OPERATIONAL FORECASTS OF RUNOFF FOR THE UPPER AMERICAN RIVER PROJECT 1/

By

Bernard S. Clorety 2/ and Jack F. Hannaford 3/

Introduction and Project Description

This paper describes the development of a continuously updated water supply forecast system designed to be used for operational decisions on Sacramento Municipal Utility District's Upper American River Project.

The Upper American River Project is located on the Middle and South Forks of the American River in El Dorado County, California. The project was designed as a single purpose power project. Principle development is in the Silver Creek drainage basin which totals about 180 square miles. Diversions into Silver Creek are made from approximately 85 square miles of the Rubicon River, a tributary to the Middle Fork of the American. In addition, one major powerhouse is located below the confluence of Silver Creek and the South Fork of the American River adding about 320 square miles to the over-all project area. The project area is outlined in Figure 1.

The District's Project consists of three major storage reservoirs with a combined capacity of 393,000 acre feet, and numerous smaller reservoirs for diversion and regulation purposes. Five powerhouses with a combined capacity of 500,000 kilowatts are located throughout the system.

Precipitation in the area averages from 40 to 60 inches. Elevations range from 1,500 to approximately 10,000 feet, so that a major portion of the annual water supply passes through storage as snowpack.

The District currently produces only a portion of the power required for distribution. As a consequence certain power commitments must be made very early in the water year and the economic consequences of failing to meet these power requirements pose a critical problem in operation of the project.

Snowpack storage and early-season power commitments combine to make long range water-supply forecasts both necessary and feasible for this project.

Forecast Requirements

Prior to initiation of this study, the District set up a series of criteria pertinent to development of long-range forecasting procedures. Conditions encountered dictated only a few changes in these criteria as the development progressed.

First, it was required that long range forecasts could be made from any date in the water year through June 30th. In operation, the procedure might be updated after each storm or after periods of excessive snowmelt or snow accumulation. The requirements for daily operation dictated not only the type of procedure to be used in analysis, but also the basic data which could reasonably be used in preparing the actual forecast.

Second, the basic forecast was to be prepared using median conditions of precipitation, temperature and runoff after the date of forecast, but additional forecasts were to be prepared for runoff which could be expected to be exceeded at specific probability levels.

1/ Presented at Western Snow Conference, Lake Tahoe, Nevada, April 16 - 18, 1968.

2/ Sacramento Municipal Utility District, Sacramento, California

3/ Sierra Hydrotech, Placerville, California.

Third, forecasting procedures were to be prepared for seven watersheds. With one exception, forecast procedures were based upon estimated full natural flow of the streams. Pre-existing storage and diversion facilities in the excepted watershed had established a long-term pattern consistent with available water supply, so measured flow was to be used in preparation of a forecast procedure.

Data Analysis

Although there are a number of precipitation stations with daily records in the lower elevation of the Sierra near the latitude of the Upper American River Project, many of these records have missing data or inconsistencies in measurement during the period proposed for analysis. The period of 1926 through 1967 was used in preliminary analysis of precipitation records. First step in analysis was to double-mass plot annual precipitation at individual stations against a representative group of stations. Approximately 75% of the potential stations in the area were eliminated as a result of either insufficient length of record or inconsistencies during the period of record which could not be corrected. After analysis, eleven precipitation stations were selected which appeared representative of the Upper American River Project area. These stations were tested singly and in groups to determine which stations had the greatest likelihood of producing a reliable forecast procedure. In final analysis it was determined that seven of the stations provided the best description of conditions in the project area. One of these stations was normally unavailable on a daily basis and often unavailable on a monthly basis for operational work. It was eliminated from the analysis with apparently little effect on the over-all forecast procedure.

There was little choice in selection of snow courses to use in analysis of the Upper American River Project. Within recent years, a number of courses have been measured within the project area, but long-term records dating back to 1930 are only available at courses outside the project area. Nine snow courses with varying lengths of record located both within and adjacent to the project area were selected for procedure development. A snowpack index was developed using all nine courses. In addition, a low elevation snow course index was developed using the four lowest snow courses, in order to provide additional information on the extent and quantity of snow cover in the District's lower watersheds.

