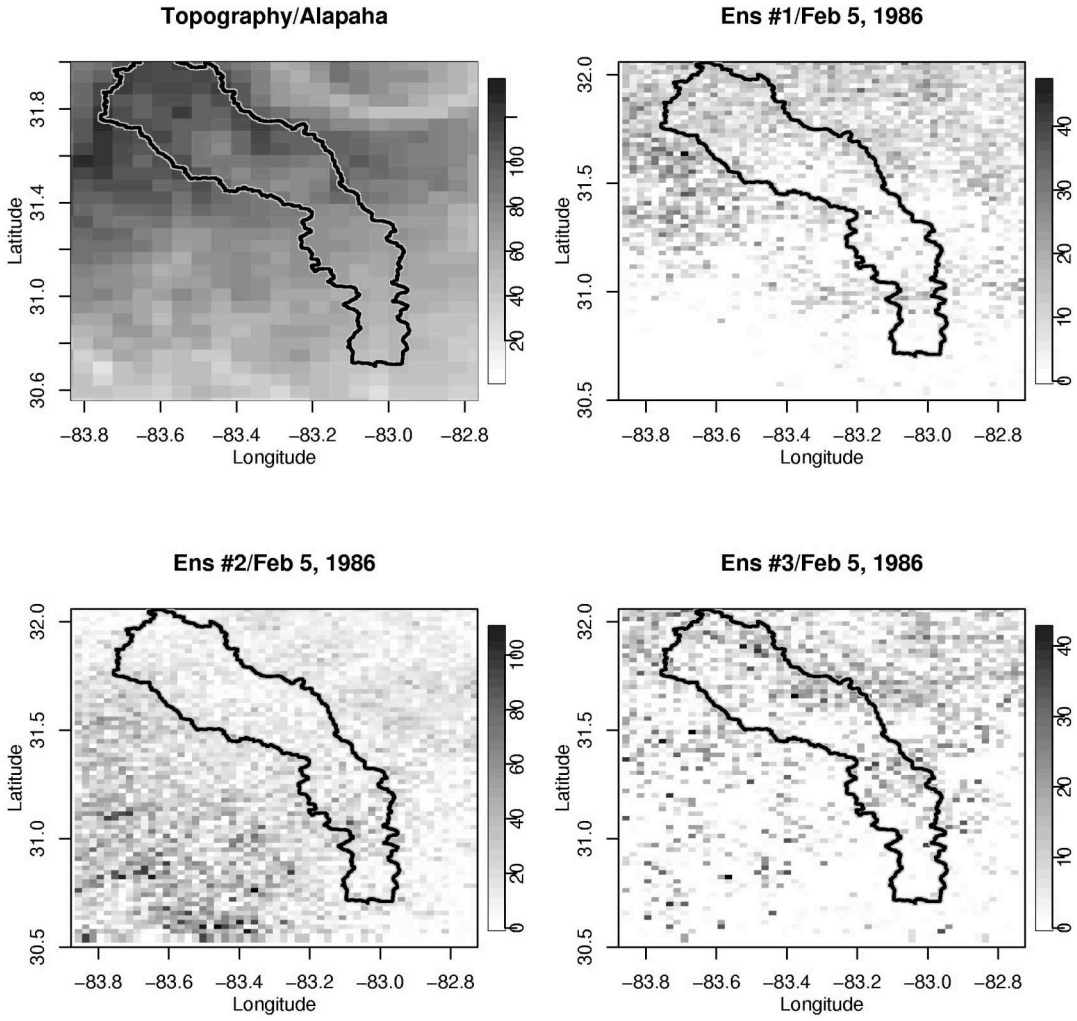

Figure 3. Generated sample ensemble members on February 5, 1986 in Alapaha basin. Top-left graph is the topography plot of the basin.

