

NOTES AND CORRESPONDENCE

Comments on “The Window Probability Matching Method for Rainfall Measurements with Radar”

GRZEGORZ J. CIACH AND WITOLD F. KRAJEWSKI

Iowa Institute of Hydraulic Research, University of Iowa, Iowa City, Iowa

JAMES A. SMITH

Department of Civil Engineering and Operations Research, Princeton University, Princeton, New Jersey

9 January 1995 and 26 February 1996

1. Introduction

The probability matching method (PMM) appeared in radar meteorology as a remedy to overcome difficulties with relating radar reflectivity (Z) and rain gauge precipitation (R) measurements. The typically large differences between radar and rain gauge observations are often attributed to the errors in time and space synchronization of the two measurements. The main idea behind the PMM is to construct a Z - R transformation by relating the separate cumulative frequencies of both signals, rather than the synchronized pairs of the measured values. In statistical language this means that the PMM is looking for a relation between two variables based on their marginal distributions only, without using their joint distribution. Analysis of some statistical properties of this method was presented by Krajewski and Smith (1991). They discovered very slow convergence of the PMM and its tendency to produce biased estimates. Here, we will elaborate more on the issue of the validity of the methods based on distribution matching.

Recently, Rosenfeld et al. (1994) modified the PMM to include synchronization of the two sensor measurements by forming pairs over space-time windows. They term the modified method the window probability matching method (WPMM). The windows are to be larger than the errors in space and time synchronization of the measurements, but small enough to maintain some level of physical coupling between the values measured. After selecting the data samples in this way, they apply probability matching relating reflectivities and gauge

rain intensities, which have the same unconditional (including zero values) sample cumulative probabilities.

The purpose of our comment is to analyze some of the properties that are common to both the PMM and the WPMM, and to discuss their validity. We also show that some of the effects obtained by Rosenfeld et al. (1994) have a statistical origin and their physical interpretation is doubtful. Our demonstration is based on simple simulation examples designed to illustrate some specific features of the methodology. This comment is organized as follows. First, we introduce a mathematical definition of the method and show that it always gives some relationship, even for completely unrelated physical variables. Next, we discuss the WPMM evaluations and their interpretations in Rosenfeld et al. (1994). A summary and conclusion section completes this comment.

2. Basic properties of the methodology

Let X and $Y = g(X)$ be two random variables related by a monotone transformation g . Our purpose is to find the transformation g based on two samples of observed values of X and Y . Usually, both X and Y observed sample values are corrupted with significant random errors, and the relationship between the random variables exists only in a statistical sense. Traditionally, we are looking for an estimate \hat{g} of the relation g , which optimizes some desired criterion. For example, a regression function defined by the conditional mean of random variable Y given X minimizes the mean square error. Another estimate might be the conditional median, minimizing the average absolute error. Note that this statistical approach can be applied only if some information about the joint distribution of X and Y is known.

Corresponding author address: Prof. Witold F. Krajewski, Iowa Institute of Hydraulic Research, University of Iowa, Iowa City, IA 52242.

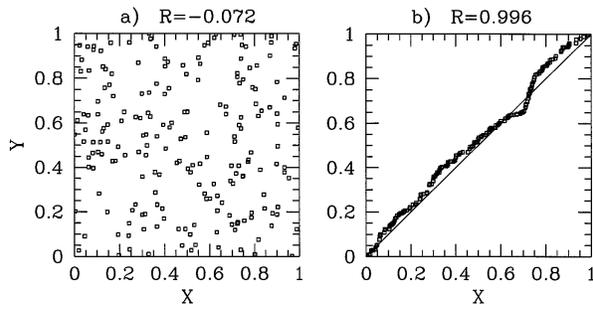


FIG. 1. The results of the application of the probability matching to two completely unrelated variables for sample size $N = 200$. (a) Original unsynchronized pairs of data points. (b) Pairs of values related to each other by using the transformation (1). Erroneous inference of a relationship between X and Y is evident. The parameter R is the correlation coefficient. The samples of X and Y values are generated from the uniform distribution over the $[0,1]$ interval.

The PMM and WPMM use quite different estimation schemes. Given two measurement samples, $\{X_i, i = 1, \dots, n_x\}$ and $\{Y_i, i = 1, \dots, n_y\}$, one can estimate the marginal cumulative distribution functions (CDF) separately for X and Y . The estimation of the relationship g in both PMM and WPMM can be generally formulated as follows:

$$\hat{y} = \hat{g}(x) = \hat{F}_y^{-1}[\hat{F}_x(x)], \quad (1)$$

where \hat{F}_z denotes the estimated CDF of the variable z . The expression says that $\hat{g}(x)$ is estimated by the sample quantile function of Y , evaluated at the cumulative probability $\hat{F}_x(x)$.

