P3.3 ESTIMATION OF THE DROP SIZE DISTRIBUTION PARAMETER IN THE RAIN PROFILING ALGORITHM FOR THE TRMM PRECIPITATION RADAR

Toshio Iguchi*
NICT, Tokyo, Japan

1. INTRODUCTION

The standard rain profiling algorithm for the TRMM Precipitation Radar, which is known as 2A25, uses the Hitschfeld-Bordan (H-B) method for the rain attenuation correction (Hitschfeld and Bordan, 1954). Unless the path integrated attenuation (PIA) is very small, the attenuation estimate from the H-B method is compared with the PIA estimate given by the surface reference technique (SRT). By taking into consideration of measurement errors, the specific attenuation assumed in the H-B method is adjusted so that the PIA from the H-B method with the modified specific attenuation is consistent with the PIA estimate from the SRT. Adjusting the specific attenuation is equivalent of adjusting the initially assumed drop size distribution (DSD). Thus, when the PIA is significant, one parameter in the DSD model which is assumed to be constant at all ranges along the radar beam can be estimated. This information is used to adjust the Z-R relationship with which the attenuation corrected radar reflectivity factor Z_e is converted into rainfall rate R in the current algorithm.

In practice, however, the deviation of the model DSD function from the real DSD is not the only factor that causes the discrepancy between the attenuation estimates from the H-B method and those from the SRT. There are several other factors that also create the discrepancies in the attenuation estimates. In the current algorithm, all such possible factors are ignored and the discrepancy is totally attributed to the deviation of the model DSD. In this paper, effects of ignoring other possible factors are examined in terms of resultant biases in the estimated DSD parameter. Such factors include biases in the assumed vertical profile of precipitating particles and their properties, biases in the PIA estimates by SRT, non-uniform beam filling effect, and biases caused by a particular choice of estimation method.

2. STRUCTURE OF THE ALGORITHM

In this section, the structure of the rain profiling algorithm 2A25 is described (Iguchi et al. 2000, Iguchi 2007). In order to show the essential issues conspicuously, the algorithm described here ignores almost all minor details which include, for example, the attenuation correction for cloud liquid water and water vapor, and handling of data in surface-cluttered range bins. Similarly, to have a clearer insight into the structure, range dependent variables such as the measured radar reflectivity factor Z_m as a function of range r are treated as a continuous function of range rather than a function defined only at discrete range bins at which the actual data are sampled.

In 2A25, measured radar reflectivity factor $Z_m(r)$ is converted into rainfall rate $R(r)$ in two steps. The first step is to correct for the attenuation to obtain the effective radar reflectivity factor $Z_e(r)$ and the second step is to convert $Z_e(r)$ into $R(r)$.

The first step of attenuation correction is carried out by dividing $Z_m(r)$ by the estimate of attenuation factor $A(r)$.

$$Z_e(r) = \frac{Z_m(r)}{A(r)} \tag{1}$$

The problem is to find the best estimate of $A(r)$. We denote the true attenuation to surface r_s by A_s ($A_s \stackrel{\text{def}}{=} A(r_s)$) and define y_t as the true attenuation to surface expressed in dB:

$$y_t \stackrel{\text{def}}{=} -10\log_{10}(A_s). \tag{2}$$

If we know the relationship between the specific attenuation k and Z_e, we can estimate the attenuation to range r from the profile of Z_m by the Hitschfeld-Bordan method. The Hitschfeld-Bordan estimate of $A(r)$ is given by

$$A_{\text{HB}}(r) = [1 - \zeta_m(r)]^{1/\beta} \tag{3}$$

where

$$\zeta_m(r) \stackrel{\text{def}}{=} 2q\int_0^r \alpha_0 Z_m^\beta(s)\,ds \tag{4}$$

and $q = 0.1\beta\ln 10$. α_0 and β are coefficients in the assumed k–Z_e relationship $k = \alpha_0 Z_e^\beta$.

