Magnitude of Error Factors in Estimates of Snow-Particle Masses from Images

ANDREW G. DETWILER

Institute of Atmospheric Sciences, South Dakota School of Mines and Technology, Rapid City, South Dakota

NANCY C. KNIGHT AND ANDREW J. HEYMSFIELD

Mesoscale and Microscale Meteorology Division, National Center for Atmospheric Research, Boulder, Colorado*

3 January 1992 and 29 May 1992

ABSTRACT

Ice particles are captured, photographed, melted, and then photographed again. Mass is estimated from the size of the melted drop. Based on a sample of 640 particles, the standard error in estimating particle mass solely from the maximum dimension of the particle is found to be a factor of 4. The standard error in estimating mass concentration M in a cloud from a sample of n well-characterized particles recorded by an optical array probe is estimated to be approximately a factor of $10^{0.6/\sqrt{n}}$.

1. Introduction

Since one of the most important characteristics of convective clouds is their ability to convert water vapor to precipitation, it is of obvious importance to understand the process through which this is accomplished and to determine the efficiency with which it may occur. Precipitation efficiency has been simply defined by Cotton and Anthes (1989) as "the ratio of the measured precipitation rate at the surface to the water vapor flux through the base of a cloud system." The difficulties inherent in the task of measuring the quantities required to calculate such a ratio approach the Sisyphean. Many factors that may be relevant are often omitted from such calculations because they have not been reliably measured or because the scope of the task exceeds the facilities presently available (Fankhauser 1988; Heymsfield and Miller 1988; among others).

Among the myriad difficulties is the problem of accurately determining the total mass of liquid water and ice in the cloud—those particles that will form the precipitation. A variety of methods have been used, both in situ and remote, including evaporators, optical array probes, and microwave measurements of cloud reflectivity, transmissivity, and emissivity. No one technique has been found applicable in all situations.

We concern ourselves here with a discussion of mass concentration estimates from optical array probe data. Estimates of the possible errors in particle mass concentration determinations from in situ optical array probe data have generally been about a factor of 2 (e.g., Holroyd 1987; Heymsfield and Palmer 1986; Heymsfield and Miller 1988; Heymsfield et al. 1990; Braham et al. 1992). The technique generally adopted is to use the image data to estimate an apparent maximum dimension D for each sampled particle and then apply an empirically determined power-law prediction equation relating the particle mass M to D. In some cases, the desired parameter is mass flux (i.e., precipitation rate), so an additional relationship relating D to fall speed V is applied (e.g., Holroyd 1987; Braham et al. 1992).

Holroyd (1987) compared directly measured mass flux (snowfall) at the ground to the mass flux calculated from a fixed Particle Measuring Systems (PMS) 2D-C mounted near the ground. He concluded that with proper choice of M versus D and V versus D relationships, agreement between measured and calculated mass flux to within a factor of 2 is possible. He did not separately consider the magnitude of the error in the mass concentration estimate.

The M versus D relationships available in the literature (e.g., Locatelli and Hobbs 1974; Mitchell et al. 1990) vary considerably with particle habit, as will be discussed further below. Heymsfield and Palmer (1986) and Heymsfield et al. (1990) qualitatively addressed the uncertainty in estimating M from D, which is attributable to imperfect knowledge of particle habit. They concluded that the uncertainty in ice water mass

* The National Center for Atmospheric Research is sponsored by the National Science Foundation.

Corresponding author address: Andrew G. Detwiler, Institute of Atmospheric Sciences, South Dakota School of Mines and Technology, 501 East St. Joseph Street, Rapid City, SD 57701-3995.

integrated over a measured particle spectrum **M**, could be as great as a factor of 2. Heymsfield and Miller (1988) concluded that this uncertainty could be reduced to approximately 25% if collocated radar reflectivity data were available and a refined estimate of **M** was produced by the method developed by Plank et al. (1980) using both images and radar data.

Braham et al. (1992) compared mass flux in snowfall estimated from airborne 2D probe data to flux based upon reflectivity data from ground-based radar. They concluded that agreement of fluxes within 25% was possible with the proper choice of an M versus D relationship.