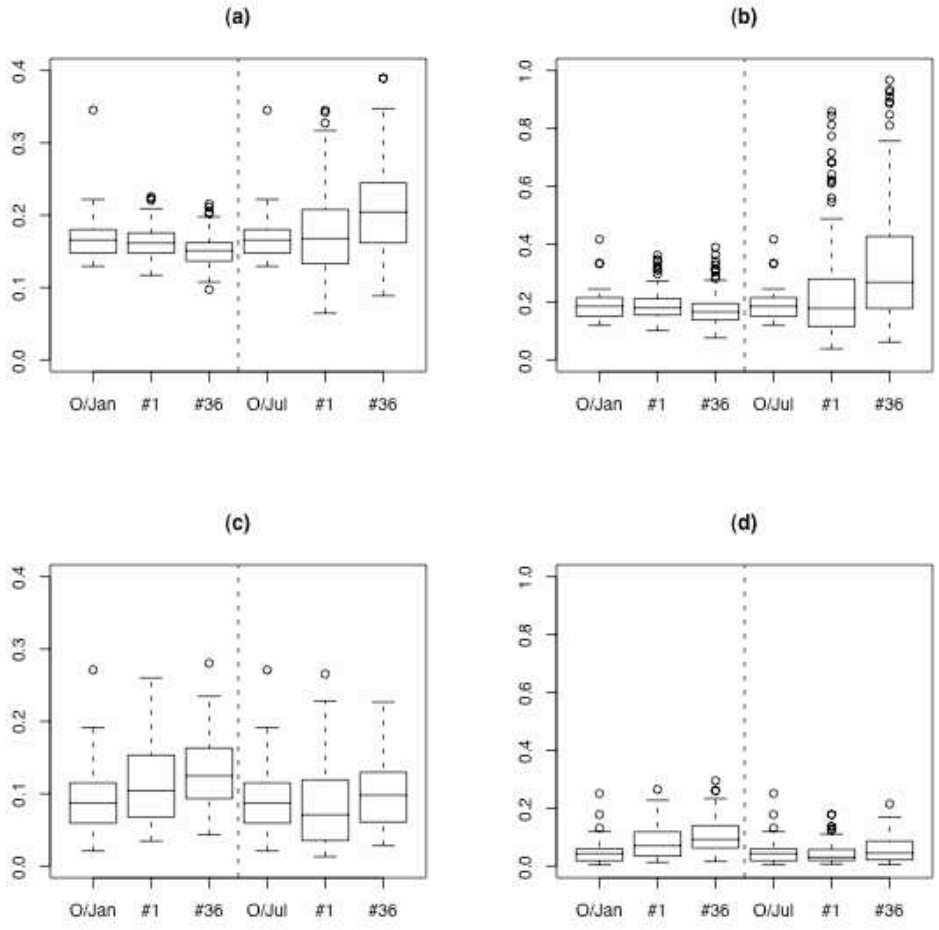

Figure 4. Comparison of mean and variance of stations and grid cells in Alapaha (graphs (a) and (b)) and Animas (graphs (c) and (d)) basin. Left column shows mean and right column shows variance statistics. 'O' stands for observations and '#1' and '#36' stands for ensemble number 1 and 36, respectively. Ensemble #1 is used for model parameter calibration.

