ACCURACY OF HYDROLOGIC FORECASTS

by

John A. Dracup[1], Dana L. Haynes[2], and Seth D. Abramson[2]

INTRODUCTION

In the first five months of each year the National Weather Service (NWS) and the Soil Conservation Service (SCS) issues runoff forecasts in a joint report entitled "Water Supply Outlook for the Western United States." Besides their use in the operation of the major river projects, the localized forecasts are utilized by irrigation districts, municipal water managers, and any other interested parties, in determining their plan of operation for the subsequent month or season. For example, in the case of the Colorado River the computed inflow forecast for Lake Powell is used as input to a Flood Control Algorithm in order to compute the release rate for Hoover Dam.

The forecasts, which are made five times a year (January-May), predict the water supply outlook for the current year or season in terms of runoff in millions of acre-feet (MAF). The forecast periods vary according to state. Precedent to an improvement in the runoff forecasting system is a refinement of the techniques used to assess current forecasting methods and to judge their accuracy. The ability to accurately check current forecasting methods will enable forecasters to select the best suited forcasting method, and to determine when and if improvements need to be made. This project evaluated the accuracy of current hydrologic forecast techniques by comparing the monthly forecasts with the recorded streamflow data. Streamflow data for 29 stations in Arizona, Utah and Colorado were compared with the corresponding forecast over a ten year period for 1973-1982.

Analyses of the correlations between forecasted and observed runoffs were made using four techniques: Spearman rank-order correlation coefficients; Product-Moment correlation coefficients; average forecast errors; and a derived statistical measure referred to as the Coefficient of Prediction. Each type of coefficient was calculated for each station for each month over a ten year period. In addition, an overall average coefficient was calculated for each state and the entire data base. The correlation coefficients were calculated with the use of the Statistical Analysis System (SAS, 1979, 1981) which is available on UCLA's IBM 3033. The other statistical measures were calculated through programs that were developed at UCLA (Haynes, 1985).

FORECAST MODELS

The NWS and the SCS use statistical and conceptual models to produce their forecasts. The statistical models are regression models that are based upon two periods of record, 1948-1977 and 1948-1967. The models correlate a number of different variables, such as precipitation, snowpack, water equivalent, and carry over influence from the previous year's runoff, with the

---

Paper presented Western Snow Conference 1985.

1. Professor, Civil Engineering Department, University of California, Los Angeles, CA

2. Graduate Student, Civil Engineering Department, University of California, Los Angeles, CA

predicted runoff for the forecasted period. The variable with the best correlation in the period of record is first selected. Added to this variable is the next variable that improves the derived prediction equation. The process is repeated until all of the statistically significant variables are taken into account. Regression models are developed for each of the forecast points. The Regression models use the current conditions as of the forecast date, and then assume normal weather conditions from the forecast date through the end of the forecast period.

The conceptual models integrate observed precipitation and temperature data into a combination of mathematical equations and decision trees that simulate the runoff process. The models are calibrated using observed data for a 20 year period (1948-1967). These models are run with the current site conditions, and produce 20 year runoff sequences with the probability of each occurrence indicated.

DESCRIPTION OF STATISTICAL MEASURES

Spearman Coefficients

Wallis (1977) presented the use of Spearman correlation coefficients in a National Academy of Science (NAS) analysis of the accuracy of long range weather forecasts. The NAS analysis, which determined the accuracy of forecasted monthly precipitation amounts, showed the usefulness of the Spearman coefficient in comparing forecasted and observed values. Based upon the NAS analysis, this project applied the Spearman coefficient to runoff forecasts.

The Spearman correlation coefficient is one of two types of rank correlation coefficients, the other being the Kendall correlation coefficient. Spearman's coefficient of rank correlation (ρ_s) is given by,

$$\rho_s = 1 - \frac{6 \Sigma d^2}{n(n^2-1)} \qquad [1]$$

where

Σd^2 is the sum of squares of the differences between corresponding pairs of ranks

n is the number of pairs of rank of x and y.

Product-Moment Coefficient

The Product-Moment coefficient of correlation, better known simply as the correlation coefficient, is the most common type of correlation determination method. This coefficient (ρ_p) is given by the expression,

$$\rho_p = \frac{\text{cov}_{(x,y)}}{\sigma_x^2 \sigma_y^2} = \frac{\Sigma (y-y)(x-x)}{\Sigma (y-y)^2 \ \Sigma (x-x)^2} \qquad [2]$$

where

σ_x^2 = variance of x

σ_y^2 = variance of y

Both the Spearman and Product-Moment correlation coefficients can range from -1.0 (perfect disagreement) to 1.0 (perfect agreement).

Coefficient of Prediction

In a recent paper Lettenmaier (1984) stated that "Past model comparison studies have been flawed by improper model verification and, in some cases, by the use of misleading statistics such as the correlation between recorded and forecasted runoff." Furthermore, Lettenmaier (1984) stated that "Methods of assessing forecast accuracy are not very advanced, ... and that claims for improvement in forecast accuracy realizable from application of particular models and related forecast model comparison studies, have been based on erroneous assumptions."

