專 論

Frequency Analysis of Annual Debris Volume

岩滓與泥沙年沉積量之頻率分析

Raymond I. Jeng*

美國加州洛杉機加州州立大學土木系副教授

鄭 英 松

ABSTRACT

The annual debris volume, the sum of debris produced by individual storms occurring in one year, is considered as a kind of problem involing a random number of random variables. The number of debris producing storms occurring in one year is assumed to be a Poisson variate and the debris produced by a single storm is assumed to follow the normal probability distribution. A triple parameter, compound probability function of the annual debris volume is derived in an implicit form. The mean, variance, and skewness of annual debris volume are also derived from the moment-generating function. The derived compound probability function is used to fit debris data and a comparison with the log-Pearson distribution is presented.

INTRODUCTION

In many urban areas mud-debris flow due to heavy rainfall very often causes property damage or even disaster. Sometimes open channels fail to convey design flood flow because channels are clogged with debris. The capacity of reservoirs decreases due to debris accumulation after heavy rainfalls, and as a consequence, flood control and water conservation functions of the reservoirs are diminished. The mud-debris flow is a very complex phenomenon. It involves many factors still yet unknown. To date, a good approximate physical model describing mud-debris flow has not been successfully developed. Nor has a reliable mathematical equation been derived to relate debris volume to its causative factors. There are several ways, such as check dams, which are employed to try to reduce debris production. However, without a good understanding of the mud-debris flow, it is difficult to evaluate the effectiveness of check dams on the reduction of debris production in the long run. Statistical analysis of debris data alone cannot prove nor disprove the effectiveness of check dams. As a consequence, retaining debris at convenient places along the watershed and then removing accumulated debris at a more convenient time has become a common practice in dealing with debris problems.

* **Associate Professor, Department of Civil Engineering, California State University, Los Angeles**

To decide the capacity of debris basins to retain debris and tos et up the schedule to remove accumulated debris require a knowledge of the frequency of the debris volume produced by a watershed. In this instance, the total debris volume produced by a storm or the total debris volume during a given period of time is more important than the peak instantaneous debris flow rate. The purpose of this paper is to present a compound probability function for analysis of the frequency of the annual debris volumes produced by a watershed. In this study, the debris volume produced is considered as a stochastic process (i. e. debris volume is a time function and governed by some probability law). Debris volume produced by a storm is considered as a continuous random variable and assumed to follow the normal probability function. The number of storms which produce debris occurring in one year is also considered as a discrete random number and assumed to follow the Poisson probability law. The annual debris volume which is the sum of debris volume produced by individual storms occurring in one year is, therefore, a kind of problem involving a random number of random variables. A compound proqability function with three parameters is derived for the annual debris volume based on the Poisson and the normal probability distribution and is used to fit historical debris data. The comparison of this compound probability function and the log-Pearson distribution in fitting data is presented. Statistical parameters of the annual debris volume, such as mean, variance, and coefficient of skewness are obtained from the moment-generating function.

CHARACTERISTICS AND ASSUMPTIONS OF DEBRIS TIME SERIES

Debris phenomenon is stochastic in the sense that on the basis of the present state only probabilities of the future events may be estimated. This study of debris phenomenon follows in principle this probabilistic approach. With respect to the nature of the phenomenon, this approach is the most logical for the analysis and prediction of the future characteristics of the time series of debris.

For a debris stochastic process, if N_t stands for the number of storms producing debris in the interval of time from 0 to t, which gives the present state of the process, the number of storms, $N_{t,t+\Delta t}$, in the interval of time from t to $t+\Delta t$, can never be predicted with certainty for any $\Delta t>0$. In other words, the number of debris producing storms occurring in the time interval of t to $t+\Delta t$ is a random variable defined over some probability space for every $\Delta t>0$. Since the number of debris producing storms can only be a non-negative integer number, $N_{t,t+\Delta t}$ is, therefore, a discrete random variable. Only the probabilities.

$$Pr\{N_{t,t+\Delta t}=k\},\ k=0,\ 1,\ 2,\ 3,\cdots\cdots$$

where $\sum_{k=0}^{\infty} Pr\{N_{t,t+\Delta t}=k\}=1$ for all $t\geqq 0$, and $\Delta t>0$, may be estimated.

It should be noted here that in this study only debris producing storms are considered. Those storms which do not produce debris are not taken into account.