We have to stress that the estimator (1) is the same for PMM and WPMM. So the difference in data sample selection, which is the fundamental difference between PMM and WPMM, does not affect their properties (discussed below). One basic property is that a relationship is obtained even for completely independent variables, as all variables have some marginal distributions that can be plugged into (1). To illustrate how it works in this situation, let us generate two independent random samples of X and Y with 200 values each from a uniform distribution over the interval $[0,1]$ and apply (1) to them. The results before and after using the method are presented in Fig. 1. First we see that since the samples are independent, the points are scattered randomly in Fig. 1a, without any pattern. However, application of the quantile matching results in a strong relationship close to $\hat{Y} = \hat{g}(X) = X$, with a correlation coefficient for the matched pairs of X and Y of 0.996 (Fig. 1b). Clearly, the solution is misleading, and the relationship obtained does not reflect any physical connection between those variables.

The estimator (1) does not differentiate between the deterministic, partially correlated, and fully independent situations. In all cases some function is obtained, regardless of the existence or nonexistence of any actual physical dependence. Certainly, the estimator (1) cannot distinguish between signal and noise. In most situations

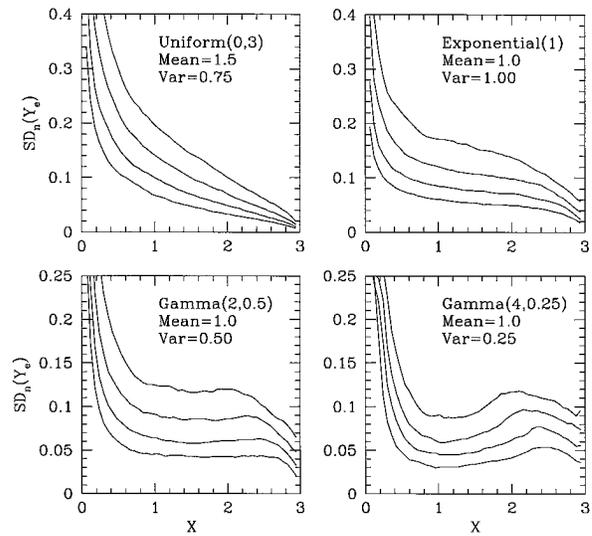


FIG. 2. Normalized standard deviations of probability matched relationships between two independent random variables X and Y for different marginal distributions and sample sizes. The normalized standard deviations express the relative error of Y estimation. Note the rapid growth of the error for small X at all panels and the local maxima at the two lower panels.

the shape of the relationship will depend on the error structure of observations, as all sources of uncertainty load to the shape of the variable marginal distributions. To make things worse, the method also accepts any measurements and is incapable of distinguishing outliers from valid data. However, based on the methodology, we are unable to tell how much we are misled.

3. Tests of the methodology

By repeating the simulation experiments independently many times for given marginal distributions, we can assess the uncertainty of a relationship estimated by (1). We have to stress that we do not give here any indication about the degree of pairwise organization that we maintained in our simulation. The sample correlation could have any value between zero and almost one (Fig. 1) without affecting the marginal CDFs. The aspect of data synchronization is simply irrelevant for the properties discussed here.

We present the results in Fig. 2 in terms of the normalized standard deviation $SD_n(\hat{Y})$, which expresses the relative estimation errors exactly the same way as in Rosenfeld et al. (1994). To obtain reliable and stable results, we repeated the experiment 1000 times for each distribution and sample size. The normalization was performed by dividing the sample variance of the estimate \hat{Y} by its average value. This experiment was performed for four examples of probability distributions of the two variables X and Y , as shown in Fig. 2. The normalized standard deviations are presented as functions of the measured values of the variable X for the four sample sizes for each distribution. The sample sizes are 100,

200, 400, and 800 data points, and the normalized standard deviation decreases with increasing sample. The shape of the curves depends on the distribution type, but also reproduces fairly well some general features of the corresponding results obtained by Rosenfeld et al. (1994) and presented in Figs. 2 and 6 of their paper. Below, we discuss two of these features.

Rosenfeld et al. (1994, 686) draw conclusions from Fig. 2 of their paper. They write, "The most striking result of this exercise is the finding that the Darwin radar can measure rain intensities with reasonable accuracy for reflectivities greater than about 25 dBZ, but that the accuracy falls sharply below this reflectivity threshold." Our simulations have no underlying physics; nevertheless, they exhibit exactly the same behavior: $SD_n(\hat{Y})$ grows rapidly in the region of small X for all distributions. This feature is clearly a general mathematical property of the transformation (1), which appears regardless of any physical reality or actual measurement errors.

On page 689 Rosenfeld et al. write, "According to Figs. 2 and 6 there is another broad minimum in the accuracy, or maximum $SD_n(R)$, around 35 dBZ," and then, "It is suggested here that the possible coexistence of both stratiform and convective rainfall in this range of reflectivities, but with different rain intensities, contributes to the increased $SD_n(R)$ around 35 dBZ." Again, the lower two panels in Fig. 2 show that such maxima can appear for some marginal distributions without any physical causes. Explaining this feature as evidence of some desired physical factors is highly speculative, as it can also appear without those factors for purely statistical reasons.

