A Study of Fitting the Generalized Lambda Distribution to Solar Radiation Data[1]

A. ÖZTÜRK[2] AND R. F. DALE

Department of Agronomy, Purdue University, West Lafayette, IN 47907

(Manuscript received 10 July 1981, in final form 5 March 1982)

ABSTRACT

The increased interest in the climatology of solar radiation dictates a need for a distribution to fit daily solar radiation totals which tend to have negatively-skewed probability distributions. Even daily mean solar radiation for weekly periods tends to have non-normal distributions. The generalized lambda distribution, which includes a wide variety of curve shapes, is discussed for fitting these data. The underlying probability distribution is a generalization of the lambda distribution from three to four parameters. Using the weekly averages of daily solar radiation totals for each of 12 weeks during the growing season and daily totals for the week 5–11 July at West Lafayette, Indiana, it is shown that the generalized lambda distribution model fits the data well. Some results concerning percentiles and quantiles, parameter estimates, and goodness-of-fit tests are also discussed.

1. Introduction

Practically all energy utilized on the earth, except for nuclear power, arises directly or indirectly from solar radiation. Solar radiation is the basic forcing function for diurnal and annual soil and air temperature and evapotranspiration patterns, it directly limits the photosynthetic process in crop and other biomass production, and it is a non-polluting source for energy systems. As both food and energy supplies have become more costly, interest has increased in better assessment and use of our solar radiation resource. To better assess this resource, it is not only necessary to expand the observational network used for the measurement of solar radiation and to improve the quality control and archiving of these data, but it is also necessary to use the proper statistical distributions in preparing solar radiation climatologies.

In most seasons and areas of the United States the distributions of daily solar radiation totals are non-normal, primarily because of a physical upper limit on the daily incident radiation for any given location and day of the year. Bennett (1967) prepared frequencies of daily solar radiation totals for 57 U.S. and 27 Canadian stations for the 1950–64 period. He showed that in June, with the exception of stations along the North Pacific Coast, daily insolation distributions were negatively skewed, i.e., a clustering of daily amounts above the mean and a long tail below. He found the greatest skewness in areas of highest insolation, the skewness statistic (α_3 defined later in this paper) being −2.29 at Davis CA and −2.43 at Phoenix AZ. Yet, even distributions in areas of high cloudiness, such as the northeastern United States, showed negative skewness. He concluded that for such non-normal distributions the median was more representative than the mean and prepared charts of the median daily insolation for each month.

Recognizing this non-normality, most climatologists have used empirical distributions of the daily incident radiation totals or, utilizing the central limit theorem, daily means for weekly or 10-day periods. McQuigg and Decker (1958) for Columbia MO, Waite and Shaw (1961) for Ames IA, and Changnon (1959, 1978) for Urbana IL, presented average daily solar radiation values for 10-day periods. Rosenberg (1964) used 7-day averages for climatological weeks. Baker and Haines (1969) provided empirical frequencies of daily solar radiation totals and percent possible sunshine within each climatological week for all stations in the North Central region of the United States. Baker and Klink (1975) extended this regional study through 1970, presented histograms of daily solar radiation totals for representative stations for summer and winter weeks, and charted the difference between the median and mean daily solar radiation for each week of the year. All stations in the North Central region showed strong negative skewness during the summer months. This skewness disappeared in the winter months in the eastern part of the region, but continued all year in the western

[1] Journal Paper No. 8642. Agricultural Experiment Station, Purdue University, West Lafayette, Indiana 47907. Partially supported by Fulbright Scholars Program and National Science Foundation Atmospheric Sciences Section Grant ATM-8021075.

[2] Fulbright visiting professor, October 1980–September 1981. Present affiliation: Ege University, Institute of Statistics and Computer Sciences (EHBE), Bornova, Izmir, Turkey.

0021-8952/82/070995-10$06.50

portion. Getz and Nicholas (1979), using the SOLMET (USDC, 1977) data and the Baker and Klink methodology, determined empirically the daily solar radiation at each probability decile for each climatological week for 26 U.S. locations, generally based on the period of record July 1952–December 1975. They also computed the daily mean and standard deviation of the mean for each week, but did not discuss the limitations of these statistics.

For predicting the average performance of certain solar energy systems, Bendt *et al.* (1981) computed the "utilizability," defined as the fraction of the time the long-term average insolation incident on a solar collector aperture is above a specified threshold. They used measured daily solar radiation totals for 90 stations in the contiguous United States with ~20 years of observations. They did not discuss the underlying distribution, but converted the results into the ratio of terrestrial to extraterrestrial irradiation, defined as a "clearness index." Engels *et al.* (1981) separated the deterministic component of the diurnal solar radiation pattern from the stochastic component by analyzing the ratio of the measured hourly total insolation to that computed for a clear sky condition. They reported that the probability distribution of this hourly attenuation also demonstrated strong negative skewness, indicating that knowledge of only the mean and variance of irradiation may be of little utility in solar collector design. Engels (1980) showed that in both January and July at Madison WI, North Omaha NE, Bismark ND, and Columbia MO, the mode of the negatively-skewed hourly attenuation was between 0.90 and 1.05. He introduced a time series–Markov model of this hourly attenuation, which included a compound probability density function consisting of two normal distributions separated by a "cut point" and enhanced with a "delta function" located near the mode.

The objective of this study is to describe a probability distribution model introduced by Ramberg *et al.* (1979), namely the generalized lambda distribution, and point out its effectiveness in fitting daily solar radiation totals. In the following section the general definition of the distribution and its properties are reviewed. Parameter estimation procedures are discussed in Section 3, and the application of the model on a data set for 12 weekly periods during the summer at West Lafayette is illustrated in Section 4.

