DISTRIBUTIONS OF EXTREME SNOWFALL IN WEST VIRGINIA

Mingteh Chang

Stephen F. Austin State University
School of Forestry
Nacogdoches, Texas 75962

Douglas G. Boyer

USDA-SEA-AR, North Atlantic Area
Plant Sciences, West Virginia University
Morgantown, West Virginia 26506

ABSTRACT

The goodness of fit of six probability distributions, Pearson type III, Gumbel, log-normal, cube-root normal, gamma, and Lieblein, for extreme snowfalls was investigated at 18 stations in West Virginia using the coefficient of determination (R^2), Friedman's two-way analysis, and plotting the theoretical versus observed values. All six models underestimated the observed extreme snowfalls for longer return periods. The overall deviation was smallest for the Pearson type III and largest for the Lieblein methods, which were, respectively, the best and poorest approaches for fitting the extreme snowfalls in the study area. A simple empirical equation was developed using station elevation and latitude for estimating the extreme snowfalls of different return periods at ungaged stations. This equation, applicable to portions of West Virginia at less than 730 m elevation, accounted for 95% of the variation with 16.5% average absolute error.

INTRODUCTION

Above the south and north temperate zones of the earth, snow functions in a complex manner with our environment. For example, the winter accumulation of snow may be a source of water for irrigation, navigation, and hydroelectric projects in one area, but may produce early spring flooding in another area. Snow may provide farmers an insulation to prevent soils from freezing, but it may be a condemned material to people on city sanitation and highway transportation. A knowledge of the frequency of extreme snowfall is of interest to many fields and disciplines.

Studies on the occurrences of extreme snowfall are relatively few compared with studies of extreme rainfall. Dunlap (1970) used Lieblein's (1954) approach of order statistics to fit the extreme snowfall at 120 stations in the Northeastern United States including 12 in West Virginia. Dunlap described this approach as "the most efficient currently known for fitting a straight line to extreme values." Vance and Whaley (1971) and Reese et al. (1973) found, however, that the Log-Pearson type III and the Pearson type III models, respectively, fit snowfall distributions satisfactorily in the West.

The study objectives were to compare the goodness of fit of Lieblein's method with that of other probability models, and to derive an equation that could be used to estimate expected extreme snowfalls of different return periods at ungaged sites. The procedure included: 1) repetition of Dunlap's work with a longer record (21 vs. 15 years) and more stations in West Virginia (18 vs. 12), 2) comparison of the Lieblein method with five other distribution models, i.e., Pearson type III, Gumbel, log-normal, gamma, and cube-root normal, for the goodness of fit, as a basis for selecting a best overall model to estimate the extreme snowfalls for different return periods in the study area, and 3) derive an empirical equation to estimate expected extreme snowfalls of different return periods at ungaged sites.

STUDY AREA AND DATA

West Virginia, located between 37° and 40° N latitude and 78° and 83° W longitude, has a complex topography with severe elevation changes within short distances. For example, the elevation ranges from 73 m at Harpers Ferry to 1,482 m at Spruce Knob within a horizontal distance of about 160 km; elevation differences greater than 900 m occur within individual counties. The state has about 21,000 square km, or 33% of its total area, at or above 600 m, and about 5,200 square km, or 8%, at or above 900 m. Its average elevation (505 m) is about 150 m greater than that of the second highest state (Pennsylvania) east of the Mississippi River (Lee et al., 1973). Coupling these high mountains with their axis (NE-SW) oriented perpendicular to the normal winter airflow from the Continental Arctic and the Great Lakes areas (also perpendicular to the few heavier storms from the southeast off of the Atlantic), the orographic lifting and resulting condensation causes the Allegheny Plateau in West Virginia to experience some of the heaviest snowfalls on the Atlantic slope of the United States (Leffler and Foster, 1974; Chang and Lee, 1975).

The annual maximum daily snowfall data used in this study were obtained from the U.S. National Weather Service records at 18 stations in West Virginia (Figure 1) for the 21 year period, 1950-1970. The station elevations range from 164 m (Martinsburg) to 724 m (Bayard), and the observed extreme snowfall from 13 mm (London Locks) to 711 mm (Glenville). Since the elevation of the highest station is less than half of the state's maximum, the unrecorded daily extreme snowfalls greater than 711 mm probably occurred at elevations above 724 m (Table 1).

