Nonlinear Estimation of Spatial Distribution of Rainfall
- An Indicator Cokriging Approach

Dong-Jun Seo

Hydrologic Research Laboratory, National Weather Service, Silver Spring, MD 20910

Submitted to Stochastic Hydrology and Hydraulics

October, 1994

Nonlinear Estimation of Spatial Distribution of Rainfall - An Indicator Cokriging Approach

Dong-Jun Seo

Hydrologic Research Laboratory, National Weather Service, Silver Spring, MD 20190

ABSTRACT

Indicator cokriging (Journel 1983) is examined as a tool for real-time estimation of rainfall from rain gage measurements. The approach proposed in this work obviates estimation of time-varying statistics of rainfall by using ensemble or climatological statistics exclusively, and reduces computational requirement attendant to indicator cokriging by employing only a few auxiliary cutoffs in estimation of conditional probabilities. Due to unavailablity of suitable rain gage measurements, hourly radar rainfall data were used for both indicator covariance estimation and a comparative evaluation. Preliminary results suggest that the indicator cokriging approach is clearly superior to its ordinary kriging counterpart, whereas the indicator kriging approach is not. The improvement is most significant in estimation of light rainfall, but drops off significantly for heavy rainfall. The lack of predictability in spatial estimation of heavy rainfall is borne out in the integral scale of indicator correlation: peaking to its maximum for cutoffs near the median, indicator

correlation scale becomes increasingly smaller for larger cutoffs of rainfall depth. A derived-distribution analysis, based on the assumption that radar rainfall is a linear sum of ground-truth and a random error, suggests that, at low cutoffs, indicator correlation scale of ground-truth can significantly differ from that of radar rainfall, and points toward inclusion of rainfall intermittency, for example, within the framework proposed in this work.

1. INTRODUCTION

In recent years, a number of works using kriging has been reported in the area of estimating point or mean areal rainfall using rain gage measurements (Chua and Bras 1982, Creutin and Obled 1982, Bastin et al. 1984, Tabios and Salas 1985, Lebel et al. 1987, Lebel and Laborde 1988, Barancourt et al. 1992, and others). Although recognized as superior to heuristic or ad-hoc ones, kriging approaches in rainfall estimation, specifically those using linear kriging, has not been as readily and widely accepted in operational hydrology as once thought. The main reasons for this are seen as: 1) kriging requires time-varying statistics of each individual rainfall field at hand, but in real-world situations reliable estimation of them is practically impossible due to lack of data, and 2) because of highly skewed nature of rainfall distribution, the margin of improvement that linear kriging can consistently provide may not be large enough to warrant its routine application.

With respect to the first point, a number of researchers (Bastin et al. 1984, Lebel et al. 1987) have used so called 'climatological kriging' in which, instead of estimating time-varying semi-variogram for each realization of rainfall field, climatological or ensemble semi-variograms are estimated from long-term records, parameterized with respect to seasonality and rainfall depth, and then used in the framework of ordinary kriging for estimation of mean areal rainfall. With respect to the second point, nonlinear estimation may be attempted. For example, in the context of rainfall estimation using rain gage measurements and radar rainfall data, disjunctive cokriging has been shown to provide significantly better estimates of hourly rainfall than its linear counterparts (Azimi-Zonooz et al. 1987).

This work is an attempt to address the above points in rainfall estimation via indicator cokriging (Journel 1983, Journel and Posa 1990, Suro-Perez and Journel 1991). Making use of the conditional probability interpretation of the conditinal expectation of an indicator random variable, indicator cokriging allows nonparametric, nonlinear estimation of spatial random functions. It may be viewed as an equivalent of disjunctive kriging (Matheron 1975) in that bivariate probability distribution is utilized following its discrete approximation by a set of indicator covariances at various cutoffs (Journel 1983). Unlike lognormal (Journel and Huibjbregts 1978) or disjunctive kriging, however, indicator cokriging requires neither distributional assumptions nor variable transformations.

In this work, particular attentions are paid to 1) use of climatological or ensemble indicator statistics in place of time-varying statistics of rainfall, and 2) reduction in computational requirement by reducing the number of auxiliary cutoffs used in estimation of conditional probabilities. To evaluate its performance, the indicator cokriging approach is compared with its indicator and ordinary kriging counterparts. Due to unavailability of suitable rain gage measurements, however, radar rainfall data were used for both indicator covariance estimation and performance evaluation.

This paper is organized as follows: description of aspects of indicator cokriging which make the proposed approach attractive in rainfall estimation, description of indicator cokriging, presentation of indicator covariance structure of rainfall as obtained from radar rainfall data, description of issues and approaches taken in implemention of indicator cokriging, description of comparative evaluation, and presentation of conclusions and future research recommendations. Because radar rainfall data suffer from various types of error, indicator covariance structure of radar rainfall may not represent that of ground-truth rainfall very well. In Appendix, we examine how the two may differ by assuming that radar rainfall is a linear sum of ground-truth and a random error.

2. INDICATOR COKRIGING APPROACH IN RAINFALL ESTIMATION

Indicator cokriging utilizes bivariate probability distribution in the form of covariances of indicator variables at various cutoffs (see the following section). Estimation of indicator covariances from rainfall data calls for construction of experimental indicator variables at various cutoffs, which amounts to a natural stratification of rainfall variability with respect to rainfall depth. Accordingly, indicator statistics are well-suited for ensemble or climatological representation.

Since both indicator cokriging and disjunctive kriging utilize bivariate probability distribution, it may be expected that they perform comparably to each other given that the statistics required are perfectly known. When ensemble or climatological statistics are used, however, disjunctive kriging is seen as less suitable for real-time estimation in that, regardless of the sample (i.e., posterior) statistics of the data at hand, the unconditional mean of its estimate will be given by the ensemble or the climatological (i.e., a priori) mean because of its unbiasedness property (Journel and Huijbregts 1978, p575).

Linear cokriging under second-order homogeneity includes ordinary and simple types (see Journel and Huijbregts, 1978, for difference between ordinary and simple kriging, and, e.g., Seo et al., 1990, for an extension to cokriging), the choice depending on whether the spatially constant mean of the variables is known or not. Analogously, depending on whether the spatially constant mean of the indicator variables (or, equivalently, the cumulative distribution of the original variable) is known or not, two different formulations of indicator cokriging is possible, leading to the ordinary cokriging analogue (Suro-Perez and Journel 1991) and the simple cokriging analogue (Journel 1983). What makes indicator cokriging attractive in real-time estimation is that it does have the ordinary cokriging analogue, for which one may forego the extremely difficult (if not impossible) task of estimating time-varying cumulative distribution of rainfall from a very small number of data (while using ensemble statistics for indicator covariances), and still achieve the unbiasedness in the (unknown but spatially constant) mean. Disjunctive kriging, on the other hand, may be viewed as a simple kriging analogue in that the mean must be known a priori.

Indicator cokriging assumes second-order homogeneity of indicator variables, and hence, at least on theoretical grounds, we are limited to dealing with rainfall fields that satisfy at least the bivariate version of the strong homogeneity (Karlin and Taylor 1975). Checks on such a condition, however, require an independent analysis of a large number of different realizations, and were not within the scope of this work. In the following, we present a brief description of indicator cokriging in a general context for reference purposes (see also Journel, 1983, Journel and Posa, 1990, Suro-Perez and Journel, 1991), which may seem somewhat redundant to those who are already familiar with ordinary cokriging. Though estimates of mean areal rainfall is of more direct interest in operational hydrology, we pay in this work our attention mainly to estimation of point rainfall so that distributional characteristics of the estimates may also be examined.

3. ESTIMATION VIA INDICATOR COKRIGING

Define the indicator random variable, $I(u;z_c)$, as follows:

$$i(u;z_c) = \begin{cases} 1 \text{ if } z(u) \le z_c \\ 0 \text{ if } z(u) > z_c \end{cases} \qquad (1)$$

where $z(u)$ is the measurement at location u, and z_c is the cutoff. Under the assumption that indicator variables are second-order homogeneous, indicator cokriging amounts to performing ordinary cokriging of collocated experimental indicator variables at multiple cutoffs a number of times to estimate conditional cumulative probability distribution function (cdf) of the original variable, from which conditional expectation may be obtained via numerical integration.

Indicator Cokriging

For each cutoff z_{ck}, $k=1,\ldots,K$, estimate the conditional expectation $E[I(u_0;z_{ck})|I(u_i;z_{cj})=i(u_i;z_{cj}), i=1,\ldots,N, j=1,\ldots,M]$ at an arbitrary location u_0 using the following linear estimator, where N and M are the number of measurements at u_i's and the number of cutoffs z_{cj}'s used in the estimation, respectively:

$$i^*(u_0;z_{ck}) = \sum_{j=1}^{M} \sum_{i=1}^{N} w_i(z_{ck},z_{cj})\; i(u_i;z_{cj}) \tag{2}$$

where z_{ck} is the main cutoff, and z_{cj}, $j=1,\ldots,M$, $j\neq k$, are the auxilary cutoffs. In Eq. (2), $w_i(z_{ck},z_{cj})$'s are the optimal weights to be determined by minimizing the error variance, $E[\{I(u_0;z_{ck}) - i^*(u_0;z_{ck})\}^2]$ under the following M unbiasedness constraints:

$$\sum_{i=1}^{N} w_i(z_{ck},z_{cj}) = \delta_{kj} \qquad \text{for } j=1,\ldots,M \tag{3}$$

where

$$\delta_{kj} = \begin{cases} 1 & \text{if } j=k \\ 0 & \text{if } j\neq k \end{cases}$$

The optimal weights are obtained by solving the following ordinary cokriging analogue of the indicator cokriging system:

$$\begin{bmatrix} Q(z_{c1},z_{c1}) & \dots & Q(z_{c1},z_{cM}) & U^T \dots & \mathbf{0} \\ \vdots & & & & \\ Q(z_{cM},z_{c1}) & \dots & Q(z_{cM},z_{cM}) & \mathbf{0} \dots & U^T \\ U & \dots & \mathbf{0} & 0 \dots & 0 \\ \vdots & & & & \\ \mathbf{0} & \dots & U & 0 \dots & 0 \end{bmatrix} \begin{bmatrix} W(z_{ck},z_{c1}) \\ \vdots \\ W(z_{ck},z_{cM}) \\ \lambda_1 \\ \vdots \\ \lambda_M \end{bmatrix}$$

$$= \begin{bmatrix} Q_0(z_{ck},z_{c1}) \\ \vdots \\ Q_0(z_{ck},z_{cM}) \\ \delta_{k1} \\ \vdots \\ \delta_{kM} \end{bmatrix} \tag{4}$$

where $Q(z_{cm},z_{cn})$ is the (NxN) indicator covariance matrix whose ij-th entry is $Cov[I(u_i;z_{cm}),I(u_j;z_{cn})]$, $W(z_{ck},z_{cj})$ is the (Nx1) optimal-weight vector whose i-th entry is $w_i(z_{ck},z_{cj})$, $Q_0(z_{ck},z_{cj})$ is the (Nx1) indicator covariance vector whose i-th entry is $Cov[I(u_0;z_{ck}),I(u_i;z_{cj})]$, U is the (1xN) unit vector, U=[1,...,1], **0** is the (1xN) zero vector, **0**=[0,...,0] and λ_1 through λ_M are the Lagrange multipliers. The estimation variance associated with $i^*(u_0;z_{ck})$ is given by:

$$\mathrm{Var}[i^*(u_0;z_{ck})]$$

$$= \mathrm{Var}[I(u_0;z_{ck})] - \sum_{j=1}^{M} W^T(z_{ck}, z_{cj})\, Q_0(z_{ck}, z_{cj}) - \lambda_k \tag{5}$$

The indicator covariance, $\mathrm{Cov}[I(u_i;z_{cm}), I(u_j;z_{cn})]$, is given by:

$$\mathrm{Cov}[I(u_i;z_{cm}), I(u_j;z_{cn})]$$

$$= E[I(u_i;z_{cm})\ I(u_j;z_{cn})] - E[I(u_i;z_{cm})]\ E[I(u_j;z_{cn})] \tag{6a}$$

$$= \Pr[z(u_i) \le z_{cm}, z(u_j) \le z_{cn}] - \Pr[z(u_i) \le z_{cm}]\ \Pr[z(u_j) \le z_{cn}] \tag{6b}$$

where z_{cm} and z_{cn} are two generally different cutoffs. The corresponding indicator correlation is written as:

$$\mathrm{Cor}[I(u_i;z_{cm}), I(u_j;z_{cn})] = \frac{\mathrm{Cov}[I(u_i;z_{cm}), I(u_j;z_{cn})]}{\sigma_{Im}\ \sigma_{In}} \tag{7}$$

The indicator variance σ^2_{Im} is written as:

$$\sigma^2_{Im} = \Pr[z(u_i) \le z_{cm}]\ (1-\Pr[z(u_i) \le z_{cm}]) \tag{8}$$

The indicator cokriging estimate, $i^*(u_0;z_{ck})$ in Eq.(2), also has the following conditional probability interpretation:

$$E[I(u_0;z_{ck}) \mid I(u_i;z_{cj})=i(u_i;z_{cj}), i=1,\dots,N, j=1,\dots,M]$$

$$= \Pr[z(u_0) \le z_{ck} \mid I(u_i;z_{cj})=i(u_i;z_{cj}), i=1,\dots,N, j=1,\dots,M] \tag{9}$$

Then, once $i^*(u_0;z_{ck})$'s are obtained for all z_{ck}'s, $k=1,\dots,K$, the conditional estimate of the original variable at u_0, $z^*(u_0)$, can be obtained, e.g., from the 8ufollowing discrete approximation of the expectation operation (Journel 1983):

$$z^*(u_0) = \int_0^\infty x \, di^*(u_0;x)$$

$$\approx \sum_{k=1}^{K-1} z_{ck'} \cdot [i^*(u_0;z_{c,k+1}) - i^*(u_0;z_{ck})] \qquad (10)$$

where $z_{ck'}$ denotes the mid-point between z_{ck} and $z_{c,k+1}$.

