RAINFALL SPATIAL AND SEASONAL VARIABILITY ANALYSIS IN SEMI - ARID WATERSHEDS

by

Daquan Tian

A Thesis Submitted to the Faculty of the

DEPARTMENT OF HYDROLOGY AND WATER RESOURCES

In Partial Fulfillment of Requirements

For the Degree of

MASTER OF SCIENCE

WITH A MAJOR IN HYDROLOGY

In the Graduate College

THE UNIVERSITY OF ARIZONA

1993

STATEMENT BY AUTHOR

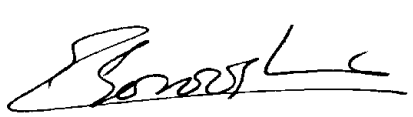
The thesis has been submitted in partial fulfillment of requirements for an advanced degree at the university of Arizona and is deposited in the University Library to be made available to borrowers under rules of the library.

Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction this manuscript in whole or in part may be granted by the head of the major department of the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED : 

APPROVAL BY THESIS DIRECTOR

This thesis has been approved on the date shown below.



Soroosh Sorooshian, Professor
Hydrology and Water Resources

7/14/93
Date

ACKNOWLEDGMENTS

Many people have contributed to this study. Special thanks to Dr. Soroosh Sorooshian for his constant guidance, attention, patience, and encouragement. I am grateful to Dr. Donald E. Myers for his expert guidance, and his enthusiasm in sharing his knowledge. Comments and critiques by Dr. David C. Goodrich and Dr. Jene D. Michaud have contributed greatly to this study. Special thanks are extended to Dr. Jene D. Michaud for her review of the manuscript. At the USDA-ARS Walnut Gulch Watershed in Tombstone, AZ, Howard Larson, Jim Smith, Art Dolphin and John Smith contributed to the success of the field work. I am thankful to the Tucson USDA-ARS staff for providing Walnut Gulch rainfall data.

Financial support was provided by the National Science Foundation (grant BCS - 8920 - 851), and the National Weather Service, under cooperative agreement NA90AA - H - HY505.

I would like to express my deepest thanks to my wife Mofen Jiang, for her unending encouragement, assistance, and constant support.

TABLE OF CONTENTS

TABLE OF CONTENTS - Continued

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS - Continued

Figure Page

LIST OF TABLES

ABSTRACT

A simple procedure for the real-time estimation of the average areal rainfall over a 4 km by 4 km grid was developed. The method proposed here was obtained by a straightforward extension of the well-known block Kriging approach. In order to take into account the seasonal influences, the monthly variogram models were built by least square fit for a ten-year period of rainfall data collected at USDA-ARS Walnut Gulch Experimental watershed. To investigate the rainfall spatial variability on a small scale, a small scale experiment was conducted within the watershed. Sensitivity analysis was also applied between the raingage densities and the relative estimated errors.

A portable program MATCORS, for correspondence analysis was developed. A series of graphs and various diagnostic results generated by this program are provided to aid in interpreting the results. It was applied to a monthly rainfall data set recorded statewide in Arizona.

CHAPTER ONE

INTRODUCTION

1.1 The Spatial and Seasonal Variability of Rainfall

The rainfall process can be viewed as a spatiotemporal phenomenon. It is measured by means of radar, raingage networks or combinations of the two. Orlanski (1975) summarized that thunderstorms (convective storms) have characteristic scales between two and twenty kilometers spatially and between two minutes and four hours temporally. Precipitation is the primary factor in runoff generation. Furthermore, most rainfall-runoff models require rainfall to be spatially interpolated from measurement points to areal values over the watershed or over several sub-watersheds. An important improvement upon existing rainfall-runoff models would result from incorporation of the spatial and seasonal characteristics of rainfall.

Although rainfall recorders and raingage networks have provided numerous experimental data about spatial and seasonal rainfall distribution, there are a lack of data collection networks with sufficiently small resolution and of large area coverage, such as a typical GCM grid square (10^4 km^2). This is mainly due to the operational costs and difficult access to some locations. Therefore, developing a geostatistical method to better characterize and recognize precipitation systems would be quite valuable.

Geostatistics (Matheron, 1971; Journel et al. 1978) and multivariate statistical analysis methods have the advantage of being able to handle spatial and seasonal variability of parameters. Geostatistical methods can present not only

spatial structure but also the random characteristics of a spatial variable. It should be pointed out, however, that rainfall random fields have multi-realizations at the same spatial location. This has important implications for rainfall estimation using a time-invariant geostatistical model (e.g. Kriging). Applying those spatial domain methods to space-time variant rainfall random field raises a problem: how to embody seasonal information into the spatial space so that well-developed geostatistical methods can be used.

The purpose of the present study is to gain insights into the spatial and seasonal variability of rainfall, as well as to better characterize them. The method for real-time estimation of the areal rainfall over a radar pixel grid (4 km by 4 km) is proposed here and can be compared with the Doppler radar signal in the future. To better understand the spatial characters of a rainfall field, a small scale experiment was conducted in a semi-arid environment within the USDA-ARS Walnut Gulch Experimental Watershed near Tombstone, Arizona. A general location map of the overall watershed is presented in Figure 1.1. The entire Walnut Gulch experimental watershed is monitored by a dense recording raingage network. Data collected by 84 raingages from 1980 to 1990 (except 1989) are used in this study. In order to ascertain the seasonal variation in rainfall, correspondence analysis (Benzecri, 1973) was applied to a monthly raingage data set from Arizona. The monthly variograms were built according to the average of the ten-year records at the Walnut Gulch watershed. The rainfall data were partitioned into three seasons: the primary wet season, the secondary wet season, and the dry season.

The monthly variograms, which take into account the influences of the seasonal variations, were developed to compute an average areal rainfall over a

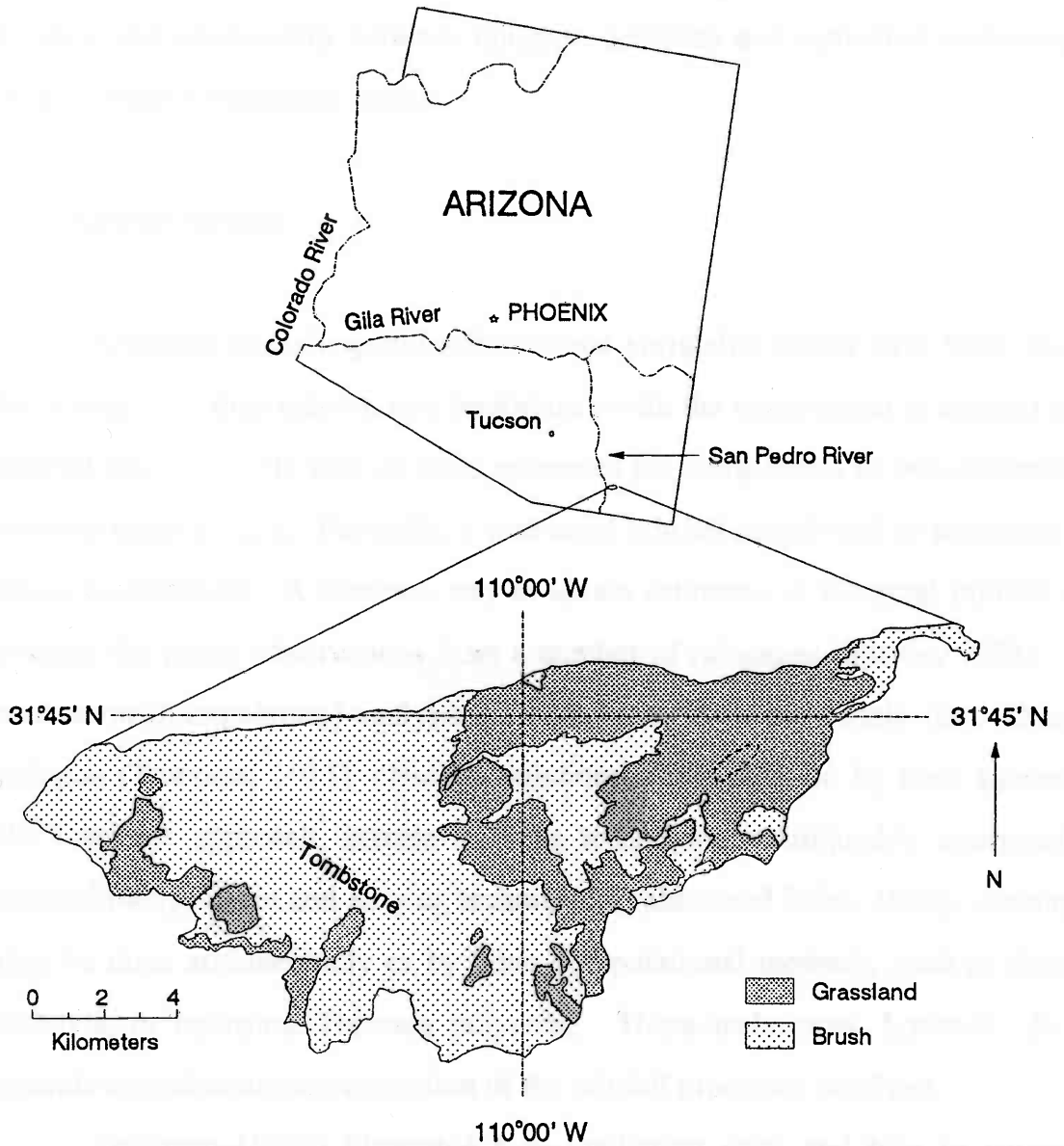


Figure 1.1 USDA-ARS Walnut Gulch Experimental Watershed Location Map

radar pixel using the Kriging technique. A sensitivity analysis was alsoconducted to show the relationship between raingage densities and estimated variances, as well as relative estimated errors.

1.2 Literature Review

Although rainfall spatial and seasonal variability issues have been studied for a long time, they take on new importance with the requirement to convert point rainfall measurements into an areal estimates for comparison to two-dimensional weather radar images. Factually, a true areal rainfall (depth and or intensity) can never be measured. A common way to obtain estimates of the areal rainfall is to average the points observations from a number of raingages (Sharon, 1972). The estimation (interpolation) analysis methods include the commonly used Thiessen polygon (Thiessen, 1911), classical polynomial interpolation by least squares or the Lagrange approach, inverse distance techniques, multiquadric interpolation, optimal interpolation and Kriging techniques (Tabios and Salas, 1985). Averaging may be done arithmetically or by other computational methods, such as drawing isohyets, or by using Thiessen polygons. These techniques, however, do not provide a continuous representation of the rainfall processes involved.

Troutman (1983) illustrated how prediction error and bias in parameter estimation are the direct result of spatial variability of rainfall. The earliest work to analyze the spatial variations of a process over an area were done by Drozdov and Sephelevskii (1946) in studies to establish an error criterion for spatial interpolation and network density design. Gandin (1965) and Belousov, et al.

(1971) developed a technique called optimal interpolation to describe spatial variations of a process.

Matheron (1971) introduced the theory of regionalized variables to estimate areal averages considered as realizations of stochastic processes. This theory led to the development of the Kriging technique, now universally known as geostatistics. It is similar to the optimal interpolation except that the spatial correlation function is replaced by the so-called variogram.

Geostatistics is superior to the other techniques mentioned above, because it remedies a defect of classical statistics. It takes into consideration the random characteristics and the spatial structure of a spatial variable at the same time. It is based on a concept that regionalized samples are not independent, but that a certain correlation is often associated with them. This correlation relationship changes with distance and the relative orientation between the samples.

Delfiner and Delhomme (1975) and Delhomme (1978) applied the Kriging technique to spatial interpolation and areal averaging. Mejia and Rodriguez-Iturbe (1974), and Rodriguez-Iturbe and Mejia (1974) presented extensions of the use of optimal error criteria for areal averaging, generation or synthesis of spatial data and network design focusing on the rainfall process. Similar papers were presented by Lenton and Rodriguez-Iturbe (1977a, 1977b) dealing with rainfall averages and rainfall network analysis using a multidimensional model.

Convective thunderstorms are known to show very important spatial and temporal variability. For example, as can be found in Chapter 4, the difference in rainfall depth for a storm from two stations 1360 meters apart was about 78.7% at the Walnut Gulch watershed on August 2, 1991. In such regions where thunderstorms prevail, a high network density is necessary.

As stated by Hutchinson (1970): " no one definition of areal variability has universal acceptance". In general, the standard error of measurement is still taken as a reference to deal with the spatial and temporal variability in rainfall random fields. If the autocorrelation function begins to fall within the standard error region, the beginning lag (distance) represents the correlation distance between two samples. Chapter 4 uses this method to examine the rainfall depth on a small scale. The correlation coefficient in a multilinear correlation analysis method is another criterion generally used in the study of the variability between raingages (Sharon, 1972, Dean and Snyder, 1977, Shih, 1982).

Point rainfall are usually transformed to an areal average by using a correction factor developed by the U. S. Weather Bureau (Rodriguez-Iturbe and Mejia, 1974). This correction factor was designed to transform the point rainfall frequency curves into the areal frequency curves. These corrections are normally applied to the maximum rain depth at the center of the storm. In order to construct those areal frequency curves, a large amount of historical rainfall data (at least 20 years) with a variety of rainfall duration are needed. This makes its application somewhat questionable except for the region in which it was developed. Roche (1963) proposed a method which converts the point rainfall to an areal average according to the level of the rainfall probability. It is extremely cumbersome because of the large number of graphs which must be constructed and integrated (Rodriguez-Iturbe Mejia, 1974). In Rodriguez-Iturbe and Mejia's (1974) method, the reduction factor used to compute an areal average depends solely on the spatial correlation structure of the point rainfall process. A correlation coefficient is used to estimate the spatial correlation structure by evaluating it at the characteristic correlation distance.

The above methods obviously provide no measure of their accuracy. It is often useful for a hydrologist to evaluate the errors involved in the areal averaging. However, a major difficulty arises in validating the calculated variance from experimental data, since no direct measurement of the areal rainfall are available. Block Kriging can provide a method to compute the areal average with a minimum variance unbiased estimate.

In practice, the block Kriging system calculates block-to-block and point-to-block covariance by separating the estimated area into several sub-blocks. Isaaks and Srivastav (1989) compared the results by using block Kriging and the inverse distance squared average methods. They found the average of the Kriging block estimates is closer to the true mean than the inverse distance squared estimates, and the correlation of the block Kriging estimates with the true values is also larger. Furthermore, the mean absolute error (MAE) and mean square error (MSE), which provide measures of the combination of bias and spread, both favor block Kriging estimates.

In fact, there are very few studies dealing with rainfall spatial variability at a very small scale (few hundred meters). Most studies in the past were either performed at very large scales (>200 km^2) to examine orographic effects (Peck, 1973), or conducted at the watershed scale (few hundred km^2) in conjunction with the runoff studies (Huff, 1970; Merva et al., 1971; and Goodrich, 1990). The spatial variability of precipitation has also been studied with the intention of optimizing gage density for the purpose of estimating the areal average depth (Osborn et al., 1972; and 1979). Fennessey et al. (1986) determined the spatial distribution of total rainfall depth using spatial correlation.

Faures (1990) conducted an experimental study of small-scale rainfall variability at Walnut Gulch. In his study, raingages were installed on a 30 meter by 30 meter square grid covering a subwatershed of about 4.4 hectares. He concluded that spatial variations of the rainfall field exist at a scale smaller than larger scale Walnut Gulch networks can resolve. Hershfield (1969) reported a large rainfall variation in his small scale study as well.

No reference for using the correspondence analysis method to study the rainfall spatial and seasonal variability was found in the literature. However, in recent years hydrologists have shown an increased interest in research on the spatial and temporal properties of precipitation (Georgakakos and Kavvas, 1987, and Gupta, 1985). In particular, further advances in stochastic modeling of rainfall process demand a more detailed statistical description of precipitation variability and dynamics on different temporal and spatial scales (Berndtson, 1988). It is difficult to apply geostatistical techniques with respect to both temporal and spatial properties of rainfall at the same time. This is because of fundamental differences between spatial and spatiotemporal processes (Rouhani and Myers, 1990).

Bastin et al. (1984) have shown the seasonal behavior of spatial variability of the rainfall process in two river basins. They found that the variogram appears much larger in summer than in autumn. Berndtson (1988) used a cross correlation for lagged time steps to indicate the temporal dependence for different spatial scales and possible translation of rainfall areas.

1.3 Problem Statement and Approach

Precipitation fields may be regarded as realizations of a space-time random function, but most published works dealing with estimation of rainfall depth and intensity de-emphasize the role of the time dimension in order to deal with the spatial model (Rouhani and Myers, 1990). For instance, the rainfall depth of a storm event is often considered in space as a regionalized variable. Although rainfall distribution is closely related to the topographical features of the surrounding area (Kassim and Kottegoda, 1991), applying such space-oriented approaches to spatiotemporal processes may lead to the loss of valuable information in the time dimension.

To deal with a spatiotemporal variable, it is customary to determine the dependent dominator first, then either use time-series methods with a spatial average or use a spatial analysis method (e.g. Geostatistics) and treat the seasonal variable as a periodic function. It is extremely difficult to estimate a rainfall event with respect to both time and space. Rouhani and Myers (1992) indicated "this is due to the traditional application of geostatistics and time-series methods wherein only spatial or only seasonal dependence, but not both, need to be considered". Other problems include qualitative differences between spatial and seasonal processes and imbalances between quantities of seasonal and spatial information.

The purpose of this study is to describe and better characterize the rainfall fields both on seasonal and spatial scales. It also attempts to establish a reliable real time-areal estimation method. Correspondence analysis and geostatistics (Kriging) methods were used in this study to verify and represent the seasonal and

spatial trends of rainfall random field and to optimally estimate the average areal rainfall. Geographically, this study is restricted to the state of Arizona.

In the near future, the NEXRAD (NEXt generation RADar) data will become available. It is generally expected that average areal rainfall would be more easily obtained from NEXRAD Doppler radar measurements than through raingage measurements. But the radar measurements cannot be compared with a "true" average rainfall, because the latter is unknown. The purpose of this study was also to demonstrate a simple procedure for the real-time estimation of the average rainfall over a radar pixel. Because the average areal rainfall is estimated by a linear unbiased minimum variance estimation method (Kriging), it can be used as a ground measurement reference to compare with the Doppler radar measurements. Primarily, this study addresses the variance of the local average process.

The rainfall spatial variability at extremely small scales (30-180 m) is also closely examined in this study. The objective is to develop a general description of the small-scale spatial characteristics of rainfall fields.

CHAPTER TWO*

CORRESPONDENCE ANALYSIS FOR RAINFALL DATA

2.1 Introduction

Correspondence analysis (Benzecri, 1973), is a multivariate statistical analysis method that simultaneously produces R (variable) and Q (sample) mode analyses. Similar to principal component analysis (PCA), correspondence analysis (CA) uses principal factors to extract the significant information from the variable space and from the sample space by projecting into lower dimensional space and by quantifying the correlation between the variation and between samples. The principal factors G_1, G_2, ... G_n in the variable space (R^p) have the same contribution to the total variation as the principal factors F_1, F_2 ... F_p in the sample space (R^n). Unlike principal component analysis, correspondence analysis treats variables and samples in a symmetrical fashion. These principal factors can be used to graphically represent the variables and samples in a manner that geometrically describes the correlation structure. In Correspondence Analysis the principal factors for variables are dual to the principal factors for samples, hence it is possible to combine the plots using the same axes. Several aspects of the correlation structure of a data set may be seen in such graphical representations:

1. Variables that plot close together are more highly correlated than those far

* This chapter is a paper titled "Correspondence Analysis with Matlab" by Daquan Tian, Soroosh Sorooshian and Donald E. Myers, Computers & Geosciences, in press.

apart. Note that variables that plot close together with one pair of principal factors may not for another pair, hence close proximity may not always indicate high correlation.

2. Samples that plot close together are more highly correlated than those far apart. Note that samples that plot close together for one pair of principal factors may not for another pair, hence close proximity may not always indicate high correlation.
3. The combined plots provide some indication of the degree of correlation between a variable and a group of samples or between a sample and a group of variables, but this graphical indication of correlation is less reliable than that between samples or between variables. This is discussed in greater detail in LeBart, Morineau and Warrick (1984).

There are several commercial programs (Dual3, MAPWISE, PC-MDS, SimCA) for correspondence analysis, but none of these are portable, and do not incorporate high-quality graphics (Goldstein, 1991). The objective of this Chapter is to present a computer program that can perform correspondence analysis easily and quickly with high-quality graphic plots. To perform a correspondence analysis for precipitation data set is also extended to find its seasonal and spatial patterns. The program is written with Matlab (version 3.5), a software package for matrix operations which shares many of the characteristics of a high-level programming language. It operates interactively on a UNIX system (workstation) in the user-friendly Matlab environment. It also can be used on a personal computer (IBM, Macintosh, etc.). Drivers for different hard-copy devices are available within Matlab. There is almost no limitation on matrix size

(depending on the machine memory). This program checks the input data matrix to ensure there are no negative entries. It also allows users to specify supplementary rows and columns. Thus the CA solution is determined by the active points only. The supplementary variables and samples then will be projected on the factorial loading plots. The plotting feature, of Matlab, provides an easy and immediate visual inspection of the results. This particular attraction is the results of the fact that correspondence analysis is a geometric approach to data analysis (Goldstein, 1991). The point label also is included on the plots and the high quality output plots are acceptable for use in a presentation or paper. The output graphics can be saved as a PostScript file to take advantage of the capability of a laser printer.

2.2 Method

Let the input matrix be represented by the non-negative entries x_{ij} (i = 1,...., n; j = 1,...., p). Every element in the input matrix is divided by the sum of the (nxp) data elements to form a new matrix which may be referred top as the relative frequency matrix $F = X / \sum_{i=1}^{n}\sum_{j=1}^{p} x_{ij}$. The variable weights are defined by the row vectors $f_{0j} = \sum_{i=1}^{n} f_{ij}$, and sample weights by column vectors $f_{i0} = \sum_{j=1}^{p} f_{ij}$, respectively. There are two normalizing matrices, $D_n = diag[f_{i0}]$ and $D_p = diag[f_{0j}]$. Thus the sample coordinate matrix is obtained by weighting the frequency matrix as $D_n^{-1}F = f_{ij} / f_{i0}$. The coordinates of every sample in this matrix are in proportion to the values of the variables in the samples. So the relationship between variables can be represented by the relative location of n sample points in the variable space (R^p). The usual Euclidean distance between two variables *l* and *m* is given by:

$$d(l,m)=\sqrt{\sum_{j=1}^{p}\left(\frac{f_{lj}}{f_{l0}}-\frac{f_{mj}}{f_{m0}}\right)^2} \tag{2.1}$$

Because the distance (2.1) gives the same weight to each variable, the chi-squared distance is used in place of the usual Euclidean distance (Rhodes and Myers, 1991).

$$d(l,m)=\sqrt{\sum_{j=1}^{p}\left(\frac{f_{lj}}{f_{l0}\sqrt{f_{0j}}}-\frac{f_{mj}}{f_{m0}\sqrt{f_{0j}}}\right)^2} \tag{2.2}$$

When the principal axes for these distances are computed and plotted on the usual Cartesian system, the relationship between the samples and variables can be demonstrated visually.

The probability average of j^{th} variable in the sample space (R^n) is

$$\sum_{i=1}^{n}\frac{f_{ij}}{f_{i0}\sqrt{f_{0j}}}f_{i0}=\frac{1}{\sqrt{f_{0j}}}\sum_{i=1}^{n}f_{ij} \tag{2.3}$$

Let $\mathbf{W}=(w_{ij})_{nxp}$ and $\mathbf{A}=\mathbf{W}^T\mathbf{W}_{(pxp)}$;

$$w_{ij}=\frac{f_{ij}-f_{i0}f_{0j}}{\sqrt{f_{i0}f_{0j}}}=\frac{x_{ij}-x_{i0}x_{0j}/\sum_{i=1}^{n}\sum_{j=1}^{p}x_{ij}}{\sqrt{x_{i0}x_{0j}}} \tag{2.4}$$

The factors **U** for variables can be computed by the eigen decomposition of the symmetric matrix **A**($[\mathbf{A}]=[\mathbf{U}][\mathbf{D}][\mathbf{U}^T]$). **U** is a matrix of eigenvectors and a diagonal matrix **D** contains the eigenvalues. The principal axes for variables (R-mode factor loading) then are obtained from the scaled factors $\mathbf{R}=\mathbf{D}^{1/2}\mathbf{U}$.

Similarly, the principal axes for samples (Q-mode factor loading) are $\mathbf{Q}=\mathbf{D}^{1/2}\mathbf{V}$, where **V** is obtained from eigen decomposition of symmetric matrix **B**

$= \mathbf{WW}^{T}_{(nxn)}$.

The original element can be reconstructed from a bilinear form (Avila and Myers, 1991):

$$f_{ij} = f_{i0} f_{0j} (1 + \sum_{k=1}^{p-1} \sqrt{d_k} R_{ik} Q_{jk}) \qquad (2.5)$$

This program calculates several quantities that aid in the interpretation of the output (Avila and Myers, 1991).

