2272

JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION
VOL. 34, NO. 3 AMERICAN WATER RESOURCES ASSOCIATION JUNE 1998

EVALUATION OF PROBABILITY DENSITY FUNCTIONS IN PRECIPITATION MODELS FOR THE PACIFIC NORTHWEST[1]

Jinfan Duan, John Selker, and Gordon E. Grant[2]

ABSTRACT: Recent work has found that a one-parameter Weibull model of wet day precipitation amount based on the Weibull distribution provides a better fit to historical daily precipitation data for eastern U.S. sites than other one-parameter models. The general two-parameter Weibull distribution was compared in this study to other widely used distributions for describing the distribution of daily precipitation event sizes at 99 sites from the U.S. Pacific Northwest. Surprisingly little performance was sacrificed by reducing the two-parameter Weibull to a single-parameter distribution. Advantages of the single-parameter model included requiring only the mean wet day precipitation amount for calibration, invertibility for simulation purposes, and ease of analytical manipulation. The fit of the single-parameter Weibull to the 99 stations included in this study was significantly better than other single-parameter models tested, and performed as well as the widely endorsed, more cumbersome, two-parameter gamma model. Both the one- and two-parameter Weibull distributions are shown to have L-moments that are consistent with historical precipitation data, while the ratio of L-skew and L-variance in the gamma model is inconsistent with the historical record by this measure. In addition, it was found that the two-parameter gamma distribution was better fit using the method of moments estimators than maximum likelihood estimates. These findings suggested that the distribution in precipitation among sites in the Pacific Northwest with dramatically different settings are nearly identical if expressed in proportion to the mean site event size.
(KEY TERMS: meteorology/climatology; precipitation model; probability density functions; Pacific Northwest; surface water hydrology; simulation.)

INTRODUCTION

A variety of models can be found in the literature that describe daily precipitation (Smith and Schreiber, 1974; Todorovic and Woolhiser, 1974; Chin, 1977; Buishand, 1978; Roldan and Woolhiser, 1982; Richardson and Wright, 1984). Due to the complex nature of precipitation processes, these models have been selected by goodness-of-fit criteria rather than being derived from knowledge of the underlying physical processes. Although the ability to reproduce the historical distribution of precipitation amounts is paramount (Richardson, 1982), additional salient model characteristics include the ability to use in a simulation mode (e.g., invertibility), the ease of estimating parameter values, and computational flexibility (e.g., existence of closed form expressions for moments, ability to integrate and differentiate).

The most common approach for describing the distribution of precipitation amounts on days with precipitation or wet days (here we consider a day with total rainfall of 0.0254 centimeters (0.01 inch) or more as a wet day) is to assume the precipitation amounts are serially independent and to fit an analytical distribution to the precipitation depths (Woolhiser *et al.*, 1973; Smith and Schreiber, 1974; Todorovic and Woolhiser, 1974, 1975; Richardson, 1982). Various probability density functions requiring from one to three parameters have been proposed to describe the distribution of precipitation depths (Smith and Schreiber, 1974; Todorovic and Woolhiser, 1974, 1975; Richardson, 1982; Woolhiser and Roldan, 1982; Pickering *et al.*, 1988; Selker and Haith, 1990). Woolhiser and Roldan (1982) compared the chain-dependent (i.e., precipitation amounts are independent given the state of the previous day), exponential, gamma, and three-parameter mixed exponential distributions. They found that the three-parameter mixed exponential was the best for five U.S. stations.

[1]Paper No. 96177 of the *Journal of the American Water Resources Association*. **Discussions are open until February 1, 1999.**
[2]Respectively, Consultant, 15 Atherwood Avenue, No. 5, Redwood City, California 94061; Associate Professor, Department of Bioresource Engineering, Oregon State University, Room 2401, Gilmore Hall, Corvallis, Oregon 97331-3906; and Research Hydrologist, USDA, Pacific Northwest Research Station, USDA, 3200 Jefferson Way, Forest Science Lab, Corvallis, Oregon 97331 (E-Mail/Duan: duan@hydromodel.com).

Multi-parameter distributions such as the two-parameter gamma, the three-parameter gamma, the three-parameter mixed exponential, and others also have been used in precipitation simulation (e.g., Mielke and Johnson, 1974). These multiparameter models are generally assumed to fit the distribution of precipitation amounts more closely than do one-parameter distributions. On the other hand, single-parameter models are more appealing for their simplicity and ease of parameterization, but at this time there is no consensus on which family of models is inherently more reasonable or suitable. For a given application, the choice of models is often driven by the feasibility of obtaining the required parameters. Richardson (1982) suggests that unless the three-parameter mixed-exponential distribution has a clear advantage over the two-parameter gamma distribution, the gamma distribution should be the appropriate choice of model for most applications.

The Weibull family of distributions has several very desirable features, including invertibility, integrability, and closed form expressions for all moments. The Weibull model was used by Selker and Haith (1990) as a one-parameter model for daily precipitation amounts for the Eastern United States. A recent study using a slightly different calibration of the Weibull distribution also provided superior precipitation generation (in the sense of better chi-square fit to historical data) for the H. J. Andrews Experimental Forest in Oregon in comparison to other one-parameter distributions (Duan *et al.*, 1995). The study by Selker and Haith (1990) was limited to Eastern U.S. sites, and the calibration procedure employed only summary statistics for extreme precipitation events. In the study presented here, 99 full precipitation records selected from sites west of the Cascade Range in Washington and Oregon and the Klamath Mountains in northern California are examined to check the performance of the calibrated Weibull distribution in comparison to other widely used models. The study area included a wide variety of precipitation regimes, including coastal rain forest and moderate inter-montane and arid sites. The possibility of more precise calibration of the Weibull models based on geographic region, elevation, and annual precipitation amount also were explored. To accomplish this, five probability distributions for the precipitation amount (the exponential, calibrated Weibull, calibrated beta-P, and two-parameter gamma distribution with two different estimators) were compared.

