ANALYSIS OF WATER SUPPLY FORECAST ACCURACY

BY

J. C. Schaake, Jr. [1/] and E. L. Peck [2/]

INTRODUCTION

As part of a Federal interagency study to consider placing hydrometeorologic stations in wilderness areas of the Colorado River Basin, an analysis is being made of the factors contributing to errors in the forecasts. These factors include errors because of uncertainty in future climate (i.e. uncertainty in long-range weather events), limitations of data networks and limitations of forecast models. A statistical rationale is being developed to estimate these errors and to predict how these errors might be reduced by various improvements in data networks, models, and climate forecasts. This paper presents the early stages of development of the rationale and discusses the current status of the analytical methods being used.

The primary interest of the interagency study is to assess the possible improvements in forecasts that could be achieved by having additional data. Only forecast points down stream from wilderness areas in the Upper Colorado River Basin are being studied. Potential hydrometeorological data sites inside and outside of wilderness are being considered. The study is aimed at assessing what forecast improvements would occur for any particular gaging alternative, why is it reasonable to expect these particular improvements and what difference does it make (quantitatively) if gages are located in one place versus another.

CRITERIA FOR MEASURING FORECAST ACCURACY

To make a practical study of forecast accuracy, it is necessary first to define what forecasts are to be studied and then to define some measure of accuracy. In this case, the volume forecast of runoff for the period April 1 through July 31 was selected for study. The initial work reported in this paper was to understand the error components of the forecasts for inflow to Lake Powell on the assumption that results learned for Lake Powell would apply at least in principle to forecasts upstream.

The accuracy of a forecast may be described in terms of the statistical properties of the error

$$E_Q = Q - \hat{Q} \qquad (1)$$

where Q is the observed runoff and $\hat{Q}$ is the forecast. Possible statistical measures are: bias, standard deviation, variance, absolute error, and root

---

Presented at the 53rd Annual Western Snow Conference, April 1985, at Boulder, Colorado.

1/ Hydrologic Services Division, National Weather Service, Silver Spring, Maryland.

2/ HYDEX, Inc., Fairfax, Virginia.

mean square (RMS) error. In this case, it was found that the historical bias in forecasts for Lake Powell is very small and can be neglected in this study. The low forecast bias is not too surprising because the historical forecasts have been made using regression-based forecast equations. Much more significant is the variability of the forecast from year to year and the improvements that occur as any one year progresses from January 1 to the end of the forecast period. The important error properties to be studied can be measured in terms of the standard deviation, but the measure chosen for analysis in this study was the error variance. The error variance was chosen because there are well known rules in the theory of statistics (Mood, 1950) for analyzing the components of variance.

The error variance can be expressed in units of runoff volume squared and as a percentage of the variance of the runoff volume being predicted. For example, the standard deviation of the April-July runoff volume to Lake Powell for the period 1947 to 1984 was 4.2 million cubic dekameters (MCD) (3.4 million acre feet (MAF)). The variance is 17.5 MCD^2 (11.5 MAF^2). If a forecast were made (hypothetically) many years in advance, there would be no information to use to distinguish that forecast from any other. The best that could be done would be to forecast the long-term mean. And the error variance of that forecast would equal to 100 per cent of the variance in the runoff volume being forecast. As information becomes available, the forecast accounts for more of the variability of the runoff and this can be expressed as a percentage of the total runoff variance. The accuracy of forecasts for the volume inflow to Lake Powell from 1947 to 1984 is illustrated in figure 1. By January 1, enough is known to account for about one-third of the total runoff variance; by April 1, about two-thirds is accounted for.

COMPONENTS OF FORECAST ERROR

This study is essentially an investigation of possible sources of forecast error. There is no absolute or unique way to approach such an investigation because there are many perspectives and degrees of detail with which to view the situation. In this case, a goal was to keep the analysis as simple as possible, using subjective judgment where necessary as long as that judgment represented the corporate view of the work group.