Lettenmaier (1984) proposed that if the forecasts are statistically unbiased then a more appropriate forecast error measure is the expected value of the squared forecast error, or alternately known as the Coefficient of Prediction (C_p). The Coefficient of Prediction was first discussed by Tangborn (1977), and is defined as,

$$C_p = 1 - \frac{E(e(t_i,t_f)^2)}{Var(Y_r(t_i,t_f))} \qquad [3]$$

where Y_r is the recorded runoff, $Var(Y_r(t_i,t_f))$ is the variance of the recorded runoff in the period (t_i,t_f) and $E(e(t_i,t_f)^2)$ is the expected value of the squared forecast error.

If it is assumed that the forecasts are unbiased, it is necessary to only consider the ratio of the standard deviation of the forecasted and observed sequences, since by definition the expectations are identical. Letting σ_f^2 be the variance of Y_r (forecasted runoff) and σ_r^2 be the variance of Y_r, with $\gamma = \sigma_f/\sigma_r$ then

$$C_p = 1 - (\sigma_f^2 + \sigma_r^2 - 2\sigma_f\sigma_r\rho_{fr})/\sigma_r^2 = \gamma(2\rho_{fr} - \gamma) \qquad [4]$$

where ρ_{fr} is the correlation between the forecasted and recorded runoff. (Two sets of Coefficient of Predictions were generated. One used the Spearman coefficient (C_{ps}) and one used the Product-Moment coefficient (C_{pp}).)

The maximum value of C_p, corresponding to a perfect forecast, is 1.0; a value of zero results for forecasts always equal to the long term mean. C_p is negative for forecasts less accurate than the long term mean.

DATA

Twenty-nine streamflow stations in Arizona, Utah, and Colorado were used in the analysis. Each particular station was selected because of its location on a direct or indirect tributary of the Colorado River. The selection of streamflow stations was limited by the existence of verifiable hydrological data which corresponded exactly with the stations used in the forecasts. Because different agencies prepare the forecasts and collect the recorded data, many stations which had ten years of forecasts did not have ten years of recorded hydrological data. The inverse was also true.

Another problem that was encountered in ensuring that forecasted and recorded stations corresponded exactly, was when despite the stations corresponding by name and location, an obvious discrepancy existed between the forecasted and recorded data. For example, when the recorded data was much lower than the forecasts in every case, it was assumed that the forecasts did not consider a diversion for irrigation, and/or some other use, upstream from the streamflow station, and therefore this station was not used in the analysis. An effort was made to ensure that the observed and forecasted figures were based on the same premise. To expedite future comparison studies, and to give researchers an opportunity to readily gauge the accuracy of forecasts, it will become essential that the United States Geological Survey (USGS), which prepares the recorded data, and the NWS/SCS, which prepares the forecasts, coordinate their efforts to ensure that identical comparisons can be made.

The forecast periods that were used varied according to state and location. In Utah and Colorado most of the stations had forecast periods that included January through September. Arid regions, such as Arizona, had forecast periods that only included through July, since almost all of the Spring runoff takes place by this time. Previous to the 1979-1980 water year the forecast periods were shortened in each successive forecast. For example, the January forecast would include a runoff forecast for the period January to September, the February forecast would include a runoff forecast for the period February to September, and so on. After 1980 the forecast periods remained fixed at January to September.

RESULTS

A comparison of the Spearman and Product-Moment correlation coefficients between states revealed some expected and noticeable trends. Arid regions, such as Arizona and parts of Utah, which experience more variability in weather conditions, and where snowmelt is less of a factor in runoff magnitudes, did not have as accurate forecasts as compared to Colorado and northern Utah. Figures 1 - 3 which show the spread of forecast/recorded comparisons about the line of "Perfect Forecast" confirm this observation. Figure 3 (Colorado) shows much more of a clustering of the forecasts about the line of "Perfect Forecast" as compared to Figure 2 (Utah) and much more so than Figure 1 (Arizona).

While Spearman correlation coefficients of close to 0.80, and Product-Moment correlation coefficients of close to 0.90 were readily achieved after February of each year for Colorado; Arizona's data indicated Spearman and Product-Moment coefficients only in the range of 0.70. Utah's forecast accuracy roughly fell in between that of Arizona and Colorado. The same trend discussed above was followed by the Coefficients of Prediction.