THE MODELS

In this study, seasonal variations in daily precipitation processes were assumed to be constant within a month but to differ among months. Twelve sets of parameters were used for each station. Although more complex techniques (e.g., finite Fourier series) have been proposed to model seasonal variations (Roldan and Woolhiser, 1982), the technique used here is adequate for most applications (Richardson, 1981).

Calibration of the Weibull Distribution

A member of the Weibull family of distributions was given by Rodriguez (1977) as:

$$F(x) = 1 - e^{\left[\frac{kx}{\lambda}\right]^c} \tag{1}$$

with the density function in the form of

$$f(x) = c\left(\frac{k}{\lambda}\right)^c x^{c-1} e^{-\left(\frac{kx}{\lambda}\right)} \tag{2}$$

$$k = \Gamma\left(1 + \frac{1}{c}\right) \tag{3}$$

where x is the daily precipitation depth, F(x) is the probability of events less than x, f(x) is the density function, λ = E(X) is the expected value of daily precipitation, c is a constant, and Γ is the gamma function.

The constant c affects the general shape of the distribution (Figure 1). In the special case of c = 1 the Weibull distribution reduces to the single-parameter exponential distribution. In comparison to the exponential model, values of c > 1 have higher probability of precipitation amounts around the daily expected value and lower probabilities for either tail. The opposite is true for values of c < 1, which have higher probabilities for both low and high precipitation extremes (Figure 1).

Selker and Haith (1990) obtained a Weibull-based model through a regional calibration of a three-parameter beta-P distribution. Selker and Haith (1990) optimized the distribution to fit extreme high precipitation probabilities by using summary statistics from 31 Eastern U.S. locations. They derived the optimal one-parameter probability distribution for wet day precipitation with c = 0.75 so that

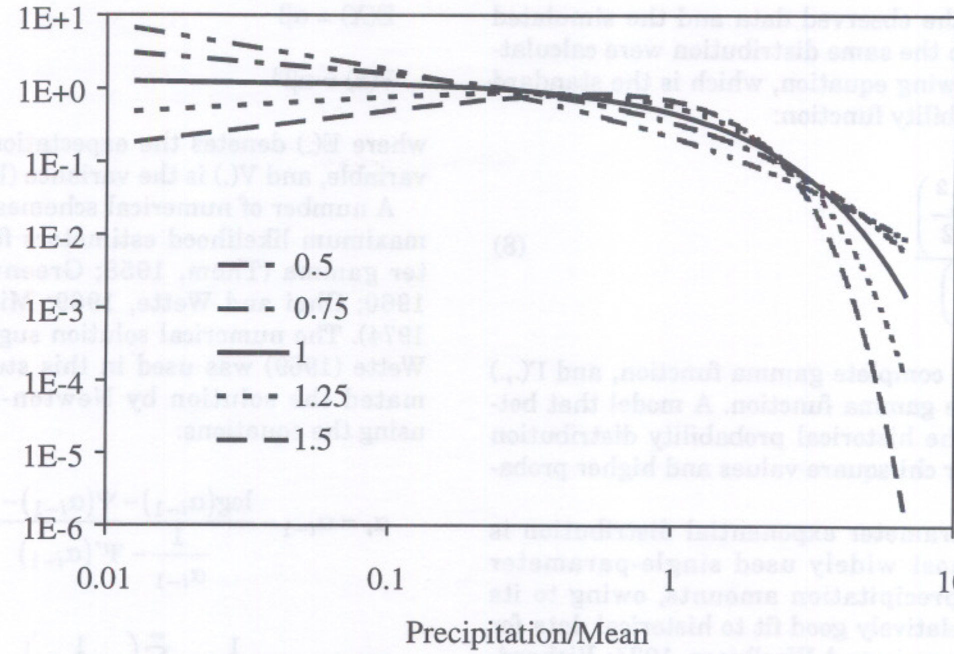

Figure 1. Weibull Density Functions for Various c Values; Precipitation Amounts Are Normalized by the Mean.

$$F(x) = 1 - e^{-\left(1.191\frac{x}{\lambda}\right)^{0.75}} \tag{4}$$

In this study, we calibrated the family of distributions described in Equations (1) to (3) with records from each of 99 stations to obtain optimized values of the parameters based on minimizing chi-squared statistics. In this way, each site was assumed to have one value of c for the entire period of record without distinguishing among months; the result was a model requiring one site parameter (c) and one parameter for each month (λ). Model distributions were obtained for a range of c between 0.5 and 1.5, incremented by 0.01. The chi-square values were computed as:

$$\chi^2 = \sum_{i=1}^{n} \frac{(observed_i - simulated_i)^2}{observed_i} \tag{5}$$

where *simulated* and *observed* are the number of precipitation events in a combined bin i and n is the total number of bins constructed. The bins were constructed to assure that at least five occurrences were observed in each bin category. The value of c that gave the least chi-square was selected as the optimal value. Figure 2 shows a typical relation between the chi-square statistic and c. Root mean squared errors (RMSE) were not used to evaluate models because Selker and Haith (1990) found very little difference in results when they compared RMSE to chi-square techniques.

