APPLICATION OF A NEW HYDROMETEOROLOGICAL 625-77

STREAMFLOW PREDICTION MODEL

By

Wendell V. Tangborn 1/

"The Lord made things simple to confound the wise" Corinthians I.

Introduction

The bases for this hydrometeorological streamflow prediction model (henceforth called the HM model) are these simple but rational premises:

1) That the catch of a comparatively low-altitude precipitation gage is proportional to the precipitation into a nearby mountain basin.

2) That a drainage basin's temporary water storage, which, in addition to snow, includes ground water and soil moisture, is approximately equal to cumulative precipitation (all forms) less runoff which has occurred during the precipitation season.

3) That runoff subsequent to the day for which storage is calculated is proportional to the amount of storage on that day.

4) That the prediction error in a short test season (before the main prediction season) is closely related to the error in calculating basin storage and that this relationship can be utilized to reduce the error in the main prediction season.

In this report three topics are emphasized: the test season approach to improve prediction accuracy, the method by which precipitation stations used in the model are selected and the application of the HM model as an operational tool. A complete description of the model has been given earlier (Tangborn and Rasmussen, 1976; Tangborn and Rasmussen, 1977). A brief summary of the derivation given in these two earlier reports is presented here to serve as a guide for those that may wish to test the HM model in other areas.

Foundation for the model

Assume that a prediction of runoff is to be made for the June 1-September 30 season (t_3-t_4) as in the example shown in figure 1. For this particular basin, it has been found that the optimum starting date for the winter season is October 1 (t_1) and that a one-month test season, or May in this case, is optimum for the prediction season (t3-t4) for this basin.

Storage (S_2) at the end of the winter season (t_2) is then approximately

$$S_2 = Pw-Rw \qquad (1)$$

where

P_w = winter-season basin precipitation (rain and snow)

R, = winter-season basin runoff

and at the end of the test season (t_3) storage (S_3) is approximately,

$$S_3 = Pw+Pt-Rw-Rt \qquad (2)$$

1/ Hydrologist, U.S. Geological Survey, Tacoma, WA where

P_t = test-season basin precipitation

R_t = test-season basin runoff.

However, at the end of the test season basin precipitation is unmeasured and assumed to be equal to some multiple of gage precipitation (or an average of several gages).

$$P_w + P_t = A_s(p_w + p_t) \tag{3}$$

where,

A_s = coefficient of gage to basin precipitation at the end of the test season,

p_w = winter-season gage precipitation.

and

p_t = test-season gage precipitation

In the example given in figure 1, the precipitation station used was Longmire, which is within the drainage basin and appears to closely represent total basin precipitation.

Therefore,

$$S_3 \simeq A_s(p_w+p_t)-R_w-R_t. \tag{4}$$

Summer (prediction season) runoff is approximately equal to the storage (S_3) plus summer precipitation less losses due to summer evaporation and transpiration.

$$R_s \simeq S_3 + P_s-L_s \tag{5}$$

where

R_s = summer-season basin runoff

P_s = summer-season basin precipitation

L_s = summer-season basin evaporation and transpiration losses.

Substituting equation 4 for S_3 in equation 5,

$$R_s = A_s(p_w+p_t)-R_w-R_t+P_s-L_s \tag{6}$$

The net effect of summer precipitation and evaporation is assumed to be constant.

$$B_s = P_s-L_s$$

where B_s is a constant.

Therefore,

$$R_s = A_s(p_w+p_t)-R_w-R_t+B_s \tag{7}$$

or

$$R_s+R_w+R_t = A_s(p_w+p_t)+B_s \tag{8}$$

resulting in a linear regression equation relating the total of annual runoff ($R_s+R_w+R_t$) to winter and test season precipitation (p_w+p_t) observed in a gage. Values for A_s and B_s can then be determined if at least 5 years of historical data are available for both runoff and precipitation and an equation to predict summer-season runoff results,

$$R_s* = A_s(p_w+p_t)-R_w-R_t+B_s \quad (9)$$

where R_s* equals the predicted runoff during the main prediction season.

