MULTIPLE-GRAPHICAL CORRELATION FOR WATER SUPPLY FORECASTING

by
Jack F. Hannaford[1/]

Introduction

The primary purpose of this paper is to acquaint you with a method of multiple-graphical correlation for use in water supply forecasting. The method has been used quite successfully by the California Division of Water Resources in the preparation of forecasts for the annual Water Conditions in California reports. These reports are published covering the conditions as of February first, March first, April first, and May first, with the most emphasis being placed on the April first report.

Normally, the forecast is based upon the April through July spring runoff period, during which time the majority of snowmelt runoff occurs.

In order to provide a workable method, the forecasting procedure must fulfill several requirements. It must be based upon data which are readily available at the time of forecast, and if portions of these data are not available, the procedure should have a means of developing an approximate forecast with the data which are available. It should be in such a form that it can be worked successfully by persons not familiar with the watershed or conditions, as often during the rush of forecasting, when many streams must be forecasted in a very short time, adequate skilled personnel are not available. The procedure should have provisions for several or continuous forecasts during the year, so that the same scheme may be used in March or May without excessive computational work. Last, but not least, the forecasting procedure must be reliable, for without this last qualification, the most flexible method would indeed be useless.

The methods as outlined here should provide a logical approach to forecasting almost any stream, not only in the Sierra, but in British Columbia or Nepal.

General Parameters

Since the most emphasis is placed upon the April first forecast, the forecasting procedure is based upon the data as of this date. The water in the basin on April first which is available for the spring runoff exists in several different forms in the Sierra basins of California. One may look at all of these forms as "basin wetness". The recognition of the existence and roles of the more important indexes of wetness is of primary importance in successful water supply forecasting.

At the relatively barren, rocky, high-elevation reaches of the typical Sierra basin, the most obvious form of basin storage is the snowpack. The measured water content of the pack varies with topography, latitude, and more pronounced within the basin, altitude. In most of the Sierra basins the snowpack is the main contributor to the spring and early summer runoff, even though the area covered by snow may not constitute a major portion of the basin. Snowpack, then, is a good index of the basin storage or wetness at the time of forecast.

The lower elevation areas of the basin also contribute some to the spring runoff, whether they are covered with snow or not at the time of the survey. We might assume that the contribution from this area would be roughly proportional to the contribution as determined by snow surveys. However, we also know that the runoff will be equal to the precipitation in the area minus the runoff to date (less basin losses). The precipitation information is usually readily available from gages in the lower, more populated reaches of the basin. There are not usually enough runoff data, however, to separate the runoff from the various portions of the basins. For this reason, total runoff from the beginning of the season to the date of forecast might be taken as an index of the depletion of storage in the basin. Precipitation minus runoff to date, even with such limitations, gives a fairly reliable index to the storage or wetness of the basin.

Another form of storage is the carryover from the previous year or years, usually in the form of ground water. In the relatively impermeable granitic type basins this form of storage may prove to be unimportant. However, in the permeable volcanic areas which are found in the central and northern Sierra, carryover storage from year to year may become a large factor in developing a reliable forecasting method. The previous year's runoff, or portions thereof, has been found to be a good index of this carryover effect, and has been used in the Division's forecasting schemes because of the ease of measurement and computation. Precipitation or snowpack from a previous year, or perhaps the minimum flow during a late summer month may give an equally good index.

The water which is stored in the basin at the time of the April first forecast has been generally covered, but there still remains one important contribution to the spring runoff - the spring precipitation occurring after the date of the April first forecast. Since the snow survey program is greatly reduced in California after the April first measurement is made, the later season snowpack measurements are not used to evaluate the new water content added to the basin by late season storms. At some future date it may be possible to use the late season snowpack measurements directly. The late season precipitation as measured at the regular precipitation stations, when added to the

---

1/ Assistant Hydraulic Engineer, Division of Water Resources, State of California