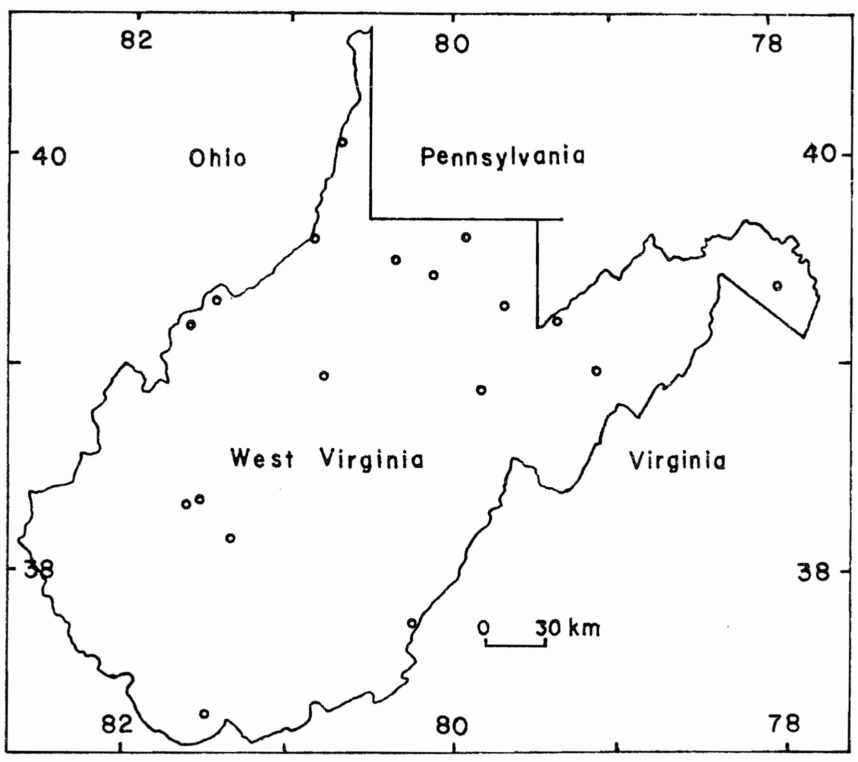

Figure 1. Extreme snowfall station location map.

Table 1. Snowfall Station, and the Mean, Standard Deviation (SD), Minimum (Min), and Maximum (Max) of the Extreme Snowfall, 1950 - 1970.

Station	Index	Elevation	Latitude (degrees)	Mean	SD	Min	Max
Bayard	0527	724	39.27	269.7	124.6	101.6	584.2
Charleston AP	1570	290	38.37	159.8	80.2	66.0	320.0
Charleston	1575	183	38.35	129.8	81.7	58.4	365.8
Elkins	2718	594	38.92	187.2	95.9	86.4	452.1
Fairmont	2920	396	39.47	193.4	93.6	88.9	442.0
Gary	3353	435	37.37	110.1	65.7	12.7	228.6
Glenville	3544	226	38.93	181.2	139.3	76.2	711.2
London Locks	5365	190	38.20	98.0	56.2	12.7	215.9
Mannington 1N	5621	297	39.55	173.1	97.1	76.2	523.2
Martinsburg AP	5707	164	39.40	192.0	105.6	76.2	533.4
Morgantown AP	6202	379	39.63	162.8	90.6	76.2	449.6
New Martinsville	6467	194	39.65	138.2	62.9	6.4	317.5
Parkersburg AP	6849	255	39.35	126.2	71.7	25.4	304.8
Parkersburg City	6859	187	39.27	142.0	90.8	50.8	398.8
Petersburg	6954	309	39.00	212.3	150.1	50.8	660.4
Rowlesburg	7785	419	39.35	200.5	70.3	76.2	381.0
Wheeling Warwood Dam	9492	201	40.10	166.7	114.5	38.1	482.6
White Sulphur Springs	9522	583	37.80	174.5	92.2	50.8	406.4

Note: All snowfall statistics are in millimeters.

FITTING THE SIX PROBABILITY MODELS

The six probability models, i.e., Pearson type III, Gumbel, log-normal, gamma, cube-root normal, and Lieblein, used in this study have frequently been fitted to hydrologic and climatic data in numerous regions with satisfactory results. Details of these six models can be found in textbooks of hydrology, Barger and Nyhan's (1960) work, and elsewhere.

