PRESENTED AT THE 28TH ANNUAL CONFERENCE AND SYMPOSIUM ON MANAGING WATER RESOURCES DURING GLOBAL CHANGE, RENO, NEVADA, NOVEMBER 1-5, 1992.

Verification of the National Weather Service Extended Streamflow Prediction Procedure

Gerald N. Day[1], Larry E. Brazil[2], Cara S. McCarthy[1], and Don P. Laurine[3]

ABSTRACT: The National Oceanic and Atmospheric Administration (NOAA) National Weather Service (NWS) is working with a number of other federal, state, multi-state, quasi-governmental, and private sector organizations toward cooperative efforts in applying advanced hydrometeorological data collection and forecasting techniques to address water management problems. As part of these efforts, NOAA is planning a new initiative called the Water Resources Forecasting System (WARFS), which will capitalize on advanced technologies to provide improvements in hydrological prediction services.

The infrastructure for WARFS is the current NWS River Forecast System and its Extended Streamflow Prediction (ESP) procedure. The ESP procedure assumes that meteorological events that occurred in the past are representative of events that may occur in the future. Each year of available historical meteorological data is input to conceptual hydrologic/hydraulic models to simulate a possible streamflow trace. Probabilistic forecasts can be produced from these simulated traces for maximum flow, minimum flow, volume of flow, reservoir state, etc., for any period in the future.

It is impossible to judge the procedure's performance based only on how closely the forecast value matches the observed value for a few forecast years. ESP's performance should be judged by whether or not the generated probabilistic forecasts accurately represent the true probability distributions of the forecast variables. The purpose of this paper is to use historical data on a real basin to determine if, in a case study, ESP can produce representative probabilistic forecasts.
KEY TERMS: forecasting, water management, verification.

INTRODUCTION

As development, pollution, and climate variability stress our water resources, long-range forecasts for water management become extremely important. There are,

[1] Research Hydrologist, Hydrologic Research Laboratory, National Weather Service 1325 East-West Highway, Room 8338, Silver Spring, MD 20910 (301)713-0640

[2] Senior Water Resources Engineer, Riverside Technology, inc., 2821 Remington St., Fort Collins, CO 80525 (303)223-2944

[3] WARFS Implementation Leader, Colorado Basin River Forecast Center, National Weather Service, 337 North 2370 West, Salt Lake City, UT 84116 (801)524-5130

inherently, large uncertainties associated with these long-range forecasts because of the chaotic nature of the atmosphere and our current limited skill in forecasting the resulting weather patterns more than a few days into the future. In spite of these limitations, the hydrologic system usually provides the forecaster interested in the water management problem several months of memory. In snow basins or during extreme drought memory from hydrologic conditions may extend several seasons. This hydrologic memory translates into forecasting skill.

It is becoming prohibitively expensive to design water management facilities to handle worst case scenarios. Instead, risk-based decision systems are being used to optimize water management operations in order to increase effective water yields. These systems are designed to handle probabilistic forecast information such as reservoir inflows. The value of this forecast information is a function of its uncertainty, as well as the forecaster's ability to define this uncertainty.

Extended Streamflow Prediction (ESP)

ESP is the portion of the National Weather Service River Forecast System (NWSRFS) which enables a hydrologist to make extended probabilistic forecasts of streamflow and other hydrological variables (Day, 1985). A schematic of the ESP procedure is shown in Figure 1. ESP assumes that historical meteorological data are representative of possible future conditions and uses these as input data to hydrologic models along with the current states of these models obtained from the daily forecast component of NWSRFS. A separate streamflow time series trace is simulated for each year of historical data using the current conditions as the starting point for each simulation. The streamflow time series traces can be analyzed for peak flows, minimum flows, flow volumes, etc., for any period in the future. A statistical analysis is performed using the values obtained from each year's simulation to produce a probabilistic forecast for the streamflow variable.

