A Further Study of the Estimation of Area Rain-Rate Moments by the Threshold Method

By Kunio Shimizu

Department of Applied Mathematics, Science University of Tokyo, Shinjuku-ku, Tokyo 162, Japan

and

Koichi Kayano

Technical Administration Division, Toyota Motor Corporation, Toyota City, Aichi 471, Japan

(Manuscript received 21 April 1994, in revised form 12 September 1994)

Abstract

The mixed distribution, a mixture of the discrete mass at the origin and the continuous distribution, conditional on rain, is adopted as a rain rate distribution. The k th moment of rain rate is linearly dependent on the probability that rain rate exceeds a fixed threshold level. The paper provides an asymptotic (large-sample) property of the maximum likelihood (ML) estimator of the proportionality constant, called slope. It numerically shows that the ML estimators of the slopes for the first and second moments are mutually asymptotically almost independent for the GATE I parameter if the thresholds are chosen to be optimal in the sense of minimizing the normalized asymptotic variances. The property provides a reason why the optimal thresholds for estimating the first and second moments of area rain rate should be used for estimating the variance by the double threshold method. The paper also proposes a method of choosing thresholds simultaneously for the estimation of the first and second moments of area rain rate.

1. Introduction

It has been observed (Chiu, 1988; Deneaud *et al.*, 1984; Rosenfeld *et al.*, 1990; Short *et al.*, 1993b) from radar and raingage network data that the area-average rain rate and the fractional area covered by rain rate exceeding a fixed threshold level are highly linearly related. The fact provides the basis of the so-called threshold method. The method is attractive because area-average rain rate is estimable from the fractional area that rain rate exceeds a fixed threshold level, and may be useful for measuring rainfall from an orbiting spacecraft as well as from ground-based systems.

Several authors have studied the statistical explanation of the empirical fact. Among others, Kedem *et al.* (1990a, b) derived a linear relationship between the mean of rain rate and the probability that rain rate exceeds a fixed threshold under the statistical setup that rain rate forms a mixed distribution. Kedem and Pavlopoulos (1991) introduced the maximum likelihood (ML) approach to choose

an optimal threshold. An extension to the case of the rain rate k th moment was given by Short *et al.* (1993a), who noted a linear relationship between the k th moment and the probability that rain rate exceeds a fixed threshold level. Using the Global Atmospheric Research Program (GARP), Atlantic Tropical Experiment (GATE), Phase I data set, they showed a good agreement between theoretically and empirically determined optimal thresholds for estimating the first and second moments. The variance, a dispersion measure, of rain rate is quadratically expressed by the probability that rain rate exceeds a fixed threshold level. Optimal thresholds for estimating the variance by the ML approach were studied by Shimizu *et al.* (1993). Some knowledge of the variance of rain rate within an area may be useful for distinguishing convective and stratiform regimes.

The subject of the present paper is closely connected with those considered by the papers above, especially the last two, but there is some difference as follows. After a brief description of preliminary results in Section 2, the main result of this paper will be presented in Section 3. It will be numerically

shown that the ML estimators of the slopes for the first and second moments are mutually asymptotically almost independent for the GATE I parameter if the thresholds are chosen to be optimal in the sense of minimizing normalized asymptotic variances as in the paper by Short *et al.* (1993a). The fact provides a reason why optimal thresholds for estimating the first and second moments of area rain rate should be used for estimating the variance by the double threshold method introduced by Shimizu *et al.* (1993). Based on the result stated in Section 3, we shall propose a method of the simultaneous choice of thresholds for the estimation of the first and second moments of area rain rate in Section 4, while Short *et al.* (1993a) have treated the individual choice of optimal thresholds for the estimation of the k th moment. The method may appeal to one's intuition because the thresholds are chosen in such a way that the correlation between the estimates of two slopes for the first and second moments is zero. Investigation of its practical performance, however, is beyond the scope of the present paper.

2. Preliminaries

This section briefly describes the statistical setup and preliminary results.