Both in structure identification and computation, indicator cokriging can be very costly as 1) indicator covariances at all pairwise cutoff combinations have to be modelled, and 2) a linear system of dimension $\{(N+1)xM\}$ has be solved K times. If all the cross terms vanish in the indicator cokriging system, i.e., $Cov[I(u_i;z_{cm}), I(u_j;z_{cn})]=0$ if $m \neq n$ (see Solow, 1986, and Hu, 1988, for examples of such spatial processes), indicator cokriging is reduced to indicator kriging (Journel 1983, Young 1987, Journel and Posa 1990, Alli et al. 1990, Suro-Perez and Journel 1991), which greatly reduces computational requirement. In spatial prediction, zero indicator cross-covariances imply that, relative to the reference cutoff, the knowledge of how much smaller or larger the nearby measurements are of no additional value, so long as they are known to be either smaller or larger. Rainfall processes, however, are of diffusion-type, for which indicator cross-covariances do not vanish. As will be seen later in this paper, significant differences in performance exist between the indicator cokriging and the indicator kriging approaches.

Analogously to ordinary block kriging, indicator block cokriging can also be performed for estimation of block averages, by cokriging for spatial averages of indicator variables over a range of cutoffs. These estimates constitute an estimate of the cdf of the point variable because they may be interpreted as fractions or percentages, within the block, of the point values that are less than the corresponding cutoffs. The block estimate may then be obtained by evaluating the spatial mean from the cdf thus estimated.

Ensemble Versus Real-Time Statistics

As noted, the approach taken in this work was to use ensemble statistics of indicator covariance in cokriging. The price for using ensemble covariances instead of time-varying ones may vary depending on: 1) whether estimation variances are sought or not, 2) choice between indicator cokriging and indicator kriging, and 3) effectiveness of stratifying ensemble statistics with respect to important climatological and meteorological factors.

Of the three, the last factor is difficult to assess because extensive analyses of meteorological observations have to accompany. No such efforts were made in this work for the sake of simplicity, and hence ensemble statistics used in this work are simply the sample statistics obtained from using all the available data of the same type regardless of seasonality, type of storm, etc.

A few specific comments may be made on the first two factors. If one is interested only in estimates and not in estimation variances, it can be easily shown that indicator kriging (as well as simple or ordinary kriging) may be performed using correlation (or, in the case of indicator kriging, conditional probability) instead of covariance. In the case of hourly rainfall, it is observed that spatial correlation scale exhibits much smaller temporal variability than spatial mean or variance does (see, e.g., Seo and Smith 1991). It is seen to suggest that, if estimation variances are not sought, substituting ensemble correlation for time-varying correlation should not significantly reduce the kriging performance.

In the case of indicator cokriging, however, covariance, and not correlation, must be used (i.e., conditional probability does not suffice, and bivariate probability must be used) even if estimation variances are not sought. It is due to the fact that indicator variables at different cutoffs have varying degrees of variability (analogously to ordinary cokriging where different variables have different variabilities). The mechanics of cokriging is such that, with other conditions being equal, a relatively larger/smaller weight will be given to the experimental value of a variable (indicator or original) that has a smaller/larger variability. In indicator cokriging, this variability is specified by the indicator variance, for whose evaluation the univariate cdf (i.e., indicator mean) must be known.

Compared to climatological statistics of indicator correlation, climatological statistics of indicator variance may be considered as a less potent substitute in real-time estimation for time-varying statistics because cdf of rainfall fields can have substantial temporal variability. It is important to reiterate, however, that this substitution may result in increase in estimation variance but does not affect the unbiasedness in the mean.

4. INDICATOR COVARIANCE STRUCTURE OF RAINFALL

In general, the estimation domain of a drainage basin or a sub-basin frequently includes not only rain area but also no-rain area. The indicator cokriging approach offers a unified framework in which rainfall intermittency may also be dealt with.

Consideration for Intermittency

Conditioning on the intermittency of rainfall, indicator covariance may be decomposed as follows:

$$\mathrm{Cov}[I(v;z_l), I(u;z_c)]$$
$$= \Pr[z(v) \le z_l, z(u) \le z_c] - \Pr[z(v) \le z_l] \Pr[z(u) \le z_c] \tag{11a}$$
$$= \{ \Pr[0 < z(v) \le z_l, 0 < z(u) \le z_c] - \Pr[0 < z(v) \le z_l] \Pr[0 < z(u) \le z_c] \}$$
$$+ \{ \Pr[z(v)=0, 0 < z(u) \le z_c] - \Pr[z(v)=0] \Pr[0 < z(u) \le z_c] \}$$
$$+ \{ \Pr[0 < z(v) \le z_l, z(u)=0] - \Pr[0 < z(v) \le z_l] \Pr[z(u)=0] \}$$
$$+ \{ \Pr[z(v)=0, z(u)=0] - \Pr[z(v)=0] \Pr[z(u)=0] \} \tag{11b}$$

The four terms in Eq.(11b) represent indicator covariances for rain area, between rain and no-rain areas, and for no-rain area. The last three terms reflect the nature of intermittency, and depend on various factors such as geometry of rain area, fractional coverage of rain area, and rainfall distribution.

In indicator cokriging, if all the joint probability terms in Eq.(11b) can be estimated with accuracy, taking intermittency into account only amounts to adding a cutoff at zero rainfall, and hence adds no difficulty. In our problem, however, substantial additional data and effort are needed in structure identification to handle intermittency. The reason is that, unlike the non-intermittency term of $\Pr[0<z(v)\le z_1, 0<z(u)\le z_c]$, intermittency terms such as $\Pr[z(v)=0, 0<z(u)\le z_c]$ and $\Pr[z(v)=0, z(u)=0]$ are functions of fractional coverage of rain area, which is highly variable in time. The implications are that fractional coverage has to be estimated in real-time, and that the climatological estimates of the intermittency terms have to be obtained over a range of values of fractional coverage (ideally from zero to 100 percent), or parameterized in terms of fractional coverage. For a direct estimation of intermittency terms, one may adopt an approach similar to Barancourt et al. (1992), in which rainfall fields are stratified with respect to the fractional coverage for an ensemble characterization of rainfall intermittency. On the parameterization approach, work is under way and the results will be reported in the near future. In this work, for the sake of simplicity, we assume that it rains everywhere in the estimation domain (by choosing it to include only the rain area), and thus all the intermittency terms vanish. Hence, in the developments to follow, it is to be understood that by indicator covariance we in fact mean indicator covariance conditional on occurrence of rain.

Data Used

Data in the form of rain gage measurements suitable for covariance estimation were unavailable as typical operational rain gage networks are not dense enough. As an alternative, we used radar rainfall data to estimate ensemble statistics of indicator covariance. They are from the Korea Meteorological Administration (KMA)'s weather radar (DWSR-88C by EEC) at Mt. Kwanak (629 m above MSL) near Seoul, one of the network of five covering the southern Korean peninsula and vicinity. The radar has a wavelength of 5.6 cm and beamwidths of less than 1.2 degrees. At the base-scan (elevation angle of 0.5°), digital dBZ data were available on a 4x4 km grid, and not in the polar form, at every 5 minutes. The data cover July through September in 1989, and include significant rainfall events from warm fronts, extratropical cyclones and typhoons (see Kim et al., 1989, for detailed meteorological descriptions).

Because indicator cokriging utilizes bivariate probability distribution, it is important that the radar rainfall data possess the essential distributional characteristics of rainfall as observed by rain gages on the ground. For this, we used a simplified version of the range-dependent climatological Z-R conversion (Calheiros and Zawadzki 1987). The conversion function fcn( ) in R=fcn(Z), where Z is the raw radar reflectivity factor in dBZ and R is the radar rainfall, satisfies:

$$\int_R^\infty f^i_{zg}(u)\, du = \int_Z^\infty f^i_{zr}(v)\, dv \tag{12}$$

where $f^i_{zg}(\ )$ and $f^i_{zr}(\ )$ are the experimental probability density functions of 10-minute gage rainfall and radar reflectivity factor of spatial resolution of 4x4 km^2, respectively, for the i-th annulus of the range stratification (see Seo et al., 1992, for details).

Radar rainfall data thus obtained preserve ensemble univariate cdf of 10-minute gage rainfall at every range ring. They are then integrated to yield hourly rainfall data, and verified against hourly rain gage measurements from the KMA's operational rain gage network. It was found that the hourly radar rainfall data thus obtained have little bias in the ensemble mean and are reasonably well-correlated with rain gage data (ensemble cross-correlation coefficient ranging from 0.50 to 0.61) between the ranges of 60 and 240 km, (see Seo et al., 1992, for details). It is noted here that the magnitude of the cross-correlation is comparable to that of RADAP II data from a WSR-57 radar (2.2° beamwidth, 10 cm wavelength) at Oklahoma City (Seo and Smith 1991). In the following, we present examples of ensemble indicator correlation as obtained from the hourly radar rainfall data, and describe how they may be parameterized for application of indicator cokriging in rainfall estimation.

Indicator Covariance Structure of Radar Rainfall

Figures 1 and 2 show examples of the conditional probability as a function of the lag distance h, $Pr[z(u+h) \leq z_l | z(u) \leq z_c]$, at different cutoff combinations as estimated from hourly radar rainfall data. Also shown in the figures is the weighted least-squares fit of the exponential model in the following form:

$$\Pr[z(u+h) \le z_1 | z(u) \le z_c]$$
$$= \{ 1 - n(z_1, z_c) - \Pr[z(u) \le z_1] \} e^{-h/L(z1,zc)} + \Pr[z(u) \le z_1] \quad \text{for } z_1 \ge z_c \tag{13}$$

where h is the Euclidian distance $|u-v|$, $L(z_1, z_c)$ is the correlation distance, and $n(z_1, z_c)$ is the nugget effect term. The corresponding indicator covariance is written as:

$$\mathrm{Cov}[I(u+h;z_1), I(u;z_c)]$$
$$= \Pr[z(u) \le z_c] \{ 1 - n(z_1, z_c) - \Pr[z(u) \le z_1] \} e^{-h/L(z1,zc)} \quad \text{for } z_1 \ge z_c \tag{14}$$

Hence, for the exponential model, indicator covariance is completely specified by univariate cdf, nugget effect and correlation distance. The integral scale of indicator correlation is given by:

$$L_c(z_1, z_c) = \rho_{0+}(z_1, z_c) \, L(z_1, z_c) \tag{15}$$

In Eq.(15), $\rho_{0+}(z_1, z_c)$ is the lag-0 correlation coefficient, which can be written in terms of the nugget effect as follows:

$$\rho_{0+}(z_1, z_c)$$
$$\equiv \lim_{h \to 0} \mathrm{Cor}[I(u, z_1), I(u+h, z_c)] \tag{16a}$$
$$= \Pr[z(u) \le z_c] \{ 1 - n(z_1, z_c) - \Pr[z(u) \le z_1] \} / (\sigma_{I1} \, \sigma_{Ic}) \quad \text{for } z_1 \ge z_c \tag{16b}$$

Figure 3 displays $L_c(z_1, z_c)$ as a function of cutoffs as obtained from the hourly radar rainfall data. The indicator correlation scale is at the maximum of 18.5 km near the median rainfall depth of 2 mm (i.e., $z_1=z_c=2$ mm), and decreases as either cutoff deviates from the median (destructuration effect, Journel and Posa 1990). Small and steadily decreasing correlation scale at higher cutoffs indicates lack of spatial predictability in large rainfall amounts. It is well known that rainfall distribution is approximately lognormal, and Figure 3 is indeed similar to that of a bivariate lognormal variable with comparable statistics, as obtainable from theoretical indicator covariances of a bivariate standard normal variable (Journel and Posa 1990). The most notable difference is the presence of a significant nugget effect in the radar rainfall data (see Figures 1 and 2), which can be attributed, at least partly, to various sources of error (of random nature in particular) associated with radar observation of rainfall and post-processing.

It is not readily clear how the indicator correlation scale of ground-truth rainfall might differ from Figure 3. When both rain gage measurements and radar rainfall data are used in estimation, it is often assumed that radar rainfall is a linear sum of gage rainfall and a random error (Creutin et al. 1988, Seo and Smith 1991). In Appendix, we follow a similar approach to examine how indicator correlation scale of ground-truth may differ from that of radar rainfall.

5. IMPLEMENTATION OF INDICATOR COKRIGING

Because of the computational burden of solving K linear systems of dimension (N+1)xM, where K is the number of discretization points used in numerical integration, and N and M are the numbers of surrounding measurements and cutoffs used in estimation of conditional probabilities, respectively, only a small number of cutoffs may be used in actual implementation. Given that, selecting cutoffs bears great importance as it pertains to 1) how closely the set of indicator covariances sampled at the limited number of cutoffs approximates the information content of a full bivariate distribution, and 2) how accurately the numerical integration evaluates the conditional expectation.

Literature reports an application of indicator cokriging using nine cutoffs (Suro-Perez and Journel 1991) for comparison against indicator principal component kriging (see Suro-Perez and Journel, 1991, for its description, and Lajaunie, 1992, for possible theoretical inconsistencies). In that application, apparently the same set of cutoffs used in estimation of conditional probabilities is used again for numerical integration of conditional expectation. Such a practice, however, is seen to suffer from the conflicting problem that the cutoffs chosen for the estimation of conditional probabilities cannot in general be suitable for the numerical integration, and vice versa.

From the viewpoint of estimation of conditional probabilities, we would like our choice of cutoffs to minimize the estimation variance $Var[i^*(u_0;z_{ck})]$ when cokriging for $i^*(u_0;z_{ck})$, preferably at all cutoffs z_{ck},

$k=1,\ldots,K$. From the viewpoint of numerical integration, on the other hand, the choice of cutoffs (i.e., discretization points) depends strictly on the shape of the conditional cdf being estimated. In light of these observations, we took the approach in this work to use as many discretization points in numerical integration as the desired accuracy calls for, but to use only a very small number of optimally selected (in the sense that estimation variance is minimized) cutoffs for estimation of conditional probabilities.

Choosing Cutoffs for Estimation of Conditional Probabilities

In cokriging for $i^*(u_0;z_{ck})$ using a total of M cutoffs, where z_{ck} belongs to the set of K discretization points used in numerical integration, the problem of finding optimal cutoffs may be stated as follows: find M-1 cutoffs of the auxiliary indicator variables, which, along with the cutoff z_{ck} of the main indicator variable, minimize the estimation variance associated with $i^*(u_0;z_{ck})$.