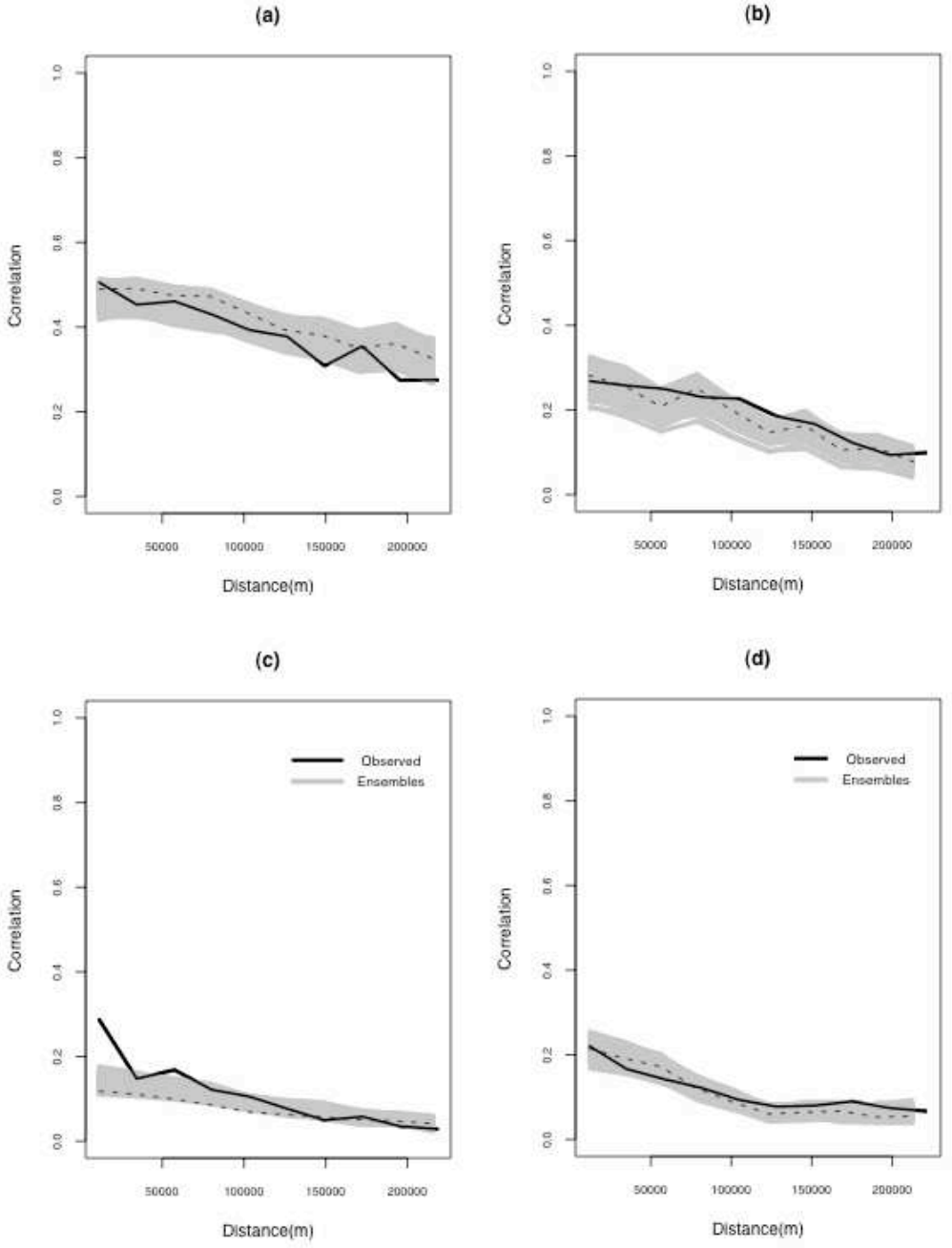

Figure 5. Lagged distance correlation in Alapaha (graphs (a), (c)) and Animas (graphs (b), (d)) basin. Solid and dashed black lines represent observations and ensemble #1 (used for runoff model calibration), respectively.