A collection of non-negative random variables $\{X_t : t \in [0, \infty)\}$ forms a stochastic process defined on Ω , an abstract sample space, where t denotes time. For each t , X_t can be interpreted as either instantaneous rain rate at a fixed geographic point or instantaneous area-average rain rate over a fixed region, and it is assumed that X_t follows a mixed-type distribution. The probability that X_t takes the value 0 (no rain) is $1 - p_t$ ($0 < p_t < 1$), which forms the discrete part of the distribution. For given $X_t > 0$ (conditional on rain), the distribution of X_t is continuous. The continuous part is characterized by the continuous distribution function $F_t(x, \theta_t)$ with (vector) parameter θ_t . Thus, the unconditional distribution function, $G_t(x)$, of X_t is

$$G_t(x) = \Pr(X_t \leq x) = (1 - p_t)H(x) + p_t F_t(x, \theta_t),$$

where $H(x)$ denotes a step function such that $H(x) = 1$ if $x \geq 0$ and 0 if $x < 0$.

Short *et al.* (1993a) show that the linear relationship between the k th moment of rain rate and the probability that rain rate exceeds a fixed threshold level τ (≥ 0):

$$E(X_t^k) = \beta_t(k, \tau, \theta_t) \Pr(X_t > \tau) \quad (1)$$

provided that X_t admits the population k th moment, where

$$\beta_t(k, \tau, \theta_t) = \frac{E(X_t^k | X_t > 0)}{\Pr(X_t > \tau | X_t > 0)},$$

which is called the slope, independent of the probability of rain. Equation (1) includes a special case of the expected value ($k = 1$), whose case (Kedem *et al.*, 1990a, b; Kedem and Pavlopoulos, 1991) was used to statistically explain Chiu's (1988) empirical fact. We assume the existence of the distribution function G on $[0, \infty)$ as the weak limit of G_t , $t \rightarrow \infty$. In this case the slope is written $\beta(k, \tau, \theta)$. Some conditions for the existence were given by Pavlopoulos and Kedem (1992).

Short *et al.* (1993a) treated two-parameter log-normal, gamma and inverse-Gaussian distributions as the continuous part of the population distribution. All of these distributions are positive, continuous and skewed. The frequency curves resemble each other after the parameter adjustment by the method of moments. The genesis of the distributions, however, is quite different. This leads to a particular tail property, and the distributions can represent different rain patterns. In this paper we shall add the two-parameter Weibull distribution to these distributions. Thus, we assume that θ is a two-dimensional vector, $\theta = (\theta_1, \theta_2)'$; the continuous part of the distribution has two parameters. The Fisher information matrix of the positive continuous distribution with vector parameter θ is denoted by $\mathbf{I}(\theta) = \mathbf{I}(\theta_1, \theta_2)$, which is a 2×2 symmetric matrix. Noting that $\beta(k, \tau, \theta)$ is dependent only on the continuous part of the distribution, we consider the ML estimator, $\hat{\beta}(k, \tau, \theta) = \beta(k, \tau, \hat{\theta})$, of $\beta(k, \tau, \theta)$, which can be obtained from a random sample, $Y_1, \dots, Y_n$, of size n whose distribution function is $F(x, \theta)$, the continuous part of G . Here $\hat{\theta}$ is the ML estimator of θ . From the univariate δ method (Kedem and Pavlopoulos, 1991), $\hat{\beta}(k, \tau, \theta)$ asymptotically has a normal distribution in the sense that

$$\sqrt{n} \left[\hat{\beta}(k, \tau, \theta) - \beta(k, \tau, \theta) \right] \xrightarrow[n \rightarrow \infty]{L} N(0, v(k, \tau, \theta)),$$

where $v(k, \tau, \theta)$ stands for the asymptotic variance defined by

$$v(k, \tau, \theta) = \left(\frac{\partial \beta(k, \tau, \theta)}{\partial \theta_1} \frac{\partial \beta(k, \tau, \theta)}{\partial \theta_2} \right) \mathbf{I}^{-1}(\theta) \times \left(\frac{\partial \beta(k, \tau, \theta)}{\partial \theta_1} \frac{\partial \beta(k, \tau, \theta)}{\partial \theta_2} \right)'$$

and $\mathbf{I}^{-1}(\theta)$ the inverse matrix of the Fisher information matrix $\mathbf{I}(\theta)$. Following the ML approach of Kedem and Pavlopoulos (1991), Short *et al.* (1993a) sought the optimal threshold that minimizes the normalized asymptotic variance $v(k, \tau, \theta) / \{\beta(k, \tau, \theta)\}^2$ with respect to τ .