(1) The cumulative percentage of variation: it is a global measure of fit when K factors are retained. The term variation in here does not refer to "variance'" in the normal statistical sense. It is expressed as a cumulative percentage of explained variation which is similar to the measure used in principal component analysis, and is given by:

$$\sum_{k}^{K} d_k \; / \; \sum_{k}^{p-1} d_k \qquad (2.6)$$

(2) Relative contributions: these are measures of sample or variable variation explained by a particular factor. If all factors are retained, the sum is equal to 100 for a particular sample or variable. It is computed by:

For every variable j=1, ..., p

$$RC^{k}(j) = d_k u_{jk}^2 \; / \; \sum_{l=1}^{p-1} d_l u_{jl}^2 \qquad \text{k=1,.....,p-1} \qquad (2.7)$$

and, for every sample i=1,.....,n

$$RC^{k}(j) = d_k v_{jk}^2 \; / \; \sum_{l=1}^{p-1} d_l v_{jl}^2 \qquad \text{k=1,.....,p-1} \qquad (2.8)$$

(3) Absolute contributions: these are the contributions of variables or samples to a factor. For a particular factor it is equal to 100.

For every factor k=1, ..., p-1

$$AC^k(j) = (f_{i0}(j)\ (diag(f_{0j}(j))^{-1}u(j,k))^2 \qquad \text{variable j=1,....,p} \qquad (2.9)$$

$$AC^k(i) = (f_{0j}(i)(diag(f_{i0}(i))^{-1}u(i,k))^2 \qquad \text{sample i=1,.....,n} \qquad (2.10)$$

(4) Error profiles: these provide a measure of the errors when the original data matrix is "reconstructed" by K principal factors. They are defined as:

Error profile for variable j=1, ..., p

$$EP(j) = \sum_{i=1}^{n} f_{i0}\left(\sum_{k=K+1}^{p-1}\sqrt{d_k}\, v_{ik}\, u_{jk}\right)^2 \qquad (2.11)$$

Error profile for sample i=1, ..., n

$$EP(i) = \sum_{i=1}^{n} f_{0j}\left(\sum_{k=K+1}^{p-1}\sqrt{d_k}\, v_{ik}\, u_{jk}\right)^2 \qquad (2.12)$$

(5) Supplementary elements: these are inactive elements in the data matrix, as they are not used in CA to determine the factors. Their projections onto the factors determined only by the active elements are computed by the formula:

For a supplementary row s

$$SP = \sum_{j=1}^{p} \frac{f_{sj}}{f_{s0}} u_{kj} \qquad (2.13)$$

For a supplementary column s

$$SV = \sum_{i=1}^{n} \frac{f_{is}}{f_{0s}} v_{ik} \qquad (2.14)$$

2.3 Program Description

A Matlab code, MATCORS, is composed of three functions as listed in Appendix A. The main function MATCORS echoes with input options, checks negative entries in the input matrix, and calls the subfunction EIGV to form a symmetric matrix (**A**) and to perform eigen decomposition. Matlab built-in function (eig) can get the same eigenvectors, but opposite sign. If supplementary elements are selected, it will delete the supplementary elements to form an active data matrix, and calculate the supplementary projections. It also enables EIGV to operate again with a new active data matrix. Then MATCORS obtains the factor loading for both variables and samples. The program first calculates the number of principal factors according to the cumulative percent of variation which is specified by the user, then compares it with the number of principal factors that user wants to keep. The larger one is used for the factor loading. Several quantities used for interpreting the results are also calculated by this program. These quantities include the variable and sample weights, the relative and absolute contributions for every factor, as well as error profiles for each variable and sample, etc. Finally, a series of graphics is generated by this program, which includes a two-dimensional factor loading plot, a combined plane for both Q-mode and R-mode, and bar graphics for variable and sample weights, their absolute contributions and error profiles.

The subfunction EIGV constructs a real symmetric covariance matrix of variables and performs eigen decomposition based on the Jacobi's orthogonal rotation method. The subfunction INDEX, which is called by EIGV, locates the row and column number of maximum off-diagonal element, as well as its absolute

value in a real symmetric matrix.

The input data matrix is read in free-format. Several parameters are input from the keyboard:

pre : cumulative percent variation to determine the number of principal factors retained.

nb : desired number of principal factors to be retained.

nc : the number of supplementary variables.

cc : list of supplementary variables.

nr : the number of supplementary samples.

cr : list of supplementary samples.

f1 & f2 : two principal factors to be plotted.

The program output consists of the following matrices, which can be saved as eight MATLAB files or ASCII files for later use.

r : R-mode factor loading pxp matrix.

q : Q-mode factor loading nxp matrix.

ev : Eigenvalues, relative and cumulative variation explained by factors.

c_ra : p by 2p-2 variables contribution matrix; the first p-1 columns are the relative contributions, and the rest are the absolute contributions.

s_ra : n by 2p-2 samples contribution matrix; the first p-1 columns are the relative contributions, and the rest are the absolute contributions.

evv : weights and error profiles of variables.

ess : weights and error profiles of samples.

sv : Supplementary variable projections.

sp : Supplementary sample projections.

2.4 Application

Two examples are presented in this section. Table 2.1 lists an example data set from Luo and Xing (1987) for which a detailed analysis is given to verify the program. Figure 2.1 shows the plots using the original data set, while Figure 2.2 is a graphic plot in which variable 4 and samples 4 and 8 are made supplementary. Appendix B lists the output for this example. Both the factors loading matrix and factor plane plot are identical to Luo at al.'s results, except for the supplementary projection plots (his program does not have this type of function).

Table 2.1 Soil sample data (Luo and Xing, 1987)

Sample No.	Sand Content	Silt Content	Clay Content	Organic Matter	pH
1	77.3	13.0	9.7	1.5	6.4
2	82.5	10.0	7.5	1.5	6.5
3	66.9	20.6	12.5	2.3	7.0
4	47.2	33.3	19.0	2.8	5.8
5	65.3	20.5	14.2	1.9	6.9
6	83.3	10.0	6.7	2.2	7.0
7	81.6	12.7	5.7	2.9	6.7
8	47.8	36.5	15.7	2.3	7.2
9	48.6	37.1	14.3	2.1	7.2
10	61.6	25.5	12.9	1.9	7.3
11	58.6	26.5	14.9	2.4	6.7
12	69.3	22.3	8.4	4.0	7.0
13	61.8	30.8	7.4	2.7	6.4
14	67.7	25.3	7.0	4.8	7.3
15	57.2	31.2	11.6	2.4	6.3
16	67.2	22.7	10.1	3.3	6.2
17	59.2	31.2	9.6	2.4	6.0
18	80.2	13.2	6.6	2.0	5.8
19	82.2	11.1	6.7	2.2	7.2
20	69.7	20.7	9.6	3.1	5.9

Table 2.2 lists the monthly average precipitation in inches for 64 weather stations in Arizona (EarthInfo, 1992). The columns represent the months January-December, and the rows are the station numbers as samples. All of the data in this table have at least a 10 years record. Appendix C lists the weather station descriptions.

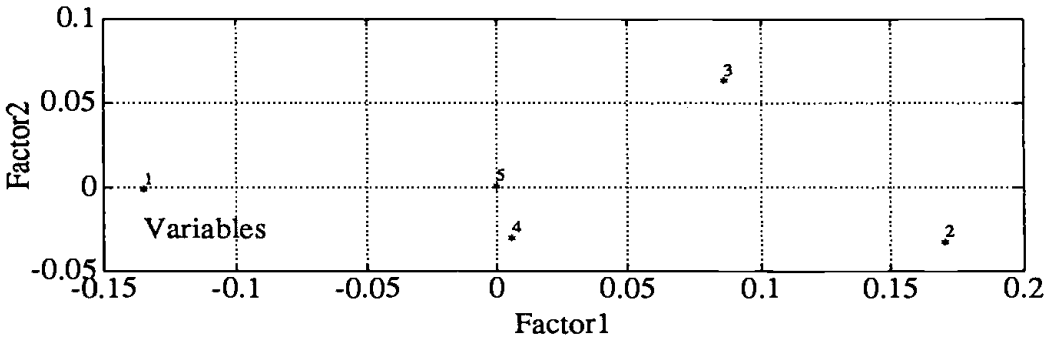
Table 2.2 Monthly average rainfall in Arizona (inch)

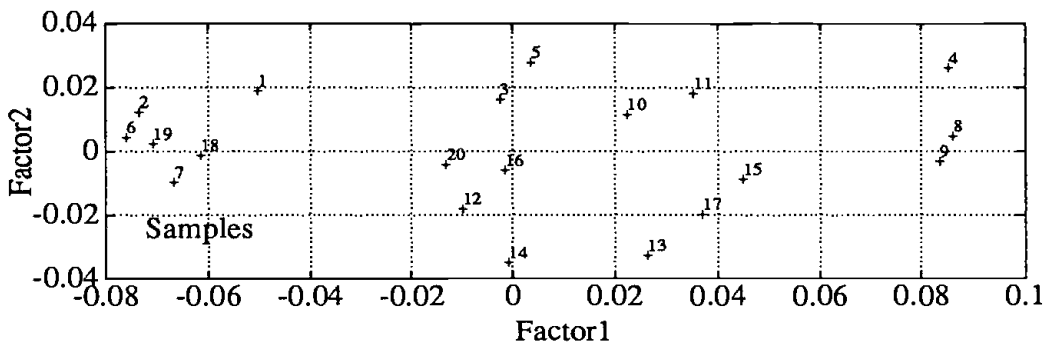
No.	Jan.	Feb.	Mar.	Apr.	May.	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
1	0.63	0.48	0.66	0.21	0.09	0.07	1.16	1.84	0.68	0.56	0.49	0.81
2	0.25	0.22	0.82	0.69	0.05	0.18	0.72	0.64	0.23	0.47	0.93	0.35
3	0.63	0.51	0.59	0.35	0.13	0.03	0.59	1.21	0.58	0.57	0.66	0.81
4	0.90	1.03	1.12	0.55	0.45	0.37	1.56	2.44	1.15	0.94	1.03	1.02
5	0.93	1.03	0.41	1.30	0.27	0.48	1.41	0.78	0.56	0.36	1.44	0.72
6	1.17	0.83	0.77	0.33	0.18	0.53	4.05	3.74	1.57	0.70	0.55	1.65
7	1.64	1.37	1.63	0.71	0.44	0.56	3.22	3.01	1.76	1.30	1.14	1.63
8	0.76	0.59	0.49	0.17	0.09	0.57	2.03	2.54	0.87	0.76	0.33	0.84
9	0.75	0.63	0.86	0.30	0.11	0.12	0.86	1.01	0.67	0.78	0.62	1.02
10	1.47	1.38	1.91	0.81	0.38	0.40	1.83	3.27	2.88	1.57	2.23	1.57
11	0.76	0.48	0.47	0.24	0.23	0.38	2.53	2.33	1.22	0.71	0.41	0.85
12	2.46	2.00	3.02	1.34	0.46	0.52	2.82	3.94	2.63	1.31	3.26	3.04
13	0.72	0.49	0.53	0.24	0.27	0.60	3.75	2.80	1.39	0.99	0.52	0.86
14	1.08	0.99	0.64	0.34	0.34	0.46	2.93	1.60	1.32	1.02	0.79	1.13
15	2.04	1.91	2.21	1.36	0.68	0.56	2.63	2.85	1.78	1.57	1.87	2.13
16	1.38	1.17	1.52	0.28	0.22	0.17	0.56	1.48	0.88	1.36	0.75	0.78
17	1.04	1.25	1.24	1.03	0.57	0.42	1.16	1.88	1.15	1.08	1.05	1.36
18	1.77	1.48	1.87	0.90	0.57	0.48	1.66	1.12	1.00	0.96	1.22	1.04
19	1.23	0.49	0.88	0.68	0.25	0.25	1.61	1.71	0.49	0.59	0.46	0.63
20	0.65	1.08	0.71	0.57	0.16	0.29	0.95	1.70	0.80	0.51	0.83	1.02
21	0.72	0.69	0.82	0.61	0.36	0.34	1.06	1.40	0.75	0.91	0.66	0.66
22	1.06	0.90	0.95	0.64	0.23	0.13	0.85	1.34	0.64	0.62	0.67	0.81
23	0.69	0.93	1.02	0.50	0.23	0.18	0.86	1.36	0.68	0.52	0.64	0.73
24	1.54	1.89	2.07	0.51	0.41	0.36	2.47	2.54	1.77	0.96	1.44	1.25
25	1.17	1.37	1.46	0.80	0.27	0.32	1.48	2.46	1.43	1.31	1.02	1.18
26	0.91	0.74	0.77	0.24	0.15	0.37	4.49	3.48	1.57	1.09	0.63	1.18

Table 2.2, Continued

27	1.96	1.43	1.82	0.87	0.30	0.32	3.18	3.13	1.74	1.85	1.52	2.25
28	0.34	0.43	0.66	0.35	0.39	0.20	0.44	0.60	0.53	0.72	0.52	0.34
29	0.56	0.54	0.52	0.18	0.08	0.06	0.55	1.17	0.72	0.58	0.47	0.76
30	1.91	1.69	2.12	0.97	0.51	0.40	2.57	3.04	1.84	1.43	1.57	1.72
31	1.04	0.76	1.36	0.82	0.42	0.54	1.46	2.06	0.52	0.67	1.72	0.27
32	0.52	0.55	0.66	0.33	0.38	0.34	1.33	1.56	0.99	1.05	0.51	0.56
33	0.76	0.63	0.80	0.23	0.11	0.14	0.85	0.99	0.74	0.68	0.53	0.91
34	0.73	0.53	0.63	0.37	0.05	0.14	0.83	1.26	0.57	0.53	0.44	0.62
35	0.77	0.52	0.66	0.32	0.12	0.37	2.28	2.10	0.89	0.62	0.46	0.82
36	1.40	0.93	1.03	0.60	0.23	0.55	2.33	2.97	1.27	0.92	0.95	1.85
37	0.97	0.71	0.77	0.61	0.33	0.44	2.74	2.48	1.01	0.62	0.65	1.01
38	2.10	1.71	1.52	0.81	0.35	0.64	3.23	3.98	1.90	2.35	1.51	2.21
39	0.91	0.74	0.39	0.35	0.15	0.13	0.87	0.87	0.62	0.14	0.95	0.36
40	1.51	1.09	1.20	0.62	0.18	0.65	4.54	3.74	1.87	1.45	0.97	1.78
41	2.18	2.27	2.23	1.22	0.69	0.33	1.59	1.46	1.69	1.37	1.97	1.02
42	1.96	1.14	1.90	0.76	0.28	0.33	2.01	2.73	0.87	1.46	0.94	1.28
43	3.38	2.79	2.88	1.12	0.95	0.24	2.22	3.25	2.53	1.69	2.98	2.92
44	1.00	0.54	0.65	0.26	0.15	0.01	0.43	1.13	0.68	0.21	0.70	0.52
45	1.31	1.00	1.32	0.47	0.20	0.37	1.99	2.41	1.31	1.46	1.08	1.78
46	2.56	1.77	2.13	1.09	0.33	0.33	2.09	2.59	2.17	1.91	3.35	2.43
47	1.55	1.38	1.65	0.52	0.27	0.22	1.47	2.54	1.33	1.54	1.14	2.41
48	3.15	2.40	2.95	0.82	0.48	0.14	2.05	2.86	1.67	1.47	2.39	2.66
49	1.30	0.86	0.91	0.46	0.13	0.30	1.24	1.14	0.39	0.75	0.56	1.41
50	0.58	0.48	0.39	0.16	0.05	0.31	1.71	1.85	0.81	0.47	0.21	0.74
51	1.27	0.56	0.38	0.33	0.56	0.13	0.79	0.72	0.33	0.26	2.67	0.52
52	1.01	0.97	1.40	0.61	0.37	0.32	1.36	1.50	1.10	0.77	0.95	0.78
53	0.87	0.59	0.58	0.48	0.05	0.55	2.52	2.05	0.88	0.68	0.49	0.68
54	0.88	0.66	0.70	0.33	0.16	0.24	2.44	2.23	1.40	0.93	0.62	0.94
55	2.18	2.21	1.89	1.00	0.29	0.58	3.70	4.12	1.81	2.29	1.74	2.83
56	1.14	0.82	1.14	0.60	0.46	0.40	1.30	1.65	0.83	0.76	0.94	1.14
57	2.70	1.83	2.45	1.11	0.33	0.47	2.52	3.73	2.00	1.67	1.87	2.65
58	1.61	1.16	1.09	0.46	0.39	0.26	2.45	3.23	1.64	1.42	0.99	1.11
59	1.51	1.64	1.72	0.68	0.56	0.26	1.89	2.81	1.53	0.94	1.20	1.44
60	1.58	1.23	1.72	0.85	0.44	0.51	2.65	3.19	1.59	1.59	1.26	1.54
61	0.51	0.44	0.32	0.24	0.05	0.16	1.55	1.24	0.93	0.48	0.22	0.59
62	0.46	0.48	0.52	0.31	0.31	0.36	1.24	1.42	0.85	0.86	0.49	0.61
63	4.07	2.61	3.52	1.50	0.57	0.56	3.29	4.13	2.68	2.27	3.35	3.66
64	0.42	0.24	0.20	0.12	0.04	0.01	0.26	0.49	0.22	0.32	0.17	0.36

Computer output for example 2 is omitted. Sellers and Hill (1974) indicated that there are two wet seasons per year in Arizona. The primary wet season starts in June and ends in September in a broad sense; the secondary wet season extends from December through the middle of March. Spatial patterns of precipitation can be determined by the correspondence analysis. From the factor plane plots (Fig. 2.3), the precipitation is approximately divided into 4 types statewide. In Group 1 (including stations 6, 7, 8, 11, 13, 14, 26, 35, 36, 37, 40, 50, 53, 54, 58, and 61), the primary wet season is a dominant. In this region more than 60% of annual precipitation falls during the monsoon season. The summer moisture normally flows from the Gulf of California and Gulf of Mexico over strongly heated mountainous terrain, which generates convective thundershowers. All Group 1 stations (except for stations 36 and 37, which are located in central Arizona between the forested plateaus to the northeast and the arid desert region to the southwest) are located in southeastern Arizona. This result tallies with Sellers and Hill's (1974) conclusion. Most of Arizona has a second rainy season in midwinter. The correspondence analysis results tell us that Group 2 (stations 3, 9, 12, 15, 16, 17, 18, 22, 23, 28, 41, 43, 44, 46, 47, 48, 49, 57, and 63) weather stations are associated mostly with the winter rainy season. Figure 2.3 shows that secondary wet season (December to February) precipitation is much more important at these stations. This precipitation is generated by the eastward movement of storm systems from the Pacific Ocean, and some of it falls in the form of snow. Most of the stations in this group are located between the northwest and northeast. The third group (stations 2, 5, 31, 39, and 51) is an intermediate type. These stations are associated with November, April and May precipitation (Fig. 2.4) and are located in the northwestern part of Arizona. Station 52 is anomalous; November's





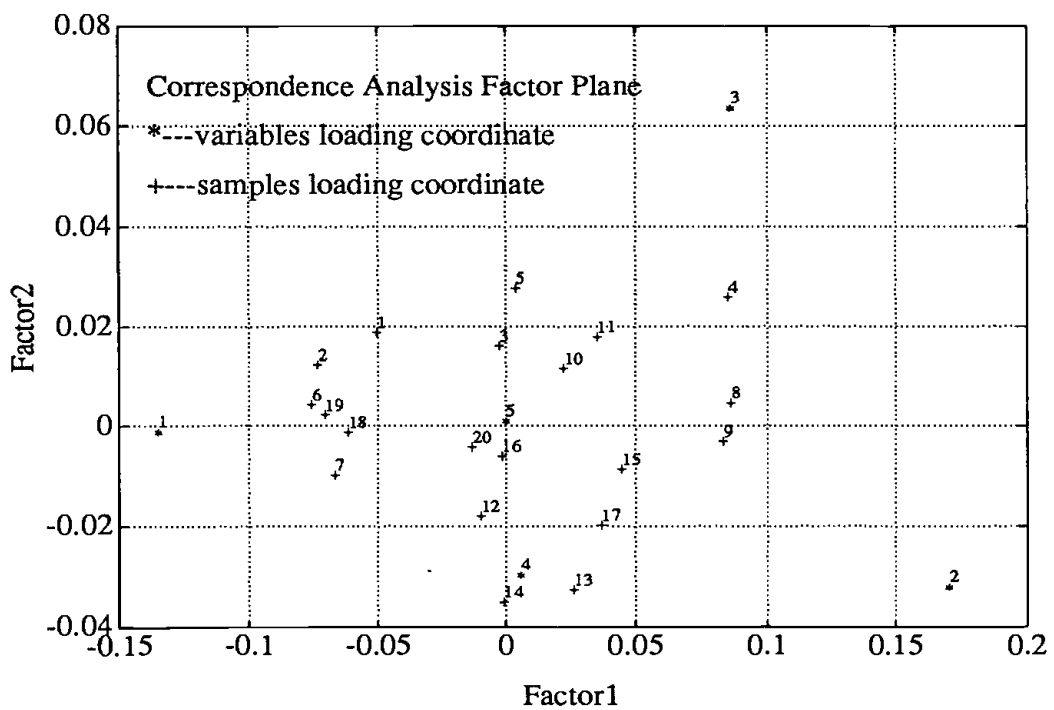
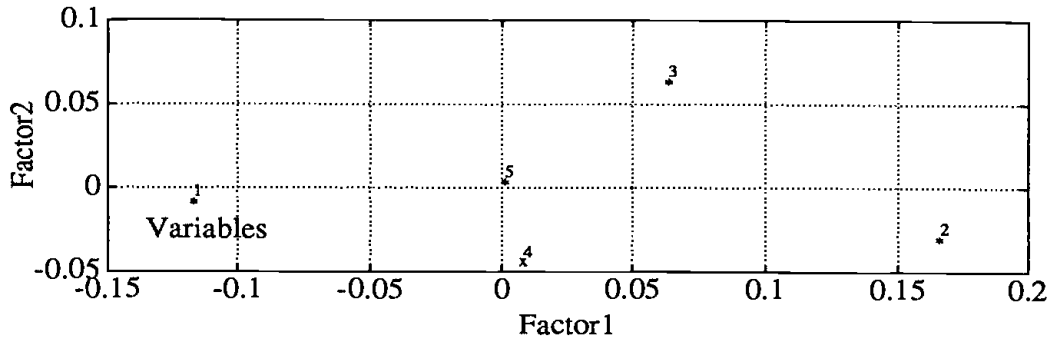
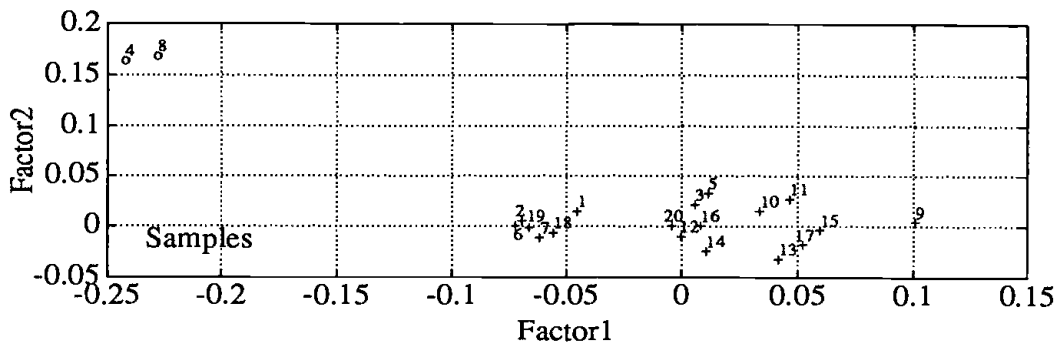