2. The generalized lambda distribution and its properties

The generalized lambda distribution (GLD) is a generalization of Tukey's (1960) three-parameter lambda distribution. Ramberg and Schmeiser (1974) generalized Tukey's lambda distribution by introducing one more parameter. They used the underlying distribution to include unimodal asymmetric distributions and to provide an algorithm for generating asymmetric random variables. Ramberg *et al.* (1979) indicated the usefulness of the distribution for representing data, especially when the underlying model is unknown, and much of the following description, up to the fitting of solar radiation data, is a summary of their study.

The GLD is obtained by the transformation of the uniform distribution. Let P be a uniformly distributed random variable in the interval zero to one. The random variable which has a GLD is defined as

$$X = R(P) = \lambda_1 + [P^{\lambda_3} - (1 - P)^{\lambda_4}]/\lambda_2, \quad 0 \leqslant P \leqslant 1, \tag{1}$$

where λ_1, λ_2, λ_3 and λ_4 are constants. The probability density function of X may be obtained by a change of variables as

$$f(x) = f[R(p)] = 1/R'(p) = \lambda_2[\lambda_3 p^{\lambda_3 - 1} + \lambda_4(1 - p)^{\lambda_4 - 1}]^{-1}, \quad 0 \leqslant p \leqslant 1, \tag{2}$$

where $R'(p)$ is the derivative of $R(p)$. Although λ_1 disappeared in (2), the density function $f(x)$ is also a function of λ_1, since it is included in the definition of the random variable X. Therefore, the distribution has four parameters: λ_1 is the location parameter, λ_2 the scale parameter, and λ_3 and λ_4 the shape parameters. In general, the mean and variance are determined by λ_1 and λ_2, respectively. Skewness and kurtosis are determined by λ_3 and λ_4 together.

There are four regions of parameters where GLD is legitimate. These regions, for which $f(x) \geqslant 0$ and $\int_{-\infty}^{\infty} f(x)dx = 1$, were tabulated by Ramberg and Schmeiser. Since $R(p)$ is an increasing function of p, the lower and upper bounds of X are $R(0)$ and $R(1)$, respectively. Obviously, these bounds also depend on the values of all parameters.

In order to illustrate the general shape and the flexibility of the density function, various plots were drawn in Fig. 1 for the respective combinations of parameter values shown in Table 1. The coefficient of skewness (α_3) and kurtosis (α_4) corresponding to each density function are also given in the last two columns of Table 1. Each density function has zero mean and unit variance. For checking computer program consistency, 5 of the 9 selected sets of parameter values in Table 1 and of the 13 sets used by Ramberg *et al.* (1979, Table 1) were common. Since the "simple closed" form of the distribution does not exist, x values were computed as a function of p. For the given values of λ_1, λ_2, λ_3 and λ_4, x values may be obtained by (1) and the function values of $f(x)$

by (2) for selected values of p. Then $f(x)$ is plotted on the y-axis versus $x = R(p)$ on the x-axis.

As it may be seen from Fig. 1, the four parameter distributions include a wide range of curve shapes. By choosing the appropriate values of the parameters, it is possible to approximate the J- or U-shaped distributions, as well as the symmetric or asymmetric distributions. For example, Schmeiser (1977) has shown that the limiting distribution of GLD is exponential with parameter θ as $\lambda_4 \rightarrow 0$ when $\lambda_1 = \lambda_3 = 0$ and $\lambda_2 = \lambda_4/\theta$. Schmeiser also mentioned that the distribution with $\lambda_1 = 0$, $\lambda_2 = 0.1975$ and $\lambda_3 = \lambda_4 = 0.1349$ results in an approximation to the standard normal distribution for which the maximum absolute difference between the cumulative distribution functions is $\sim$0.001.

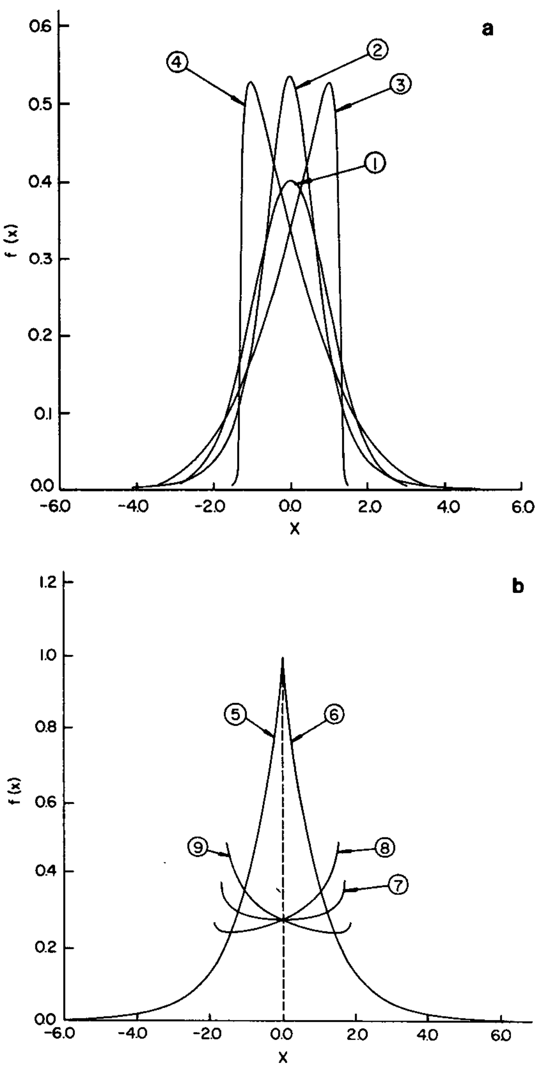

FIG. 1. GLD density functions for specified parameter values given in Table 1.