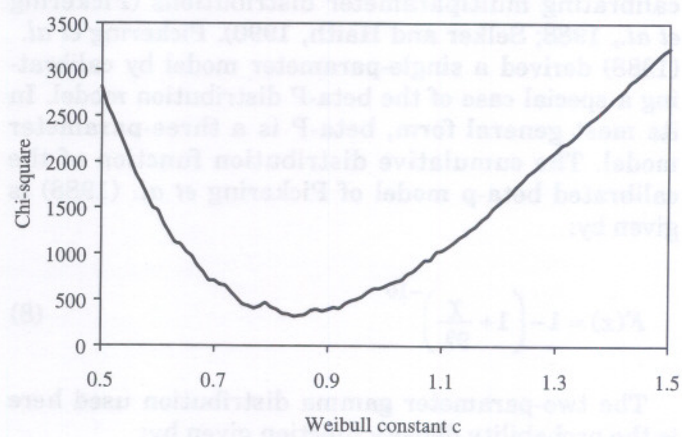

Figure 2. The Chi-Square Values and c-Parameters for the Corvallis, Oregon, Meteorological Station (minimum chi-square is at c = 0.84).

Models Compared

The Weibull models discussed above were compared with other single-parameter and multiparameter models by using the chi-square statistic for each model as calculated from Equation (5). The probabilities of the observed data and the simulated data belonging to the same distribution were calculated from the following equation, which is the standard chi-square probability function:

$$\text{Prob} = \frac{\Gamma\left(\frac{n}{2}, \frac{\chi^2}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \quad (8)$$

where $\Gamma(.)$ is the complete gamma function, and $\Gamma(.,.)$ is the incomplete gamma function. A model that better reproduces the historical probability distribution should give lower chi-square values and higher probability values.

The single-parameter exponential distribution is probably the most widely used single-parameter model of daily precipitation amounts, owing to its simplicity and relatively good fit to historical data for many sites (Todorovic and Woolhiser, 1974; Richardson, 1981; Pickering *et al.*, 1988). The cumulative distribution function of the exponential model is given by:

$$F(x) = 1 - e^{-\frac{x}{\lambda}} \quad (7)$$

Probability distributions also have been derived by calibrating multiparameter distributions (Pickering *et al.*, 1988; Selker and Haith, 1990). Pickering *et al.* (1988) derived a single-parameter model by calibrating a special case of the beta-P distribution model. In its most general form, beta-P is a three-parameter model. The cumulative distribution function of the calibrated beta-p model of Pickering *et al.* (1988) is given by:

$$F(x) = 1 - \left(1 + \frac{\chi}{9\lambda}\right)^{-10} \quad (8)$$

The two-parameter gamma distribution used here is the probability density function given by:

$$f(x) = \frac{x^{\alpha-1} e^{-\frac{x}{\beta}}}{\beta^{\alpha}\Gamma(\alpha)} \quad (9)$$

where α and β are the shape and scale parameters, respectively, of the gamma distribution. The parameters of the gamma distribution have commonly been fitted through two approaches: the method of moments and the maximum likelihood method. We employed both methods in this study. The moment estimators are given by:

$$E(X) = \alpha\beta \quad (10)$$

$$V(x) = \alpha\beta^2 \quad (11)$$

where E(.) denotes the expectation or average of the variable, and V(.) is the variance (Devore, 1987).

A number of numerical schemes will determine the maximum likelihood estimators for the two-parameter gamma (Thom, 1958; Greenwood and Durand, 1960; Choi and Wette, 1969; Mielke and Johnson, 1974). The numerical solution suggested by Choi and Wette (1969) was used in this study, which approximated the solution by Newton-Raphson iteration using the equations:

$$\alpha_i = \alpha_{i-1} - \frac{\log(\alpha_{i-1}) - \Psi(\alpha_{i-1}) - Y}{\frac{1}{\alpha_{i-1}} - \Psi'(\alpha_{i-1})} \quad (12)$$

$$\Psi(\alpha) = -\gamma - \frac{1}{\alpha} + \alpha \sum_{i=1}^{\infty}\left(\frac{1}{i(i+\alpha)}\right) \quad (13)$$

$$\Psi'(\alpha) = \sum_{i=0}^{\infty} \frac{1}{(i+\alpha)^2} \quad (14)$$

where γ is Euler's Constant with value of 0.57722157; Ψ is the Digamma function and Ψ' is the derivatives of the Digamma function; and i is the numerical step in the above scheme. Y is an intermediate value calculated as:

$$Y = \ln(E(X)) - E(\ln(X)) \quad (15)$$

β is then calculated by using Equation (10).

The above infinite summations are approximated by an arbitrary 10^{-7} criterion. If the difference between the calculated parameter value at iteration i and iteration i-1 is less than the criterion, then the value obtained on the i-th iteration was used. Because this numerical solution is always convergent with any initial value such that $0 < \alpha_0 < \infty$ (Choi and Wette, 1969), the starting value employed is immaterial.

GAUGE STATIONS AND DATA

The complete meteorological records from 99 gauge stations in the Pacific Northwest west of the Cascade Range were used (Figure 3, Table 1). They cover several physiographic provinces and subprovinces, including the Olympic rain forest, Coast Ranges, the