In general, for large n's (pairs of x and y), and unbiased data, the Spearman and Product-Moment correlation coefficients can be very similar. In this project which used ten (10) years of data, and 5 months per year, (n=50) the Spearman and Product-Moment coefficients differed noticeably. The differences may be due to an 'n' that was not large enough to remove any inherent variations between the Spearman and Product-Moment correlation coefficients, and to some possible biases in the data. Besides the difference in magnitude, the Spearman and Product-Moment coefficients also had different trends in a few cases. Figure 4 plots the average monthly correlation coefficients (Spearman and Product-Moment) for each state. This figure shows that while the coefficients derived from the Colorado data followed very similar trends, the different coefficients for Utah and Arizona do not show trends that match quite as well. Therefore, it can be concluded that the Spearman and Product Moment coefficients are equally suited correlation coefficients for this type of analysis. The average monthly Coefficients of Prediction for each state showed a similar pattern as to that of the correlation coefficients. (See Figure 5)

To determine if a bias existed in the forecast procedures a program was developed to evaluate the forecast errors. Overall, 56.8% of the forecasts were higher than the recorded data. This indicates an apparent balance in the forecasts. In Arizona the forecasts were high 41% of the time, in Utah, 69% of the time, and in Colorado, 57% of the time.

The average forecast error for all months for stations in Arizona was 57.3%, for stations in Utah it was 24.8%, and for stations in Colorado it was 22.1%. The average forecast error was calculated by dividing the absolute value of the difference (observed-forecast) by the average runoff. All five months of forecasts were grouped together to determine the overall average forefast error.

To show the relationship between the different statistical measures that have been previously discussed, two graphs were prepared. Figure 6 is a comparison by station of the overall Spearman and Product-Moment correlation coefficients and the average forecast errors.

Figure 6 graphically illustrates the inverse relationship between the correlation coefficients and the forecast errors, which is that the forecast error increases when the correlation coefficient decreases. This was generally true for most stations. So while the correlation coefficients do not directly measure the magnitude of the forecast errors, forecasts that accurately predict trends generally have greater accuracy in terms of magnitude. Figure 6 also shows how well the two different correlation coefficients match except for the first 9 stations which were in Arizona.

SUMMARY

As a result of the analysis, several conclusions can be drawn about the effectiveness of the water supply predictions. These conclusions can be organized according to the four questions that are stated below:

1. Are the forecasts of value ?

The NWS/SCS has developed somewhat surprising accuracy in the preparation of their runoff forecasts. The analysis in this project showed that even as early as January of each year worthwhile forecasts are being developed. Coefficients of Correlation between the observed and forecasted runoffs in the range of .70 to .80 are being achieved as early as February and March of each year. The Coefficients of Prediction also indicate that forecasts much more accurate than the long term mean are being generated.

2. Does the accuracy of the forecast improve significantly over time ?

As expected, the accuracy of the forecast does improve significantly over time. This is due to the inclusion of increasingly more information on hydrologic parameters affecting runoff as the year progresses.

3. What statistical measure is best suited for assessing forecast accuracy ?

The results of this analysis do not agree with the claim made by Lettenmaier (1984) that the correlation coefficient is a false measure of forecast accuracy. In fact, the Coefficients of Prediction and the correlation coefficients compare favorably when used to judge which forecast stations had relatively higher forecast accuracy. The correlation coefficients and the Coefficients of Prediction also each compare well with the average forecast errors, and verify that the two sets of measure are inversely related. The coefficients increase when the average forecast error decreases.

4. How does the forecasts' accuracy compare between different stations, and what factors explain the difference ?

A comparison between stations of the different statistical measures revealed some expected and noticeable trends. Arid regions such as Arizona and southern Utah which experience more variability in weather conditions, and where snowmelt is less of a factor in runoff magnitudes, did not have as accurate of forecasts as compared to Colorado and northern Utah. Of the two components of runoff, snowmelt and precipitation, the largest uncertainty is due to the unpredictability of the remaining seasonal precipitation. Therefore, watersheds where snowmelt constitutes a larger percentage of the runoff will have more accurate forecasts. Other factors which explain the differences in forecast accuracy between different stations include the amount of accurate hydrologic data available for that watershed and how well the models have been calibrated to fit that particular watershed.

Together, the correlation coefficients or Coefficients of Prediction, along with the average forecast error can be used to give a worthwhile assessment of a station's forecasting performance. Future research should be geared toward the development of possibly an overall weighted coefficient of forecasting accuracy which takes into consideration the correlation coefficients to measure trends, the average forecast error to measure differences, and the Coefficient of Prediction. Finally, to facilitate future verification studies which look at the effectiveness of the water supply forecasts, the NWS/SCS and the USGS need to coordinate their efforts to ensure that identical comparisons can be made. Forecasting stations and observation points should be close together as possible, and the runoffs measured should be based on the same criteria.

REFERENCES

Lettenmaier, Dennis P. "Some Issues in Assessing the Accuracy of Hydrologic Forecasts," Draft 1984.

Haynes, Dana L., Unpublished M.S. Thesis, UCLA, School of Engineering and Applied Science, Civil Engineering Department, March, 1985.

National Weather Service and National Ocean and Atmospheric Administration, "Water Supply Outlook for the Western United States," Silver Springs, Maryland, 1973-1982.

SAS Users Guide, SAS Institute, Cary, North Carolina, 1979.

SAS/GRAPH Users Guide, SAS Institute, Cary, North Carolina, 1981.