If $V_{i,j}$ stands for the volume of debris produced by the jth storm of the ith year, $V_{i,j}$ can not be predicted with certainty by the present state of the process. Since debris volume, $V_{i,j}$, produced by a storm is a non-negative real number, it is a continuous random variable. Only probability density.

$$\Pr\{x \leqq V_{i,j} \leqq x+\Delta x\} = f(x)dx,\ x \geqq 0$$

where $\int_0^\infty f(x)dx = 1$, may be determined.

In this study, the following properties of basic stochastic processes are assumed: Both N_t and $V_{i,j}$ are stationary stochasic processes. This implies that

$$\Pr\{N_{t,t+\Delta t}=k\} = \Pr\{N_{t+\Delta t,\ t+2\Delta t}=k\}$$
$$= \Pr\{N_{t+2\Delta t,\ t+3\Delta t}=k\} = \Pr\{N_{t+3+\Delta t,\ t+4\Delta t}=k\} = \cdots\cdots\cdots$$

and
$$\Pr\{V_{i,1} \leqq x\} = \Pr\{V_{i,2} \leqq x\} = \Pr\{V_{i,3} \leqq x\} = \cdots\cdots\cdots$$
$$= \Pr\{V_{i+1,1} \leqq x\} = \Pr\{V_{i+1,2} \leqq x\} = \cdots\cdots\cdots$$
$$= \Pr\{V_{i+2,1} \leqq x\} = \Pr\{V_{i+2,2} \leqq x\} = \cdots\cdots\cdots$$

It is assumed that no significant change in characteristics of a watershed are expected in the future. The results of this analysis can be applied only when these assumptions are indeed valid since the time-invariant parameters of the probability function are estimated from historical data.

In this study, it is assumed that the number of debris producing storms occurring in one year, $N_{t,t+1}$' a non-negative integer number, where t, which denotes years, follows the Poisson distribution. Therefore, the probabilities of the number of debris producing storms occurring in one year are

$$\Pr\{N_{t,t+1}=k\} = \frac{e^{-\lambda}\lambda^k}{k!} \quad \text{for } k=0,\ 1,\ 2,\ \cdots$$

The probability density function of the debris volume produced by one storm, $V_{i,j}$' is assumed to be the normal probability density function or

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/2\sigma^2}$$

The normal probability function may have a small probability for negative $V_{i,j}$', particularly when μ is small and σ is large. This small probability, however, may be ignored since only large debris volumes are of most concern.

COMPOUND PROBABILITY FUNCTION OF THE ANNUAL DEBRIS VOLUME

The total debris volume produced by one storm or during a given period of time rather than instantaneous peak debris flow rate during a storm is of primary importance for many debris problems. For convenience, one year period is selected as a unit of time. Therfore, the annual debris volume, V_i, which is the total debris volume produced in the ith year and is equal to the sum of debris produced by one or more storms occurring in that year, may be expressed as

$$V_i = \sum_{j=1}^{k_i} V_{i,j} = V_{i,1} + V_{i,2} + \cdots + V_{i,k_i} \cdots\cdots\cdots\cdots (1)$$

where k_i is the number of debris producing storms in the ith year, or $N_{i,i+1}=k_i$. Therefore, in this case, the number of random variables contributing to the sum is itself a random variable.

By using the conditional probability, the probability function of the annual debris volume, V, may be expressed as

$$Pr\{V \leqq x\} = \sum_{k=1}^{\infty} Pr\{V \leqq x / N_{t,t+1} = k\}\ Pr\{N_{t,t+1} = k\}$$

From the properties of the normal probability distribution, it is also known that if $V_{i,j}$ are independent normal random variables distributed with a mean μ and a variance σ^2, then the sum, V_i defined in equation 1, will also be normally distributed with mean $k\mu$ and variance $k\sigma^2$(2). Thus, the compound probability function of the annual debris volume is given by

$$F(x) = Pr\{V \leqq x\} = \sum_{k=1}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} \int_{-\infty}^{x} \frac{1}{\sqrt{2k\pi}\ \sigma} e^{-(x-k\mu)^2/2k\sigma^2} dx \cdots\cdots(2)$$

and the compound probability density function of the annual debris volume is given by

$$f(x) = \sum_{k=1}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} \frac{1}{\sqrt{2k\pi}\ \sigma} e^{-(x-k\mu)^2/2k\sigma^2} \cdots\cdots(3)$$

An unbiased estimator of three parameters of the compound probability function can be obtained in the following manner. The average number of debris producing storms in one year, λ, may be computed by the equation

$$\lambda = \sum_{k=1}^{\infty} kY_k \Big/ \sum_{k=0}^{\infty} Y_k \cdots\cdots(4)$$