The normalized standard deviations do not reflect the performance of the method in terms of radar rainfall estimation accuracy, but only the stability of the distribution transformation fitting, which has quite different meaning. Associating them with measurement errors and other physical effects is a misinterpretation of the method, which by itself has no mechanism for separating signal from noise; $SD_n(R)$ depends mainly on the sample size and, to a smaller degree, on the marginal distribution type. It does not depend in any way on the pairwise organization (synchronization) of the data points.

The last result that we want to comment on is the comparison of the $Z-R$ functions obtained by WPMM with the power-law relationship based on disdrometer data taken from Short et al. (1990). Figure 9 in Rosenfeld et al. (1994) indicates that this disdrometer relationship, used as a reference, is generally biased and underestimates the rainfall intensities by a factor of about 2. The bias itself does not affect the correlations; however, it is a clear indicator that the reference $Z-R$ relationship used for the comparison test has some problems. It is a recognized fact that disdrometer information explains only a small portion of uncertainties involved in radar measurements (Zawadzki 1984), and the $Z-R$

relationships obtained from a disdrometer should not be applied directly to radar reflectivities without recalibration for the specific radar and preprocessing system. Thus, a possible explanation why the correlations between the estimated and measured accumulations are higher for the WPMM estimate than for the power law estimate is that the latter is questionable. One should also note that the correlation test was applied to the same samples on which the WPMM $Z-R$ relationships were based, whereas the reference relationship was calibrated on another, independent dataset. This obviously favors the WPMM.

4. Summary and conclusions

In this comment we have demonstrated two significant drawbacks of the methods based on transformation (1). The first is their ability to build a relation between any two random variables, regardless of whether they are physically related or not. As a result, the $Z-R$ functions obtained reflect the formal distribution transformation only, and their predictive and explanatory potential in many practical situations might be doubtful. For any two random variables their marginal distributions bear no information about the existence and the nature of any dependencies between them. Consequently, it is virtually impossible to explain and verify the physical meaning of the effects expressed by the relationships obtained using (1).

The second major problem with the methods based on (1) is their lack of internal diagnostics. This means that they must accept all data points with equal weight and cannot filter information from noise. This is contrary to modern statistical estimation methods, for which extracting signal from noise is the main task. It is very likely that for a certain kind of noise structure the methodology can give a relationship that is far from any physical effect that it is trying to describe.

Conceptually, the methods using "probability matching" (1) arise from the assumptions that there exist almost one-to-one relationships between radar reflectivities and rain gauge intensities, and that the predominant difficulty, after reducing the uncertainties caused by the drop size distribution variations, is the lack of good synchronization of both measurements in space and time. This is obviously not the case in radar rainfall estimation, where uncertainties are numerous and arise from many other sources. The fundamental difference between point, time-integrated sampling (rain gauges) and space-integrated, instantaneous sampling (radar) of a very complex spatiotemporal stochastic process (precipitation) cannot be represented by this crude simplification. Unfortunately, the existing validation methodologies and the observational capabilities are not adequate to assess different estimation schemes and to test appropriate hypotheses. In the absence of sufficiently reliable techniques to measure aerial rainfall, application of computer simulations and/or modern statistical in-

ference offer alternative ways to at least conceptualize possible pathologies.

Acknowledgments. This work was supported by National Aeronautics and Space Administration (NASA) Grant NAG 5-2084. The first author was supported by NASA under a Graduate Student Fellowship in Global Change Research, Reference 4146-GC93-0225, and by the United States Agency for International Development under Grant HRN-5600-G-00-2037-00. The authors gratefully acknowledge this support. The authors also acknowledge helpful discussions with Richard Dykstra from the Statistics Department of the University of Iowa and with V. Chandrasekar from the Electrical Engineering Department of Colorado State University.

REFERENCES

- Krajewski, W. F., and J. A. Smith, 1991: On the estimation of climatological $Z-R$ relationships. *J. Appl. Meteor.*, **30**, 1436–1445.
- Rosenfeld, D., D. B. Wolff, and E. Amitai, 1994: The window probability matching method for rainfall measurement with radar. *J. Appl. Meteor.*, **33**, 682–693.
- Short, D., T. Kozu, and K. Nakamura, 1990: Rainfall and raindrop size distribution observations in Darwin Australia. *Proc. URSI-F Open Symp. on Regional Factors in Predicting Radiowave Attenuation Due to Rain*, Rio de Janeiro, Brazil, Union of Radio Science International, 35–40.
- Zawadzki, I., 1984: Factors affecting the precision of radar measurements of rain. Preprints, *22nd Radar Meteorology Conf.*, Zurich, Switzerland, Amer. Meteor. Soc., 251–256.