3. Asymptotically almost independent property

3.1 Asymptotic distribution of the ML estimators for slopes

To evaluate the asymptotic variance-covariance matrix between $\hat{\beta}(k_1, \tau_1, \theta)$ and $\hat{\beta}(k_2, \tau_2, \theta)$, we use the multivariate δ method (cf. Bishop *et al.*, 1980, p. 492; see also Shimizu *et al.*, 1993). Thus, $(\hat{\beta}(k_1, \tau_1, \theta), \hat{\beta}(k_2, \tau_2, \theta))'$ asymptotically has a bivariate normal distribution:

$$\sqrt{n} \left[\begin{pmatrix} \hat{\beta}(k_1, \tau_1, \theta) \\ \hat{\beta}(k_2, \tau_2, \theta) \end{pmatrix} - \begin{pmatrix} \beta(k_1, \tau_1, \theta) \\ \beta(k_2, \tau_2, \theta) \end{pmatrix} \right] \xrightarrow{L} N_2(0, V(k_1, k_2, \tau_1, \tau_2, \theta)), \quad n \rightarrow \infty,$$

where the asymptotic variance-covariance matrix is

$$V(k_1, k_2, \tau_1, \tau_2, \theta) = \begin{pmatrix} \frac{\partial \beta(k_1, \tau_1, \theta)}{\partial \theta_1} & \frac{\partial \beta(k_1, \tau_1, \theta)}{\partial \theta_2} \\ \frac{\partial \beta(k_2, \tau_2, \theta)}{\partial \theta_1} & \frac{\partial \beta(k_2, \tau_2, \theta)}{\partial \theta_2} \end{pmatrix} \times \mathbf{I}^{-1}(\theta) \begin{pmatrix} \frac{\partial \beta(k_1, \tau_1, \theta)}{\partial \theta_1} & \frac{\partial \beta(k_2, \tau_2, \theta)}{\partial \theta_1} \\ \frac{\partial \beta(k_1, \tau_1, \theta)}{\partial \theta_2} & \frac{\partial \beta(k_2, \tau_2, \theta)}{\partial \theta_2} \end{pmatrix},$$

whose (i, j) -elements, $i, j = 1, 2$, are denoted by $v_{11}(k_1, \tau_1, \tau_2, \theta)$, $v_{12}(k_1, k_2, \tau_1, \tau_2, \theta) = v_{21}(k_1, k_2, \tau_1, \tau_2, \theta)$ and $v_{22}(k_2, \tau_1, \tau_2, \theta)$. Actually, it holds that $v_{ii}(k_i, \tau_i, \tau_2, \theta) = v(k_i, \tau_i, \theta)$, $i = 1, 2$.

The asymptotic correlation coefficient (*AsCorr*) between $\hat{\beta}(k_1, \tau_1, \theta)$ and $\hat{\beta}(k_2, \tau_2, \theta)$ is defined by

$$\text{AsCorr}(\hat{\beta}(k_1, \tau_1, \theta), \hat{\beta}(k_2, \tau_2, \theta)) = \frac{v_{12}(k_1, k_2, \tau_1, \tau_2, \theta)}{\sqrt{v_{11}(k_1, \tau_1, \tau_2, \theta)v_{22}(k_2, \tau_1, \tau_2, \theta)}}, \quad (2)$$

which is regarded as a measure of linear dependence between the two slopes. The ML estimators, $\hat{\beta}(k_1, \tau_1, \theta)$ and $\hat{\beta}(k_2, \tau_2, \theta)$, are said to be mutually asymptotically independent or to have the mutually asymptotically independent property when the right side of (2) vanishes, because, under the asymptotic normality assumption, *AsCorr* = 0 if and only if $\hat{\beta}(k_1, \tau_1, \theta)$ and $\hat{\beta}(k_2, \tau_2, \theta)$ are asymptotically independent.