Tangborn, Wendel W., "Application of a New Hydrometeorological Streamflow Prediction Model," Proceedings of the 45th Annual Western Snow Conference, Albuquerque, New Mexico, 1977, pp. 35-43.

Wallis, James, Climate, Climatic Change and Water Supply, Washington, D.C.: National Academy of Sciences, 1977.

FIG (1) STREAMFLOW FORECAST VS. OBSERVED 1973–1982

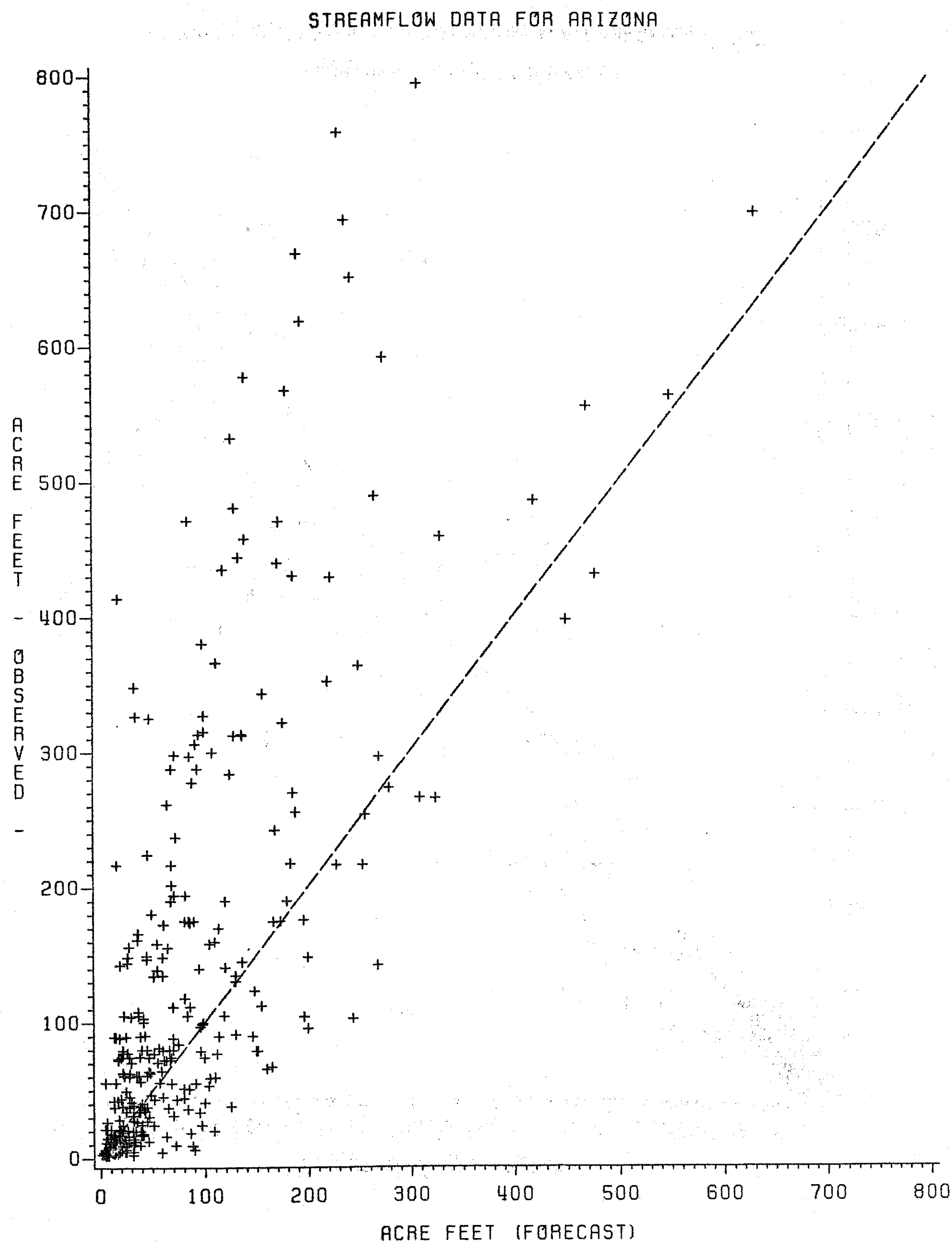


FIG (2) STREAMFLOW FORECAST VS. OBSERVED 1973-1982