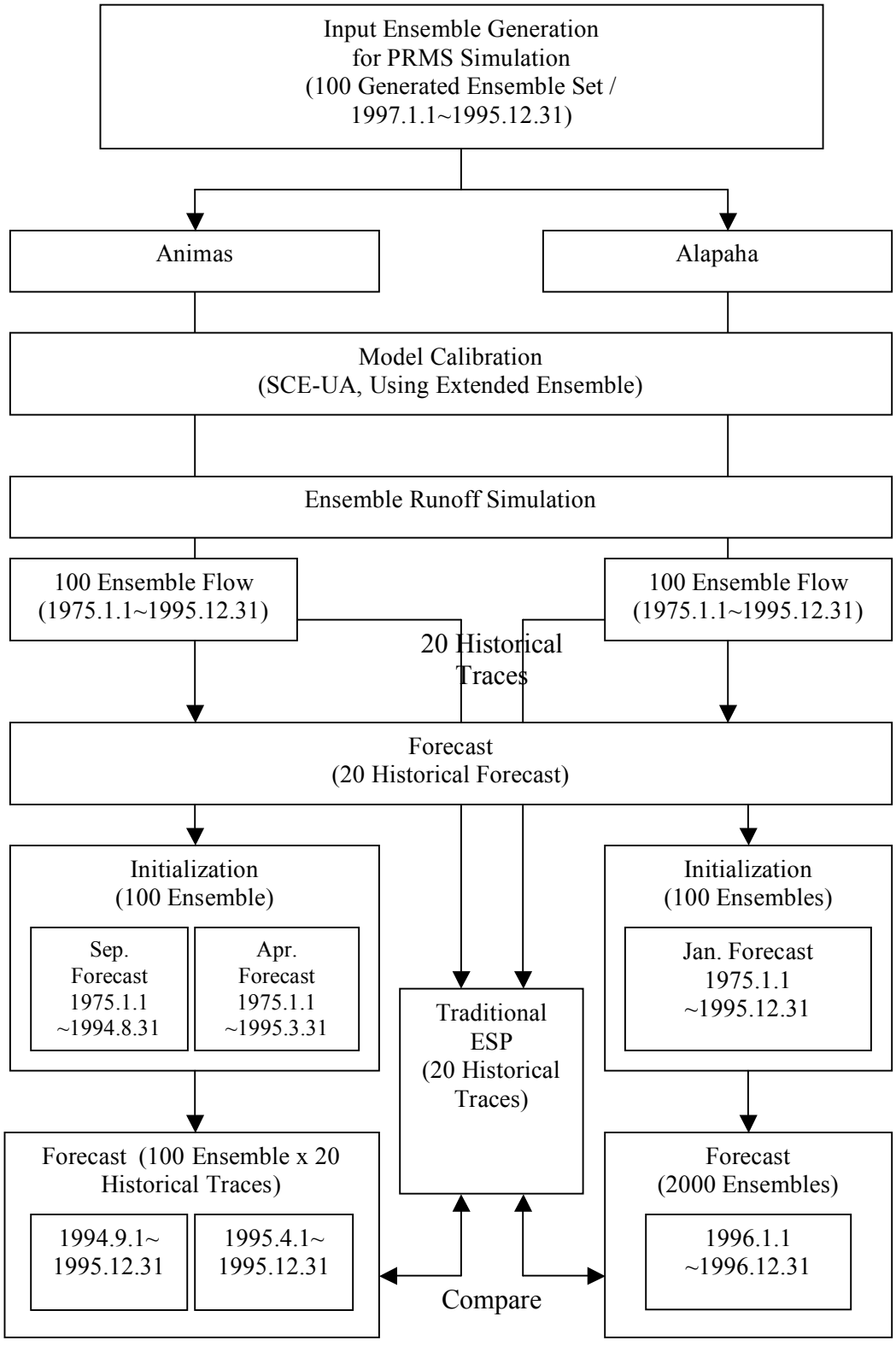

Figure 6. Ensemble flow simulation and forecast workflow

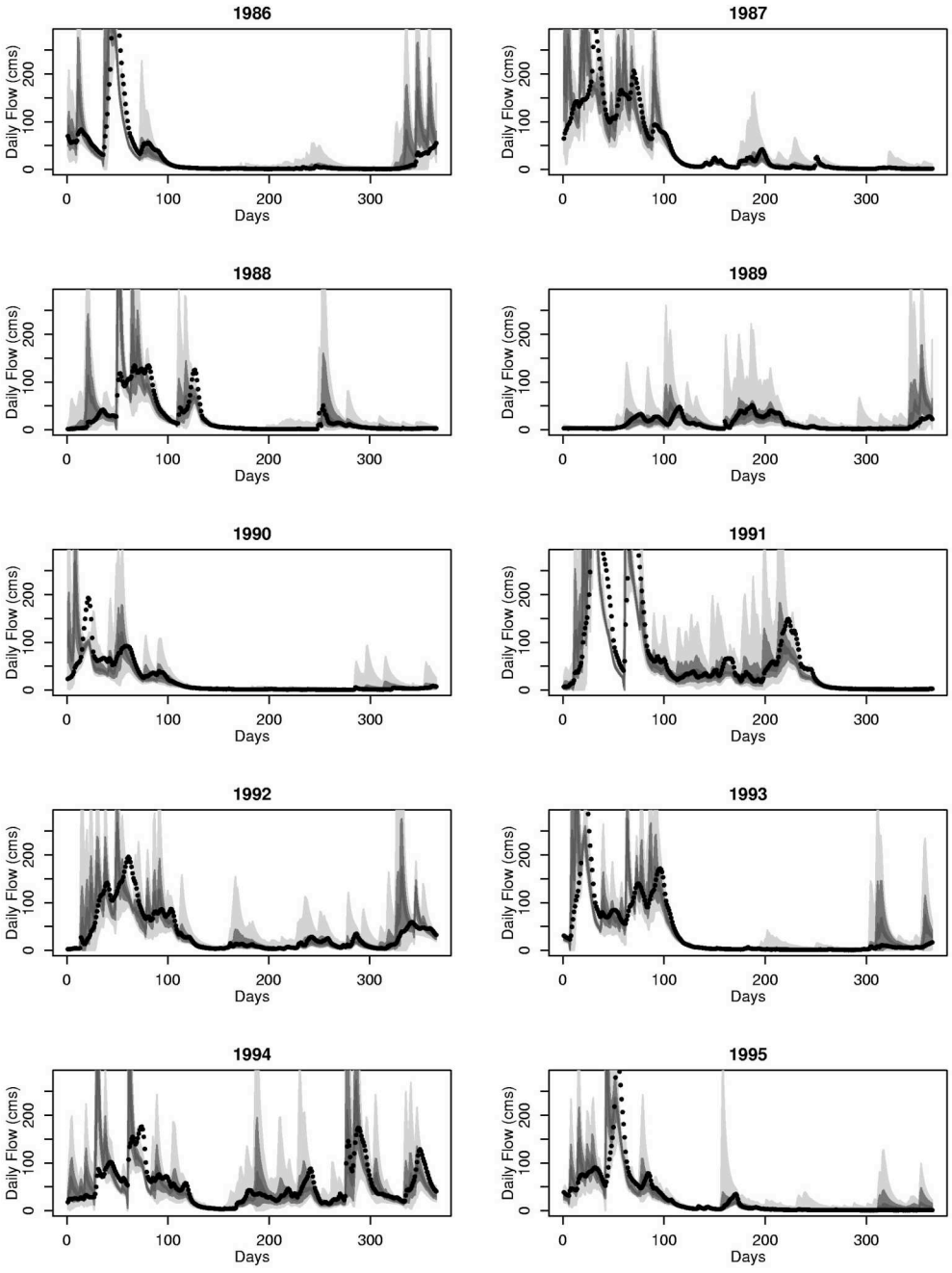

Figure 7. Simulated and observed daily runoff in Alapaha basin. Shades represent ensemble stream flows and solid dots represent observations.