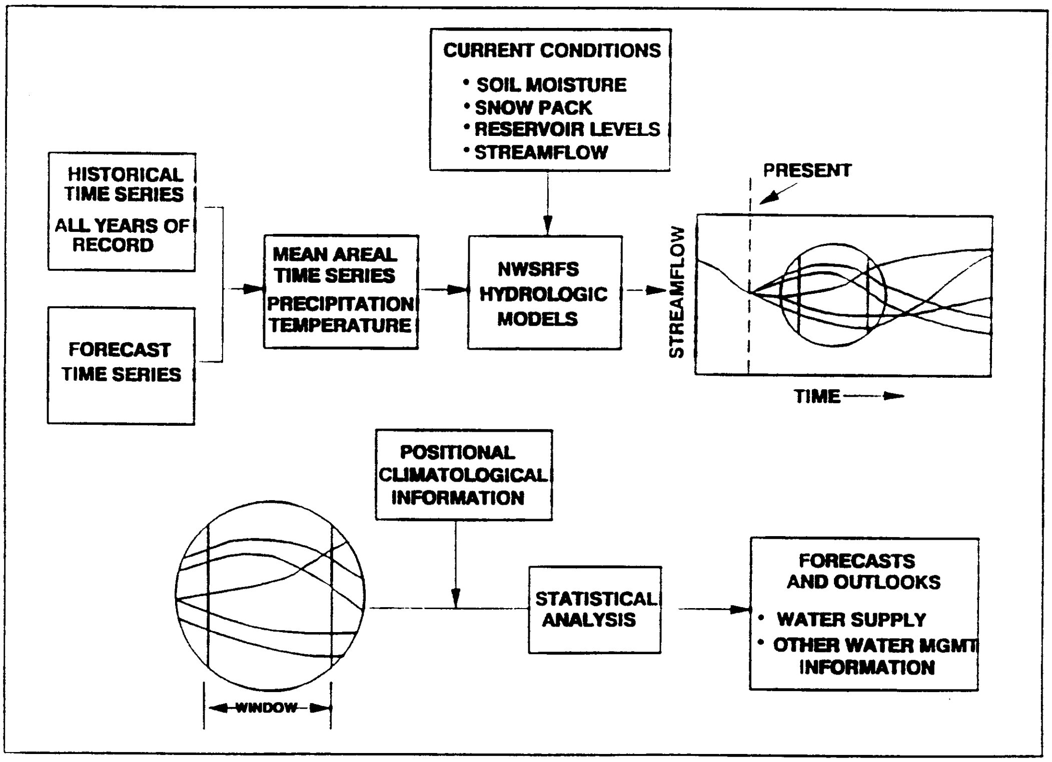


Figure 1. The ESP Procedure

This analysis can be repeated for different forecast periods and additional streamflow variables of interest. Short-term quantitative forecasts of precipitation and temperature can be blended with the historical time series to take advantage of any skill in short-term meteorological forecasting. In addition, knowledge of the current climatology can be used to weight the years of simulated streamflow based on the similarity between the climatological conditions of each historical year and the current year.

ESP allows flexibility in the streamflow variables which can be analyzed, the capability to make forecasts over both short and long time periods, and the ability to incorporate forecast metereorological data into the procedure. Because of the flexibility and conceptual basis of ESP, it has many water management applications, including water supply forecasts, flood control planning, drought analysis, hydropower planning, and navigation forecasts. The ESP probabilistic forecasts provide uncertainty information needed by water managers for risk-based decisions. The streamflow time series generated by ESP can be output as products, so that they can be used in reservoir simulation/optimization models to investigate how operations might be improved. The ESP forecast information is based on our best estimate of current hydrometeorological conditions, as well as an analysis of the local historical climatological variability.

Verification Problem

Water managers interested in using ESP forecasts typically want to know how good the information is. In the world of short-term forecasting, comparison of a deterministic forecast to the associated observed flow parameter for a single event provides an indication of the accuracy of the forecast model and value of the information. By monitoring the performance of a model over a number of events, error statistics can be compiled to give a more complete picture of the model's forecast skill. How does one judge the skill of a probabilistic forecast technique? One event provides very little information for verification since the forecast does not predict what will happen, but only defines the probabilities of what could happen. Only by observing the performance of a probabilistic forecast technique over many events can conclusions be drawn about the technique's skill. This is made difficult by the fact that there are typically fewer long-range forecast events available for verification. For the case of snowmelt water supply forecasts there is only one event per year, but it is important that the skill of the technique be quantified so that users might incorporate this information into operational decisions.