The optimal auxiliary cutoffs depend in general on the particular spatial configuration of the surrounding measurements in reference to the point of estimation, the indicator covariance structure, the univariate cdf, and the magnitude of the main cutoff, z_{ck}. In this work, we chose M=3 (i.e., two auxiliary cutoffs plus the main cutoff), based on a numerical experiment indicating that increasing the number of cutoffs brings only a minor additional reduction in estimation variance.

The two optimal auxiliary cutoffs can be found by actually indicator-cokriging with various pairwise combinations of cutoffs, and locating the minimum for each main cutoff z_{ck}. Numerical experiments indicate that, given the same bivariate probabilistic structure, the optimal auxiliary cutoffs, when expressed in fractions of the main cutoff z_{ck}, show little dependence on z_{ck}. For the cases examined in this work, typical optimal auxiliary cutoffs are about 0.6 z_{ck} and 1.5 z_{ck}, for all $k=1,\ldots,K$. They are in accordance with the intuition that deploying both a smaller and a larger cutoffs makes a better sampling strategy than having either two smaller or larger ones. Also, as one might expect, two auxiliary cutoffs shift farther apart for a sparser network, and closer for a denser one.

Because the discretization points in numerical integration, z_{ck}'s, $k=1,\ldots,K$, can vary depending on the integration method used and the accuracy desired, the approach described above requires evaluation of indicator covariance at arbitrary cutoffs. This, however, can be easily done by interpolating the nugget effect, $n(z_1,z_c)$, and the correlation distance, $L(z_1,z_c)$, on contour surfaces similar to Figure 3.

Numerical experiments also indicate that optimal auxiliary cutoffs resulting from the above two-dimensional minimization can also be obtained from a one-dimensional version of indicator-cokriging with only one auxiliary indicator variable (i.e., M=2). In the latter case, the two cutoffs yielding the two smallest estimation variances constitute the optimal set of auxiliary cutoffs. In this work, the one-dimension minimization is performed, with the main cutoff set equal to the ensemble median, for a dozen or so points to obtain individual optimal multiplicative factors, which are then averaged to yield average factors. Note that this adds only a minimal amount of computation of solving a linear system of dimension (N+1)x2 a dozen or so times.

6. VALIDATION

In order to evaluate the indicator cokriging approach using real-world data via, e.g., cross-validation, the data have to come from a rain gage network that is sufficiently dense so that, relative to the correlation scale of rainfall, surrounding measurements used in point estimation are not too distant. Also, the rain gage measurements have to be of good resolution and quality in order to evaluate performance at both ends of the rainfall distribution. Unfortunately, no satisfactory network-measurement combinations were available, and, as an alternative, we used hourly radar rainfall data as ground-truth rainfall fields, from which pseudo-rain gage measurements were generated by sampling at random locations. Due to this artificiality, however, performance evaluation was limited to only a few cases, and thus results are only preliminary.

Description of the Comparative Evaluation

Three approaches using indicator cokriging (ICK), indicator kriging (IK), and ordinary kriging (OK) are compared. From the pool of hourly radar rainfall fields of 1989, we selected two fields (Hour 15, July 26, 1989, and Hour 15, August 22, 1989), for which the radar umbrella captures most of the areal extent of the rainfall system. They are then assumed to be ground-truth rainfall fields. Each field represents an area of $384x448\ km^2$, each radar rainfall datum representing a $4x4km^2$ bin. Within each field, 330 gage locations ($1/520km^2$) were randomly selected. In Korea, for example, this is between the density of telemetry networks for real-time flood forecasting in selected river basins, and the density of the current nationwide meteorological gage network. Then, for each field, point estimation was performed at the center of each bin over the entire rain area using 10 nearest neighbors. In all estimation approaches, ensemble statistics obtained from the radar rainfall data of 1989 were used. As intermittency was not considered, rain gage measurements reporting no rainfall were excluded from estimation.

Results

Table 1 shows statistics of the ground-truth and the estimates. The most notable difference is with the variance of estimates. In terms of percentage of the variance of ground-truth explained by estimates, ICK brings about 7 to 8 percent improvement over OK, whereas improvement by IK over OK is only about 1 to 2 percent. Figures 4 through 6 show scatter-plots of ICK-, IK- and OK-estimates versus ground-truth, respectively, for Case 1. All three approaches similarly underestimate large rainfall amounts. The most notable difference is at small amounts, where ICK-estimates are clearly better than OK-estimates. Similar features are observed for Case 2 as well. Figures 7 through 10 show reductions in absolute mean error (AME) and mean square error (MSE) by ICK and IK over OK for various rainfall amounts. Levels 1 through 10 in the x-axes correspond to hourly rainfall amounts in mm of 1) 0.0 ~ 0.4, 2) 0.4 ~ 0.8, 3) 0.8 ~ 1.4, 4) 1.4 ~ 2.1, 5) 2.1 ~ 3.0, 6) 3.0 ~ 4.0, 7) 4.0 ~ 5.3, 8) 5.3 ~ 7.0, 9) 7.0 ~ 10.0, and 10) 10.0 ~ 35.8, respectively. For both ICK and IK, improvements are most significant for light rainfall. For heavy rainfall, however, only ICK is seen to provide any improvement. Similar performance characteristics can also be observed in Figures 11 and 12, where empirical cdf's of ground-truth, ICK- and OK-estimates are shown together. In each figure, the cdf of ICK-estimates is closer to that of assumed ground-truth. The cdf's of IK-estimates (not shown) are very similar to those of OK-estimates.

In Korea, for example, operational rainfall-runoff models used for real-time flood forecasting typically take estimates of mean areal rainfall (MAR) over sub-basin areas of about 350 km^2. To examine relative performance in MAR estimation at this scale, MAR estimates over subareas of 16x16 km^2 are compared with ground-truth. Figures 13 and 14 show reductions for Case 2, by ICK and IK over OK, respectively, in absolute error in MAR estimates and in mean square error in point estimates within each subarea. In each figure, the reference point of zero reductions is marked by a cross. It is seen that ICK clearly outperforms OK with a reduction in MAR error as high as 1.8 mm for hourly rainfall. IK, on the other hand, is seen to bring little improvement over OK.

7. CONCLUSIONS AND FUTURE RESEARCH RECOMMENDATIONS

Indicator cokriging is used to estimate spatial distribution of rainfall from rain gage measurements. The proposed approach uses ensemble statistics exclusively in order to avoid real-time estimation of time-varying statistics of rainfall, and uses only a few, but optimally selected, auxiliary cutoffs to reduce computational requirements. A comparative evaluation indicates that the indicator cokriging approach is clearly superior to its ordinary kriging counterpart. The improvement is most significant for light rainfall, but drops off significantly for heavier rainfall. Due to the use of radar rainfall data in place of rain gage measurements in covariance estimation and validation, however, the results are only preliminary.

The lack of predictability in spatial estimation of heavy rainfall is borne out in the integral scale of indicator correlation, as obtained from hourly radar rainfall data: peaking to its maximum for cutoffs near the median, indicator correlation scale becomes increasingly smaller for larger cutoffs of rainfall depth. The indicator correlation scale also exhibits similarity to that of a bivarirate lognormal variable, which reflects the diffusional nature of spatial rainfall process. Hence, as evidenced in this work, indicator kriging may not be considered as an alternative to indicator cokriging in rainfall estimation.

A derived-distribution analysis based on linearity between radar rainfall and ground-truth indicates that indicator correlation scale of ground-truth rainfall can be significantly greater than that of radar rainfall at low cutoffs, and points toward inclusion of rainfall intermittency, for example, following the approach proposed in this work.

ACKNOWLEDGEMENTS

This paper is based on works supported in parts by the Korea Ministry of Science and Technology, the Korea Meteorological Administration, by the Korea Water Resources Corporation, and NOAA (Climate and Global Change Research Program, Grant NA36690419). These supports are gratefully acknowledged. The author would also like to thank the Water Resources Division of the Korea Institute of Construction Technology for providing with rain gage data, and an anonymous WRR reviewer for pointing out an inaccurate statement in the manuscript.

APPENDIX

Here, we examine how a random error introduced in the rainfall observation process might affect indicator covariance structure of rainfall data. Let us assume the following observation equation:

$$z(v) = \beta\, x(v) + \varepsilon(v) \qquad \text{(A1)}$$

where $z(v)$ is the observed rainfall at location v, $x(v)$ is the true rainfall, $\varepsilon(v)$ is the random error, and β is a known constant representing a systematic bias. Denoting the experimental values of the random variables, $z(v)$, $x(v)$ and $\varepsilon(v)$ as z_v, x_v and e_v, respectively, we have $e_v \geq -\beta x_v$ due to the nonnegativity of z_v. The conditional probability, $Pr[z(v) \leq z_c | x(u) \leq x_c]$, where z_c and x_c are arbitrary cutoffs for observed and true rainfall, respectively, is given by:

$$Pr[z(v) \leq z_c | x(u) \leq x_c]$$

$$= \int_0^\infty \int_{-\beta x_v}^{z_c - \beta x_v} f_{\varepsilon(v), x(v)|} (e_v, x_v | x(u) \leq x_c) \, de_v \, dx_v \tag{A2a}$$

$$= \int_0^\infty \int_{-\beta x_v}^{z_c - \beta x_v} f_{\varepsilon(v)|x(v)}(e_v | x_v) \, de_v \frac{f_{x(v)}(x_v)}{Pr[x(u) \leq x_c]} \, dx_v \qquad \text{for } x(u) \leq x_c \tag{A2b}$$

where $f_{\varepsilon(v), x(v)|\bullet}(e_v, x_v | x(u) \leq x_c)$ is the conditional joint pdf of $\varepsilon(v)$ and $x(v)$, and $f_{x(v)}(x_v)$ is the marginal pdf of $x(v)$. Assuming that $\varepsilon(v)$ is independent of $\{x(u) \leq x_c\}$ for $u \neq v$, we may write:

$$Pr[z(v) \leq z_c | x(u) \leq x_c]$$

$$= \int_0^\infty \phi(-\beta x_v, z_c - \beta x_v) \, f_{x(v)|\bullet}(x_v | x(u) \leq x_c) \, dx_v \tag{A3a}$$

$$= \int_0^\infty -\phi'(-\beta x_v, z_c - \beta x_v) \, Pr[x(v) \leq x_v | x(u) \leq x_c] \, dx_v \tag{A3b}$$

where

$$\phi(-\beta x_v, z_c - \beta x_v) \equiv \int_{-\beta x_v}^{z_c - \beta x_v} f_{\varepsilon(v)|x(v)}(e_v | x_v) \, de_v \tag{A4}$$

$$\phi'(-\beta x_v, z_c - \beta x_v) \equiv d\phi(-\beta x_v, z_c - \beta x_v)/dx_v \tag{A5}$$

Analogously, $Pr[z(u) \leq z_c]$ is given by:

$$Pr[z(u) \leq z_c]$$

$$= \int_0^{\infty} -\phi'(-\beta x_v, z_c - \beta x_v)\ Pr[x(v) \leq x_v]\ dx_v \tag{A7}$$

To obtain $Pr[z(v) \leq z_1 | z(u) \leq z_c]$, we first write:

$$Pr[z(v) \leq z_1 | z(u) \leq z_c]$$

$$= \int_0^{\infty} \int_{-\beta x_v}^{z_1 - \beta x_v} f_{\varepsilon(v), x(v) | \bullet}(e_v, x_v | z(u) \leq z_c)\ de_v\ dx_v \tag{A8}$$

$$= \int_0^{\infty} \int_{-\beta x_v}^{z_1 - \beta x_v} f_{\varepsilon(v) | x(v)}(e_v | x_v)\ de_v\ \frac{f_{x(v)}(x_v)}{Pr[z(u) \leq z_c]}\ dx_v \qquad \text{for } z(u) \leq z_c \tag{A9}$$

Assuming $\varepsilon(v)$ is independent of $\{z(u) \leq z_c\}$ for $u \neq v$ (i.e., $\varepsilon(v)$ is now a white-noise process in space), we may write:

$$Pr[z(v) \leq z_1 | z(u) \leq z_c]$$

$$= \int_0^{\infty} \phi(-\beta x_v, z_1 - \beta x_v)\ f_{x(v) | \bullet}(x_v | z(u) \leq z_c)\ dx_v \tag{A10}$$

$$= \int_0^{\infty} -\phi'(-\beta x_v, z_1 - \beta x_v)\ Pr[x(v) \leq x_v | z(u) \leq z_c]\ dx_v \tag{A11}$$

where $\phi(-\beta x_v, z_1 - \beta x_v)$ and $\phi'(-\beta x_v, z_1 - \beta x_v)$ are as defined in Eqs.(A4) and (A5). If $u = v$, we have by definition:

$$\Pr[z(u) \le z_1 | z(u) \le z_c] = \begin{cases} 1 & \text{if } z_1 \ge z_c \\ \dfrac{\Pr[z(u) \le z_1]}{\Pr[z(u) \le z_c]} & \text{if } z_1 < z_c \end{cases} \tag{A13}$$

Given the bounded nature of the random error (i.e., $e_v \ge -\beta x_v$), a reasonable candidate for $f_{\varepsilon(v)|zr(v)}(e_v|z_{rv})$ is the truncated normal:

$$f_{\varepsilon(v)|zr(v)}(e_v|z_{rv})$$

$$= \frac{1}{(2\pi)^{1/2}\,\sigma\{1-\phi(-\beta z_{rv})\}} \exp\{-(e_v-m)^2/2\sigma^2\} \tag{A14}$$

In Eq.(A14), m and σ^2 are the mean and variance of the untruncated normal variable, and $\phi(-\beta z_{rv})$ is its cumulative probability defined as:

$$\phi(x) = \int_{-\infty}^{x} \frac{1}{(2\pi)^{1/2}\,\sigma} \exp\{(x-m)^2/2\sigma^2\}\, dx \tag{A15}$$

Then, $\phi(-\beta x_v, z_1-\beta x_v)$ in Eq.(A10) is given by:

$$\phi(-\beta x_v, z_1-\beta x_v) = \frac{\phi(z_1-\beta x_v) - \phi(-\beta x_v)}{1 - \phi(-\beta x_v)} \tag{A16}$$

Using the Leibnitz rule, $-\phi'(-\beta x_v, z_1-\beta x_v)$ in Eq.(A11) is given by:

$$-\phi'(-\beta x_v, z_1-\beta x_v)$$

$$= \frac{\beta}{(2\pi)^{1/2}\,\sigma\,\{1-\phi(-\beta x_v)\}} \exp\{-(z_1-\beta x_v-m)^2/2\sigma^2\}$$

$$- \frac{\beta \, \{1-\phi(z_1-\beta x_v)\}}{(2\pi)^{1/2} \, \sigma \, \{1-\phi(-\beta x_v)\}^2} \exp\{-(-\beta x_v - m)^2/2\sigma^2\} \tag{A17}$$

We are now in a position to evaluate $Cov[I(u+h;z_1), I(u;z_c)]$ via numerical integration for arbitrary values of h, z_1 and z_c given the bivariate pdf of the random function x( ). In this work, it is assumed that x( ) has a bivariate lognormal distribution with parameter values that are comparable to ensemble statistics of the hourly radar rainfall data used. Also, parameter values used for the truncated normal density function are based on error statistics of radar rainfall data compared against rain gage measurements.