In the α-adjustment method, ζ_m is adjusted by multiplying α_0 by ϵ in such a way that the attenuation to surface from the H-B method agrees with the true attenuation y_t.

$$y_t = -\frac{10}{\beta}\log_{10}[1 - \epsilon\zeta_m(r_s)] \tag{5}$$

* *Corresponding author address:* Toshio Iguchi, National Institute of Information and Communications Technology, Tokyo 184-8795, Japan; e-mail: iguchi@nict.go.jp

or

$$\epsilon = \frac{1 - 10^{-\beta y_t/10}}{\zeta_m(r_s)} \tag{6}$$

Once ϵ is determined, we can calculate the attenuation factor $A(r)$ at any range r by the α-adjustment method:

$$A_\alpha(r) = [1 - \epsilon\zeta_m(r)]^{1/\beta} \tag{7}$$

If we define x_t by

$$x_t \stackrel{\text{def}}{=} \ln[1 - \exp(-q y_t)], \tag{8}$$

and x_m by

$$x_m \stackrel{\text{def}}{=} \ln(\zeta_m(r_s)), \tag{9}$$

then (5) is equivalent of

$$x_m = x_t - \ln(\epsilon). \tag{10}$$

If we rewrite $-\ln(\epsilon)$ as e_x, (10) becomes

$$x_m = x_t + e_x \tag{11}$$

This equation shows that the observable x_m is the sum of its true value x_t and the error term e_x. Since x_t is related to y_t as in (8), x_m is an observable that enables us to estimate the attenuation.

The surface reference technique (SRT) also gives an estimate of the attenuation to the surface. We denote this estimate by y_m. If we denote the error associated with SRT by e_y, then y_m is related to y_t as follows:

$$y_m = y_t + e_y \tag{12}$$

The combination of (11) and (12) together with (8) gives the coupled set of equations that define the structure of the problem of attenuation correction. In other words, we have two observables, x_m and y_m, for a given true value of attenuation y_t. (Note that x_t is a function of y_t.) Both x_m and y_m contain some unknown amount of errors of which magnitudes are known only statistically. The problem is to find the best estimate of y_t for a given set of x_m and y_m.

Since the radar beam loses energy as it propagates, the attenuation (y_t) is always positive. This is already implicitly assumed when we define x_t by (8). Equation (8) also indicates that x_t must be negative, which guarantees that the H-B solution does not blow up.

$$x_t < 0, \tag{13}$$

$$y_t > 0. \tag{14}$$

These conditions limit the domain of solution.

Once the estimate y_t is obtained, we can calculate the value of e_x from x_m and x_t in each case. As will be described in the next section, the error e_x consists of several factors. But in both V5 and V6 of 2A25 e_x is assumed to be equal to $-\ln(\epsilon)$ and the deviation of x_m from x_t is totally attributed to the deviation of true α from the initially assumed value of α_0 in the k–Z_e relationship. Since the adjustment of the k–Z_e relationship corresponds to the modification of the DSD, the Z_e–R relationship with which the attenuation corrected Z_e is converted into R is also modified in conformity with the modified k–Z_e relationship.

If the deviation of x_m from x_t is caused by a different error source rather than the DSD difference, the modified Z_e–R relationship does not reflect the true DSD. Similarly, if x_t (and hence y_t) is estimated with some bias, we again end up with a biased Z_e–R relationship and hence a biased estimate of R. In other words, the DSD parameter can be estimated correctly only when the assumptions on the causes of errors e_x and e_y are valid. It is the major objective of this paper to examine these assumptions and to evaluate the effect of errors if there is any deviation from the assumptions.

To carry out the above program, we need to know the characteristics of errors e_x and e_y. However, knowing the characteristics of e_x and e_y is not enough to determine the best estimate of y_t or x_t in a unique way. We need to define the meaning of the "best" estimate and select the appropriate estimator. For example, depending on whether we regard y_t as an unknown but fixed parameter or as a random variable, the solution changes. In the former case, the classical maximum likelihood estimate will give the solution, and in the latter case, Bayesian method should give a statistically better solution. Before we go into this philosophical issue, we look at the error characteristics of e_x and e_y first. We will come back to this issue in section 4.

3. ERROR SOURCES

Since x_m is the logarithm of $\zeta_m(r_s)$, and since ζ_m is defined by (4), the error consists of two factors.