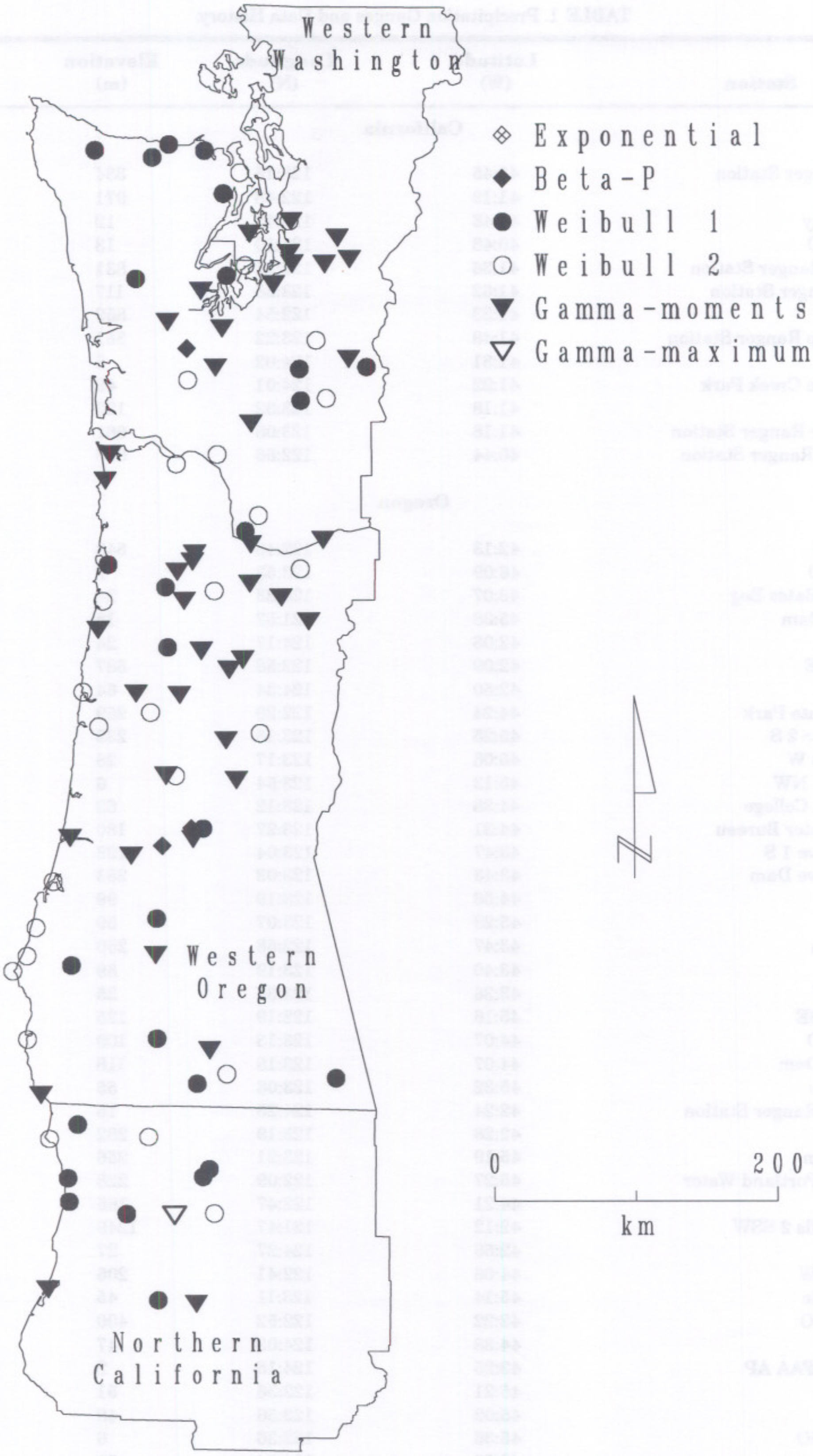

Figure 3. Station Sites and Best Fits of the Five Distributions.

TABLE 1. Precipitation Gauges and Data History.

Station	Latitude (W)	Longitude (N)	Elevation (m)	Data Length (yr)
California				
Big Bar Ranger Station	40:45	123:15	384	44
Callahan	41:19	122:48	971	44
Crescent City	41:46	124:12	12	44
Eureka WSO	40:48	124:10	13	44
Fort Jones Ranger Station	41:36	122:51	831	44
Gasquet Ranger Station	41:52	123:58	117	44
Greenview	41:33	122:54	859	44
Happy Camp Ranger Station	41:48	123:22	351	62
Klamath	41:31	124:02	8	44
Orick Prairie Creek Park	41:22	124:01	49	44
Orleans	41:18	123:32	123	44
Sawyers Bar Ranger Station	41:18	123:08	661	37
Weaverville Ranger Station	40:44	122:56	625	44
Oregon				
Ashland 1 N	42:13	122:43	543	44
Astoria WSO	46:09	123:53	2	40
Bandon 1E-Bates Bog	43:07	124:23	24	44
Bonneville Dam	45:38	121:57	18	44
Brookings	42:03	124:17	24	62
Buncom 2 SE	42:09	122:58	587	44
Cape Blanco	42:50	124:34	64	26
Cascadia State Park	44:24	122:29	259	61
Cherry Grove 2 S	45:25	123:15	238	34
Clatskanie 3 W	46:06	123:17	28	44
Cloverdale 1 NW	45:13	123:54	6	43
Corvallis St. College	44:38	123:12	69	103
Corvallis Water Bureau	44:31	123:27	180	44
Cottage Grove 1 S	43:47	123:04	198	44
Cottage Grove Dam	43:43	123:03	253	49
Dallas	44:56	123:19	99	57
Dilley 1 S	45:29	123:07	50	44
Dorena Dam	43:47	122:58	250	44
Drain	43:40	123:19	89	44
Elkton 3 SW	43:36	123:35	35	44
Estacada 2 SE	45:16	122:19	125	45
Eugene WSO	44:07	123:13	109	54
Fern Ridge Dam	44:07	123:18	118	49
Forest Grove	45:32	123:06	55	65
Gold Beach Ranger Station	42:24	124:25	15	44
Grants Pass	42:26	123:19	282	65
Haskins Dam	45:19	123:21	256	44
Headworks Portland Water	45:27	122:09	228	62
Holley	44:21	122:47	165	44
Klamath Falls 2 SSW	42:12	121:47	1249	43
Langlois 2	42:56	124:27	27	36
Leaburg 1 SW	44:06	122:41	206	44
MC Minnville	45:14	123:11	45	54
Medford WSO	42:22	122:52	400	65
Newport	44:38	124:03	47	59
North Bend FAA AP	43:25	124:15	2	62
Oregon City	45:21	122:36	51	44
Otis 2 NE	45:02	123:56	48	44
Portland WSO	45:36	122:36	6	44
Powers	42:53	124:04	70	44
Reedsport	43:42	124:07	18	32
Rex 1 S	45:18	122:55	149	40
Riddle 2 NNE	42:58	123:21	202	44
Roseburg KQEN	43:12	123:21	142	27

TABLE 1. Precipitation Gauges and Data History (cont'd.)