Runoff data proved one of the biggest pitfalls of the whole project. Reasonably consistent runoff records have been maintained for Silver Creek at Union Valley and South Fork of Silver Creek near Ice House from 1925 through the date when project construction began. During the construction years, estimated data are subject to considerable question. At the present time it is believed that reliable records are being obtained at the Ice House site. At Union Valley, imports to the reservoir and estimates of reservoir releases from powerhouse generation have made some of the record questionable, but it is felt that the record is reasonably good during the last few years.

On the South Fork of the American River near Camino, the forecast area was to include the entire drainage of the basin with the exception of the Silver Creek drainage. Flows had previously been estimated using the process of subtraction resulting in some discrepancies in the over-all figures. Analysis of this record was made using flows from an upstream tributary near the Camino station which included a major portion of the watershed above Camino. Although early season records were not corrected as a result of this analysis, runoff during the snowmelt period was corrected where inconsistencies were apparent between South Fork of the American River at Kyburz and South Fork of the American River near Camino. Errors in runoff which could be tolerated on Febrary 1st were far too large to appear in a May 1st or June 1st forecast. On the South Fork, the Pacific Gas and Electric Company operates several small reservoirs, a power canal and powerhouse, as well as a small out-of-basin diversion. Discussion with PG&E personnel indicated that with few exceptions, the operation of their project on the South Fork has been consistent with available water supply since before 1930. Corrections were made for exceptions as required, and operational activities were included as part of the over-all probability in analysis. Although the final analysis probably represents a wider probability range than would be expected for an analysis based on full natural flow, future use of the forecasting procedure will not require adjustment for normal PG&E operations. Any abnormal operating procedure would require adjustment in the South Fork forecast procedure.

Runoff records for remaining project basins were estimated for the period of analysis by subtraction of available records or from very sketchy records predating the period of analysis. Records collected since the project went into operation indicate that the preliminary estimates were apparently quite good.

Part of the data analysis included development of a means for estimating runoff on a monthly or daily basis for watersheds where data might not be available on an operational basis. Estimates were prepared using data from a station or groups of stations where the data would most likely be available on a short term operational basis. These studies have been used in analyzing runoff for periods where data are questionable as a result of missing record or project construction influences.

Approach

The approach to development of long-term forecast procedures for the Upper American River Basin was to develop an analysis for a single basin on which all data, or a major portion of the data, were reliable and available. Silver Creek at Union Valley was selected as the initial basin.

Analysis began with a precipitation-based procedure which started operation on October 1 and continued through June 1, forecasting runoff from the date of forecast through June 30, July 31 and September 30. Forty-two years of record, 1926-1967 were used in analysis. Data were analyzed on a monthly basis so that nine monthly forecasts were made for each water year. Interpolation between monthly values was made on a smooth curve basis to permit preparation of daily forecasts.

A computerized multiple linear regression program was used to develop a forecasting relationship between precipitation, runoff to date and carry-over from the previous year. Correction for non-linearity and interdependency of data was made as required. All precipitation data available for the entire year, September through June, were used in this analysis, regardless of the forecast date. The effectiveness of precipitation was weighted seasonally according to the results of basic data analysis.

After the final forecast procedure was developed, the procedure and the basic data were put back into another computer program which developed forecasts as of the first of each month for the entire period of record. Forecasts were developed using only that data which would have been available at the time of forecast. Forecast error was arrayed and centered around a median probability. Computed probabilities based on the forty-two year record were then plotted on probability paper in order to smooth the effects of the basic data. Probabilities at each of the required specific levels were then plotted on an error versus time scale (Figure 2) to create a funnel type diagram. This diagram delineates the probability of forecast error resulting from future weather events after the date of forecast based on historic record. Procedure error is included also, but it is a relatively small factor early in the season. By May 1, procedure error becomes significant in operational forecasts. For operational purposes, daily forecast probabilities are interpolated on a smooth curve between first-of-month figures. The resulting forecast system provided an operational forecast procedure and probabilities from October 1 through June 1, and was used as a basic standard to evaluate improvements in forecasting procedure resulting from analysis of additional forecast parameters.