Test season approach

The basin water storage at t_2 (fig. 1) is approximately:

$$S_2 \simeq A_t p_w - R_w \quad (10)$$

where A_t equals a coefficient of gage to basin precipitation at the end of the winter period. Test season runoff is then proportional to storage at t_2 plus precipitation less evaporative losses.

$$R_t \simeq S_2+P_t-L_t \quad (11)$$

where L_t = precipitation less evaporation during the test season.

$$R_t = A_t p_w - R_w + p_t - L_t \quad (12)$$

where $(P_t - L_t)$ is assumed to be a constant.

$$B_t = P_t - L_t$$

thus

$$R_t+R_w = A_t p_w + B_t \quad (13)$$

which is similar to equation 8 and relates winter and test season runoff to winter precipitation. An equation to predict test season runoff can then be formed.

$$R_t* = A_t p_w - R_w + B_t \quad (14)$$

where R_t* is the predicted runoff during the test season.

The error for the test season (e_t) is given by:

$$e_t = R_t*-R_t. \quad (15)$$

The error for the prediction season (e_s) is:

$$e_s = R_s*-R_s. \quad (16)$$

For most seasons and basins tested thus far with the HM model, a significant correlation is observed between the test season and prediction season errors (fig. 2). This relationship seems to imply that the cause of these errors are the same, namely an error in the estimate of basin storage.

When the relationship between e_s and e_t is determined using historical observations of runoff and precipitation, then the estimated prediction season error is

$$e_s* = c\ e_t \quad (17)$$

where (by regression) $$c = \frac{\Sigma e_t e_s}{\Sigma e_t^2} \quad (18)$$

and e_s* equals predicted summer season error (fig. 2).

It is possible to improve the prediction season forecast by reducing the initial prediction by the estimate of e_s*.

$$R_s** = R_s*-e_s* \quad (19)$$

or

$$R_s^{**} = R_s^{*} - ce_t \tag{20}$$

where R_s^{**} is the revised prediction.

In figure 3 the results of improvement in prediction accuracy as a function of prediction date is shown for several drainages in widely varied environments. The improvement, I, is the reduction in the standard error of estimate resulting from predictions made over a 10 to 15 year period,

$$I = \frac{SE_{(o)} - SE_{(1)}}{SE_{(o)}}$$

where

$SE_{(o)}$ = Standard error without a test season

$SE_{(1)}$ = Standard error with a 1 month test season

It is apparent from figure 3 that the inclusion of a one-month test-season in the prediction model reduces the standard error for a May 1 to July 31 prediction made on May 1 by as much as 50 percent. For predictions on March 1 or earlier, the test season reduces the standard error by less than 12 percent. One reason for this significant improvement is believed to be the difference in the areal coverage of snow between mid-winter and spring. That is, the test season revises the estimate of the amount of snow available for ablation. However, when the landscape is nearly covered as it usually is on January 1, differences in the volume of snow are difficult to detect. As spring approaches and sizable patches of bare ground appear, snow cover will then control how much runoff from snowmelt will occur. Thus, a heavy snowpack will have a higher ablation potential, because of a greater areal coverage and will produce more runoff than would a sparce snowpack that covered less area. This would explain the difference in the value of the test season between different basins (a more rugged topography would logically show this effect more than a level one). There is evidence that the test season improves accuracy more in some years than in others because of differences in snow distribution from year to year. It then follows that the incorporation of an areal measurement of snow cover (for example by satellite imagery) may strengthen the test-season approach.

Another possible reason for the increased value of the test season as summer approaches is that there is much less snowmelt runoff during midwinter as compared to spring and summer; thus, an early season prediction error will not reveal much information on basin snow storage.

Improvements in the test season approach

Several possible means of producing greater prediction accuracy through the test season updating process are apparent. One obvious way would be to reduce the variance in the test season prediction error by adding temperature as a variable--both to account for runoff caused by snow ablation and for the effect of the freezing level during storms. For example, a low freezing level would produce more precipitation as snow; consequently predicted runoff would be less than if temperature were included in a predictive equation.

Another possibility would be to add another even earlier test season, which would be used to reduce the error in the initial test season (this suggests that even more test seasons could be added, each one utilized to reduce the prediction error in the subsequent one).

However, before any of the above ideas are examined in detail both runoff and precipitation should be applied at much shorter time intervals than the 1-month increments used thus far with this model. Shortening the test season length would also be possible if this were done. It is obvious that because of the immense quantities of data involved in such an analysis, all of these methods would need to be automated.

Selection of precipitation stations

Selecting precipitation stations in or near the drainage for which predictions are needed and whose catches appear to represent basin precipitation is a difficult task. Most weather stations with a sufficiently long record are at a particular location for such reasons as the availability of someone who will act as an observer. In addition, stations are often moved or abandoned for unavoidable reasons. Consequently, it is necessary to use one or more of the established stations that apparently represent the weather occurring over large areas and, in mountainous regions, over large altitude differences.