The observed extreme snowfalls at the 18 stations were fit by the six models. For comparing the goodness of fit among the six probability models on the distribution of extreme snowfall in West Virginia, three procedures were used: 1) computing the coefficients of Determination (R^2), as suggested by Gupta (1970), of the six probability models for each station, 2) using Friedman's (1937) two-way analysis by ranking these R^2

values among the six models at each station, and 3) plotting the theoretical values computed by the six models versus observed values on probability paper for each station. The R^2 value, measuring the goodness of fit between the theoretical and observed values at a single station, was computed by

$$R^2 = 1 - \frac{\Sigma(X_m - \hat{X}_m)^2}{\Sigma(X_m - \overline{X}_m)^2} \quad (1)$$

where $\hat{X}_m$ and $\overline{X}_m$ are the theoretical values and the mean of the samples, respectively, and X_m is observed values. Friedman's analysis starts by ranking the R^2 values from lowest to highest at each station and summing the ranks of all stations for each distribution. The Chi-square (χ_r^2), was obtained by the following equation used to test the significance of the hypothesis at a selected probability level:

$$\chi_r^2 = \frac{12(\Sigma r^2)}{nk(k+1)} - 3n(k+1) \quad (2)$$

where n is the number of stations, k is the number of distributions, and Σr^2 is the sum of the squared values of the total ranks of each distribution. The critical values can be found from Chi-square tables using (k - 1) degrees of freedom. The technique was first used to test the differences among the six models. After it was confirmed that at least one model was significantly different from the others, the test was repeated to analyze the difference between any two models.

The total values for the 18 stations ranked as described above were highest for the Pearson type III and lowest for the Lieblein methods (92 vs. 34). Table 2 reveals: 1) within each model, about 83% of the stations were best or next to best fit to the Pearson type III distribution and 83% were poorly fit to the Lieblein method; 2) within the highest rank (i.e., 6), the Pearson type III method appeared at 61% of the stations and Lieblein at 11%; and 3) the R^2 values of the Pearson type III method ranked in the lower two categories 11% of the time. From these analyses it seems logical to conclude that, of the six models, the Pearson type III distribution generally had the highest R^2 value and, consequently, the best fit at the most stations. Friedman's technique of rank analysis confirmed that the six distribution methods differed significantly at an alpha level less than 0.001.

For the comparison between any pair of the six models for fitting the extreme snowfalls, the analysis of Friedman's technique showed that the Pearson type III method differed from the log-normal, cube-root normal, gamma, and Lieblein's distribution models at the probability level (P) greater than 99% and that there was little statistical difference (P = 66%) between the Pearson type III and the Gumbel (Table 3). The Lieblein model was statistically different (P > 94%) from the other five models, and there was no significant difference between the cube-root normal and the gamma models (P = 84%).

Generally, the predicted values of the six models were lower than the observed values for longer return periods, and particularly so for the Lieblein method. Average deviation of the predicted extreme snowfall from the maximum observed values for the Lieblein method was 114 mm, whereas mean deviations were 43 mm for the Gumbel, 78 mm for the Pearson type III, 79 mm for the log-normal, 97 mm for the cube-root normal, and 97 mm for the gamma. The prediction with the greatest deviation was at Glenville, Gilmer County, where the Lieblein method prediction was 368 mm (52%) below the observed maximum while the deviation for the Gumbel method prediction was 203 mm (28%) (the smallest) below the observed maximum and 234 mm (33%) below for the Pearson type III prediction. The Gumbel method gave a higher predicted value for long return periods than the other five models, but the Pearson type III predicted extreme snowfall with the smallest deviations from the observed values for all but the greatest return period. This contributed to the higher R^2 value for the Pearson type III, which identified it as the best overall model for fitting the extreme

Table 2. Percent of Stations When the Coefficients of Determination (R^2) of the Six Probability Models Were Ranked by Ascending Orders[1/], and Total Value of Ranks for Each Distribution.

Model	Rank of R^2 6	5	4	3	2	1	Total value of ranks
Pearson type III	61.1	22.2	0.0	5.6	5.6	5.6	92
Gumbel	22.2	38.9	33.3	5.6	0.0	0.0	86
Log-normal	5.6	33.3	22.2	11.1	5.6	22.2	70
Cube-root normal	0.0	5.6	27.8	33.3	33.3	0.0	56
Gamma	0.0	0.0	11.1	44.4	38.9	5.6	47
Lieblein	11.1	0.0	5.6	0.0	16.7	66.7	34

[1/] Ranks 6 and 1 are the highest and the lowest R^2 values, respectively.

Table 3. Probability Levels of Friedman's Analysis of Ranks.

Model	Pearson	Gumbel	Log-normal	Cube-root	Gamma
Gumbel	0.66				
Log-normal	>0.99	0.66			
Cube-root normal	>0.99	>0.99	0.84		
Gamma	>0.99	>0.99	0.94	0.84	
Lieblein	>0.99	>0.99	0.94	>0.99	>0.99

daily snowfall in West Virginia. The goodness of fit of the Pearson type III and Lieblein methods for the extreme snowfall at Glenville and London Locks is shown in Figure 2.