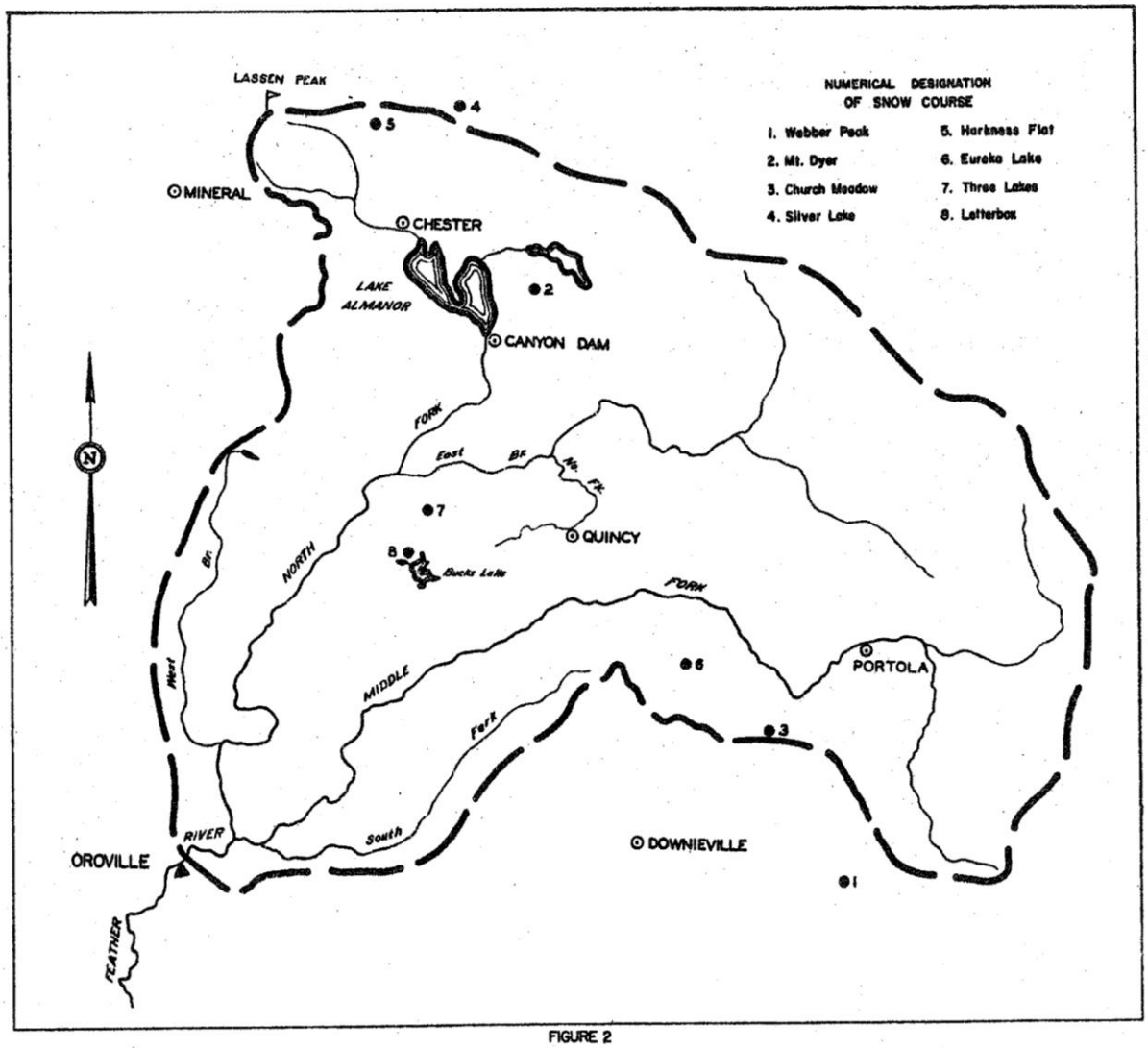


FIGURE 2

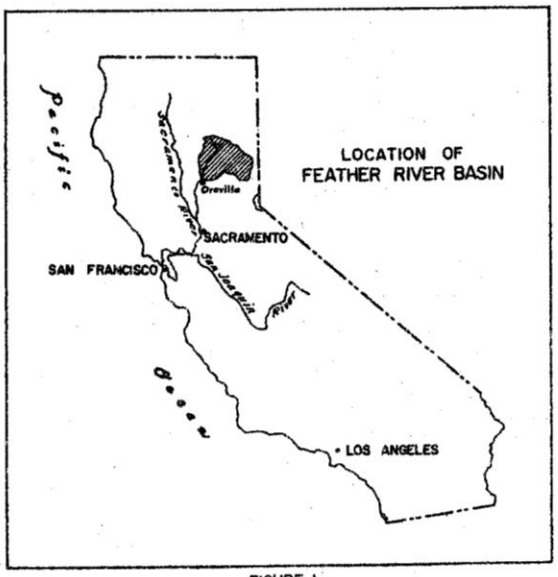


FIGURE 1

LEGEND

▲ STREAM GAGING STATION

⊙ PRECIPITATION STATION

●7 SNOW COURSE

STATE OF CALIFORNIA
DEPARTMENT OF PUBLIC WORKS
DIVISION OF WATER RESOURCES

FEATHER RIVER BASIN

SCALE OF MILES
5 0 5 10 15 20

basin storage parameters mentioned above, has proved to be a suitable adjustment factor for the scheme. In actual practice, it has been found that the type of storm occurring after April first normally changes gradually from the general coverage winter type storm to the more localized thunder-storm type. The significance of storms in determining the runoff seems to decrease as the season progresses, partially due to the more localized conditions, and partially due to the fact that the losses from each storm become progressively greater as the season advances. The precipitation falling on the drier ground of the late season is partially lost in refilling field capacity. By the last of June, the effects of late season precipitation become almost negligible.

Although one can no doubt think of a number of other parameters which could affect the runoff, the Division has found that the above variables give satisfactory forecasts, almost within the limits of field measurements. It should be noted here that all of the above mentioned data are readily available at the date of forecast (which is actually the fourth or fifth of the month).

Approach and Method

The above parameters must be related by some type of an equation to produce the desired forecast. The forecaster may control the form of the equation by selection of the method of solution. For instance, a common mathematical solution, the linear regression, assumes that the final forecast bears a straight-line relationship to each of the variables introduced, or at least, to each of the variables upon combination in such a forecast. Some forecasters[1] have altered the linear regression by the introduction of inverse functions, powers, roots, etc., incorporated directly into the parameter itself. These methods, however, require some foreknowledge of the form which the equation must take.

A forehand knowledge of the form of the equation is difficult to acquire without a great deal of study on the individual basin, but fortunately it is possible to develop a forecasting scheme making only very general assumptions as to the form of the equation. The Division of Water Resources has found that the relationship of the independent variables can most readily be determined by graphical means. The multiple-graphical correlation takes the form of the Feather River Forecast Diagram in Figure 3. The actual mechanics of the multiple-graphical correlation are quite completely written up elsewhere[2/3].