Fig. 2.1 Example 1 Factor Loading and Factor Plane





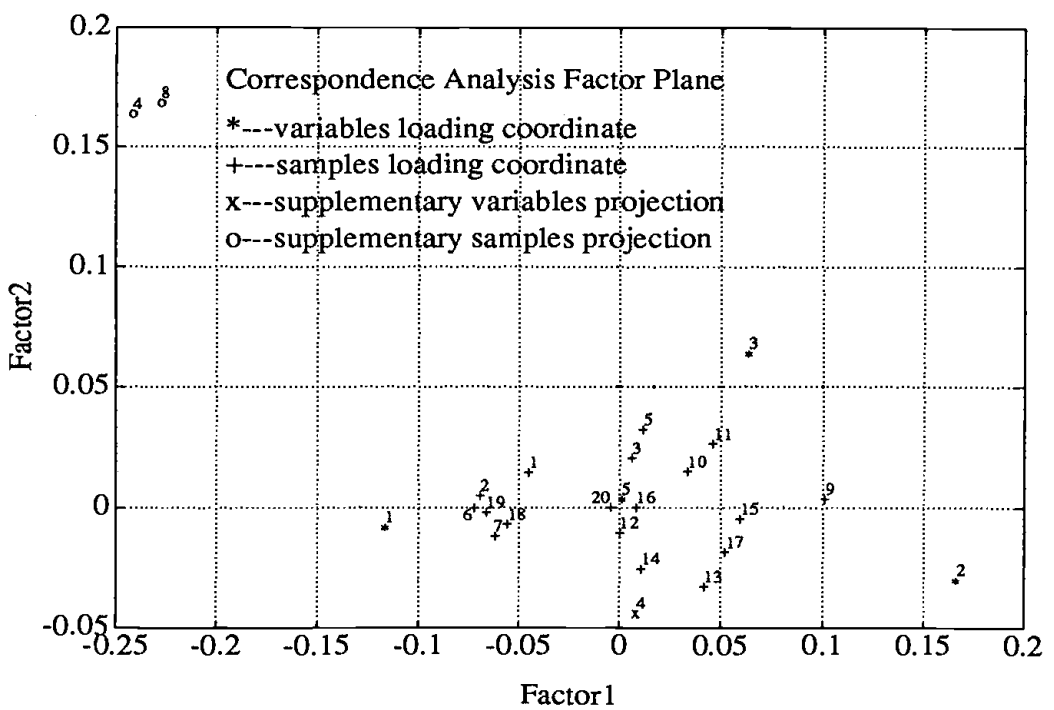


Fig. 2.2 Example 1 Factor Loading and Factor Plane with supplementary elements

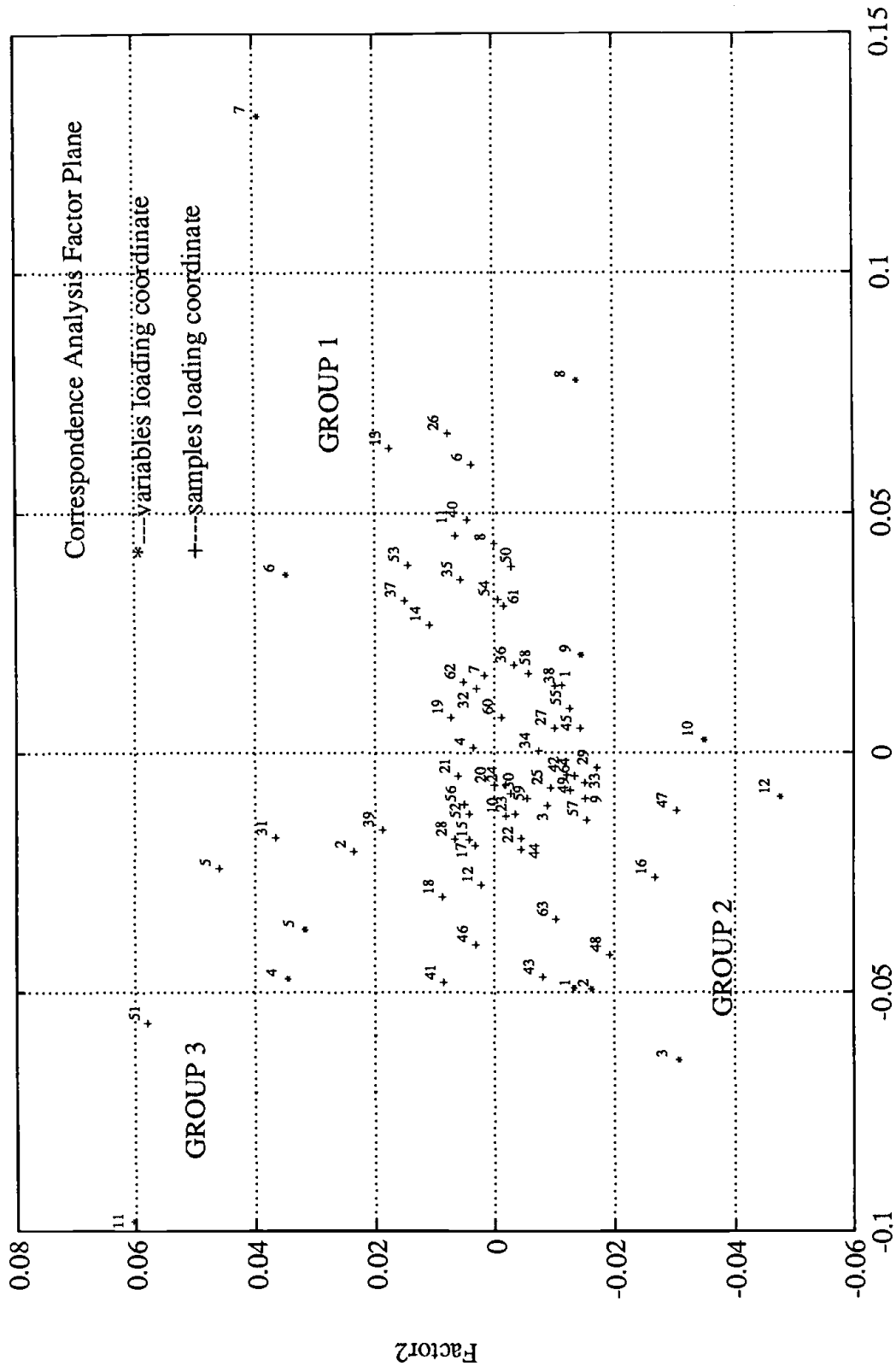


Fig. 2.3 Example 2 Facrot Plane

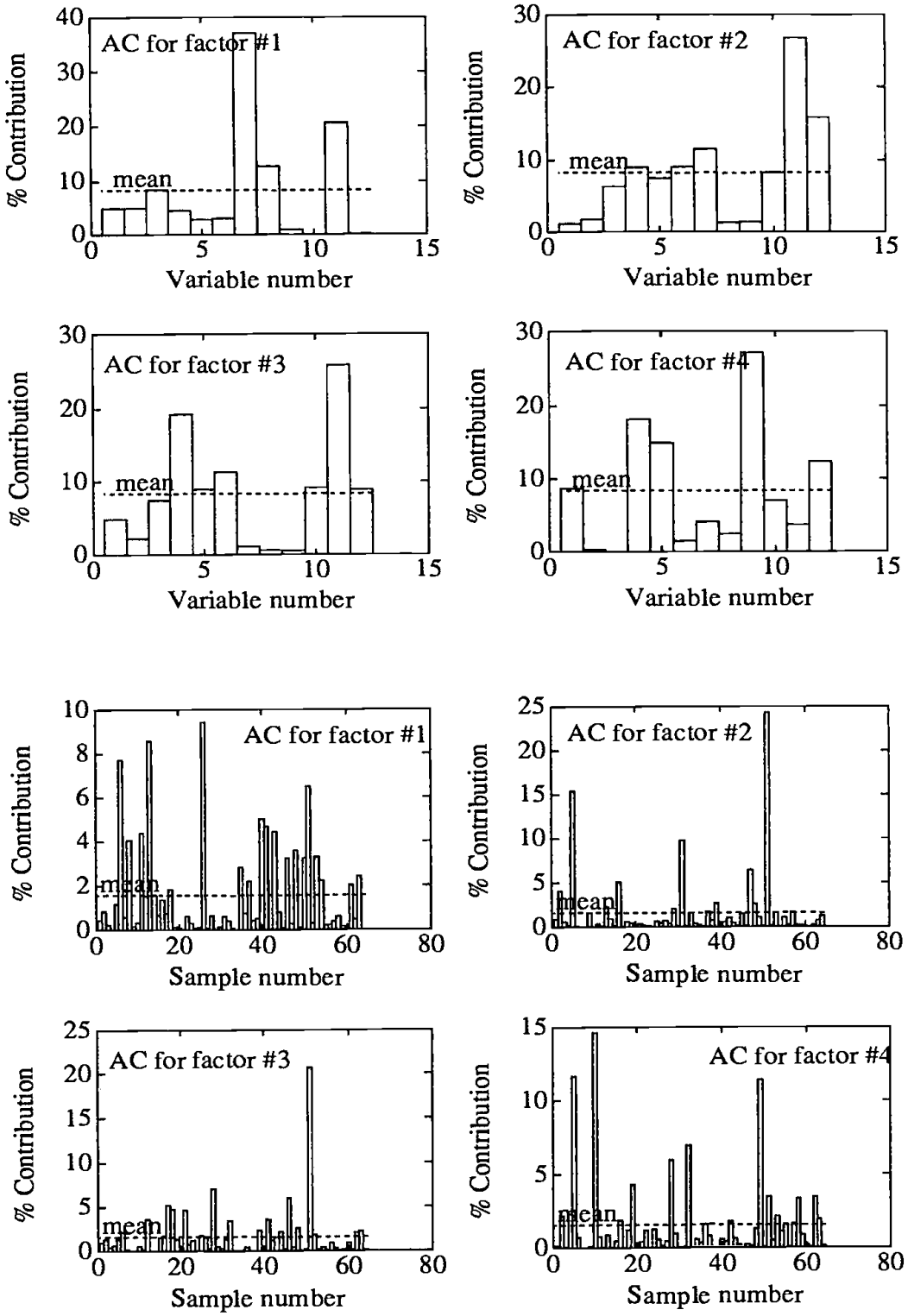


Fig. 2.4 Example 2 Absolute Contributions

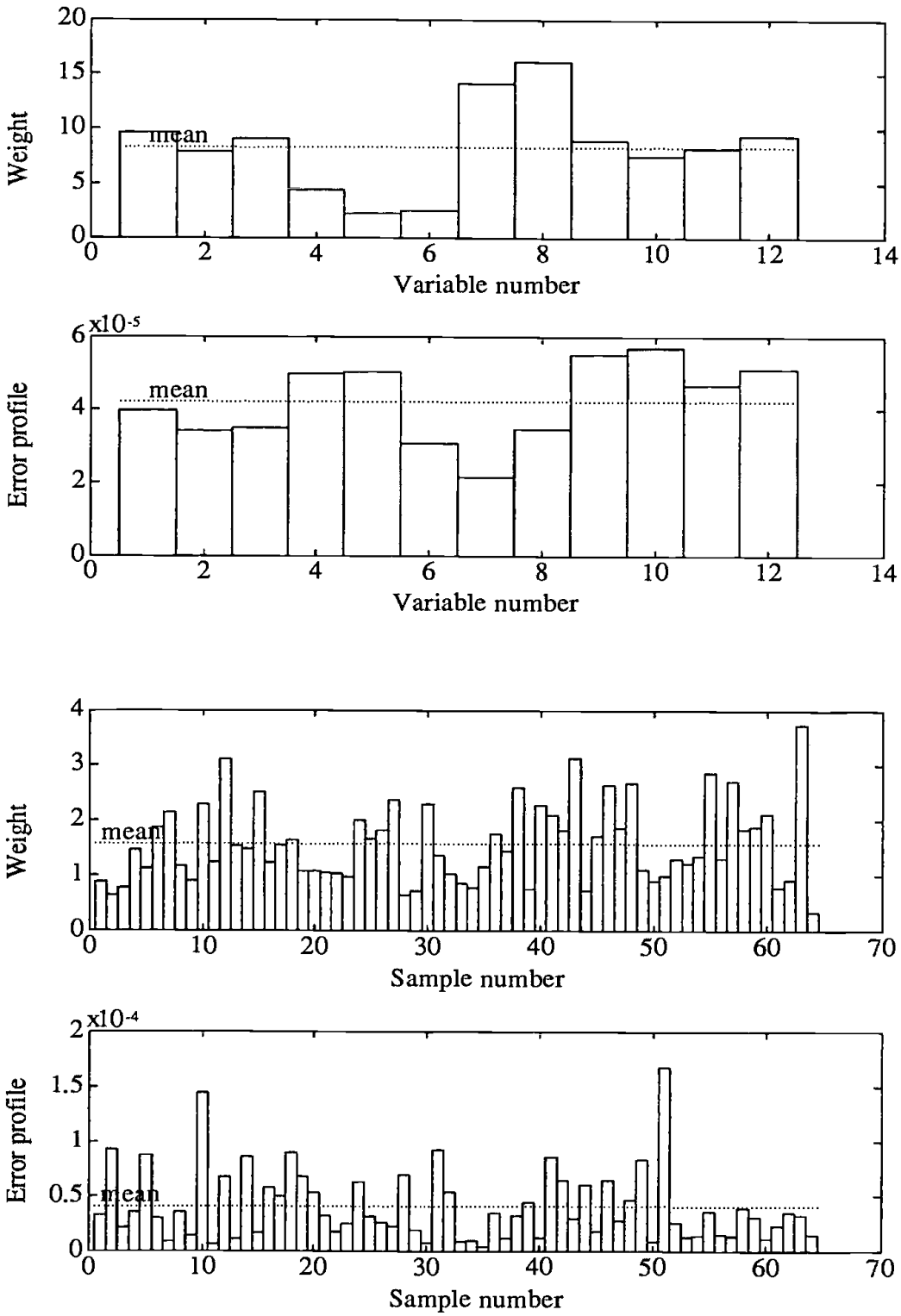


Fig. 2.5 Example 2 Weights and Error Profiles

average precipitation is three times greater than all other monthly averages. The rest of the stations do not present a clear pattern from the view of monthly average rainfall depth. Figure 2.4 are graphics for absolute contributions of the first four factors respect with variables and samples, whereas Figure 2.5 plots variable and sample weights and error profiles, respectively.

2.5 Conclusions

"Graphics are arguably the most important part of the program, since CA is a geometric approach to data analysis" (Goldstein, 1991). MATCORS provides varied and high-quality graphic plot functions. This program works in the user-friendly Matlab environment, which makes the program portable and easy to use. It can operate with any number of variables and with any number of samples. This program allows users to specify any number of variables or samples as supplementary elements, and then will project them on the plots. If these elements are not desired, they can be deleted within Matlab to form a new data matrix.

CHAPTER THREE

RAINFALL SPATIAL AND SEASONAL VARIABILITY ANALYSIS

3.1 Introduction

The seasonal and spatial variability of the rainfall process play an important role in many hydrological applications, such as evaluation of hydraulic balances, management of surface water resources, and real-time runoff forecasting. The rainfall input is obviously the most important factor in rainfall-runoff simulation models in non-snow dominated regions. Prediction errors and bias in rainfall-runoff simulation are the direct results of spatial and seasonal variations of rainfall. Many investigations have shown that the lack of knowledge about the spatial and seasonal variation of precipitation is the greatest source of errors in rainfall-runoff simulation (Niemczynowicz and Jonsson, 1981; Shilling and Harms, 1983; Shilling, 1984 a, b). Therefore, how to choose relevant rainfall input to rainfall-runoff simulation models has become a crucial question to hydrologists. An obvious solution to diminish errors in runoff simulation due to erroneous rainfall input is to increase the density of raingages and sampling intervals. However, this solution requires more time and increased costs.

Precipitation can be regarded as a spatiotemporal variable. This chapter attempts to better describe the characteristics of the rainfall fields both on the seasonal and spatial scales. The selected watershed is Walnut Gulch Experimental Watershed, Tombstone, Arizona, which is operated by U.S. Department of Agriculture, Agricultural Research Service (USDA-ARS) since 1957 (Osborn et.

al., 1979). As Neyman and Scott (1972) reported, "Arizona is not very densely populated either by people or by raingages". It is important to have a good understanding of thunderstorms throughout the southwestern United States. The Walnut Gulch basin is similar to other places of the southwestern area of the United States, where the lower two thirds of the basin is mostly covered with brush, and the remaining higher part is mainly covered by native grasses (Osborn and Lausen, 1973). The area of the watershed is adequate (150 km^2) for studying the structure of convective storms and the location is ideal for this rainfall spatial and seasonal variability study. High quality, long term rainfall data are available at this watershed as well. Numerous previous studies provide valuable information dealing with the spatial and seasonal aspects of convective storms, so the hydrology is well known in this watershed.

A total number of 84 raingages was used in this study. The average density of the raingages is approximately one per square mile (2.6 km^2). The greatest and shortest distances between a pair of raingages are 24.61 and 0.65 km, respectively. The raingages are dual-reverse weighted recording type with a depth resolution of 0.15 mm $\pm$ 0.30 mm approximately. The time resolutions of the raingages are 2.5 min. $\pm$ 5 min. for a 24-hour chart rotation, and 0.5 min. $\pm$ 1 min. for a 6-hour chart rotation (Faures, 1990). The raingage rim is located about one meter above the ground.

Ten-years of records (from 1980 to 1990, except 1989) for these 84 raingages were used in order to understand the spatial and seasonal variability of the precipitation in the region of Walnut Gulch, Arizona. It should be pointed out, for a given storm, gages receiving at least 2.54 mm were included in the data set, but gages receiving less than 2.54 mm were excluded. Because most of the gages

were shut down during the dry season (from January to May), dry season rainfall was also excluded in the data set. A basic statistical analysis was applied to these ten-years of records. Table 3.1 gives the minimum, maximum, average, standard deviation and coefficient of variation of the total rainfall depth in each year. To eliminate the seasonal variation effect, the average rainfall depth during the ten-year period was used. Figure 3.1 illustrates the spatial variability of this average data set, in which the difference in annual rainfall depth between the wet region and dry region can reach as high as 100 mm. Some of this variability could be due to raingages that were turned off for a short time. This study, however, did not check every gage. In order to take into account the seasonal variations of the rainfall intensity, experimental monthly variograms were established based on the entire ten years of data. The exponential model was used to fit these experimental variograms, in which the sill and range were determined by least squares fitting. A detailed analysis for these variogram models was performed in section 3.3, from which the seasonal and spatial patterns were clearly exposed. An overall index method was proposed in section 3.4 to evaluate the spatial variability in a random rainfall field. This overall index can represent not only the rainfall variation speed (spatial gradient), but also the correlation scale and the spatial anisotropy. An average areal rainfall estimation method, linear unbiased minimum variance estimation (Kriging), was applied for the real-time estimation, in which the influence of the seasonal variation was taken into account. Several observation data sets at Walnut Gulch watershed were used to illustrate its application.

Table 3.1 Basic Statistics of Rainfall Depth at Each Year
(Exclude dry season)

Year	Min. (mm)	Max. (mm)	Avg. (mm)	Std. (mm)	CV (%)
1980	62.0	204.7	116.6	29.5	0.26
1981	151.6	283.5	222.1	27.9	0.13
1982	129.3	258.3	185.5	26.8	0.14
1983	225.8	396.5	324.8	36.4	0.11
1984	212.6	411.2	303.9	46.6	0.15
1985	157.0	419.9	258.9	47.1	0.18
1986	84.6	352.8	242.8	42.8	0.18
1987	113.5	240.8	169.3	26.4	0.16
1988	106.9	331.0	214.1	43.0	0.20
1990	192.5	342.4	264.1	27.7	0.10
Total	1958.6	2668.3	2297.2	138.7	0.06

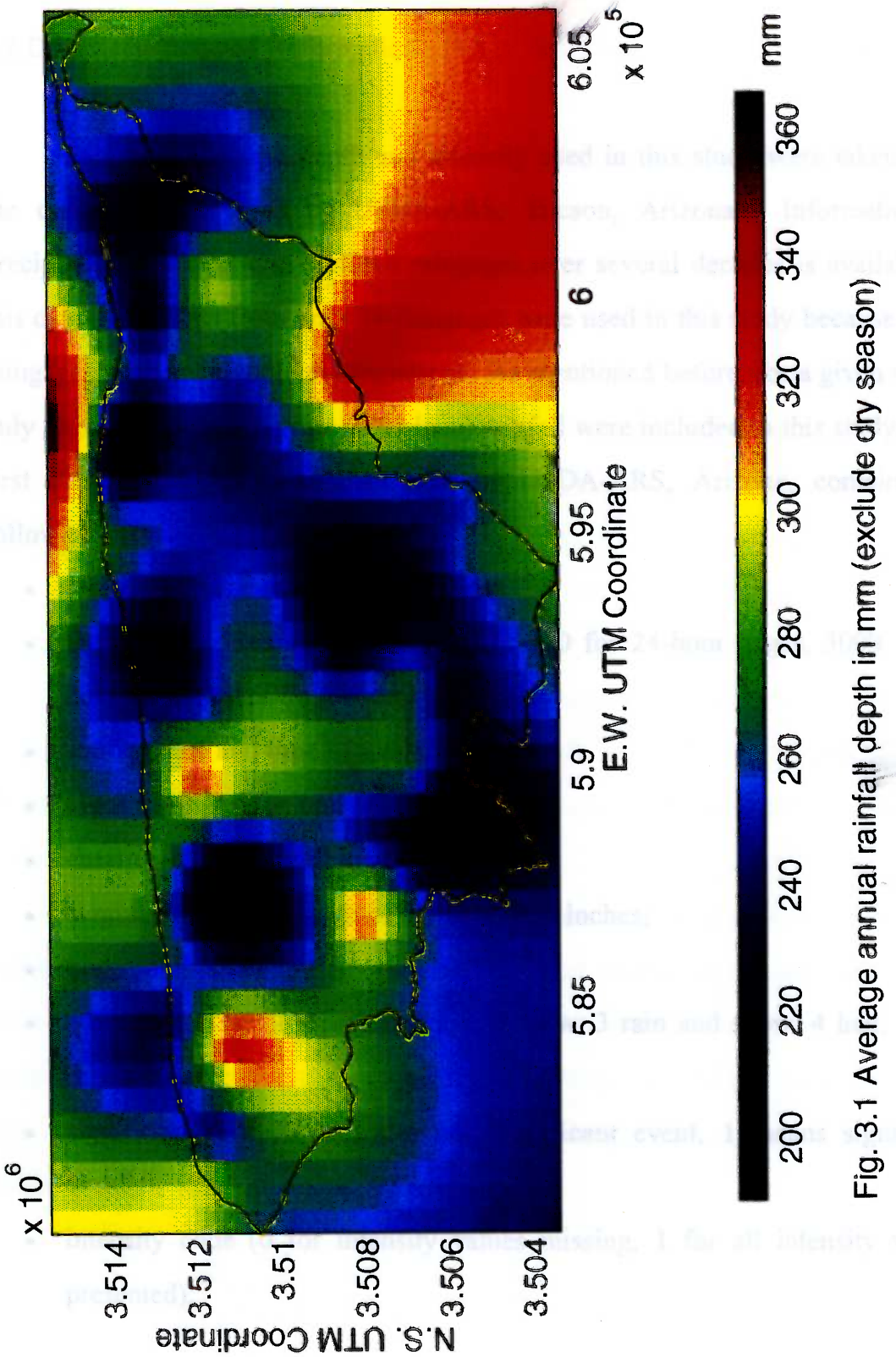


Fig. 3.1 Average annual rainfall depth in mm (exclude dry season)

3.2 Data Reduction and Reformation

Raw data of storm depth and intensity used in this study were taken from the computer data base of USDA-ARS, Tucson, Arizona. Information of precipitation for over one hundred raingages over several decades is available in this data base. Data from only 84-raingages were used in this study because some raingages were not completely registered. As mentioned before, for a given storm, only gages receiving 2.54 mm (0.01 inch) rainfall were included in this study. The first line of the computer data base of USDA-ARS, Arizona, contains the following items:

- the watershed identification number;
- the raingage identification number (1-100 for 24-hour gages, 300's for 6-hour gages and 500's for weekly gages);
- starting date and time of a rainfall event;
- event identification number for each year;
- duration of the rainfall in minutes;
- cumulative or total depth of the rainfall in inches;
- breakpoint number;
- event type code (1 represents rain, 2 snow, 3 rain and snow, 4 hail, and 5 rain and hail);
- significant code (0 indicates non significant event, 1 means significant event);
- intensity code (0 for intensity values missing, 1 for all intensity values presented);

- midnight crossing code (0 for events not crossing midnight, 1 for events crossing midnight);
- maximum depth table code (0 for depth table not computed, 1 for depth table computed); and
- last code 'p' denotes header record.

There are six watershed data in this computer data base with watershed identification numbers of 45 for Safford, AZ; 47 for Albuquerque, NM; 63 for Tombstone (Walnut Gulch), AZ; 64 for Santa Rosa, NM; 73 for Fort Stanton, NM; and 76 for Santa Rita, AZ. In case of missing data, a code following the beginning time will indicate whether the beginning time was estimated or not (0 for not estimated, 1 for estimated).

The second line of the computer data base gives the maximum depth table in inches at the following time duration: 2, 4, 5, 10, 15, 20, 30, 40, 45, 60, and 120 minutes. The remaining lines includes the elapse time, accumulated rainfall, rainfall depth and intensity code, etc..

In order to examine the seasonal characteristics of rainfall at the Walnut Gulch watershed, rainfall hourly intensity in each month for the ten-year period was retrieved. For each storm, the total storm depth was divided by the total storm duration, and expressed in units of mm/hr. This data is called the hourly intensity data. Table 3.2 gives the basic statistics for the monthly averages of the 84 raingages over a ten-year period.