THE REGIONALIZED PREDICTION EQUATION

Based on data computed by the Pearson type III distribution model at the 18 stations, the equation derived for predicting the extreme snowfall (S, in mm) of different return periods (T_i in years) at ungaged locations in West Virginia was in the form:

$$S = 0.2240e^{(0.0742*Z + 0.1871*\phi + T_i)} \qquad (3)$$

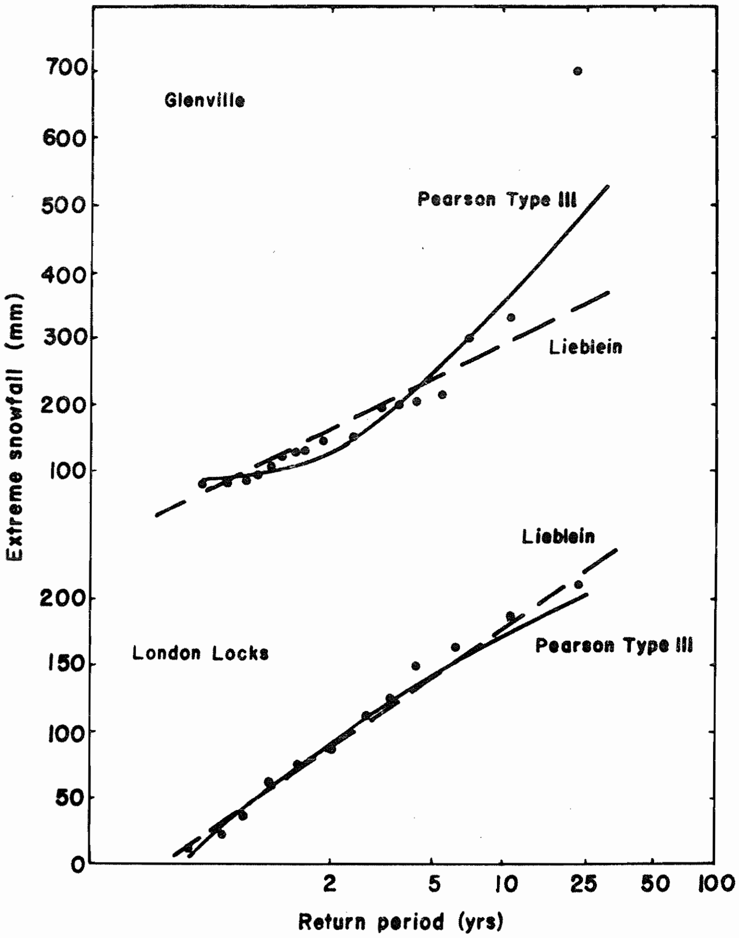

Figure 2. Goodness of fit of the Pearson type III and Lieblein methods to the observed extreme snowfall at Glenville and London Locks.

where Z and ϕ are station elevation in hundreds of meters and latitude in degrees, respectively, and e is the base of natural logarithms. Equation 3 is a covariance-analysis model, with T_i used as a classification parameter. Values of T_i for six different return periods are shown in Table 4, and the solution of Equation 3 can be read directly from Figure 3.

Equation 3 accounted for about 95% of the variability of the extreme snowfalls with an average absolute error of 53 mm or 16.5% of the mean. The residuals ranged from <1 mm (25-year snowstorm) at Parkersburg to 328 mm (100-year snowstorm) at Glenville, and 77% of the residuals were below 75 mm. The percentage of average absolute error was about identical among all return periods.

It is well known that the probability of precipitation in mountainous areas is enhanced by the orographic effects of mountain barriers on air mass lifting, cooling, and condensation processes; the lower air temperature at higher elevations increases the probability that precipitation will occur as snowfall. Results of the present study, as indicated by Equation 3 that extreme snowfall increases with elevation and latitude, agree with our expectation. In West Virginia the effects of 100 m in elevation were about equal to 0.40° in latitude. However, since no extreme snowfall samples from above 730 m were

Table 4. Values of Classification Parameter (T_i) for Six Different Return Periods.