Figures A1 and A2 show indicator covariances of z( ) and x( ) at selected cutoffs of 0.5 and 2.2 mm when β=1 (i.e., no bias). Note that the random error (truncated normal in this case) significantly reduces the integral scale of indicator correlation at the lower cutoff. We may interpret the linear observation equation in Eq.(A1) as describing the radar observation of rainfall, where now x( ) is the ground-truth, z( ) is the radar rainfall, β is the bias factor, and ε( ) is the random error. The above results suggest that the indicator correlation scale of radar rainfall (Figure 3 in the text) does represent that of ground-truth except at very low cutoffs.

REFERENCES

Alli, M. M., E. A. Nowatzki, and D. E. Myers, Probabilistic analysis of collapsing soil by indicator kriging, Math. Geol., 22(1), 1990.

Azimi-Zonooz, A., W. F. Krajewski, D. S. Bowles, and D.-J. Seo, Spatial rainfall estimation by linear and non-linear co-kriging of radar-rainfall and raingage data, Stoch. Hydrol. Hydraul., 3, 1989.

Barancourt, C., J. D. Creutin, and J. Rivoirard, A method for delineating and estimating rainfall fields, Water Resour. Res., 28(4), 1992.

Bastin, G., B. Lorent, C. Duque, and M. Gevers, Optimal estimation of the average areal rainfall and optimal selection of rain gauge locations, Water Resour. Res., 20(4), 1984.

Calheiros, R., and I. Zawadzki, Reflectivity-rain rate relationships for radar hydrology in Brazil, J. Climate Appl. Meteor., 26, 118-132, 1987.

Chua, S. H., and R. L. Bras, Optimal estimators of mean areal precipitation in regions of orographic influence, J. Hydrol., 15(6), 23-48, 1982.

Creutin, J. D., and C. Obled, Objective analysis and mapping techniques for rainfall field: An objective comparison, Water Resour. Res., 18(2), 413-431, 1982.

Creutin, J. D., G. Delrieu, and T. Lebel, Rain measurements by raingage-radar combination: A geostatistical approach, J. Atmos. Ocean. Tech., 5, 102-115.

Hu, L. Y., Comparing gamma isofactorial disjunctive kriging and indicator kriging for estimating local spatial distributions, Geostatistics, Vol. 1, M. Armstrong (ed.), Kluwer Academic Publishers, 1989.

Journel, A. G., and Huijbregts, C. J., 1978: Mining Geostatistics, Academic Press, San Diego, CA.

Journel, A. G., Nonparametric estimation of spatial distributions, Math. Geol., 15(3), 1983.

Journel, A. G., and D. Posa, Characteristic behavior and order relations for indicator variograms, Math. Geol., 22(8), 1990.

Karlin, S. and H. M. Taylor, A first course in stochastic processes, 2nd Ed., Academic Press, New York, 1975.

Kim, I.-S., and 10 others, Accuracy evaluation of radar rainfall measurements, Res. Report, MR 89-3, Meteor. Res. Inst., Korean Meteor. Admin., Seoul, Korea, 1989.

Lajaunie, Ch., Comment on "Indicator principal component kriging" by V. Suro-Perez and A. G. Journel, Math. Geol., 24(3), 1992.

Lebel, T., G. Bastin, C. Obled, and J. D. Creutin, On the accuracy of areal rainfall estimation: A case study, Water Resour. Res., 22(11), 1987.

Lebel, Th., J. P. Larborde, A geostatistical approach for areal rainfall statistical assessment, Stoch. Hydrol. Hydraul., 2(4), 1988.

Matheron, G., A simple substitute for conditional expectation: Disjunctive kriging, Geostat 1975, Guarascio et al. (Eds.), NATO ASI: Reidel Publ. Co., Dordrecht, Netherlands, 1975.

Schweppe, F. C., Uncertain Dynamic Systems, Prentice-Hall, Englewood Cliffs, NJ, 1973.

Seo, D.-J., W. F. Krajewski, and D. S. Bowles, Stochastic interpolation of rainfall data from rain gages and radar using co-kriging 1. Design of experiments, Water Resour. Res., 26(3), 1990.

Seo, D.-J., and J. A. Smith, Rainfall estimation using raingages and radar - A Bayesian approach: 2. An application, Stoch. Hydrol. Hydraul., 5, 1991.

Seo, D.-J., M.-S. Seo, H.-O., Nam, and S.-J. Choi,, Rainfall estimation using rain gages and radar (I), Res. Report (in Korean with English summary), BS E20910, Systems Eng. Res. Inst., Korea Inst. Sci. Tech., Taejon, Korea, 1992.

Seo, D.-J., and 8 others, A study on real-time rainfall estimation using data from meteorological satellite and radar, Joint Res. Report (in Korean with English summary), WRRI-WR-92-3, Water Resour. Res. Inst., Korea Water Resour. Corp., and Systems Eng. Res. Inst., Korea Inst. Sci. Tech., Taejon, Korea, 1992.

Solow, A. R., Mapping by simple indicator kriging, Math. Geol. 18(3), 1986.

Suro-Perez, V., and A. G. Journel, Indicator principal component kriging, Math. Geol., 23(5), 1991.

Tabios, G. Q., III, and J. D. Salas, A comparative analysis of techniques for spatial interpolation of precipitation, Water Resour. Bull., 21(3), 365-380, 1985.

Young, D. S., Indicator kriging for unit vectors: Rock joint orientations, Math. Geol., 19(6), 1987.

Table 1. Statistics for Ground-Truth and Three Types of Estimates

	Ground-Truth	ICK-Estimate	IK-Estimate	OK-Estimate
Mean (mm)	3.4 (5.3)	3.7 (5.5)	3.6 (5.6)	3.8 (5.7)
Variance (mm^2)	10.6 (26.4)	5.9 (11.3)	5.2 (9.6)	5.1 (9.2)
Maximum (mm)	27.5 (35.8)	12.6 (18.0)	11.6 (17.2)	12.0 (18.1)
ρ_0		0.65 (0.79)	0.62 (0.78)	0.61 (0.78)

Unparenthesized numbers are for Case 1, the parenthesized for Case 2.

ρ_0 is the cross-correlation coefficient with ground-truth.

Numbers of data are 3597 for Case 1 and 3803 for Case 2.

List of Figures and their Captions

Fig. 1. Ensemble experimental conditional probability as a function of spatial lag distance and the exponential-model fit at cutoffs z_1=z_c=2.2 mm.

Fig. 2. Same as Fig. 1 but at cutoffs z_1=7.2 mm and z_c=2.2 mm.

Fig. 3. Indicator correlation scale of hourly radar rainfall displayed as a function of cutoffs.

Fig. 4. Scatter-plot of ICK-estimates vs. ground-truth for Case 1.

Fig. 5. Scatter-plot of IK-estimates vs. ground-truth for Case 1.

Fig. 6. Scatter-plot of OK-estimates vs. ground-truth for Case 1.

Fig. 7. Reductions in absolute mean error by ICK (solid dots) and IK (empty dots) over OK for various rainfall amounts for Case 1.

Fig. 8. Same as Fig. 7, but in mean square error.

Fig. 9. Same as Fig. 7, but for Case 2.

Fig. 10. Same as Fig. 8, but for Case 2.

Fig. 11. Empirical cumulative distribution functions of ground-truth (dotted line), ICK- and OK-estimates (solid lines) for Case 1.

Fig. 12. Same as Figure 11, but for Case 2.

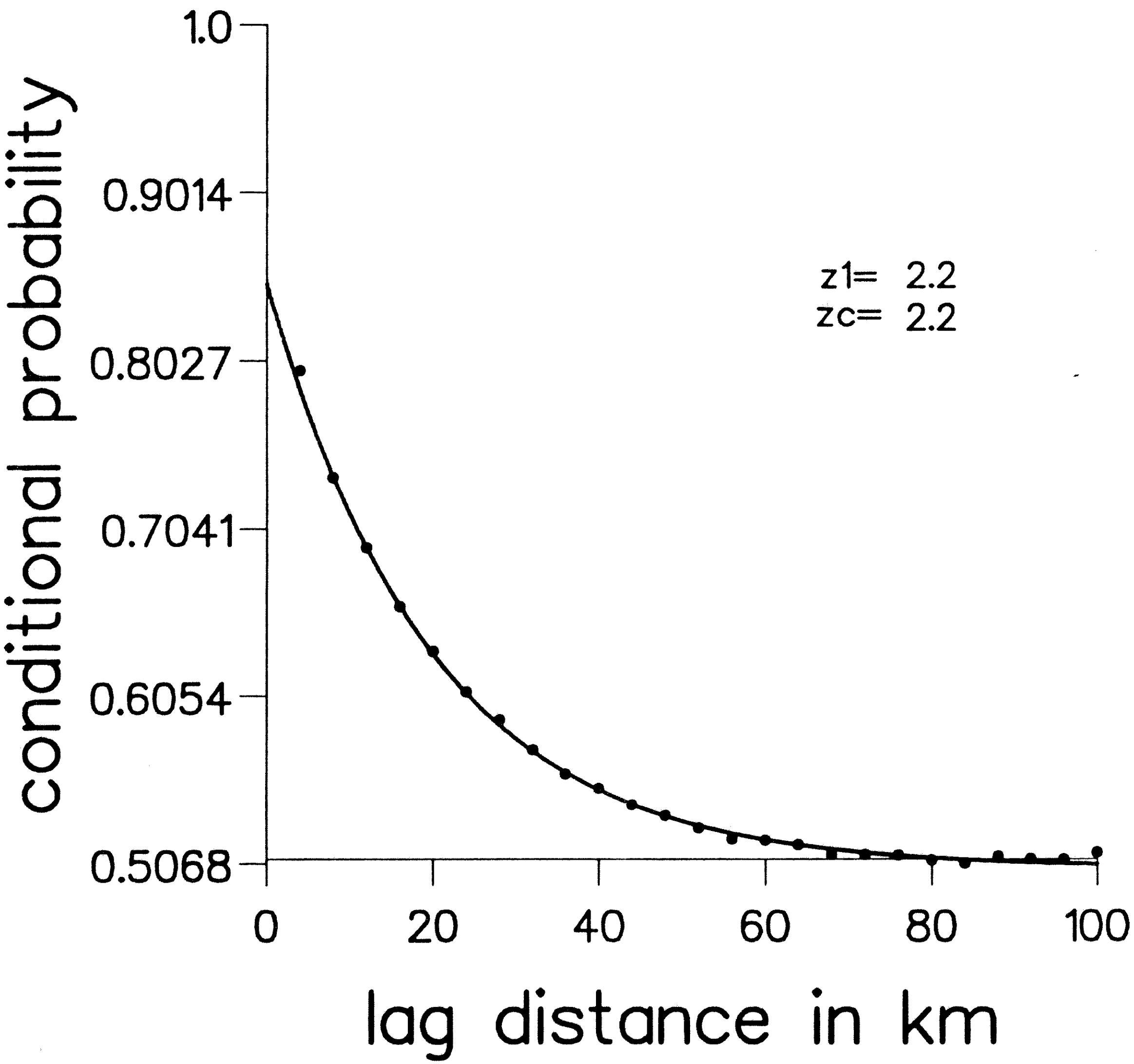
Fig. 13. Reductions in absolute error in MAR estimates and in mean square error in point estimates within each subarea by ICK over OK for Case 2.

