FREQUENCY DISTRIBUTION OF ANNUAL PEAK WATER-EQUIVALENT ON WYOMING SNOW COURSES

By

Ronald D. Tabler[1/]

INTRODUCTION

The first snow courses in Wyoming were established in 1919, about 13 years after Dr. James Church established the first western snow course on Mt. Rose. The snow course network was expanded after the drought of 1934, and now provides an historical record that will find many applications in addition to the original objective of providing a basis for forecasting streamflow. Information on the yearly variation of annual peak water equivalent is useful for designing water development and flood control projects, particularly in areas having insufficient streamflow data.

Probabilities associated with snowpack variables have not been studied as extensively as have other hydrologic variables, but future improvements in mathematical models for snowmelt runoff will require more effort in this area. Examples of past applications of frequency distributions related to snow include snow load computations (Brown, 1970; George and McAndrew, 1973), streamflow forecasting (Twedt et al., 1977), flood and drought planning (Doesken et al., 1981), avalanche hazard evaluation (LaChapelle, 1966; Armstrong, 1981), and snowpack augmentation (Tabler, 1968; Engelen, 1972). The study reported here was prompted by a need to determine if storage capacity of snow fence systems to protect highways in Wyoming should be designed for a winter with average precipitation, or if some other design year might improve the benefit/cost ratio. This problem is typical of those for which probability data are needed, while raising the additional question of how much variation might be expected in empirical frequency distributions over areas as large as a state.

DESCRIPTION OF DATA

Definition and Data Source.

Annual peak water-equivalent is defined for this study as the maximum water content reported on a snow course each winter, without regard to date. Because the snow courses are measured only monthly (or biweekly in a few cases), actual peak water-equivalents may differ somewhat from reported values. The data set consisted of the annual peak water-equivalent observed on each snow course in Wyoming, both active and inactive, as published by the Soil Conservation Service in Summaries of Snow Survey Measurements, and Water Supply Outlook for Wyoming, for the period 1919 to 1981, inclusive. These values were recorded on a magnetic disk and were analyzed with a 23,000-byte desk-top computer.

Characteristics of Wyoming Snow Courses.

Characteristics of the 172 snow courses with published records are shown in figure 1. Mean length of record is about 29 years, with 8 snow courses having more than 60 years of records. Average elevation is about 2530 m. Annual peak water-equivalent averaged for all courses is about 430 mm, with one course having a mean of 1500 mm. Locations of snow courses within the state are shown in figure 2.

Presented at the Western Snow Conference, April 20-23, 1982; Reno, Nevada

1/ Hydrologist, USDA Forest Service, Rocky Mountain Forest and Range Experiment Station, Laramie, in cooperation with the University of Wyoming. Station headquarters is at Fort Collins, in cooperation with Colorado State University.

"Reprinted Western Snow Conference 1982"