VERIFICATION METHODOLOGY

An ESP Verification System has been developed to provide quantitative estimates of the forecast skill provided by ESP for a particular basin. The system is made up of two components: a trace generation component and a trace analysis component. In ESP a streamflow time series trace is generated for each year of historical meteorological data with the current hydrologic conditions (e.g. snow pack, soil moisture, etc.). In the ESP Verification System the streamflow traces are generated for historical dates so that ESP forecasts can be reconstructed for these historical dates. The assumption is made that the years are independent events and that a representative ESP forecast can be generated for any historical date by excluding that year from the historical data base when the traces are generated. In order to generate an ESP forecast for a date in historical year i, traces are generated using the meteorological data for years 1,2,..., i-1, i+1, i+2, ...,n where n is the number of years in the historical data base. The trace generation program is designed to step through the historical record generating ESP traces for each historical date of interest using hydrologic conditions for that date. Traces are one-year in length and can be generated at weekly or monthly intervals throughout the historical period.

Streamflow Trace Analysis

Once the historical streamflow traces have been generated for a particular date, ESP forecasts can be computed for a particular streamflow variable, e.g. volume, and a particular forecast window. The conditional mean and variance of the forecast streamflow variable can be computed directly from the streamflow traces. A probability distribution must be assumed in order to compute probabilities of exceedance. In this study the empirical distribution was used where sample point plotting probabilities are calculated using Weibull's formula: (Linsley et al., 1975)

$$p = \frac{m}{n + 1}$$

where p is the exceedance probability, n is the number of years of data, and m is the rank of the sample point. The sample point with the largest value is assigned m = 1. Linear interpolation between sample points was used to compute particular exceedance probability values.

Two factors must be considered in verification of ESP:
1) Are the probabilistic statements from ESP correct?
2) How much skill is contained in the forecasts?

If one simply forecasts the marginal distribution of the streamflow variable , the probabilistic statements would be accurate, but the procedure would possess very little forecast skill. In this study, the forecast skill was assessed by computing a forecast root mean square (RMS) error assuming that the forecast is the conditional mean estimated by ESP. This value was compared to the standard deviation of the marginal distribution which is the same as the root mean square error using the mean of the marginal distribution as the forecast. The resulting measure of forecast skill is expressed as a percent reduction in RMS error.

Verification of the exceedance probability values estimated by ESP is more difficult. A visual comparison can be made between the forecast and the observed distributions. Plots were constructed which show the forecast exceedance probability versus the actual observed exceedance probability. For a given forecast exceedance probability P_f, the observed exceedance probability is computed as:

$$p_o = \frac{k}{l}$$

where P_o is the observed exceedance probability, k is the number of ESP forecast years in which the observed value exceeded the forecasted value with exceedance probability P_f, and l is the total number of ESP forecast years. In addition, each forecast can be transformed to a standardized deviate:

$$z = \frac{o - \bar{x}}{\sigma}$$

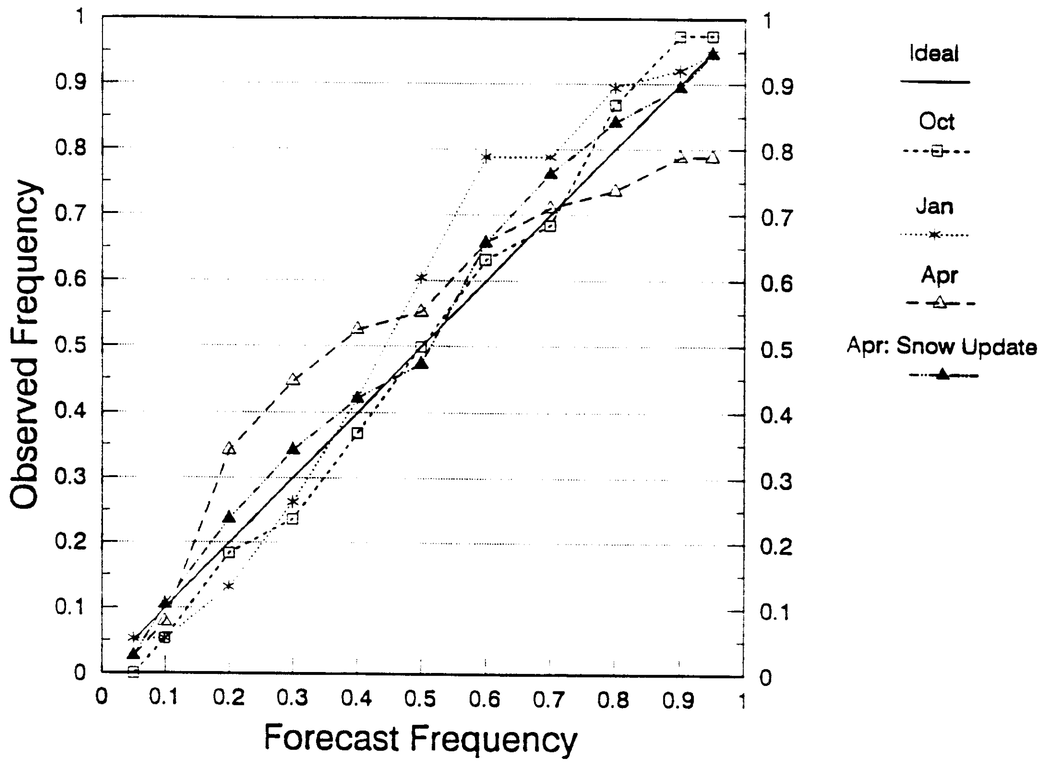
where z is the standardized deviate, o is the observed streamflow, $\bar{x}$ is the mean of the forecast distribution, and σ is the standard deviation of the forecast distribution. If x and σ are truly the mean and standard deviation of the conditional streamflow distribution, then the transformed variable z will have a mean of zero and a variance of one. Formal hypothesis testing requires that an assumption is made about the conditional streamflow distribution. If the conditional distribution is normal, a t-statistic can be used to test if the mean of z is zero. Even when the conditional distribution is not normal the test becomes valid for large n. If the distribution is known (or can be assumed), other tests can be used to verify the ESP forecasts, e.g. Chi-Square Goodness of Fit test or Kolmogorov-Smirnov test.

VERIFICATION RESULTS

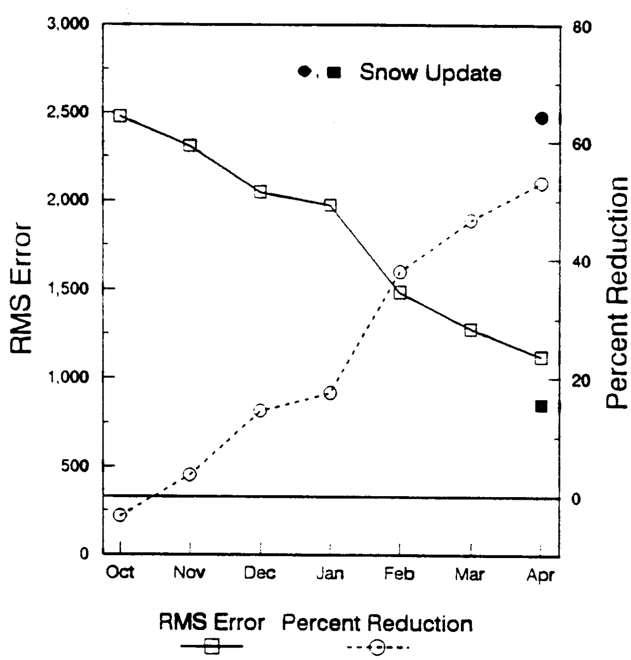
One of the most important applications for ESP-type forecasts is water supply forecasting in areas with significant snow accumulation. Two basins in the Upper Colorado Basin were used to test the ESP Verification System and to develop some preliminary indication of the accuracy of ESP forecasts. The Animas Basin above Durango, Colorado is in the headwaters of the San Juan River Basin. The Animas was chosen because it is relatively free of diversions, the models used to forecast the Animas have been well calibrated, and a procedure for incorporating observations of snow water equivalent has been implemented for the basin. The basin draining into Dillon Reservoir on the Blue River was chosen because it is being used as part of a demonstration project to show the value of WARFS technologies and ESP type forecast information in water management applications (Greer et al., 1992). Plans are being made to recalibrate this basin, but preliminary ESP verification results would provide useful input to the demonstration project.