The graphical solution does not limit the form of the equation, as long as it may be represented graphically in a plane. Curved and fanning lines which develop in the various quadrants may represent inverse functions, roots, powers, etc. If the relationship does prove to be linear, the resulting lines on the chart will be straight, parallel, and evenly spaced. If the relationship is not linear, it will be indicated by the shape of the lines.

The multiple-graphical correlation provides no one "best" solution. (The interaction of variables allows one to use many variations in form and still get comparable results.) Engineering judgment must be used to secure a relationship which adheres rigidly to the basic hydrologic principles which we know to exist. The form and trend of the variables may be observed visually, and possible deviations from a sound logical approach are often easily spotted. A relationship, even though apparently strong, should not be relied upon unless it can be justified through the basic principles of hydrology.

A "percentage of normal" index is used to convert the raw data into a form readily adaptable for use in the multiple-graphical relationship. The percentage of normal indexes permit rapid computation for an approximate forecast, even though some precipitation or snowpack data may be missing. Normals are based upon a 50-year period, and adjusted where records are shorter. Actually, the forecast depends in no way upon what "normal" is used, as long as consistency is maintained within each index (ie., Snowpack Index). The primary purpose of the normal is to simplify the reduction of data to a common form and standardize the weighting (in this case equal weighting) for each snow course and precipitation gage.

The variables involving runoff were used in the forecast scheme directly in acre-feet. This not only simplified computation, but gave the resulting forecast in acre-feet units.

Multiple-graphical correlations for forecasting the April through July runoff have been completed for sixteen Central Valley streams and are available upon request to the California Snow Survey Program, Division of Water Resources, P. O. Box 1079, Sacramento, California.