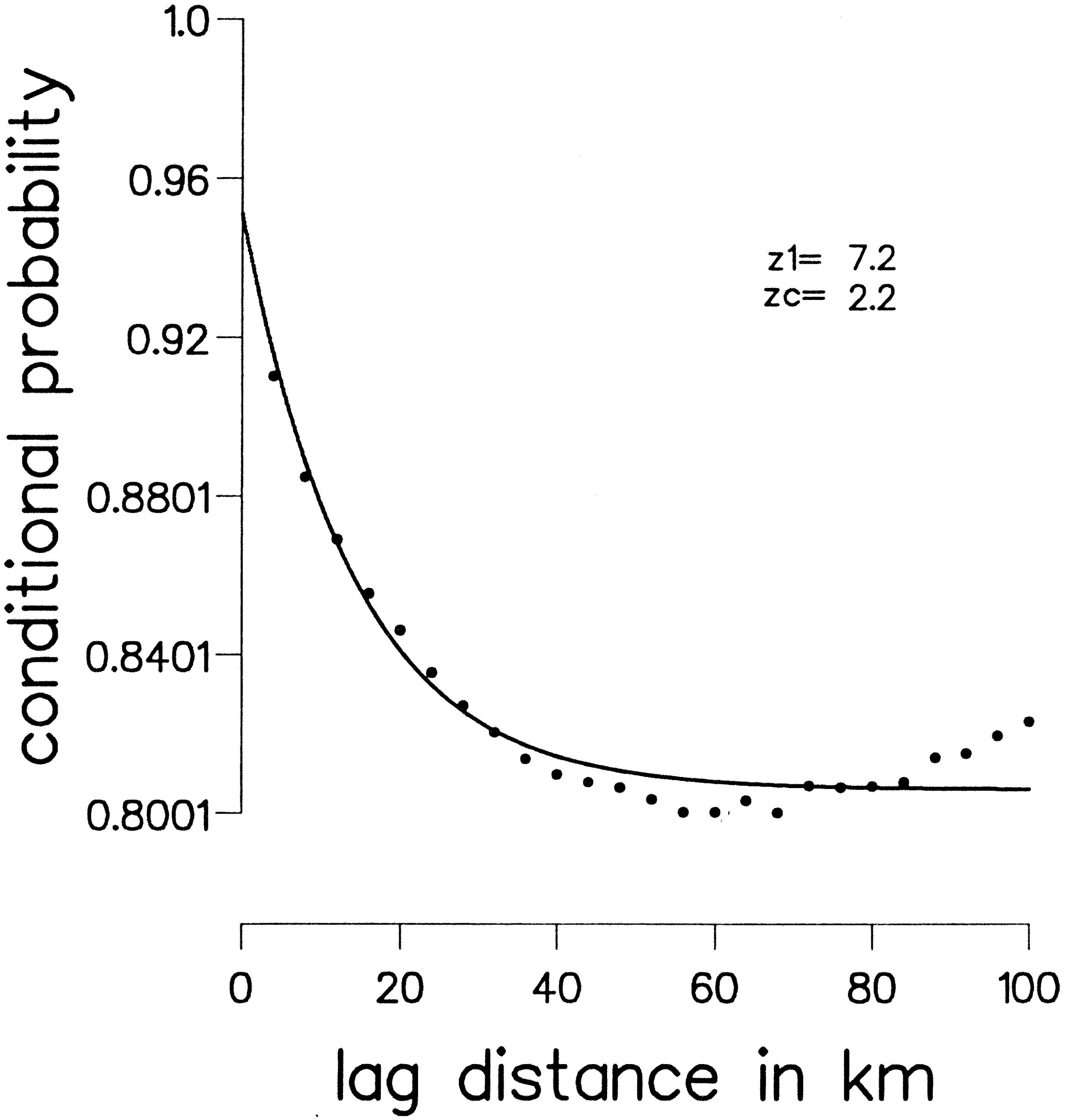
Fig. 14. Same as Fig. 13, but by IK over OK.

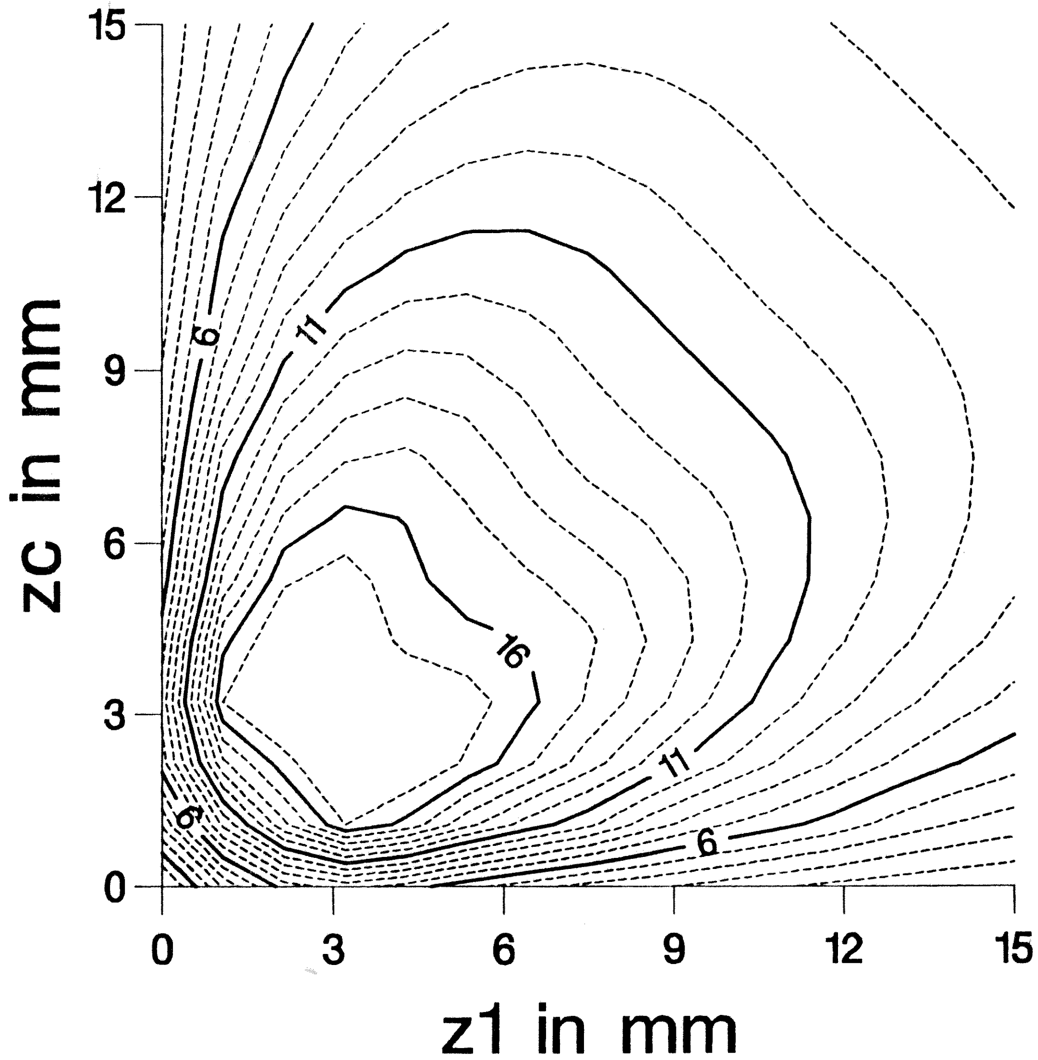
Fig. A1. Indicator covariance of z( ) (empty circles) and x( ) (in solid circles) at cutoffs $z_1=z_c=0.5$ mm.

Fig. A2. Same as Fig. A1, but at cutoffs $z_1=z_c=2.2$ mm.

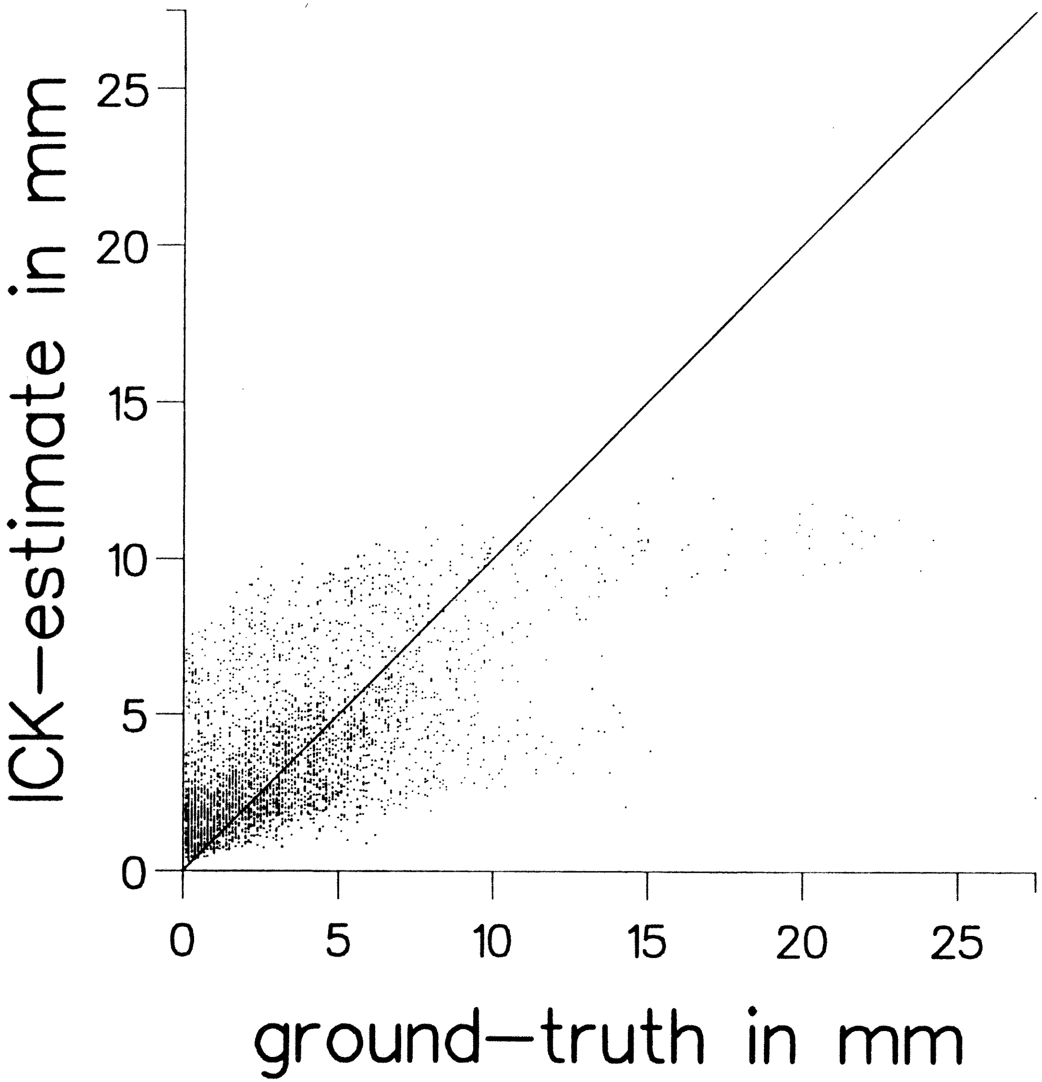
Fig 1

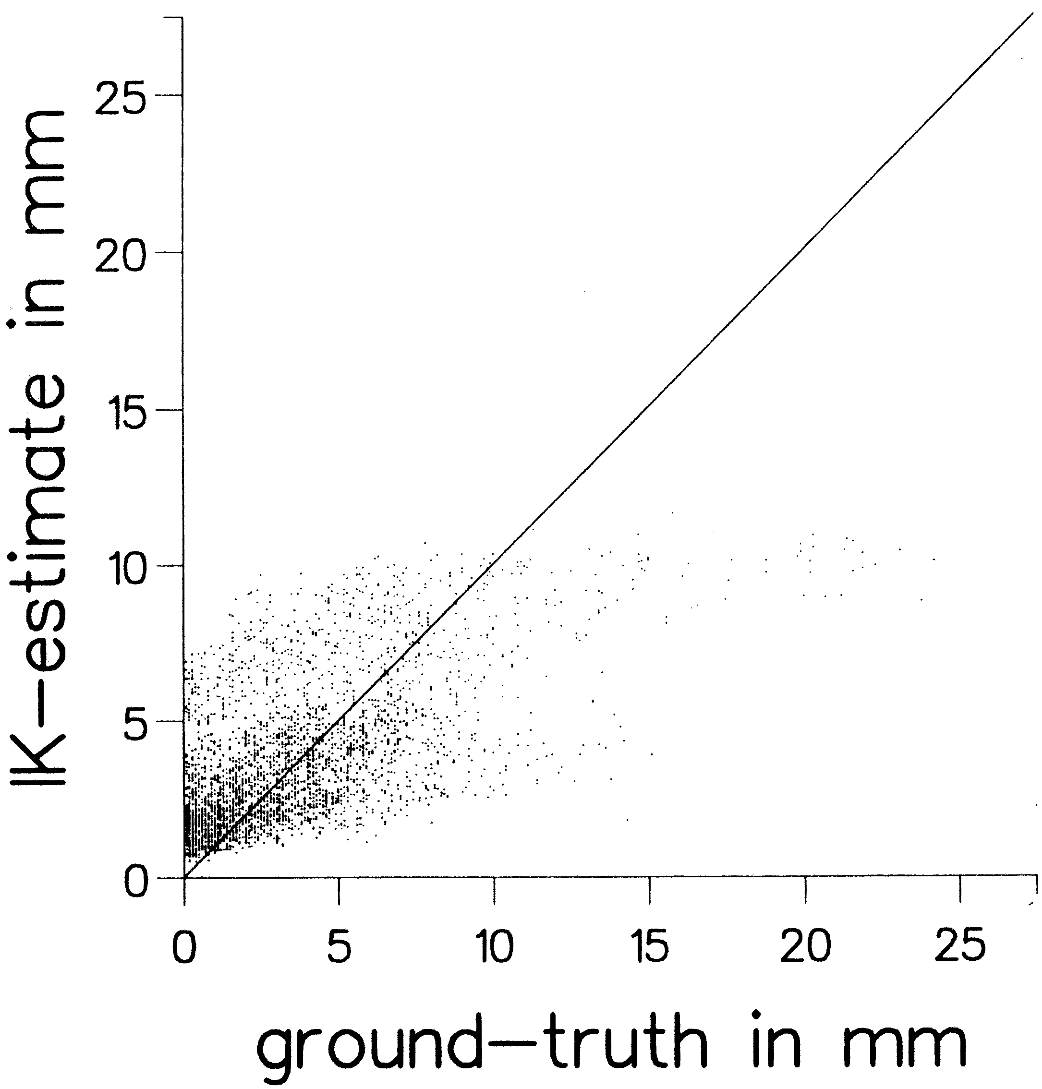

conditional probability
1.0
0.9014
0.8027
0.7041
0.6054
0.5068
0
20
40
60
80
100
lag distance in km
z1= 2.2
zc= 2.2