Since the first snow surveys are made on February 1st, development of a forecast procedure using snow data was initiated as of the February 1st forecast date. The procedure for development was similar to the procedure for precipitation in that the final analysis included the development of a funnel diagram from February 1 through the end of the forecast period. Updating was to be done on a daily basis through the use of precipitation and runoff data. Monthly updating was accomplished as new snow surveys were available to put into the forecasting scheme at the first of each month. A funnel diagram similar to that plotted for the precipitation scheme was developed for the snow forecasting procedure. (Figure 2)

Three important items were noted. First, there was no substantial increase in forecast accuracy as a result of including snow in the forecasting procedure until about April 1. Even though the forecasting procedure using snow may possibly be more accurate after all data are available, the independent error introduced by variation in precipitation after the date of forecast far exceeds and masks the error in the procedure itself.

In addition, only limited snow data have been historically collected in this area before April 1, this complicating development of probabilities. Second, it was noted that except in two cases during the forty-two year period, the sign of the forecast error was the same for either the snow or the precipitation procedure. In those two cases, although the sign was different the magnitude of the difference was small. Third, as the season progresses, the procedure using snow closes much faster and to a smaller final error than the precipitation procedure. (Figure 2)

As a result of this analysis it was decided to use the precipitation scheme until March and then convert to the procedure using snowpack. During the March - April period weight was to be gradually transferred from one procedure to the other, dependent upon (1) the date, and (2) the number of snow courses which had been measured as of that date. By April 10th full weight was placed on the snow procedure even though the snow courses might not all be available. An attempt was made to develop an analysis procedure using the May 1st snow measurements. As in the case of the February and March snow measurements, there is little historic data available in the local area. Although procedures were developed that give some promise of improving the late season forecasts, analysis of probabilities became impossible as a result of missing data. As a consequence, the May 1st snow measurements were not used in the final analysis. It should by no means be concluded from this that it would be impossible to use late season snow measurements to increase the accuracy of late season forecasts. In this particular case, loss of adequate probability range outweighed any probable increase in late season forecasting accuracy that might have been obtained.

The original probability requirement was for seven probability levels including the median (i.e., that forecast which will be exceeded 5%, 10%, 25%, 50%, 75%, 90%, or 95% of the time) and for extremes of record. With approximately 40 years of record, the extremes of record would represent approximately the 2% level and the 98% level. At these levels and with this short period of record, the range should be viewed in terms of the extremes of record rather than actual 2% and 98% levels. Probabilities may be applied to economic decisions, knowing that there is a 50 per cent chance that runoff will exceed (or fall below) the median forecast, a 90 per cent chance runoff will exceed the 90% forecast, and an 80 per cent chance that runoff will fall between the 10% and 90% forecasts.

Operational Forecast

The operational forecast procedures have been prepared for use on the District's IBM 360 computer system. Input is by card. Flexibility of input data was of primary concern in preparation of the operational forecasting program.

Following is a brief description of how the operational forecast system works.

Data Read-In and Analysis. All data are read in on cards which may be updated as required.

1. Forecast date is read in.

2. Runoff data.

   a. All available runoff is read in. Monthly values are used up through the last complete month. Daily values may be read up to 45 days after last complete month.

   b. Missing data at key stations are estimated by interpolation or extrapolation.

   c. Missing data at remaining stations are estimated from key stations.

3. Precipitation data. All available precipitation data are read in. Monthly and daily values are read as with runoff above, but:

   a. A missing daily card indicates no precipitation on that date.

   b. Missing data at any station for a day or month on which precipitation occurred at other stations flags that data so that it may be acknowledged later in the computations.

4. Snowpack water content data.

   a. All available snow data are read in.

   b. Data adjusted to first of month through precipitation occurring between the date of measurement and the first of month -- or probability of precipitation during that period, whichever is appropriate.

   c. High and low elevation snow indexes are computed and adjusted, depending upon the number of surveys completed on date of forecast.

Data Output. All data read in are then listed, either as read in or as estimated by the program, to provide for visual check of input. (Figure 3.)

Forecast Computation.