Feather River Forecast

As a typical example of the forecasting method used by the Division of Water Resources, let us take the well-known Sierra Nevada stream, the Feather River. This river heads in the northernmost section of the Sierra Nevada proper, and has the highest runoff of any Sierra stream. Lassen Peak on the northern boundary is usually considered the dividing point between the Sierras and the Cascades. It is 10,465 feet high, and is the only active volcano

---

1/Blanchard, F. B., Mokelumne River Streamflow Forecasting Analysis, East Bay M.U.D., 1954.
2/Linsley, R. K., Kohler, M. A., and Paulhus, J. L. H., Applied Hydrology, McGraw-Hill Book Co., N.Y., 1949 Appendix A, p. 643.
3/Ezekiel, M., Methods of Correlation Analysis, John Wiley & Sons, Inc., N.Y., 1941.

in the continental United States, although it has not erupted since 1921. The watershed consists of both granitic and volcanic areas, with the latter predominating. This physical characteristic of the basin has a pronounced effect on the annual distribution of the runoff.

The Feather River basin is drained by three main forks, the North, Middle, and South. The total drainage area above Oroville is 3,611 square miles. The normal April-July runoff at the Oroville gaging station is about 1,920,000 acre-feet, and since the beginning of snow surveys in 1930 has varied from 525,000 to 4,638,000 acre-feet. Figure 2 gives the general basin layout, with snow course and precipitation gage sites.

As stated previously, the snowpack parameter uses the percentage of normal snowpack as an index. Although many methods for weighing snow courses might be suggested, we have decided on an arithmetical average for ease of computation. An original network of existing snow courses with the longest period of record was selected, and the average percentage of snowpack for a given year for all courses was taken as the basin index for that year. At a later step in the development of the scheme, each course was compared with the results to test the compatibility of that course with the scheme. The least compatible courses were dropped as being unrepresentative of the basin as a whole.

The average percentage normal index for several stations was used to compute the October 1-April 1 Precipitation Index. The station percentage normals for any year in any one Sierra basin are usually very consistent. The storm patterns are very general during this period so that large percentage differences are not likely to occur. Again, as in the snowpack indexes, the individual station indexes are compared with the final results, and any stations which appear to be inconsistent with the scheme are dropped. In this scheme, out of an original nine stations, six were used in the final forecast.

In the next quadrant is the October 1-April 1 Runoff which correlates negatively with the forecasted runoff. It should be noted here that the correlation is not on a one to one basis, giving at least some clue to the interrelationship between snowpack on one hand and precipitation minus runoff on the other.