3.2 Asymptotically almost independent property: an illustration

We illustrate the case where the population distribution is assumed to be lognormal. The probability density function of the lognormal distribution is defined by

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp \left[-\frac{(\log x - \mu)^2}{2\sigma^2} \right], \quad x > 0$$

with $\theta = (\mu, \sigma)'$, $-\infty < \mu < \infty$, $\sigma > 0$. Short *et al.* (1993a) have shown that the mixed lognormal distribution is most representative between the lognormal, gamma and inverse-Gaussian distributions in comparing empirical results from the GATE I data to theoretical predictions of optimal thresholds for

the first and second moments of area rain rate. In the lognormal case, we have

$$\beta(k, \tau, \theta) = \frac{\exp(k\mu + k^2\sigma^2/2)}{1 - \Phi(u)},$$

where $u = (\log \tau - \mu)/\sigma$, $\tau > 0$, and Φ denotes the distribution function of the standard normal distribution. Since the Fisher information matrix is

$$\mathbf{I}(\theta) = \mathbf{I}(\mu, \sigma) = \frac{1}{\sigma^2} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix},$$

the elements of the asymptotic variance-covariance matrix are, for $u_i = (\log \tau_i - \mu)/\sigma$, $\tau_i > 0$ ($i = 1, 2$),

$$\begin{aligned} v_{ii}(k_i, \tau_i, \tau_2, \theta) &= v(k_i, \tau_i, \theta) \\ &= \frac{\sigma^2 \exp(2k_i\mu + k_i^2\sigma^2)}{\{1 - \Phi(u_i)\}^4} \left[\left\{ k_i(1 - \Phi(u_i)) - \frac{1}{\sigma}\phi(u_i) \right\}^2 \right. \\ &\quad \left. + \frac{1}{2} \left\{ k_i^2\sigma(1 - \Phi(u_i)) - \frac{u_i}{\sigma}\phi(u_i) \right\}^2 \right] \end{aligned}$$

for $i = 1, 2$, and

$$\begin{aligned} v_{12}(k_1, k_2, \tau_1, \tau_2, \theta) &= \\ &= \frac{\sigma^2 \exp\{(k_1 + k_2)\mu + (k_1^2 + k_2^2)\sigma^2/2\}}{\{1 - \Phi(u_1)\}^2 \{1 - \Phi(u_2)\}^2} \\ &\quad \times P(k_1, k_2, \tau_1, \tau_2, \theta), \end{aligned}$$

where

$$\begin{aligned} P(k_1, k_2, \tau_1, \tau_2, \theta) &= \\ &= \left\{ k_1(1 - \Phi(u_1)) - \frac{1}{\sigma}\phi(u_1) \right\} \\ &\quad \times \left\{ k_2(1 - \Phi(u_2)) - \frac{1}{\sigma}\phi(u_2) \right\} \\ &\quad + \frac{1}{2} \left\{ k_1^2\sigma(1 - \Phi(u_1)) - \frac{u_1}{\sigma}\phi(u_1) \right\} \\ &\quad \times \left\{ k_2^2\sigma(1 - \Phi(u_2)) - \frac{u_2}{\sigma}\phi(u_2) \right\}. \end{aligned}$$