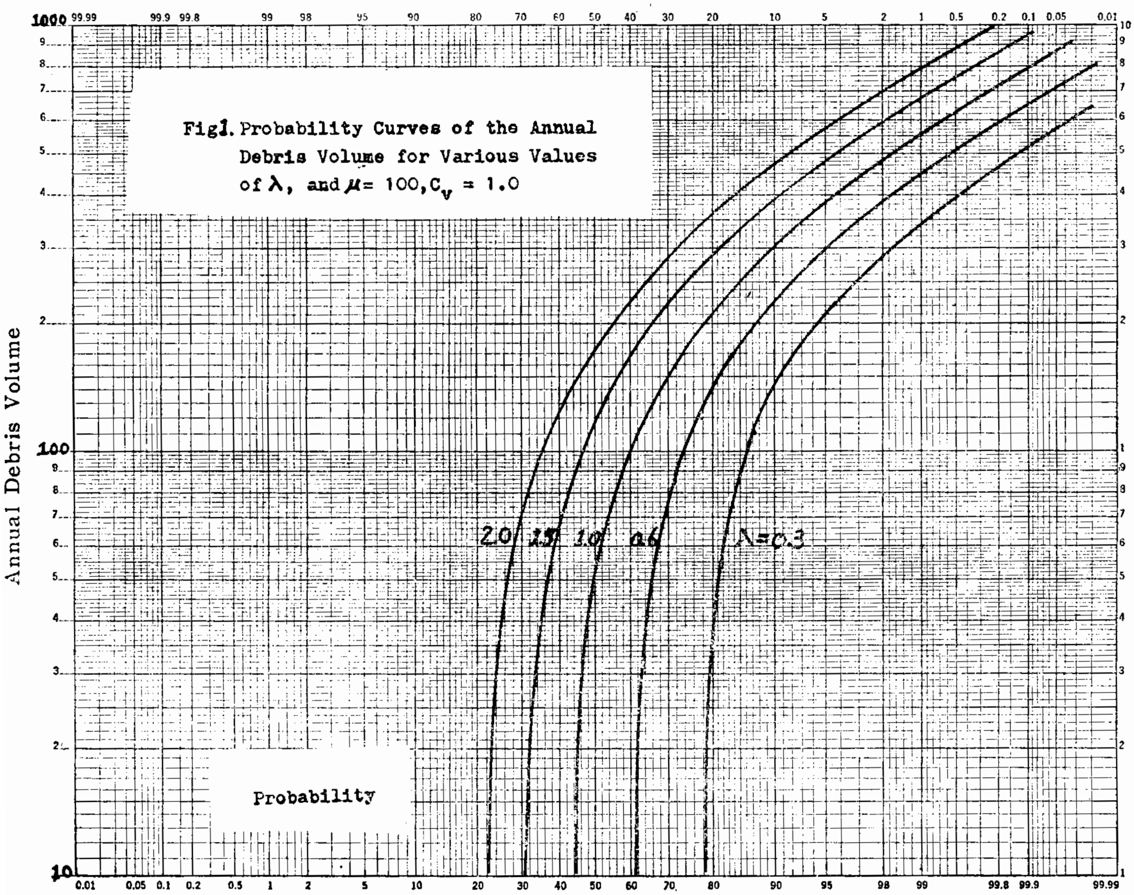

Fig1. Probability Curves of the Annual Debris Volume for Various Values of λ, and μ= 100, C_v = 1.0

where k is the number of debris producing storms in one year, and Y_k is the number of years having k number of debris producing storms. The average debris volume produced by a single storm, μ, may be estimated by

$$\mu = \sum_{i=1}^{n} \sum_{j=1}^{k_i} V_{i,j} \Big/ \sum_{i=1}^{n} k_i \qquad (5)$$

where $V_{i,j}$ is the debris volume produced by the jth storm of the ith year, k_i is the number of debris producing storms which occurred in the ith year, and n is the total length of the historical record in years. The variance of debris volume produced by a single storm, σ^2, may be determined by

$$\sigma^2 = \sum_{i=1}^{n} \sum_{j=1}^{k_i} (V_{i,j} - \mu)^2 \Big/ \left(\sum_{i=1}^{n} k_i - 1\right) \qquad (6)$$

Based on the present state of process, three parameters may be estimated from equations 4 to 6. Using the resulting triple parameter compound probability function,