The problem of estimating particle mass concentrations in clouds can be seen to have two components. The first is acquiring the number and sizes of particles in a known volume. The second is finding the total mass of the particles by applying an appropriate M versus D relationship. Section 2 gives results from investigations concerning the accuracy of estimates of M from D. Section 3 contains rough estimates of the uncertainty in the computation of mass concentration from a sample containing many particles.

2. Individual particles

Data were obtained in the National Center for Atmospheric Research cold laboratory in Boulder, Colorado, over three winter seasons from 1983 through 1986. The ice particles sampled were primarily from clouds close to the surface in upslope situations. Particles fell directly into the laboratory through an insulated chimney and were first photographed in an array of strobe lights as they fell (see Fig. 1). Some of the particles fell through the sample volume of a PMS 2D-C or 2D-P optical array probe and were imaged by the probe as well. All particles were finally captured in a hexane-filled petri dish. The captured particles were photographed, then melted, and the resulting drop photographed (all while in hexane). Particle size and habit were determined from the photographs and the mass from the melted drop size. In obtaining particle mass, only those particles that melted in single drops were used. The error resulting from this method was calculated by using drops of known volume and was found to be negligible. Although 640 particles passed through the probe and were captured, it was only possible to positively identify a small number (35) as being both imaged and collected, and a second measurement of D was obtained from the probe image of these particles. During the first season of the experiment, the 2D-C probe was damaged and a 2D-P was substituted while the 2D-C was being repaired.

Figure 2 gives the mass of the particles as determined from their melted drop diameters as a function of their apparent maximum dimension from photographs. The observed trend, $M \propto D^2$, agrees with those obtained in other similar studies of snow-particle size and mass (e.g., Locatelli and Hobbs 1974; Mitchell et al. 1990).

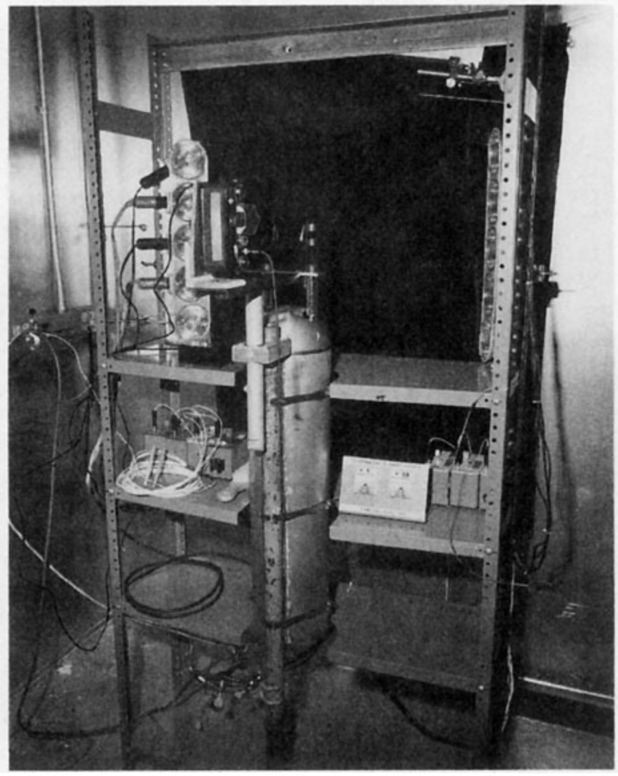

FIG. 1. Experimental apparatus in cold laboratory.

One commonly used method for examining the correlation between M and D in data similar to that shown in Fig. 2 is a least-squares linear regression (LSLR) of the logarithm of the dependent variable, here M, versus the logarithm of the independent variable, here D. The result of such analysis is a linear relationship between $\log M$ and $\log D$, or equivalently, a power-law relationship between M and D. It is not obvious in such data which parameter is independent and which dependent; therefore, M and D might be reversed, giving a different best-fit line. Another approach is to take the bisector of the two best-fit lines as a measure of the association between M and D.

The best fit line in Fig. 2 was computed by LSLR using $\log D$ as the independent variable and $\log M$ as the dependent variable. As the figure shows, there is fairly uniform scatter about the line and, for any value of D, there is at least an order of magnitude range in the variation of M. The LSLR relationships for $\log D$ on $\log M$ and the equation of the line bisecting these two lines were also computed.