There exists $\tau_{k,opt} (> 0)$, as in the case of $k = 1$ (Kedem and Pavlopoulos, 1991), which minimizes the normalized asymptotic variance $w(\tau) = v(k, \tau, \theta)/\{\beta(k, \tau, \theta)\}^2$ with respect to τ . From the equation $d w(\tau)/d \tau = 0$ and the fact that $1 - \Phi(u) \sim \phi(u)/u$, as $u \rightarrow \infty$, where ϕ denotes the density function of the standard normal distribution, we approximately have a solution $u = k\sigma$, i.e., $\tau = \exp(\mu + k\sigma^2)$, if solutions are sought in the region $u > 0$, i.e., $\tau > e^\mu$ (median). On the other hand, we have from the same approximation that

$$\begin{aligned} P(k_1, k_2, \tau_1, \tau_2, \theta) &\sim \frac{\phi(u_1)\phi(u_2)}{2\sigma^2 u_1 u_2} (u_1 - k_1\sigma)(u_2 - k_2\sigma) \\ &\quad \times \{2 + (u_1 + k_1\sigma)(u_2 + k_2\sigma)\}. \end{aligned}$$

Thus, approximate solutions of the equation $v_{12}(k_1, k_2, \tau_1, \tau_2, \theta) = 0$ are $u_1 = k_1\sigma$ and $u_2 = k_2\sigma$ if they are sought in the region $u_1, u_2 > 0$. Hence, at least approximately, the value of *AsCorr* between

(a)

(b)

Fig. 1. The absolute values of $AsCorr(\hat{\beta}(1, \tau_1, \theta), \hat{\beta}(2, \tau_2, \theta))$ (a) on the (u_1, u_2) plane and (b) on the (τ_1, τ_2) plane for the GATE I parameter. Darker pixels mean that the values are close to zero.

$\hat{\beta}(k_1, \tau_1, \theta)$ and $\hat{\beta}(k_2, \tau_2, \theta)$ is zero if τ_1 or τ_2 is chosen to be optimal in the sense of minimizing normalized asymptotic variances.

To study exact solutions, we fix the parameters of the distribution, which are chosen by the method of moments to match conditional first and second rain rate moments, 4 mm/hr and 65 (mm/hr)², for the GATE I data to the theoretical ones; $\mu = 0.685$ and $\sigma = 1.184$. From the practical point of view, we confine ourselves to the case $k_1 = 1$ (first moment) and $k_2 = 2$ (second moment). We shall numerically show that the value of $AsCorr$ is close to zero for the case of the GATE I parameter if $\tau_1 = \tau_2 = \tau_{1,opt}$ or $\tau_1 = \tau_2 = \tau_{2,opt}$. This implies that $\hat{\beta}(1, \tau, \theta)$ and $\hat{\beta}(2, \tau, \theta)$ are mutually asymptotically almost independent (AAI) if τ is one of the optimal thresholds chosen by minimizing the normalized asymptotic variances. Figure 1a pictures the absolute values of $AsCorr(\hat{\beta}(1, \tau_1, \theta), \hat{\beta}(2, \tau_2, \theta))$ on the (u_1, u_2) plane and Fig. 1b on the (τ_1, τ_2) plane, from which we can see that there are many combinations of (τ_1, τ_2) satisfying the fact that the value of $AsCorr$ is close to zero. Figure 2 shows the behavior of $AsCorr$ between $\hat{\beta}(1, \tau_1, \theta)$ and $\hat{\beta}(2, \tau_2, \theta)$ along the line $\tau_1 = \tau_2 (= \tau)$, while Fig. 7 of Short *et al.* (1993a) shows the behavior of the values of the normalized asymptotic variances $v(i, \tau_i, \theta) / \{\beta(i, \tau_i, \theta)\}^2$, $i = 1, 2$. It is noted that the optimal thresholds are $\tau_{1,opt} = 5.5$ and $\tau_{2,opt} = 26$, and that the values of $AsCorr$ are zero at 5.7 and 26; the AAI property holds.

We omit the probability density function, $\beta(i, \tau_i, \theta)$, $i = 1, 2$, and the Fisher information matrix for each of the gamma and inverse-Gaussian cases because they can be seen in the papers by Short *et al.* (1993a) and Shimizu *et al.* (1993).