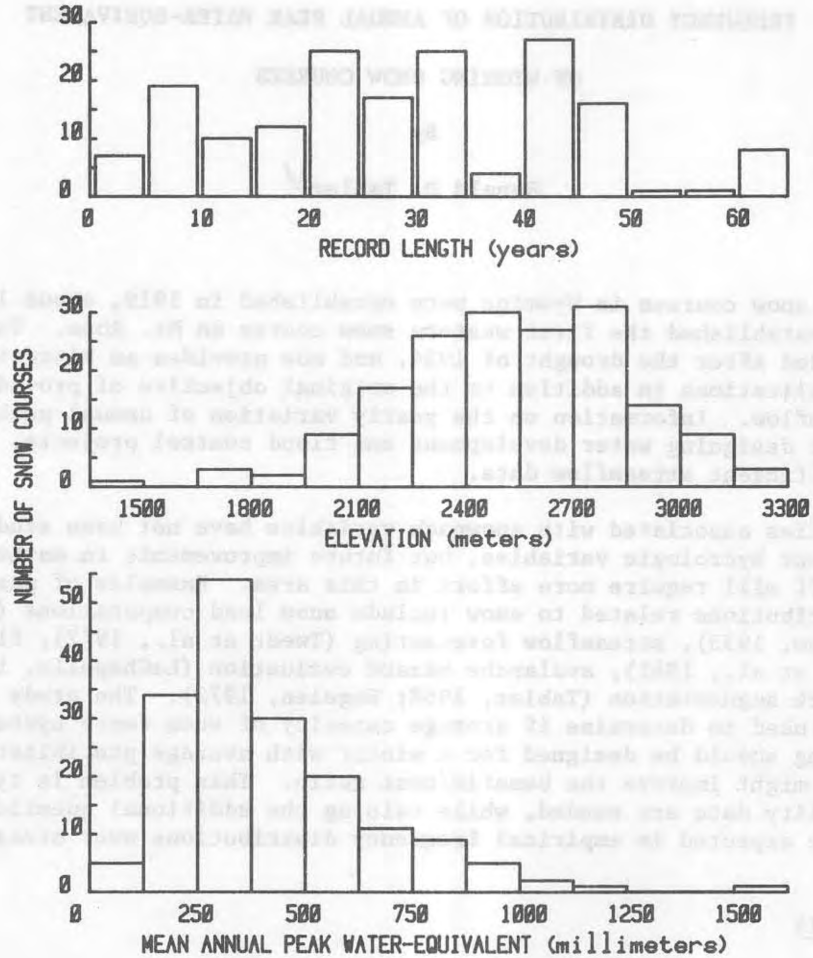

Figure 1. Characteristics of all 172 Wyoming snow courses with published records, as of 1981.

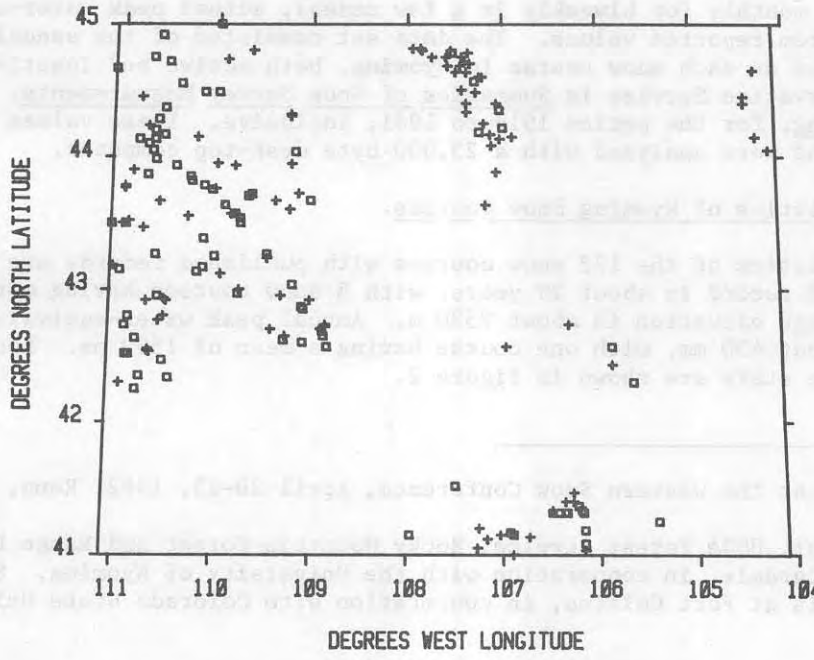

Figure 2. Location of snow courses in Wyoming. Squares denote stations having 30 or more years of record.

Restriction on Record Length.

Frequency analyses were arbitrarily restricted to the 82 courses having 30 or more years of data, consistent with common practice. This set has about the same mean elevation and water-equivalent, with an average record length of 43 years. The relationship of sample variance to record length (Fig. 3) suggests that the minimum could have been reduced to 20 years.