The precipitation stations are selected for the HM model simply on the basis of how well they serve when used to predict runoff on a year-by-year basis. In figure 4, the Nisqually River near National (#12-0825) is shown as an example of this procedure. Fourteen different National Weather Service stations were used, two within the basin (Longmire and Paradise), the remaining were from 2 to 150 kilometers from the Nisqually River basin (fig. 5). Predictions were made for each May-July season during the 1966-75 period. For example, the coefficients determined by the 1952-65 period were used to calculate the 1966 runoff, those of 1952-66, to calculate that of 1967 and so on.

When an average of two or more of these stations is made, a significant improvement in accuracy results. The lower curve in figure 4 shows the reduction in the standard error when the first two stations, Packwood and Longmire are averaged, and for each consecutive set of stations up to an average which uses all 14 stations. The station averages are weighted equally by the equation

$$p_n = \frac{p_1 + p_2 ---- p_n}{n}$$

where P_1, P_2-----P_n is precipitation for stations, 1, 2-----n, ranked in order of increasing standard error of estimation. The results given here also indicate that the average of 3 or 4 stations gives the lowest standard error. This may be due simply to the lack of representative stations and that if more low-error stations were available, averages of these would be even more beneficial.

One conclusion gathered from these results is that additional precipitation gages placed within or very near the drainage would likely improve the estimate of basin precipitation. It should be noted that high altitude stations, such as Paradise on Mount Rainier in this example, are not necessarily the most representative. The reason for this is believed to be the difficulty in catching precipitation as snow at these exposed, windy sites.

Prior tests

Application of the HM model to predict seasonal streamflow in three distinctly different regions, the Sierra Nevada, the Cascades and in northern Norway, indicates that the accuracy usually is superior to the current snow survey system. It should be emphasized that the HM predictions are based only on knowledge up to the date of the predictions.

There are three possible reasons for the apparent lower prediction error produced by this model.

1) Basin precipitation may be better represented by a low-altitude precipitation gage than by a higher altitude snow course.

2) A test season is used to revise the prediction.

3) The model implicitly incorporates soil moisture and ground water as part of basin storage.

All three factors may contribute to improving overall accuracy in some of the basins tested. The test season could, of course, also be applied to the snow survey method. Some hydrologic models do account for soil moisture and ground water storage so that the third factor is not significant in these instances. It appears that the first factor is valid for many drainages because the accuracy is improved by this method even when a test season is not used, e.g., on January 1.

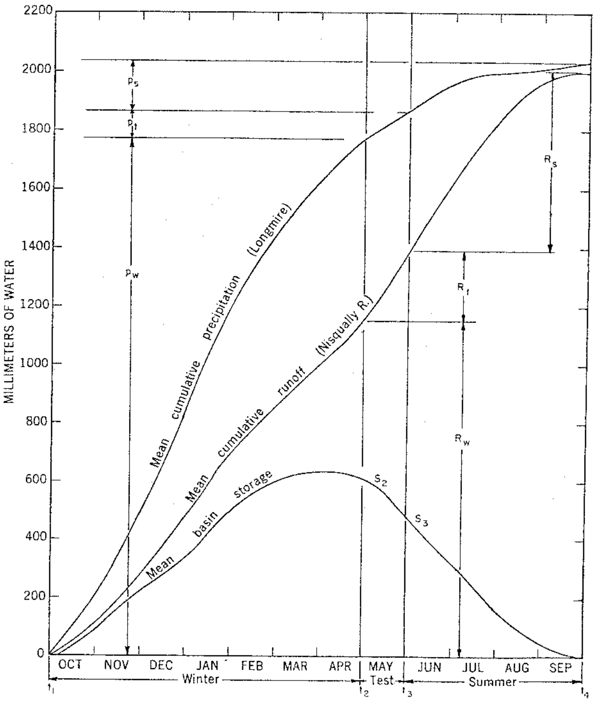

FIGURE 1.--Mean water storage as snow, cumulative runoff from and approximate precipitation in the Nisqually River drainage basin. This diagram demonstrates the basic ideas of the HM model - namely, the test season and the determination of basin storage from gage precipitation and runoff.