Table 3.2 Basic Statistics of rainfall intensity at each month

Month	Average intensity (mm/hr)	Average standard deviation (mm/hr)	Average Coefficient of variation (%)
January	1.15	0.53	0.33
February	1.78	1.23	0.68
March	1.29	0.94	0.67
April	2.02	1.27	0.59
May	3.24	2.59	0.67
June	6.29	5.77	0.90
July	6.16	6.32	0.99
August	6.35	6.16	0.97
September	5.69	5.90	1.02
October	3.73	3.93	0.87
November	1.16	0.53	0.38
December	1.98	1.34	0.65

As can be seen in Table 3.2, the mean of the coefficient of variation for all raingages during the wet season (July, August, and September) is bigger than that of the other seasons. This phenomenon indicates that the rainfall intensities are more dispersive during the wet season. Figure 3.2 is a scatterplot of standard deviations versus means for the typical months of August, January, and April. These three months represent three typical seasons: the primary wet season, the secondary wet season, and the dry season. The data sets used in Figure 3.2 are the average values of precipitation for ten-years at each raingage. Only 9 raingages were operated during the dry season: gages 4, 12, 41, 43, 45, 59, 67 75, and 76. The scatterplot of the means and the standard deviations for August shows that the means and the standard deviations are almost proportional to one another. This indicates the relationship between the local mean and local standard deviation for the primary wet season is stronger than the other seasons. Figure 3.2 shows that this relationship decreases in an order from the primary wet season, the secondary

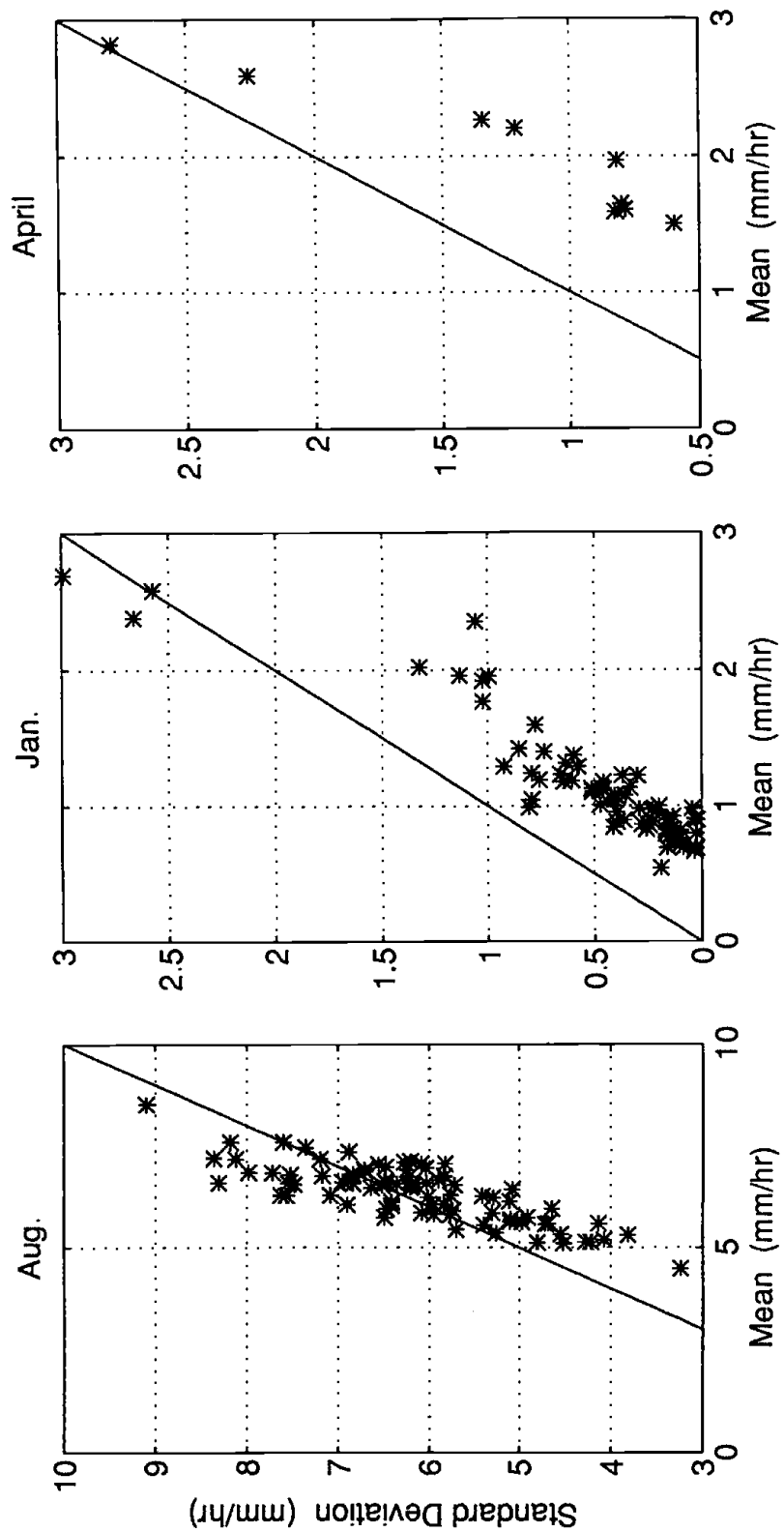


Fig. 3.2 Scatterplot of the means and the standard deviations

wet season, and to the dry season. The relationship between the means and the standard deviations for the dry season also appears nonlinear.

Several AWK (Aho, et al., 1988) programs were written to reduce and reformulate the data sets from the raw data base. AWK is good at manipulation of huge data sets. It generates new data files automatically by matching line patterns such as watershed ID, year, event number and raingage ID by searching the raw data. When a matching line is found, the corresponding action is performed to generate the different form of the data files. The data files include the total rainfall at each gage, the total rainfall intensity for a specific month, or the distribution of a specific rainfall event, etc..

3.3 Rainfall Variogram Establishment

Variogram functions are the foundation of the geostatistical calculation, because they not only describes the space structure of a spatial variable, but also its random characteristics. They are a basic tool peculiar to geostatistics.

By the definition, variogram is the half of a variance of a random variable z(x) at two points x_i and x_j (Journel and Huijbregts, 1978). That is

$$\begin{aligned}\gamma_{ij} &= \frac{1}{2}Var[z(x_i)-z(x_j)] \\ &= \frac{1}{2}E[z(x_i)-z(x_j)]^2-\frac{1}{2}\{E[z(x_i)-z(x_j)]\}^2\end{aligned} \tag{3.1}$$

In traditional geostatistics applications, especially in mining engineering, only one pair of data $z(x_i)$ and $z(x_j)$ can be observed at the location of x_i and x_j. It is impossible to get more than one observation at the same point to get the expectation values in (3.1). In order to overcome this problem, the second order

stationary or the intrinsic stationary assumptions are invoked. This makes $E[z(x_i)-z(x_j)] = 0$; thus the variogram depends only on the space increment h (between point x_i and x_j), not x. Each data pair $\{z(x_i), z(x_{i+h})\}$ (i=1, 2, ..., N(h)) broken up by vector h, can be treated as a different realization of $\{z(x), z(x_h)\}$. Therefore, a variogram can be calculated in practice:

$$\gamma^*(h)=\frac{1}{2N(h)}\sum_{i=1}^{N(h)}[z(x_i)-z(x_i+h)]^2 \qquad (3.2)$$

Rainfall random fields, however, do not always exhibit secondary stationary or even intrinsic stationary behavior. Furthermore, it is really a misnomer to talk about variogram estimates since they are estimates of a value of the variogram for a particular h (Myers, 1991).

The rainfall variograms were established exactly according to their original definition in this study, because multi-realization of rainfall measurements can be obtained at the same observation locations. Expectation values in (3.1) can be found by taking the average of those multi-realizations rather than using the experiment equation (3.2). Another big difference with the traditional variogram calculation method is with the spatial increment vector h. In the traditional variogram calculation method (3.2), vector h is always equally spaced. Normally, the increment Δh is equal to the search radius r. Every measurement point falling within distance r is treated as the same spatial lag. Thus, the traditional variogram calculation method describes the changing trend of a spatial variable along space increment Δh, rather than the exact variance in space. Experimental variograms established in this study used the exact distance between the pairs:

$$\hat{\gamma}^m(d_{ij}) = \frac{1}{2K}\{\sum_{k=i}^{K}[z(k,x_i) - z(k,x_j)]^2 - \frac{1}{K}[\sum_{k=i}^{K}(z(k,x_i) - z(k,x_j))]^2\} \quad (3.3)$$

Such an estimate has been computed for every pair of raingages on a monthly basis at the Walnut Gulch watershed, in which K is the total number of rainfall events received by the both raingages (i and j) according to the rainfall event ID number at the same year in month m (m=1, 2, ..., 12). During the summer, up to 3486 gage pairs ($=\sum_{n=1}^{83} n$) were available for variogram estimation. There were only 45 pairs of measurements during the dry season, because most raingages are shut down as mentioned before. The results are graphically presented in Figure 3.3, in which the variograms represent the ensemble variograms produced from many realizations (many storms) using equation (3.3). For comparative purpose, variograms were also produced for individual storms using equation (3.2). Two individual storm variograms are shown in Figure 3.4, in which the upper portion used widely spaced (1.5 km) raingages while the lower portion used 60 meter apart raingages (see Chapter Four). They show that individual storms have their own characteristics. Generally, the correlation scale of individual storms were smaller than the monthly variogram. Because the ensemble variograms combined the different storms, the correlation scale was extended. The sill varies from storm to storm. Faures (1990) stated that spatial variance of rainfall field exists at a scale smaller than the large network can resolve.

In order to estimate the unknown values of a spatial variable, an experimental variogram is needed to fit a corresponding theoretical variogram model which will directly participate in a linear unbiased minimum variance estimation method (Kriging). Under the assumption of anisotropic condition of

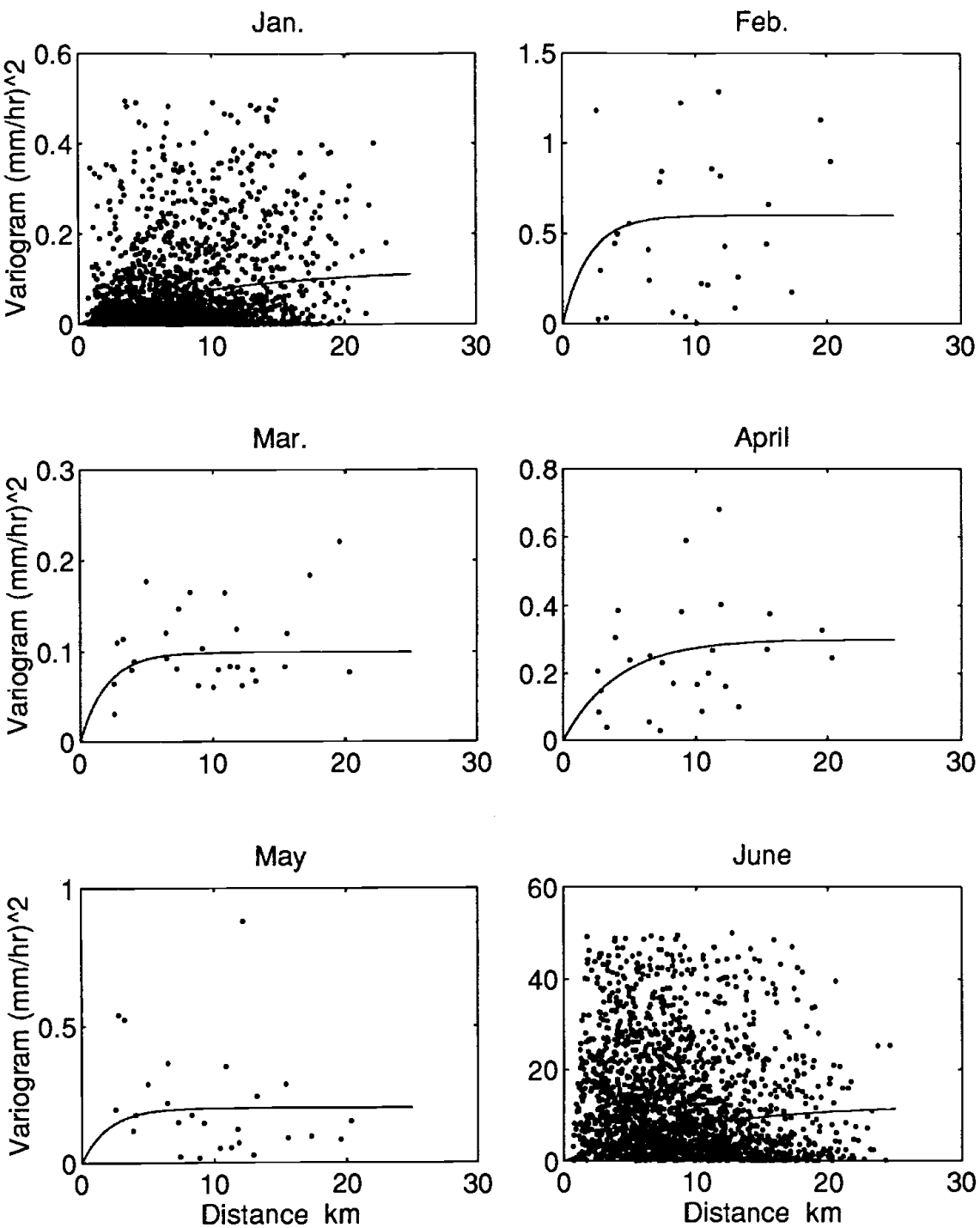


Fig. 3.3 Variogram for each month

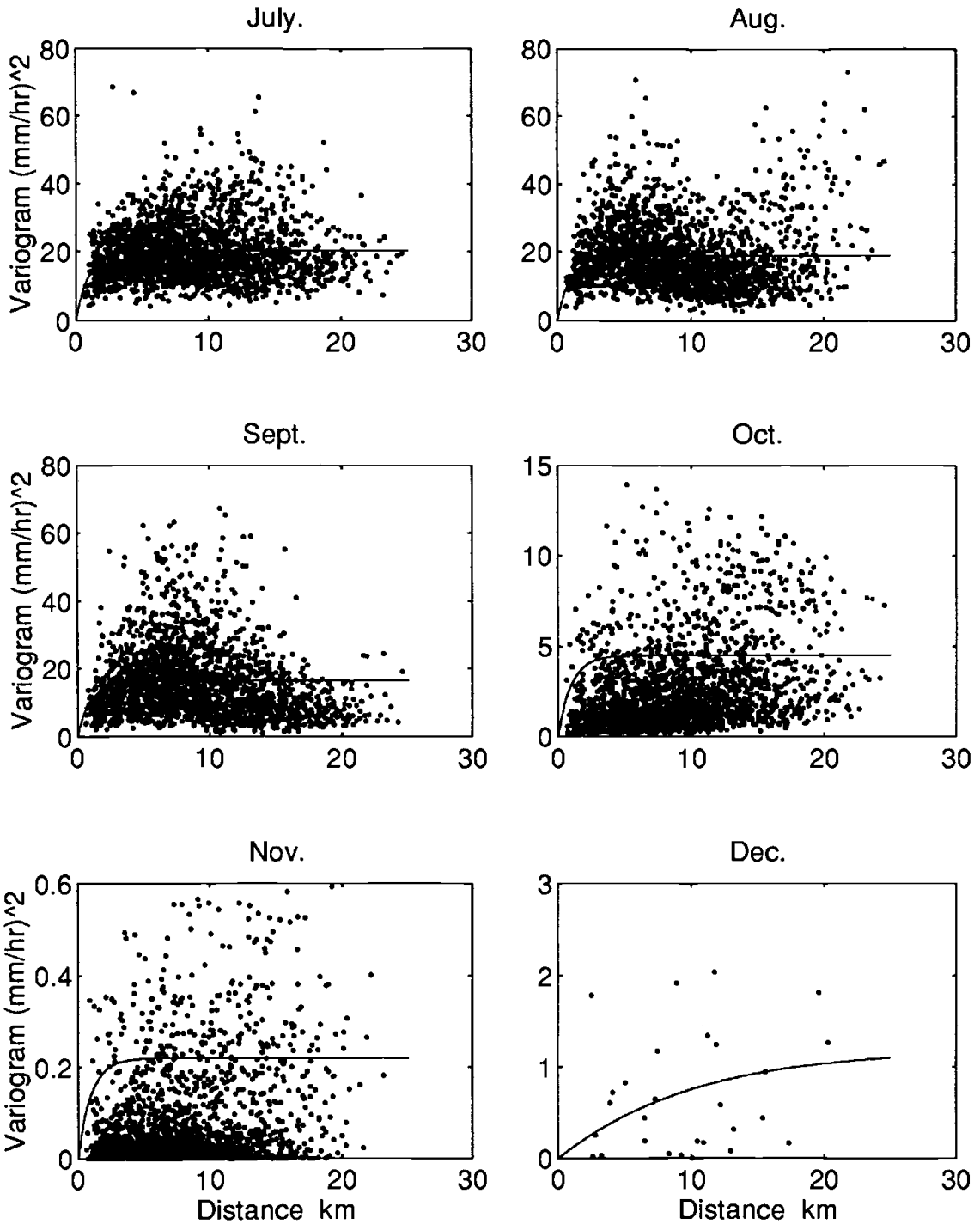


Fig. 3.3 Continued

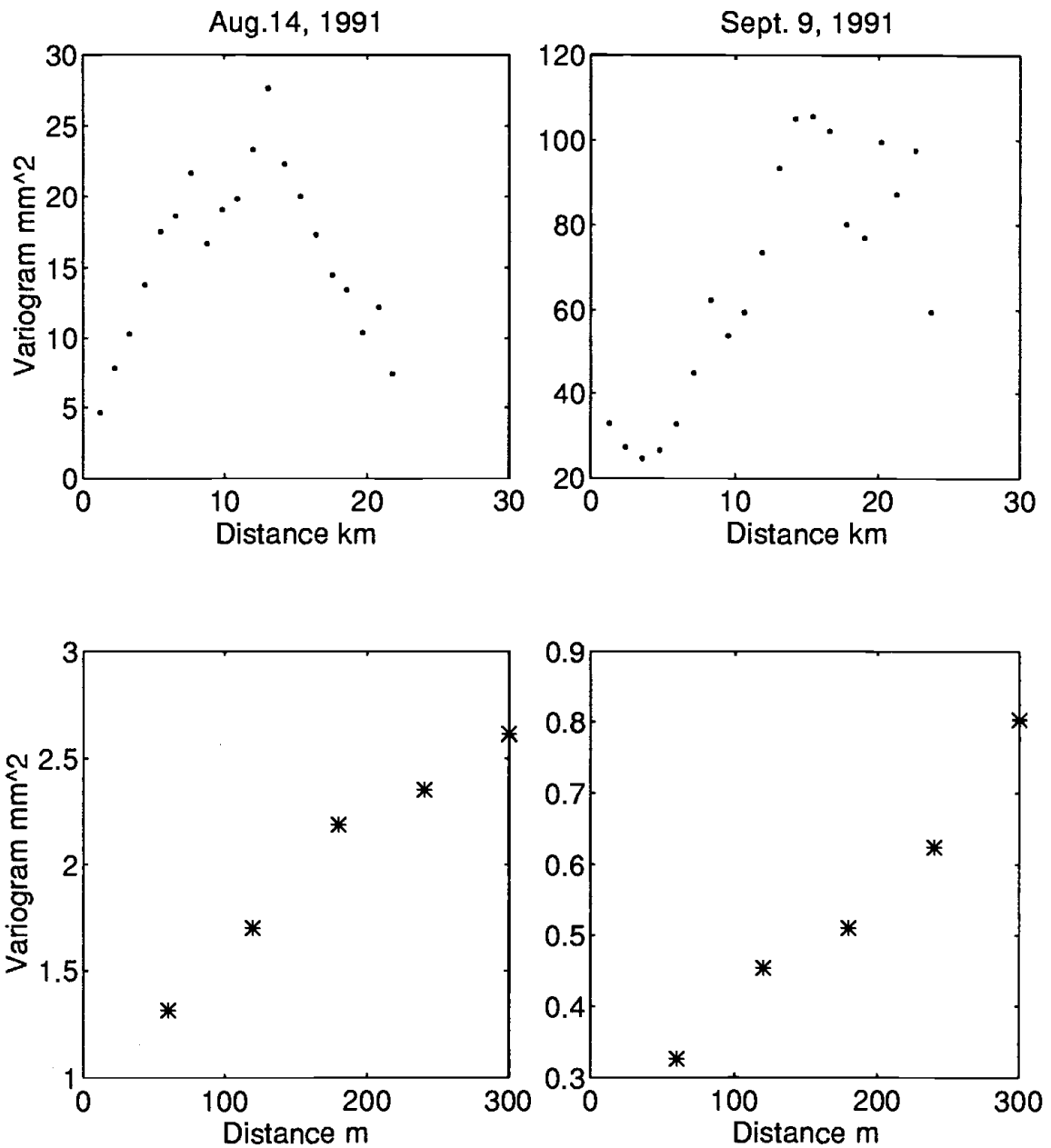


Fig. 3.4 Individual storm variogram at two spatial resolutions

the rainfall field, that is $\gamma(d) = \gamma(|d|) = \gamma(-d)$, a zero nugget constant exponential model was selected:

$$\gamma^m(d) = C(1 - e^{-\frac{3d}{\alpha}}) \qquad (3.4)$$

Where constant C and α are the sill and range, respectively. The range can also be considered the correlation scale. When d is greater than α, the variogram approaches its sill, or variance of the process.

Other forms of theoretical variograms (such as the Gaussian, or Spherical model) could also be used. Bastin and Gevers (1984) indicated that the types of variograms (in the least squares sense) do not necessarily lead to the best linear interpolation (such as Kriging). In addition, Bastin and Gevers (1984) verified thatthe estimates are almost unchanged by the choice of variogram model, provided they have the form $C\gamma(d, \alpha)$.

The variogram function is held in great respect in geostatistics, because it is the foundation of many geostatistical calculation (e.g. estimated variance, discrete variance and regularizing variable's variogram). It can also reflect a lot of important characteristics of a spatial variable. In the exponential model, the variogram increases with the distance $|d|$ between 0 to α. Where the distance $|d| \geq \alpha$, the variogram asymptotically approaches the limit $\gamma(\infty)$. Any data sets that fall into the region with α as radius, and x as center, correlate with z(x) spatially. In the other words, the data at any two points in the region with α as diameter, are always spatially correlated. The degree of correlation decreases as the distance increases between two points; when the distance $|d|$ greater than α, the correlation relationship disappears. Therefore, range (α) can be considered a correlation scale or maximum distance of influence. The sill ($= \gamma(\infty) = \mathrm{Var}[z(x)]$) represents the range of variation a spatial variable.

To identify a exponential variogram model, a non-linear least squares fitting method was used. By using all of the observation data in each month to fit the theoretical model (3.4), the values of Table 3.3 were obtained (for distance d expressed in meters).

Table 3.3 Estimates of variogram parameter

Month	Jan.	Feb.	Mar.	April	May	June
C $(mm/hr)^2$	0.12	0.6	0.1	0.3	0.2	12.42
α (m)	12000	12000	6000	12000	6000	30000
Month	July	Aug.	Sept.	Oct.	Nov.	Dec.
C $(mm/hr)^2$	18.10	18.92	16.48	4.51	0.22	1.2
α (m)	3000	3000	3000	30000	30000	30000

3.4 Rainfall Field Evaluation Method

An overall index of spatial variability has been developed by Sun (1990). Here, it is adopted and applied to evaluate the spatial variability of rainfall fields. This dimensionless index comprehensively respects all the characteristics of rainfall phenomena. It can represent the rainfall spatial gradient, coefficient of variation, and the spatial anisotropy. It is expressed as follows:

$$H = \frac{B^2 R^2}{\alpha^2 m^4 + B^2 R^2} \tag{3.5}$$

Where B : sill (including the nugget constant);

R : radius of rainfall field;

α : range (correlation scale);

m : mean of observation data.

Due to $\alpha^2 m^4 \geq 0$, the denominator is greater than the numerator, therefore H varies between 0 and 1. Because the dimension of rainfall intensity is L/T, the dimension of B^2 is equal to the dimension of m^4 (L^4/T^4); the dimension of the radius of rainfall field (R) is the same as the dimension of the variance distance of variogram. Therefore, H is dimensionless.

The value of sill represents the variance of the rainfall intensity, but the average value m is different in different storms. It is necessary to introduce a value to represent the relative variation for a comparison, value of B/m^2 satisfies the requirement and it is dimensionless. On the other hand, the bigger the correlation scale α, the smaller the spatial gradient of rainfall intensity. In practice, however, α should be considered in association with the radius of the rainfall field, so that R/α is used to represent the relative correlation scale, and it is also dimensionless. The value $BR/\alpha m^2$ can be obtained by combining the above two factors. This combination not only respects the variance, but also the spatial gradient of rainfall intensity, and it is dimensionless. Because of the range of this combination value is [0, ∞], it is not convenient for comparison. So, let

$$x = arctg\frac{BR}{\alpha m^2} \quad \text{and} \quad tg\,x = \frac{BR}{\alpha m^2} \tag{3.6}$$

in which x is in the range [0, $\pi/2$]. In order to get a dimensionless index which varies within [0, 1], let $H = \sin^2 x$. Then equation (3.5) can be derived as follows:

$$H = \sin^2 x = \frac{tg^2 x}{1+tg^2 x} = \frac{B^2R^2}{\alpha^2 m^4 + B^2R^2} \tag{3.7}$$

Compared to the normally-used coefficient of variation, this overall index has many advantages. The overall index H can be built according to the variance and spatial gradient at different directions in a rainfall field, so it can represent

anisotropic characteristics, while the normal coefficient of variation does not represent any spatial meaning. It is very convenient for comparing and classifying the rainfall random field by using the overall index, because it changes within the range of [0, 1]. If the variogram is close to a pure nugget effect type, that means α is close to zero. Then H is close to 100%, which indicates the greatest spatial variation. Figure 3.5 gives the overall index H for all 12 monthly variograms.

There are many potential uses for this overall index. It can efficiently and conveniently evaluate the spatial variability of a rainfall field, as well as optimize a raingage network design. The raingage network is usually designed according to a long-term average in order to get the confidence of the measurement. It, however, may be misleading in an individual rainfall analysis process because small-scale variability is ignored. Taking advantage of the overall index method, a relationship between the overall index and raingage interval as well as network shape can be obtained. Another usage of this overall index method may be in the converting a point rainfall measurement to an areal estimate. The smaller the overall index, the more homogenous the area is.

