in all parts of the province. Since most districts are scattered, communication facilities are very limited, consequently the Engineer is hard pressed to provide road condition information accurately - especially about some road in his district that he himself rarely sees in winter. However, most districts have radio or telephone communication with all foremen and major snow-plowing units on the road and the district office can at least supply a general report - if not a detailed one.

There are, of course, many other problems confronting the District Engineer in his snow removal work. Problems dealing with pressure groups; with isolated settlers; with mining concerns and the logging industry; with municipality bodies and community associations. But perhaps one more problem should be mentioned before closing and that is one dealing with personal snowplowing requests. To cite an example - one of our district men received a request to plow five miles of road beyond the district's customary terminal point on that road. A plow was already heading to the terminal point but the operator was almost dead beat, having bucked several slides on route which kept him busy for 18 hours. On reaching the terminal point, he made for his bed after servicing his machine, thinking only of the four hours sleep he could get before he had to be on the road for the return journey.

The party who had requested the additional plowing arrived shortly after the operator had retired, barged into his bedroom and demanded that he arise and get to it. The operator refused, explaining that his instructions were to plow only to this particular point and to start his return trip in approximately five hours. Whereupon the party jumped onto his bed and began undressing and stating that she would start screaming if he wouldn't agree to plowing the five miles to her home.

Problems such as these are not too frequent but they do occur.

Incidentally the extra five miles were plowed out.

GRAPHICAL FORECAST ERRORS

By

Joseph I. Burns and Fred A. Strauss[1/]

At the Portland meeting of the Western Snow Conference in April 1955, a paper presented to the group by Francis Blanchard entitled "Operational Economy Through Applied Hydrology" very adequately set forth the need for accurate water supply forecasting and showed how statistical levels of forecast variance were required if the water manager was to get the "most" out of his operational program. A mathematically devised forecast scheme was used in that paper to illustrate the method of solution to certain problems of a water system operator. The California Division of Water Resources uses a graphical forecast scheme, which necessitates special steps in order to determine statistical levels of forecast departure which are necessary for the making of operational decisions. Many forecasters have avoided use of the graphical method only because of inability to accurately define statistical levels of departure of forecast. However, there is good authority that a graphical forecast method may well be as valuable as a mathematically derived scheme. For instance, Mordecai Ezekiel had this to say on the relative value of a mathematical or graphical solution to a problem:

> "When there is some logical basis for the selection of a particular equation, the equation and the corresponding curve may provide a definite logical measurement of the nature of the relationship. When no such logical basis can be developed, a curve fitted by a definite equation yields only an empirical statement of the relationship and may fail to show the true relation. In such cases a curve fitted freehand by graphic methods, and conforming to logical limitations on its shape, may be even more valuable as a description of the facts of the relationship than a definite equation and corresponding curve selected empirically."

1/ Joseph I. Burns, Senior Hydraulic Engineer, California Department of Water Resources, Sacramento, California.
Fred A. Strauss, Senior Hydraulic Engineer, California Department of Water Resources, Sacramento, California. (Now with Althouse-Strauss, Engineering Services, Porterville, California)

Ezekiel goes on to state as to the acceptability of a graphical solution:

"With proper care in analyzing the data for interrelationships and in carrying through the successive approximations, graphic methods will ordinarily give results about as significant, within their error zone, as results obtained by the more laborious methods of fitting mathematical curves by extensive arithmetic calculations."

But a word of caution is introduced by Ezekiel on regression curves fitted graphically:

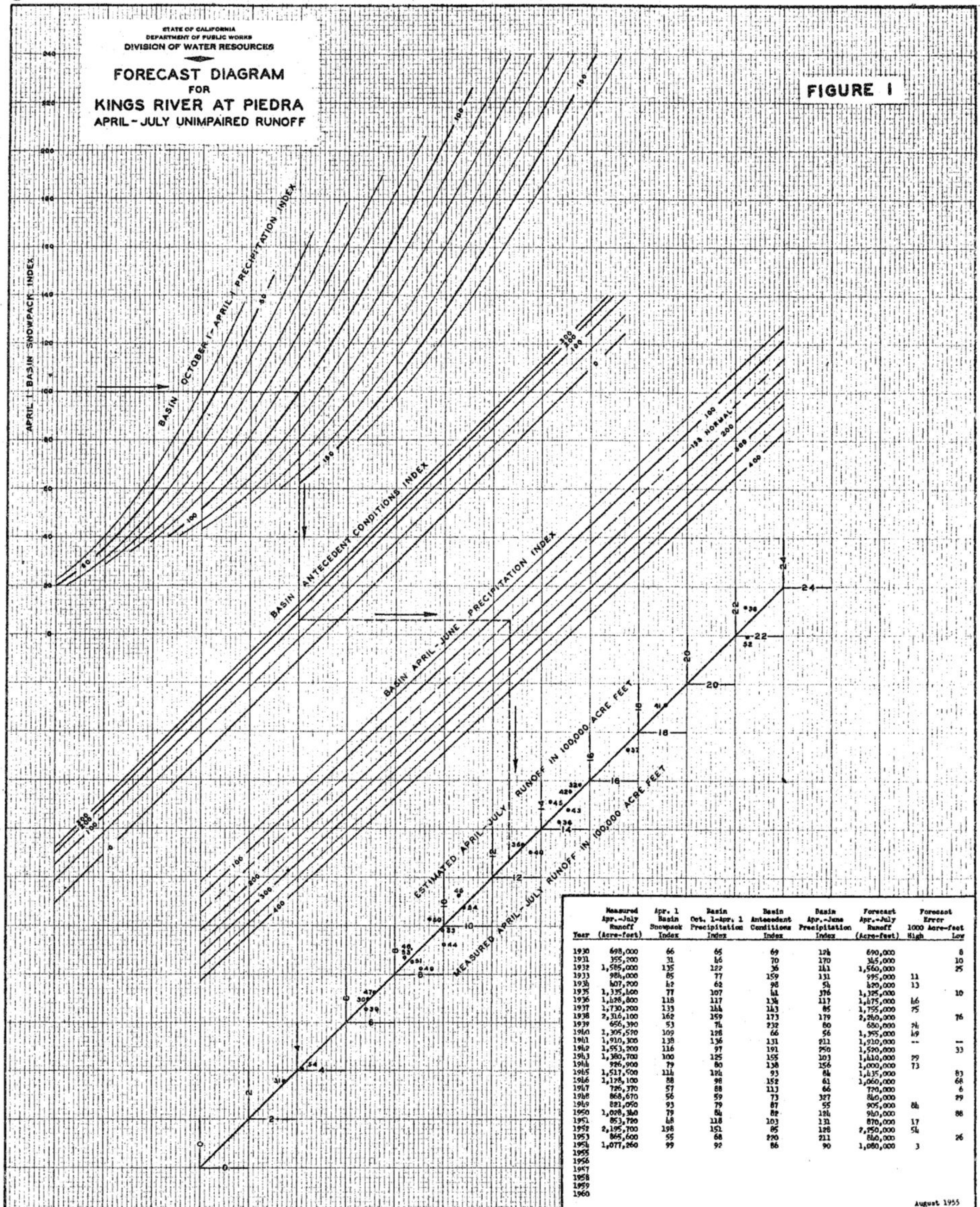
"In all problems based on random sampling, where any generalization as to the relations in the universe are to be based on the shape or slope of the final curves obtained by the graphic method, the error zones should be computed and should be given due consideration when the data are presented. This is just as important for curvilinear regressions as is the use of standard error values for linear regression coefficients."

In the discussion following, it is well to remember that most statistical arguments applicable to forecast methods apply equally to the mathematically and graphically developed schemes.

It is impossible to discuss levels of forecast error without touching upon "degrees of freedom" For example, using only two points, a straight line, which does not need to be plotted through the origin and which may be sloped to any degree, can be fitted to the data with the appearance that all past record fits the "curve". However, no one in the forecast field would wish to risk his reputation on using this curve in forecasting future events unless he has reason to believe that these two points do truthfully represent all conditions relative to the problem. In similar fashion, graphical schemes increase the liberties that may be taken to explain past data. Each additional liberty that is taken in spreading, fanning, and/or more abruptly changing the degree of curvature of any line, correspondingly increases the degrees of freedom lost in the development of the forecast scheme. Each parameter introduced unfortunately also increases the degrees of freedom lost. It is therefore necessary that a relatively large number of years of data be available before a complex graphical forecast scheme can be produced. The more complex the scheme, the greater the number of degrees of freedom lost. The California Department of Water Resources is indeed fortunate that the pioneers of water supply forecasting in 1930 set in operation the comprehensive program of data collection which now makes available for forecast analyses some 26 years of record.