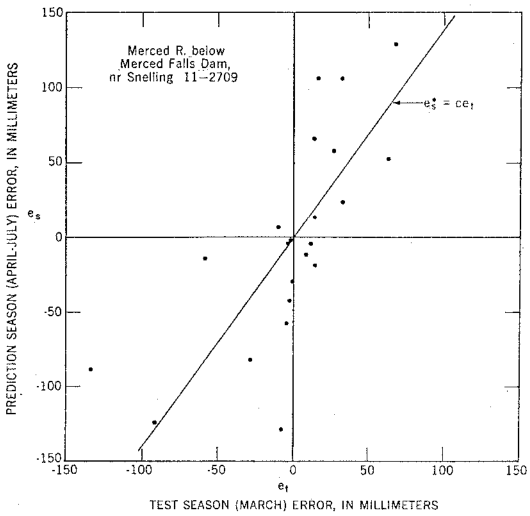

FIGURE 2.--The relationship between prediction season (April-July) and the test season (March) errors for the Merced River below Merced Falls Dam, California. The linear function determined by the fit between these errors is used to reduce the final prediction error. The correlation coefficient for this sample is 0.79.

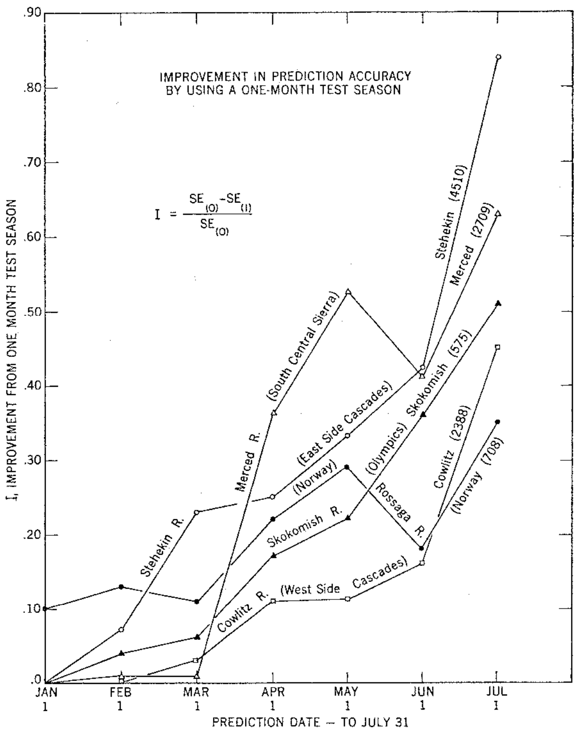

FIGURE 3.--Improvement in accuracy from using a 1-month test season, as a function of prediction day. The five drainage basins shown here are in widely different environments.

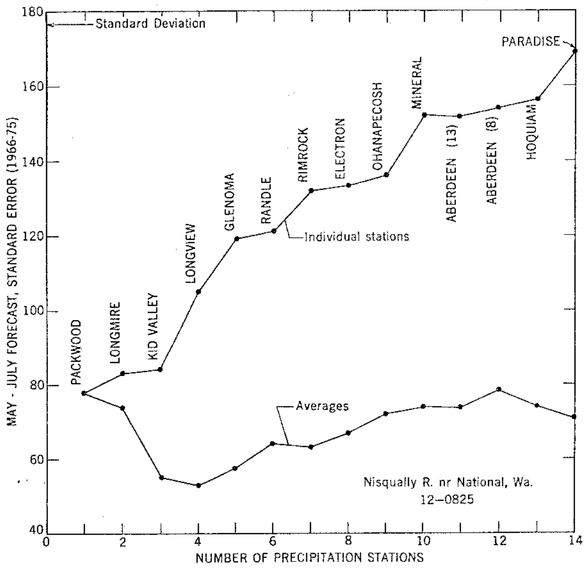

FIGURE 4.--Prediction accuracy of the Nisqually River of a May-July season using data from different weather stations (upper curve) in southwest Washington (see fig. 5) in the model. Averages of these (lower curve) were made by equal weighting each station in consecutively ranked order; for example, the first average is the two highest ranked stations, the second average is the top three stations and so on up to an average of all 14 stations.