In addition to using LSLR, a 1D principal component analysis, or PCA (e.g., Preisendorfer 1988), of our $\log M$ versus $\log D$ data was computed as well.

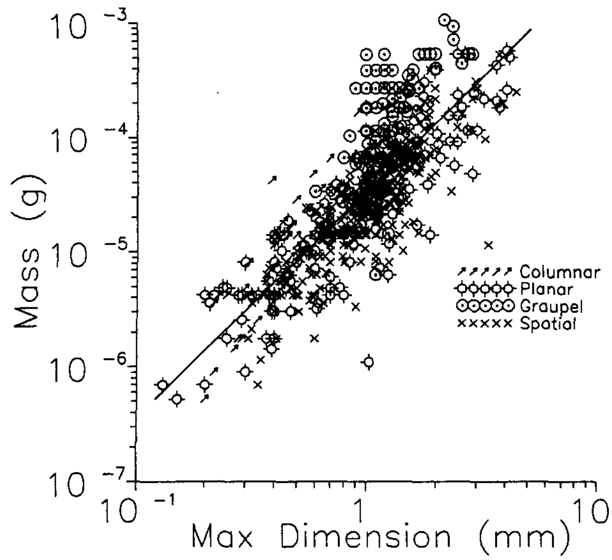

FIG. 2. Masses are plotted for 640 particles as a function of their maximum dimension as measured from photographs. Dashed line shows the LSLR of $\log M$ on $\log D$ and is defined by $M = 3.41 \times (10^{-5}) D^{1.97}$, where D is in millimeters and M is in grams.

Least-squares linear regression analysis minimizes the squared vertical deviation of the dependent variable from the best fit. The fundamental difference in PCA is that the squared deviation of the data from the best-fit line is measured orthogonally (see Fig. 3). Linear regression analysis assumes that the independent variable is perfectly known and that observational uncertainty exists with respect to the dependent variable. Principal component analysis assumes that uncertainty exists with respect to both observed variables and provides a measure of the association of the two.

The collected particles were also subjectively characterized as belonging to one of four categories: graupel, planar, spatial, or columnar. The power law relationships, computed by LSLR and PCA, are given in Table 1 for each of the categories, as well as for the entire set grouped together. In general, the exponent in these relationships is smallest for planar particles and largest for graupel, but it varies substantially for all categories depending upon the method used to compute it. Planar particle masses increase roughly as the square of their apparent maximum dimension. Other particle types increased in mass by a power higher than 2, and sometimes higher than 3, in proportion to their apparent maximum dimension.

The most interesting quantity to the authors is not the precise M versus D relationship but the standard deviation of the data from the best-fit line, however determined. Correlation coefficients are given by Locatelli and Hobbs (1974) and by Mitchell et al. (1990), but the data included in these papers do not allow one to calculate the standard deviation of the data about the power law describing the best-fit line. This standard deviation can be used to derive a measure of the accuracy with which the mass of a snow crystal may be estimated from its maximum dimension and some knowledge of its habit. The measure used is $\sigma/\sqrt{n-2}$, where σ is the standard deviation of the data about the regression line and n is the number of particles in the sample. This measure is called the *standard error,* following Snedecor and Cochran (1967, p. 386). These standard errors are given in Table 2 as the antilog of the standard errors from the LSLR and PCA $\log M$–$\log D$ analyses. Here the estimated error is seen to be a factor of approximately 2 using LSLR and 1.3 using PCA.

We interpret these results in the following way. One-dimensional PCA shows that M and D are quite closely associated. Maximum dimension D might be thought of as an imperfect proxy for an ideal size measure that would be even more closely associated with M. To the degree that the standard error associated with the LSLR line is larger that the standard error associated with the PCA line, the use of D as a proxy for this ideal size measure degrades the accuracy with which M can be estimated.