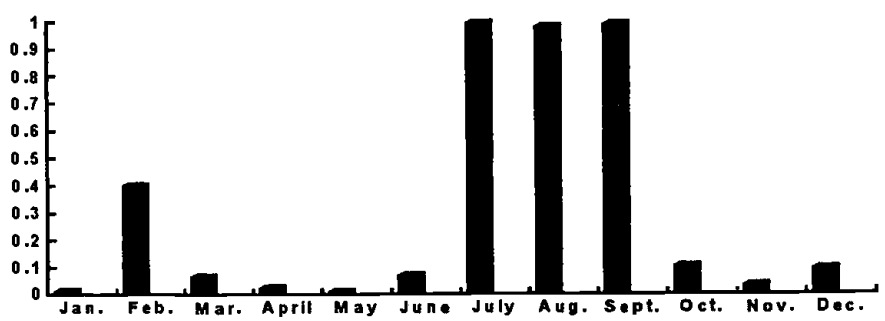


Fig. 3.5 Overall Index

3.5 Areal Rainfall Estimation

An average areal rainfall estimating method, a linear unbiased minimum variance estimation (Kriging), was applied to the problem of estimating real-time rainfall fields based on gage data. The method proposed here is a straightforward extension of the well-known Block Kriging approach (Delfiner and Delhome, 1975; Journel and Huijbregts, 1978; and Delhome, 1978).

The average areal rainfall intensity or depth A(m) is defined as follows:

$$A(m) = \frac{1}{|\Omega|}\int_{\Omega} z_m(x)dx \tag{3.8}$$

Where $|\Omega|$ is the area considered, and $z_m(x)$ is the rainfall at location x. A(m) and $z_m(x)$ may represent either rainfall depth or rainfall intensity over a specific interval of time which is in month m. The variable m (month) is used to stress the seasonality of the variogram. As is well known (Journel and Huijbregts, 1978), if the rainfall field is second order stationary an optimal (unbiased, minimum variance) estimate of A(m) can be computed by:

$$A^*(m) = \sum_{i=1}^{N} \lambda^m z_m(x_i) \tag{3.9}$$

Where $z_m(x_i)$ is observation data at raingage i in month m (i=1, 2, ..., N), and the coefficients λ^m are the solution of the Kriging system:

$$\begin{aligned} &\sum_{j=1}^{N} \lambda_i{}^m \gamma'^m(x_i, x_j) + \mu = \bar{\gamma}'^m(x_i, \Omega) \\ &\sum_{i=1}^{N} \lambda_i{}^m = 1 \quad (i = 1,2,.....N) \end{aligned} \tag{3.10}$$

with μ a Lagrange parameter, and the variance of the estimate is given by:

$$\sigma_m^2 = \sum_{i=1}^{N} \lambda_i{}^m \bar{\gamma}^m(x_i,\Omega) - \bar{\gamma}^m(\Omega,\Omega) + \mu \qquad (3.11)$$

Four by four km square grids were superimposed on the entire Walnut Gulch watershed as shown in Figure 3.6, in which the square grids were named by row from bottom to top, and column from left to right, respectively. The grid size was chosen to be 4 km by 4 km, because the standard data products from the NEXRAD Doppler radar measurements will have a spatial resolution of 4 km by 4 km and a time interval of 1 hour (Michaud and Sorooshian, 1992). NEXRAD's raw measurements have a spatial resolution of 1 km square and a temporal resolution of 15 minutes. As mentioned before, one of the main purposes of this research is to develop a reliable method of estimating the average areal rainfall according to the ground measurement; and to compare them with the radar estimates measured in space. A computer program was written to compute the areal rainfall intensity or depth for each grid. The grid size, of course, can be changed if necessary. To compute $\bar{\gamma}(x_i,\Omega)$ and $\bar{\gamma}(\Omega,\Omega)$, estimation grid Ω was equally divided to several sub-grids, in which all the precipitation properties are assumed to be homogenous. Then, the monthly variogram model established in the previous section was used to find $\bar{\gamma}(\Omega,\Omega)$ which is the average variogram between sub-grids. The variogram between the estimation grid and raingage points are also computed by this sub-grid average method. Obviously, the smaller the sub-grid, the more accurate the estimate. This, however, will require a longer computing time. The sub-grid size used in this study was 400 by 400 meters.

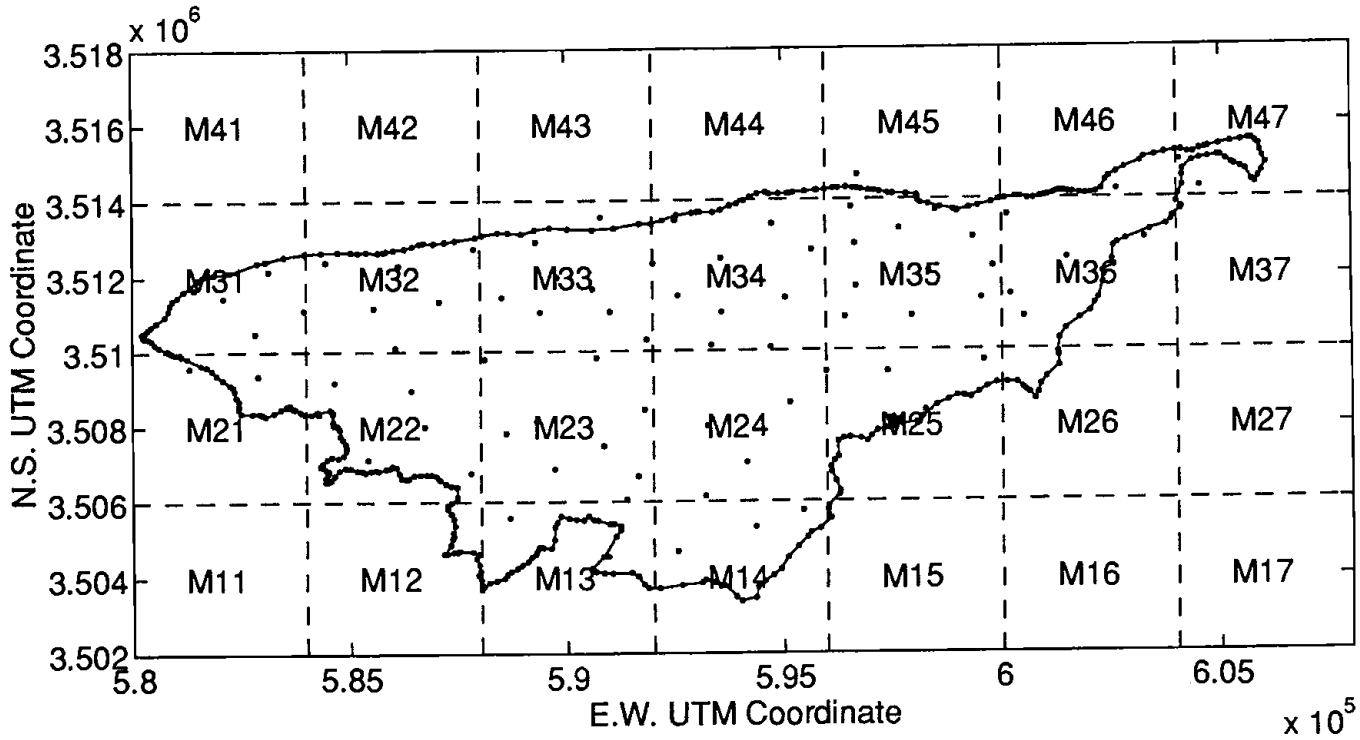


Fig. 3.6 Grid modes with raingage locations for watersheld

As can be found in the next section, the variograms differ only in their scale during the primary wet season. Assuming the exponential model, the variogram may be expressed as ($\gamma = C(1-e^{-\frac{1}{3\alpha}d})$), where C is a scaling factor and $1-e^{-\frac{1}{3\alpha}d}$ is the general variogram formulation. Issaks and Srivastava (1989) pointed out that scaling factor does not affect the ordinary Kriging weights or the ordinary Kriging estimate; however, it affects the variance of the ordinary Kriging estimate. Thus, one can use the general variogram formulation to compute the Kriging weights once and for all, if all the raingages maintain the same geometrical location. Issaks and Srivastava (1989) also found that the Kriging error variance changes by the same factor that was used to scale the variogram, which means the estimation error variance can be simply written as:

$$\ddot{\sigma}_m^2 = C\sigma_m^2 \tag{3.12}$$

with σ_m^2 obtained by using general variogram formulation. This procedure is simple and concise, and can be implemented in real time as well. Once the Kriging weights are found, the areal rainfall intensities or depths and their error variance can be obtained directly from equation (3.9) and (3.12). The next section will give some examples for this real time estimation for the average areal rainfall intensity or depth.

3.6 Results and Conclusions

The theoretical variograms of the form $C(1-e^{-\frac{1}{\alpha}d})$ for each month were computed by least squares fitting for all the data collected in that month during the ten-year period. The results are shown in Table 3.3 and Figure 3.7, and they clearly show the seasonal patterns of the rainfall. An assumption is made here that

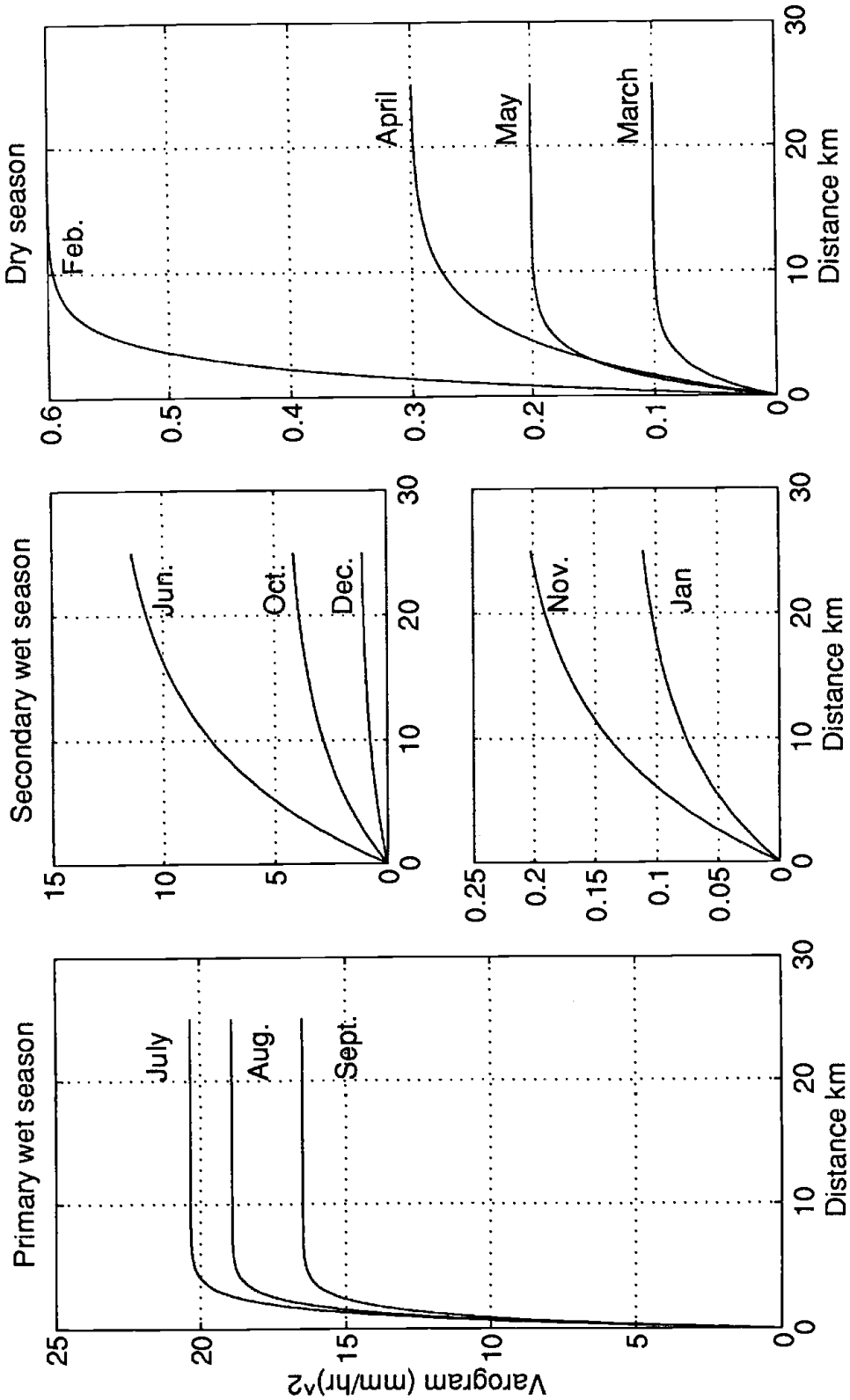


Fig. 3.7 Exponential monthly variogram models

the variogram is time-invariant during a month, but not necessarily from one month to another. As can be seen in the following, this assumption is reasonable; it takes into account the seasonal trends of the rainfall phenomenon and would yield a different estimated variance of the average areal rainfall intensity or depth for different months.

Figure 3.7 clearly shows the seasonal behavior of the spatial variability of the rainfall process in Walnut Gulch watershed: the variogram appears much larger in wet than dry seasons. During the primary wet season (July, August, September), the variograms are directly proportional to the average rainfall intensity, and the ranges are fixed at 3 km. Thus, the variogram models can be separated into two factors : C(m), which is time-varying but space-invariant, and $(1-e^{-\frac{3}{3000}d})$, which is time-invariant but space dependent. The range of 3 km is the smallest one compared to the other monthly variograms, so rainfall intensity during the primary wet season has the biggest spatial variation. Figure 3.5 verifies this finding. To eliminate the directly proportional effect, a stable and universal rainfall intensity variogram can be obtained for the primary wet season:

$$\gamma_w(d) = CKm^*(1-e^{-0.001d}) \tag{3.13}$$

where m* is the average rainfall intensity, and CK is the proportional constant (here equal to 2.938).

Another way to remove this directly proportional effect is to use a relative variogram function, which can be obtained through a variogram divided by a proportional constant.

During the secondary wet season and dry season, however, rainfall intensity did not show a uniform spatial relationship. This is mainly caused by the variant climatic conditions and the atmospheric moisture sources in those two seasons. Observations by Osborn et al. (1979) point out that the summer rainfall at Walnut Gulch stems from air-mass thunderstorms while precipitation during the winter season results from frontal storm systems. These summer air-mass thunderstorms typically result from intensive daytime heating of the land surface. Fennessey (1986) assumed that the air-mass thunderstorms in this area are horizontally motionless in space. Our results show that summer rainfall has the smallest correlation scale. This finding is supported by Fennessey's assumption and the work of Fletcher (1960) who found that air-mass thunderstorms in Arizona have little transitional motion. It is difficult to characterize the precipitation in secondary wet season and dry season due to the storm movements. In addition, the average duration of winter precipitation is twice as long as summer storms. The correlation scale in the secondary wet season changed from 12 kilometers to 30 kilometers (Fig. 3.7), indicating the storms were correlated within a large region.

A computer program was developed to estimate the areal rainfall intensity or depth, its estimated variance and their relative errors in each grid. Rainfall intensity from the storm on August 2, 1991, as an example, was used to compute the estimates in each grid. The results are shown graphically on Figure 3.8, from which the comparison between point measurements and block estimate can be easily obtained. The block estimates are much smoother than the original point measurements, and closer to the average intensity of the whole watershed. Three different raingage densities were used here to compute their block estimates, and their estimated variance and relative errors as listed in Table 3.4, in which the grid

Table 3.4 Estimated results at different raingage densities. For a given gage density, each block in the table represents a different grid (see Fig. 3.6) . The top left block is pixel M41; the bottom right block is pixel M17.

a). 43 raingages used

A*	18.8363	18.8430	18.8631	18.8655	18.7485	18.8361	17.6056
σ	0.9168	0.9167	0.9161	0.9134	0.9153	0.9167	0.8966
σ/A*	0.0487	0.0486	0.0486	0.0484	0.0488	0.0487	0.0509
A*	16.0189	27.4815	22.1528	16.4610	18.8225	18.7075	18.8346
σ	1.2449	0.8556	0.8447	0.8586	0.9164	0.9148	0.9167
σ/A*	0.0492	0.0311	0.0381	0.0522	0.0487	0.0489	0.0487
A*	18.6301	19.0747	18.1676	18.7880	18.8369	18.8369	18.8369
σ	0.9117	0.8542	0.8674	0.9158	0.9168	0.9168	0.9168
σ/A*	0.0489	0.0448	0.0477	0.0487	0.0487	0.0487	0.0487
A*	18.8345	17.8687	18.8078	18.8369	18.8369	18.8369	18.8369
σ	0.9167	0.8913	0.9163	0.9168	0.9168	0.9168	0.9168
σ/A*	0.0487	0.0499	0.0487	0.0487	0.0487	0.0487	0.0487

b). 22 raingages used

A*	18.0694	18.0738	18.0976	18.0609	17.9898	18.0687	16.9011
σ	1.3192	1.3191	1.3182	1.3189	1.3170	1.3191	1.2881
σ/A*	0.0730	0.0730	0.0728	0.0730	0.0732	0.0730	0.0762
A*	16.1927	23.8919	19.8439	17.0429	18.0690	17.9467	18.0673
σ	1.2673	1.2631	1.2589	1.2742	1.3191	1.3161	1.3191
σ/A*	0.0783	0.0529	0.0634	0.0748	0.0730	0.0733	0.0730
A*	17.9216	18.9746	17.6791	18.0378	18.0695	18.0695	18.0695
σ	1.3132	1.2748	1.2722	1.3183	1.3192	1.3192	1.3192
σ/A*	0.0733	0.0672	0.0720	0.0731	0.0730	0.0730	0.0730
A*	18.0684	18.1774	18.0553	18.0695	18.0695	18.0695	18.0695
σ	1.3192	1.3165	1.3188	1.3192	1.3192	1.3192	1.3192
σ/A*	0.0730	0.0724	0.0730	0.0730	0.0730	0.0730	0.0730

c). 11 raingages used

A*	12.9581	12.9581	12.9804	12.9531	12.9097	12.9581	12.9581
σ	1.8618	1.8618	1.8611	1.8615	1.8587	1.8618	1.8618
σ/A*	0.1437	0.1437	0.1434	0.1437	0.1440	0.1437	0.1437
A*	12.7179	13.0095	13.4125	12.9580	12.9580	12.9581	12.9581
σ	1.8297	1.8558	1.8377	1.8618	1.8618	1.8618	1.8618
σ/A*	0.1439	0.1427	0.1377	0.1449	0.1437	0.1437	0.1437
A*	12.8478	14.4568	12.3876	12.9393	12.9581	12.9581	12.9581
σ	1.8564	1.7976	1.8263	1.8607	1.8618	1.8618	1.8618
σ/A*	0.1445	0.1243	0.1474	0.1438	0.1437	0.1437	0.1437
A*	12.9581	12.9568	12.9490	12.9581	12.9581	12.9581	12.9581
σ	1.8618	1.8617	1.8613	1.8618	1.8618	1.8618	1.8618
σ/A*	0.1437	0.1437	0.1437	0.1437	0.1437	0.1437	0.1437

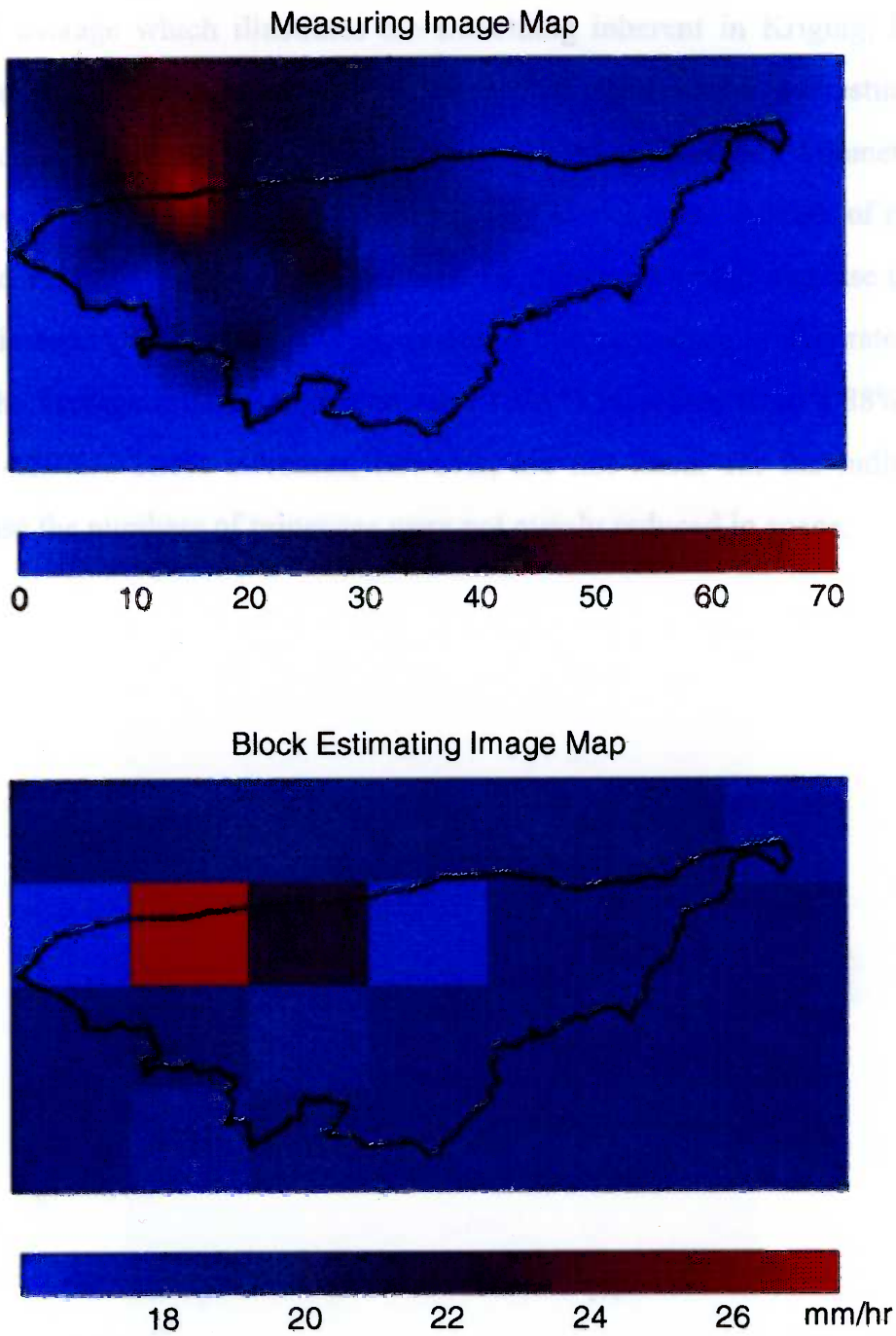


Fig. 3.8 Comparison between measuring image and block estimating image

numbers are the same as Figure 3.6. All of the block estimates close to their global average which illustrates the smoothing inherent in Kriging; high values tend to be underestimated while low values tend to be overestimated. As mentioned before, the raingages are spaced about one and half kilometer apart on the average when 43 raingages are used. By reducing the number of raingages to 22 and 11, the average interval between the raingages would increase to three and six kilometers, respectively. The results of this procedure is illustrated in Figure 3.9, the average relative estimated error (σ/A^*) increases from 4.88% to 7.21%, and 14.30%. These increases, however, are not linear for the individual grid, because the numbers of raingages were not evenly reduced in space.