Station	Latitude (W)	Longitude (N)	Elevation (m)	Data Length (yr)
Oregon (cont'd.)				
Salem WSO	44:55	123:01	60	65
Seaside	45:59	123:55	3	63
Silver Creek Falls	44:52	122:39	411	44
Stayton	44:48	122:46	142	41
Summit	44:38	123:35	227	44
Three Lynx	45:07	122:04	341	62
Tidewater	44:25	123:54	15	44
Tillamook 1 W	45:27	123:52	3	45
Washington				
Aberdeen 20 NNE	47:16	123:42	133	62
Battle Ground	45:47	122:32	90	44
Bremerton	47:34	122:40	49	44
Cedar Lake	47:25	121:44	475	62
Centralia	46:43	122:57	56	62
Chimacum 4 S	47:57	122:46	76	45
Doty 3 E	46:38	123:12	79	15
Electron Headworks	46:54	122:02	527	31
Elma	47:00	123:24	21	14
Elwha Ranger Station	48:02	123:35	110	44
Glenoma 1 W	46:31	122:10	265	25
Grapeview 3 SW	47:18	122:52	9	44
Kent	47:23	122:16	10	37
Kid Valley	46:22	122:37	210	31
Landsburg	47:23	121:58	163	62
Longview	46:10	122:55	4	62
Mineral	46:43	122:11	451	31
Oakville	46:50	123:13	26	44
Olympia WSO	46:58	122:54	59	44
Port Angeles	48:07	123:26	30	44
Quilcene 2 SW	47:49	122:55	37	44
Rainier Ohanapecosh	46:44	121:34	594	44
Rainier Paradise RS	46:47	121:44	1654	38
Randle 1 E	46:32	121:56	274	44
Sappho 8 E	48:04	124:07	232	40
Seattle EMSU WSO	47:39	122:18	6	20
Seattle-Tacoma WSO	47:27	122:18	122	44
Seattle U of W	47:39	122:17	29	34
Sequim	48:05	123:06	55	41
Shelton	47:12	123:06	7	44
Snoqualmie Falls	47:33	121:51	134	62
Tacoma City Hall	47:15	122:26	81	33
Vancouver 4 NNE	45:41	122:39	64	93
Wauna 3 W	47:22	122:42	5	44

Cascade Range, Willamette Valley, Klamath Mountains, and others. Most of the stations began observation in either 1948 (61 of 99 stations) or 1931 (15 of 99), and generated an observed history of about 44 and 62 years, respectively. Observed daily precipitation datum quality is good with relatively few missing data. The first years of observed data were often omitted because of irregular observations in this period.

RESULTS AND DISCUSSION

The calibration of Weibull distributions for the 99 stations showed that the optimal value of c ranged between 0.6 and 1.06. One of the motivations in carrying out this study was to develop a method of estimating the value of c from data other than the full daily weather record. Toward this end, multiple regression analysis was carried out on the value of the cs and log(c) against elevation, mean annual precipitation, latitude, and longitude. The models obtained generally had R^2s of less than 0.10, thereby indicating no significant relation between these variables and the value of c. The possibility of regionalizing the c-parameter by using the values from the 99 sites was then considered, because complex spatial patterns would not be amenable to analysis under a simple regression approach. The variogram shown in Figure 4 was obtained from the EPA geostatistical software GEO-EAS. Remarkably, there is essentially no spatial correlation in values of c.

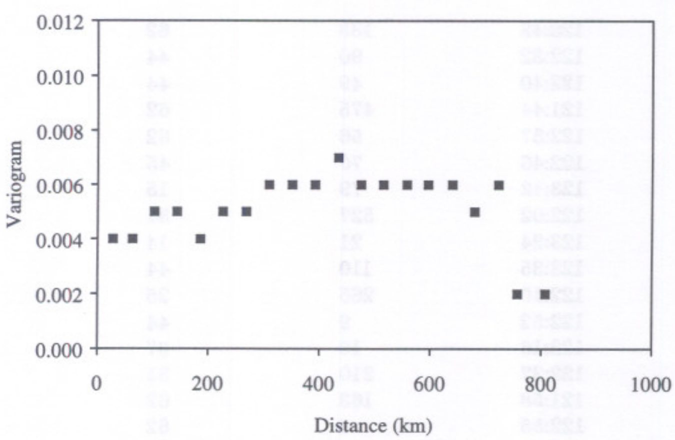

Figure 4. Spatial Variogram for c-Values for the 99 Stations.