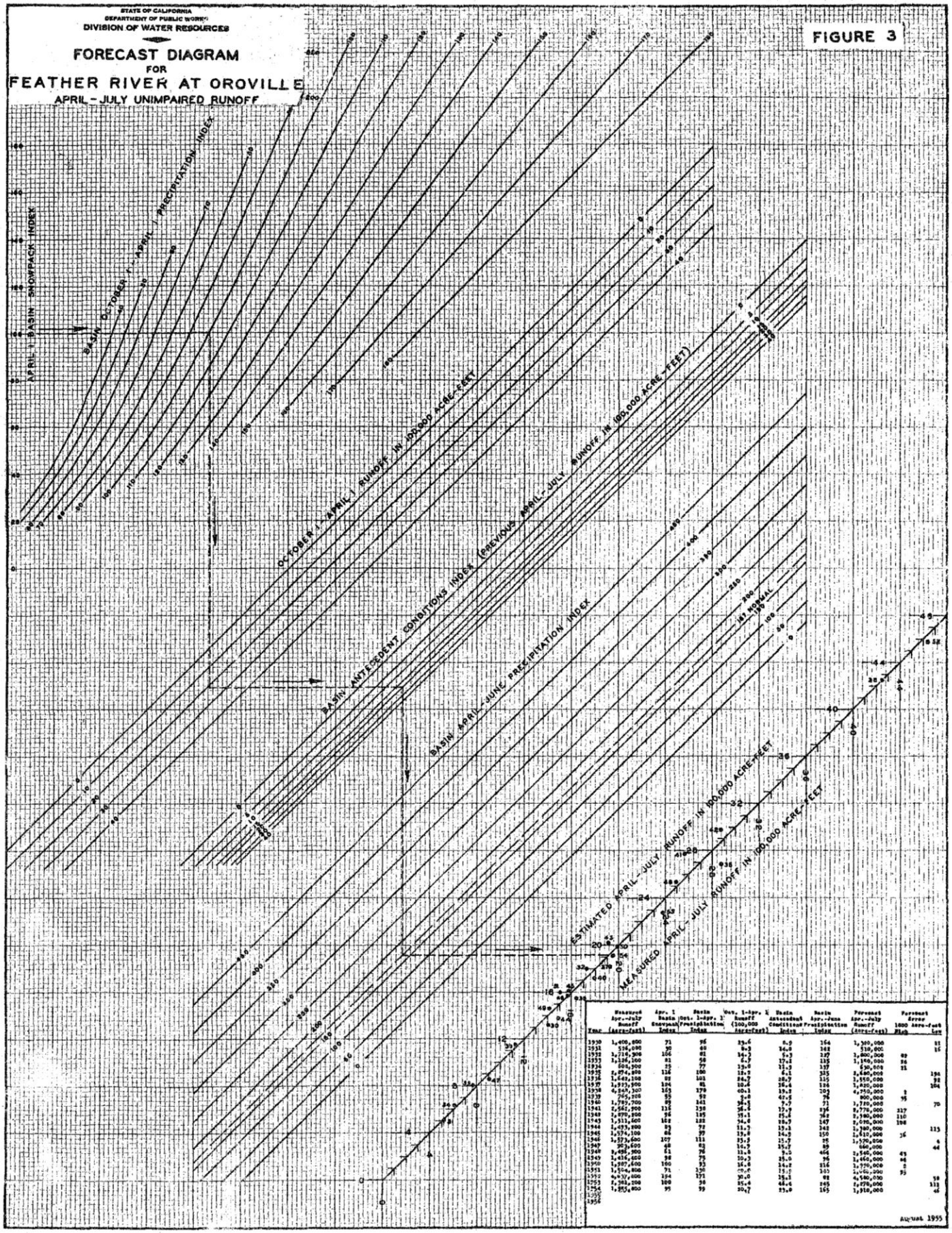
The Basin Antecedent Conditions Index is merely an indication of the portion of water contributed by carryover from the previous year, or perhaps it may be considered as a measure of the groundwater storage which must be filled each year before runoff may occur. Whichever way you may look at it, the previous year's spring runoff seems to provide a good index of the carryover effects for this basin. Note the spread of the lines, indicating that the forecast varies with, but not directly proportional to, the previous April-July runoff. The basin might be compared to a sponge, whose capacity to absorb water at any particular time is not only a function of the water available, but also of the remaining pore spaces which may take up water. When there is more than enough water to fill the pore spaces, or storage, further amounts of water do not increase the later yield of the sponge, or the basin.

The last independent parameter is the April-June Precipitation Index. It has been found that by the last of June the effects of late season precipitation become almost negligible. In general, if the effectiveness in producing runoff of the April precipitation is taken as 100%, the May precipitation is about 80%, the June precipitation is 50%, and the July precipitation is zero. (See Figure 4). The index is computed on the basis of the percentage of normal April precipitation. In this particular case, the normal April-June Precipitation Index is 167, as it includes April ($\frac{100 \text{ April precip.}}{\text{Norm. April precip.}}$ = 100 normally), May ($\frac{80 \text{ May precip.}}{\text{Norm. April precip.}}$ = 52 normally), and June ($\frac{50 \text{ June precip.}}{\text{Norm. April precip.}}$ = 15 normally). This method permits easy adjustment of the index and the forecast as the season progresses and more data become available.

Referring to Figure 4, it will be noted that the forecast scheme has been developed primarily for the April first forecast, with adjustments for May first and June first. However, a continuous forecast may be made from, say, February first to June first, if the forecaster has some means of estimating the probable occurrence of weather phenomena following the date of forecast. Boxes (d), (i), (n), and (s) in Figure 4 give the "normal" snowpack or precipitation to occur after the forecast date as determined from cumulative curves derived from past records (Figure 5). The earlier forecasts (March and February) are of course subject to progressively greater error or deviations, since major portions of the precipitation usually occur during these months. It is felt that the uncertainty of weather conditions does not warrant further refinement of the early season forecasts. Fred Strauss, in the following paper, will discuss more thoroughly the errors or deviations which may be expected by this forecasting method as the season progresses, and the application of exceedence intervals in making use of the early season forecasts.