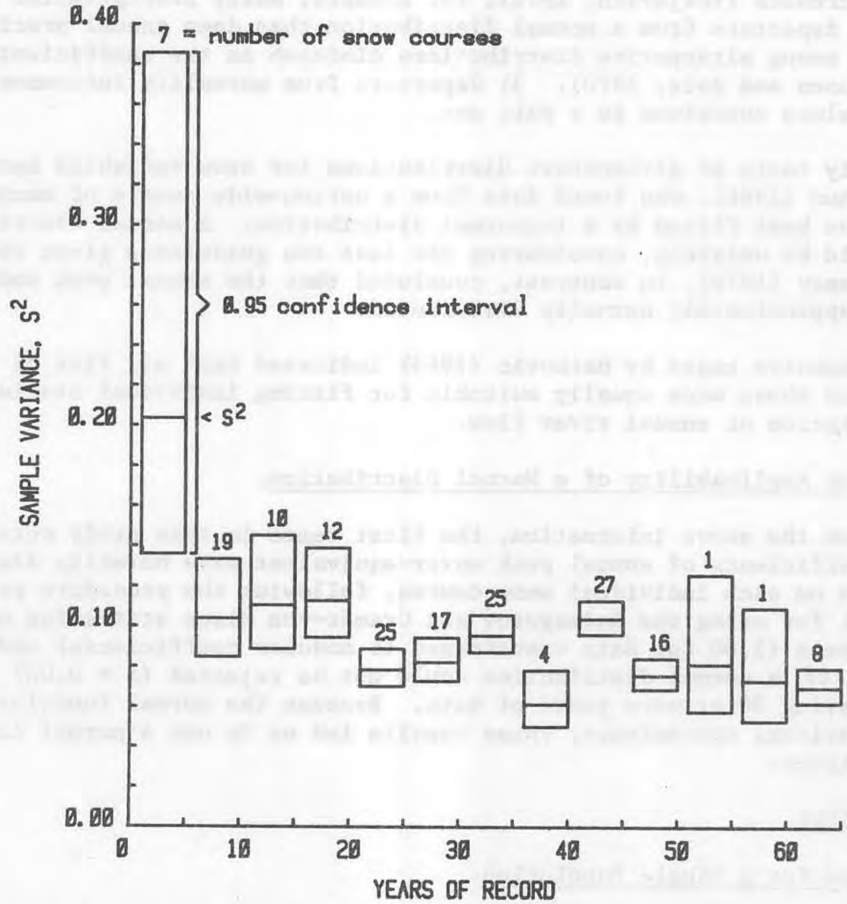

Figure 3. Sample variance of annual peak water-equivalent on Wyoming snow courses, in relation to years of record.

Transformation of Data to Modular Coefficients.

To allow data from different snow courses to be compared and combined for grouped analyses, each observed value of annual peak water-equivalent (W_i) was transformed into a modular coefficient, K_i, by dividing by the long-term station mean ($\overline{W}$)

$$K_i = \frac{W_i}{\overline{W}} \qquad (1)$$

where the subscript i refers to a particular year. As an example, the modular coefficient of an annual peak water-equivalent of 500 mm on a snow course having a long-term mean of 400 mm, would be 500/400 = 1.25.

The use of modular coefficients for frequency analysis of hydrologic variables has been noted by Markovic (1965) and Yevjevich (1972). Considering the utility of dimensionless frequency distributions, it is surprising that modular coefficients have not been used more in analyzing snowpack variables. Examples known to the author are Rakhmanov's (1969) use of modular coefficients to describe spatial variability of snow water-equivalents on selected drainage basins in the USSR, and an analysis of annual variability of April 1 water-equivalents for selected drainages in the Colorado River Basin (Doesken et al., 1981).

SELECTING A DISTRIBUTION FUNCTION

Probability functions most often used for hydrologic variables are the normal, two- and three-parameter lognormal, and two- and three-parameter gamma functions. The primary basis for selecting a particular function is how well it fits the data, with the following guidelines helping to indicate the more likely choices. 1) In general, a frequency curve for a hydrologic variable approaches a normal distribution more closely as the period of observation increases (Yevjevich, 1972); for example, daily precipitation typically exhibits a greater departure from a normal distribution than does annual precipitation. 2) Differences among alternative distributions diminish as the coefficient of variation decreases (Johnson and Kotz, 1970). 3) Departure from normality increases with the proportion of zero values contained in a data set.

The only tests of alternative distributions for snow variables known to the author are those by Thom (1966), who found data from a nation-wide sample of maximum annual water-equivalent to be best fitted by a lognormal distribution. A normal distribution for this population would be unlikely, considering the last two guidelines given above. Engelen (1972) and Kopanev (1976), in contrast, concluded that the annual peak snow depths they analyzed were approximately normally distributed.

Comprehensive tests by Markovic (1965) indicated that all five of the probability functions listed above were equally suitable for fitting individual station samples of annual precipitation or annual river flow.

Test for Applicability of a Normal Distribution.

Based on the above information, the first tests in this study were to determine if the modular coefficients of annual peak water-equivalent were normally distributed. These tests were made on each individual snow course, following the procedure proposed by Stephens (1974) for using the Kolmogorov and Cramér-von Mises statistics modified for the case of known mean (1.00 for data transformed to modular coefficients) and unknown variance. The hypothesis of a normal distribution could not be rejected ($\alpha = 0.05$) for any of the 82 snow courses having 30 or more years of data. Because the normal function offers matematical and computational convenience, these results led us to use a normal function for all subsequent analyses.