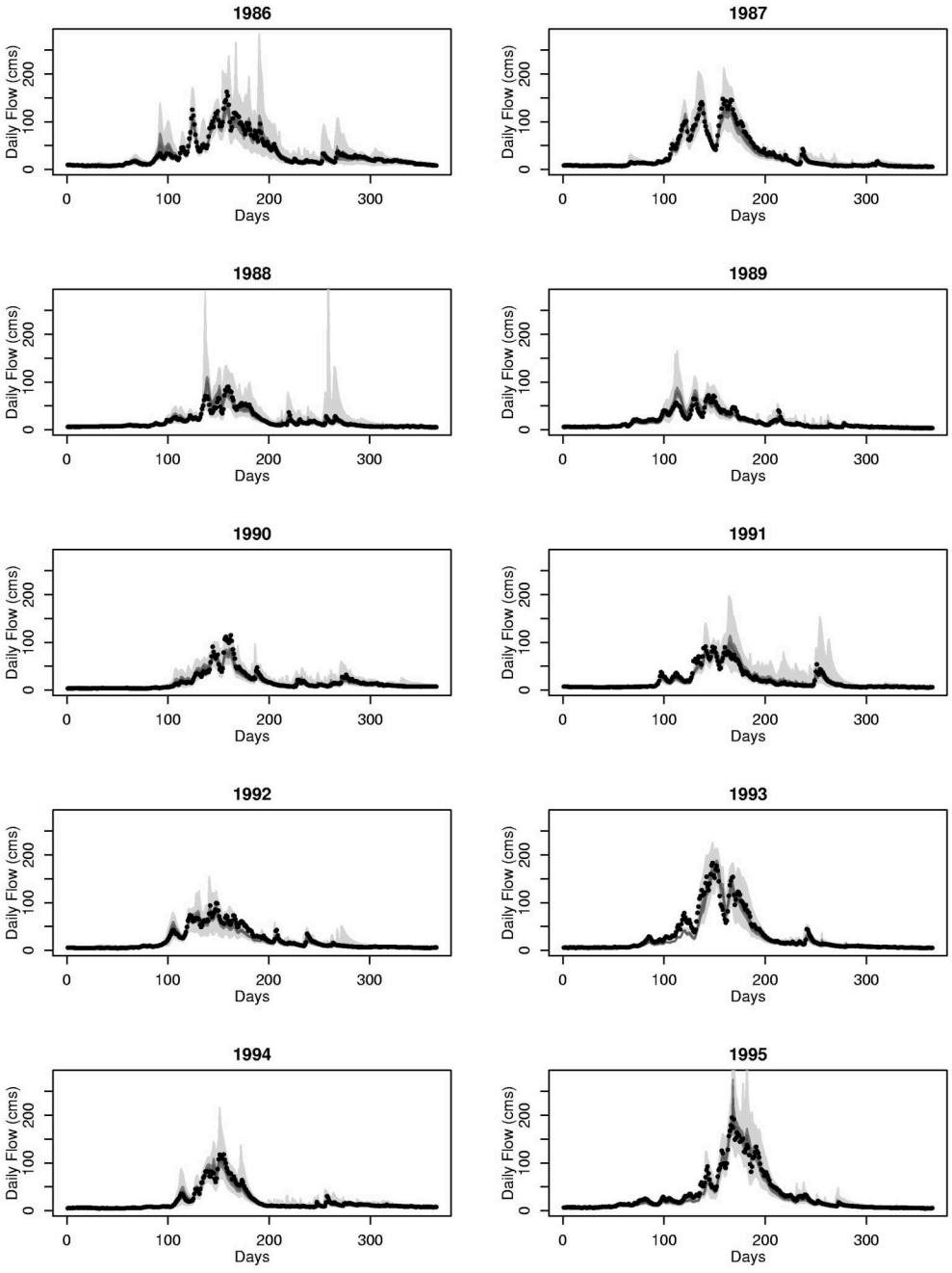

Figure 8. Same as Figure 4.6 but in Animas basin.