1. Prepares a date-of-forecast through July 31 forecast for first basin at median level.

2. Prepares eight remaining levels of forecast (2%, 5%, 10%, 25%, 75%, 90%, 95%, 98%) based upon median level.

3. Prepares date-of-forecast through June 30 forecast at median level, using July 31 forecast.

4. Adjusts for bias resulting from timing of snowmelt.

5. Prepares eight remaining levels of forecast based upon median level. Makes same adjustments as in 4. above.

6. Prepares date-of-forecast through September 30 forecasts as in steps 3, 4, and 5 above.

7. Goes through steps 1 - 6 for each of remaining six basins.

Forecast Output. Forecasts for all seven basins are printed out as indicated on Figure 3.

Actual computational and output time is approximately one minute, but it varies slightly as a result of whether the forecast is based on (a) the procedure using precipitation alone, (b) the procedure using snowpack and precipitation, or (c) a combination of the above two procedures during the intermediate period.

The program has the ability to make a forecast as of any date, even though data received after that date are introduced into the system. For example, a forecast may be made as of April 1, even though precipitation through April 16 is included. The result is that actual data are substituted for unknown data, giving an April 1 forecast and reducing the probability range to that of April 16. Hypothetical data can be put into the program at any time of year to answer, within the computer's capabilities -- "what would happen if ---?"

Interesting Observations

During development of the forecasting procedure for Union Valley and extension of this procedure to the remaining six basins, a number of rather interesting items were observed.

Note the convergence of the forecast probability diagram in Figure 2. On April 1st the snow procedure is about the same as the precipitation procedure on the positive side, but it closes more rapidly on the negative side. As expected, over-all error is smaller on the negative side because it is possible to receive several hundred per cent of average precipitation after April 1, but it is impossible to receive less than zero precipitation. The negative side closes more rapidly, probably as a result of procedure error controlling on the negative side while weather after date of forecast controls on the positive side.