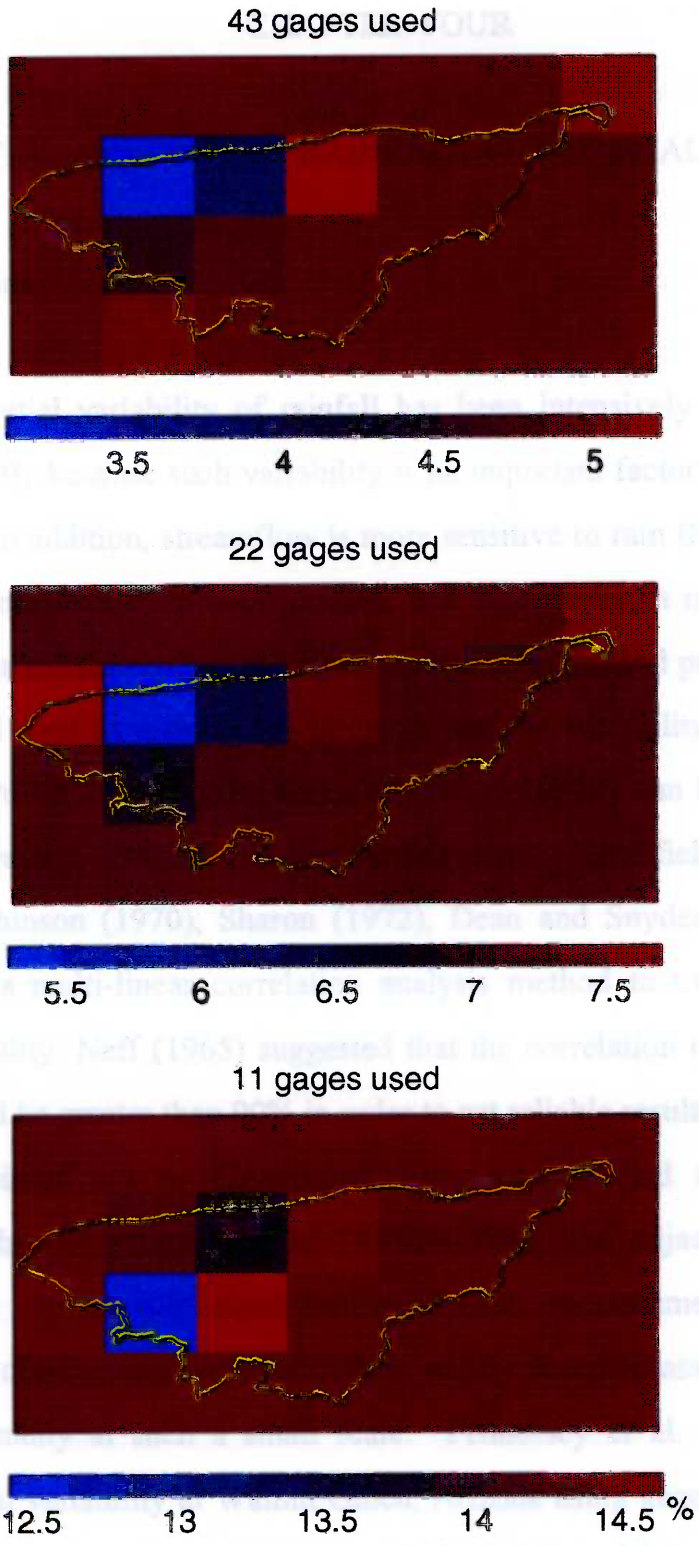


Fig. 3.9 Relative estimated errors at different raingage densities

CHAPTER FOUR

SPATIAL ANALYSIS OF RAINFALL AT THE SMALL SCALE

4.1 Introduction

The spatial variability of rainfall has been intensively investigated for a long time, partly because such variability is an important factor in rainfall - runoff predictions. In addition, streamflow is more sensitive to rain than other variables, such as soil moisture, evapotranspiration, and topography in many hydroclimatic regions. The standard error of the mean precipitation method proposed by Rycroft (1947) is still used as a reference in comparing the variability between stations. Conrad and Pollak (1950) stated that a relative variability can be judged from the mean and average deviation of the rainfall data. Hershfield (1965), Matalas (1967), Hutchinson (1970), Sharon (1972), Dean and Snyder (1977), and Shin (1982) used a multi-linear correlation analysis method to evaluate the rainfall spatial variability. Neff (1965) suggested that the correlation coefficient between gauges should be greater than 90% in order to get reliable results.

As pointed out by Goodrich (1990), who studied the behavior of a distributed physically based model (KINEROSR), two adjacent raingages 300 meters apart can provide significantly different measurements of convective rainfall. His observations have refuted the widely accepted assumption of rainfall field homogeneity at such a small scale. Fennessey et al. (1986) studied the rainfall spatial variability at Walnut Gulch, Arizona using gages spaced from 1 to 1.5 kilometers apart. Faures (1990) studied rainfall variability at the same sub-

watershed as Goodrich (1990), using raingages that were 30 meters apart; he found that spatial variations of the rainfall field exist at a scale smaller than the large network can resolve. There was no information, however, about variability at a scale intermediate from 30 meter to 1000 meter. For a better understanding of the multi-scale behavior of rainfall, Faures (1990) suggested that a field study involving such an intermediate network is required.

The purpose of this chapter is to provide information about rainfall variability at a scale that has not been studied before, at least in Arizona. The objective of this chapter is to closely examine the rainfall spatial variability at a extremely small scale (30-180m) and to better characterize the rainfall spatial structure at such small distances. This study also attempts to get some first-hand information toward developing general description of the spatial characteristics at the small scale field. Primarily, this chapter addresses:

1. the relationship of rainfall spatial variability at different scales
2. the variance of the local average process

4.2 Experimental Design

As mentioned above, there are not many studies addressing the spatial variability of rainfall at very small scales. Because the dense networks of raingages are very expensive, a single gage is often used to represent the average rainfall in small watersheds. It is often questionable to use such measurements to represent the entire small watersheds. A very dense network seems to be the best way to establish such small scale variogram structures.

A total of 54 non-recording raingages were installed along a gas pipeline road in the Lucky Hills area of Walnut Gulch Watershed in Arizona. The starting point was at the Flume 4 road turn-off and the ending point at Lucky road turn-off, respectively. Figure 4.1 gives the location of the raingage network, while Figure 4.2 is the cross-section of the raingage network. There were 15 non-recording raingages in the gully area with 30 meter spacing; the remaining 26 raingages had 60 meter spacing. Thirteen tilted non-recording raingages were co-located with vertical gages in the gully area. The tilted gages were perpendicular to the mean slope and also were spaced at 30 meters. The rims of both tilted and non-tilted raingages were approximately one meter from the ground.

All of these non-recording raingages had a 16.92 cm^2 rain aperture area and were similar to those described by Kalma et al. (1969). Variations due to manufacture can be assumed to be negligible. Readings were made by transferring the water from the raingages into a graduated cylinder, because the scale on the exterior of the gages was not considered sufficiently accurate. Even though this process induces a loss due to the water adhering to the walls of the gages, the error is systematic, i.e. linearly proportional to the amount of rainfall received for all non-recording gages. Spatial heterogeneity, therefore, can still be assessed (Faures, 1990).

Two graduated cylinders were used for two different ranges of rainfall depth. Their reading accuracy was estimated as follows:

- From 0.0 mm to 15.0 mm: 0.07 mm $\pm$ 0.15 mm,
- From 15.1 mm to 60.0 mm: 0.15 mm $\pm$ 0.30 mm

Measurements of the non-recording raingages were made after each storm if possible. Losses by evaporation are assumed to be negligible.

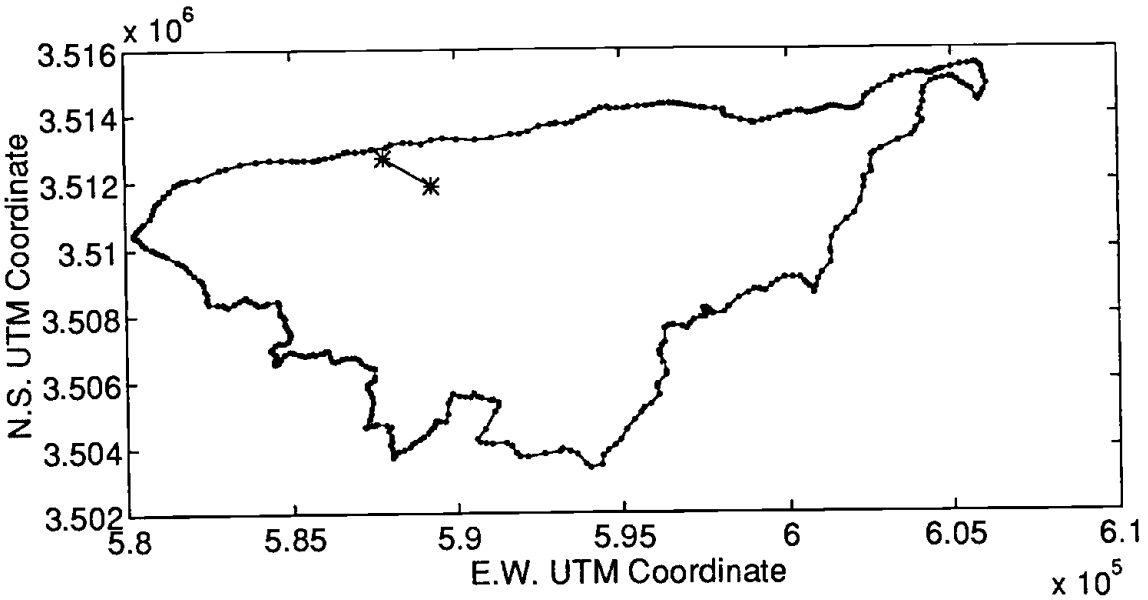


Fig. 4.1 Experimental Network Location

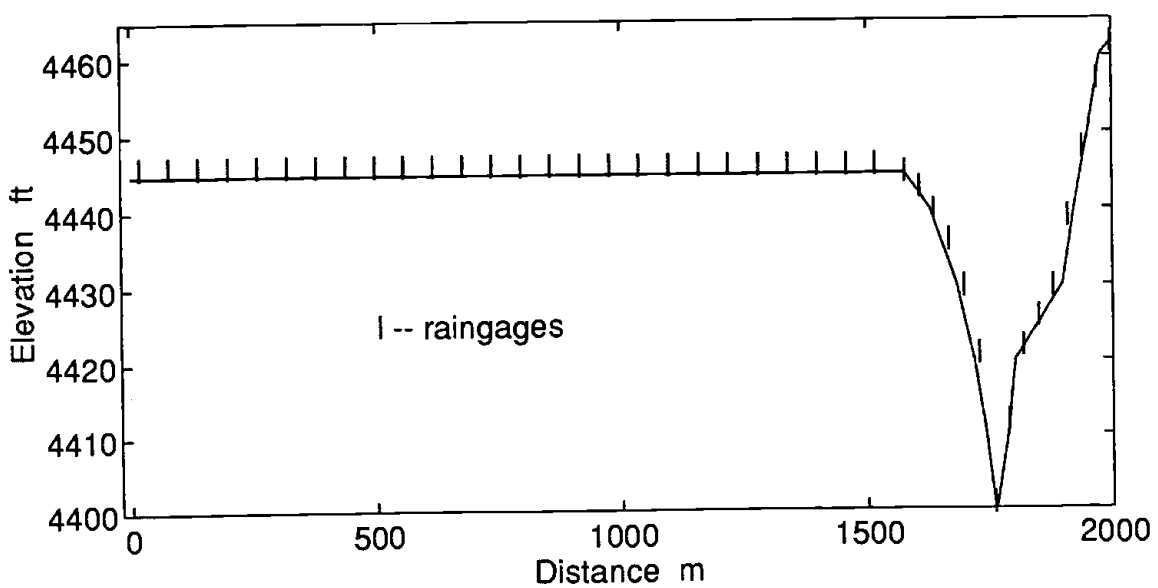


Fig. 4.2 Cross Section of Network

4.3 Results and Analysis

Precipitation data (total rainfall depth) sets of 5 summer thunderstorms (from August 2 to September 9, 1991) are listed in Appendix D in mm. Appendix E contains the total rainfall depth recorded by the tilted raingages in the gully area.

Geostatistical analysis performed here is used to investigate the spatial characteristics of the rainfall field at small scales. The theoretical basis and a description of the terms used below are described in Journel and Huijbregts (1978).

Based on experimental work at the same watershed, Faures (1990) concluded that elevation and vegetation index do not play a significant role in the distribution of rainfall over the 4.4 hectares Lucky Hill 106 sub-watershed. It was then assumed that variations in rainfall measurements between gages are caused by two factors: measurement errors and spatial variations in rainfall field. Statistics, including the sample mean, standard deviation and coefficient of variation, were computed for these 5 sub-sets of the data (Table 4.1), for 4 sampling intervals. Statistics were computed at 30, 60, 120, and 180 m spacing. In general, the average standard deviation and coefficient of variation increases as the sampling interval increases (Table 4.1). This result indicates the sensitivity of the sample statistics to the sampling intervals. Such dispersion should be taken into account when areal rainfall depth is estimated by the point measurements. The coefficient of variation was plotted against sampling distance in Figure 4.3 for each storm.

Table 4.1 Statistics at different sampling interval for 5 storms (mm)

	Aug.2	Aug.9	Aug.14	Aug.27	Sept.9
a. Raingage interval = 30 m (15 sample points in gully area)					
Mean	36.72	19.29	9.69	34.07	11.58
Std	2.66	1.14	1.31	1.75	0.42
C.V.	0.07	0.06	0.14	0.05	0.04
b. Raingage interval = 60 m (34 sample points including gully area)					
Mean	31.18	16.15	12.03	30.77	13.01
Std	5.48	2.31	2.77	2.69	1.22
C.V.	0.18	0.14	0.23	0.09	0.09
c. Raingage interval = 120 m (17 sample points including gully area)					
Mean	31.66	16.26	12.07	30.77	13.01
Std	5.62	2.39	3.05	2.22	1.26
C.V.	0.18	0.15	0.25	0.07	0.10
d. Raingage interval =180 m (8 sample points including gully area)					
Mean	31.57	16.32	12.08	30.99	12.85
Std	6.50	2.46	3.10	2.98	1.44
C.V.	0.21	0.15	0.26	0.10	0.11
e. Tilted raingage with interval = 30 m (13 sample points in gully area)					
Mean	34.70	19.23	9.56	33.99	11.32
Std	4.81	1.20	0.91	1.29	0.64
C.V.	0.14	0.06	0.10	0.04	0.06

The results also show that on average tilted raingages caught about 2.25% less rainfall than the vertical raingages. For the tilted raingages, standard deviations and coefficient of variation were larger than vertical raingages. The

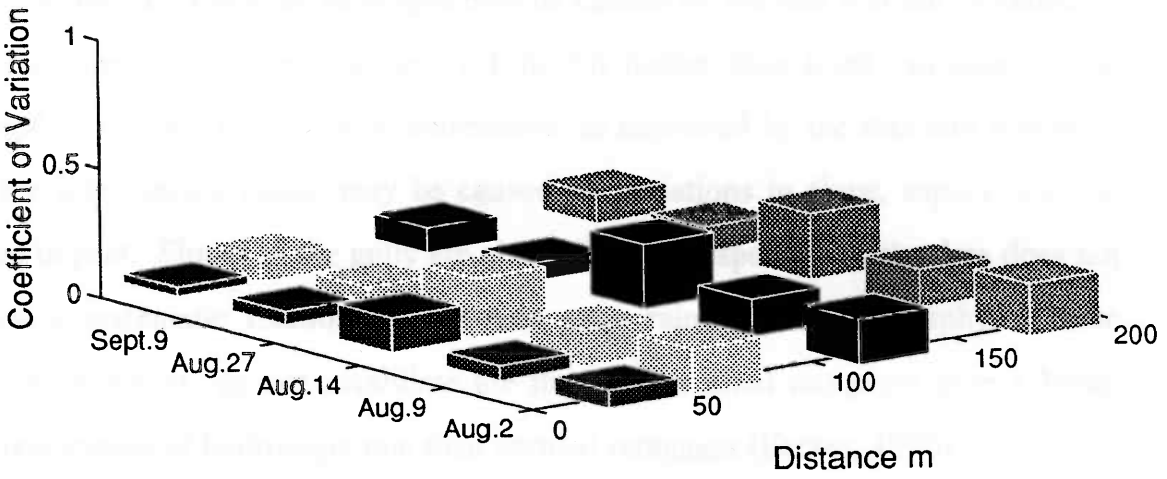


Fig. 4.3 Coefficient of Variation at different storms

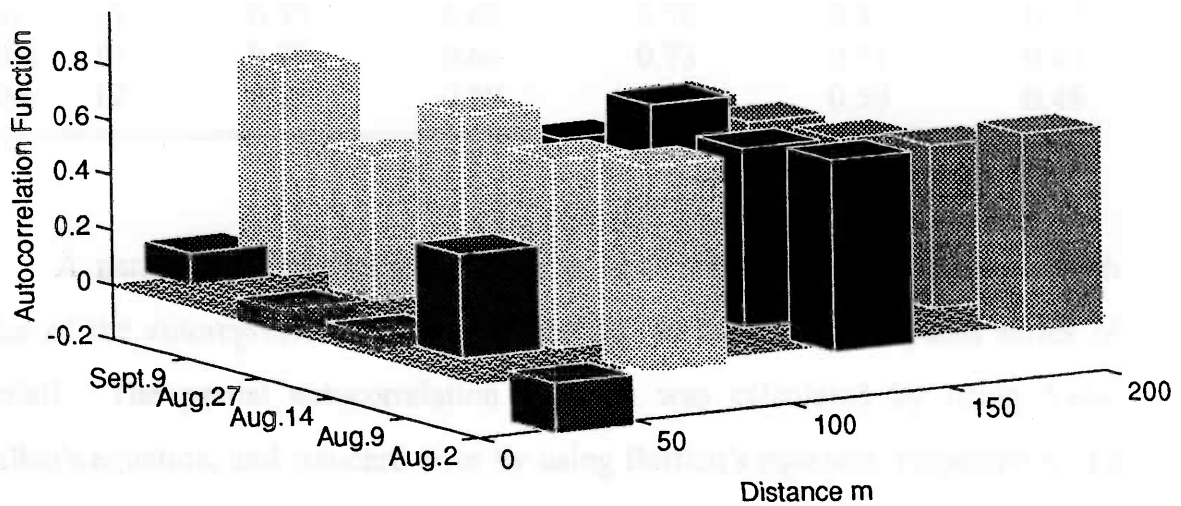


Fig. 4.4 Autocorrelation function at different storms

overestimation of vertical raingages may be caused by the fact that the elevation of its rain aperture area was always a little bit higher than tilted raingages. The spread from tilted raingages measurements, as measured by the standard deviation or the interquartile range, may be caused by variations in slope, aspect, and the wind impact. Slopes in the gully are about 7.75%. Inspection of the data does not show a systematic relationship between gage rainfall and topography. These results, however, do not invalidate the theory that tilted raingages give a better representation of hydrologic rain than vertical raingages (Faures, 1990).

Table 4.2 lists the lag one autocorrelation function for every storm at different sampling intervals. Figure 4.4 gives plots.

Table 4.2 Autocorrelation function (lag 1) for different sampling interval

Interval	n	Aug.2	Aug.9	Aug.14	Aug.27	Sept.9
30m	15	-0.19	0.38	-0.06	-0.09	0.12
60m	33	0.75	0.69	0.76	0.52	0.75
120m	17	0.70	0.66	0.73	0.51	0.41
180m	12	0.73	0.59	0.53	0.50	0.48

A partial autocorrelation function is also used here to determine which order of the autoregressive process will be fit to the observed spatial series of rainfall. The partial autocorrelation function was calculated by using Yule-Walker's equation, and standard error by using Bartlett's equation, respectively. To identify an autoregressive model, the standard error of autocorrelation estimates is a very useful and visible approach, because it is a check on whether the estimated autocorrelation is effectively zero beyond a certain lag. Figure 4.5 exhibits the

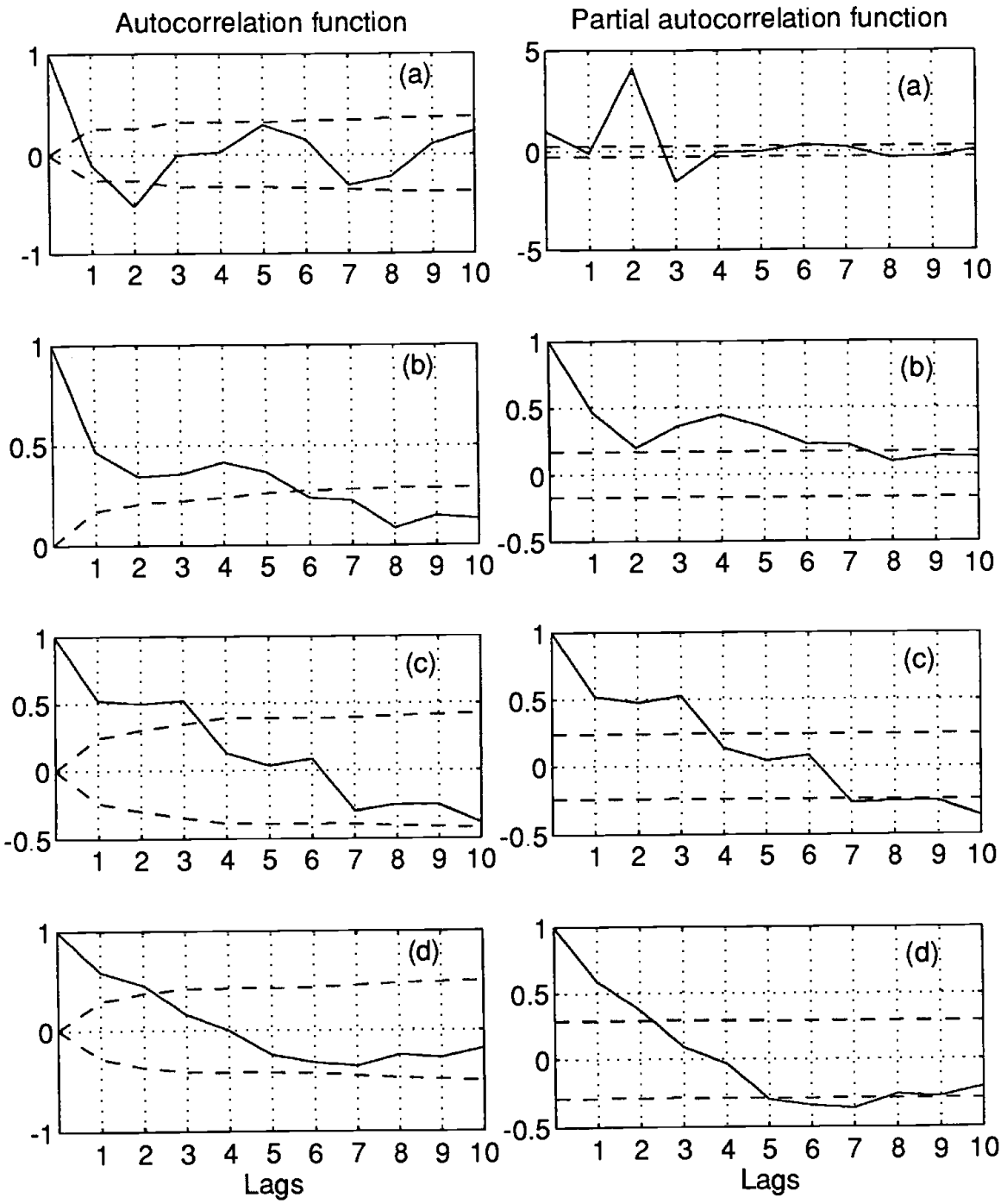


Fig. 4.5 Autocorrelation and partial autocorrelation function at different sampling intervals (a) 30m, (b) 60m, (c) 120m, (d) 180m Dashed lines show standard errors.

autocorrelation function and partial autocorrelation function at different sampling intervals, in which the data were processed by taking the average of total 5 storms. Obviously, all of the autocorrelation fell within the standard error region for all the lags at the 30 meter sampling interval. Therefore, the spatial series of rainfall data do not seem to behave as a first order autoregressive model checked by this method at 30 meter sampling interval. At the other sampling intervals, however, the autocorrelation did not fall within the standard error region at all lags. These results indicate that the precipitation data taken from 60-180 meter intervals are correlated to some degree. However, the order of these autoregressive processes is difficult to detect because both the autocorrelation and the partial autocorrelation function tail off. Whereas the autocorrelation function of an autoregressive model of order n tails off, its partial autocorrelation function has a cutoff after lag n. Conversely, the autocorrelation function of an m order moving average model has a cutoff after lag m, while its partial autocorrelation tails off after the same lag. All of data here did not show these characteristics. Therefore, a mixed process may be suggested. Variograms of the five storms are shown in Appendix F, in which the measurements were taken at 60 meter intervals. It is difficult to analyze those Variograms due to the sampling population and measurement errors. The sill and correlation scale, however, were much smaller than those from the large scale's variogram. It is expected that the larger the scale the bigger the variance (sill). Also, the range (correlation scale) apparent in the small scale data is different than the range apparent in the large scale data.

The most significant conclusion that can be drawn from this study is that within the gully region precipitation does not exhibit significant correlation at the

30 meter interval. Figure 4.5 shows that the autocorrelation is greater than the standard error at several lags for all sampling intervals except for 30 meter. The data used in this graph is the total rainfall depth of all 5 storms. This phenomenon can also be seen by using an individual rainfall storm in three dimensional bar graph as shown in Figure 4.4. This is opposite to the traditional theory of autocorrelation decreases as distance increasing, which will be discussed in the following section.

The entire 150 km^2 Walnut Gulch watershed is also monitored by recording raingages when the experiment was conducted. An example is the storm of August 2, 1991, which started about 12:30 p.m. in which 43 raingages recorded. The upper portion of Figure 4.6 shows the contour map of the total depth of the storm while the lower portion of the graph is a partial enlargement of the area around the line of small-scale measurements. Figure 4.7 compares observations from the closely spaced (30 - 60 meter) raingage to values interpolated from the widely spaced (2 km) raingage network (dashed line). Contours in Figure 4.6 and 4.7 were drawn by the Matlab program, which used the inverse distance technique for interpolation. In Figure 4.7, the solid and dashed lines are different lengths because of the difficulty of registering the two raingage networks.

Generally, the interpolated rainfall depths were about 20% smaller than the observational depths on average. As can be seen in Figure 4.7, the interpolated data were much smoother than the observation data. The average standard deviation of the interpolated rainfall depths was only 47% of that of the data from the closely spaced gages. As a matter of fact, the contour mapping technique makes the raw data smoother and more continuous. This smothing should be taken into account when large scale data are used to interpolate small scale data. It is

realized that the linear interpolation will introduce a spatial dependence which is not present in the observations. Therefore, the biased estimates of the moments and of the spatial correlation as described before are expected.