TABLE 1. Values of parameters used to plot the frequency curves in Fig. 1.

Curve no.	λ_1	λ_2	λ_3	λ_4	α_3	α_4
1	0.0000	0.1975	0.1349	0.1349	0.000	3.000
2	0.0000	−0.3203	−0.1359	−0.1359	0.000	9.000
3	1.1872	0.1761	0.2683	0.0048	−0.957	3.452
4	−1.1872	0.1761	0.0048	0.2683	0.957	3.452
5	0.0000	−0.00058	−0.00058	0.0000	−2.000	9.000
6	0.0000	−0.00058	0.0000	−0.00058	2.000	9.000
7	0.0000	0.5943	1.4501	1.4501	0.000	1.750
8	−0.1660	0.5901	1.1773	1.7680	−0.200	1.800
9	0.1660	0.5901	1.7680	1.1773	0.200	1.800

3. Parameter estimation

The maximum likelihood estimates of the parameters for GLD are difficult to obtain since it has no "simple closed" form. However, moment estimates of the parameters may be obtained. Although the sample moments are sensitive to extreme observations, this method has been proposed because of its simplicity and widespread use.

Ramberg and Schmeiser derived the following expressions for the mean, variance, third and fourth central moments of the GLD:

$$\left.\begin{aligned} \mu &= \lambda_1 + A/\lambda_2 \\ \sigma^2 &= (B - A^2)/\lambda_2^2 \\ \mu_3 &= (C - 3AB + 2A^3)/\lambda_2^3 \\ \mu_4 &= (D - 4AC + 6A^2B - 3A^4)/\lambda_2^4 \end{aligned}\right\}, \quad (3)$$

where

$$\left.\begin{aligned} A &= 1/(1 + \lambda_3) - 1/(1 + \lambda_4) \\ B &= 1/(1 + 2\lambda_3) + 1/(1 + 2\lambda_4) \\ &\quad - 2\beta(1 + \lambda_3, 1 + \lambda_4) \\ C &= 1/(1 + 3\lambda_3) - 3\beta(1 + 2\lambda_3, 1 + \lambda_4) \\ &\quad + 3\beta(1 + \lambda_3, 1 + 2\lambda_4) - 1/(1 + 3\lambda_4) \\ D &= 1/(1 + 4\lambda_3) - 4\beta(1 + 3\lambda_3, 1 + \lambda_4) \\ &\quad + 6\beta(1 + 2\lambda_3, 1 + 2\lambda_4) \\ &\quad - 4\beta(1 + \lambda_3, 1 + 3\lambda_4) + 1/(1 + 4\lambda_4) \end{aligned}\right\}. \quad (4)$$

In the equations above, β stands for the beta function.

Coefficients of skewness and kurtosis which are the standardized third and fourth moments, respectively, are given by

$$\alpha_3 = \mu_3/\sigma^3, \quad (5)$$

$$\alpha_4 = \mu_4/\sigma^4. \quad (6)$$

TABLE 2. Mean, standard deviation, and coefficient of skewness and kurtosis for weekly averages of daily mean solar radiation for indicated weekly period, West Lafayette, Indiana, 1958–80.

Period no.	Period	Mean (ly day^{-1})	Standard deviation	α_3	α_4
1	5–11 July	528.478	83.264	0.182	3.436
2	12–18 July	539.547	62.420	−0.699	3.983
3	19–25 July	499.118	76.199	−0.803	2.819
4	26 July–1 Aug.	498.068	81.004	−0.716	2.870
5	2–8 Aug.	467.727	74.979	0.093	2.622
6	9–15 Aug.	461.460	74.064	−0.224	2.331
7	16–22 Aug.	467.323	73.132	−1.199	3.702
8	23–29 Aug.	454.441	62.928	−0.485	2.563
9	30 Aug.–5 Sep.	415.913	68.229	−0.281	2.446
10	6–12 Sep.	411.087	57.388	−0.204	3.615
11	13–19 Sep.	345.497	55.416	−0.237	2.410
12	20–26 Sep.	329.360	59.958	−0.249	2.901

The coefficients α_3 and α_4 depend only on the parameters λ_3 and λ_4. This property of the distribution provides an important simplicity in obtaining the estimates of the parameters.

As in the usual case, moment estimates of the GLD may be obtained by equating the first four sample moments to the corresponding moments of the distribution. The sample moments about the mean may be expressed as

$$m_r = n^{-1} \sum_{i=1}^{n} (x - \bar{x})^r, \quad r = 2, 3, \ldots, \tag{7}$$

where n is the total number of observations and $\bar{x}$ the sample mean. From (5) and (6) sample estimates of α_3 and α_4 are obtained as

$$\hat{\alpha}_3 = \frac{m_3}{(m_2)^{3/2}}, \tag{8}$$

$$\hat{\alpha}_4 = \frac{m_4}{m_2^2}. \tag{9}$$

After having obtained the sample moments, the estimation procedure may be summarized as follows (Ramberg and Schmeiser, 1974):