In the following discussion reference is made to the forecast diagram for the Kings River at Piedra (see Figure 1). Note that this forecast diagram does not incorporate as many variables as the forecast diagram for Feather River at Oroville which has just been presented to this conference in the paper entitled "Multiple-Graphical Correlation for Water Supply Forecasting" by Mr. Jack Hannaford. (1956 Proceedings) Nevertheless, the following discussion is also applicable to the Feather River and similar type analyses. The following analysis also makes use of the forecast computation form for the Kings River at Piedra (see Figure 2) which is rigidly adhered to in the computation of historical forecasts. This adherence to the computation form is, of course, necessary if "intellectual honesty" is to govern the placement of the probability levels for the forecast accuracy evaluation. The computation form may not be as closely followed by the forecaster in making forecasts since there are numerous short-cuts and appoximations which may be necessary if the forecast is to be published according to schedule. After the publication schedule deadline has passed and all obtainable basic data have been obtained, a forecast utilizing the computation form should be made to insure that the short-cuts have not led the forecaster astray. Statistical levels of forecast departure as of any date will incorporate errors due to (1) the development of the forecast scheme, and (2) the inability to forecast long-term weather. In Mr. Blanchard's paper the error due to the forecast scheme has not been discussed and levels of forecast departure have been determined solely from the standpoint of the inability to forecast future weather occurrences. In the following analysis an attempt has been made to include both the scheme and weather forecast errors in the computations.

Ezekiel points out the need of drawing confidence lines in order to determine the strength of the forecast relationship in any particular range of data. Confidence lines should be developed for any type of forecast scheme when adequate data are available. However, it can be readily seen that there is insufficient years of snow survey data to develop confidence lines of high significance. To alleviate this shortcoming it is necessary that all forecast relationships be sound from a hydrologic standpoint. Graphical forecast schemes developed for adjacent basins of similar hydrologic characteristics should have closely related forecast scheme patterns. This factor is an independent check and is used in determining the acceptability of a newly devised forecast relationship in a manner similar to the use of confidence levels.



Year	Measured Apr.-July Runoff (Acre-feet)	Apr. 1 Basin Snowpack Index	Basin Oct. 1-Apr. 1 Precipitation Index	Basin Antecedent Conditions Index	Basin Apr.-June Precipitation Index	Forecast Apr.-July Runoff (Acre-feet)	Forecast Error 1000 Acre-feet High	Forecast Error 1000 Acre-feet Low
1930	698,000	66	65	69	124	690,000		8
1931	355,200	31	46	70	170	345,000		10
1932	1,585,000	135	122	36	141	1,560,000		25
1933	984,000	85	77	159	131	995,000	11	
1934	407,200	42	62	98	54	420,000	13	
1935	1,335,400	77	107	41	376	1,325,000		10
1936	1,428,800	118	117	134	117	1,475,000	46	
1937	1,730,200	133	144	143	85	1,755,000	25	
1938	2,316,100	162	159	173	179	2,240,000		76
1939	656,390	53	74	232	80	680,000	24	
1940	1,305,520	109	128	66	56	1,355,000	49	
1941	1,910,300	138	136	131	211	1,910,000	--	--
1942	1,553,200	116	97	191	250	1,520,000		33
1943	1,380,700	100	125	155	103	1,410,000	29	
1944	926,900	79	80	138	156	1,000,000	73	
1945	1,517,500	114	124	93	84	1,435,000		83
1946	1,128,100	88	98	152	61	1,060,000		68
1947	726,370	57	88	113	66	720,000		6
1948	868,670	56	59	73	327	840,000		29
1949	821,050	93	79	87	55	905,000	84	
1950	1,028,340	79	84	82	124	940,000		88
1951	853,720	48	118	103	131	870,000	17	
1952	2,195,700	198	151	85	128	2,250,000	54	
1953	865,600	55	68	220	211	840,000		26
1954	1,077,260	99	92	86	90	1,080,000	3	
1955								
1956								
1957								
1958								
1959								
1960								

August 1955

Method

The forecast scheme accuracy levels can usually be determined from the equation that standard deviation is equal to the square root of the summation of the square of individual forecast errors divided by the number of data less the degrees of freedom lost in the development. Statistical level constants may then be multiplied by this standard deviation in order to determine probability levels.