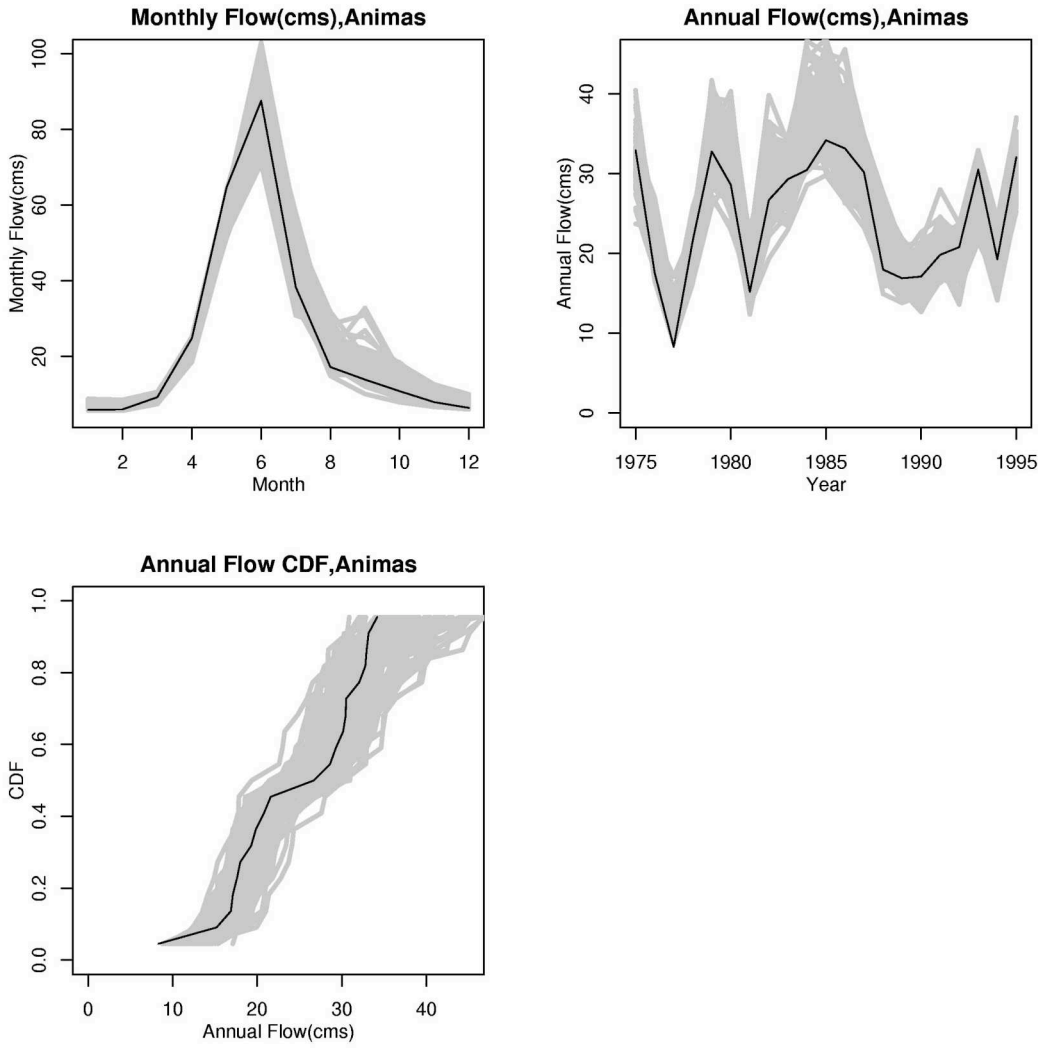

Figure 9. Selected statistics of simulated ensemble runoffs in Animas basin. Top-left: Monthly mean flow. Top-right: Annual mean flow. Bottom-left: Cumulative probability of annual flow. Thick gray lines represent 100 ensemble flows.