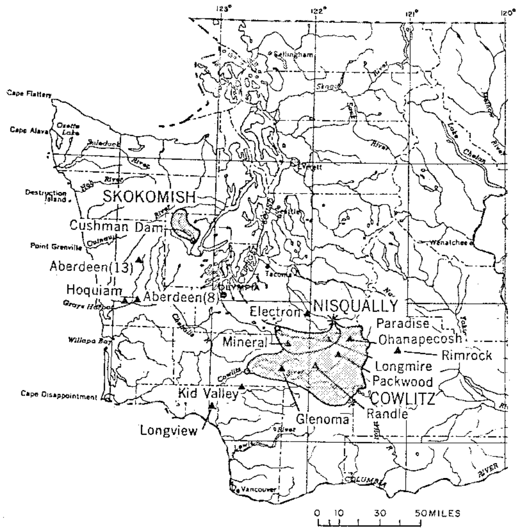

FIGURE 5.--Western Washington showing the three drainages used by Tacoma City Light for hydroelectric power generation. Also shown are the precipitation stations tested for use in the model to predict runoff of the Nisqually River.

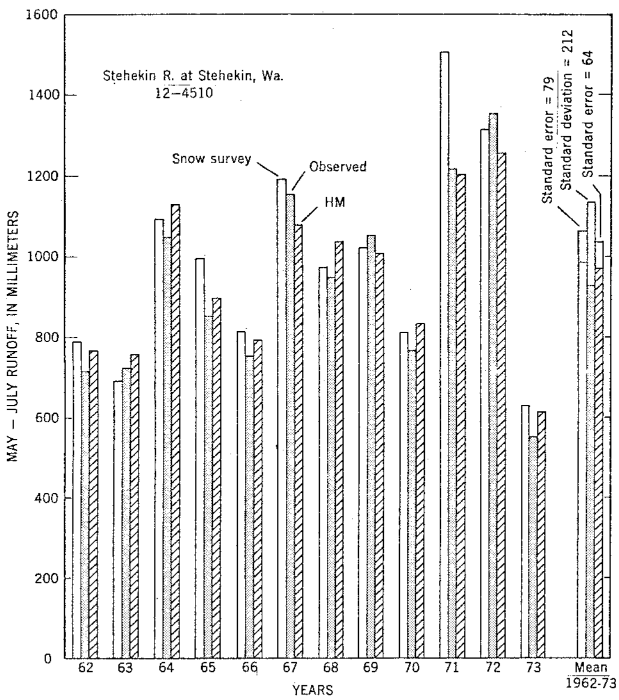

FIGURE 7.--Prediction comparisons for the Stehekin River near Stehekin, eastern Washington, for the May-July season.

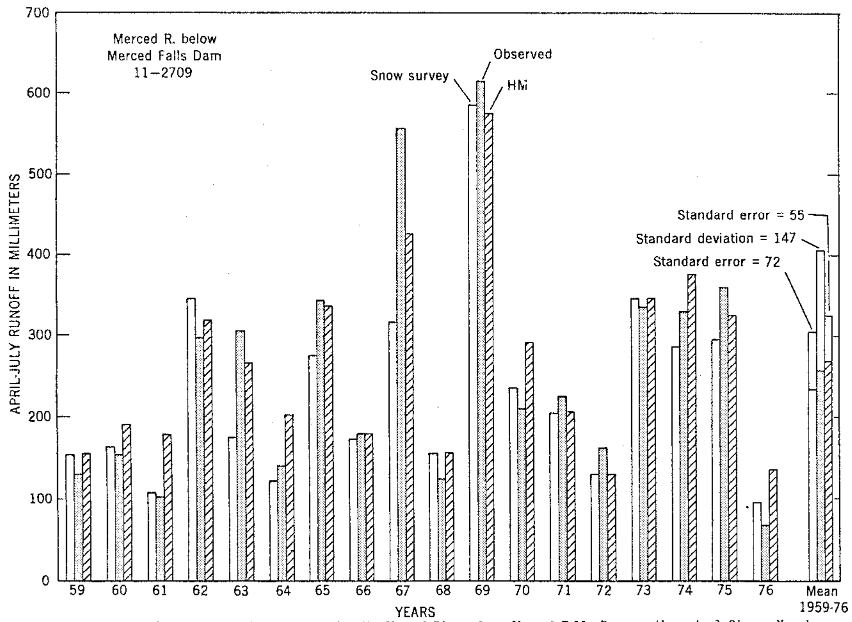

FIGURE 6.--Prediction comparisons for the Merced River above Merced Falls Dam, south-central Sierra Nevada, California. The snow survey method (clear) and the HM model (hatched) predictions are shown with the observed April-July runoff (solid).