The first factor is the deviation of the k–Z_e model profile assumed in 2A25 from the true one. The assumption that this relationship can be represented by a power law $k = \alpha Z_e^\beta$ is already an approximation. In our model, β is assumed constant and independent of range. The dependence of α on the altitude is also assumed for each type of rain. The value of ζ_m calculated with these assumptions may create a deviation from the true value in individual cases even if they cause no bias statistically.

The second factor is the error in the measurement of Z_m. It includes both the fading noise of 0.7 dB due and the calibration error. We assume that the correction for attenuation due to water vapor (WV) and cloud liquid water (CLW) has already been included in Z_m. Since the vertical profile models of WV and CLW do not perfectly match with the true profiles in individual cases, the differences may contribute to the total error in Z_m.

Since y_m is obtained by subtracting the measured apparent surface cross section from the reference cross section, there are two kinds of errors in y_m. The first one is the error in the measurement of the apparent surface cross section itself. The second one is the error in the reference. The effect of attenuation due to WV and CLW is also considered and corrected for in V6, but this correction is also a source of error in y_m. This error can be regarded as the second kind of error that affects the reference cross section. The error of the second kind mainly consists of the fluctuation of sampled data taken for reference. In the spatial reference method, it consists of the fluctuation of surface cross sections measured at 8 footprints in non-raining region adjacent to the rain area. If the temporal reference is used, it is the standard deviation of the cross sections measured at the same incidence angle in the same x-degree grid box that includes the location in the previous month. y_m may also contain some bias error because of the difference in surface characteristics between the surface in question and the reference surface (Seto and Iguchi, 2007). However, both V5 and V6 of 2A25 assume that e_y consists of only random errors and follows a normal distribution.

4. BAYESIAN AND ML ESTIMATES

From the Bayesian point of view, the parameter y_t is regarded as a random variable, and the problem is to find the a posteriori probability density function (pdf) of y_t (or x_t) for given x_m and y_m. By applying the Bayes theorem, we obtain

$$p(y_t|x_m, y_m) \propto p(x_m|x_t(y_t))p(y_m|y_t)p(y_t) \qquad (15)$$

$p(x_m|x_t(y_t))$ is equal to the pdf of e_x shifted by x_t, and $p(y_m|y_t)$ is equal to the pdf of e_y shifted by y_t.

Once the a posteriori pdf is obtained, we still have some freedom to choose the estimator, $\hat{y}_t$, of y_t. Probably the most commonly used estimator in Bayesian statistics is the mean of $\hat{y}_t$, i.e., the expected value of $\hat{y}_t$ with respect to $p(y_t|x_m, y_m)$. We will denote the expected value of any function $f(y_t)$ of y_t by $\langle f \rangle$. Then,

$$\begin{aligned} \langle y_t \rangle &\stackrel{\text{def}}{=} E[y_t|x_m, y_m] \\ &= \int y_t p(y_t|x_m, y_m)\, dy_t \end{aligned} \qquad (16)$$

Since the expected value $\langle y_t \rangle$ is the integral of y_t weighted by $p(y_t|x_m, y_m)$, it depends on the entire distribution of $p(y_t|x_m, y_m)$.

Another popular method is the maximum a posteriori (MAP) estimation in which the mode of the posterior pdf is chosen as the estimator.

$$\hat{y}_{t,\mathrm{MAP}} \stackrel{\text{def}}{=} \arg\max_{y_t} p(y_t|x_m, y_m) \qquad (17)$$

Note that MAP estimation is not generally seen as a Bayesian method because Bayesian methods are characterized by the use of distributions to summarize data and draw inferences.

The likelihood function of y_t for given x_m and y_m is

$$l(y_t|x_m, y_m) = p(x_m, y_m|y_t) = p(x_m|x_t(y_t))p(y_m|y_t) \qquad (18)$$

and the maximum likelihood (ML) estimator $\hat{y}_{t,\mathrm{ML}}$ is defined as

$$\hat{y}_{t,\mathrm{ML}} = \arg\max_{y_t} l(y_t|x_m, y_m) \qquad (19)$$

The comparison between (15) and (18) shows that the ML estimator is a special case of the MAP estimator with a uniform pdf of y_t from a Bayesian point of view.