1) Find the values of λ_3 and λ_4 by solving the two simultaneous equations

$$\left.\begin{aligned} \hat{\alpha}_3 &= \alpha_3(\lambda_3, \lambda_4) \\ \hat{\alpha}_4 &= \alpha_4(\lambda_3, \lambda_4) \end{aligned}\right\}, \tag{10}$$

where $\alpha_3(\lambda_3, \lambda_4)$ and $\alpha_4(\lambda_3, \lambda_4)$ are the coefficient of skewness and kurtosis as defined in (3), (5) and (6). The solution of these equations involves some difficulties. Tadikamalla (1975) computed $\hat{\lambda}_3$ and $\hat{\lambda}_4$ by using the subroutine ZSYSTM in the IMSL (the International Mathematical and Statistical Library) system. Ramberg *et al.* (1979) obtained an extended and corrected version of Tadikamalla's tables by using the Nelder and Mead simplex method for function minimization where the objective function

$$f(\lambda_3, \lambda_4) = [\hat{\alpha}_3 - \alpha_3(\lambda_3, \lambda_4)]^2 + [\hat{\alpha}_4 - \alpha_4(\lambda_3, \lambda_4)]^2, \tag{11}$$

was minimized over λ_3 and λ_4 subject to the constraint that $\lambda_3\lambda_4 > 0$.

2) Set the second central sample moment equal to the variance of GLD to obtain $\hat{\lambda}_2$ as

$$\hat{\lambda}_2 = \pm[(B - A^2)/m_2]^{1/2}, \tag{12}$$

where the sign of $\hat{\lambda}_2$ is the same as the sign of λ_3 and λ_4.

3) Set the sample mean equal to the GLD mean to obtain $\hat{\lambda}_1$ as

$$\hat{\lambda}_1 = \bar{x} - A/\hat{\lambda}_2. \tag{13}$$

As it may be seen from the above algorithm, the estimation procedure becomes rather easy once the simultaneous equations are solved for λ_3 and λ_4. It is also possible to obtain the estimates of the parameters by using the tables prepared by Ramberg *et al.* (1979). They also suggested the use of nonlinear least squares to minimize the squared distance between the observations (x_i) and the corresponding function values given by (1).

4. Computation of probabilities and quantiles

One of the basic goals of fitting a probability distribution model for a given set of data is to estimate probabilities and/or quantiles. For example, one might be interested in finding the probability of getting less than x_0 amount of solar radiation for a given period and/or location. Inversely, given a probability level p_0, estimating the amount of solar radiation corresponding to this probability level might be of interest.

TABLE 3. Parameter estimates and Kolmogorov-Smirnov (K-S) test statistics for daily solar radiation totals for indicated weekly period, West Lafayette, Indiana, 1958–80.

Period no.	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\lambda}_3$	$\hat{\lambda}_4$	K-S
1	517.410	0.0014	0.0636	0.0819	0.1380
2	568.283	0.0015	0.0789	0.0313	0.1582
3	598.149	0.0028	0.3829	0.0000	0.1046
4	589.374	0.0026	0.3444	0.0156	0.0887
5	450.374	0.0039	0.1705	0.2698	0.1568
6	522.577	0.0040	0.4700	0.0815	0.0754
7	554.645	0.0023	0.2453	0.0000	0.1105
8	515.923	0.0039	0.3892	0.0410	0.0622
9	468.070	0.0041	0.4070	0.0811	0.1108
10	418.557	0.0016	0.0600	0.0469	0.1211
11	386.418	0.0052	0.4221	0.0899	0.1057
12	349.353	0.0037	0.2045	0.1060	0.0799