Given the nearly complete lack of explanatory variables, the unavoidable conclusion is that the value of c is spatially random. A typical plot of the chi-square statistic as a function of c demonstrated that most sites had very little variation in goodness of fit over the range of $0.7 < c < 0.9$ (Figure 2), suggesting that use of a fixed value for c in this range might not significantly reduce the quality of fit of the model. For simplicity, the value of c that gave the average best fit for the 99 sites was selected to form a one-parameter model that could be included in the overall comparison of candidate models. The optimum overall value obtained from our calibration was 0.78 (Table 2 and Table 3). This value differed only slightly from the value of 0.75 obtained by Selker and Haith (1990). This correspondence is somewhat surprising in that the method of obtaining the values employed by Selker and Haith (1990), was quite different from that employed here. First, Selker and Haith (1990) calibrated the Weibull distribution by using only the probability distribution of extreme events (> 12 mm) from the Weather Bureau Report 57 (Miller and Fredrick, 1966), reflecting the fact that they desired a model that was most precise in the simulation of extreme events, as well as their lack of full weather records for a significant number of sites. Secondly, Selker and Haith (1990) made use of sites only in the Eastern U.S. for their calibration. In addition to these differences in methodology, it is clear both here and in Selker and Haith (1990) that the quality of fit of the data is relatively insensitive to values of c between $0.7 < c < 0.9$. Taken together, these results suggest that c will be in the vicinity of 0.75 for sites with a wide range of physiographic conditions. This result suggests that all sites studied have the same distribution of precipitation if scaled by the mean wet day precipitation and furthermore, that these distributions are monthly invariant. This result buttresses that of Selker and Haith (1990) and implies that a very simple analysis may be used to predict the probability of extreme events. In the matter of quality control checks on precipitation data, a simple check on number of recorded vs. predicted events can be carried out by using only the mean site precipitation, a method that appears to be very promising based on initial tests (George Taylor, State of Oregon Climate Service, Personal Communication).

Of the six models tested (the exponential, the calibrated beta-P, the method of moments fit two-parameter gamma, the maximum likelihood two-parameter gamma, the two-parameter Weibull, and the calibrated one-parameter Weibull), the Weibull and Gamma fit by method of moments gave the best results (Table

TABLE 2. Number of Stations That Best Fit for Data Using Different Models.

Model	Monthly Best Fit	Fixed Weibull Parameter For All Stations					
		0.75	0.78	0.80	0.82	0.84	0.85
Exponential	0	0	0	0	1	1	1
Weibull	53	35	36	31	27	22	18
Beta	3	4	3	3	3	2	1
Gamma Moments	42	59	60	65	68	74	79
Gamma ML	1	1	0	0	0	0	0

2, and Table 3). A box plot of the distribution of model p-values, which represent the the closeness of the theoretical distributions match the observed data, (Figure 5) shows that the two-parameter Weibull distribution has a median p-value of 0.184, and the one-parameter version has a median p-value of 0.163. Clearly, very little was lost in the reduction to a single-parameter model. The two-parameter Gamma distribution fit using the moment estimators provided a median p-value of 0.184. It is surprising that the performance of this widely used two-parameter model provides performance that is only as good as a far simpler one-parameter alternative. The spatial characteristics of these results may be visualized by plotting the number of sites where each of the models provided the best fit to the data (Figure 3). Neither the 53 of stations where the two-parameter Weibull gave best fits nor the 36 stations where the single parameter Weibull gave best fit show any particular spatial pattern. Given the variability of regional patterns of precipitation in the study area, it seems that the relative quality of fit is not a strong function of precipitation regime.

TABLE 3. Average Probabilities of Model and Observed Data.

Model	Average Probability of Model and Observed Data Belong to the Same Population
Exponential	0.031
Beta	0.055
Gamma Moments	0.184
Gamma ML	0.095
Weibull: Monthly Best Fit	0.184
Weibull: Fixed 0.75	0.161
Weibull: Fixed 0.78	0.163
Weibull: Fixed 0.80	0.158
Weibull: Fixed 0.84	0.137
Weibull: Fixed 0.85	0.130

Why is it that the very flexible two-parameter Gamma distribution does not provide a significantly better fit than the more constrained single-parameter Weibull model? L-moments diagrams provide a useful alternative approach to differentiate the goodness-

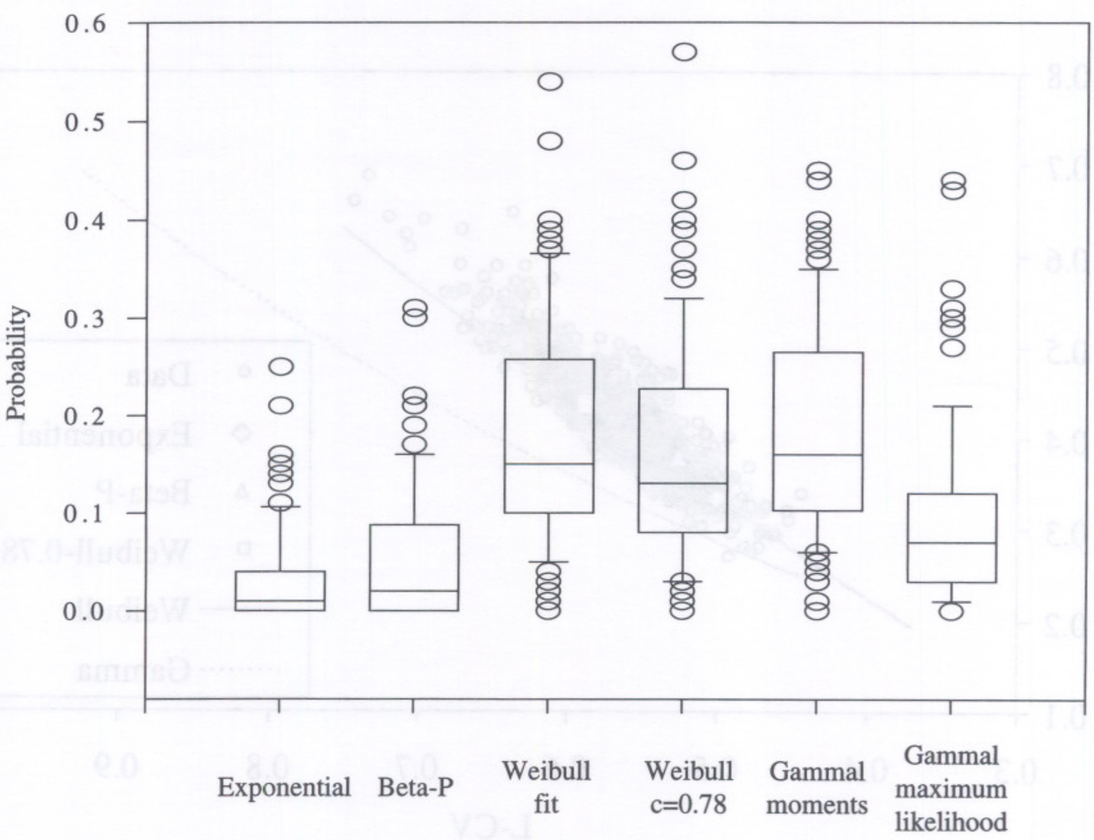

Figure 5. Box Plot of p-Values for Chi-Square Evaluation of Models.