This fact implies that if the a priori pdf $p(y_t)$ is nearly uniform in the region in which $l(y_t|x_m, y_m)$ takes dominant values, the ML estimate and the MAP estimate are nearly the same. Even in the same condition, however, the ML estimate and the mean may differ substantially if the a posteriori pdf is skewed. This is actually the case in the current problem. Because of the non-linear relationship between x_t and y_t as defined in (8), $p(x_m|x_t(y_t))$ in (15) is skewed as a function of y_t. In the region where $p(x_m|x_t(y_t))$ is the determining factor of (15) or (18), i.e., when y_t is small, the a posteriori pdf $p(y_t|x_m, y_m)$ and the likelihood function $l(y_t|x_m, y_m)$ are very skewed, and the mean $\langle y_t \rangle$ differs from the mode $y_{t,\mathrm{MAP}}$ and the ML estimate $y_{t,\mathrm{ML}}$.

An important question is which estimator is a better estimator, $\langle y_t \rangle$ or $y_{t,\mathrm{ML}}$? V5 of 2A25 uses $y_{t,\mathrm{ML}}$, and V6 adopts $\langle y_t \rangle$. Although the choice between them does not create a large difference in the total rain amount or the attenuation correction when the attenuation is large, the values of ϵ estimated in these two different ways differ when the attenuation is small. In fact, the deterministically calculated value of $\epsilon(y_{t,\mathrm{ML}})$ is used in V5, whereas the expected value $\langle \epsilon \rangle$ is calculated in V6. Specifically, in the former case,

$$\epsilon(y_{t,\mathrm{ML}}) = \exp[x_t(y_{t,\mathrm{ML}}) - x_m], \qquad (20)$$

whereas in the latter case

$$\langle \epsilon \rangle = \int \epsilon(y_t) p(y_t|x_m, y_m)\, dy_t. \qquad (21)$$

Note that (21) is different from $\epsilon(\langle y_t \rangle)$.

In the region where $p(x_m|x_t(y_t))$ is the determining factor of (15) or (18), $x_{t,\mathrm{ML}} \stackrel{\mathrm{def}}{=} x_t(y_{t,\mathrm{ML}})$ is nearly equal to x_m, and $\epsilon(y_{t,\mathrm{ML}})$ remains close to the initially assumed value of unity if the pdf of error model $p(x_m|x_t(y_t))$ is properly chosen in conformity with the assumption of the ML estimation. However, $\langle \epsilon \rangle$ may differ from unity even in the same circumstances. This happens because the a posteriori distribution of ϵ is skewed. The skewness comes not only from the nonlinear dependence of ϵ on x_t, but from the condition that x_t must be negative, i.e., $\zeta_t \stackrel{\mathrm{def}}{=} \exp(x_t)$ must be less than 1 in order that the H-B solution exists. This can be easily seen if we transform the independent variable of integration from y_t to x_t and use the approximation $p(y_t|x_m, y_m) \approx p(y_t|x_m) \propto p(x_m|x_t)(dx_t/dy_t)$.

$$\begin{aligned} \langle \epsilon \rangle &= \int \epsilon(x_t, x_m) p(x_t|x_m, y_m)\, dx_t. \\ &\approx \int_{-\infty}^{0} \exp(x_t - x_m) p(x_m|x_t)\, dx_t. \end{aligned} \tag{22}$$

In V6, $p(\epsilon) \stackrel{\mathrm{def}}{=} p(\exp(x_t - x_m))$ without any constraint is assumed to follow a lognormal distribution with mean 1.

$$p(\epsilon) = \frac{1}{\sqrt{2\pi}\sigma\epsilon} \exp\left(-\frac{(\ln(\epsilon) - m)^2}{2\sigma^2}\right) \tag{23}$$

with $\exp(m + \sigma^2/2) = 1$, or if the variance of ϵ is σ_ϵ^2, then $m = -(1/2)\ln(1 + \sigma_\epsilon^2)$ and $\sigma^2 = \ln(1 + \sigma_\epsilon^2)$. Since

$$\begin{aligned} p_{x_m}(x_m - x_t) &= p_{\ln(\epsilon)}(\ln(\epsilon)) \\ &= \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(\ln(\epsilon) - m)^2}{2\sigma^2}\right) \\ &= \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x_m - (x_t - m))^2}{2\sigma^2}\right) \\ &= p(x_m|x_t) \end{aligned} \tag{24}$$