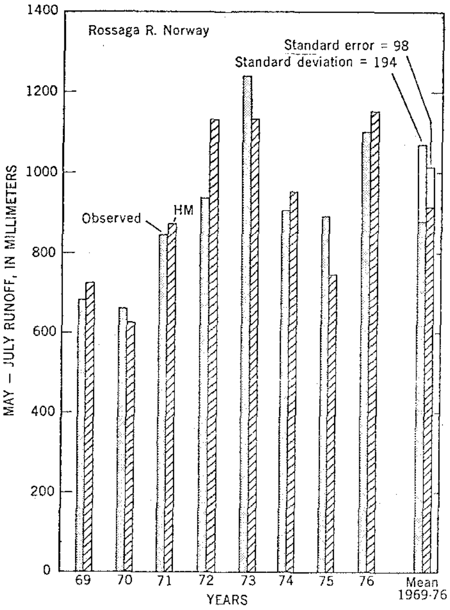

FIGURE 8.--Predictions for the Røssaga River in northern Norway (near the Arctic Circle), part of a large hydroelectric development. The prediction season in this example is May-July.

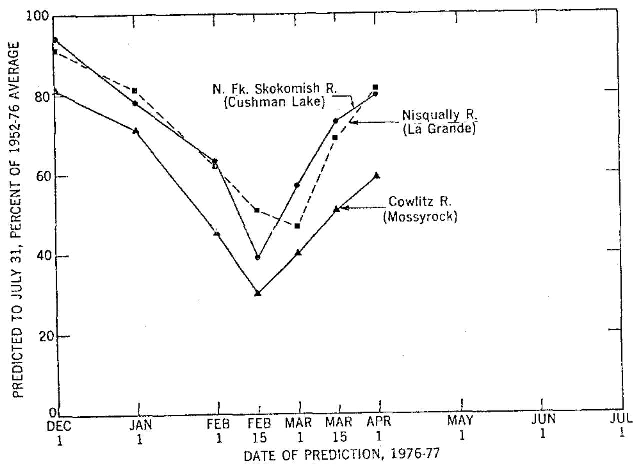

FIGURE 9.--Predictions of the Cowlitz, Nisqually and Skokomish River, Washington during the 1977 season, given as percent of the 1952-76 average runoff for each of these seasons.

TABLE 1.--Calculation of estimated April-July runoff with the HM model, in millimeters. Merced River below Merced Falls Dam, near Snelling, California (11-2709). Precipitation is an average of Balch Power House, Grant Grove and North Fork Ranger Station

	$R_t^* = A_tP_w + B_t - R_w$ Test season prediction						$R_s^* = A_s(P_w+P_t) + B_s - (R_w+R_t)$ Initial prediction						$e_s^* = c \cdot e_t$ $R_s^{**} = R_s^* - e_s^*$ Final prediction				Observed
Year	A_t	P_w	B_t	R_w	R_t^*	e_t	A_s	P_w+P_t	B_s	R_w+R_t	R_s^*	e_s	c	e_s^*	R_s^{**}	e_s^{**}	R_s
1959	0.386	432	-112	41	14	-11	0.695	439	-97	66	138	9	1.457	-16	154	25	129
60	.382	360	-108	35	-5	-32	.689	445	-101	62	143	-11	1.456	-47	190	36	154
61	.366	329	-95	24	2	-11	.695	434	-97	37	163	60	1.410	-16	179	76	103
62	.361	781	-91	82	109	76	.704	932	-116	115	425	129	1.306	105	319	23	296
63	.336	515	-83	105	-15	-42	.671	655	-104	132	203	-101	1.501	-63	266	-38	304
64	.332	363	-77	48	-5	-18	.665	490	-92	61	173	34	1.587	-29	201	62	139
65	.326	745	-73	229	-59	-90	.672	826	-100	260	196	-146	1.548	-139	335	-7	342
66	.349	549	-79	87	25	-4	.692	565	-103	116	172	-7	1.562	-6	178	-1	179
67	.348	772	-79	111	79	4	.691	1,041	-102	186	432	-125	1.561	6	425	-132	557
68	.340	336	-79	41	-3	-24	.731	413	-122	62	118	-5	1.575	-38	155	32	123
69	.297	1,480	-73	280	151	-78	.729	1,603	-121	353	695	81	1.557	121	574	-40	614
70	.296	575	-51	132	-12	-61	.693	687	-98	181	197	-12	1.532	-93	290	81	209
71	.296	540	-46	73	40	13	.693	605	-97	100	222	-4	1.387	18	203	-23	226
72	.296	453	-47	50	37	1	.692	453	-97	86	131	-30	1.377	1	130	-31	161
73	.292	819	-47	109	86	34	.688	1,025	-93	161	452	117	1.374	47	405	70	335
74	.293	667	-46	110	38	-21	.672	898	-86	169	349	-19	1.436	-30	378	48	330
75	.290	493	-46	70	28	-30	.671	741	-86	128	283	-78	1.406	-42	325	-36	361
76	.291	331	-43	38	15	1	.671	405	-83	52	137	68	1.438	1	136	67	69
Mean		603		94		-8		730		131	257	0			269	12	257
Standard deviation																	147
Standard error of estimate											75				55		