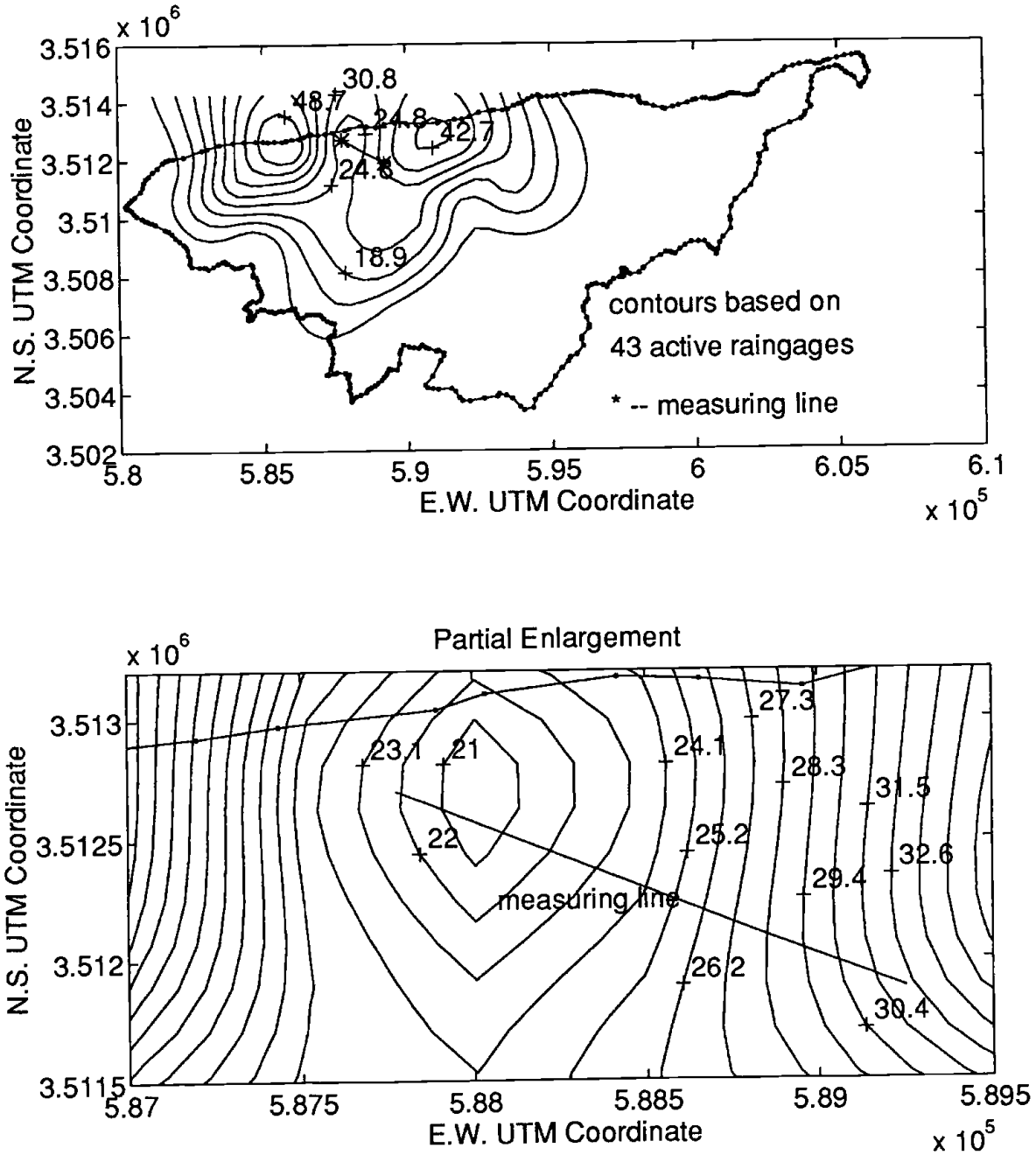


Fig. 4.6 Contour map of total depth on Aug. 2 and its partial enlargement

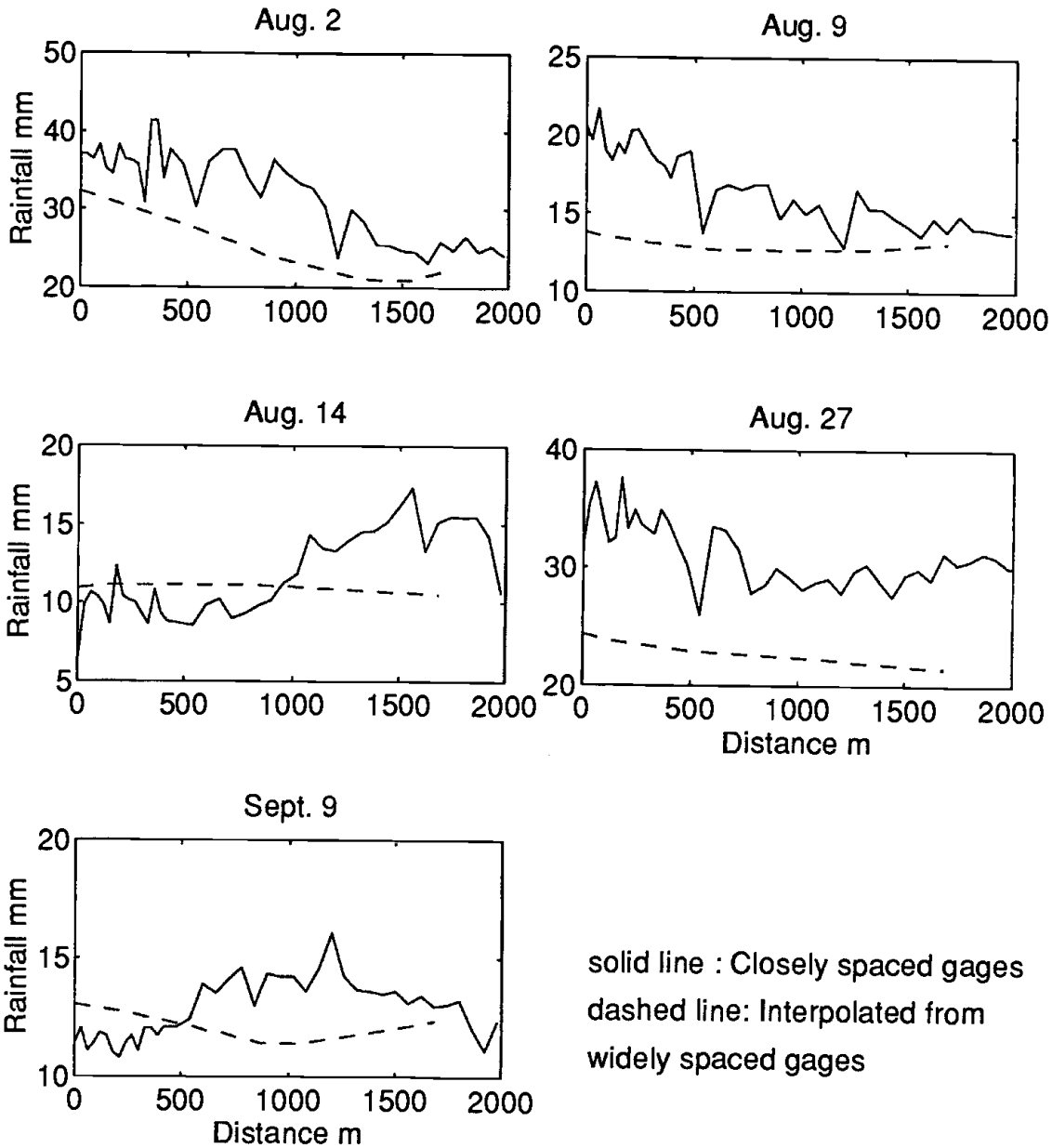


Fig. 4.7 Comparison between observations at two spatial resolutions

4.4 Discussion

The experimental results generally agree with the normal expectation that the rainfall correlation decreases with distance, except for the samples taken at 30 meter intervals. Fennessey et al. (1986) used 2.0 km interval rainfall data and Mcconkey et al. (1990) took the data from 800 to 4400 m apart to study the rainfall correlation between the gages. Both works show that the longer the distance the weaker the correlation. The "characteristic correlation distance", proposed by Mejia et al. (1973), is the expected distance between two independent randomly chosen points within the particular watershed. Fennessey et al. (1986) used this method to find that the characteristic correlation distance of the Walnut Gulch watershed was about 7.0 km. The sampling process in this study, however, is different from that of Fennessey et al. Individual storm depth was used in this study rather than the storm day total depth as used by Fennessey et al. (1986). The storm day total depth may be composed of the components of variance across several storms. This inter storm variance results from the different strengths of the separate storm-producing instabilities. It may also result from incomplete (and variable) coverage of the watershed by the storm (Eagleson et al., 1985).

As mentioned above, all of the autocorrelation functions calculated from 30 meter raingage intervals were not statistically significant. This indicates that there are no correlation between raingages at such small scales. This seems counter-intuitive. A possible explanation follows:

1) High level of variability in the sampling processes.

The random variation includes the observation error and the instrumentation error. The sampling error is caused by a variety of factors such as wind speed, rain type and intensity, as well as exposure (Sharon, 1972). The sampling error is not so important since it has an effect for all samples. The instrumentation error has no effect on the spatial variability because it induces a systematic error for all raingages.

2) Topographic influence

Perhaps, topographic influence is the most important reason why there is no correlation in 30 meter interval in this study. It may be due to the fact that samples at 30 meter intervals were all taken from the gully area. Generally, the recorded rainfall increases with altitude uprising (orographic effect). Weak correlation of rainfall at 30 meter interval may also be due to the turbulent rainfall field caused by irregular topography in the gully. The autocorrelation function calculated by the data with 60 meter spacing in the gully area further verify this conclusion. Although the sample population is considerably smaller in this case (total 8 measurements). This weak correlation is not caused by the slope impact, because the same results were obtained with both titled and vertical gages.

4) Event population limitation

The number of rainfall events in this study may be too small to characterize the variogram structure at such small scale. Further investigation with more data sets is necessary, especially the parallel investigation taking place simultaneously both in the gully area and in the flat area.

CHAPTER FIVE

CONCLUSIONS AND RECOMMENDATIONS

5.1 Major Conclusions and Contributions

The object of this study was to analyze the seasonal and spatial variability of rainfall and to examine how these variations would affect the estimation of areal rainfall. The motivations underlying this research were generated by the necessities such as: to understand rainfall process, to interpolate the gage data to rainfall distributed runoff models, and to have ground measurement references to compare with weather radar measurements.

To better understand and characterize the rainfall spatial and seasonal variability, three major approaches were taken in this research. First, correspondence analysis was applied to a long-term (10 years) monthly rainfall data set which was recorded statewide in Arizona. A portable computer program for the correspondence analysis was developed. Various diagnostic results are provided in the program to aid in interpreting the results (e.g. relative and absolute contributions, error profiles, and supplementary element projections). A series of graphs generated by this program is helpful for the analysis of the results. High quality graphs are easily produced because the program operates within the user-friendly Matlab environment. Second, a geostatistical method was applied to hourly rainfall data from a ten-year raingage data set from the USDA-ARS Walnut Gulch Experimental Watershed, Tombstone, Arizona. An overall index method was developed to evaluate rainfall spatial variability more efficiently and

comparably. This index was based on monthly variograms which were built by the ten years of observations of rainfall intensity. These variograms represent the spatial structure of the rainfall at the Walnut Gulch Watershed and the seasonal variations of the spatial structure. Estimates of areal rainfall were computed using block Kriging. The effect of raingage density on the estimation was examined. The techniques developed in this study can be adopted to other watersheds in a general sense, but the variograms must be rebuilt if the climate and scale are significantly different. Third, in order to better understand and characterize the rainfall spatial variability at a small scale (30 - 180 m), a small scale experiment was conducted within the Walnut Gulch Watershed. The spatial variability from the closely spaced (60 m) gages was greater than would be expected based on an examination of the widely spaced (2 km) data. The major conclusions drawn from this research are as follows:

1. Correspondence analysis is an appropriate method to examine the rainfall spatial and temporal variation. However, it is difficult to define the variables and samples in practice. For a large region with a temperate climate, monthly rainfall may be a suitable variable. The seasonal patterns can be exposed by the correspondence analysis method. For a small region or a watershed, however, rainfall duration or intensities may be a suitable variable. The spatial patterns (wet and dry regions) can be exposed by the correspondence analysis method. Applying the correspondence analysis to the raingage data collected statewide in Arizona, spatial patterns of precipitation can be grouped as follows:

a) In southeastern Arizona, the primary wet season (summer) is dominant.

b) Between the northwest and northeast part of Arizona, the secondary wet season (midwinter) is dominant.

c) And, in the northwest part of Arizona, the dry season is dominant.

2. A portable computer program (MACTORS) developed in this study can perform correspondence analysis easily and quickly. Compared to the available commercial programs, MACTORS provides high quality output plots and many diagnostic results to aid in interpreting the results.

3. Monthly variograms built from 10 years of hourly records at Walnut Gulch Watershed clearly show the rainfall seasonal and spatial variability. The range of the variogram, or correlation scale, is smallest in the primary wet season (3 km). The correlation scale ranges from six to thirty km during the secondary wet season and dry season, respectively. Another important phenomenon is that the monthly variograms are directly proportional to the average rainfall intensity during the primary wet season. The variograms have the same shapes but different scales. This is convenient because the Kriging weights (used for rainfall interpolation) are a function only of rainfall location and the variogram shape: Kriging weighting coefficients λi can be just computed once for all, and all of the estimated variances are directly proportional to each other by the scaling factors.

4. A block Kriging computer program was written to find the average areal rainfall over each four kilometer by four kilometer grid at the Walnut Gulch Watershed. When 43 raingages (one raingage per 2.6 square kilometers approximately) were used, the average relative estimated error (σ / A*) was

4.88%. By reducing the raingage network densities to approximately one raingage per 5.2 square kilometers and 10.4 square kilometers, the average relative estimated errors increased to 7.21% and 14.30%, respectively.

5. An overall index was developed to evaluate the rainfall spatial variability. It is very useful for comparing and classifying the random rainfall field because it changes from zero to one and it is dimensionless. This overall index showed that Walnut Gulch rainfall has the biggest spatial variations during the primary wet season (July, August and September).

6. The small scale rainfall measurement experiment conducted by this study shows that large rainfall spatial variations exist at small scales. Results, which support Faures (1990) conclusions, have refuted the widely accepted assumption that convective rainfall fields in the Southwest are homogeneous at the scale of 30 - 60 m.

A turbulent rainfall field caused by a steep gully was observed by this experiment. It was found that this turbulent effect eliminated the rainfall correlation relationship, even at very short distances.

7. The interpolated rainfall depths taken from the widely spaced (2 km) raingage network were approximately 20% smaller than observed depths from the closely spaced (60 m) gages. Linear interpolated methods make the raw data smoother and more continuous. The average standard deviation of the interpolated rainfall depths was 47% of that computed from the closely spaced gages.

5.2 Recommendations for Future Research

The above conclusions are based upon the data sets collected from the semi-arid environment within the USDA-ARS Walnut Gulch Experimental Watershed, Tombstone, Arizona. Therefore, the results of this study should not be extrapolated to other watersheds without using extreme prudence. The major recommendations for further study are as follows:

1. More data sets with different variables and samples are needed to perform the correspondence analysis. Further experience with correspondence analysis may demonstrate its value in hydroclimtic analyses.
2. MACTORS, a portable computer program, developed by this study needs to be revised under the Matlab new version (4.1). Thus, the program should be more attractive by taking advantage of the Matlab new version.
3. A comparison between the weather radar measurements and the areal rainfall estimates (such as estimates used by this study) should be carried out. Furthermore, a method to evaluate the accuracy of the weather radar measurements should be developed based on the ground measurements and their estimations.
4. The relationship between the raingage densities and the average relative estimation errors needs to be studied more completely, so that an optimal raingage network density can be selected according to this relationship.
5. Further rainfall measurement at 30 meter spacing (along both gullies and flat areas) would help resolve whether the 30 meter results in this study are due to gully-related erratic winds.

APPENDICES

APPENDIX A

MATLAB CODE OF MATCORS

```
function [r,q,ev,c_ra,s_ra,evv,ess,sv,sp]=matcors(x)

% MATCORS performs correspondence analysis for a contingency matrix x with n
% samples and p variables.

% Syntax [r,q,ev,c_ra,s_ra,evv,ess,sv,sp]=matcors(x)


% Missing values are coded as 'NaN' (not-a-number).

% Input Description
% x  : contingency matrix with n samples and p variables
% Echo input
% pec: Cumulative percent variation to determine the amount of principal factors.
% nb : Selected number of principal factors.
% nc : The number of supplementary variables.
% cc : Supplementary variables.
% nr : The number of supplementary samples.
% cr : Supplementary samples.
% f1 & f2 : Two principal factors to be plotted.

% Output description
% r  : R-type main factors loading matrix.
% q  : Q-type main factors loading matrix.
% ev : Eigenvalues, relative and cumulative variation explained by factors
% c_ra : Relative and absolute contributions of variables.
% s_ra : Relative and absolute contributions of samples.
% evv  : Weights and error profiles of variables.
% ess  : Weights and error profiles of samples.
% sv   : Supplementary variables projection values.
% sp   : Supplementary samples projection values.

% Echo of data and input options

disp('Input data matrix')
disp(x);

tot=sum(sum(x));x=x/tot;
```

```
% Check negative entries in the data matrix
nz=find(x<0);
if nz>0
error('Data matrix has negative entries !!!')
end

disp('Strike any key when ready'), pause

% Call function EIGV to form a real symmetric covariance matrix of variables
% and compute all eigenvalues and eigenvectors by Jacobi rotation method.
[n,p,evl,evc,evr,tco,tro]=eigv(x);
nn=n;pp=p;

kk=menu('Do you want to select Supplementary Elements?','Yes','No');
if kk==1

  nc=input('How many variables (column)?');

  for i=1:nc
     cc(i)=input(['Please give supplementary column no.',num2str(i),'number: ' ]);
       % Check input choice
       if (nc>p) | (cc(i)>p)
       error('Wrong number--Larger than maximal column number')
       end
  end

  nr=input('How many samples (row)?');

  for i=1:nr
     cr(i)=input(['Please give supplementary row no.',num2str(i),'number: ' ]);
       % Check input choice

       if (nr>n) | (cr(i)>n)
       error('Wrong number--Larger than maximal row number')
       end
  end

% Supplementary element projection values
   sv=((x(:,cc))'*evr)./(diag(tco(cc))*ones(nc,pp));
   sp=(x(cr,:)*evc')./(diag(tro(cr))*ones(nr,pp));

% Eliminate the supplementary elements from original data matrix [x]
% to form new matrix [xc]
```

```
xc=x;xc(:,cc)=[];xc(cr,:)=[];

% Performs EIGV function again for new matrix [xc].
[n,p,evl,evc,evr,tco,tro]=eigv(xc);

end
% Calculate and display eigenvalues and their relative and cumulative
% percentage of variation.
ev(:,1)=evl;ev(:,2)=100*evl/sum(evl);ev(:,3)=100*cumsum(evl)/sum(evl);
disp('Eigvenvalues Rel.--%   Cul.--%');
disp(ev);

 pre=input('How many cumulative percent variations do you want? ');
   if pre>100
   error('Wrong number -- Cumulative percentage larger than 100')
   end

 nb=input('How many principal factors do you want? ');
   if nb>p-1
   error(['Wrong number -- larger than the total number ',num2str(p)-1])
   end

% The number of principal factors are determined by cumulative percent variation
  for i=1:p-1
      if ev(i,3) >= pre
      m=i;
      break,end
  end

% The amount of principal factors are determined by the selected number
  if nb > m
  m=nb;
  end

% R-mode factor loadings;
r=evc*diag(sqrt(evl));r=r(:,1:p-1);

% Q-mode factor loadings;
q=evr*diag(sqrt(evl));q=q(:,1:p-1);

% Relative and Absolute contributions of variables and samples for main factors.
for k=1:p-1
```

```
% Relative contributions
 c_ra(:,k)=100*(evl(k)*evc(:,k).^2)./(evc.^2*evl);
 s_ra(:,k)=100*(evl(k)*evr(:,k).^2)./(evr.^2*evl);
 % Absolute contributions
 c_ra(:,p+k-1)=100*(tco'.*(inv(diag(tco))*evc(:,k))).^2;
 s_ra(:,p+k-1)=100*(tro.*(inv(diag(tro))*evr(:,k))).^2;
end


% Error profiles of all variables and samples for the remaining factors.
for j=1:p
epv(j)=(((diag(sqrt(evl(m+1:p))*evc(j,m+1:p)))'*(evr(:,m+1:p))').^2)*tro;
end
for i=1:n
epps(i)=(((diag(sqrt(evl(m+1:p))*evr(i,m+1:p)))'*(evc(:,m+1:p))').^2)*tco';
end

% Variables and samples' weights and their error profiles.
    vw=(tco/sum(tco))';sw=tro/sum(tro);
    evv(:,1)=100*vw;evv(:,2)=epv';
    ess(:,1)=100*sw;ess(:,2)=epps';

% Echo of principal factors to be plotted
f1=input('Which two main factors do you want to plot? First--please!');
f2=input('Second--please!');
  if (f1 >m) | (f2 >m)
  error('Wrong number -- larger than the number of principal factors')
  end

%  R-mode factor loading plot
if nc>0
tt=1:pp;tt(cc)=[];
rm(cc,1:m)=sv(:,1:m);rm(tt,1:m)=r(:,1:m);
subplot(211),plot(r(:,f1),r(:,f2),'*',sv(:,f1),sv(:,f2),'x');
gtext('x -- Supplementary variable projections');
else
rm=r;
subplot(211),plot(r(:,f1),r(:,f2),'*')
end
xlabel(['Factor',num2str(f1)]);
ylabel(['Factor',num2str(f2)]);
gtext('Variables');
```

```
% Mark down the variables number.
   for i=1:pp
   text(rm(i,f1),rm(i,f2),num2str(i));
   end
grid;

% Q-mode factor loading plot
if nr>0
tt=1:nn;tt(cr)=[];
qm(cr,1:m)=sp(:,1:m);qm(tt,1:m)=q(:,1:m);
subplot(212),plot(q(:,f1),q(:,f2),'+',sp(:,f1),sp(:,f2),'o');
gtext('o -- Supplementary sample projections');
else
qm=q;
subplot(212),plot(q(:,f1),q(:,f2),'+')
end
xlabel(['Factor',num2str(f1)]);
ylabel(['Factor',num2str(f2)]);
gtext('Samples');
   for i=1:nn
   text(qm(i,f1),qm(i,f2),num2str(i));
   end
grid;pause;clg;

% factor plane
if kk==2 % No supplementary elements
plot(r(:,f1),r(:,f2),'*',q(:,f1),q(:,f2),'+');
else
if (nc>0) & (nr>0) % Both variables and samples have supplementary elements.
plot(r(:,f1),r(:,f2),'*',q(:,f1),q(:,f2),'+',sv(:,f1),sv(:,f2),'x',sp(:,f1),sp(:,f2),'o');
elseif (nc>0) & (nr==0) % Only variables have supplementary elements.
plot(r(:,f1),r(:,f2),'*',q(:,f1),q(:,f2),'+',sv(:,f1),sv(:,f2),'x');
elseif (nr>0) & (nc==0) % Only samples have supplementary elements.
plot(r(:,f1),r(:,f2),'*',q(:,f1),q(:,f2),'+',sp(:,f1),sp(:,f2),'o');
end;end;end;end

% Mark down both variable and sample number on the plot;
   for i=1:pp
   text(rm(i,f1),rm(i,f2),num2str(i));
   end
   for j=1:nn
   text(qm(j,f1),qm(j,f2),num2str(j));
   end;grid;
```

```
xlabel(['Factor',num2str(f1)]);
ylabel(['Factor',num2str(f2)]);
gtext('Correspondence Analysis Factor Plane');
gtext('*---variables loading coordinate');
gtext('+---samples loading coordinate');
if kk==1
gtext('x---supplementary variables projection');
gtext('o---supplementary samples projection');
end
pause;clg;

% Plot bar graph of varibles' and samples' absolute contributions for
% first 4 factors
if p>=5
np=4;
else
np=p-1;
end
for nk=1:np;
tw1=['Variable number';'Sample number  ';'% Contribution '];
subplot(220+nk),bar(1:p,c_ra(:,p-1+nk))
xlabel(tw1(1,:));
ylabel(tw1(3,:));
gtext(['AC for factor #',num2str(nk)]);
hold on
me=mean(c_ra(:,p-1+nk));
plot([.5 p+.5],[me me],'--b');
text(1,me,'mean');
hold off
end
pause;clg

for nk=1:np;
subplot(220+nk),bar(1:n,s_ra(:,p-1+nk));
xlabel(tw1(2,:));
ylabel(tw1(3,:));
gtext(['AC for factor #',num2str(nk)]);
hold on
me=mean(s_ra(:,p-1+nk));
plot([.5 n+.5],[me me],'--b');
text(1,me,'mean');
hold off
end
```

```
pause;clg

% Plot bar graph of variables' and samples' weights and error profiles
tw2=['Weight      ';'Error profile'];
for nk=1:2;
subplot(210+nk),bar(1:p,evv(:,nk));
ylabel(tw2(nk,:));
xlabel(tw1(1,:));
me=mean(evv(:,nk));
hold on
plot([.5 p+.5],[me me],':w');
text(1,me,'mean');
hold off
end
pause;clg

for nk=1:2;
subplot(210+nk),bar(1:n,ess(:,nk));
ylabel(tw2(nk,:));
xlabel(tw1(2,:));
me=mean(ess(:,nk));
hold on
plot([.5 n+.5],[me me],':w');
text(1,me,'mean');
hold off
end
pause;clg




function [n,p,evl,evc,evr,tco,tro]=eigv(x)

% Forms a real symmetric covariance matrix of variables, and computes
% all eigenvalues and eigenvectors by Jacobi rotation mathod.

pre=1.0e-6;%Jacobi iteration precision;
total=pre+1; %Initial first rotation;

[n,p]=size(x);
tot=sum(sum(x)); % the sum of every element in matrix [x];
tco=sum(x);   % a row vector with the sum over each column of matrix [x];
```

```
    tro=sum(x')';  % a column vector with the sum over each row of matrix [x];

    z=(x-tro*tco/tot)./sqrt(tro*tco);
    a=z'*z;     % covariance matrix of variables

p=length(a);
v=eye(p);u=v;

while total > pre


       [cul,row,total]=index(a);

       st=(a(cul,cul)-a(row,row))./(2.*a(row,cul));

       if st==0
       c=1/sqrt(2);
       s=c;

       else

         t=sign(st)./(abs(st)+sqrt(st.^2+1));
         c=1/sqrt(t.^2+1);
         s=t.*c;

       end

     v(row,row)=c; v(row,cul)=s;
     v(cul,row)=-s;v(cul,cul)=c;

a=v'*a*v;
u=u*v;v=eye(p);
end

eigva=a;
eigvc=u;

% Sort the eigenvalues and eigenvectors in descending order;
    [h,k]=sort(diag(-eigva));
    evl=(-h); % eigenvalues;
    evc=eigvc(:,k); % eigenvectors;

%Compute eigenvectors of covariance matrix of samples
```

```
   evr=(z*evc)./(ones(n,p)*diag(sqrt(evl)));


return



function [cul,row,tot]=index(x)
% Find the row and column number of maximum off-diagonal element of a
% real symmetric matrix [x]

z=triu(x,1);
p=length(x);

[y,i]=sort(abs(z));
[a,b]=sort(abs(y(p,:)));
m=max(a);
cul=b(p);
row=i(p,cul);

tot=abs(x(row,cul));

return
```