Return period, years	T_i
2	-1.1537
5	-0.6904
10	-0.4687
25	-0.2483
50	-0.1116
100	0.0000

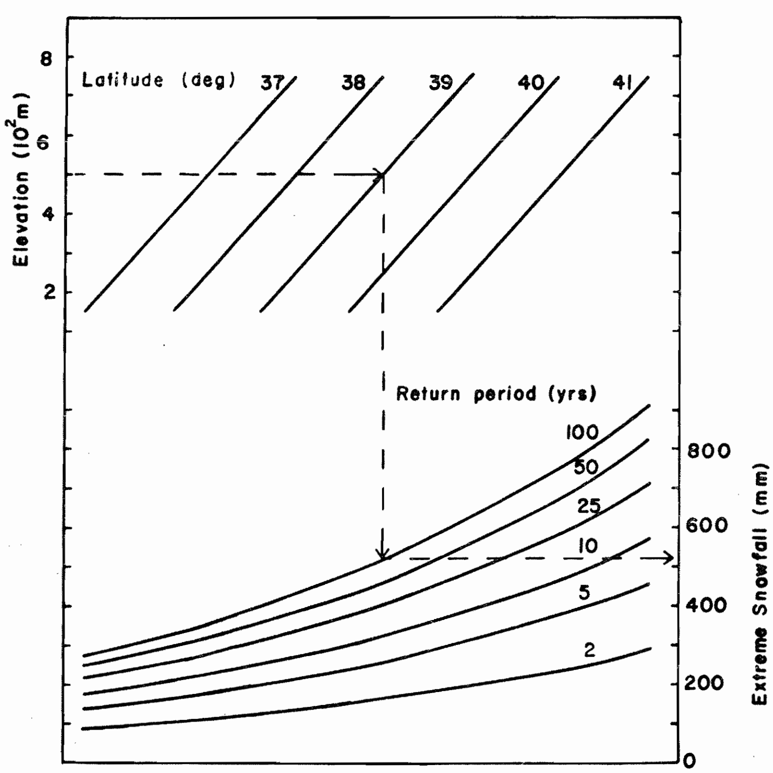

Figure 3. Nomograph for estimating extreme snowfall of different return periods in West Virginia.

used in the statistical analysis, applicability of the prediction equation at higher elevations is questionable.

The prediction equation is unique for its simplicity (i.e., it uses only two geographic parameters, which can be easily obtained from a topographic map), predictability, and plausibility. It is believed that the model can be applied for analyzing extreme snowfalls in other humid areas where winter precipitation is characterized by snowfall.

CONCLUSIONS

Among the six distribution models tested for their goodness of fit to the extreme snowfalls at 18 stations in West Virginia, the Pearson type III method was the best overall model, and the Lieblein was the poorest. No significant differences were found between Pearson type III and Gumbel, Gumbel and log-normal, log-normal and cube-root normal, and cube-root normal and gamma models. The extreme snowfalls of different return periods, computed by the Pearson type III method, can be accurately estimated using only station elevation and latitude in a multiple covariance analysis. It is believed that a similar model can be developed for other mountainous humid areas where winter precipitation is characterized by snowfall.

ACKNOWLEDGEMENTS

This paper is a contribution of the School of Forestry, Stephen F. Austin State University, Nacogdoches, Texas, and USDA-SEA-AR, North Atlantic Area, Morgantown, West Virginia. The research was supported by the Water Research Institute at West Virginia University (Project A-023-WVA), under the Water Resources Research Act of 1964 (PL 88-379) as administered by the Office of Water Research and Technology (USDI), and by the U. S. Department of Agriculture, Science and Education Administration, North Atlantic area, Morgantown, West Virginia.

LITERATURE CITED

Barger, G. L. and J. C. Nyhan. 1960. Climatology at work, measurements, methods, and machines. U. S. Weather Bureau, Asheville, NC. 109pp.

Chang, M. and R. Lee. 1975. Snowfall distribution in West Virginia. West Virginia Agriculture and Forestry, 6(2):5-8.

Dunlap, D. V. 1970. Probabilities of extreme snowfalls and snow depths. New Jersey Agricultural Experiment Station, Rutgers University, Bulletin 821. 15pp.

Friedman, M. 1937. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. Journal of the American Statistical Association, 32:675-701.

Gupta, V. L. 1970. Selection of frequency distribution models. Water Resources Research, 6(4):1193-1198.

Lee, R., M. Chang, and R. Calhoun. 1973. Elevation in West Virginia. West Virginia Agriculture and Forestry, 5(1):5-9.

Leffler, R. J. and J. L. Foster. 1974. Snowfall on Allegheny Plateau of Maryland and West Virginia. Weatherwise, 27:199-201.

Lieblein, J. 1954. A new method of analyzing extreme value data. National Advisory Committee for Aeronautics, Washington, D. C. technical Note 3053.

Reese, D. L., H. F. Mayland, and C. E. Brockway. 1973. Alternative methods of estimating snow water parameters. Water Resources Research, 9(5):1433-1439.

Vance, H. M. and B. L. Whaley. 1971. Snow frequency analysis for Oregon and Utah. In: Proceedings of 39th Annual Meeting of Western Snow Conference, April 20-22, 1971, Billing, Montana. pp. 34-38.