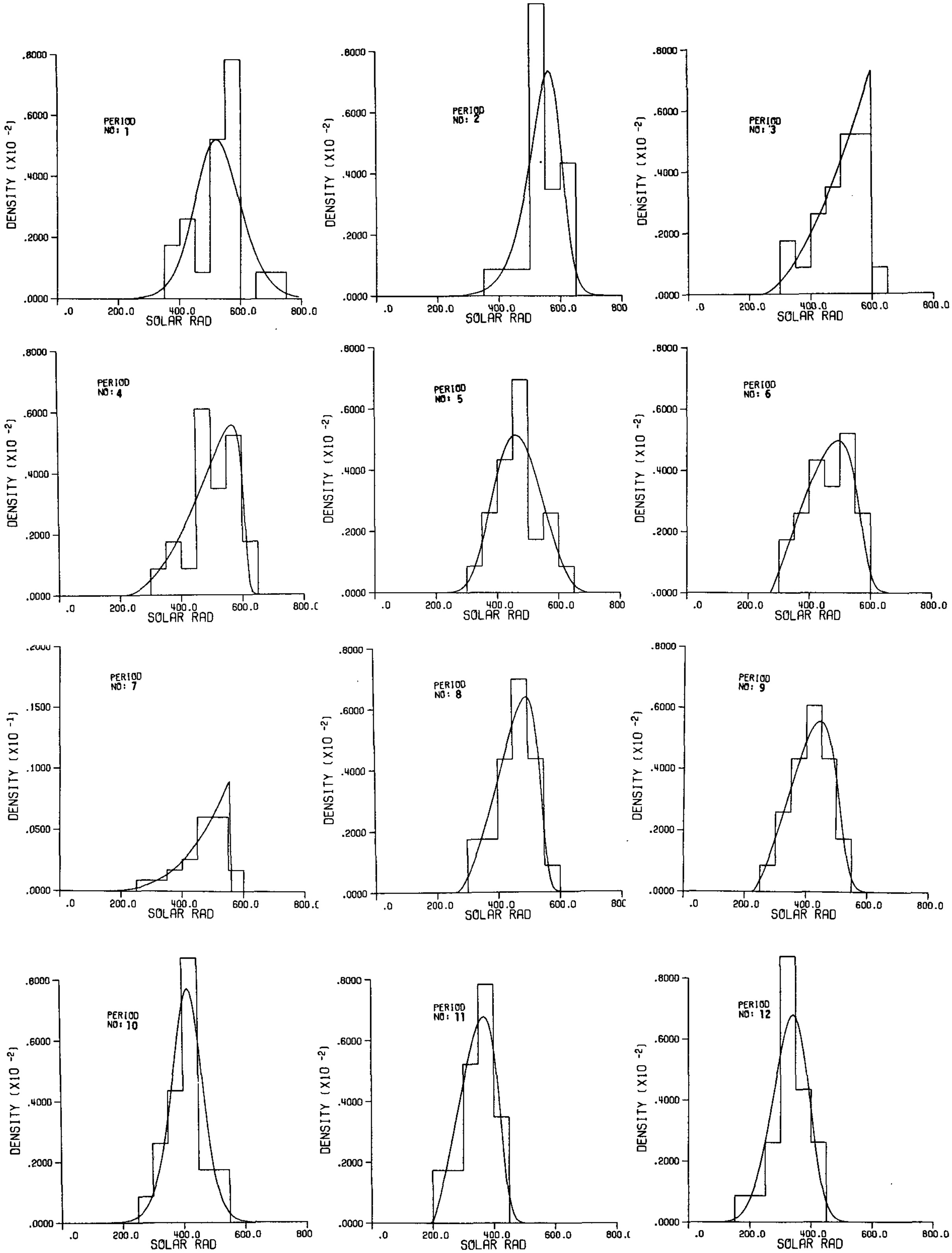

FIG. 2. Frequency histograms of daily mean solar radiation and fitted density curves for indicated weekly period (Table 2) at West Lafayette, IN, 1958–80.

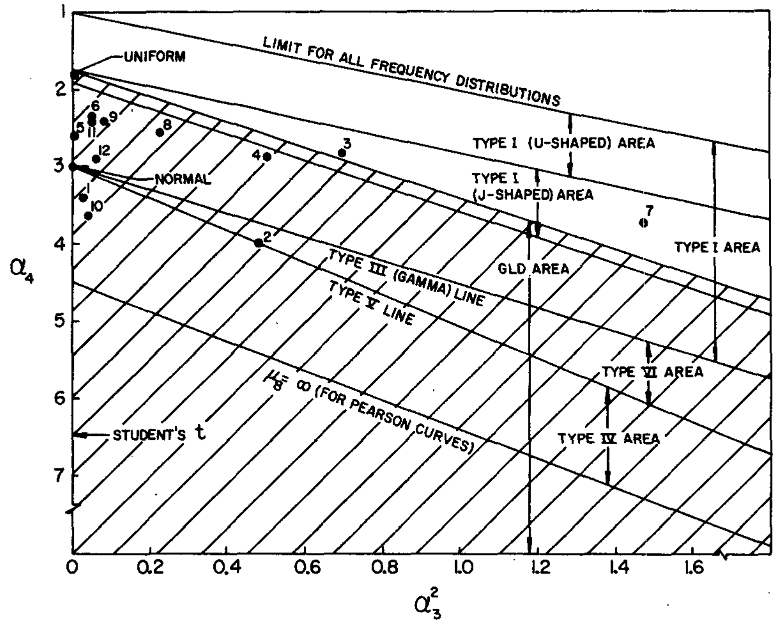

FIG. 3. Characterization of Pearson system of distributions and GLD by their skewness and kurtosis. Sample points for each week are numbered and shown with solid dots. Chart adapted from Johnson and Kotz (1970, p. 14) and Ramberg *et al.* (1979, p. 206).

The function $R(p)$ which was defined in (1) is the percentile (or quantile) function of the GLD. Given the parameter values, percentiles and quantiles of the distribution can easily be computed by using this function. For a given value of p (say p_0), a corresponding quantile is simply obtained as

$$x = R(p_0). \tag{14}$$

Hence, the quantile value (x) above satisfies the relation

$$p(X \leq x) = p_0 . \tag{15}$$

The probability of a given value of $x = x_0$ is obtained in a similar way by computing the inverse solution of

$$x_0 = R(p), \tag{16}$$

$$p = R^{-1}(x_0). \tag{17}$$

Let $G(p)$ be defined as

$$G(p) = R(p) - x_0 = 0. \tag{18}$$

Using the Newton-Raphson interative procedure, the ith approximation is obtained by

$$p_i = p_{i-1} - \frac{G(p_{i-1})}{G'(p_{i-1})}$$

$$= p_{i-1} - \frac{\lambda_1\lambda_2 + p_{i-1}^{\lambda_3} - (1 - p_{i-1})^{\lambda_4} - \lambda_2 x_0}{\lambda_3 p_{i-1}^{\lambda_3 - 1} + \lambda_4(1 - p_{i-1})^{\lambda_4 - 1}}, \tag{19}$$

where $G'(p_{i-1})$ is the derivative of $G(p)$ evaluated at p_{i-1}. Since $R(p)$ is an increasing function of p, no serious difficulty is involved with convergence. However, some protection is needed to prevent values from falling outside of the region ($0 < p < 1$).

5. Fitting GLD to solar radiation data

The GLD was fitted to the weekly averages of daily solar radiation totals from the Purdue University Agronomy Farm weather station, West Lafayette, Indiana. Twenty-three years of daily records for July, August and September were available for the period of 1958–80. One week periods in the climatological calendar were chosen to obtain 12 different sets of data. For example, 5–11 July was Period 1, and 20–26 September was Period 12. All periods are identified in Table 2. Daily solar radiation totals are in langleys per day of incident radiation on a horizontal surface.