The probability density function of the Weibull distribution is defined by

$$f(x) = \frac{c}{\alpha} \left(\frac{x}{\alpha}\right)^{c-1} \exp\left[-\left(\frac{x}{\alpha}\right)^c\right], \quad x > 0$$

with $\theta = (\alpha, c)'$, $\alpha, c > 0$. This case gives

$$\begin{aligned} \beta(1, \tau_1, \theta) &= \alpha \Gamma \left(1 + \frac{1}{c}\right) \exp\left[\left(\frac{\tau_1}{\alpha}\right)^c\right], \\ \beta(2, \tau_2, \theta) &= \alpha^2 \Gamma \left(1 + \frac{2}{c}\right) \exp\left[\left(\frac{\tau_2}{\alpha}\right)^c\right], \end{aligned}$$

where Γ stands for the gamma function. When X represents a random variable following $W(\alpha, c)$, a power transformation bX^r , where b and r are constants, follows $W(\alpha^r b, c/r)$. This means that the Weibull distribution is closed with respect to the power transformation, as in the lognormal distribution. The Fisher information matrix of the Weibull distribution is

$$\mathbf{I}(\theta) = \mathbf{I}(\alpha, c) = \begin{pmatrix} c^2/\alpha^2 & -\Gamma'(2)/\alpha \\ -\Gamma'(2)/\alpha & 1 + \Gamma''(2)/c^2 \end{pmatrix}.$$

Lognormal distributed rain rates
 $\mu = 0.685, \sigma = 1.184$

Fig. 2. The $AsCorr$ between $\hat{\beta}(1, \tau, \theta)$ and $\hat{\beta}(2, \tau, \theta)$. The value of $AsCorr$ is zero at 5.7 and 26, while the optimal values that minimize the normalized asymptotic variances $v(i, \tau_i, \theta) / \{\beta(i, \tau_i, \theta)\}^2$, $i = 1, 2$, are $\tau_{1,opt} = 5.5$ and $\tau_{2,opt} = 26$.

The results of gamma, inverse-Gaussian and Weibull cases give the same tendency as in the lognormal case. For the GATE I data, the parameters, the optimal thresholds and the points at which the values of $AsCorr(\hat{\beta}(1, \tau_1, \theta), \hat{\beta}(2, \tau_2, \theta))$ are zeroes are summarized in Table 1. The definition of the parameters for the gamma and inverse-Gaussian distributions is referred to the papers by Short *et al.* (1993a) and Shimizu *et al.* (1993).

Shimizu *et al.* (1993) have recommended, from the view point of stability, the use of the optimal thresholds for estimating the first and second moments of area rain rate to estimate the variance of rain rate by the double threshold method. The analysis of the present paper provides another reason why such thresholds should be used.

4. Simultaneous choice of optimal thresholds

This section considers the possibility of choosing thresholds simultaneously for the estimation of the first and second moments as well as for the variance of area rain rate.

Empirical evidence shows the fact that area-averages of rain rate tend to follow a lognormal distribution. This is well supported by the GATE I data. We suppose that rain rate is observed over a given area throughout a specific time period. The instantaneous rain rate in each cell is denoted by $x_1, \dots, x_N$. Here N represents the number of the samples in the investigated area. This may be given by a particular conversion of physical quantities such as radar reflectivity measured by a radar launched on a satellite from space or by a ground-based system. The sample includes positive numbers, condi-

Table 1. The parameters, the optimal thresholds $(\tau_{1,opt}, \tau_{2,opt})$, and the points at which the values of $AsCorr(\hat{\beta}(1, \tau_1, \theta), \hat{\beta}(2, \tau_2, \theta))$ are zeroes for the GATE I data.

Distributions	Parameters	$(\tau_{1,opt}, \tau_{2,opt})$	Values of $AsCorr = 0$
Lognormal	$\mu = 0.685, \sigma = 1.184$	(5.5, 26)	(5.7, 26)
Gamma	$\alpha = 0.33, 1/\lambda = 12.25$	(7.8, 18.8)	(7.8, 19)
Inverse-Gaussian	$\nu = 4, \eta = 1.306$	(7.2, 24)	(7.2, 24)
Weibull	$\alpha = 2.671, c = 0.602$	(7.3, 21)	(7.5, 21)

tional on rain, and zeroes, no rain. Let n_0 be the number of zero values and $y_1, \dots, y_n$ remaining n positive values, where $n = N - n_0$. The number n varies, from time to time.