The objectives of the verification study were to check the validity of the probabilistic information generated by ESP and to assess the overall forecast skill of the procedure. The verification system that was developed has the capability to compute verification statistics for several streamflow variables, including volumes, peaks, and low flows. The system also has the capability to analyze different forecast windows beginning on any date for which streamflow traces are available and extending through the length of the traces which is typically one-year. In a basin which is dominated by snowmelt, seasonal water supply forecasts are usually very important. In this study April through July was selected as the forecast period and verification statistics were computed to test ESP forecast skill for the first of each month October through April. A preliminary analysis of the verification results for the Animas Basin revealed a slight bias in the April through July forecast volumes. A more careful analysis indicated that the model undersimulated the volume for this window by about 5 percent for the historical period of 1949 through 1987. An adjustment factor was applied to the traces and the verification results for the seasonal volume forecast are shown in Figure 2.

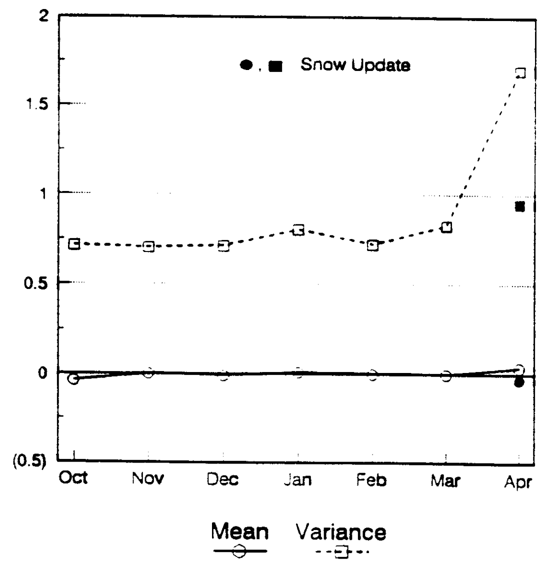
Part a) of the figure shows the forecast frequency plotted versus the frequency that was actually observed for the October, January, and April forecasts. In addition, a separate forecast was made for April after the states of the snow model had been updated using observations of snow water equivalent. It has been shown (Day, 1990) that significant improvements in water supply forecasting can be achieved by using snow cover observations to correct for data and modeling errors. In general, all the lines exhibit a small amount of scatter about the ideal 45-degree line. The amount of scatter increases as the forecast season progresses. For the April forecast there is a tendency to overestimate the lower exceedance probabilities and underestimate the higher exceedance probabilities. This is probably due to the fact that the ESP forecasts underestimate the forecast uncertainty since ESP accounts for the variability of the future weather, but it does not currently account for the uncertainty introduced through the model states, the historical data, or the models themselves. Early in the forecast season the future weather is the predominant source of uncertainty, but as the season progresses the other sources of uncertainty become more important. The April



(a) Frequency Verification



(b) RMS Error and Percent Reduction



(c) Z Value Statistics

Figure 2. Animas Basin: Streamflow statistics for April-July volume forecast.

forecast that was based on the snow update follows the 45-degree line much better than the April forecast without the update. This may be due to the fact that there is less uncertainty in the model states after the update.

Part b) of Figure 2 shows how the RMS error in the forecast changes with forecast date. It also shows forecast skill expressed as a percent reduction in RMS Error from a forecast based on the mean of the marginal distribution. The forecasts begin to exhibit skill in December and increase in skill through the season. The updated forecast in April exhibits significantly more skill than the April forecast without the update. Part c) of the figure shows the mean and variance of the standardized deviate of the forecast as defined in the Streamflow Trace Analysis section. The mean is consistently zero as indicated by the plot and by hypothesis testing using the Student's t statistic. The variance ranges from 0.71 to 0.83 through March and jumps to 1.7 in April. The z based on the updated April forecast has a variance of 0.96. A null hypothesis that the variable z is from a Normal distribution with a mean of zero and a variance of 1 is consistently accepted.