A computer program was written in FORTRAN to compute the moment estimates of the parameters and to perform the other statistical analyses. In Table 2, some statistics are given for each week. As it may be seen from the table, distributions for 10 of the 12 weeks had negative coefficients of skewness ranging from −1.199 to −0.204. The remaining two coefficients are close to zero. These results demonstrate that even the means of daily solar radiation totals within weeks tend to show negative skewness. Coefficients of kurtosis were between 2.331 and 3.983.

Moment estimates of the parameters are given in Table 3. These values were obtained by following the algorithms explained, using the Nelder and Mead

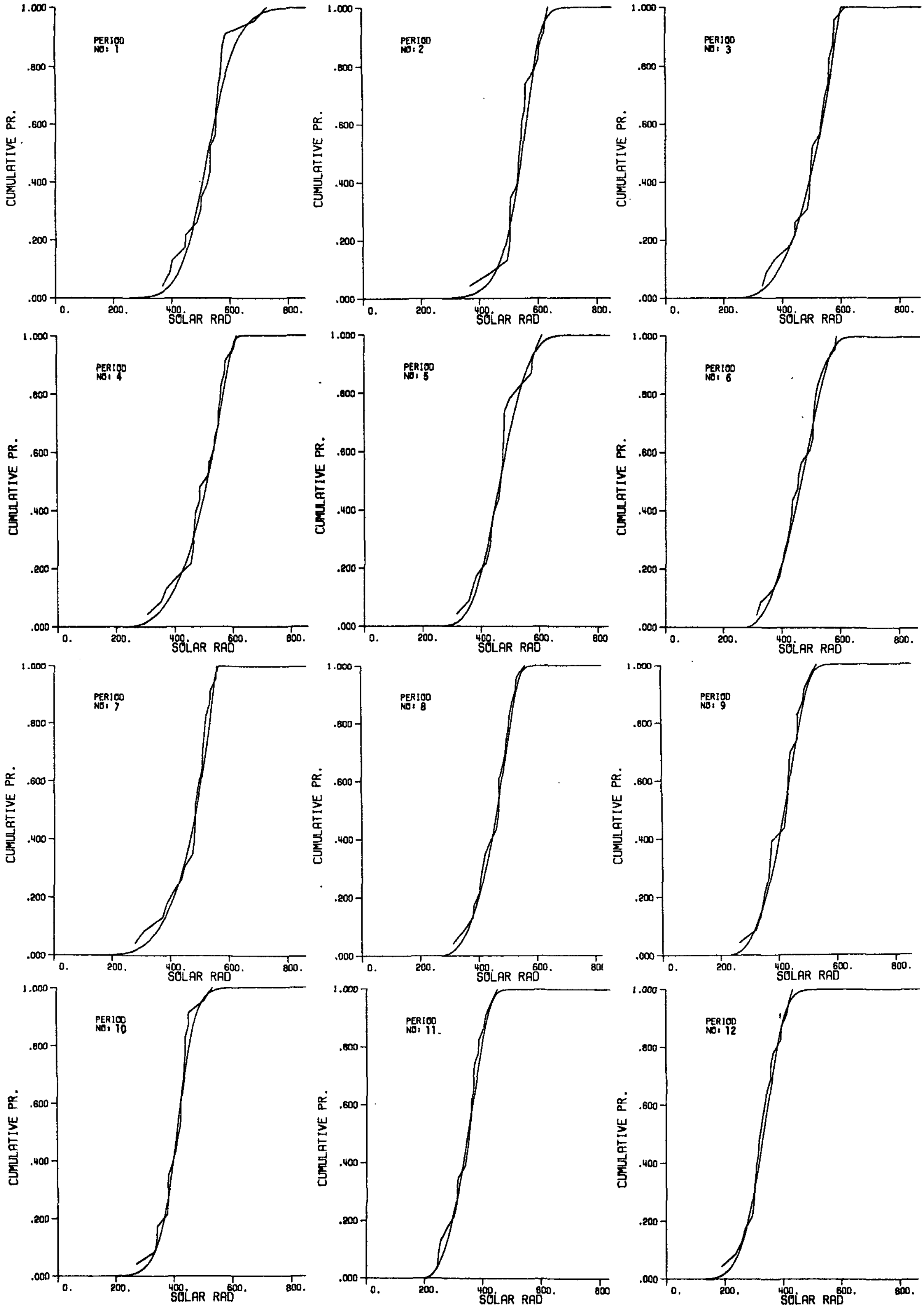

FIG. 4. Empirical and GLD-fitted cumulative probability functions of daily mean solar radiation for indicated weekly period at West Lafayette, IN, 1958–80.

TABLE 4. Probability of receiving less than the indicated mean daily solar radiation (ly day^{-1}) in the indicated weekly period at West Lafayette, Indiana.

Period no.	Probability										
	5	10	20	30	40	50	60	70	80	90	95
1	399.5	428.5	462.3	486.1	506.5	525.8	545.5	567.2	593.6	632.4	666.7
2	427.9	458.9	492.8	514.9	532.1	547.0	560.7	574.5	589.6	609.5	625.7
3	354.1	388.6	433.6	466.0	492.3	514.8	534.6	552.5	568.9	584.0	591.2
4	345.6	382.4	429.3	462.8	489.8	512.9	533.6	552.6	570.7	589.2	600.1
5	350.2	373.4	403.3	426.0	446.3	465.7	485.6	507.1	532.1	565.9	592.0
6	334.7	359.4	394.4	421.7	445.3	466.8	487.2	507.3	528.4	553.3	570.8
7	323.9	363.3	410.0	441.3	465.4	485.3	502.4	517.5	531.0	543.3	549.1
8	340.7	365.9	399.4	424.1	444.6	462.7	479.3	495.1	511.0	528.7	540.4
9	297.5	322.1	355.4	380.7	402.2	421.5	439.8	457.8	476.7	499.3	515.6
10	315.9	340.0	366.9	384.9	399.7	413.0	426.1	439.9	456.2	479.4	499.6
11	250.3	269.5	296.1	316.4	333.7	349.5	364.5	379.3	395.0	413.8	427.4
12	227.0	250.8	279.9	300.4	317.4	332.8	347.6	362.7	379.7	402.1	420.1