Assume that the positive sample, $y_1, \dots, y_n$, of size n is from a two-parameter lognormal distribution with parameter $\theta = (\mu, \sigma)'$, which is in general unknown. For fixed $\tau (> 0)$, we calculate the following quantities:

$$\hat{\beta}(1, \tau, \theta) = \frac{\exp(\hat{\mu} + s^2/2)}{1 - \Phi(u)},$$

$$\hat{\beta}(2, \tau, \theta) = \frac{\exp(2\hat{\mu} + 2s^2)}{1 - \Phi(u)}, \quad u = \frac{\log \tau - \hat{\mu}}{s},$$

where Φ stands for the distribution function of the standard normal distribution and

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n \log y_i,$$

$$s = \sqrt{\frac{1}{n} \sum_{i=1}^n (\log y_i - \hat{\mu})^2}.$$

The quantities $\hat{\mu}$ and s are the ML estimates of μ and σ from a complete sample of size n under the lognormal assumption. If there are any serious problems in converting particular physical quantities into light and heavy rain rates, one may estimate μ and σ from a truncated or censored sample, though the estimation method becomes complex. Alternatively, one can use a multiple thresholding technique (Meneghini and Jones, 1993) to estimate parameters.

If the horizontal line denotes $w = \hat{\beta}(1, \tau, \theta)$ and the vertical line $z = \hat{\beta}(2, \tau, \theta)$, each snapshot makes a point on the plane. When m denotes the number of snapshots, we can make the scatter plot of $(w_j, z_j), j = 1, \dots, m$. The correlation coefficient

$$r = \frac{\sum_{j=1}^m (w_j - \bar{w})(z_j - \bar{z})}{\sqrt{\sum_{j=1}^m (w_j - \bar{w})^2 \sum_{j=1}^m (z_j - \bar{z})^2}}$$

represents a measure of linear dependence between the variables w and z . Here $\bar{w}$ and $\bar{z}$ are respectively the arithmetic means of w and z .

From the theory given in Section 3, there should be two τ s, τ_1 and τ_2 ($\tau_1 < \tau_2$), simultaneously determined, which make r zero. We can choose τ_1 as a

threshold for the estimation of the first moment and τ_2 the second moment. These τ s should be close to the optimal thresholds individually determined for estimating the first and second moments of area rain rate. However, it should be noted that the estimates of the slopes and thus the thresholds chosen by the proposed method depend on the assumed population distribution. Comparison between lognormal, gamma, inverse-Gaussian and Weibull distributions by using real data would be an expected study in the future. The estimation method from a complete, truncated or censored sample is available (Cohen and Whitten, 1988).

5. Summary and conclusions

Empirical and theoretical aspects of the threshold method have been studied by several authors since Chiu's (1988) systematic empirical study. The present paper mainly focuses on the theoretical aspect of the method, and is closely connected with the papers by Short *et al.* (1993a) and Shimizu *et al.* (1993) among others.

Short *et al.* (1993a) have presented the ML approach for selecting optimal thresholds for the estimation of area rain-rate moments, and Shimizu *et al.* (1993) for the estimation of the variance. It is natural to study properties of slopes as well as optimal thresholds. In fact, Kedem *et al.* (1990a) have given some results of slope surfaces obtained from lognormal, gamma and inverse-Gaussian distributions. The present paper treated the two-parameter Weibull distribution, as well as these three distributions, and showed that the asymptotic correlation coefficient between the ML estimators of the slopes for the first and second moments were close to zero for the GATE I parameter if the thresholds were chosen to be optimal in the sense of minimizing normalized asymptotic variances. The result is helpful for the intuitive understanding of optimal thresholds, and provides a reason why optimal thresholds for estimating the first and second moments of area rain rate should be used for the estimation of the variance by the double threshold method.