In the forecast scheme for the Kings River at Piedra, nine degrees of freedom are assumed to have been lost in the scheme development. The nine degrees of freedom lost are attributed as follows:

1. Use of a snowpack index.
2. Use of an October 1-April 1 precipitation index parameter.
3. The October 1-April 1 precipitation index lines are curvilinear.
4. The October 1-April 1 precipitation index curves "fan".
5. The October 1-April 1 precipitation index rate of fanning is non-linear.
6. Introduction of a basin antecedent conditions index parameter.
7. The basin antecedent basin index curves have nonuniform spread.
8. Introduction of the basin April-June precipitation index parameter.
9. The curves do not necessarily pass through the origin.

Study of a plotting of standard deviation vs. degrees of freedom lost based on a 25-year period for the Kings River at Piedra (see Figure 3) indicates the relative minor effect upon the computation of the standard deviation if the nine degrees of freedom previously assumed as lost is correct or whether, in reality, five or six additional degrees of freedom have been lost in the graphical development. It is evident that a standard deviation of 55,000 acre-feet is inherent in the scheme if nine degrees of freedom have been lost and 70,000 acre-feet is the standard deviation if fifteen degrees of freedom have been lost. If six additional degrees of freedom had been lost in addition to the nine previously listed, the increase in standard deviation would have been only 15,000 acre-feet.

Data used in determination of a forecast scheme yield a "hindcast" for past record rather than a true "forecast". Hence it is necessary to adjust the error of hindcast so as to be able to estimate the true level of accuracy of future forecasts. This adjustment is made by multiplying the computed hindcast errors in this case by the ratio of 55,000 acre-feet, standard deviation for nine degrees of freedom lost, to 44,000 acre-feet, standard deviation for zero degrees of freedom lost.

Although the forecast scheme has been developed for April 1 data as modified by late season precipitation, it can readily be seen that the same basic relationship can be used on any other date of the season. Should the scheme be used at an early date in the season, approximations must be made for snowpack and precipitation variance which may yet occur. This procedure has been outlined in Mr. Hannaford's paper and need not be discussed further at this time.

Development of Accuracy Levels

In the discussion following, three arbitrary values of expected runoff have been used to determine accuracy levels for different dates of the forecast season. The values selected were 600,000; 1,300,000; and 2,000,000 acre-feet being respectively approximately 50%, 100% and 150% of normal April-July runoff expectancy.

The following steps were then performed:

(1) Values of precipitation and snowpack indices which would yield forecasts of 600,000; 1,300,000; and 2,000,000 acre-feet were arbitrarily selected for the particular date of forecast.
(2) The assumed values in (1) were adjusted to April 1 by applying randomly selected historical precipitation and snowpack index variances for the period between the date of forecast, February 1 or March 1, and April 1.
(3) Corrections were applied for end season precipitation variance in accordance with the April-June precipitation for the same season randomly selected in (2).
(4) A forecast, using the aforementioned randomly selected data, is then made from the forecast chart for this synthesized "year". The differences between the forecast selected in step (1) and the values determined by step (3) are due to the inability to forecast future weather events.

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)	February 1: Water Content of Field Measurement (Inches) (5)	February 1: Precipitation at Huntington Lake Between Date of Measurement and First of Month (Note Sign)* (6)	February 1: Col. 2 x Col. 6 (Note Sign) (7)	February 1: Col. 5 + Col. 7 (Note Sign) (8)	February 1: Col. 8 x 100 ÷ Col. 3 Snow Course Index (9)	March 1: Date of Field Measurement (14)	March 1: Water Content of Field Measurement (Inches) (15)	March 1: Precipitation at Huntington Lake Between Date of Measurement and First of Month (Note Sign)* (16)	March 1: Col. 2 x Col. 16 (Note Sign) (17)	March 1: Col. 15 + Col. 17 (Note Sign) (18)	March 1: Col. 18 x 100 ÷ Col. 3 Snow Course Index (19)	April 1: Date of Field Measurement (22)	April 1: Water Content of Field Measurement (Inches) (23)	April 1: Precipitation at Huntington Lake Between Date of Measurement and First of Month (Note Sign)* (24)	April 1: Col. 2 x Col. 24 (Note Sign) (25)	April 1: Col. 23 + Col. 25 (Note Sign) (26)	April 1: Col. 26 x 100 ÷ Col. 3 Snow Course Index (27)
Bishop Pass	0.7	34.5																		
Blackcap Basin	0.8	38.2																		
Upper Burnt Corral	0.8	37.7																		
Beard Meadow	0.7	36.2																		
Rowell Meadow	0.6	30.8																		
Woodchuck Meadow	0.7	33.5																		
Swamp Meadow	0.8	40.0																		
Helm Meadow	0.6	28.8																		
Junction Meadow	0.4	21.9																		
Long Meadow	0.6	31.1																		
Statum Meadow	0.7	33.6																		
Post Corral Meadow	0.6	29.1																		
Big Meadow	0.6	29.6																		
Horse Corral Meadow	0.4	20.7																		
Bear Ridge	0.6	30.1																		
Fred Meadow	0.5	25.4																		