P_w = winter precipitation R_w = winter runoff e_t = test prediction error R_s^{**} = final prediction
P_t = test precipitation R_t = test runoff e_s = summer prediction error R_s^* = initial prediction

TABLE II.--Comparison of predictions between the Tacoma City Light snow survey method and the HM model (1970-76 prediction period), standard error in millimeters

Prediction season	Skokomish R. (Cushman) SE TCL	HM	Nisqually R. (Alder) SE TCL	HM	Cowlitz R. (Mossyrock) SE TCL	HM	Cowlitz R. (Mayfield) SE TCL	HM
Jan-July	500	268	418	331	464	357	498	369
I		.46		.21		.23		.26
Feb-July	443	249	266	186	285	205	300	225
I		.44		.30		.28		.25
Mar-July	376	123	258	99	233	120	257	132
I		.67		.62		.48		.49
Apr-July	106	70	77	73	79	76	100	83
I		.34		.05		.04		.17
May-July	173	132	65	63	68	50	75	47
I		.24		.03		.26		.37
June-July	122	98	51	45	61	51	51	38
I		.20		.12		.16		.25
July	63	33	17	18	33	29	27	23
I		.48		-.06		.12		.15

I = prediction improvement

$$= \frac{SE_{(TCL)} - SE_{(HM)}}{SE_{(TCL)}}$$

Examples of how the predicted runoff is calculated by the HM model for the April-July season, for the Merced River below Merced Falls Dam, are shown in table 1. For each prediction, precipitation and runoff data used in the computation are given along with the predictive coefficients. These coefficients were determined from a linear fit between runoff and precipitation from the beginning of the period (1952) up through the year previous to the prediction. In this way the coefficients are revised each year as that year's data becomes available.

In figures 6 and 7 comparisons of the HM model predictions with those of operational snow surveys are shown for basins in the Eastern Cascades of Washington and the Western Sierra Nevada of California. These results are fairly typical for each of these regions. In figure 8 the results of May-July predictions are shown for the Ri6ssaga River in northern Norway.

1977 operational test

Using the HM model, operational predictions of inflow to four hydroelectric reservoirs in Washington were initiated on January 1, 1977. These plants are operated by the Tacoma City Light public utility. A comparison of accuracy between the existing snow survey system and the HM model over the past 7 years, for all seasons between January 1 and July 1, showed an average improvement in accuracy of about 28 percent (table II).

Precipitation and runoff data used for these predictions are collected on a continuous basis. Digital streamflow records are gathered from the gaging stations toward the end of each month and immediately processed as daily values for use in the model. These stations are also on a real-time network so that the records can be brought up to date at any time. Daily precipitation observations, recorded by National Weather Service observers at Longmire, Packwood, Kid Valley, Glenoma and Lake Cushman, are mailed once a week on a postcard designed for this purpose or, if needed for midmonth or extra predictions, phoned directly to the U.S. Geological Survey office in Tacoma. Thus, with these data, it is possible to make a runoff prediction at any time, a definite advantage in some water-management situations. So far this winter predictions have been made every two weeks. These results, given in figure 9, show the effect of extremely low precipitation in the Cascades and Olympics this past winter. It is, of course, yet too early to assess the accuracy of this first test.

REFERENCES

Tangborn, Wendell V., and Rasmussen, Lowell A., 1976, Hydrology of the North Cascades, Part 2. A proposed hydrometeorological streamflow prediction model: Water Resources Research, v. 12, no. 2, p. 203-216.

-----1977, Application of a hydrometeorological model to the south-central Sierra Nevada at California: Journal of Research, U.S. Geological Survey, v. 5, no. 1, p. 33-48.