STREAMFLOW DATA FOR UTAH

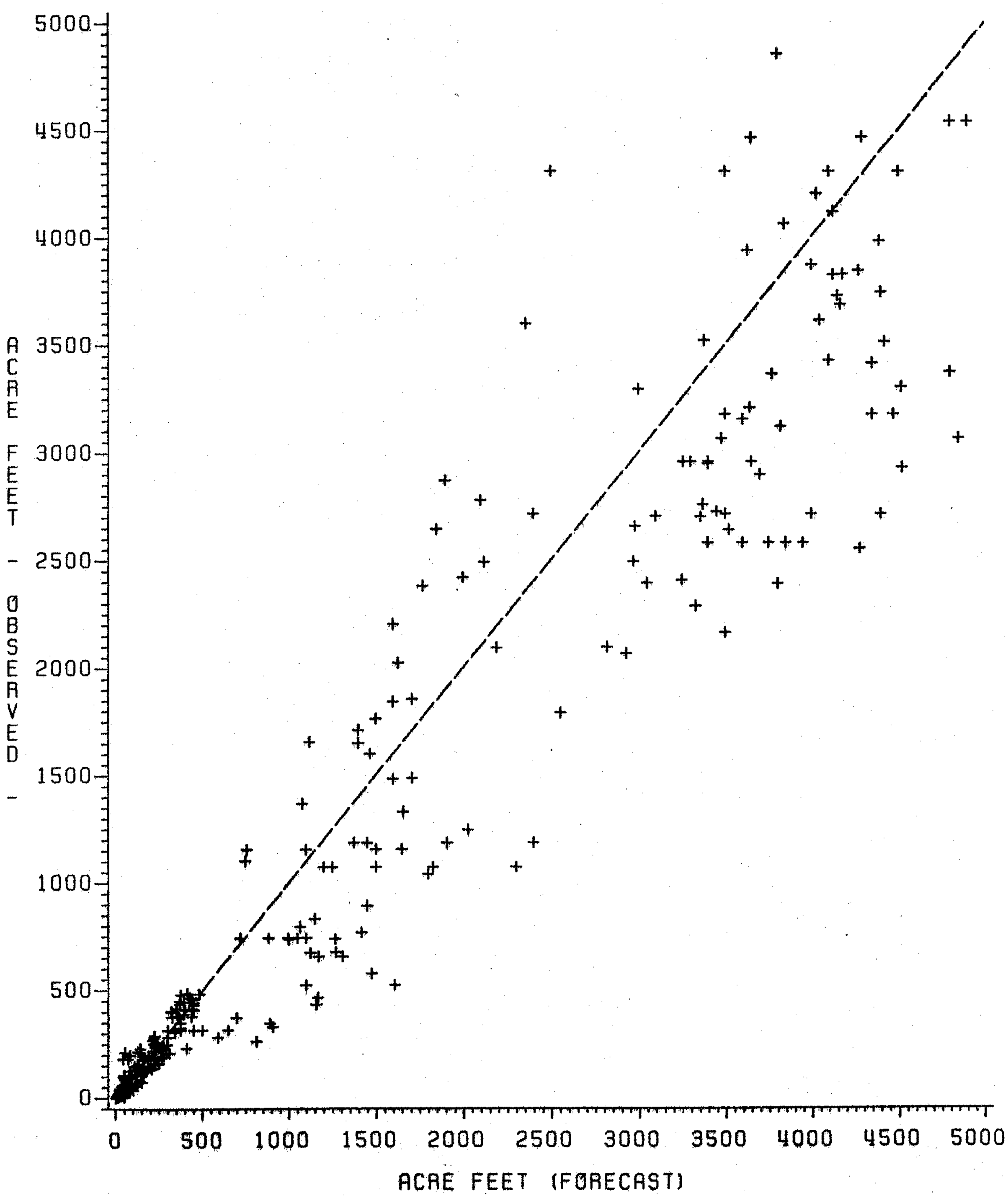


FIG (3) STREAMFLOW FORECAST VS. OBSERVED 1973–1982

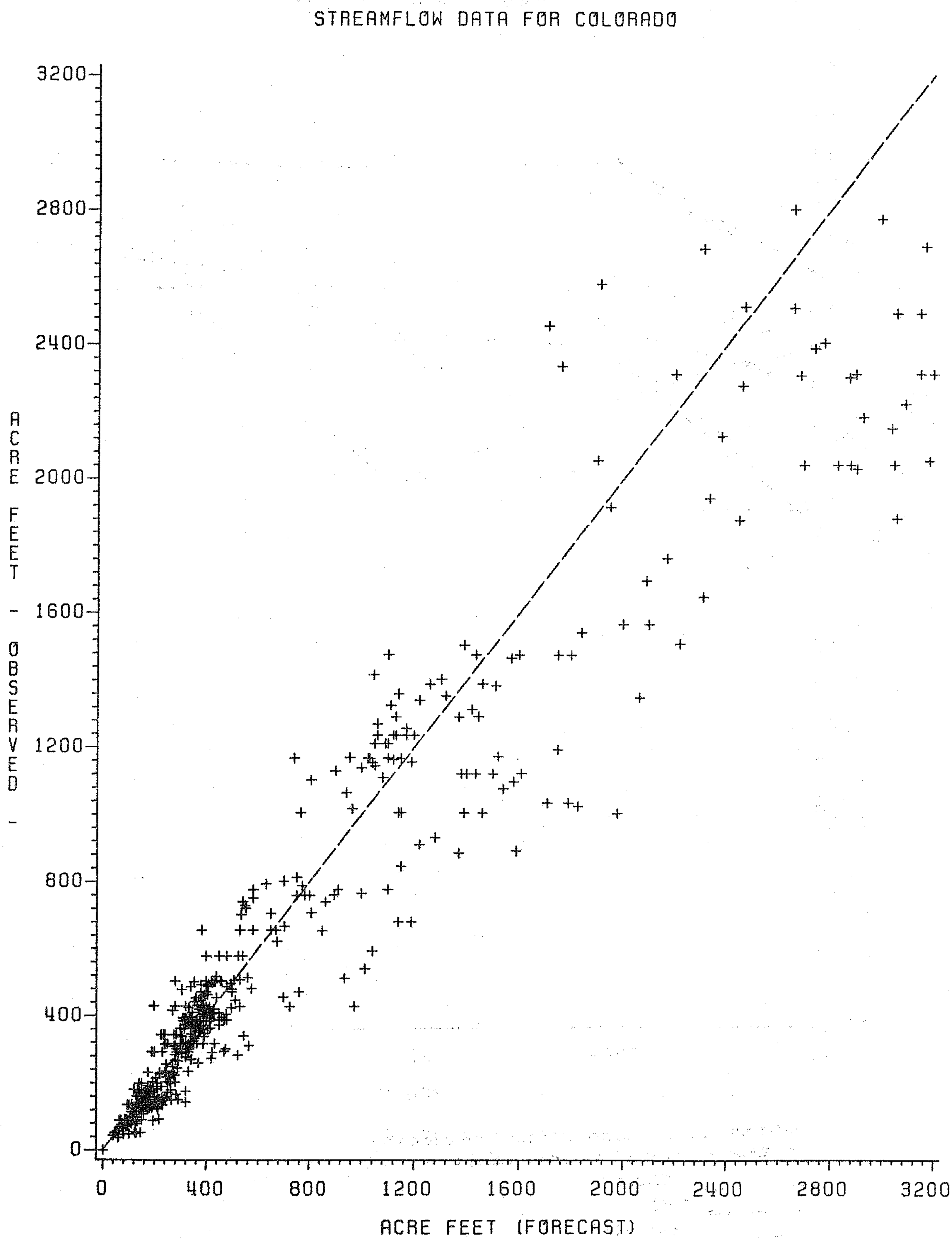
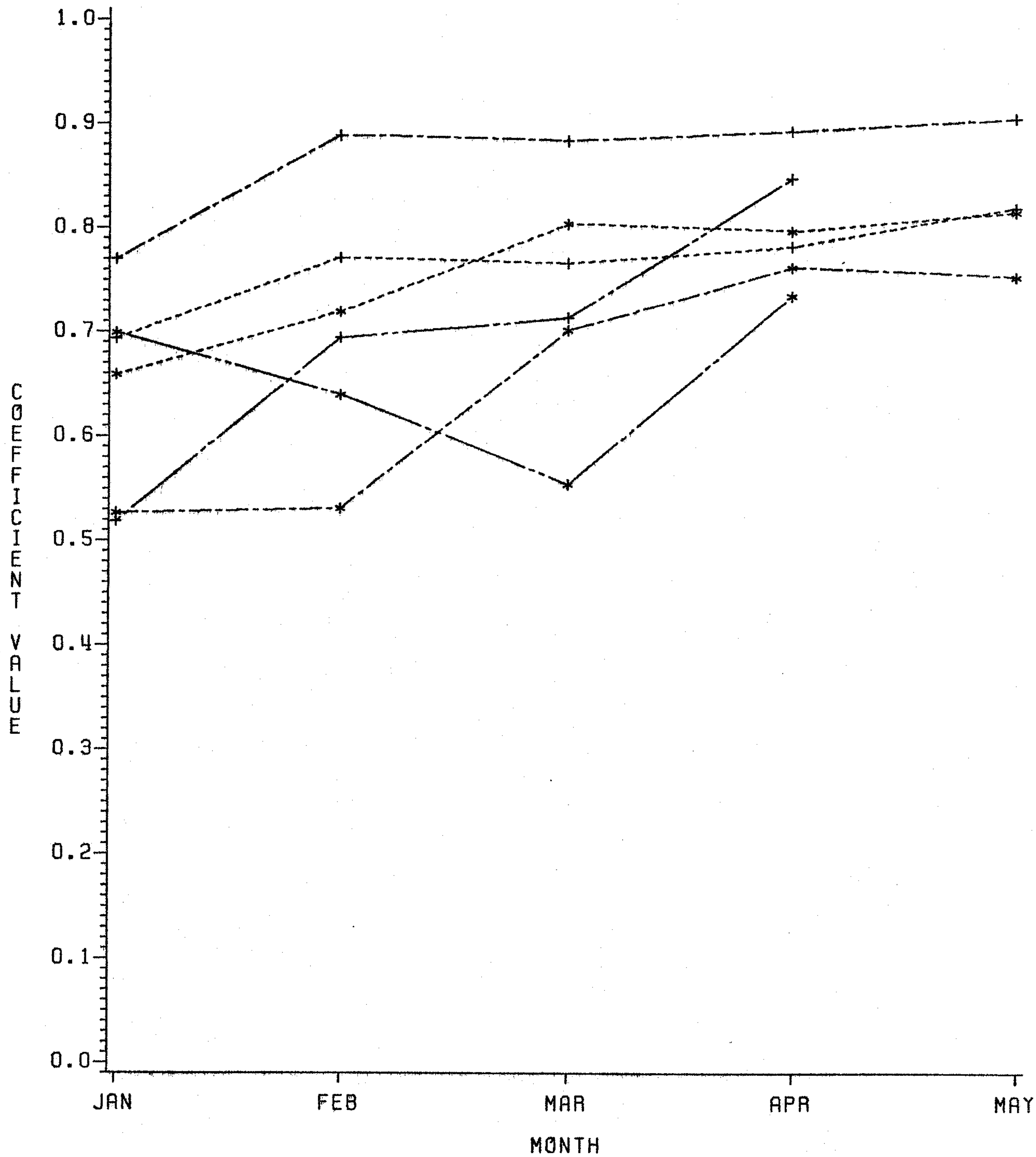