February 1		
Total	(a)	
Number of Courses Used in (a)	(b)	
February 1 Basin Snowpack Index (a ÷ b)	(c)	
Normal February 1-April Snowpack Index Increase **	(d)	35
February 1 Estimate of April 1 Snowpack Index (c + d)	(e)	

Feb. 1

March 1		
Total	(k)	
Number of Courses Used in (k)	(l)	
March 1 Basin Snowpack Index (k ÷ l)	(m)	
Normal March 1-April 1 Snowpack Index Increase **	(n)	9
March 1 Estimate of April 1 Snowpack Index (m + n)	(o)	

Mar. 1

April 1		
Total	(u)	
Number of Courses Used in (u)	(v)	
April 1 Basin Snowpack Index (u ÷ v)	(w)	

Apr. 1

* Positive if survey precedes first of the month.
Negative if survey follows first of the month.

** See Plate III

Precipitation Indexes (Normal October 1-April 1 Precipitation Index = 100; normal April-June Precipitation Index = 153)

Precipitation Stations (10)	Normal October 1-April 1 Precipitation (Inches) (11)	October 1-February 1 Precipitation (Inches) (12)	Col. 12 x 100 ÷ Col. 11 October 1-February 1 Precipitation Index (13)	October-March 1 Precipitation (Inches) (20)	Col. 20 x 100 ÷ Col. 11 October 1-March 1 Precipitation Index (21)	October 1-April 1 Precipitation (Inches) (28)	Col. 28 x 100 ÷ Col. 11 October 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 80 ÷ Col. 30 May Precipitation Index (34)	June Precipitation (Inches) (35)	Col. 35 x 50 ÷ Col. 30 June Precipitation (Index) (36)
Huntington Lake	25.98							2.77						
Piedra	15.04							1.43						
Giant Forest	37.52							4.20						
Balch Powerhouse	24.46							2.74						

Total	(f)	
Number of Stations Used in (f)	(g)	
Date of Forecast Basin Precipitation Index (f ÷ g)	(h)	
Normal Date of Forecast to April 1 Precipitation Index Increase **	(i)	41
Date of Forecast Estimate of April 1 Precipitation Index (h + i)	(j)	

February 1

(p)	
(q)	
(p÷q)=(r)	
(s)	19
(r+s)=(t)	

March 1

(x)	
(y)	
(x÷y)=(z)	

April 1

Total	(aa)	
Number of stations used in (aa)	(bb)	
April Index (aa) ÷ (bb) =	(cc)	
May Index	(dd)	41
June Index	(ee)	12
April-June Precip. Index	(ff)	

May 1

(gg)	
(hh)	
(cc)	
(gg/hh)+(ii)	
(ee)	12
(jj)	

June 1

(kk)	
(ll)	
(cc)	
(ii)	
(kk/ll)+(mm)	
(nn)	

July 1

ANTECEDENT CONDITIONS INDEX Total April-July runoff from previous year in 10,000 acre-feet

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		

FIGURE 2

STATE OF CALIFORNIA
DEPARTMENT OF PUBLIC WORKS
DIVISION OF WATER RESOURCES

FORECAST COMPUTATION FORM
FOR
KINGS RIVER AT PIEDRA
APRIL-JULY UNIMPAIRED RUNOFF

August 1955

The forecast scheme also has inherent inaccuracy which must be added to the inaccuracies occasioned by weather. The step for this correction is:

(5) Add algebraically to the values determined in (4) the adjusted "hindcast" error for the same season as for the data randomly selected in step (2).