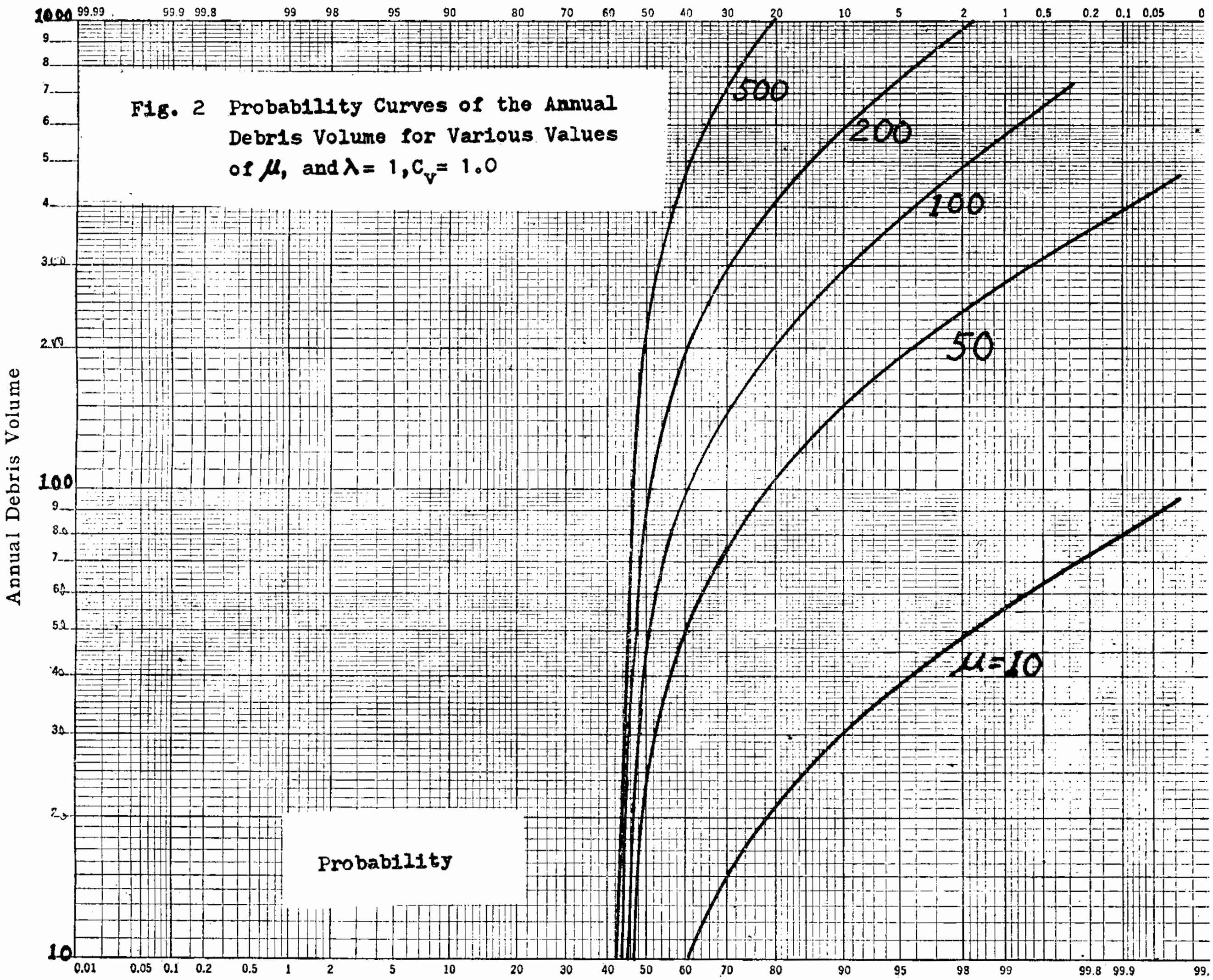

Fig. 2 Probability Curves of the Annual Debris Volume for Various Values of μ, and $\lambda = 1$, $C_v = 1.0$