(1965) simplex method for function minimization. Note that although the mean (3) of the distribution is determined by the parameter λ_1, estimated values of this parameter are different from the sample means since the distributions are not symmetric ($\hat{\lambda}_3 \neq \hat{\lambda}_4$). On the other hand, truncated distributions were obtained for Period 3 and Period 7 with $\hat{\lambda}_4 = 0$. In these cases the distributions are J-shaped (Fig. 2).

When the distributions are characterized by their skewness and kurtosis, there is a certain area where the underlying probability distribution can be determined completely. For some of the probability distributions these areas are shown in Fig. 3, adapted from Johnson and Kotz (1970, p. 14) and Ramberg *et al.* (1979, p. 206). On Fig. 3, the uniform and normal distributions are each represented by a single point. The gamma, Student-*t* and Pearson Type V distributions are each represented by a line. The GLD and other distributions of the Pearson system are each represented by a region of values. Although GLD covers a wide region, the values beyond the GLD area either will result in a truncated distribution or the distribution will be undetermined. The location of the α_3^2 and α_4 statistics have been entered on Fig. 3 for each of the 12 sample periods. It may be seen that for the third and seventh week, for which the fitted distributions are truncated, sample points fall outside the GLD region. Although the Pearson system of distributions covers larger regions of the diagram than the GLD, they require more than one functional form. GLD uses only one function and is computationally simpler.

In order to test the goodness of fit of the distribution, Kolmogorov-Smirnov statistics were computed for each period and are given in the last column of Table 3. The Kolmogorov-Smirnov statistic is defined by

$$D = \max\left|F[x_{(i)}] - \frac{i}{n}\right|, \qquad (20)$$

where $x_{(i)}$ is the ith order statistic of the sample and $F[x_{(i)}]$ is the cumulative probability obtained by

TABLE 5. Probabilities of receiving less than the selected mean daily solar radiation (ly day^{-1}) in the indicated weekly period at West Lafayette, Indiana.

Period no.	Daily mean solar radiation								
	250	300	350	400	450	500	550	600	650
1	0.000	0.003	0.013	0.051	0.158	0.367	0.622	0.820	0.929
2	0.000	0.002	0.007	0.026	0.082	0.229	0.522	0.858	0.985
3	0.000	0.009	0.045	0.121	0.247	0.433	0.685	0.999	0.999
4	0.001	0.015	0.055	0.132	0.259	0.443	0.686	0.950	0.999
5	0.000	0.006	0.050	0.187	0.419	0.668	0.858	0.961	0.996
6	0.000	0.009	0.079	0.219	0.421	0.664	0.889	0.990	0.999
7	0.009	0.031	0.080	0.174	0.334	0.585	0.958	0.999	0.999
8	0.000	0.008	0.066	0.202	0.429	0.731	0.977	0.999	0.999
9	0.004	0.054	0.181	0.389	0.657	0.902	0.994	0.999	0.999
10	0.006	0.031	0.131	0.402	0.765	0.951	0.993	0.999	0.999
11	0.049	0.218	0.503	0.829	0.990	0.999	0.999	0.999	0.999
12	0.098	0.298	0.616	0.892	0.988	0.999	0.999	0.999	0.999

$$F[x_{(i)}] = \int_{-\infty}^{x(i)} f(x)dx. \qquad (21)$$

Here, $f(x)$ is the probability function defined as in (2). From the tabulated critical levels of Kolmogorov-Smirnov test statistics it is seen the GLD fits the data well.

In Fig. 4, the fitted and the empirical probability distributions are illustrated for each period. It is noted that although GLD seems to fail to fit in the left tail areas, the overall fittings are satisfactory. By the use of these graphs, quantiles and probabilities may be obtained approximately. Using Eqs. (14) and (15), some quantiles and probabilities for the mean daily solar radiation within each week were obtained and are given in Tables 4 and 5. For example, from Table 5 and Period 1 (5–11 July), the probability is 0.367 that the mean daily solar radiation will be less than 500 ly day^{-1}.

Recognizing the need for the probability of daily values the GLD was also fitted to the daily solar radiation totals for 7 × 23 = 161 observations in each week. Although the existence of persistence in the daily solar radiation is obvious, these results still give some idea about the fitting ability of the underlying probability distribution model. Fittings for all the weeks were found to be satisfactory. One example is given in Fig. 5 for Period 1, one of the periods with the least skewness of the weekly means. Yet, the daily solar radiation totals within that week show strong negative skewness. One can see the result of the central limit theorem empirically by comparing the frequency curve for the daily totals (Fig. 5) with the frequency curve for the daily means for the same week in Fig. 2.