As previously noted, only a small number of particles could be positively identified in both the collections and probe image. One such example is shown in Fig. 4. Figure 5 gives a comparison of D based upon the probe image compared with D measured from the captured particle. Only those 33 particles in which D was less than 2 mm were included. For the particles in this

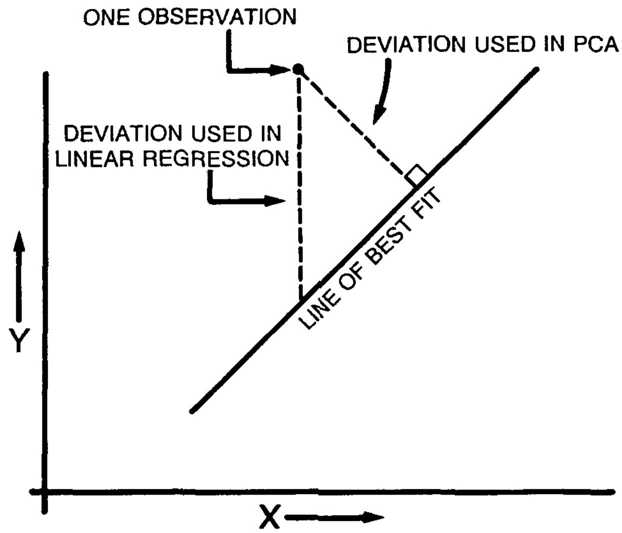

FIG. 3. Diagram comparing the quantities minimized in linear regression and principal component analysis.

TABLE 1. Power-law parameters for LSLR and PCA relationships of the form $M(g) = aD^b$ (mm).

	LSLR M vs D		LSLR D vs M*		LSLR bisector		PCA		
	a	b	a'	b'	a	b	a	b	n
Planar	2.54 (−5)**	1.79	3.00 (−5)	2.35	2.91 (−5)	2.04	2.85 (−5)	1.79	171
Graupel	1.07 (−4)	2.06	4.96 (−5)	5.29	8.50 (−5)	3.02	5.57 (−5)	4.81	87
Spatial	2.64 (−5)	1.95	2.63 (−5)	3.18	2.64 (−5)	2.34	2.64 (−5)	2.36	337
Columnar	6.03 (−5)	2.28	1.70 (−4)	3.51	9.18 (−5)	2.78	1.18 (−4)	3.07	45
All	3.41 (−5)	1.97	3.60 (−5)	3.02	3.49 (−5)	2.40	3.51 (−5)	2.51	640

* Relationship derived from regression of log$D$ on log$M$ to yield $D = aM^b$, then inverted to the form $M = a'D^{b'}$, where $a' = (1/a)^{1/b}$ and $b' = (1/b)$.
** $x.xx\ (-y)$ signifies $x.xx \times 10^{-y}$.

sample, LSLR indicates that the 2D-C typically undersizes them. If LSLR is used to predict D obtained from the particle photographs from D obtained from the probe, the standard error of estimate is 0.23 mm. Thus, if the power-law relationships given in Table 1, which are based on photographic D, are used to estimate M from a probe-derived D, additional bias and random errors are introduced. Braham et al. (1992) have recently discussed the bias for undersizing of snow particles by optical array probes.

Although the sample in Fig. 5 is small, it does give a rough estimate of the contribution of sizing errors to the uncertainty in estimating M from D. For $D \sim 1$ mm, an uncertainty of approximately 0.2 mm in D would lead to an uncertainty in M of a factor of 3/2 using expressions given in Table 1. Assuming that the uncertainty in determining D from probe data is uncorrelated with the uncertainty in determining M from D, the overall uncertainty in estimating M from probe-derived D is the product of these two independent errors. The uncertainties add logarithmically but multiply when the M–D relationship is cast in power-law form. The overall uncertainty becomes a factor of 3 at this point in our analysis. A larger sample size would be desirable in order to refine the estimates derived here, particularly the correction for the undersizing bias. It should also be noted that 1D optical array probes may have different uncertainties from 2D probes since they measure maximum size in only one dimension.

Using one power law for all particles regardless of type will introduce additional uncertainty. As is shown in Table 2, when one power law is used for all particles, the standard error in the estimate of M increases by about a factor of 1/3 over the lowest error obtained when different power laws are devised for the four subclasses using LSLR. Assuming the undersizing bias can be accounted for and all remaining errors are uncorrelated, the total error in estimating M from D accumulates as follows: a factor of 2 in estimating M from D; a factor of 3/2 in determining D from an optical array probe; and a factor of 4/3 due to uncertainty in particle type. This yields a factor of 4 for a total standard error of estimate. A factor of 4 variation in M corresponds to a variation of ±0.6021 in log$M$. Of the individual components in this overall error, our data allows us to estimate the first and last components reasonably well, but our estimate of the error in estimating D from probe data is based upon a fairly small sample and should be regarded as somewhat tentative.