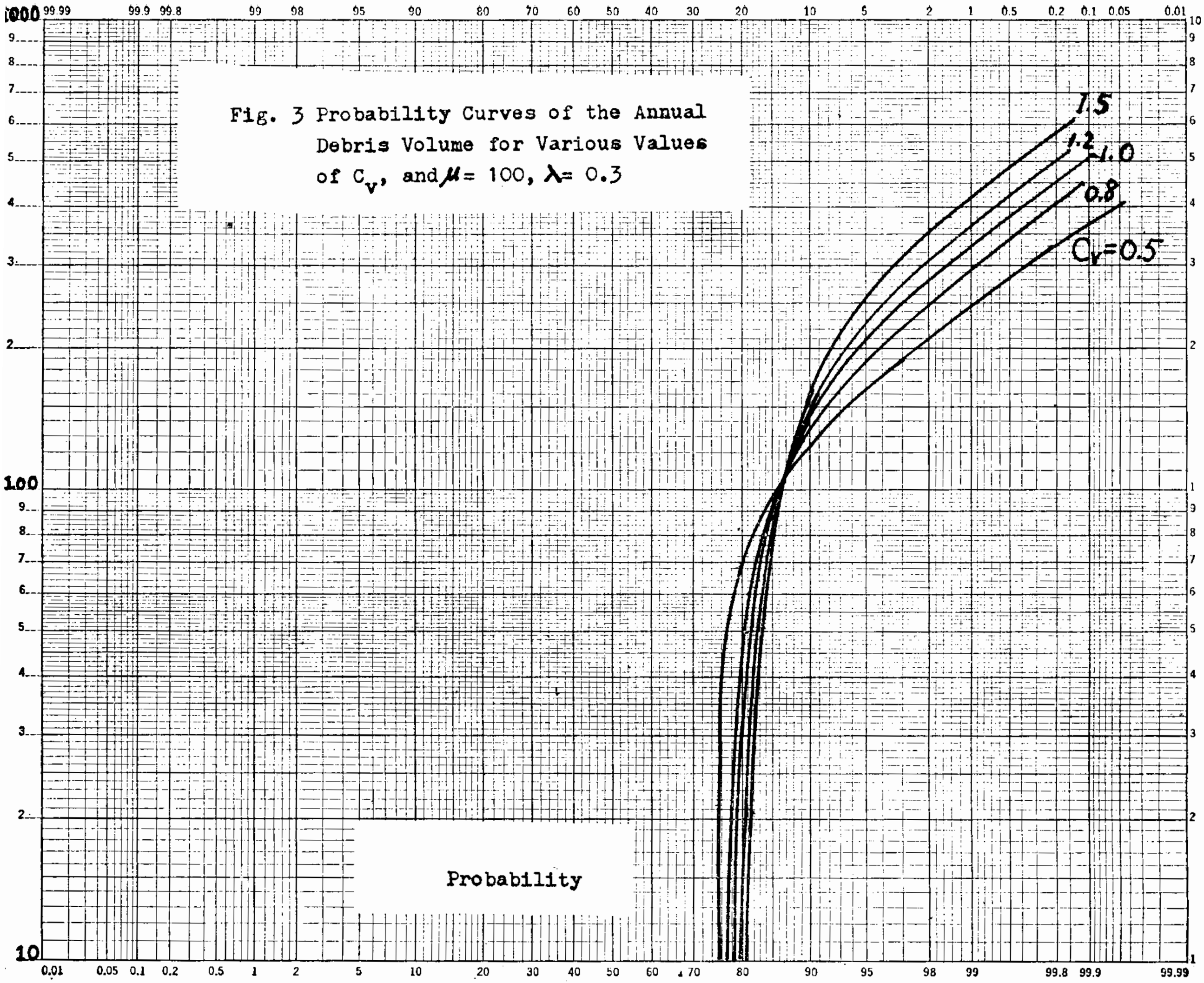

Fig. 3 Probability Curves of the Annual Debris Volume for Various Values of C_v, and μ= 100, λ= 0.3

the probabilities of the annual debris volumes for succeeding years may thus be estimated.

The compound probability function of the annual debris volume given by equation 2 is graphically shown for several values of μ, λ. and $C_v = \mu/\sigma$. coefficient of variation. Fig. 1 shows that, for given values of μ and σ, the probabilities of the annual debris volume exceeding a given value increase with increasing λ. Fig. 2 shows that that these probabilities will also increase with increasing μ for given values of λ and C_v. Finally, Fig. 3 shows that these probabilities also increase as C_v increases.

STATISTICAL PARAMETERS OF THE ANNUAL DEBRIS VOLUME

Sometimes it is desirable to be able to summarize some of the outstanding features of the compound probability function of the annual debris volume by specifying only a few parameters rather than an entire function. These parameters are the mean, the varance, and the coefficient of skewness. They are used to measure some important characteristics of the annual debris volume and may be easily obtained from the moment-generating function.