FIG (4) STREAMFLOW FORECAST VS. OBSERVED 1973-1982

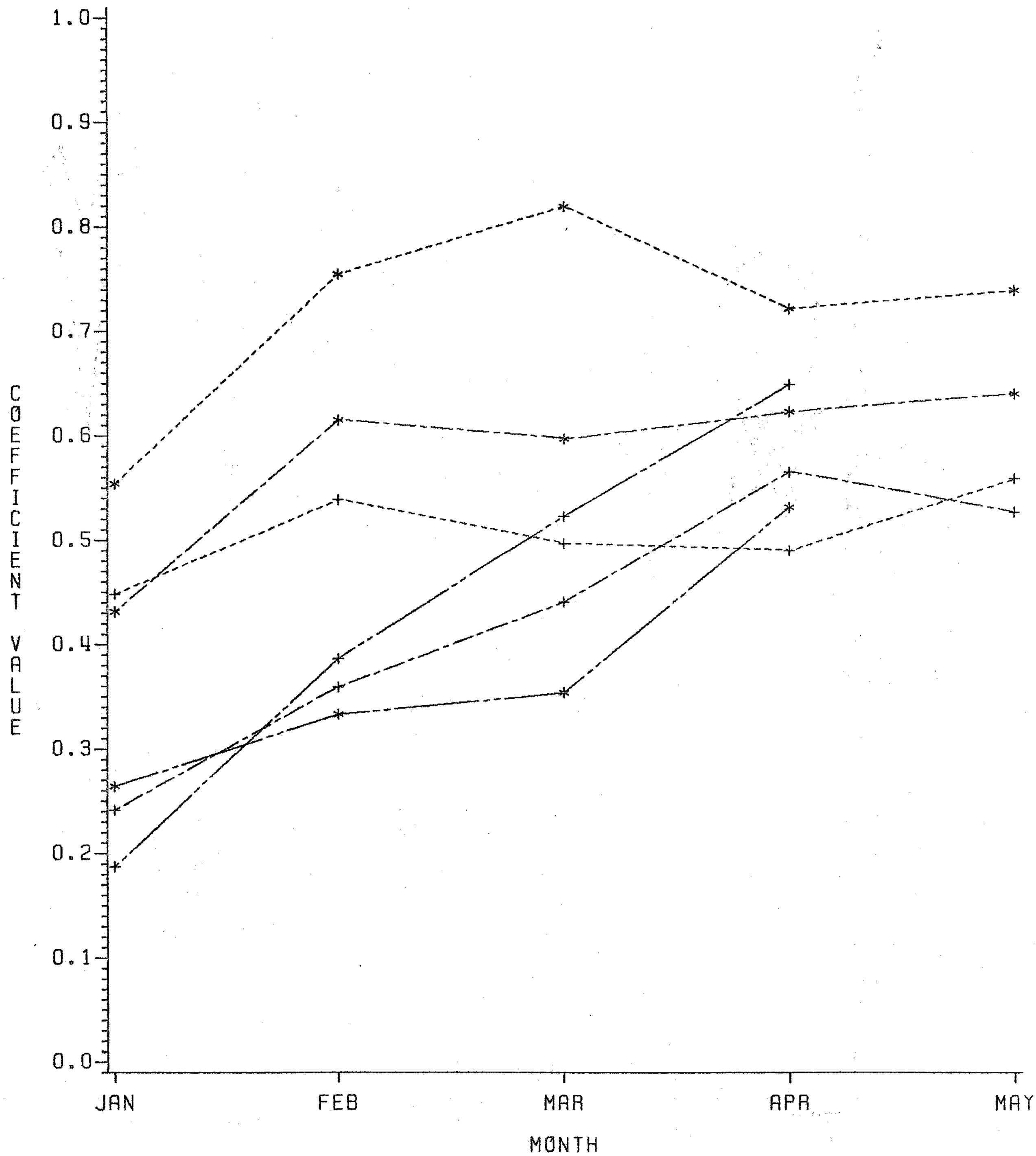
CORRELATION COEFFICIENTS



* - Product-Moment Correlation Coefficient
\+ - Spearman Correlation Coefficient
—— — — —— Arizona
—— — —— Utah
— — — — — Colorado