of-fit (Hosking, 1986, 1990; Stedinger, 1993). By comparing the relative magnitude of L-skew and L-variance presented by the historical data and the models, we can see whether the general shape of the model's probability distributions are in keeping with the data (Figure 6). In the skew-variance plane, two-parameter models define a line, and single parameter models define a single point. The L-diagram clearly shows that the Gamma model does not provide the proper ratio of skew to variance, with the line falling below all but one of the site data points on the skew-variance plane (Figure 6). In contrast, the two-parameter Weibull model is very much in keeping with the site data, falling on a line providing nearly a best straight line fit to the data. In the single-parameter models, the c = 0.78 Weibull model occupies a position close to the visual center of the data distribution, and the exponential and beta-P distributions are significantly to the lower side of the data. The beta-P would seem to be superior to the exponential model by this measure, which is in agreement with the other measures of quality of fit.

The L-moment diagram also helps make sense of the surprisingly strong performance of the method of moments fit of the Gamma relative to the maximum likelihood method. Generally speaking, maximum likelihood estimations are sufficient and less biased than moment-based estimators, yet we find that the method of moments provides a strikingly superior fit to the data in the case of precipitation. The theoretical superiority in the maximum likelihood method is based on the assumption that the data were generated from the distribution to which they were being fit. In this case, the L-moment diagram demonstrated that the gamma model is systematically inconsistent with the underlying data. With this observation, the basis for the theoretical advantage of the maximum likelihood method is void, which is demonstrated by the poor fit of the gamma models with parameters estimated in this manner. Hydrologists who subscribe to a particular parameter-estimation method based on theory that rests on stringent assumption might want to reconsider given the general lack of knowledge about the true underlying distribution.

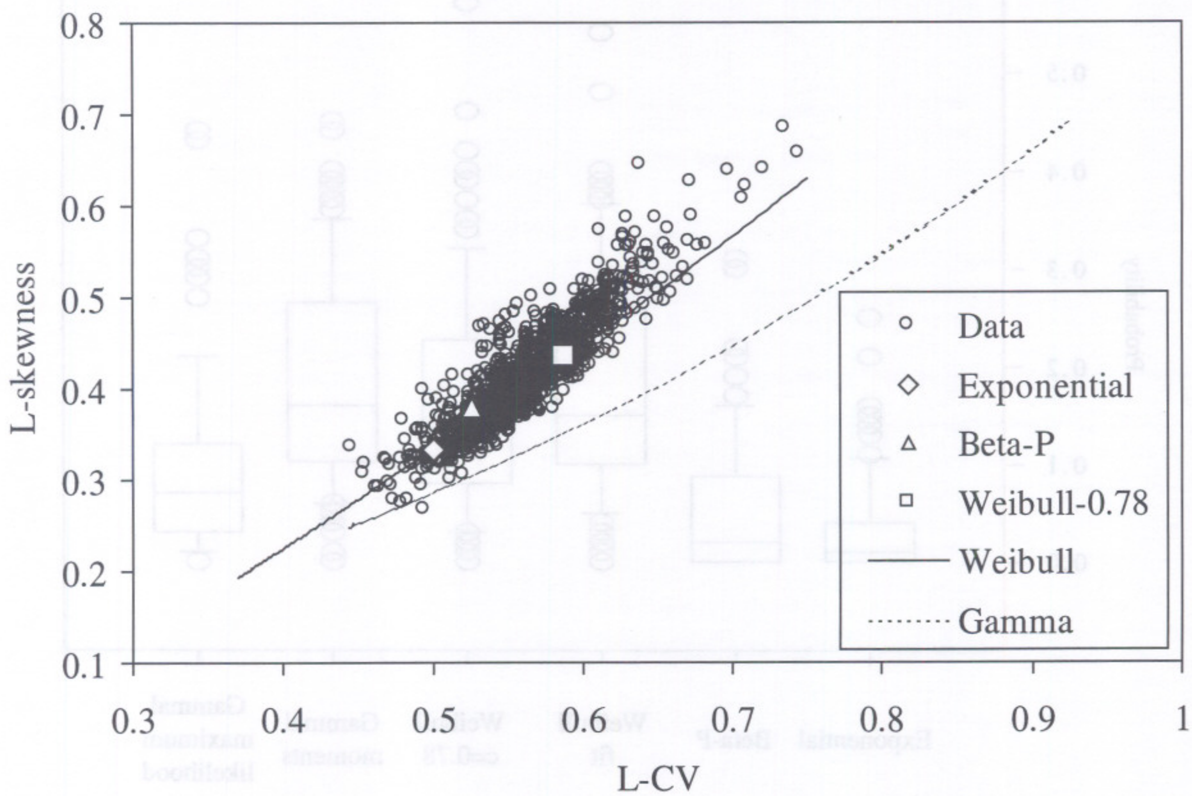

Figure 6. L-Moments Diagram Showing Observed L-Moment Ratios and L-Moment Ratio From Theoretical Distributions.