From these data it is possible to plot a frequency distribution of errors for given forecasts as depicted in Figure 4. It can be seen at a glance that the distribution of errors is non-normal. The interesting point of this distribution frequency plotting (see Figure 7) is that in the months prior to April 1 the forecast error distribution curves for 600,000 acre-feet and 1,300,000 acre-feet are coincident while the 2,000,000 acre-foot frequency distribution curve has a tendency to follow the other curves when the forecast is low, but diverges when the forecast is high in such a manner that the larger year has a smaller chance of having excessive errors. This came as a surprise to the authors as exactly the opposite effect was anticipated. However, there is a ready explanation for this anomaly. A plotting of precipitation and snowfall variations for the period February 1 through March 31 on frequency distribution paper (see Figure 5) and March precipitation and snowfall variations shows that snowpack index has a tendency to vary through the wider range. Referring back to Figure 1, it will be noted that the relationship between snowpack and October 1 - April 1 precipitation is such that snowpack index governs runoff to a greater extent during a low runoff year than in a high runoff year. This can be noted from the angle of precipitation index lines for the corresponding values of runoff. Horizontal placement of precipitation index line would mean zero influence, while vertical placement would mean zero snowpack influence. As precipitation index does not vary as widely as snowpack index, the decrease in variance probability of precipitation means that this will limit error to a greater degree when precipitation index has the greatest influence and thus explains the decreased probability of an excessively high forecast during a wet year.

Note that the variance of the precipitation indices from the assumed average index cannot be greater than the assumed index of 41 for the February 1-March 31 period and 19 for the March 1-March 31 period since the minimum possible precipitation between these dates would be zero. It is possible, however, for the variance of snowpack indicies to be greater than 35 and 9 for the two periods since it is possible to have a decrease in the water content of snow on the ground due to melt during these periods.

The probability levels for forecast departure from actual April-July runoff for the Kings River at Piedra have been plotted in Figure 7. It is to be noted that the future influence of the joint relationship of snowpack and early season precipitation index ceases as of April 1 and all curves are coincident thereafter. It will be further noted that the forecast errors of July 1 are slightly below the zero error line in the curve which is caused by slight inaccuracy of the drafting of Figure 1. A slight shift of the basin April-June index curves would correct for this bias although it was not thought warranted because of its small effect on the forecast. The June 1 forecast is almost as accurate as the July 1 forecast but does have slight distortion due to non-normal distribution of June precipitation. This distortion disappears as of July 1 since July precipitation is not included in the forecast scheme and all points should balance around the regression line.

The plotting in Figure 7 would enable an operator to utilize the graphical forecast scheme for the Kings River in the manner described in Mr. Blanchard's paper. The operator could use Figure 7 with added confidence because the forecast departure levels also include forecast scheme inaccuracies as well as the influence of possible future weather variation.

For example, an operator would use Figure 7 in the following manner: if a forecast of 1,650,000 acre-feet was made on February 1, it can be estimated from Figure 7 that there is one chance in five that the runoff would be less than 1,285,000 acre-feet (1,650,000 acre-feet less 365,000 acre-feet equals 1,285,000 acre-feet).

The authors wish to acknowledge the valuable assistance of Mr. Jack Hannaford in performing the many time-consuming computations necessary for development of the curves contained herein and in the final review of this paper.