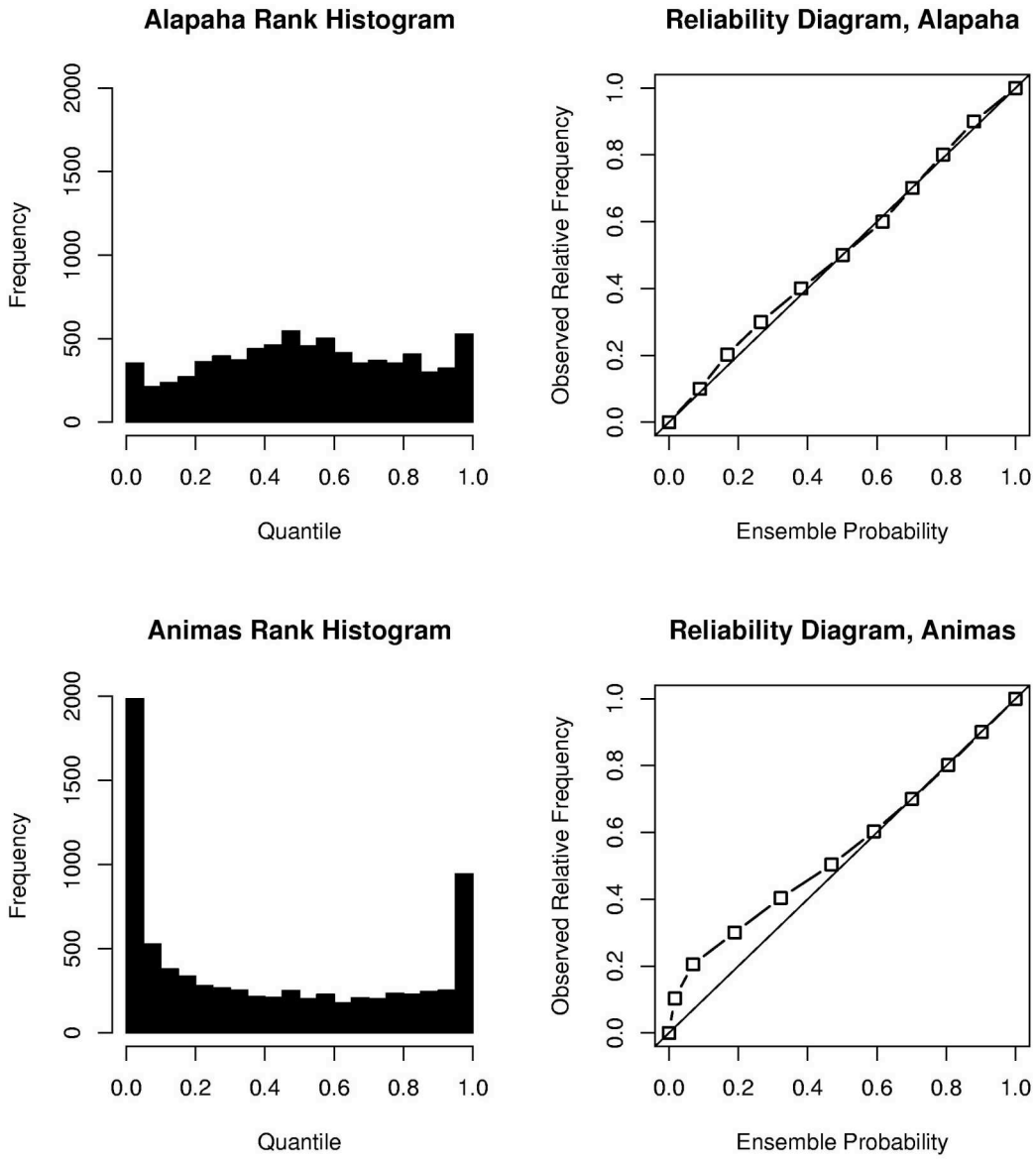

Figure 10. Rank histogram and reliability diagram for Alapaha (upper row) and Animas (lower row) basin.

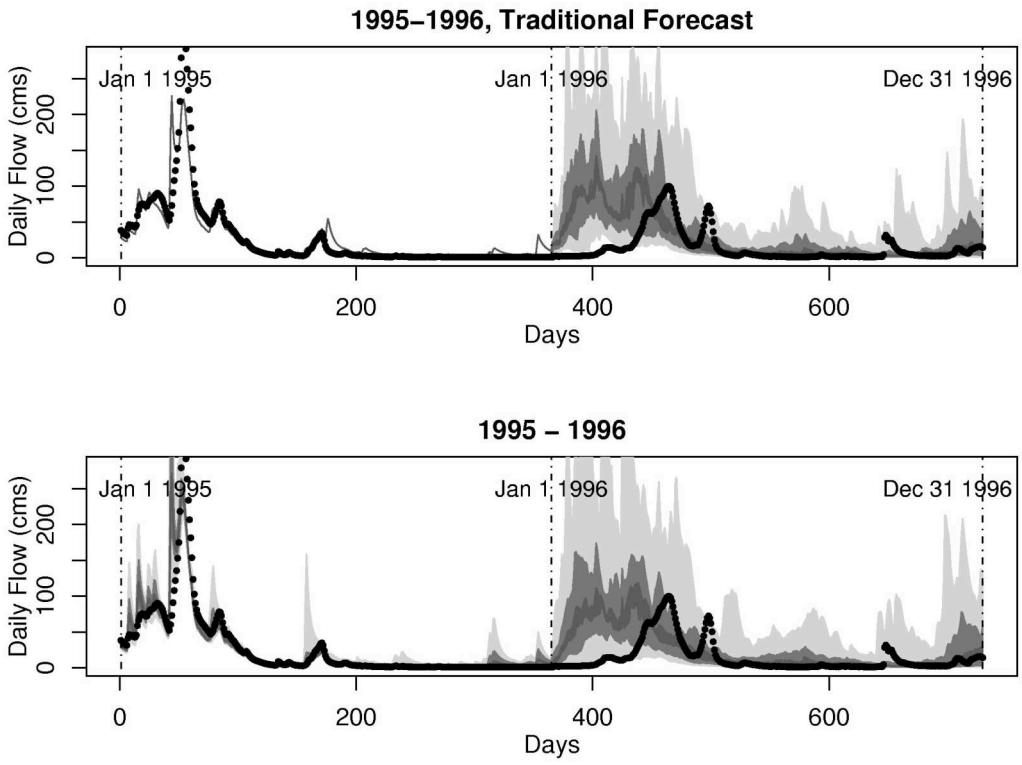

Figure 11. Forecasted and observed daily stream flow. The gray lines represent ensemble flows (20 traces for traditional ESP method and 2000 traces for input uncertainty ensembles) using the historical rainfall picked from the record. The solid circles represent observation.