CONCLUSIONS

The ability of an analytical function to fit the empirical probability distribution for wet-day precipitation amounts depends on the flexibility of the distribution and the intrinsic nature of the distribution's shape. It was shown that the underlying shape of the single-parameter Weibull distribution allowed it to reproduce the historical distribution of precipitation events as well as the otherwise more flexible two-parameter gamma model. This result is fortunate, in that the single-parameter model is far more amenable to precipitation modeling and other applications. The only parameter required to fit the Weibull model to a site is the mean wet day precipitation, which is published for thousands of sites; the two-parameter distributions require calculation of parameter values from records of daily precipitation. Further, the Weibull model is invertible, has closed-form moments, and is easily integrable and differentiable, all features of significant utility in precipitation modeling.

ACKNOWLEDGMENTS

This study was partially funded by a National Science Foundation grant for the H. J. Andrews Experimental Forest, Long-Term Ecological Research Program (BSR-9011663) and the USDA Forest Service Global Change Research Program. The authors thank Dr. George Taylor, State Climatologist of Oregon, who provided the precipitation data, and Dr. Jerry R. Stedinger, School of Civil and Environmental Engineering, Cornell University, for his very helpful review. We are also thankful to three anonymous reviewers for thoughtful comments.

LITERATURE CITED

Buishand, T. A., 1978. Some Remarks on the Use of Daily Rainfall Models, J. Hydrol. 36:295-308.

Chin, E. H. 1977. Modeling Daily Precipitation Occurrence Process With Markov Chain, Water Resources Research 13(6):949-956.

Choi, S. C. and R. Wette, 1969, Maximum Likelihood Estimation of the Parameters of the Gamma Distribution and Their Bias. Technometrics 11(4):683-690.

Devore, J. L., 1987. Probability and Statistics for Engineering and the Sciences. Brooks/Cole Publishing Company, Monterey, California, p. 672.

Duan, J., A. K. Sikka, and G. E. Grant, 1995. A Comparison of Stochastic Models for Generating Daily Precipitation at H. J. Andrews Experimental Forest. Northwest Sciences 69(4):318-329.

Greenwood, Anther J. and David Durand, 1960. Aids for Fitting the Gamma Distribution by Maximum Likelihood, Technometrics 2:55.

Hosking, J. R. M., 1986. The Theory of Probability Weighted Moments. IBM Res. Rep. RC12210, Yorktown Heights, New York, p. 160

Hosking, J. R. M., 1990. L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics. Journal Royal Statistical Soc. B. 52(2):105-174.

Mielke, P. W., Jr. and E. S. Johnson, 1974. Some Generalized Distributions of the Second Kind Having Desirable Application Features in Hydrology and Meteorology. Water Resources Research 10(2):223-226

Miller, J.F. and R. H. Frederick, 1966. Normal Monthly Number of Days with Precipitation of 0.5, 1.0, 2.0, and 4.0 Inches or More in the Conterminous United States. Tech Pap. 57, U.S. Weather Bureau, Washington, D.C.

Pickering, N. B., J. R. Stedinger, and D. A. Haith, 1988. Weather Input for Nonpoint Source Pollution Models. J. Irrig. Drain. Eng. 114(4):674-690.

Richardson, C. W., 1981. Stochastic Generation of Daily Precipitation, Temperature, and Solar Radiation, Water Resources Research 17(1):182-190.

Richardson, C. W., 1982. A Comparison of Three Distributions for the Generation of Daily Rainfall Amounts. *In:* Statistical Analysis of Rainfall and Runoff, V. P. Singh (Editor), pp. 67-78. Proc. Int. Symp. on Rainfall-Runoff Modeling, Water Resources Publications, p. 700.

Richardson, C. W. and D. A. Wright, 1984. WGEN: A Model for Generating Daily Weather Variables. U.S. Department of Agriculture, ARS-8, p. 83.

Rodriguez, R. N., 1977. A Guide to the Burr Type XII Distributions. Biometrika 64:129-134.

Roldan, J. and D. A. Woolhiser, 1982, Stochastic Daily Precipitation Models. 1. A Comparison of Occurrence Processes. Water Resources Research 18(5):1451-1459.

Selker, J. S. and D. A. Haith, 1990. Development and Testing of Single-Parameter Precipitation Distributions. Water Resources Research 26(11):2733-2740.

Smith, R. E. and H. A. Schreiber, 1974. Point Processes of Seasonal Thunderstorm Rainfall. 2. Rainfall Depth Probabilities. Water Resources Research 10(3):418-423.

Stedinger, J. R., R. M. Vogel, and E. Foufoula-Georgiou, 1993. Frequency Analysis of Extreme Events. 1993 Handbook of Hydrology, D. R. Maidment (Editor). McGraw-Hill.

Thom, H. C. S., 1958. A Note on the Gamma Distribution. Monthly Weather Rev. 86(117).

Todorovic, P. and D. A. Woolhiser, 1974. Stochastic Model of Daily Rainfall. Proceedings of the Symposium on Statistical Hydrology, U.S. Dep. Agric. Misc. Publ. 1275, pp. 232-246.

Todorovic, P. and D. A. Woolhiser, 1975. A Stochastic Model of n-Day Precipitation. J. Appl. Meteorol. 14(1):17-24.

Woolhiser, D. A. and Jose Roldan, 1982. Stochastic Daily Precipitation Models. 2. A Comparison of Distribution Amounts. Water Resources Research 18(5):1461-1468.

Woolhiser D. A., E. Rove, and P. Todorovic, 1973. Temporal and Spatial Variation of Parameters for the Distribution of n-Day Precipitation. *In:* Flood and Droughts. Proceedings of the Second International Symposium on Hydrology, Water Resources Publications, Fort Collins, Colorado, pp. 605-614.