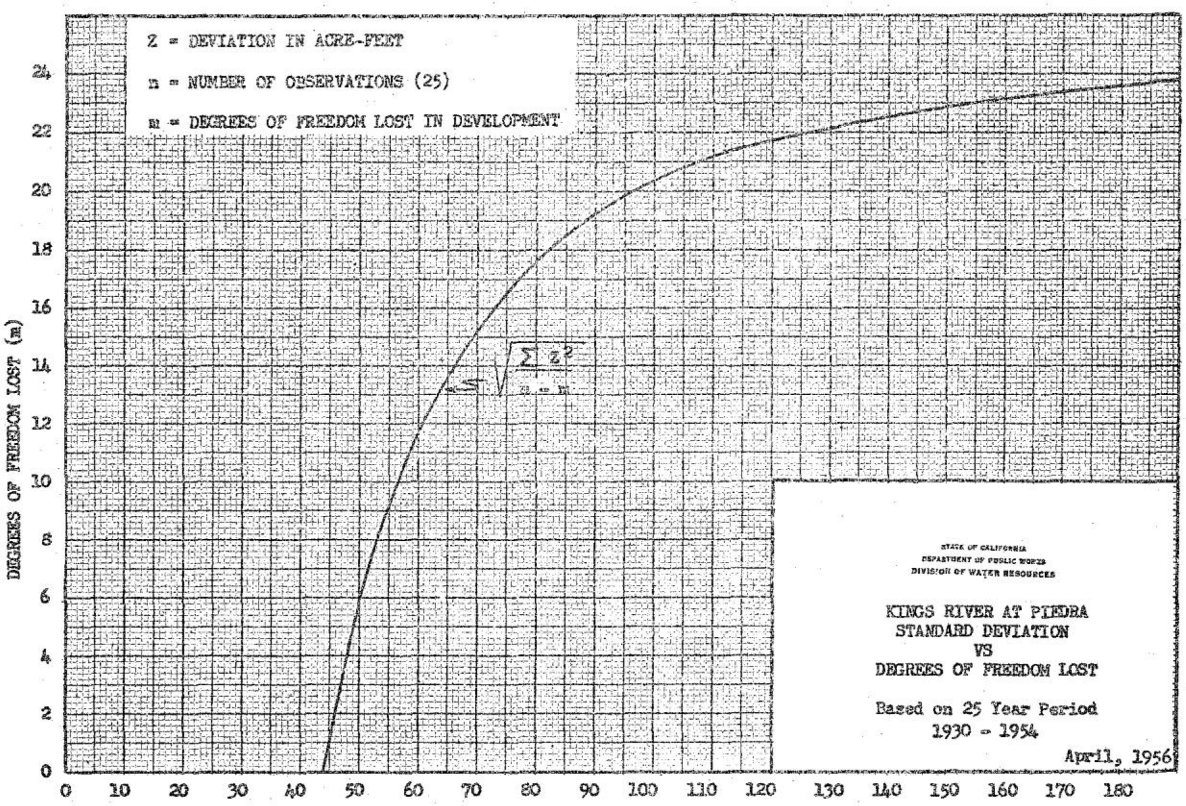


Fig. 3

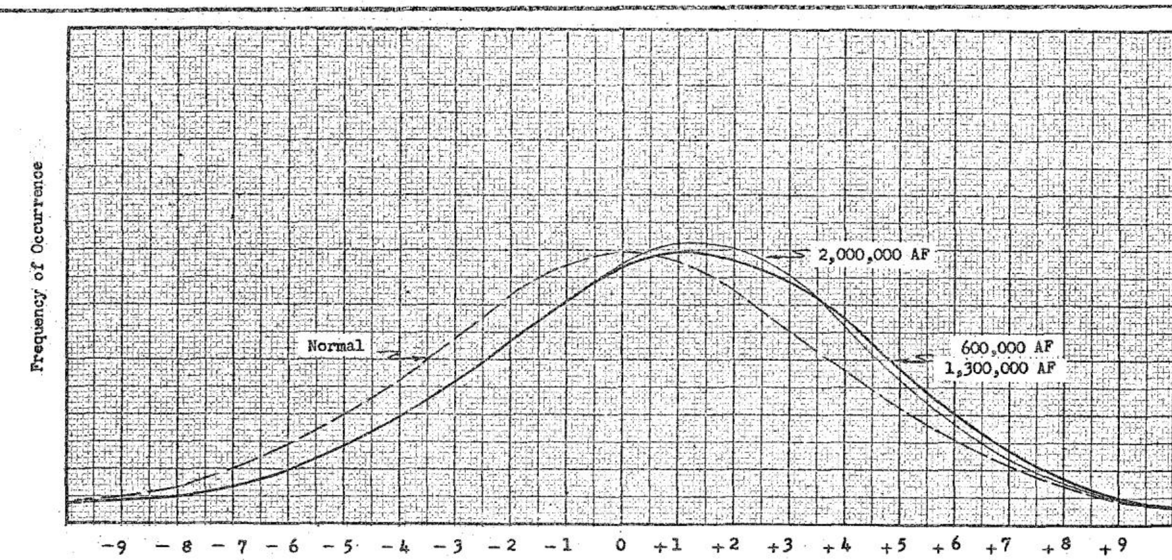


Fig. 4

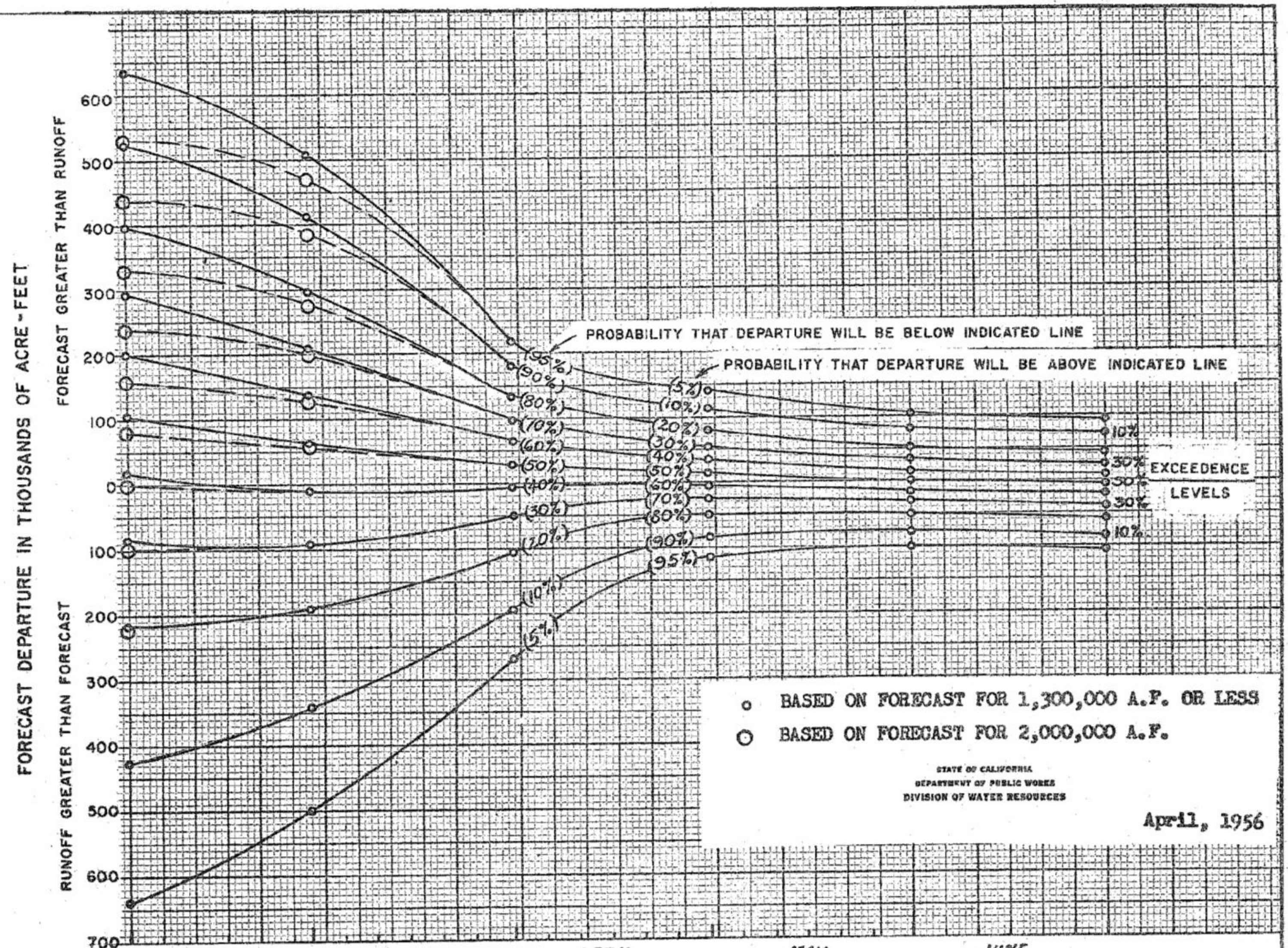
STATE OF CALIFORNIA
DEPARTMENT OF PUBLIC WORKS
DIVISION OF WATER RESOURCES

KINGS RIVER AT PIEDRA

FREQUENCY DISTRIBUTION CURVES
FEBRUARY FIRST FORECASTS

April, 1956

Figure 7



PROBABILITY LEVELS FOR DEPARTURE OF FORECAST FROM APRIL-JULY RUNOFF
KINGS RIVER AT PIEDRA