$p(x_m|x_t)$ follows the normal distribution with mean $x_t - m$ and variance σ^2. Therefore, (22) becomes

$$\begin{aligned} \langle \epsilon \rangle &\approx \int_{-\infty}^{0} \exp(x_t - x_m) p(x_m|x_t)\, dx_t. \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{0} \exp\left(-\frac{(x_t - x_m - m)^2}{2\sigma^2} + x_t - x_m\right) dx_t \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{0} \exp\left(-\frac{(x_t - x_m - m - \sigma^2)^2}{2\sigma^2}\right) dx_t \\ &\quad \times \exp(\frac{\sigma^2}{2} + m) \\ &= 1 - \frac{1}{\sqrt{\pi}} \int_{-x_m - m - \sigma^2}^{\infty} \exp\left(-t^2\right)\, dt \\ &= 1 - \frac{1}{2}\mathrm{erfc}(-x_m - m - \sigma^2) \end{aligned} \tag{25}$$

where $\mathrm{erfc}(x)$ is the complementary error function. Note that $\exp(\frac{\sigma^2}{2} + m) = 1$ from the assumption. Since as $x \to \infty$, $\mathrm{erfc}(x) \to 0$, and as $x \to -\infty$, $\mathrm{erfc}(x) \to 2$, $\langle \epsilon \rangle$ is nearly equal to 1 when $\zeta_m \ll 1$ or equivalently $x_m \ll 0$. However, as ζ_m increases, x_m increases toward 0 and when x_m becomes comparable to $-m - \sigma^2 - 1 = -(1/2)\ln(1 + \sigma_\epsilon^2) - 1$, $\langle \epsilon \rangle$ starts decreasing and deviates from 1 before the constraints by the SRT or the a priori pdf of y_t affect the a posteriori pdf. ϵ smaller than 1 will modify the original R–Z_e relationship so that a smaller rainfall rate will result than that without modification for the same Z_e. Note that since in the cases we discuss in this section, the attenuation correction is small and Z_e is nearly equal to Z_m regardless of the method used.

Under the same conditions with the same pdf, the ML estimate of ϵ becomes

$$\begin{aligned} \epsilon_{\mathrm{ML}} &= \exp(m - \sigma^2) \\ &= \exp(-(3/2)\ln(1 + \sigma_\epsilon^2)) \\ &= \frac{1}{(1 + \sigma_\epsilon^2)^{3/2}} \end{aligned} \tag{26}$$

and the estimate is independent of ζ_m as long as $\zeta_m < 1$ or $x_m < 0$. The deviation of ϵ_{ML} from unity is due to the wrong choice of the parameters in the pdf. In this case, m should be chosen to be σ^2 so that $\epsilon_{\mathrm{ML}} = 1$

5. DISCUSSION

The essence of the idea of attenuation correction in 2A25 is to use the attenuation estimate given by the surface reference techinique to constrain the final attenuation estimate. The DSD parameter is adjusted in such a way that the modified k–Z_e relationship derived from the adjusted DSD model will give a total attenuation that agrees with the final attenuation estimate.