Reliability is the real test of any forecast scheme. Standard errors, standard deviations, correlation coefficients and the other statistical quantities associated with reliability do not readily find direct application in the multiple-graphical solution. All of the above quantities are associated with the degrees of freedom lost in the development. We would certainly be first to admit that the number of degrees of freedom lost in the graphical solution would be great, considering the fanning, spread, curvature, and large number of parameters. However, according to an old California proverb, the hydrologically sound forecasting method is worth a number of degrees of freedom. The reliability of the scheme may possibly be much greater than a standard error analysis might indicate.

In the last quadrant is a 45° line upon which the historical "hindcast" errors are recorded. Points falling upon the line indicate no error, points falling above the line indicate that the forecast was high, and points falling



Year	Measured Apr.-July Runoff (Acre-feet)	Apr. 1 Basin Snowpack Index	Basin Oct. 1-Apr. 1 Precipitation Index	Oct. 1-Apr. 1 Runoff (100,000 Acre-feet)	Basin Antecedent Conditions Index	Basin Apr.-June Precipitation Index	Forecast Apr.-July Runoff (Acre-feet)	Forecast Error 1000 Acre-feet Plus	Minus
1930	1,400,000	71	96	29.6	8.9	164	1,300,000		21
1931	536,000	30	49	9.3	16.0	142	510,000		16
1932	1,716,000	106	81	34.3	6.3	197	1,800,000	82	
1933	1,126,100	61	58	6.7	17.1	135	1,140,000	26	
1934	604,900	33	77	13.0	11.3	137	630,000	21	
1935	2,074,000	116	100	18.7	6.1	115	2,400,000		194
1936	1,642,100	81	101	24.7	28.7	115	1,550,000		92
1937	1,923,000	124	81	10.6	16.4	126	1,820,000		104
1938	2,269,300	153	179	40.1	19.1	103	2,250,000	1	
1939	765,200	55	52	3.8	47.5	78	800,000	35	
1940	1,787,700	98	141	34.5	7.7	71	1,720,000		70
1941	2,542,700	118	138	36.2	17.7	174	2,778,000	237	
1942	1,870,200	96	185	35.1	15.6	362	2,380,000	110	
1943	1,511,600	102	123	24.6	28.5	147	2,070,000	198	
1944	1,073,200	53	77	11.7	15.1	142	1,380,000		115
1945	1,574,100	81	97	19.8	14.5	150	1,610,000	36	
1946	1,573,600	107	111	15.3	15.7	95	1,570,000		3
1947	703,600	48	83	14.7	35.7	89	660,000		44
1948	1,496,900	61	76	11.6	3.5	464	1,540,000	43	
1949	1,414,400	98	75	10.3	35.0	96	1,460,000	46	
1950	1,387,600	100	93	14.8	14.2	216	1,370,000	3	
1951	1,504,800	71	150	29.6	15.3	143	1,400,000	93	
1952	4,437,000	194	171	30.0	35.1	81	4,540,000		58
1953	2,361,200	100	38	25.4	46.4	245	2,478,000		111
1954	1,895,800	95	99	30.7	23.9	165	1,910,000		46
1955									
1956									

FEATHER RIVER FORECAST

SNOWPACK INDEX (Normal April 1 Basin Snowpack Index = 100)

Snow Courses (1)	Precipitation Adjustment Ratio (2)	Normal Snow Course Water Content (Inches) (3)	February 1: Date of Field Measurement (4)	Water Content of Field Measurement (Inches) (5)	Precipitation at Charter between Date of Measurement and First of Month (Note Sign) (6)	Col. 2 x Col. 6 (Note Sign) (7)	Col. 5 + Col. 7 (Note Sign) (8)	Col. 8 x 100 ÷ Col. 3 Snow Course Index (9)	March 1: Date of Field Measurement (14)	Water Content of Field Measurement (Inches) (15)	Precipitation at Charter between Date of Measurement and First of Month (Note Sign) (16)	Col. 2 x Col. 16 (Note Sign) (17)	Col. 15 + Col. 17 (Note Sign) (18)	Col. 18 x 100 ÷ Col. 3 Snow Course Index (19)	April 1: Date of Field Measurement (22)	Water Content of Field Measurement (Inches) (23)	Precipitation at Charter between Date of Measurement and First of Month (Note Sign) (24)	Col. 2 x Col. 24 (Note Sign) (25)	Col. 23 + Col. 25 (Note Sign) (26)	Col. 26 x 100 ÷ Col. 3 Snow Course Index (27)
Webber Peak	1.7	44.8																		
Mt. Dyer	1.0	35.6																		
Church Meadow	1.4	36.4																		
Silver Lake	1.2	30.7																		
Harkness Flat	1.1	28.9																		
Bucks Lake	1.4	36.2																		
Three Lakes	1.5	40.6																		
Lattimore	2.1	54.2																		