A method of the simultaneous choice of thresholds for the estimation of area rain-rate moments was proposed in the last section, which would be the base of further studies. The thresholds for the estimation

of the first and second moments of area rain rate are chosen in such a way that the sample correlation coefficient between the ML estimates of the slopes for the first and second moments is zero.

References

- Bishop, Y.M.M., S.E. Fienberg and P.W. Holland, 1980: *Discrete Multivariate Analysis: Theory and Practice*. The MIT Press, 557 pp.
- Chiu, L.S., 1988: Rain estimation from satellites: Areal rainfall-rain area relations. *3rd Conference on Satellite Meteorology and Oceanography*, Anaheim, CA, Amer. Meteor. Soc., 363–368.
- Cohen, A.C. and B.J. Whitten, 1988: *Parameter Estimation in Reliability and Life Span Models*. Marcel Dekker, 394 pp.
- Doneaud, A.A., S.I. Niscov, D.L. Priegnitz and P.L. Smith, 1984: The area-time integral as an indicator for convective rain volumes. *J. Clim. Appl. Meteor.*, **23**, 555–561.
- Kedem, B. and H. Pavlopoulos, 1991: On the threshold method for rainfall estimation: Choosing the optimal threshold level. *J. Amer. Statist. Assoc.*, **86**, 626–633.
- Kedem, B., L.S. Chiu and Z. Karni, 1990a: An analysis of the threshold method for measuring area-average rainfall. *J. Appl. Meteor.*, **29**, 3–20.
- Kedem, B., L.S. Chiu and G.R. North, 1990b: Estimation of mean rain rate: Application to satellite observations. *J. Geophys. Res.*, **95**, 1965–1972.
- Meneghini, R. and J.A. Jones, 1993: An approach to estimate the areal rain-rate distribution from spaceborne radar by the use of multiple thresholds. *J. Appl. Meteor.*, **32**, 386–398.
- Pavlopoulos, H. and B. Kedem, 1992: Stochastic modeling of rain rate processes: A diffusion model. *Commun. Statist.-Stochastic Models*, **8**, 397–420.
- Rosenfeld, D., D. Atlas and D.A. Short, 1990: The estimation of convective rainfall by area integrals, 2, The height-area rainfall threshold (HART) method. *J. Geophys. Res.*, **95**, 2161–2176.
- Shimizu, K., D.A. Short and B. Kedem, 1993: Single- and double-threshold methods for estimating the variance of area rain rate. *J. Meteor. Soc. Japan*, **71**, 673–683.
- Short, D.A., K. Shimizu and B. Kedem 1993a: Optimal thresholds for the estimation of area rain-rate moments by the threshold method. *J. Appl. Meteor.*, **32**, 182–192.
- Short, D.A., D.B. Wolff, D. Rosenfeld and D. Atlas, 1993b: A study of the threshold method utilizing raingage data. *J. Appl. Meteor.*, **32**, 1379–1387.

しきい値法による領域降雨強度モーメントの推定

清水邦夫

(東京理科大学理学部応用数学教室)

栢野浩一

(トヨタ自動車技術管理部技術システム室)

しきい値法により領域降雨強度モーメントを推定するさいの理論的最適しきい値のもつ性質について考察した。理論分布は、原点に離散確率をもち降雨の条件のもとでは連続分布であるような混合型とした。傾きの最尤推定量の規準化漸近分散を最小にするという意味で最適なしきい値が選ばれるならば、GATE I パラメータに対し、一次および二次モーメントの傾きの最尤推定量は漸近的にほぼ独立となることを数値的に示した。この事実は理論的最適しきい値の直観的理解に役立つとともに、二重しきい値法により領域降雨強度分散を推定するさいに一次および二次モーメントを推定するための最適しきい値を用いるべきであることの一つの理由を与えている。また、一次および二次モーメント推定のためのしきい値の同時選択法を提案した。