According to this idea, if the attenuation is very small and if the constraint from the surface reference is effectively negligible, we naturally think that the DSD parameter should not be modified. In the previous section, we have seen that this is not the case in the Bayesian estimate. The expected value of ϵ decreases as ζ_m increases from 0 because the upper limit of the integral changes with ζ_m. This phenomenon may be quite natural from Bayesian point of view. Nevertheless, if we imagine a situation in which a uniform light rain is measured from space at several incidence angles. Depending on the incidence angle, the maximum depth we can measure the rain echo changes due to the surface clutter. Since ζ_m defined by (4) monotonically increases with range r, its value at the bottom changes depending on how deep we can measure. The Bayesian estimate indicate our estimate of DSD parameter changes with the depth of measurement. This conclusion appears to be inconsistent with our assumption of the uniform rain. We want to choose an estimator that gives the same estimate if the same rain system is measured. The ML estimate satisfies this condition. Does this mean that the Bayesian estimate is inappropriate? The answer is "No". The discrepancy comes from the inconsistency between the assumption in the formulation of the Bayesian estimator and the assumption of the uniform rain. In fact, that the uniform rain assumption is not used in the Bayesian formulation in this example. From a Bayesian point of view, this important prior knowledge is not utilized in the formulation, and hence seemingly inconsistent estimates. If such knowledge is incorporated, the Bayesian method should have also given intuitively appealing estimates.

Therefore, it seems that the issue is not the question whether the Bayesian estimator is superior to the ML estimator. The issue seems to be in a more practical point. As we have seen in the previous section, the Bayesian estimate depends on the entire distribution of the a posteriori pdf whereas the ML estimate depends only on the distribution in the vicinity of the distribution maximum. This fact implies that to have a good Bayesian estimate, we need to model the whole distribution of pdf correctly. This seems to me a very difficult task. The ML estimation seems to be an easier way to go in a practical sense.

The case discussed in detail up to this point is probably not very important in terms of the total rain estimates. In fact, the effect of the truncation in the calculation of the expected values in the Bayesian formulation is minimal. The value of ϵ deviates from unity by a few percent at most due to this effect.

When the attenuation increases, the determining factor in (15) is taken by $p(y_m|y_t)$, and $p(x_m|x_t)$ gives almost no effect in the a posteriori pdf. In such cases, the issue discussed up to this point is immaterial. If the model assumptions are appropriate, our estimates should correctly represent the DSD characteristics. However, if the model assumptions are not totally valid, the estimated DSD parameters contain bias errors because any deviation of x_m from x_t is attributed to the difference in the DSD parameters, although x_t itself or parameters in x_m may contain a bias error if the model is not appropriate. Major error sources of this kind have been already mentioned in section 3 within the formulation of the algorithm structure described in section 2.

There is one important error source not mentioned there. That is the effect of non-uniform beam filling (NUBF). The NUBF affects both x_m and y_m. If the rain is not uniform within the field of view, ζ_m can exceed unity even when the DSD and profile models are correct. The attenuation estimate from the difference of apparent surface cross sections is smaller than the corresponding attenuation that would result if the same amount of rain is distributed uniformly in horizontal direction at each height within the filed of view. Underestimation of the attenuation results in the underestimation of Z_e, and possible overestimation of ζ_m tends to give a small value of ϵ. The overall effect of ignoring the NUBF effect in the retrieval is underestimation of rainfall rates. The magnitude of this effect is not negligibly small, especially when the attenuation is large. However, we are not going to discuss this issue in this paper any further.

6. SUMMARY

The basic structure of the TRMM/PR rain retrieval algorithm is reviewed. The essential point is the method of attenuation correction and the way it is linked to the DSD parameter estimation. With the errors in measurements taken into account, the problem is formulated as a statistical inference problem. Two possible estimators are explained. One is a classical ML estimator used in V5 of 2A25, and the other a Bayesian estimator adopted in V6 of 2A25.

Aside from the often-argued philosophical issues between classical and Bayesian statistics, an interesting difference in estimates between these statistics is shown and an advantage of the classical ML estimation is highlighted.

REFERENCES

Hitschfeld, W. and J. Bordan, 1954: Errors inherent in the radar measurement of rainfall at attenuating wavelengths. *J. Meteor.,* **11**, 58–67.

Iguchi, T., T. Kozu, R. Meneghini, J. Awaka, and K. Okamoto, 2000: Rain profiling algorithm for the TRMM Precipitation Radar, *J. Appl. Meteor.,* **39**, 2038–2052.

Iguchi, T., 2007: Spaceborne radar algorithm, in Measuring precipitation from space – EURAINSAT and the future, V. Levizzani, P. Bauer, and F. J. Turk, Eds., Springer, 705–712.