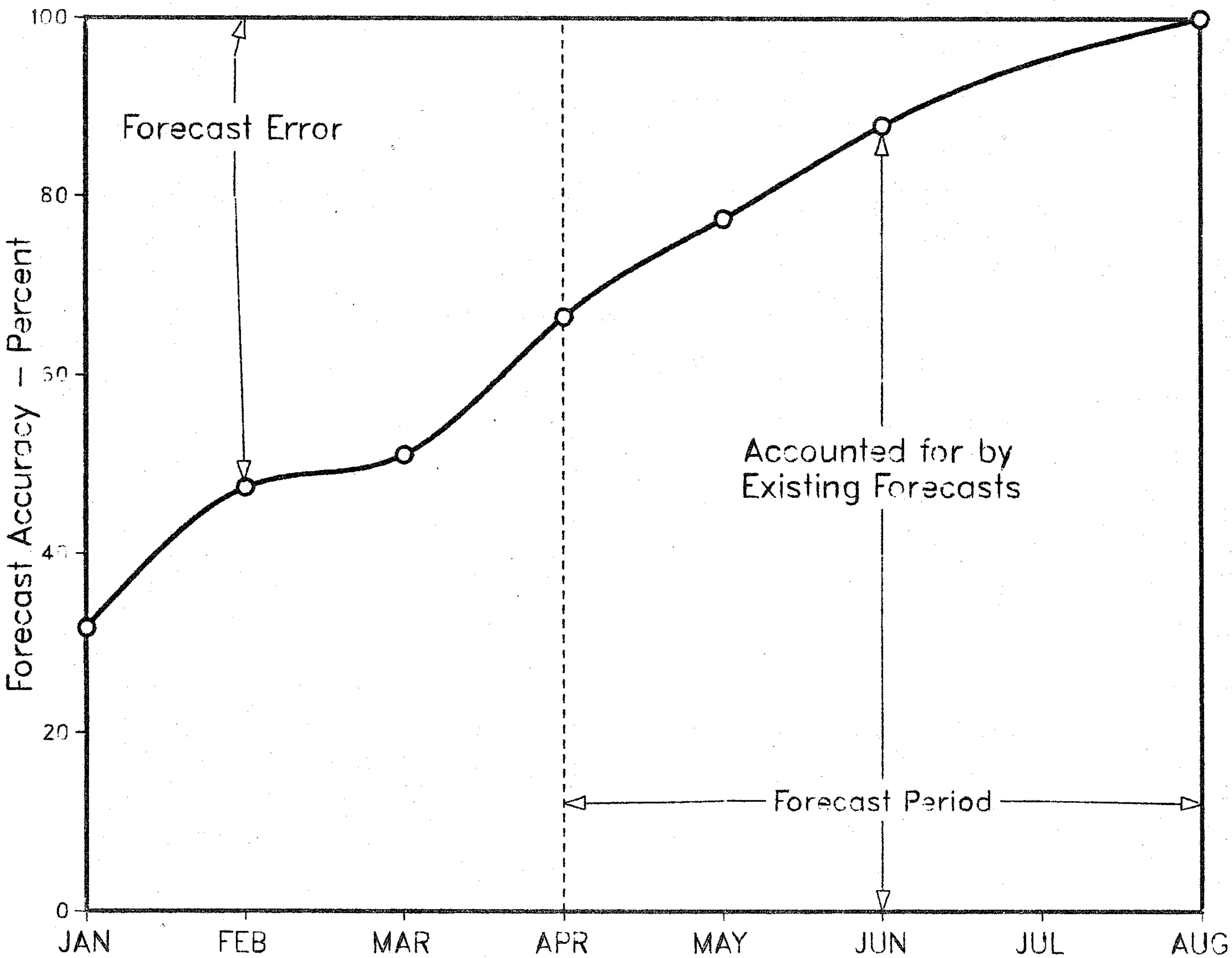
Conceptually, there are at least three main sources of forecast error. These are:

(i) Climate error
(ii) Model error
(iii) Data error

The climate error occurs because future precipitation and temperature changes important to the forecasts have not occurred and cannot be predicted with certainty. Improvements in long-range weather or climate forecasts could potentially reduce this source of error. One way to estimate the magnitude of the climate error component is to operate the forecast model with and without information on the future events. The improvement in accuracy could be attributed to the climate error. The remaining errors could be attributed to limitations of model and data.