COMPOSITE ANALYSES

Evidence for a Single Population.

Plots of the individual snow course distributions tended to exhibit about the same variance, as demonstrated by the scatter diagram of exceedance probability (Fig. 4). This suggested the possibility of fitting a single curve to all of the snow courses, and prompted tests of the homogeneity of variance among elevation and location classes (Tables 1 and 2). Using Bartlett's test (Snedecor, 1956), the hypothesis of equal variance would not be rejected ($\alpha = 0.25$) for the four elevation classes from 2130 to 3030 m. Although the variance for snow courses at elevations greater than 3030 m does not appear to come from the same population, this is dismissed as a possible artifact of the relatively small sample size within this elevation class because of the absence of a general trend in the rest of the data.

To test the effect of geographic location, the State was divided into approximately equal quadrants, with the latitudinal division across the center of the State at 43° N, and the longitudinal division at 108° W. Again using Bartlett's test, the hypothesis of homogeneous variance would not be rejected.

These results provided the justification for developing a single frequency distribution applicable to all snow courses in Wyoming.

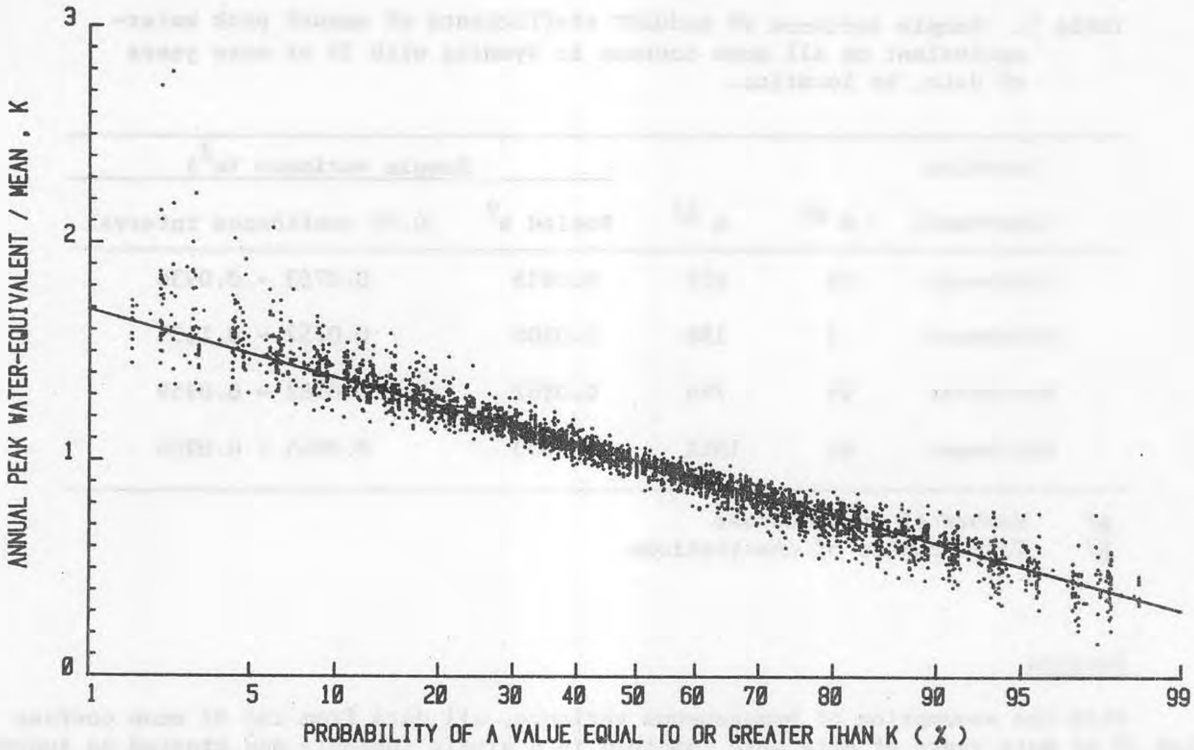

Figure 4. Frequency distributions of modular coefficients for all Wyoming snow courses having at least 30 years of data, plotted individually. Plotting positions on the normal probability scale are 100 m / (n + 1), where m is rank in order of increasing magnitude, and n is number of observations (Gumbel, 1956). The fitted line is equation (3) with $s^2 = 0.088$.

Table 1. Sample variance for modular coefficients of annual peak water-equivalent on all snow courses in Wyoming with 30 or more years of data, by elevation class.