The moment-generating function of a probability law is a function m(t) defined for all real numbers t by[2]

$$m(t)=E[e^{tx}]$$

where the operator E indicates an expectation. In other words, m(t) is the expectation of the exponential function e^{tx}. According to this definition and the derived probability density function of the annual debris volume as given in equation 3, the moment-generating function m(t) will be

$$m(t)=\int_{-\infty}^{\infty} e^{tx}f(x)dx$$

$$=\sum_{k=1}^{\infty}\frac{e^{-\lambda}\lambda^k}{k!}\int_{-\infty}^{\infty} e^{tx}\frac{1}{\sqrt{2k\pi}\,\sigma}e^{-(x-k\mu)^2/2k\sigma^2}dx$$

$$=\sum_{k=1}^{\infty}\frac{e^{-\lambda}\lambda^k}{k!}e^{k\mu t+\frac{1}{2}k\sigma^2t^2}$$

$$=e^{-\lambda}\sum_{k=1}^{\infty}\frac{(\lambda e^{\mu t+\frac{1}{2}\sigma^2})^k}{k!}$$

$$=e^{-\lambda}[e^{\lambda e^{\mu t+\frac{1}{2}\sigma^2t^2}}-1] \quad\cdots\cdots(7)$$

It can be shown that all the moments of the compound probability function exist and may be expressed in terms of the successive derivatives at t=0 of the moment-generating function.[2] Thus, the mean of the annual debris volume, $\overline{V}$, based on the compound probability function is

$$\overline{V}=E[V]=\frac{d}{dt}m(t)\Big|_{t=0}=\lambda\mu \quad\cdots\cdots(8)$$

The second moment of the compound probability of the annual debris volume is given by

$$E[V^2]=\frac{d^2}{dt^2}m(t)\Big|_{t=0}=(\lambda\mu)^2+\lambda(\mu^2+\sigma^2) \quad\cdots\cdots(9)$$

The variance of the annual debris volume is given by

$$\text{Var } V=E(V-E(V))^2=E(V^2)-[E(V)]^2=\lambda\mu^2(1+C_V^2) \quad\cdots\cdots(10)$$

Similarly, the third moment of the compound probability function is given by

$$E(V^3)=\frac{d^3}{dt^3}m(t)\Big|_{t=0}=(\lambda\mu)^3+3\lambda^2\mu(\mu^2+\sigma^2)+\lambda\mu(\mu^2+3\sigma^2) \quad\cdots\cdots(11)$$

The third central moment, $E(V-E(V))^3$, if expanded has the form

$$E(V-E(V))^3=E(V^3)-3E(V)E(V^2)+2(E(V))^3=\lambda\mu^3(1+3C_V^2) \quad\cdots\cdots(12)$$

The coefficient of skewness of the anuual debris volume, C_S, is usually defined as

$$C_S=\frac{E(V-E(V))^3}{(\text{Var } V)^{3/2}}=\frac{(1+3C_V^2)}{\sqrt{\lambda}\,(1+C_V^2)^{3/2}} \quad\cdots\cdots(13)$$

It is clearly shown in Eq. 13 that the coefficient of skewness of annual debris volume, C_S, is always positive and is maximum when the coefficient of variation, C_V, is equal to unity for a given value of λ. This maximum skewness coefficient is equal to $\sqrt{2/\lambda}$. Thus the statistical parameters of annual debris volume may be expressed in terms of λ, μ, and C_V.

THE FITTING OF DEBRIS DATA

As an application of the derived method, the compound probability function of the annnal debris volume is used to fit historical debris data. For this purpose, debris data of the San Dimas reservoir collected for more than four decades by the Los Angeles County Flood Control District will be used. For the San Dimas reservoir, it is found from debris data that, over a period of 43 years, one debris producing storm occurred during twelve years, two storms during two years, and no significant debris flow occurred during the remaining years. Thus, from equation 4,

$$\lambda = \sum_{k=0}^{\infty} \frac{kY_k}{43} = \frac{12x1+2x2}{43} = \frac{16}{43} = 0.3721$$

From the debris data and equations 5 and 6, the mean and the standard deivation of debris volume for a single storm are $\mu = 107.38$ acre-ft. and $\sigma = 78.21$ acre-ft., respectively.