FIGURE 1

Present Forecast Accuracy of
April — July Runoff Volume Forecasts
of Inflow to Lake Powell



This was done for forecasts at Lake Powell for the 1947-84 period. For each forecast, observed precipitation data for future months was used in the forecast procedure and the improvement in forecast accuracy is shown in figure 2. As expected, the importance of climate information decreases considerably from January to June. The remaining total of model plus data error remained nearly constant at about 30 percent of the total runoff variance until the runoff began to occur during the forecast period.

The next step of the error analysis was to find some way to separate the model and data error components.

The basic form of a regression forecast model is

$$Y = A + BX \ldots + B_n X_n + E_y \qquad (2)$$

where Y is the runoff to be forecast, $X_1, \ldots, X_n$ are precipitation or snow course observations and E_y is an error term. In this model, both model and data error are confounded in E_y. This brings up the interesting point that any partitioning of model vs data error should probably be viewed from the perspective of a particular model. This would seem to make sense because in the extreme case that a model did not recognize important data, all of the error could be regarded as model error.

Figure 2 shows an estimate of the balance between model error and data error for the Lake Powell forecast. This balance was estimated by the following procedure.

First, a simple form of the general regression forecast model was used

$$Y = A + B_1 \bar{X}_1 + B_2 \bar{X}_2 \qquad (3)$$

so that one coefficient, B_1, was used to apply to a composite precipitation variable, $\bar{X}_1$ (say the total or average of several individual gages) and a second coefficient B_2 was used for a composite snow water content variable $\bar{X}_2$. Then, the accuracy with which the "true" value of the precipitation and snow variables could be "measured" was assumed to depend on two factors:

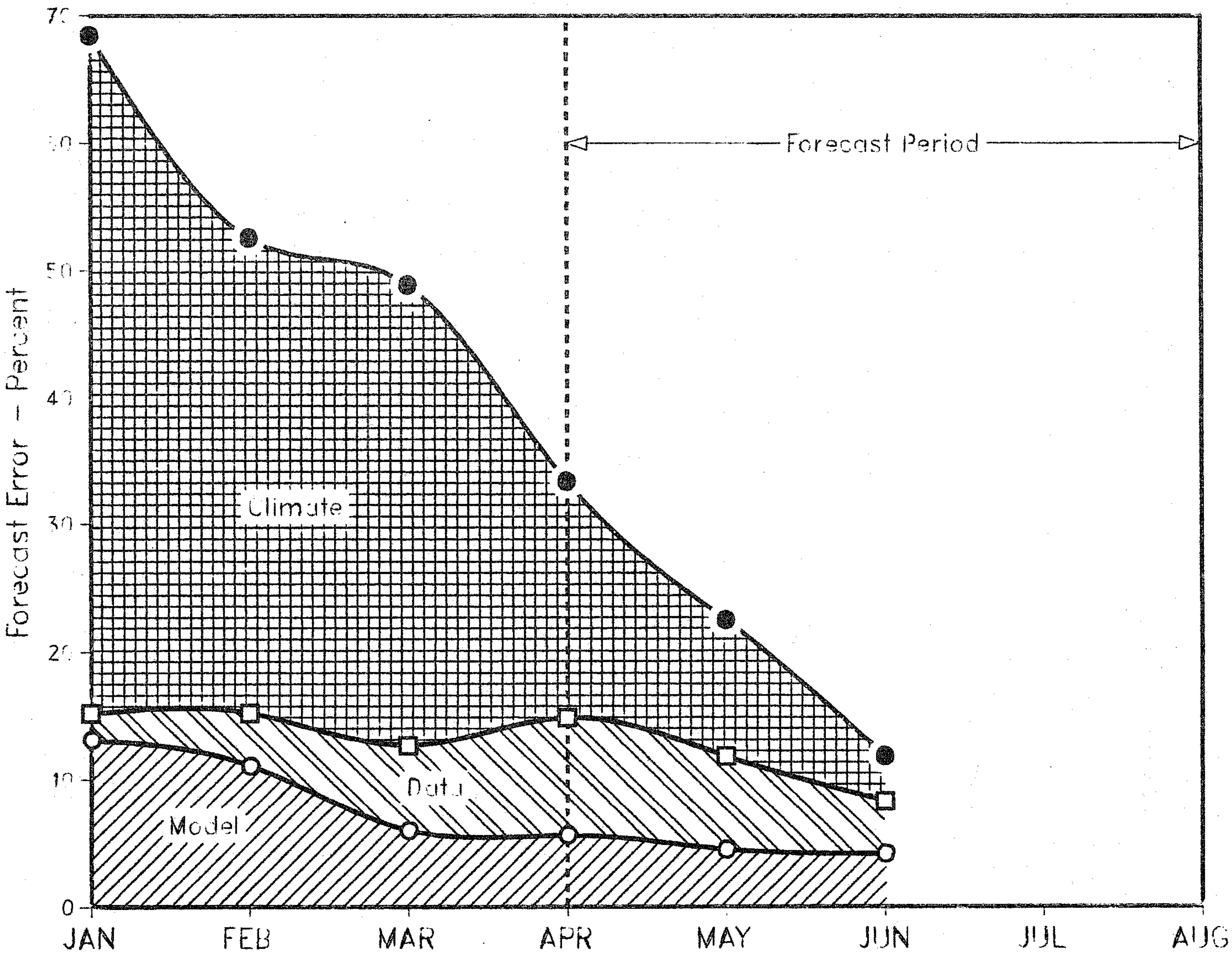
(i) the long-term climatological variance of the precipitation or snow data
(ii) the number of gages of each type used to do the forecast

The climatological variance of the precipitation, or snow data, is important because more variable phenomena measured with a given number of gages give proportionally more measurement error. To account for the influence of the number of gages on measurement error, the notion of a "number of equivalent independent gages" was introduced. The hope was that the analysis could be kept simple by working with the fundamental principle in statistics that the variance of the average of N independent observations of the same variable is equal to the variance of the observations divided by the number of observations. That is

$$S_{\bar{X}}^2 = S_X^2 / N \qquad (4)$$

FIGURE 2

Error Components for Existing System



To account for the influence of data error on forecast error, the formula

$$S_Y^2 = B_1^2 \, S_{\bar{X}_1}^2 + B_2^2 \, S_{\bar{X}_2}^2 \qquad (5)$$

was used. This formula can be derived assuming the errors in $\bar{X}_1$ and $\bar{X}_2$ are independent.

To produce the estimates shown in figure 2, it was assumed that the number of observations was equal to the number of existing gages. Also some further assumptions were made that a proportion of the variance of events in the wilderness could be measured by gages from outside the wilderness and the rest of the variance was an independent component that could only be known by placing gages in the wilderness. Of course, a central issue to the interagency study is what is that proportion. Intuitively, it would seem this proportion varies with the proportion of a basin or area that is in wilderness, and a simple relationship

$$R_w = \frac{2 \, R_q}{R_q + 1} \qquad (6)$$

was used to estimate the proportion of variance of events in wilderness, R_w, that could not be measured from outside, depending on the proportion of runoff from the wilderness, R_Q.

Equation 6 was completely assumed on the basis of intuition to apply to large areas. No data were used to support the hypothesis. It was made only in an attempt to keep the analysis simple and to avoid more complex analysis, if possible. Later, it was decided to abandon this approach and to carry-out a more complex analysis.