			Sample variance (s^2)	
Elevation m	N a/	n b/	Pooled s^2	0.95 confidence interval
< 2130	11	554	0.0855	0.0764 - 0.0968
2130 - 2429.9	26	1185	0.0911	0.0843 - 0.0991
2430 - 2729.9	26	1032	0.0937	0.0862 - 0.1024
2730 - 3029.9	14	540	0.0823	0.0735 - 0.0934
> 3029.9	5	183	0.0596	0.0494 - 0.0746

a/ Number of snow courses
b/ Total number of observations

Table 2. Sample variance of modular coefficients of annual peak water-equivalent on all snow courses in Wyoming with 30 or more years of data, by location.

Location (Quadrant)	N [a/]	n [b/]	Sample variance (s^2) Pooled s^2	0.95 confidence interval
Southeast	16	651	0.0835	0.0753 - 0.0936
Northeast	5	186	0.0906	0.0752 - 0.1133
Southwest	19	746	0.0862	0.0782 - 0.0959
Northwest	42	1911	0.0900	0.0845 - 0.0960

a/ Number of snow courses
b/ Total number of observations

Results.

With the assumption of homogeneous variance, all data from the 82 snow courses having 30 or more years of data were combined in a single ensemble and treated as independent samples from a single population. The data are plotted in figure 5, with the normal curve fitted from the computed sample variance.

The probability density function is given by

$$f(K) = \frac{1}{s\sqrt{(2\pi)}}\, e^{-(K-1)^2/2s^2} \tag{2}$$

and the probability distribution is

$$F(K) = \frac{1}{s\sqrt{(2\pi)}} \int_{-\infty}^{k} e^{-(K-1)^2/2s^2}\, dK \tag{3}$$

where K is the modular coefficient defined by equation (1) and the variance, s^2, is 0.088. Confidence intervals for the variance, and other statistical parameters, are presented in Table 3. The hypothesis of a normal distribution would not be rejected at $\alpha = 0.15$ using either the Kolmogorov or Cramér-von Mises statistics and applying the critical values given by Stephens (1974).

Discussion.

Markovic (1965) employed a similar ensemble analysis for annual precipitation and annual streamflow data from stations throughout the western United States and southwestern Canada. For the precipitation data having homogeneous variance, there was little difference in the goodness of fit provided by either the normal or lognormal-2 functions; for the normal distribution, sample variance was 0.072 and the coefficient of skewness was -0.370. Markovic's value (0.084) for the variance of annual streamflow for stations with homogeneous variance is even more remarkable for its similarity to the 0.088 value obtained from this study of annual peak water-equivalent. Such a coincidence could even lead one to consider the possibility of a "universal" approximation for "normally" distributed hydrologic variables.

Estimates provided by equation (3) with $s^2 = 0.088$, appear to approximate the average of curves presented by Doesken et al. (1981) for April 1 water-equivalent in the Colorado mountains. Although these authors used the median for their transformation to dimensionless coefficients, personal communication with one of the authors (Shafer) confirmed that the mean and median were approximately equal for their data.