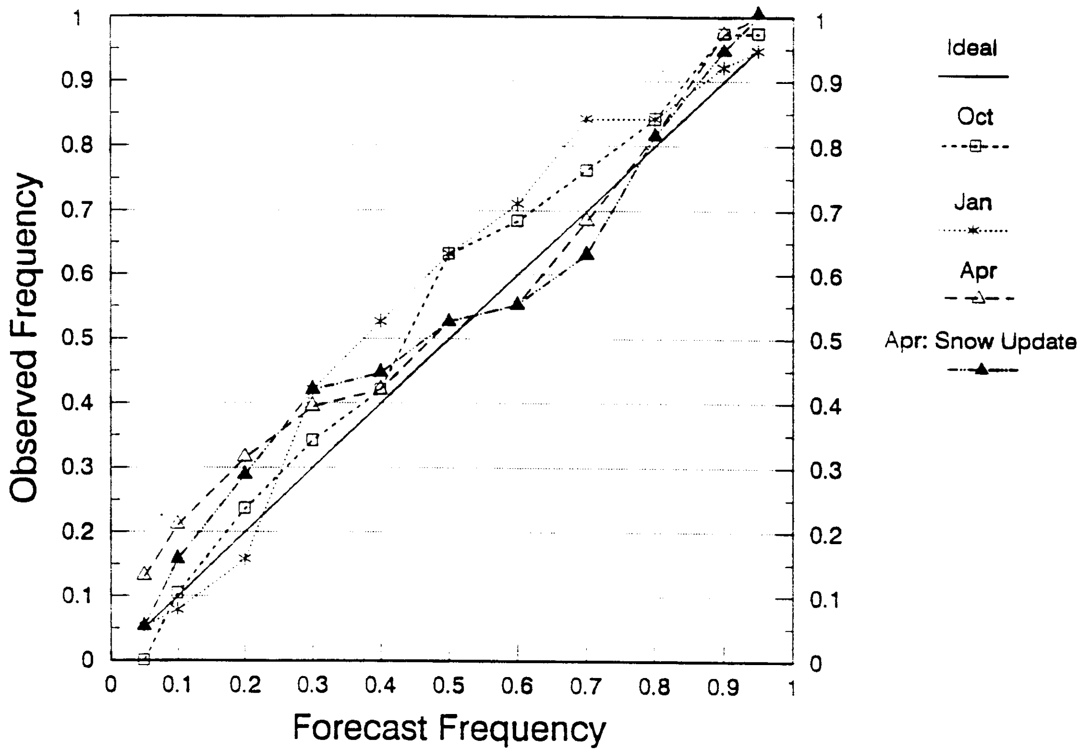
Figure 3 presents verification statistics for the forecast of the peak daily flow during the April through July window. The forecast versus observed frequency comparison in part a) of the figure is very similar to the one for volumes. The procedure does a reasonable job in estimating exceedance probabilities, but there is significant scatter about the ideal line for all months. The line based on the updated April forecast is not significantly better than the line for the April forecast without updating. Part b) of the figure indicates that the percent reduction in forecast error was significant, but it was not as large as it was for the volume forecast. In addition the improvement from updating was not as large.

Verification runs were also made for the seasonal volume inflow forecast for Dillon Reservoir. These results are shown in Figure 4. These values were not adjusted for model bias, but part c) of the figure shows that there is a consistent tendency to produce a negative standardized deviate which indicates that the models oversimulate the April through July volume. Even without the adjustment the procedure produced significant reductions in the RMS error. The large variances computed for z late in the forecast season may be due to poor model calibrations and a resulting underestimate of the uncertainty.