Total								(a)		Total					(z)		Total		(u)	
Number of Courses Used in (a)								(b)		Number of Courses Used in (z)					(1)		Number of Courses Used in (u)		(v)	
Feb. 1 Basin Snowpack Index (a÷b)								(c)		Mar. 1 Basin Snowpack Index (z÷1)					(m)		Apr. 1 Basin Snowpack Index(u÷v)		(w)	
Normal Feb. 1-Apr. 1 Snowpack Index Increase								(d)	47	Normal Mar. 1-Apr. 1 Snowpack Index Increase					(n)	20				
Feb. 1 Estimate of Apr. 1 Snowpack Index(c+d)								(e)		Mar. 1 Estimate of Apr. 1 Snowpack Index(m+n)					(o)					
								Feb. 1							Mar. 1					Apr. 1

Season of _____

Computed April-July Runoff _____ A-F.

Forecast as of	Acre-Feet	% Error
February 1		
March 1		
April 1		
May 1		
June 1		
July 1		

PRECIPITATION INDEXES (Normal October 1 - April 1 Precipitation Index = 100; Normal April - June Precipitation Index = 147)

Precipitation Stations (30)	Normal October 1 - April 1 Precipitation (Inches) (31)	October 1 - February 1 Precipitation (Inches) (32)	Col. 32 x 100 ÷ Col. 31 October 1 - April 1 Precipitation Index (33)	October 1 - March 1 Precipitation (Inches) (20)	Col. 20 x 100 ÷ Col. 31 Oct. 1 - Mar. 1 Precipitation Index (21)	October 1 - April 1 Precipitation (Inches) (28)	Col. 28 x 100 ÷ Col. 31 Oct. 1 - April 1 Precipitation Index (29)	Normal April Precipitation (Inches) (30)	April Precipitation (Inches) (31)	Col. 31 x 100 ÷ Col. 30 April Precipitation Index (32)	May Precipitation (Inches) (33)	Col. 33 x 100 ÷ Col. 30 May Precipitation Index (34)	June Precipitation (Inches) (35)	Col. 35 x 100 ÷ Col. 30 June Precipitation Index (36)
Mineral	41.02							4.85						
Chester	25.45							2.30						
Canyon Dam	31.59							2.93						
Quincy	32.55							3.51						
Portola	16.41							1.66						
Downieville	53.64							5.47						

Total			(f)		(p)		(x)		Total		(aa)			(gg)			(kk)		
Number of Stations Used in (f)			(g)		(q)		(y)		Number of Stations in (aa)		(bb)			(hh)			(ll)		
Date of Forecast Basin Precipitation Index (f÷g)			(h)		(r÷g)=(r)		(x÷y)=(z)		April Index (aa)÷(bb) =		(cc)			(ii)			(mm)		
Normal Date of Forecast to April 1 Precipitation Index Increase			(i)	40	(s)	27			May Index		(dd)	53	(gr)÷(hh)=(ii)				(ll)		
Date of Forecast Estimate of April 1 Precipitation Index (h+i)			(j)		(rss)=(t)				June Index		(ee)	35		(ee)	35		(kk)÷(ll)=(mm)		
									April-June Precip. Index		(ff)			(jj)			(nn)		
								Feb. 1					Mar. 1					Apr. 1	
													May 1					June 1	
																		July 1	

October 1 - April 1 Runoff in 100,000 A-F.			
October 1 - February 1 Runoff			
Normal February 1 - April 1 Runoff	953		
February 1 - March 1 Runoff			
October 1 - March 1 Runoff			
Normal March 1 - April 1 Runoff	540		
March 1 - April 1 Runoff			
October 1 - April 1 Runoff			

Estimated October 1-April 1 Runoff

Feb. 1

Mar. 1

Apr. 1 Actual October 1 - April 1 Runoff

ANTECEDENT CONDITIONS INDEX- April-July Runoff from previous year 100,000 A-F.