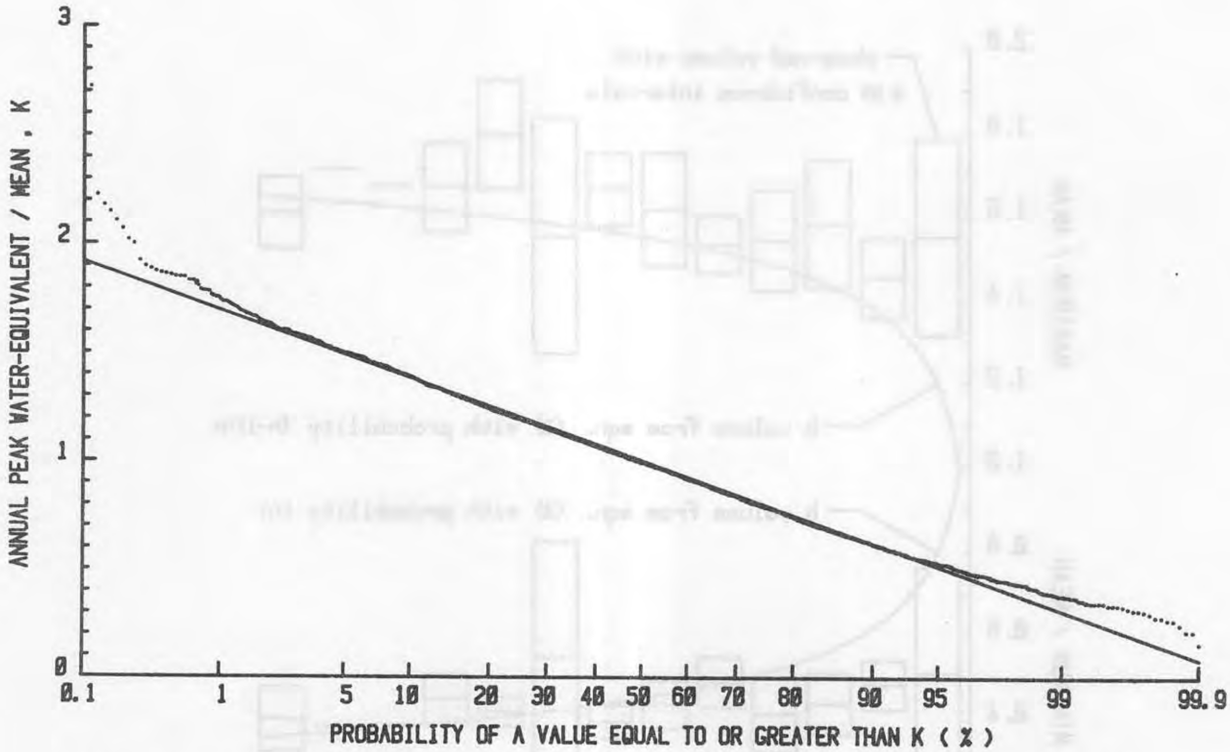

Figure 5. Frequency distributions of modular coefficients, combining all Wyoming snow courses having at least 30 years of data. Points are plotted as described in figure 4, and the fitted line is equation (3) with $s^2 = 0.088$.

Table 3. Statistical parameters for empirical frequency distribution (Eqn. (3)) of modular coefficients of annual peak water-equivalent combining all snow courses in Wyoming with 30 or more years of data.

Parameter	Value
Number of snow courses, N	82
Number of individual observations, n	3494
Mean	1.0000
Sample variance, s^2	0.0880
0.95 confidence interval	0.0840 - 0.0923
Sample coefficient of skewness, g_1	0.3445
Sample coefficient of kurtosis, g_2	3.6519
Kolmogorov statistic, D	0.0148
Cramér-von Mises statistic, W^2	0.2341

Although the calculation of actual values for expected extremes is more complex (David, 1970), observed minima and maxima for courses having more than 20 years of record (n) are in reasonable agreement with K values having probabilities 1/n and (n-1)/n, respectively, as computed from the empirical distribution (Fig. 6).