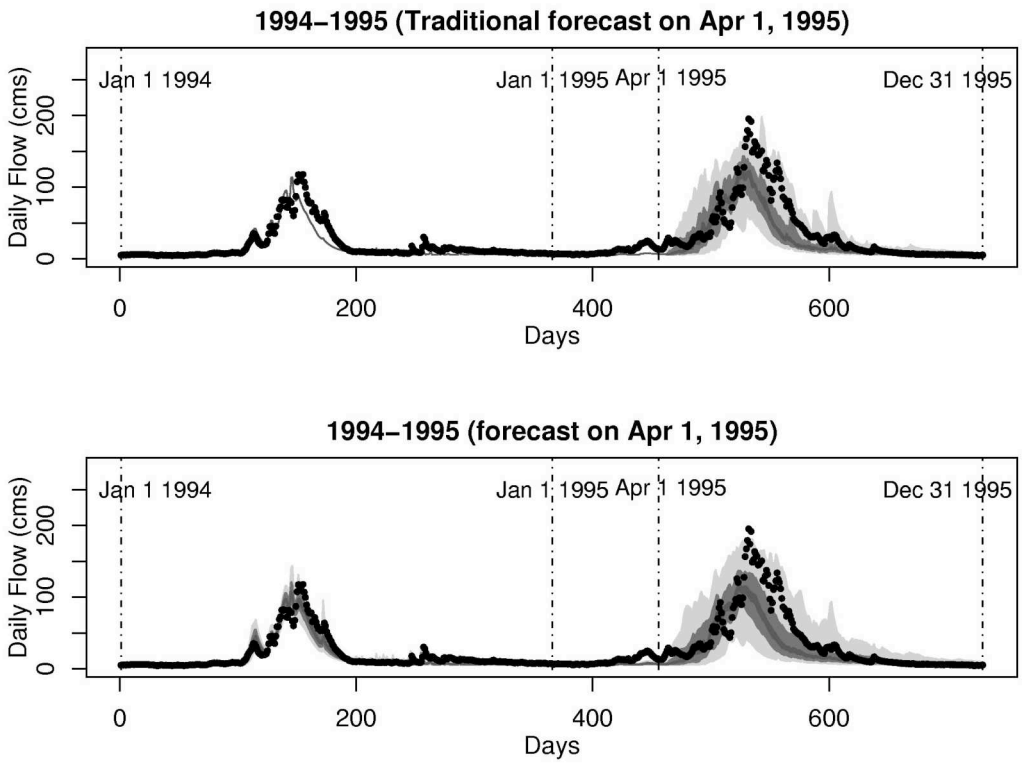

Figure 12. Comparison of traditional ESP (top) and input uncertainty added ensemble forecast (bottom) forecasted after the snow-accumulating season (April 1).

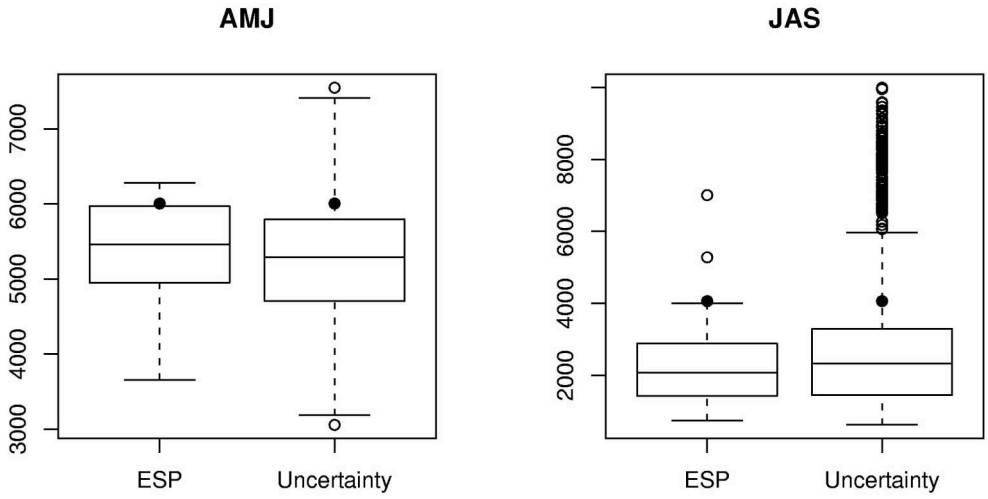

Figure 13. Boxplots of the total flow of the forecast ensembles shown in Figure 12 for in two different seasons.