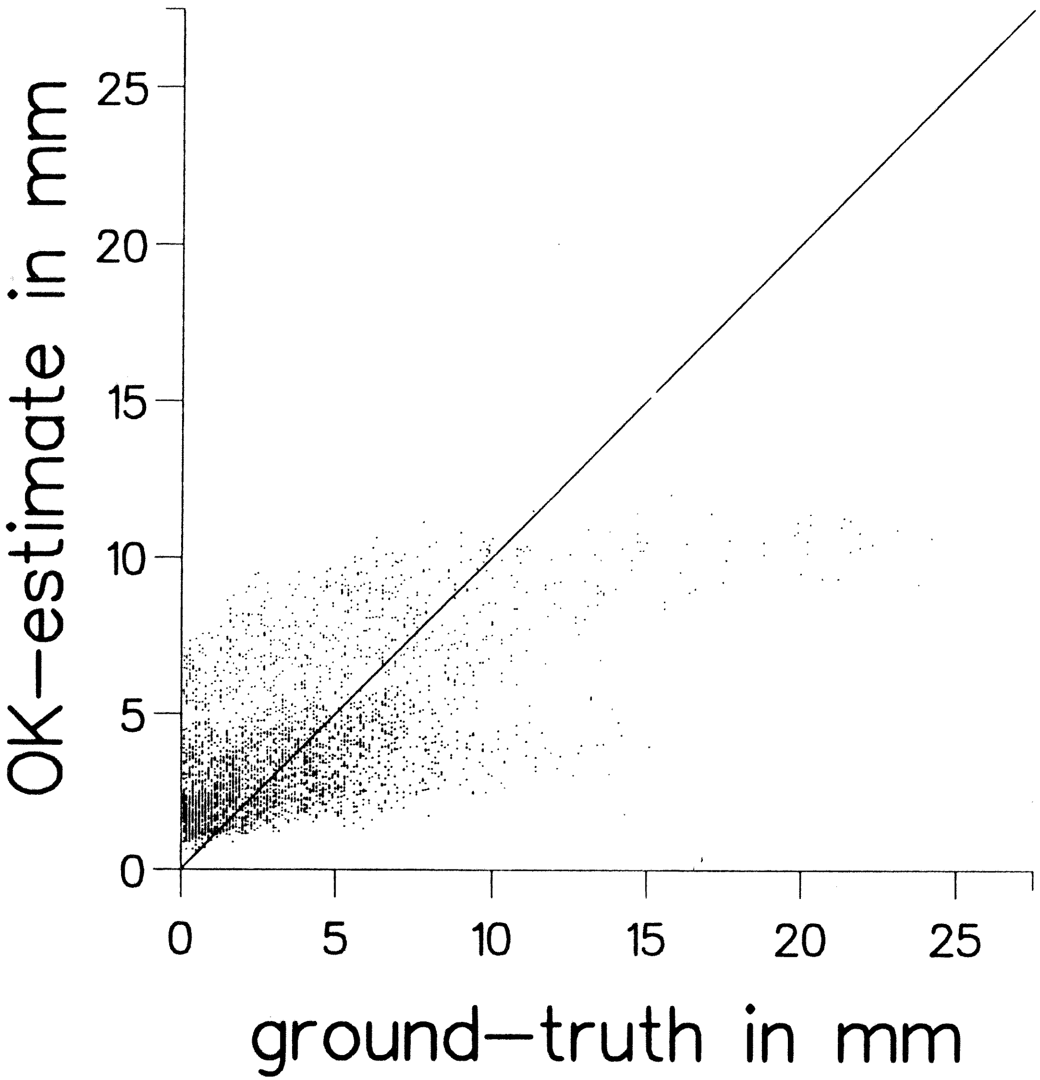

conditional probability
1.0
0.96
0.92
0.8801
0.8401
0.8001
z1= 7.2
zc= 2.2
0
20
40
60
80
100
lag distance in km