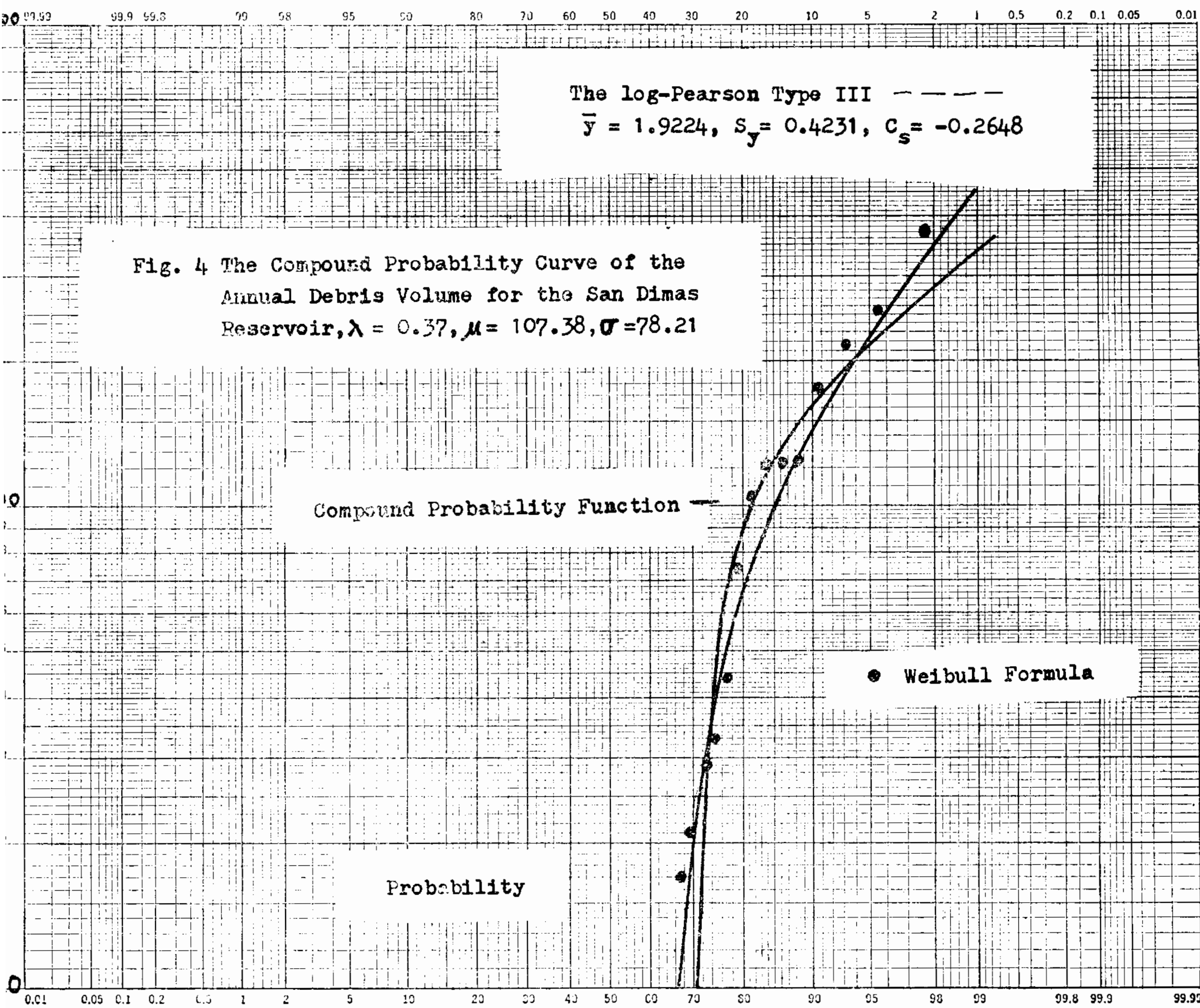

Fig. 4 The Compound Probability Curve of the Annual Debris Volume for the San Dimas Reservoir, $\lambda = 0.37, \mu = 107.38, \sigma = 78.21$

yielding a coefficient of variation of variation of $C_V=0.73$. The compound probability function of the annual debris volume for the San Dimas reservoir is, therefore, given by

$$F(x)=\sum_{k=1}^{\infty}\frac{e^{-0.372}(0.372)^k}{k!}\int_{-\infty}^{x}\frac{e^{-(x-107.38k)^2/2k(78.21)^2}}{\sqrt{2k\pi}\,(78.21)}dx$$

The probability of any annual debris volume may thus be easily estimated from the above equation by using the standarized normal table and the probability curve of the annual debris volume for the San Dimas reservoir which is shown in Fig. 4. To show the goodness of fit of the compound probability function, the Weibull formula may be used[1], that is,

$$Pr\{V_i \geq X_m\}=m/(N+1) \quad \cdots\cdots(14)$$

where X_m is the mth rank of the annual dedris volume in data in descending order and N is the total length of record.