The data used in this investigation were collected at a fixed location over a period of three years. Although

TABLE 2. Antilog standard error of estimate.

	M vs D	PCA
Planar	2.11	1.39
Graupel	2.16	1.26
Spatial	1.84	1.27
Columnar	2.05	1.27
All	2.30	1.38

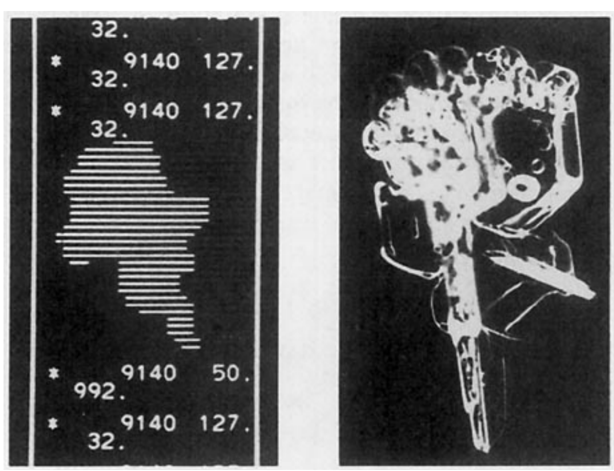

FIG. 4. Probe image (left) and identified particle (right).

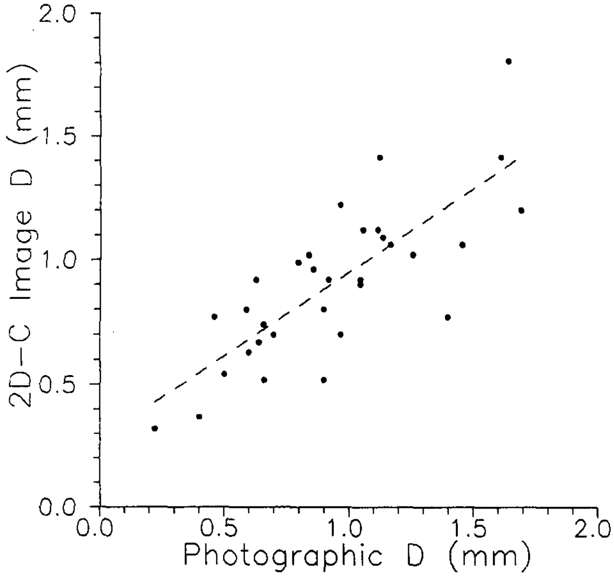

FIG. 5. Maximum dimension from probe image D, compared with D from photographs of same particle. Dashed line shows the LSLR of probe D against photographic D and is defined by [$D_{\rm probe} = 0.67 D_{\rm photo} + 0.28$ (mm)]. The particles in this sample are predominantly spatial.

the power-law fits obtained agree well with those obtained in the Sierras by Mitchell et al. (1990) and in the Cascades by Locatelli and Hobbs (1974) for similar particles, it is not positively possible to eliminate systematic variation in both power law and variability for a given particle type at another location or altitude.

3. Particle ensembles

If our error estimates are representative, then the error in an estimate of the mass loading of ice particles in a volume of cloud can be quantified. This may be done by using a power-law M versus D relationship appropriate to the particle population. As has been shown, assuming random and unbiased errors in the measurement of D leads to an uncertainty of a factor of 4 in the estimate of the mass of a particle. The uncertainty inherent in the calculations of mass in samples containing many particles will then change with the number n of particle images included in the mass-loading estimate as (see, e.g., Snedecor and Cochran 1967, p. 146)

$$\mathbf{M} = \begin{cases} (10^{0.6/\sqrt{n-2}}) \sum M_i, \\ \quad \text{plus one standard error} \\ \sum M_i, \quad \text{best estimate} \\ (10^{-0.6/\sqrt{n-2}}) \sum M_i, \\ \quad \text{minus one standard error,} \end{cases} \qquad (1)$$

where $\mathbf{M}$ is the total mass and M_i is the individual particle mass. Including more particles in a sample will result in a larger absolute but smaller relative uncertainty in the mass loading. This suggests that if the mass of a single particle can be estimated to within a factor of 4, then the mass corresponding to a sample of 100 particles can be estimated with a standard error of approximately 15%. Of course, particles too small or too large to be imaged, or too sparsely distributed to show up regularly in a given sample volume, cannot fully contribute to the estimate of mass loading.