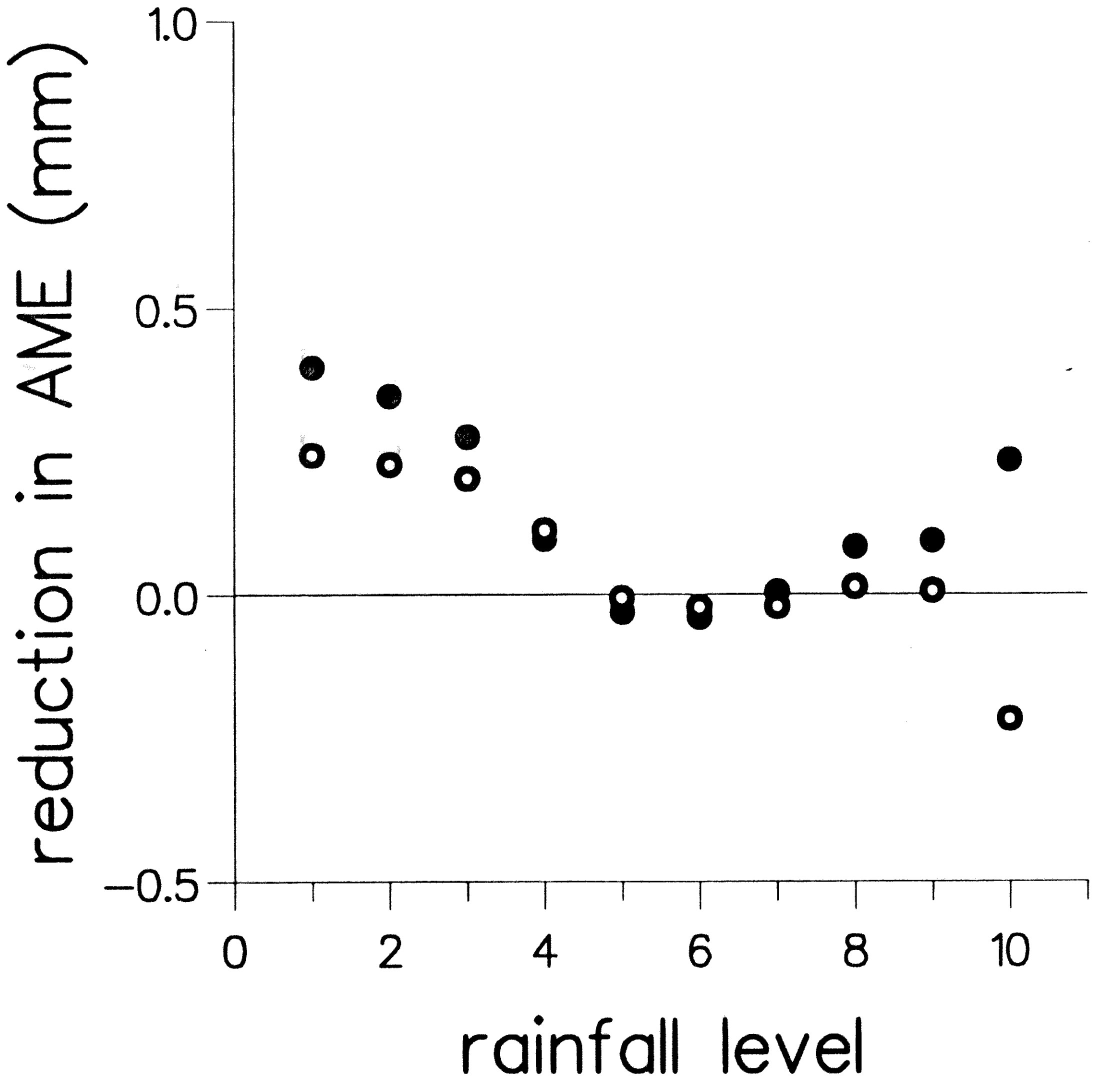


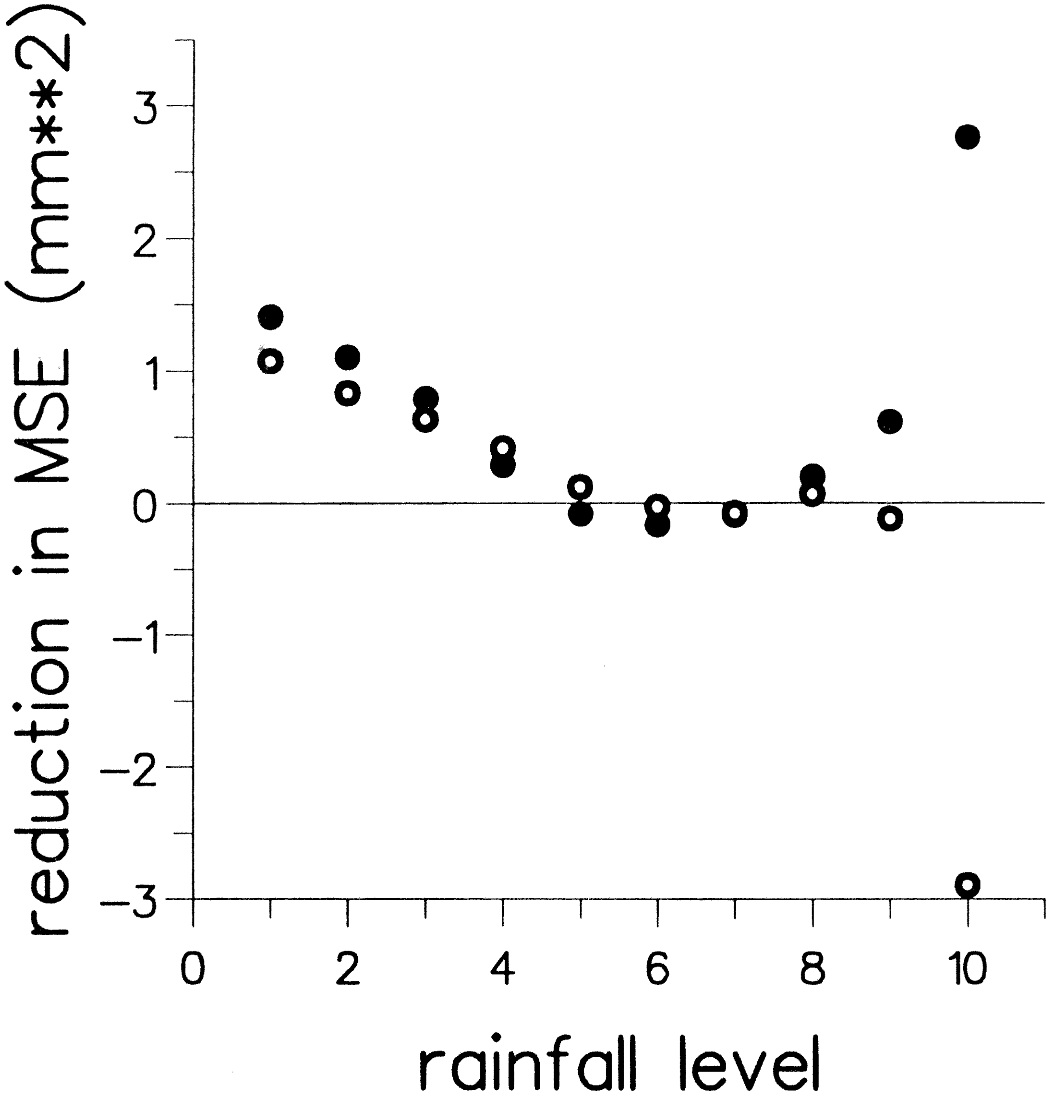
zc in mm
z1 in mm
0
3
6
9
12
15
6
11
16

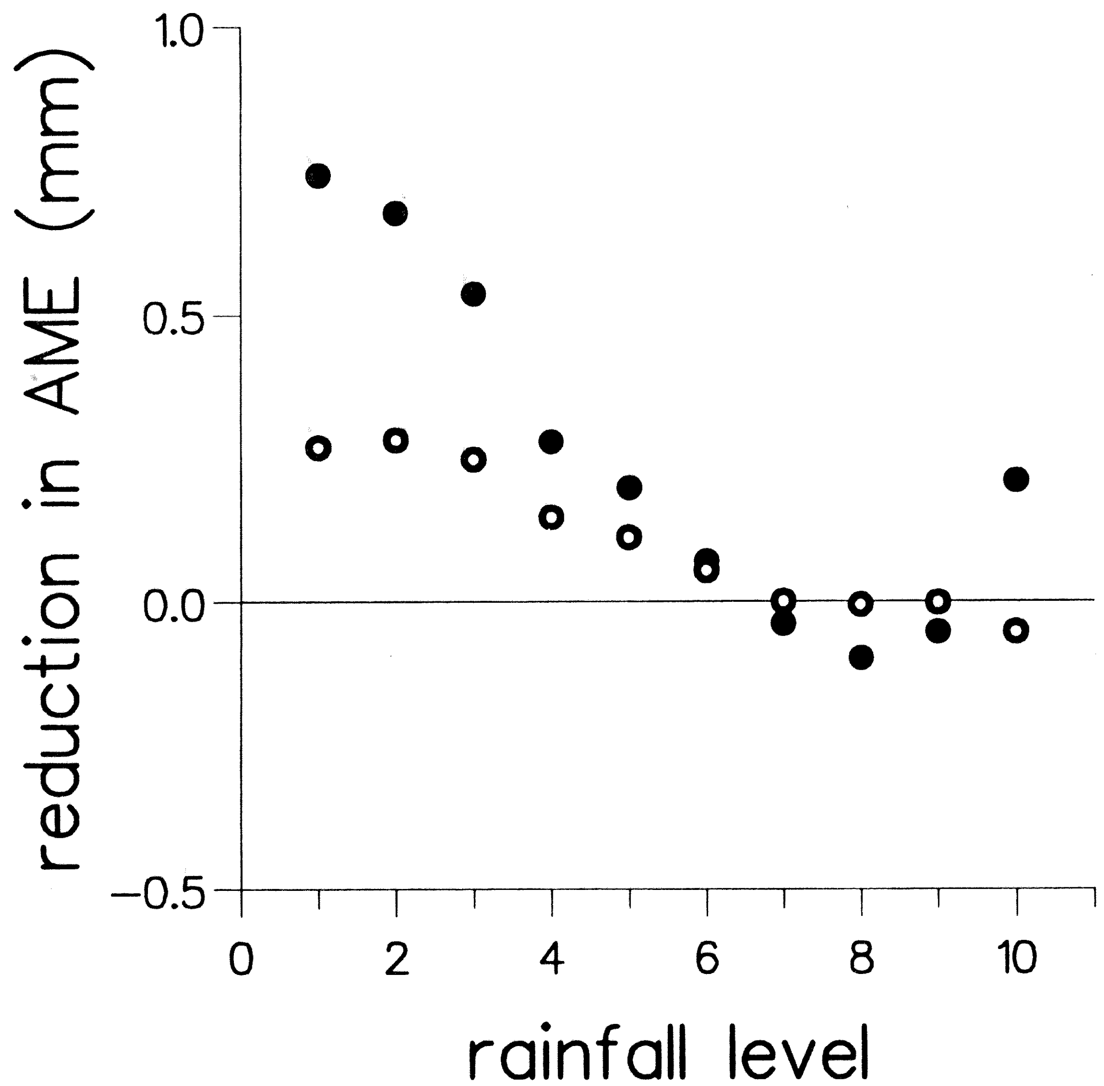

ICK—estimate in mm
ground—truth in mm
0
5
10
15
20
25

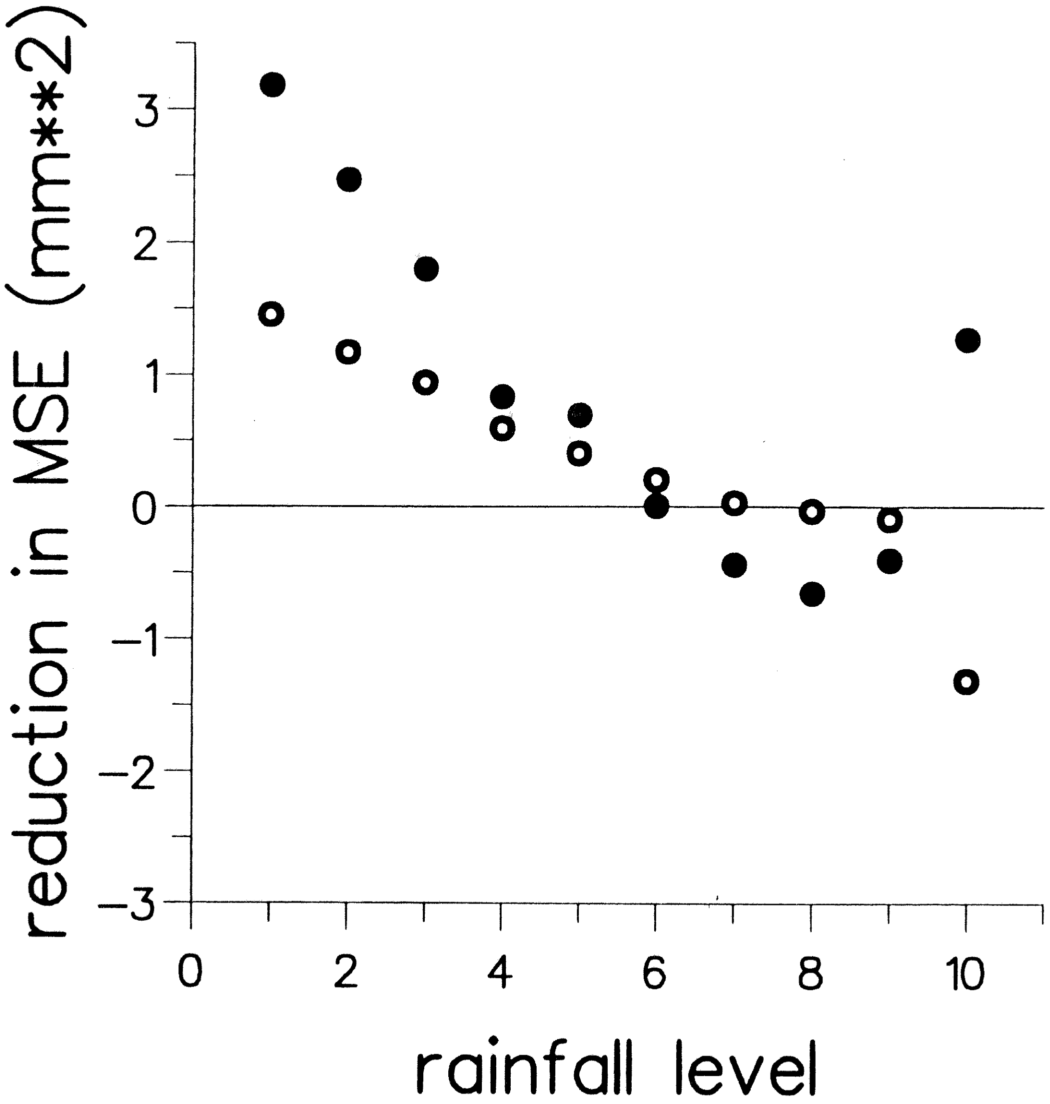

IK-estimate in mm
0
5
10
15
20
25
ground-truth in mm


OK—estimate in mm
0
5
10
15
20
25
ground—truth in mm
0
5
10
15
20
25


reduction in AME (mm)
1.0
0.5
0.0
−0.5
0
2
4
6
8
10
rainfall level

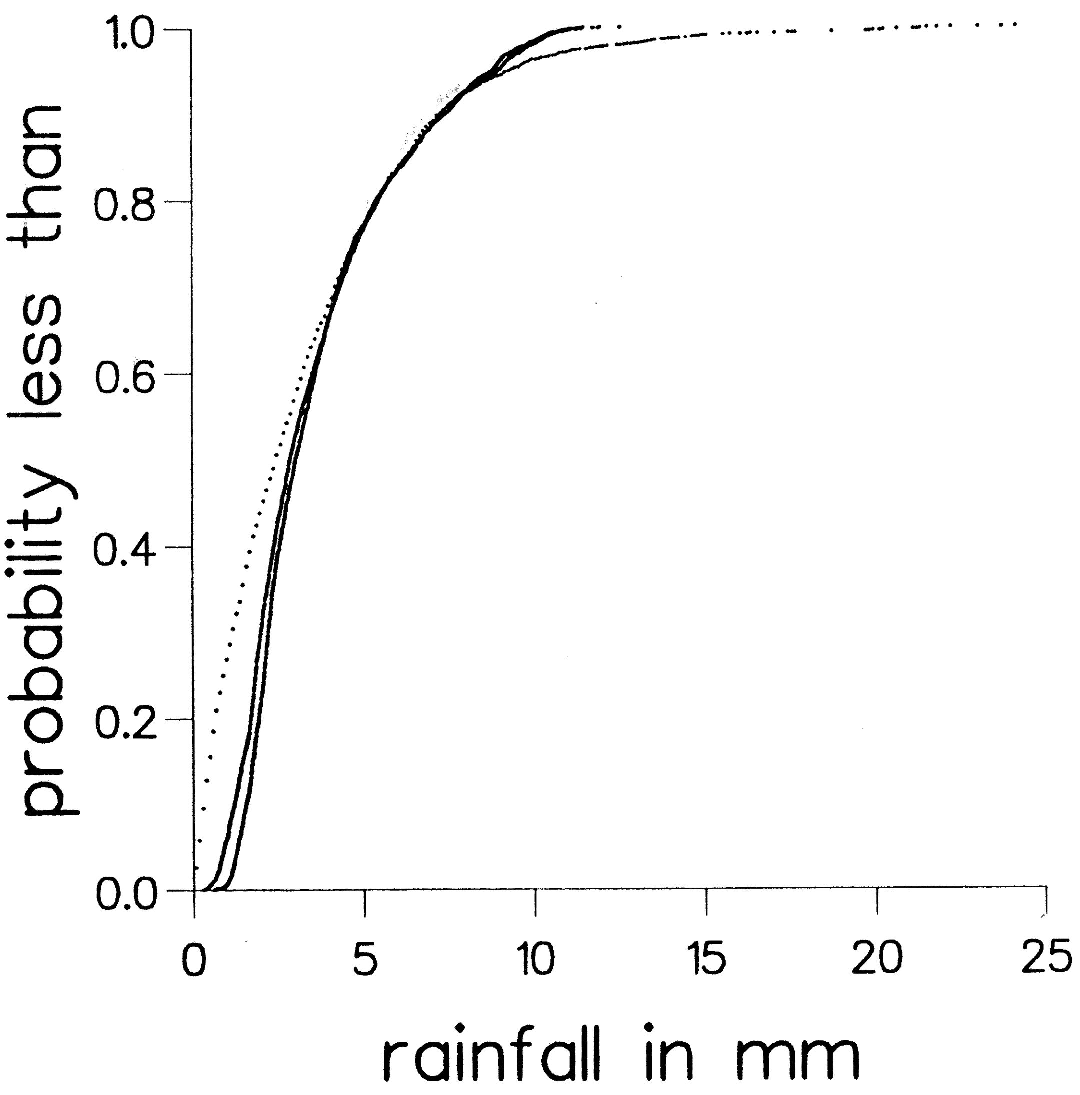

reduction in MSE (mm**2)
3
2
1
0
−1
−2
−3
0
2
4
6
8
10
rainfall level


reduction in AME (mm)
1.0
0.5
0.0
−0.5
0
2
4
6
8
10
rainfall level



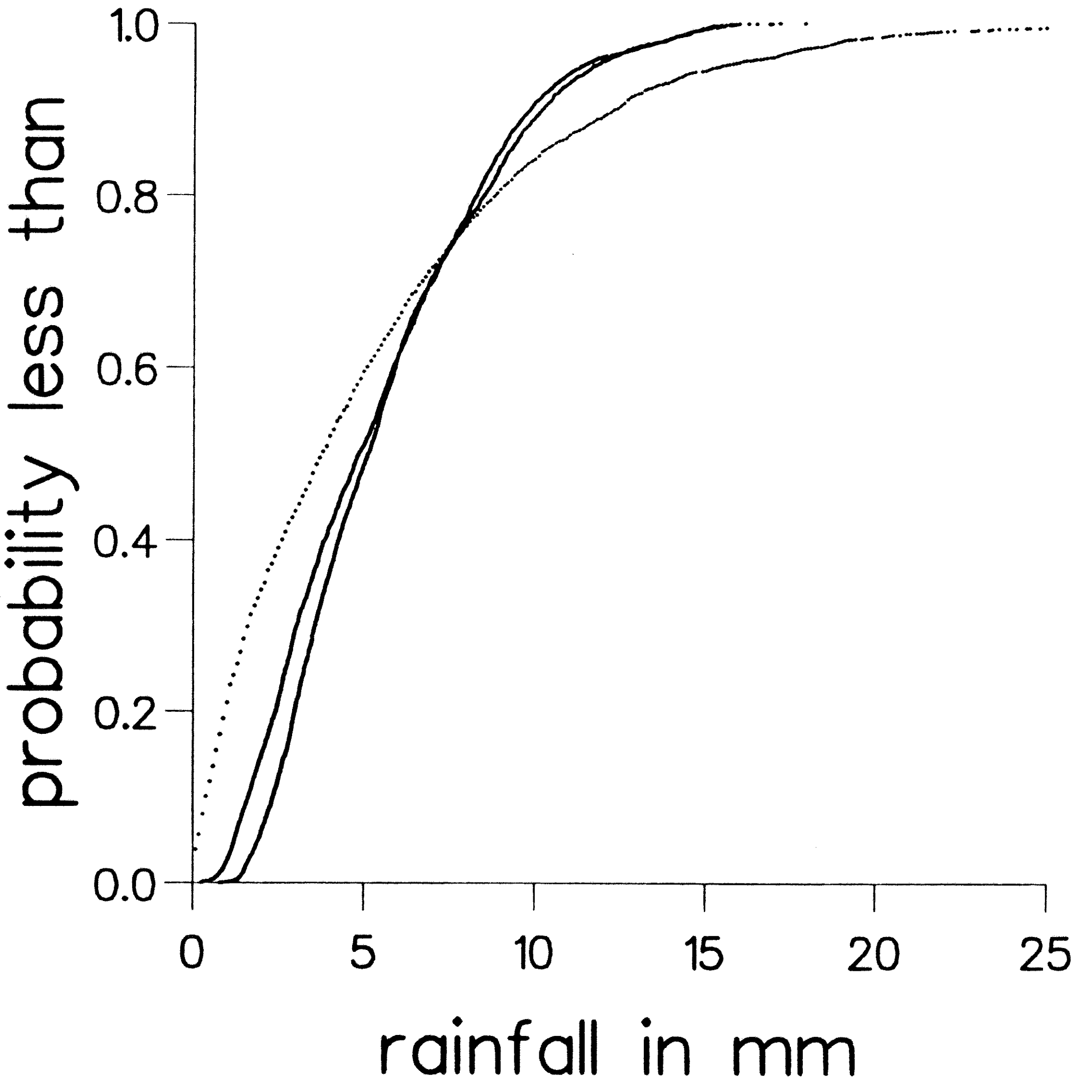
reduction in MSE (mm**2)
3
2
1
0
−1
−2
−3
0
2
4
6
8
10
rainfall level

Fig 11



probability less than
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
rainfall in mm

Figure 2



probability less than
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
rainfall in mm