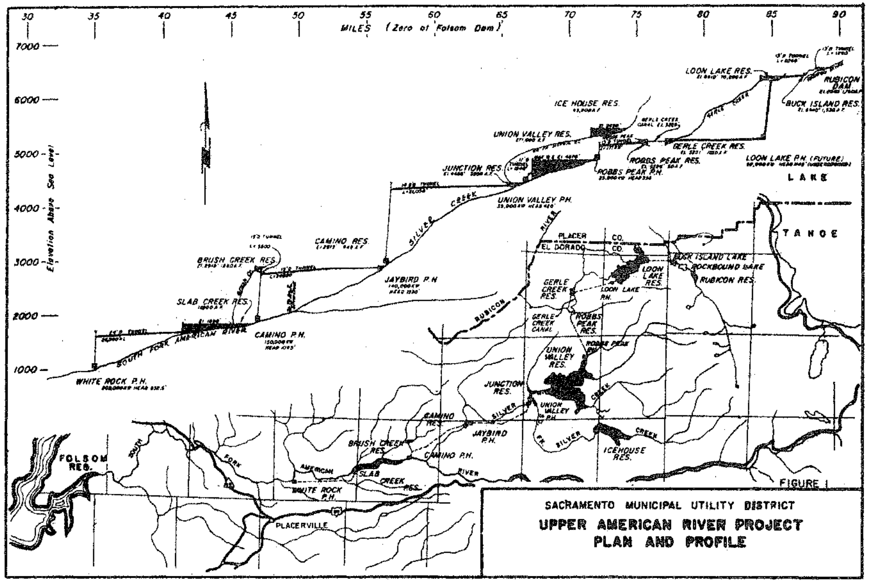

FIGURE 1

SACRAMENTO MUNICIPAL UTILITY DISTRICT

UPPER AMERICAN RIVER PROJECT

PLAN AND PROFILE

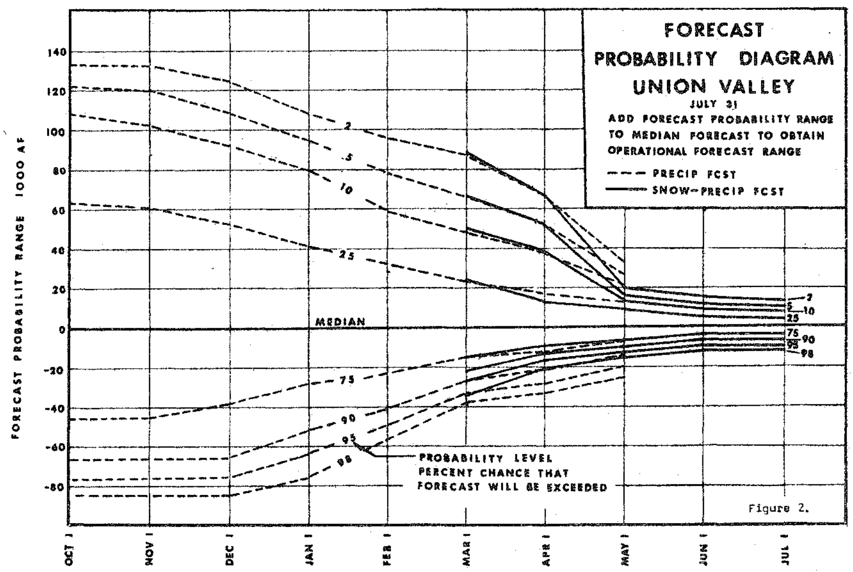

Figure 2.

COMPUTER OUTPUT - DATA CHECK

INPUT DATA -- FORECAST FOR 4 -5-68

PREVIOUS YEAR'S RUNOFF (THOUSAND AF)

UNION VALLEY	257.560
ROBBS BASIN	105.200
LOON LAKE BASIN	156.200
ICEHOUSE BASIN	83.300
S. FORK AMERICAN	748.500
AREA A	50.000
AREA B	39.000

MONTHLY DATA

	SEPT	OCT	NOV	DEC	JAN	FEB	MARCH	APRIL	MAY	JUNE
RUNOFF (THOUSAND AF)										
UNION VALLEY		4.380	2.000	4.500	6.330	23.590	18.420			
ROBBS BASIN		0.300	0.150	0.100	2.700	13.550	9.360			
LOON LAKE BASIN		0.500	0.530	0.560	3.640	15.670	5.510			
ICEHOUSE BASIN		1.000	1.200	1.200	2.330	5.520	4.560			
S. FORK AMERICAN		13.300	9.900	11.300	12.000	42.100	42.011*			
AREA A		0.320	1.410	1.460	1.040	7.100	6.347*			
AREA B		0.250	1.110	1.160	0.810	5.500	4.951*			
PRECIP (IN)										
PACIFIC HSE.	0.45	2.77	3.92	7.14	7.59	5.83	5.47			
PLACERVILLE	0.08	1.50	4.42	4.94	6.52	4.41	3.87			
TWIN LAKES	1.45	1.34	4.05	3.93	7.20	5.62				
SALT SPRINGS	0.90	1.49	3.25	4.37	7.08	6.15	4.17			
TAHOE CITY	0.85	1.37	2.18	4.32	5.26	4.89	3.31			
L. SPAULDING	0.77	2.92	5.22	8.83	12.44	10.18	7.52			

DAILY RUNOFF (THOUSAND AF)

UNION VALLEY	0.614*	0.614*	0.614*	0.614*
ROBBS BASIN	0.700	0.520	0.470	0.420
LOON LAKE BASIN	0.184*	0.184*	0.184*	0.184*
ICEHOUSE BASIN	0.220	0.190	0.190	0.170
S. FORK AMERICAN	1.400*	1.400*	1.400*	1.400*
AREA A	0.212*	0.212*	0.212*	0.212*
AREA B	0.165*	0.165*	0.165*	0.165*

DAILY PRECIP (INCHES)

PACIFIC HSE.	0.06	0.46	0.02	0.0	0.0
PLACERVILLE	0.0 *				
TWIN LAKES	0.0				
SALT SPRINGS	0.62	0.18	0.04	0.0	0.0
TAHOE CITY	0.0 *				
L. SPAULDING	0.31	0.16	0.0	0.0	0.0

SNOW COURSE WATER CONTENT (INCHES)