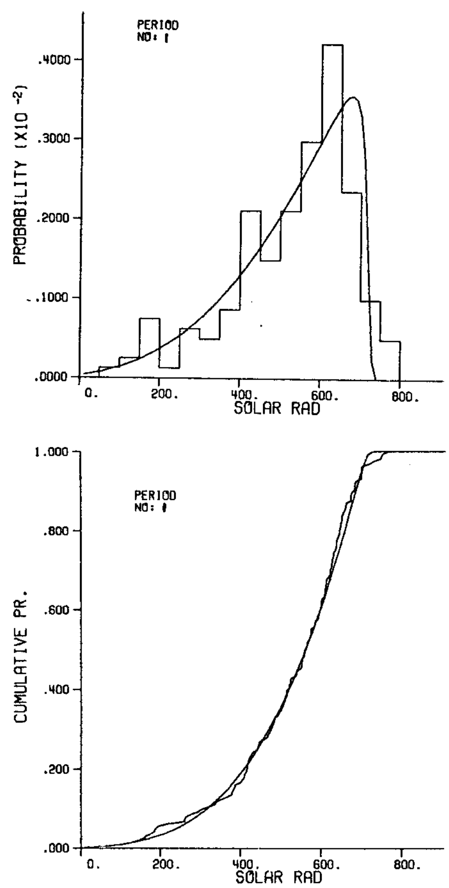

FIG. 5. Empirical and GLD-fitted frequencies and cumulated probability functions of daily solar radiation totals for weekly period 1, 5–11 July, at West Lafayette, IN, 1958–80.

6. Conclusion

The GLD is proposed as a probability distribution model to fit solar radiation data. Analysis of solar radiation data for 12 sample weeks shows that the underlying distribution fits the data well. Although there are some limitations in the use of the GLD, it can be used to fit the solar radiation data because of its flexibility and generality.

The variability of the solar radiation with respect to time results in appreciable differences between the distributions of daily totals from week to week. In some cases, these differences are so large that it becomes difficult to represent the empirical distributions by a unique probability distribution model. However, unless a certain probability model is specified to describe the general phenomenon of the terrestrial solar radiation data, it seems that the GLD could successfully be used to fit the wide variety of curve shapes observed.

REFERENCES

Baker, D. G., and D. A. Haines, 1969: Solar radiation and sunshine duration relationships in the North-Central Region (U.S.) and Alaska. NCRR Publ. 195 and University of Minnesota Agric. Expt. Sta. Tech. Bull., No. 262, St. Paul, 372 pp.

——, and J. C. Klink, 1975: Solar radiation reception, probabilities, and areal distribution in the (U.S.) North-Central Region (and Alaska). North Central Regional Res. Publ. 225 and University of Minnesota Agric. Expt. Sta. Tech. Bull., No. 300, St. Paul, 54 pp.

Bendt, P., M. Collares-Pereira and A. Rabl, 1981: The frequency distribution of daily insolation values. *Solar Energy*, **27**, 1–5.

Bennett, I., 1967: Frequency of daily insolation in Anglo North America during June and December. *Solar Energy*, **11**, 41–55.

Changnon, S. A., 1959: Summary of weather conditions at Champaign-Urbana, Illinois. Illinois State Water Survey Bull., No. 47, Urbana, 95 pp.

——, 1978: Solar energy information and data for Illinois. Illinois State Water Survey Circ., No. 133, Urbana, 61 pp.

Engels, J. D., 1980: Stochastic modeling of terrestrial solar radiation. Tech. Rep. 80-2, Dept. Industrial and Operations Engineering, University of Michigan, Ann Arbor. 85 pp.

——, S. M. Pollock and J. A. Clark, 1981: Observations on the statistical nature of terrestrial irradiation. *Solar Energy*, **26**, 91–92.

Getz, R. R., and M. M. Nicholas, 1979: Probabilities and extremes of solar radiation by climatic week. NOAA Tech. Memo.

NWS SR-98, Scientific Services Division, NOAA Southern Region, Fort Worth, 35 pp.

Johnson, N. L., and S. Kotz, 1970: *Continuous Univariate Distributions*, Vol. 1. Wiley, 300 pp.

McQuigg, J. D., and W. L. Decker, 1958: Solar energy, A summary of records at Columbia, Mo. Missouri Agric. Expt. Sta. Res. Bull., No. 671, 27 pp.

Nelder, J. A., and R. Mead, 1965: A simplex method for function minimization. *Computer J.*, **7,** 308–313.

Ramberg, J. S., and B. W. Schmeiser, 1974: An approximate method for generating asymmetric random variables. *Comm. ACM.*, **17,** 78–82.

——, E. J. Dudewicz, P. R. Tadikamalla and E. F. Mykytka, 1979: A probability distribution and its uses in fitting data. *Technometrics*, **21,** 201–214.

Rosenberg, N. J., 1964: Solar energy and sunshine in Nebraska. Nebraska Agric. Expt. Sta. Res. Bull., No. 213, 31 pp.

Schmeiser, B. W., 1977: Methods for modelling and generating probabilistic components in digital computer simulation when the standard distributions are not adequate: A survey. *Proc. of the Winter Simulation Conference*, H. J. Highland, R. G. Sargent, and J. W. Schmidt, Eds., Natl. Bur. Stnds., Gaithersburg, MD, 880 pp. (see pp. 51–57).

Tadikamalla, P. R., 1975: Modelling and generating stochastic inputs for simulation studies. Ph.D. thesis, The University of Iowa, Iowa City.

Tukey, J. W., 1960: The practical relationship between the common transformations of percentages of counts and of amounts. Tech. Rep. No. 36, Statistical Techniques Res. Group, Princeton University.

U.S. Dept. of Commerce, 1977: SOLMET users manual, Vol. 1. NOAA, Environmental Data Service, Asheville, NC, 44 pp.

Waite, P. J., and R. H. Shaw, 1961: Solar radiation and sunshine in Iowa. *Iowa State J. Sci.*, **35,** 355–365.