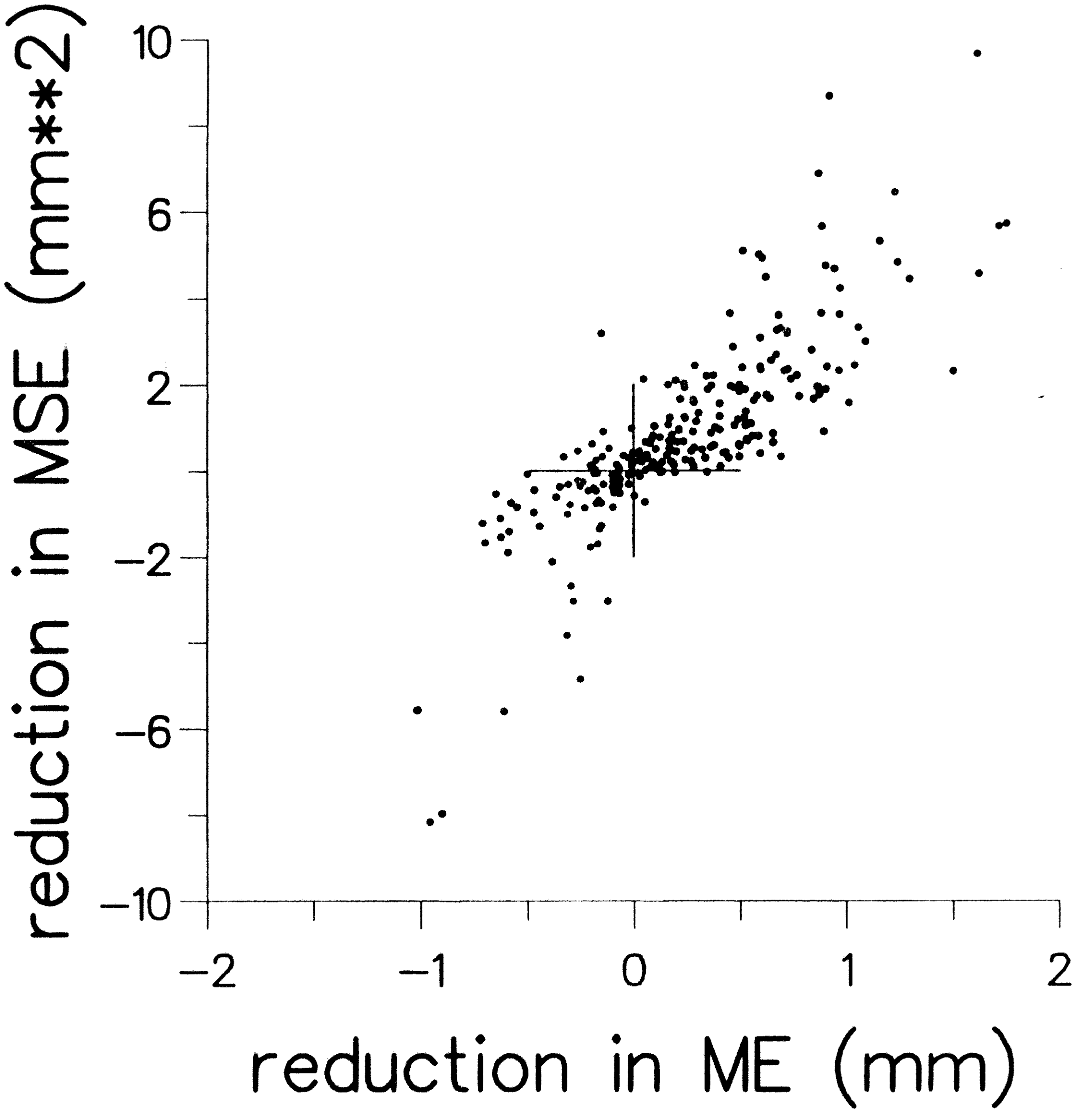

reduction in MSE (mm**2)
10
6
2
−2
−6
−10
−2
−1
0
1
2
reduction in ME (mm)

Fig 14

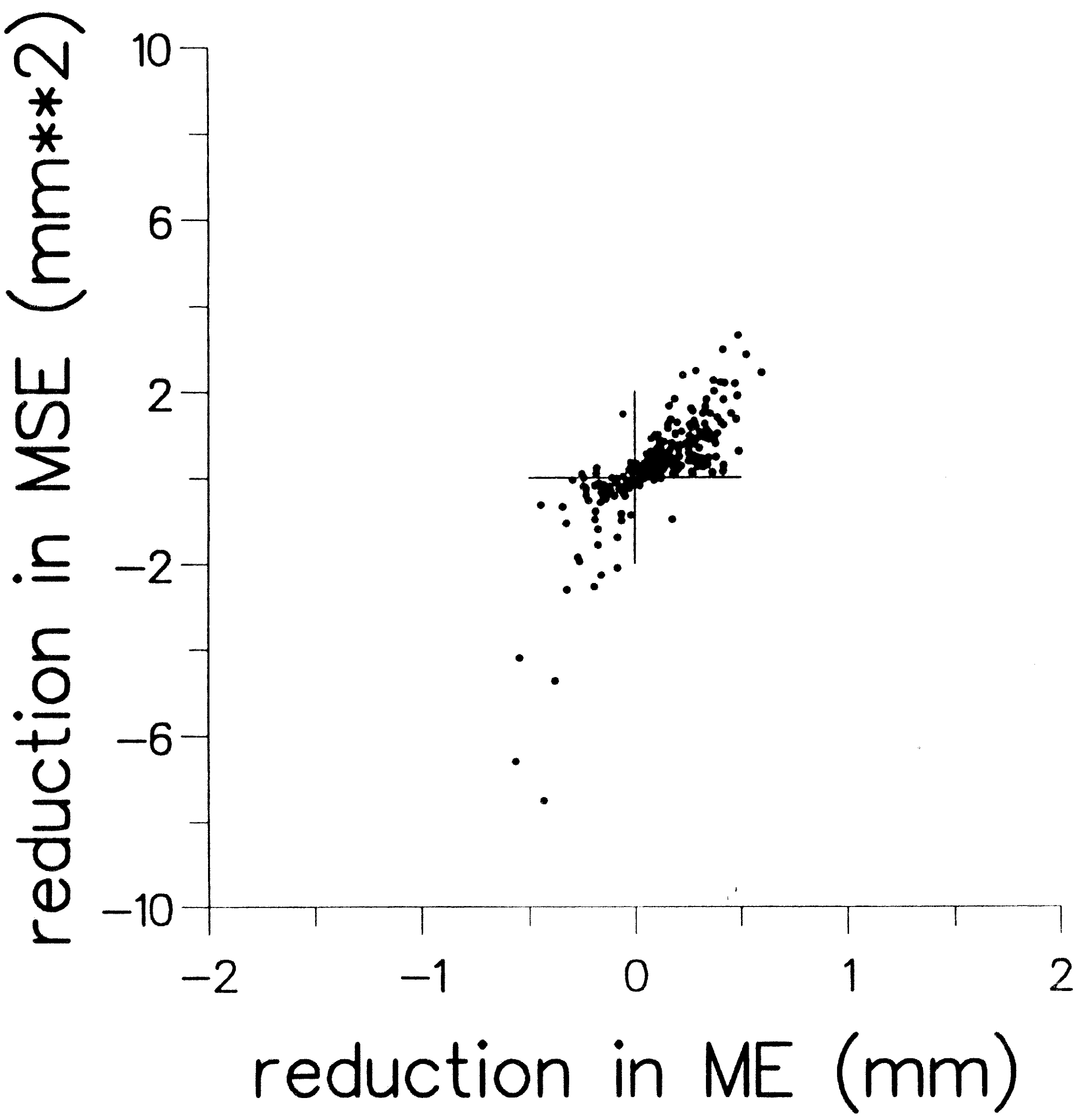


Fig A1

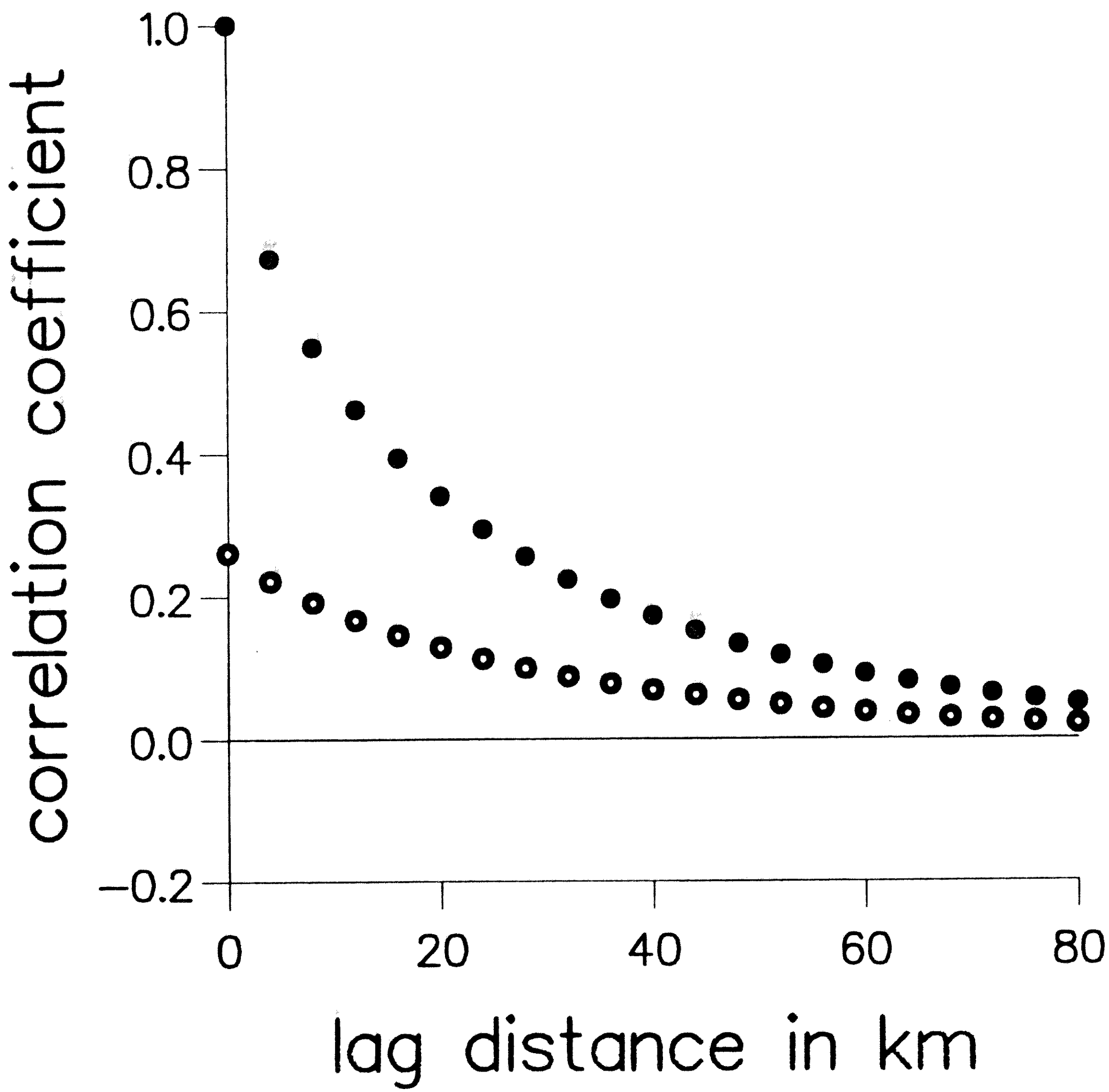
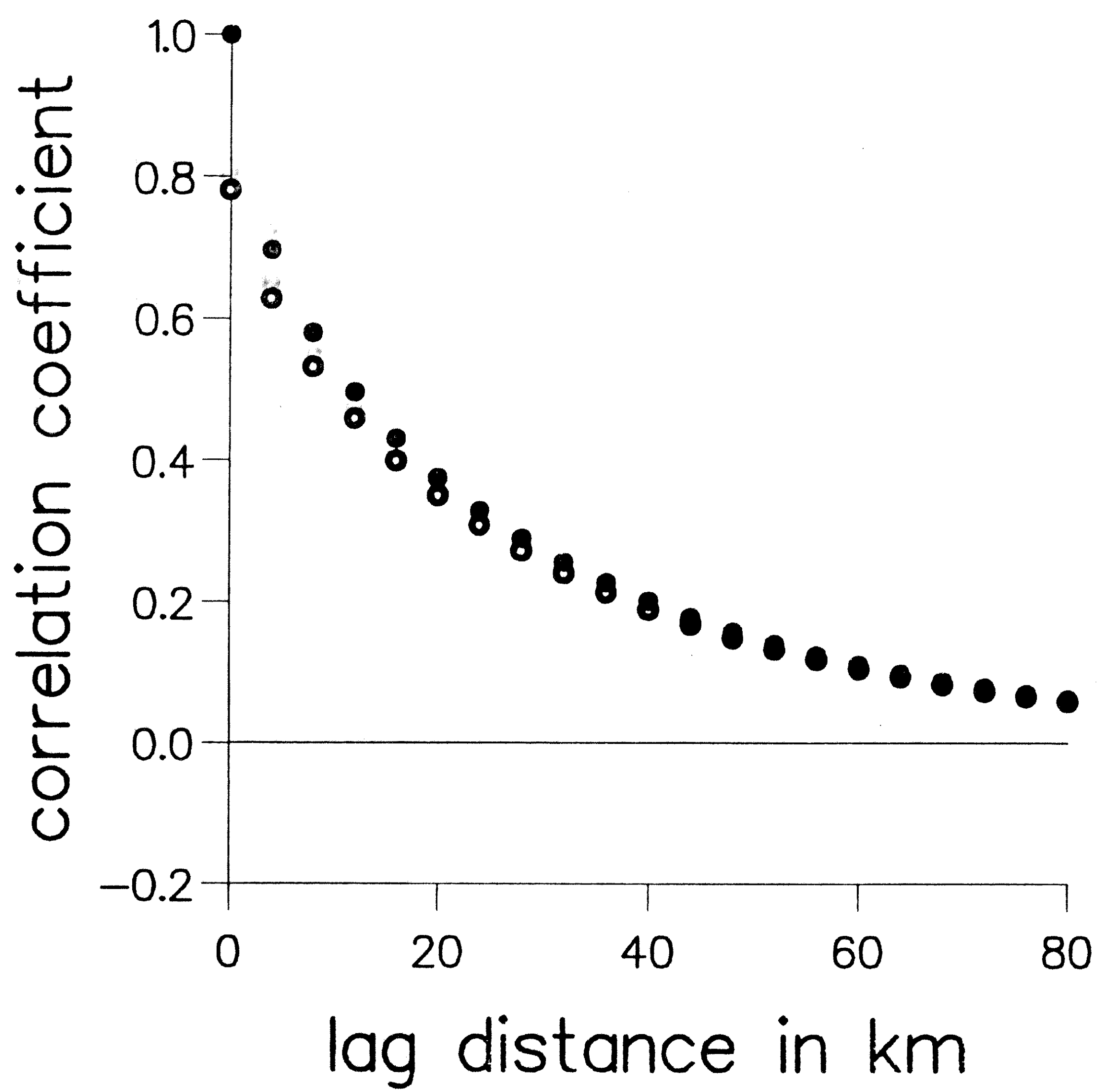

correlation coefficient
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
0
20
40
60
80
lag distance in km

Fig A2



1.0
0.8
0.6
0.4
0.2
0.0
−0.2
correlation coefficient
0
20
40
60
80
lag distance in km