It is desirable to compare the proposed compound probability function and some other probility functions that are now in common use. For this reason the log-Pearson type III distribution is selected.[4] If log V=y, then y may be expressed in terms of mean, $\bar{y}$, standard deviation, S_y, and frequency factor, k, which depends on skewness, C_S, and the probability by

$$y=\bar{y}+kS_y$$

These statistical parameters are computed by the following formula:

$$\bar{y}=\sum_{i=1}^{N_d} y_i/N_d,$$

$$S_y=\sum_{i=1}^{N_d}(y_i-\bar{y})^2/N_d-1$$

$$C_S=\left[\frac{N_d\sum_{i=1}^{N_d}(y_i-\bar{y})^3}{(N_d-1)(N_d-2)}\right]\Big/S_y^3$$

where N_d is the number of years that significant debris flow are recorded. From debris data of the San Dimas reservoir with $N_d=14$, it is found that

$$\bar{y}=1.9224,\ S_y=0.4231,\ \text{and}\ C_S=-0.2648$$

Since the probability, P_1, of the annual debris volume estimated with the log-Pearson distribution[3] is based only on the 14 years during which debris flow occurred, it must be adjusted for the entire length of the record. This is accomplished by multiplying P_1 by $N_d/N=14/43$. That is,

$$P=P_1\ (14/43)$$

where P is the probability of the annual debris volume for the entire length of record. The probability curve of the annual debris volume for the San Dimas reservoir based on the adjusted log-Person distribution is also shown on the same figure for the comparison.

As shown on the Fig., the compound probility function and the adjusted log-Pearson distribution behave very differently. The adjusted log-Pearson distribution seems to predict low annual debris volumes quite well, but it underestimates the intermediate annual debris volumes. In contrast, the compound probability function seems to fit intermediate annual debris volumes very well but underestimates low annual debris volumes. The other important difference is that the annual debris volume for the compound probability function does not increase as fast as for the adjusted log-Pearson distribution for the same probability.

CONCLUSIONS

The following conclusions may be stated based on the results of this study:

1. The assumptions of the Poisson distribution for the number of debris storms occurring in one year and of the normal probability distribution for the debris volume produced by a single storm are not necessary for establishing a compound probability function, though they are used in this study to illustrate the method.

2. The compound probability function seems to fit debris data quite well, particularly for intermediate and high annual debris volume. Since high annual debris volumes are of most concern, this approach of using a compound probability function is of practical importance in the frequency analysis of the annual debris volume.

3. This approach can also be applied to some other hydrologic variables, besides the annual debris volume when hydrologic variables are the sum of random variables and the number contributing to the sum is itself a random number.

4. Since the adjusted log-Pearson distribution seems to overestimate high annual debris volume and to underestimate intermediate annual debris volume, it should be used with extreme care when high and intermediate debris volume are of particular concern in the frequency analysis.

5. Estimation of probability based on the proposed compound distribution requires one first moment and two second moment parameters while estimation of probability based on the log-Pearson distribution requires one first moment, one second moment, and one third moment parameter. It is well known that high moment parameters estimated with limited data are not very reliable. For this reason, the proposed compound distribution is more desirable in the frequency analysis of annual debris volume, particularly when debris data is very limited.

6. Since the proposed method uses debris volume produced by individual storms to estimate parameters, it fully utilizes the limited data and abstracts the maximum information from the data. As shown in the example of San Dimas reservoir, sixteen individnal debris volumes were used for the proposed compound distribution. Only fourteen annual debris volumes were used for the log-Pearson distribution.

REFERENCES

1. Chow, V.T. *Handbook of Applied Hydrology*. Mc Graw-Hill Book Company, Inc. New York, 1964.
2. Parzen, E. *Modern Probability Theory and Its Applications*. John Wiley & Sons, New York, 1963.
3. U. S. Department of Agriculture, Soil Conservation Service, Technical Release No. 38. *New Tables of Percentage Points of the Pearson Type III Distribution.*
4. Water Resources Council, Bulletin No. 15. *A Uniform Technique for Determiniug Flood Flow Frequencies.* Washington D.C., 1967.

鄭英松博士爲臺大農工系 1963 年畢業生，1968 年美國科羅拉多州立大學土木系哲學博士，現任美國加州洛杉機加州州立大學土木系副教授。按本文爲美國土木工程師學會洛杉機分會與加州長灘 (Long Beach) 加州州立大學土木系合辦之都市水文研究班 (Urban Hydrology Work shop) 發表之講稿，蒙著者同意，在本學報刊載，以享讀者。——編者附註。