The model error in figure 2 was estimated simply to be the residual error remaining after deducting the estimated climate error and data error from the total forecast error. Different alternatives to add gages, including placing some gages within wilderness were analyzed to see if this approach could quantify the improved accuracy possible by reducing the data error. Also, if sufficient data were available, it would be possible to apply more physically based conceptual models to compute runoff which might further reduce the model error. Assuming that the residual model error could be reduced to one-third its present level and assuming some improvements could be made in 30-day temperature and precipitation forecasts, it appears as if the potential for improved forecasts at Lake Powell are as illustrated in figure 3. This figure suggests there may be significant potential for forecast improvement but it does not show which specific alternative would achieve that improvement.

To give some idea of how important improvements in forecast accuracy might be, Figure 4 was prepared. This figure shows how much more "manageable water" there would be as a result of a one percent reduction in forecast error variance, depending on the current percent forecast accuracy. These two curves in figure 4 apply to two different levels of confidence. As an example, the figure shows that when the forecasts are 70 percent accurate already (i.e., the error variance is 30 percent of the total runoff variance)

FIGURE 3

Potential Forecast Accuracy of April – July Runoff Volume Forecasts of Inflow to Lake Powell

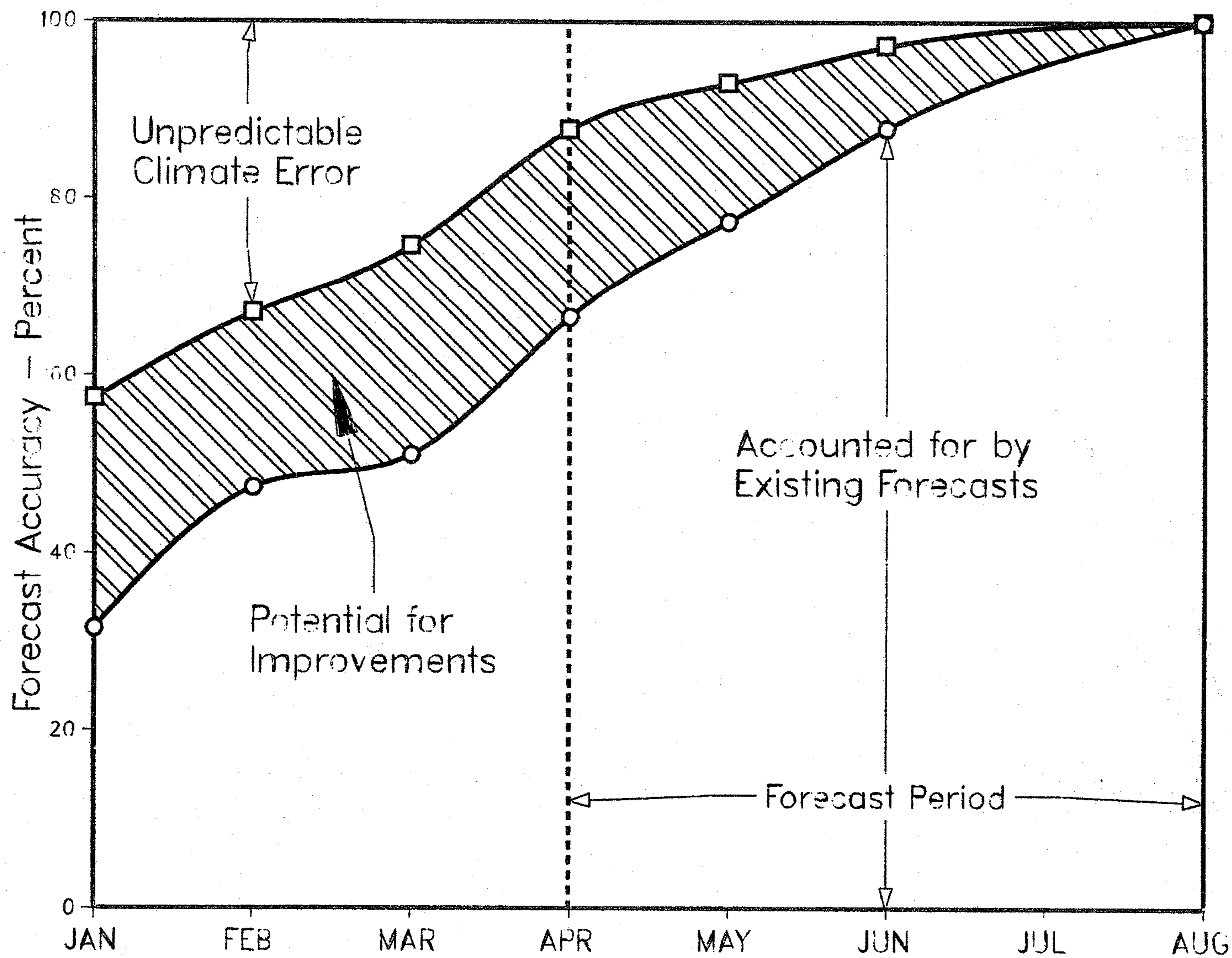
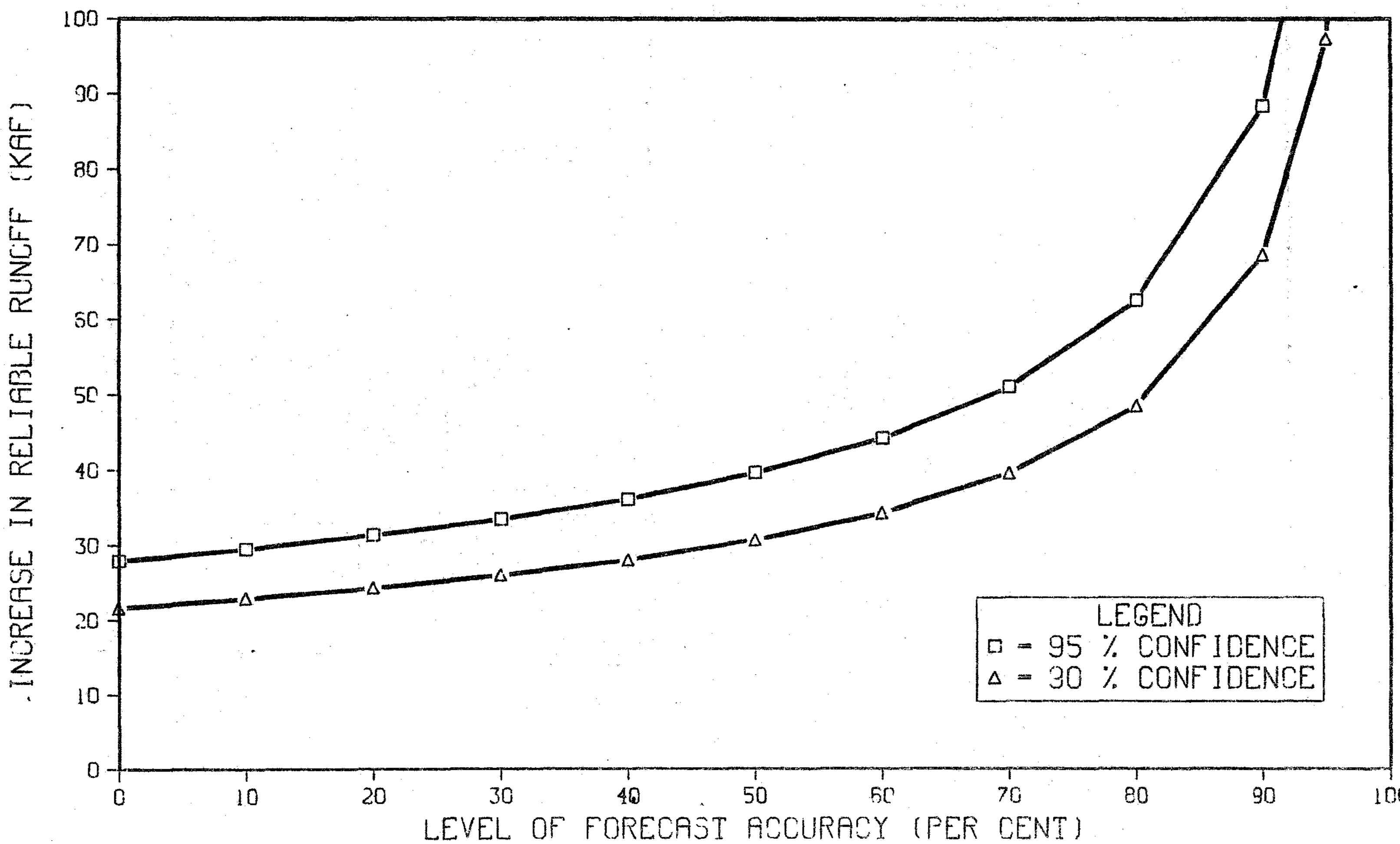