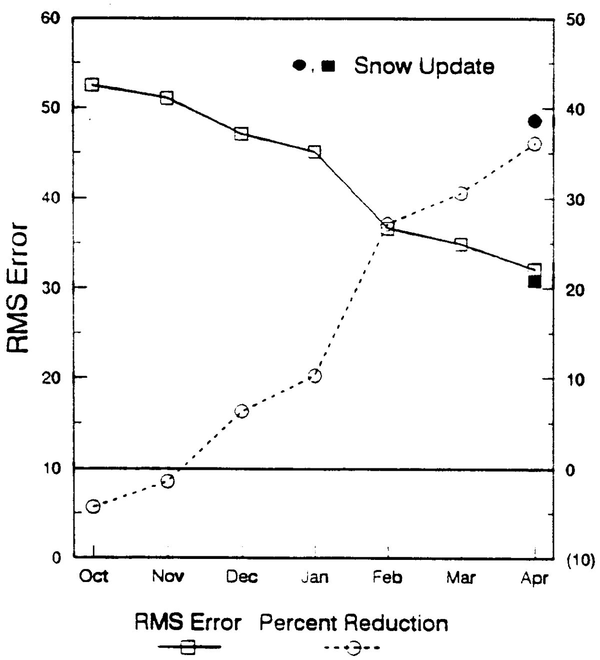
SUMMARY/CONCLUSIONS

A verification system has been developed for the NWS ESP procedure. The system includes a streamflow trace generation component and a trace analysis component. The trace generation component produces ESP forecast traces for historical dates. The trace analysis component analyzes the reconstructed historical ESP forecasts to produce forecast error information for specific forecast dates, forecast windows, and forecast output variables. This information is used to assess the ability of ESP to estimate streamflow variable exceedance probabilities as well as the overall skill of the forecast procedure.

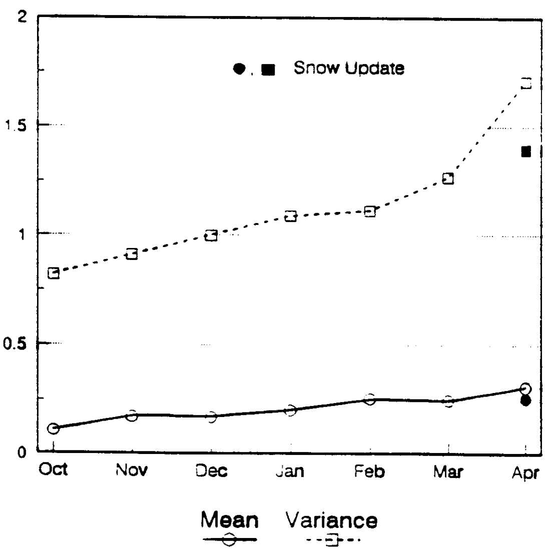
Verification results were presented for two basins. The results for the Animas Basin showed that the procedure exhibited significant skill in forecasting the seasonal volume and peak daily flow. In addition, the procedure provides a reasonable estimate of the conditional distribution of the streamflow variable. The verification results for the Dillon Reservoir seasonal volume inflow forecast were not as good as the results for the Animas Basin. Better model calibrations would probably improve these results.



(a) Frequency Verification



(b) RMS Error and Percent Reduction



(c) Z Value Statistics

Figure 3. Animas Basin: Streamflow statistics for April-July maximum mean daily flow forecast.