	DATE	WATER CONTENT	EST. APRIL 1 SC
ROBBS VALLEY	3-28-68	16.40	16.40
LYONS CREEK	3-28-68	22.20	22.20
WRIGHTS [illegible]	3-28-68	21.80	21.80
ECHO SUMMIT	3-29-68	22.50	22.50
TWIN LAKES	3-26-68	22.00	22.00
UPPER CARSON PASS	3-26-68	28.50	28.50
LAKE LUCILLE	3-26-68	45.30	45.30
RUBICON PEAK NO 1	3-76-68	41.80	41.80
RUBICON PEAK NO 2	3-26-68	26.60	26.60

* MISSING, INTERPOLATED, OR EXTRAPOLATED DATA

In years of above average snowmelt runoff, temperature played an extremely important role in forecast accuracy for the through June 30 forecast. In all years, the major portion of the snowmelt runoff occurs before July 31st. In years when snowmelt runoff is not abnormally high, almost all of the snowmelt has occurred by June 30th. However, during years when snowmelt runoff is abnormally high, temperature plays a significant role in determining just how much runoff will occur after June 30th. On the forecast through June 30th, a temperature dependent probability is superimposed on the years with far above normal snowmelt runoff, thus creating a tendency to over-forecast on the through June 30th forecast in big years. Each basin was analyzed separately for the effects of the June 30th probability bias, and, as expected, the higher elevation basins, where snowmelt represents a larger portion of the hydrologic picture, appeared to be most effected by this bias. Bias correction was made to give the same probability ranges for an abnormally high runoff year as would be expected for a normal or below normal runoff year. (Burns, 1957)

The effects of late season and summer precipitation became very apparent in this analysis (Hannaford, 1967). Analysis of daily records showed rises on the snowmelt hydrograph, probably due to localized precipitation which may or may not have been recorded at precipitation stations used in the analysis. The same situation was found in analyzing summer record after virtually all of the snowmelt had disappeared. Although quantities of water that could be directly attributed to summer precipitation were normally quite small, in several situations summer storms produced significant volumes of water. Precipitation from convective activity which occurs between May 1st and the end of snowmelt may be responsible for substantial quantities of runoff even though there is little evidence at recorded stations that such precipitation occurred.

Although some analysis was done to determine if any relationship could be detected between precipitation occurring before and after any given date, no positive results were found (Barnes, 1964). However, there does appear to be a tendency for heavier late season rainfall in the more recent years of record. Although it would be difficult to define this tendency statistically, studies performed on this project suggest that the tendency is real. Unfortunately, no work has been done to review records before 1926 in order to establish the repeatability of this tendency.

<u>Summary and Conclusion</u>

The Upper American River water supply forecast system has incorporated forecast and probability computations along with data handling and estimating procedures in an operational package requiring only minutes of computer time for each updating. If operated by hand, each updating would require from three to twenty man-hours, depending upon how much of the previous run could be utilized. This assumes that basic data would have been collected and tabulated previously for either the hand or computer runs.

A computer forecast system eliminates the very real problem of introducing human computational error, especially where computations are so involved and tedious and intermediate checks are not available. The problems caused by incorrect input data, have been alleviated through output of input data. However, we still have the frustrations of dealing with an electronic contrivance which cannot give a civil answer to even the simplest subjective questions.

REFERENCES

Barnes, S. M., 1964. An Analysis of Precipitation at Huntington Lake. Journal of Irrigation and Drainage Division, American Society of Civil Engineers, March 1964.

Blanchard, F. B., 1955. Operational Economy Through Applied Hydrology. Proceedings of the Western Snow Conference, April 1955.

Burns, J. I. and F. A. Strauss, 1957. Graphical Forecast Errors. Proceedings of the Western Snow Conference, April 1957.

Hannaford, J. F., 1959. Economics of Early Season Water Supply Forecasts. Proceedings of the Western Snow Conference, April 1959.

Hannaford, J. F. and M. C. Williams, 1967. Summer Hydrology of the High Sierra. Proceedings of the Western Snow Conference, April 1967.

Kuehl, D. W., 1962. Improving Water Supply Forecasts by Limiting Techniques. Proceedings of the Western Snow Conference, April 1962.