FIGURE 4



INCREASE IN RELIABILITY OF RUNOFF INTO LAKE POWELL
FOR 1 PER CENT ADDITIONAL FORECAST ACCURACY
DEPENDS ON LEVEL OF ACCURACY AND CONFIDENCE INTERVAL SELECTED

a one percent improvement would mean that an additional 50 thousand acre feet of runoff would be counted on within a 95 percent level of confidence.

PRESENT STATUS

As the working group of the interagency study considered whether this method of analysis could be used to assess alternatives, it became clear that further development of the simple model was needed to deal more systematically with the relation between the actual number of gages and the "effective number" and to deal explicitly with the correlation between events within wilderness vs those outside. Actually, these two topics are closely related because the reason the actual number of gages is not the same as the effective number is because events at the different gages are correlated with each other.

The problem of how to deal with gaging events that are highly spacially correlated has been studied extensively in non-mountainous areas (Rodriguez-Iturbe and Mehia, 1974) and the theory has been found to represent well the errors in estimates of point and mean aerial precipitation.

The spatial variability of precipitation in mountainous areas, and the influence of measurement errors on runoff can be studied by considering the simple model

$$Q = CP \qquad (7)$$

where Q is the runoff volume at a forecast point, P is an (Nxl) column vector of precipitation values at N grid points distributed throughout the basin and C is a (lxN) raw vector of coefficients that account for the contribution of each grid point to the total runoff downstream. In this particular study, values of C are being estimated using the SCS curve number technique and by calibrating each basin to find the apparent curve number that produces the observed mean April-July runoff from estimated values of October-April precipitation at each grid point.

Next, the precipitation vector P is analyzed as the sum

$$P = \overline{P} + p$$

of the mean $\overline{P}$ at each point and the deviation about the mean in any given year. Each term of the sum, P and p must be estimated from available gage data at each grid point. The analysis considers separately these estimates and takes into account the following considerations:

- (i) errors in the estimates at all grid points are intercorrelated
- (ii) spatial correlation in the occurrence of precipitation depends on a variety of locational parameters such as the horizontal distance between points, vertical elevation difference and differences in exposure, slope, etc.
- (iii) the standard deviation of point precipitation about the mean is a function of the mean
- (iv) knowledge of the mean precipitation at a gage is a function of the length of record at that gage.

The correlation and variability properties of precipitation needed to do this analysis can be estimated for existing precipitation and snow course records. The basic method of analysis simply extends familiar concepts of regression analysis that apply to a single dependent variable to the general multivariate case where there now is an independent variable at each grid point, the dependent variables are the gaged mean values and observed derivations about the means. The mathematics of this analysis are more involved than can be presented in this overview paper and will be presented in subsequent papers together with some results of the analysis of the correlation structure of the precipitation events.

REFERENCES

Mood, A. M., 1950: Introduction to the Theory of Statistics, McGraw-Hill, New York.

Rodriguez-Iturbe, I. and J.M. Mehia "The Design of Rainfall Networks in Time and Space," Water Resources Research, Vol. 10, No. 4, 1974, pp 713-728.