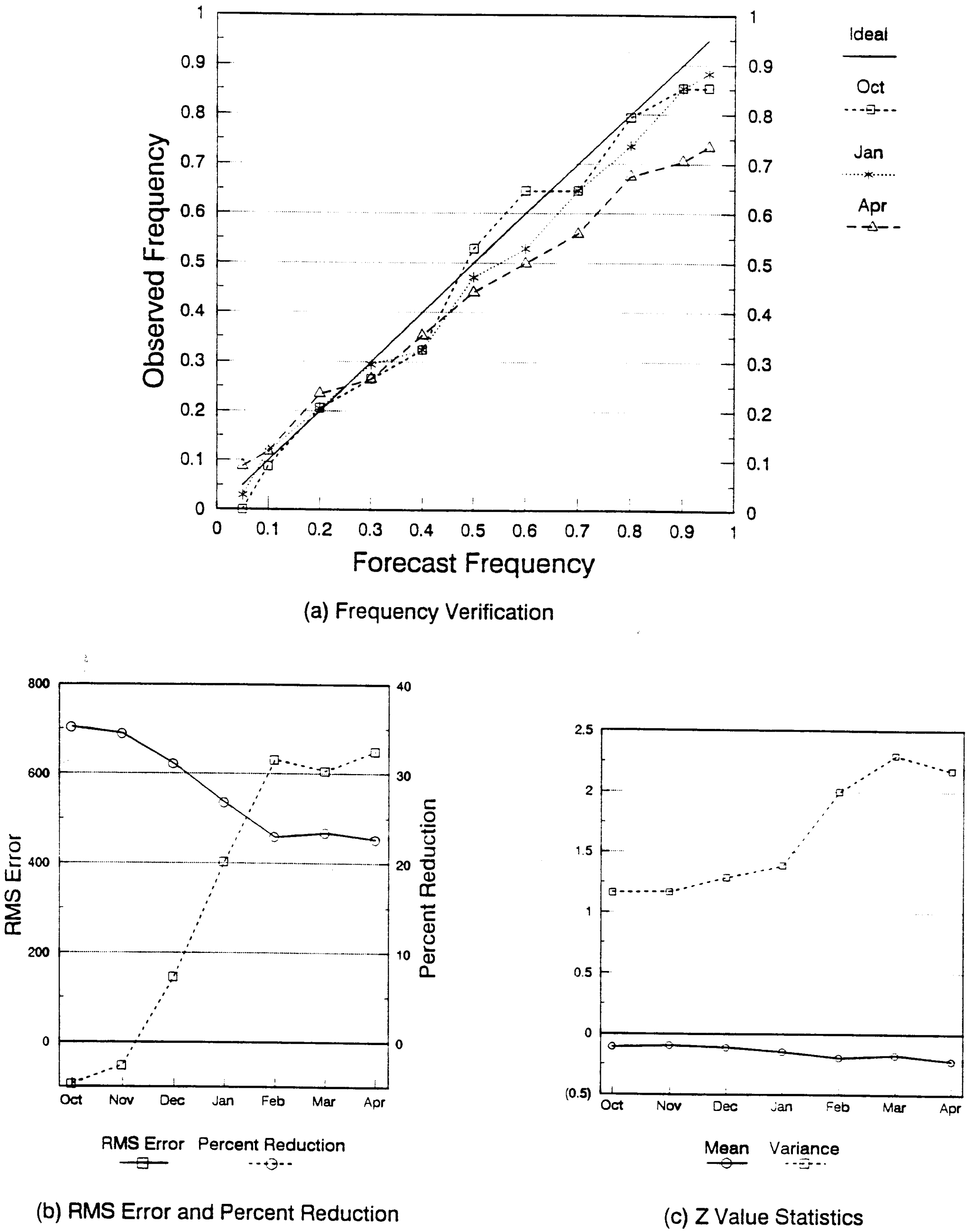


Figure 4. Dillon Reservoir: Streamflow statistics for April-July volume forecast.

The verification system should be run for additional forecast windows and forecast variables on these and other basins to establish reasonable expectations for the forecast skill provided by ESP in different parts of the country at different times of the year. In the future the system may be used routinely to estimate forecast skill after a basin has been calibrated but before it is brought on-line operationally. This will ensure integrity of NWS ESP forecast products and increase user confidence in these products.

REFERENCES

Day, G.N., 1985. "Extended Streamflow Forecasting Using NWSRFS, "Journal of Water Resources Planning and Management, ASCE, 111(2), pp. 157-170.

Day, G.N., 1990. "A Methodology for Updating a Conceptual Snow Model with Snow Measurements," NOAA Tech. Report NWS 43, U.S. Department of Commerce, NOAA/NWS, Silver Spring, MD, 133p.

Greer, D.B., M. Waage, D.P. Laurine, and W. Liu, 1992. "Role of Hydrologic Forecasting in Improved Water Management Operations," Proceedings of AWRA Symposium, Managing Water Resources During Global Change, Reno, NV.

Linsley, R.K., M.A. Kohler, and J.L.H. Paulhus, 1975. Hydrology for Engineers, McGraw-Hill, New York, 482p.