FIG (5) STREAMFLOW FORECAST VS. OBSERVED 1973-1982

COEFFICIENTS OF PREDICTION



* - C_{pp}

+ - C_{ps}

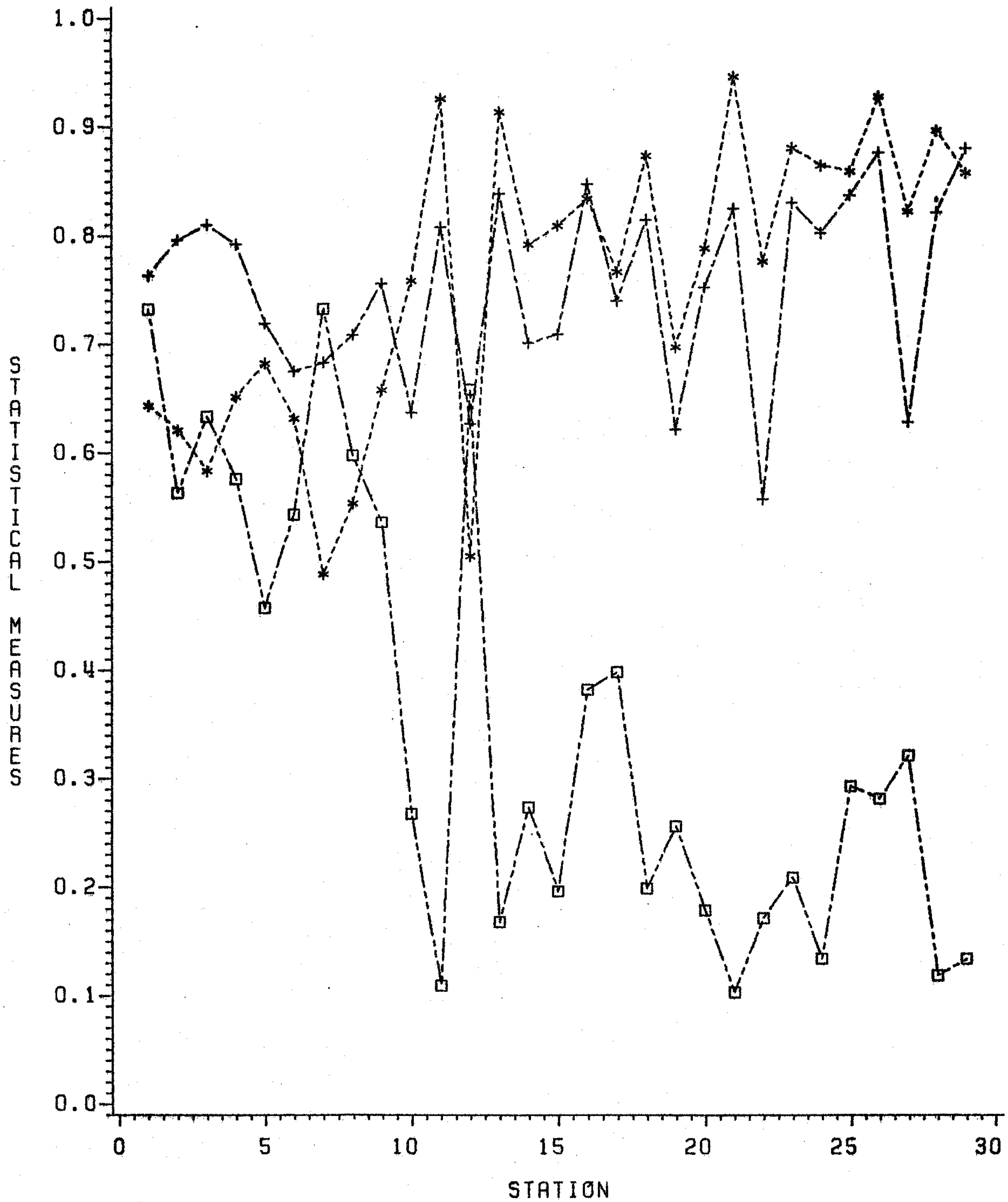
——— - - ——— Arizona

——— - ——— Utah

- - - - - - - Colorado

FIG (6) STREAMFLOW FORECAST VS. OBSERVED 1973-1982

COMPARISON OF DIFFERENT STATISTICAL MEASURES



* - Spearman Correlation Coefficient

\+ - Product-Moment Correlation Coefficient

□ - Average Forecast Error (%)