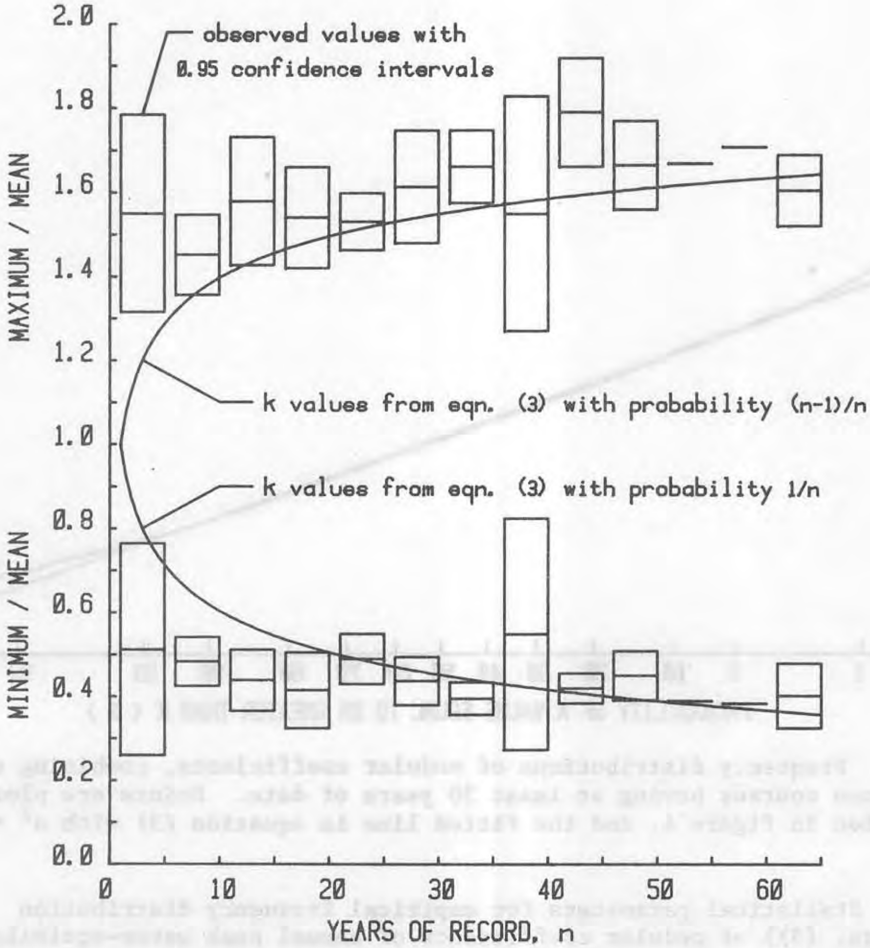

Figure 6. Extreme values of annual peak water-equivalent in relation to years of record, comparing observed values with those calculated from equation (3) with $s^2 = 0.088$.

<u>CONCLUSIONS AND APPLICATIONS</u>

The generalized frequency distribution appears to provide estimates with reasonable accuracy for most engineering applications, for all areas in Wyoming. This result was unexpected because of the variety of climate and topography within the State.

To facilitate application of the empirical frequency distribution, exceedance probabilities calculated from equation (3) as 1 - F(K), are provided in Table 4. To illustrate interpretation of tabled values, an annual peak water-equivalent greater than 0.50 times the mean would be expected to occur 95.41% of the time, whereas a value greater than 1.50 times the mean would be expected only 4.59 years out of 100. An annual peak water-equivalent greater than 200% of the mean, would be expected only 4 times in 10,000 years.

The dimensionless frequency curve can be used to generate dimensional (e.g., inches of water) distributions if the mean can be estimated. Because it is often possible to estimate mean water-equivalent from precipitation data or general topographic information, the dimensionless distribution derived here can enhance the utility of such estimates for preliminary planning.

Table 4. Probabilities of larger values for K. Values are 1 - F(K), where F(K) is given by equation (3) with $s^2 = 0.0880$.

K	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.9996	.9996	.9995	.9995	.9994	.9993	.9992	.9991	.9990	.9989
0.1	.9988	.9987	.9985	.9983	.9981	.9979	.9977	.9974	.9971	.9968
0.2	.9965	.9961	.9957	.9953	.9948	.9943	.9937	.9931	.9924	.9917
0.3	.9909	.9900	.9891	.9880	.9870	.9858	.9845	.9832	.9817	.9801
0.4	.9784	.9766	.9747	.9727	.9705	.9681	.9656	.9630	.9602	.9572
0.5	.9541	.9507	.9472	.9434	.9395	.9354	.9310	.9264	.9216	.9165
0.6	.9112	.9057	.8999	.8939	.8875	.8810	.8741	.8670	.8596	.8520
0.7	.8441	.8359	.8274	.8186	.8096	.8003	.7908	.7809	.7708	.7605
0.8	.7499	.7391	.7280	.7167	.7052	.6934	.6815	.6694	.6571	.6446
0.9	.6320	.6192	.6063	.5933	.5801	.5669	.5536	.5403	.5269	.5134
1.0	.5000	.4866	.4731	.4597	.4464	.4331	.4199	.4067	.3937	.3808
1.1	.3680	.3554	.3429	.3306	.3185	.3066	.2948	.2833	.2720	.2609
1.2	.2501	.2395	.2292	.2191	.2092	.1997	.1904	.1814	.1726	.1641
1.3	.1559	.1480	.1404	.1330	.1259	.1190	.1125	.1061	.1001	.0943
1.4	.0888	.0835	.0784	.0736	.0690	.0646	.0605	.0566	.0528	.0493
1.5	.0459	.0428	.0398	.0370	.0344	.0319	.0295	.0273	.0253	.0234
1.6	.0216	.0199	.0183	.0168	.0155	.0142	.0130	.0120	.0109	.0100
1.7	.0091	.0083	.0076	.0069	.0063	.0057	.0052	.0047	.0043	.0039
1.8	.0035	.0032	.0029	.0026	.0023	.0021	.0019	.0017	.0015	.0013
1.9	.0012	.0011	.0010	.0009	.0008	.0007	.0006	.0005	.0005	.0004
2.0	.0004	.0003	.0003	.0003	.0002	.0002	.0002	.0002	.0001	.0001
2.1	.0001	.0001	.0001	.0001	.0001	.0001	.0000	.0000	.0000	.0000

The most important use of the function, however, is in providing a basis for weighting other variables that depend on annual peak water-equivalent, in order to develop physical production functions needed for economic analyses. For example, the dimensionless frequency distribution was sufficient to answer our initial question concerning the optimum design year to use for snow fence projects, although the complexity of the procedure does not allow its description here.