FIGURE 4

STATE OF CALIFORNIA
DEPARTMENT OF PUBLIC WORKS
DIVISION OF WATER RESOURCES

FORECAST COMPUTATION FORM FOR FEATHER RIVER AT OROVILLE APRIL - JULY RUNOFF

August 1955

below the line are low. This represents how well the data fit the forecast, or the error inherent in the scheme, as all data up to July first are included.

Adherence to the 45° line, in combination with the logical construction will probably provide the most applicable measure of reliability, as long as an adequate period of record is used in the development. The logical basis for construction of the forecast diagram also inspires confidence in the scheme, giving the forecaster a measure of psychological peace of mind. The real test, however, is verification; even the logical approach is no substitute for history.

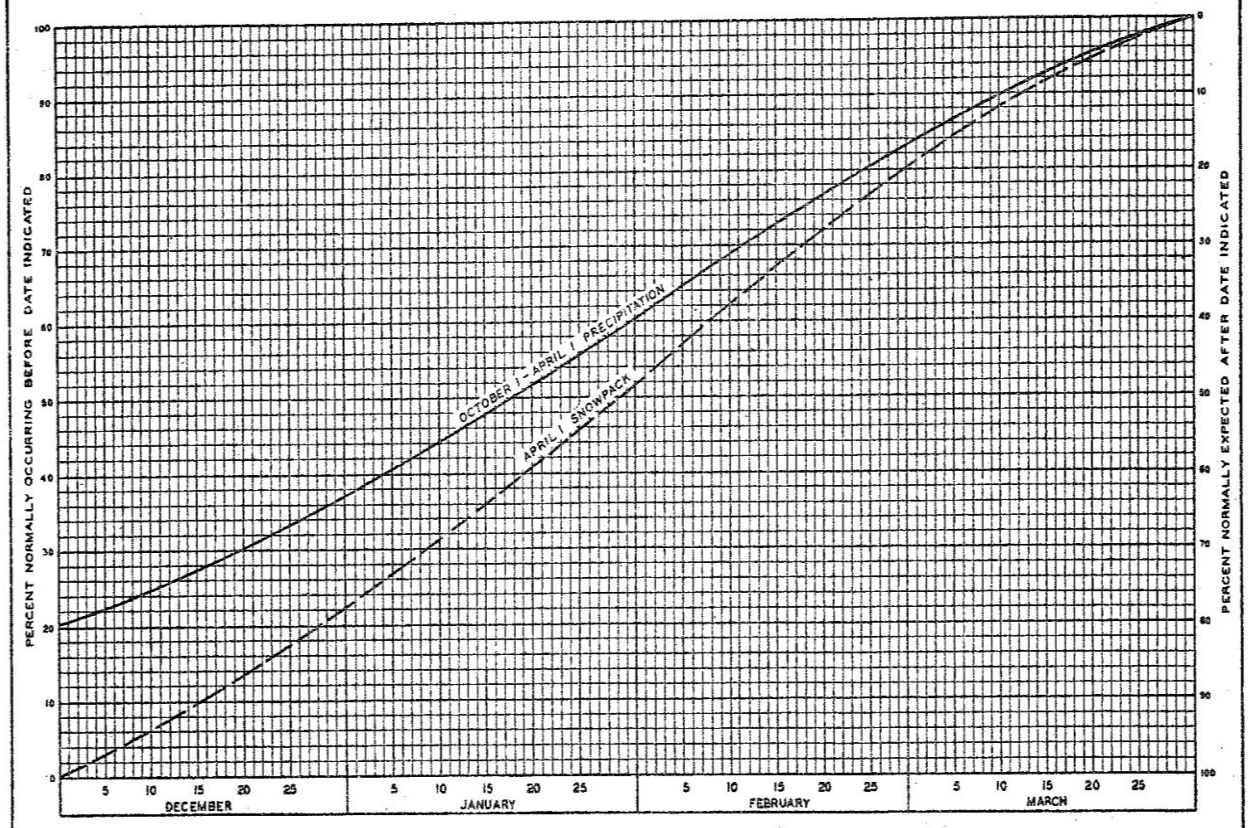


FIGURE 5

STATE OF CALIFORNIA
DEPARTMENT OF PUBLIC WORKS
DIVISION OF WATER RESOURCES

NORMAL PRECIPATION AND SNOWPACK EXPECTANCY

AUGUST, 1955