Although uncertainty may be reduced by using a larger sample, in the context of aircraft measurement with a given probe, this requires extending a long, threadlike sample volume over a greater distance in the cloud and thus the possibility of including particles from distinctly different microphysical populations.

4. Conclusion

We suggest that the limiting uncertainty with which ice particle masses can be estimated from their apparent maximum dimension as determined by a 2D optical array probe is approximately a factor of 4. Accuracy of mass estimates in a volume containing many particles will be better than this as long as measurement errors can be considered random and uncorrelated. For example, accuracy of approximately 15% should be possible for an estimate of total mass represented by a sample of 100 particles. To the extent the sample is representative, the mass concentration in the cloud can also be estimated with similar precision.

Better accuracy than this might be desired by those tracing water substance through its various transformations in clouds. As Pat Squires wrote in 1967,

> ". . . in the case of earth, we have . . . developed our knowledge of empirical facts and our partial physical explanations of these facts to the point where a new hypothesis will usually deal with details to such a degree that few or no previous observations are adequate, and an ad hoc experiment or set of observations will be needed to test the hypothesis."

Many observations and experiments have occurred since that was written, but new techniques, experiments, and observations are still required to understand many aspects of precipitation efficiency, of which accurately defining the distribution of mass in clouds is a small part.

Acknowledgments. Support of this research was provided by the National Science Foundation under Cooperative Agreements ATM-8620145 and ATM-9104474 and through sponsorship of the National Center for Atmospheric Research (NCAR). Additional support came from the National Aeronautics and Space Administration under Contract 1-98100-B to NCAR and from the state of South Dakota.

REFERENCES

Braham, R. R., Jr., P. A. R. Kristovich, and M. J. Dungly, 1992: Comparison of lake-effect snow precipitation rates determined from radar and aircraft measurements. *J. Appl. Meteor.,* **31,** 237–246.

Cotton, W. R., and R. A. Anthes, 1989: *Storm and Cloud Dynamics.* Academic Press, 883 pp.

Fankhauser, J. C., 1988: Estimates of thunderstorm precipitation efficiency from field measurements in CCOPE. *Mon. Wea. Rev.,* **116,** 663–684.

Heymsfield, A. J., and A. G. Palmer, 1986: Relationships for deriving thunderstorm anvil mass for CCOPE storm water budget measurements. *J. Climate Appl. Meteor.,* **25,** 691–702.

——, and K. M. Miller, 1988: Water vapor and ice mass transported into the anvils of CCOPE thunderstorms. Comparison with storm influx and rainout. *J. Atmos. Sci.,* **45,** 3501–3514.

——, ——, and J. D. Spinhirne, 1990: The 27–28 October 1986 FIRE IFO cirrus case study: Cloud microstructure. *Mon. Wea. Rev.,* **118,** 2313–2328.

Holroyd, E. W., III, 1987: Some techniques and uses of 2D-C habit classification software for snow particles. *J. Atmos. Oceanic Technol.,* **4,** 498–511.

Locatelli, J. D., and P. V. Hobbs, 1974: Fall speeds and masses of solid precipitation particles. *J. Geophys. Res.,* **79,** 2185–2197.

Mitchell, D. L., R. Zhang, and R. L. Pitter, 1990: Mass-dimensional relationships for ice particles and the influence of riming on snowfall rates. *J. Appl. Meteor.,* **29,** 153–163.

Plank, V. G., R. O. Berthel, and A. A. Barnes, Jr., 1980: An improved method for obtaining the water content values of ice hydrometeors from aircraft and radar data. *J. Appl. Meteor.,* **19,** 1293–1299.

Preisendorfer, R. W., 1988: *Principal Component Analysis in Meteorology and Oceanography.* Elsevier, 425 pp.

Snedecor, G. W., and W. G. Cochran, 1967: *Statistical Methods.* Iowa State University Press, 593 pp.

Squires, P., 1967: Cloud physics—Prose or poetry? *Bull. Amer. Meteor. Soc.,* **48,** 400–403.