Finally, the relative range of variation to be expected for hydrologic variables is often useful information in itself for drought and flood planning, avalanche zoning, and feasibility studies for snow management projects.

These results suggest the possibility of developing similar dimensionless distribution functions for other geographic areas and other variables.

ACKNOWLEDGEMENTS

The author is grateful for the many helpful suggestions by Rudy M. King, Biometrician, Rocky Mountain Forest and Range Experiment Station, and for the instructive counsel of Dr. Leon E. Borgman, Professor of Geology/Statistics, University of Wyoming. Special thanks are also due Edward R. Tabler, for his able assistance in entering the data in the computer.

REFERENCES

Armstrong, B. R. 1981. A quantitative analysis of avalanche hazard on U.S. Highway 550, Southwestern Colorado. Proceedings 49th Western Snow Conference (St. George, Utah, April 14-16, 1981), p. 95-104.

Brown, John Webster. 1970. An approach to snow load evaluation. Proceedings 38th Western Snow Conference (Victoria, B. C., April 21-23, 1970), p. 52-60.

David, Herbert Aron. 1970. Order statistics. John Wiley & Sons, Inc. New York. 272 p.

Doesken, Nolan J., Bernard A. Shafer, and Thomas B. McKee. 1981. Winter drought in the Colorado mountains. Proceedings, 2nd Conference on Mountain Meteorology (Steamboat Springs, Colo., November 9-12, 1981), American Meteorological Society and Forest Service, USDA. p. 57-58.

Engelen, G. B. 1972. A graphical and statistical approach to the regional study of snowpack in mountain areas, with special reference to Colorado and New Mexico. In: The Role of Snow and Ice in Hydrology, Proceedings of the Banff Symposia (September 6-13, 1972). UNESCO/WMO/IAHS. p. 885-894.

George, Tommy A. and Donald W. McAndrew. 1973. Snow load analysis and recreational uses of snow data. Proceedings, 41st Western Snow Conference (Grand Junction, Colo., April 17-19, 1973), p. 69-74.

Gumbel, Emil J. 1954. Statistical theory of extreme values and some practical applications. National Bureau of Standards Applied Mathematics Series 33. 51 p.

Johnson, Norman L. and Samuel Kotz. 1970. Continuous univariate distributions, V. 1. Houghton Mifflin Co., Boston. 300 p.

Kopanev, I. D. 1976. Probability of snow cover depth distribution. Soviet Hydrology 15 (4): 281-287.

LaChapelle, Edward R. 1966. Encounter proabilities for avalanche damage. USDA Forest Service, Miscellaneous Report 10. Alta Avalanche Study Center, Wasatch National Forest. 10 p.

Markovic, Radmilo D. 1965. Probability functions of best fit to distributions of annual precipitation and runoff. Colorado State University, Hydrology Paper No. 8. 33 p.

Rakhmanov, V. V. 1969. Snow storage distribution in river basins. Soviet Hydrology: Selected Papers 1: 49-63.

Snedecor, George W. 1956. Statistical methods. Fifth Edition. Iowa State University Press, Ames. 534 p.

Stephens, M. A. 1974. EDF statistics for goodness of fit and some comparisons. Journal of the American Statistical Association 69 (347): 730-737.

Tabler, Ronald D. 1968. Physical and economic design criteria for induced snow accumulation projects. Water Resources Research 4 (3): 513-519.

Thom, H.C.S. 1966. Distribution of maximum annual water-equivalent of snow on the ground. Monthly Weather Review 94 (4): 265-271.

Twedt, T. M., J. C. Schaake, Jr., and E. L. Peck. 1977. National Weather Service extended streamflow prediction. Proceedings, 45th Western Snow Conference (Albuquerque, N.M., April 18-21, 1977), p. 52-57.

Yevjevich, Vujica. 1972. Probability and statistics in hydrology. Water Resources Publications, Fort Collins, Colo. 302 p.