APPENDIX B

1. Example 1 output

a. (r) R-mode factor loadings (variables)

factor1	factor2	factor3	factor4
-0.135	-0.001	-0.005	-0.004
0.171	-0.032	-0.010	-0.002
0.086	0.064	0.011	-0.002
0.006	-0.030	0.034	-0.001
0.000	0.001	0.000	0.017

b. (q) Q-mode factor loadings (samples)

-0.050	0.019	-0.006	-0.002
-0.073	0.012	-0.008	-0.001
-0.002	0.016	0.003	0.003
0.085	0.026	0.013	-0.007
0.004	0.028	0.002	0.002
-0.076	0.004	-0.001	0.003
-0.067	-0.010	0.003	0.000
0.086	0.005	-0.004	0.005
0.084	-0.003	-0.010	0.005
0.023	0.011	-0.005	0.006
0.035	0.018	0.005	0.000
-0.010	-0.018	0.014	0.002
0.026	-0.033	-0.012	-0.001
-0.001	-0.035	0.017	0.004
0.045	-0.009	-0.007	-0.003
-0.001	-0.006	0.008	-0.004
0.037	-0.020	-0.011	-0.005
-0.061	-0.001	-0.007	-0.006
-0.071	0.002	-0.002	0.005
-0.013	-0.004	0.007	-0.006

c. (ev)

Eigenvalue	Relative variation	Comulative variation
0.055	87.677	87.677
0.006	9.539	97.216
0.001	2.266	99.482
0.000	0.518	100.000
0.000	0.000	100.000

d. (c_ra)

Variable rela. contri.				Variable abs. contri.			
99.791	0.007	0.133	0.069	33.203	0.020	1.710	3.904
96.234	3.396	0.357	0.012	53.155	17.243	7.628	1.168
64.095	34.913	0.967	0.024	13.583	68.005	7.931	0.857
1.572	42.135	56.265	0.028	0.060	14.716	82.723	0.179
0.000	0.322	0.039	99.639	0.000	0.016	0.008	93.892

e. (s_ra)

Sample rela.contri.				Sample abs. contri.			
86.716	12.055	1.141	0.088	4.621	5.905	2.352	0.793
96.240	2.631	1.120	0.008	9.847	2.475	4.435	0.143
1.943	91.177	4.176	2.704	0.010	4.317	0.832	2.356
89.214	8.216	1.943	0.627	13.316	11.272	11.220	15.835
1.791	96.973	0.714	0.521	0.026	12.833	0.398	1.269
99.529	0.295	0.019	0.158	10.535	0.287	0.077	2.824
97.657	2.160	0.180	0.003	8.107	1.648	0.578	0.047
99.254	0.276	0.190	0.280	13.575	0.347	1.007	6.474
98.166	0.143	1.361	0.330	12.758	0.171	6.846	7.256
73.371	18.682	3.541	4.407	0.928	2.172	1.733	9.432
78.598	20.057	1.341	0.004	2.288	5.366	1.510	0.020
15.464	53.784	29.950	0.802	0.173	5.536	12.977	1.519
36.537	55.701	7.647	0.115	1.281	17.951	10.374	0.683
0.033	79.155	19.587	1.224	0.001	20.678	21.540	5.888
93.832	3.657	2.219	0.292	3.690	1.322	3.376	1.943
1.613	30.201	55.898	12.288	0.004	0.642	5.005	4.812
71.497	20.623	6.710	1.170	2.513	6.664	9.128	6.960
97.711	0.048	1.166	1.075	6.907	0.031	3.188	12.854
99.394	0.091	0.081	0.433	9.109	0.077	0.287	6.722
62.510	6.743	16.288	14.458	0.311	0.308	3.135	12.168

f.(evv)

Variable weight	Error profile
61.16353600	0.00000185
20.80623000	0.00000562
9.62437010	0.00000577
2.32249200	0.00005884
6.08337150	0.00001517

g.(ess)

Sample weight	Error profile
4.94273930	0.00000234
4.94732020	0.00000358

5.00687130	0.00000163
4.95190110	0.00000930
4.98396700	0.00000083
5.00229040	0.00000071
5.02061380	0.00000049
5.01603300	0.00000186
5.00687130	0.00000530
5.00229040	0.00000283
4.99770960	0.00000119
5.08474580	0.00001133
4.99770960	0.00000868
5.13513510	0.00002073
4.97938620	0.00000373
5.01603300	0.00000381
4.96564360	0.00001083
4.93815850	0.00000756
5.01145210	0.00000169
4.99312870	0.00000408

2. Example 1 output with variable 4 and sample 4 and 8 as supplementary elements

a. (r) R-mode factor loadings (variables)

factor1	factor2	factor3	factor4
-0.117	-0.008	-0.003	0.000
0.166	-0.030	-0.002	0.000
0.064	0.064	-0.002	0.000
0.001	0.003	0.016	0.000

b. (q) Q-mode factor loadings (samples)

-0.045	0.014	-0.003	0.000
-0.069	0.005	-0.001	0.000
0.006	0.020	0.002	0.000
0.012	0.032	0.000	0.000
-0.072	-0.001	0.003	0.000
-0.061	-0.012	0.001	0.000
0.101	0.003	0.004	0.000
0.034	0.015	0.005	0.000
0.047	0.026	-0.001	0.000
0.000	-0.011	0.003	0.000
0.042	-0.033	-0.001	0.000
0.011	-0.026	0.007	0.000
0.060	-0.005	-0.003	0.000

0.009	-0.000	-0.004	0.000
0.052	-0.019	-0.005	0.000
-0.056	-0.007	-0.007	0.000
-0.066	-0.002	0.005	0.000
-0.004	-0.000	-0.006	0.000

c. (ev)

Eigenvalue	Relative variation	Cumulative variation
0.045	89.440	89.440
0.005	9.997	99.437
0.000	0.563	100.000
0.000	0.000	100.000

d. (c_ra)

Variable rela contri.			Variable abs. contri.		
99.425	0.493	0.082	30.120	1.335	3.944
96.778	3.212	0.010	60.919	18.088	0.970
49.913	50.032	0.055	8.958	80.336	1.570
0.498	4.354	95.149	0.003	0.241	93.516

e. (s_ra)

Sample rela contri.			Sample abs. contri.		
90.328	9.354	0.318	4.485	4.156	2.506
99.502	0.470	0.029	10.603	0.448	0.488
8.251	90.977	0.772	0.083	8.226	1.239
11.979	88.008	0.013	0.312	20.512	0.054
99.774	0.005	0.221	11.467	0.005	4.033
96.070	3.883	0.047	8.323	3.009	0.649
99.734	0.117	0.149	22.573	0.237	5.346
83.138	15.338	1.524	2.529	4.174	7.361
75.888	24.060	0.052	4.809	13.642	0.522
0.010	90.438	9.552	0.000	2.287	4.288
61.711	38.269	0.020	3.904	21.659	0.200
14.805	79.638	5.557	0.268	12.890	15.966
99.042	0.684	0.275	7.889	0.487	3.474
82.241	0.026	17.733	0.160	0.000	5.497
87.801	11.406	0.794	6.068	7.052	8.713
96.863	1.745	1.392	6.841	1.103	15.616
99.265	0.130	0.605	9.648	0.113	9.341
28.643	0.040	71.317	0.037	0.000	14.706

f.(evv)

Variable weight	Error profile
64.60047900	0.00000055
20.02291900	0.00000013
9.13636840	0.00000022
6.24023340	0.00001301

g.(ess)

Sample weight	Error profile
5.54224400	0.00000055
5.54745290	0.00000011
5.57349720	0.00000027
5.56828840	0.00000001
5.57349720	0.00000088
5.55787060	0.00000014
5.58391500	0.00000117
5.58912390	0.00000161
5.55787060	0.00000011
5.57349720	0.00000094
5.54224400	0.00000004
5.58912390	0.00000349
5.53703510	0.00000076
5.53182620	0.00000120
5.52140850	0.00000190
5.51099070	0.00000341
5.58391500	0.00000204
5.51619960	0.00000321

h. Supplementary variable projection (sv)

0.008	-0.044	0.051	-0.000	0.032

i. Supplementary sample projections (sp)

-0.242	0.164	0.479	0.059	0.044
-0.228	0.168	0.495	0.022	0.042

APPENDIX C

WEATHER STATION DESCRIPTION

No.	Station	Elevation(ft)	Latitude	Longitude
1	AJO	1800.00	32:22:00	112:52:00
2	ALAMO	1060.00	34:16:00	113:34:00
3	ALAMO DAM	1290.00	34:14:00	113:35:00
4	ASH FORK 2	5080.00	35:13:00	112:29:00
5	BAGDAD	3710.00	34:34:00	113:10:00
6	BISBEE	5310.00	31:26:00	109:55:00
7	BLACK RIVER PUMPS	6040.00	33:29:00	109:46:00
8	BOWIE JCT R15 ON W5	4720.00	32:26:00	109:42:00
9	CASA GRANDE RUINS N M	1420.00	33:00:00	111:32:00
10	CIBECUE	5050.00	34:02:00	110:29:00
11	COCHISE 4 SSE	4180.00	32:04:00	109:54:00
12	CROWN KING	5920.00	34:12:00	112:20:00
13	DOUGLAS	4040.00	31:21:00	109:32:00
14	DUNCAN	3660.00	32:45:00	109:07:00
15	FLAGSTAFF WSO AP	7010.00	35:08:00	111:40:00
16	FLORENCE JUNCTION	1880.00	33:17:00	111:22:00
17	GRAND CANYON N P	6950.00	36:03:00	112:08:00
18	GRAND CANYON NATL PK 2	6790.00	36:03:00	112:09:00
19	HACKBERRY	3580.00	35:22:00	113:44:00
20	HACKBERRY 2 SE	3700.00	35:21:00	113:41:00
21	KEAMS CANYON	6210.00	35:49:00	10:12:00
22	KINGMAN	3360.00	35:11:00	114:03:00
23	KINGMAN 2	3540.00	35:12:00	114:01:00
24	MAYER 3 NNW	4640.00	34:26:00	112:15:00
25	MONTEZUMA CASTLE N W	3180.00	34:37:00	111:50:00
26	NOGALES	3810.00	31:21:00	110:55:00
27	ORACLE 2 SE	4510.00	32:36:00	110:44:00
28	PAGE	4270.00	36:56:00	111:27:00
29	PAINTED ROCK DAM	550.00	33:05:00	113:02:00
30	PAYSON	4910.00	34:14:00	111:20:00
31	PERNER RANCH	5600.00	35:22:00	113:17:00
32	PETRIFIED FOREST NAT PK	5450.00	34:49:00	109:53:00
33	PHOENIX WSFO AP	1110.00	33:26:00	112:01:00
34	PHOENIX CITY	1080.00	33:27:00	112:04:00
35	PIMA R4 ON W2	3770.00	32:50:00	110:01:00
36	POLAND JUNCTION	4900.00	34:27:00	112:16:00
37	PRESCOTT FAA AP	5020.00	34:39:00	112:26:00
38	ROCK CREEK R S	3630.00	33:49:00	109:48:00
39	ROUND VALLEY	3740.00	35:06:00	113:38:00
40	SANTA RITA EXP RANGE	4300.00	31:46:00	110:51:00
41	SEDONA R S	4220.00	34:52:00	111:46:00
42	SENECA 3 NW	4920.00	33:47:00	110:32:00
43	SIERRA ANCHA	5100.00	33:48:00	110:58:00
44	SIGNAL 13 SW	2510.00	34:22:00	113:48:00
45	SUMMIT	3650.00	33:33:00	110:57:00

46 SUNFLOWER 3 NNW	3720.00	33:54:00	111:29:00
47 SUPERIOR	3000.00	33:18:00	111:06:00
48 SUPERIOR 2 ENE	4160.00	33:18:00	111:04:00
49 SUPERSTITION MTN	1960.00	33:22:00	111:26:00
50 TANQUE R9 ON W4	3560.00	32:37:00	109:37:00
51 TROUT CREEK STORE	2850.00	34:53:00	113:39:00
52 TRUXTON CANYON	3820.00	35:23:00	113:40:00
53 TUCSON NURSERY 4 NW	2250.00	32:18:00	111:03:00
54 TUCSON WSO AP	2580.00	32:08:00	110:57:00
55 TURKEY CREEK 1	6750.00	33:45:00	109:48:00
56 TUWEEP	4780.00	36:17:00	113:04:00
57 UPPER PARKER CREEK	5500.00	33:48:00	110:57:00
58 VAIL	3230.00	32:03:00	110:43:00
59 WALNUT CREEK	5090.00	34:56:00	112:49:00
60 WHITERIVER 1 SW	5120.00	33:50:00	109:58:00
61 WHITLOCK VLY R2 ON W1	3290.00	32:49:00	109:31:00
62 WINSLOW WSO AP	4890.00	35:01:00	110:44:00
63 WORKMAN CREEK 1	6970.00	33:49:00	110:55:00
64 YUMA WSO AP	210.00	32:40:00	114:36:00

APPENDIX D

PRECIPITATION DATA (1991 summer monsoon)

(First 15 data with 30 m spacing; the rest with 60 m)

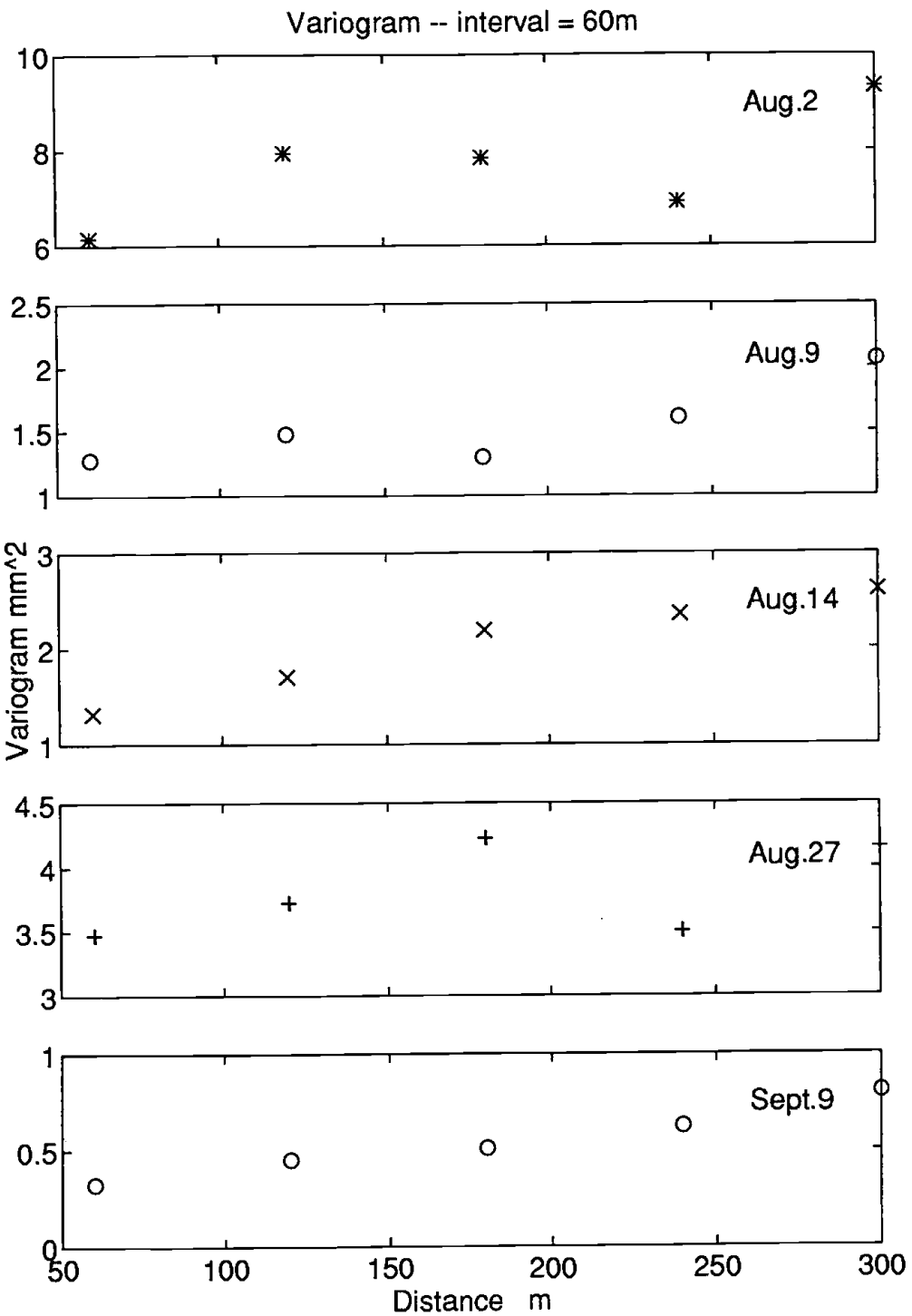
Aug.2	Aug.9	Aug.14	Aug.27	Sept.9
37.06	20.63	6.49	31.81	11.43
37.06	19.70	9.88	35.52	12.04
36.44	21.68	10.62	37.25	11.12
38.30	19.02	10.38	34.59	11.43
35.21	18.41	9.88	32.12	11.86
34.59	19.52	8.65	32.49	11.74
38.30	18.84	12.29	37.68	11.06
36.44	20.32	10.38	33.35	10.81
36.32	20.38	10.13	34.90	11.43
35.82	19.64	10.01	33.66	11.74
30.88	18.84	9.26	33.23	11.12
41.38	18.34	8.56	32.86	12.04
41.38	18.10	10.81	34.90	12.04
33.97	17.29	9.26	33.97	11.74
37.68	18.65	8.78	32.74	12.11
35.82	19.02	8.65	30.27	12.11
30.27	13.77	8.52	25.94	12.42
36.13	16.55	9.82	33.48	13.90
37.86	16.86	10.19	33.17	13.53
37.68	16.55	8.96	31.50	14.08
33.97	16.86	9.26	27.79	14.58
31.50	16.86	9.76	28.41	12.97
36.44	14.70	10.13	29.96	14.33
34.59	15.94	11.24	29.03	14.21
33.35	15.01	11.80	28.10	14.21
32.74	15.63	14.33	28.72	13.59
30.27	14.08	13.47	29.03	14.52
23.78	12.85	13.28	27.79	16.06
29.96	16.55	13.96	29.65	14.21
28.41	15.32	14.52	30.27	13.65
25.45	15.32	14.58	28.78	13.59
25.39	14.70	15.13	27.49	13.47
24.71	14.21	16.18	29.34	13.59
24.58	13.59	17.36	29.83	13.09
23.16	14.70	13.28	28.91	13.40
25.82	13.90	15.13	31.19	12.97
24.71	14.89	15.50	30.27	13.03
26.50	14.08	15.44	30.57	13.22
24.58	14.02	15.50	31.19	12.04
25.26	13.84	14.21	30.88	11.12
24.09	13.77	10.56	29.96	12.35

APPENDIX E

PRECIPITATION DATA (TILTED RAINGAGE WITH 30 METER INTERVALS)

Aug.2	Aug.9	Aug.14	Aug.27	Sept.9
35.21	19.77	8.77	34.59	11.86
24.71	21.62	8.96	33.97	12.17
38.30	20.26	10.93	35.70	12.04
39.53	18.78	10.19	35.21	11.98
37.06	20.88	8.46	33.97	11.61
42.00	18.84	8.96	31.50	10.69
38.30	18.47	8.34	32.74	11.12
33.00	18.75	9.00	33.50	10.50
29.03	19.02	10.87	33.97	10.07
33.97	19.15	9.26	33.35	11.12
31.50	19.21	9.88	32.49	11.43
37.68	18.28	10.19	35.08	11.61
30.88	16.92	10.50	35.82	10.99

APPENDIX F


Variogram -- interval = 60m
Aug.2
Aug.9
Aug.14
Aug.27
Sept.9
Variogram mm^2
Distance m
10
8
6
2.5
2
1.5
1
3
2
1
4.5
4
3.5
3
1
0.5
0
50
100
150
200
250
300

LIST OF REFERENCES

Aho, V. A., B. W. Kernighan, and P. I. Weinberger, The AWK programming language, Addison-Wesley Publishing Co., 210pp., 1988.

Avila, F., and D. E. Myers, Correspondence analysis applied to environmental data sets: a study of Chautauqua lake sediments, *Chemometrics and Intelligent Laboratory system, 11,* 229-249, 1991.

Bastin, G., and M. Gevers, Identification and Optimal estimation of random fields from scattered pointwise data, 1984.

Bastin, G., B. Lorent, C. Duque, and M. Gevers, Optimal estimation of the average areal rainfall and optimal selectinm of rain gauge location, *Water Resources Research, 20 (4),* 463-470, 1984.

Belousov, S. L., L. S. Gandin, and S. A. Mashkovich, Computing processing of meteorological data, *Isreal Program for Scientific Translation, Jerusalem,* 201pp., 1971.

Benzecri, J.P., l'Analyse des donnees, Tom II, l'Anolyse des correspondances, Dunod, Paris, 619pp., 1973.

Berndtson, R., Temporal Variability in spatial correlation of daily rainfall, *Water Resour. Res., 24 (9),* 1511-1517, 1988.

Conrad, V., and L. W. Pollak, Methods in climatology, Harvard University Press, Cambridge, Mass, 54-55, 1950.

Dean, J. D. and W. M. Snyder, Temporally and areally distributed rainfall, *J. of the Irrigation and Drainage Div. ASCE. 103 (IR2),* 221-229, 1977.

Delfiner, P., and J. P. Delhomme, Optimum interpolation by Kriging, *in Display and Analysis of Spatial Data,* Edited by J. C. Davis, and M. J. McCullagh, John Wiley, New York, 96-114, 1975.

Delhomme, J. P., Kriging in the hydrosciences, *Adv. Water Resour. I (5),* 251-266, 1978.

Drozdov, D. A., and A. A. Sephelevskii, the theory of interpolation in a stochastic field of meteorological elements and its application to meteorological maps and network regionalization problems, *Trudy NiU GUGMSI (18),* 1946.

Eagleson, P.S., and Q. Wang, Moments of catchment strom area, *Water Resources research 21 (8), 1185-1194*, 1985.

EarthInfo (Eds.), 1992, NOAA-NCDC, Hourly Precipitation - Western region, CD-ROM, Vol. 2.0, Published as CLIMATEDATA by EarthInfo Inc., Boulder Colorado.

Faures, J.M., Sensitivity of runoff to small-scale spatial variability of observed rainfall in a distributed model, M.S. Thesis, University of Arizona, Tucson, 169pp., 1990.

Fennessey, M. N., P. S. Eagleson, W. Qinliang, and I. Rodriguez-Iturbe, Spatial analysis of storm depths from an Arizona raingage network, *Department of Civil Engineering , Massachusetts Institute of Technology, Report No. 306*, 193pp, 1986.

Fletcher, J. E., Characteristics of precipitation (in the Rangelands) of the southwest, *Joint ARS-SCS Hydrology workshop*, Nwe Orleans, LA, 1960.

Gandin, L. S., Objective analysis of meteorological field, *Israel Program for Scientific Translations, Jerusalem*, 242pp. 1965.

Georgakakos, K. P., and M. L. Kavvas, Precipitation analysis, modeling, and precipitation in hydrology, *Rev. geophys, 25*, 163-178, 1987.

Goldstein, R., Satistical computing software reviews, *The American Statistician, 45(4)*, 305-311, 1991.

Goodrich, D.C, Basin scale and runoff model complexity, *Technical Report HWR No. 91-010*, 361pp., Department of Hydrology and Water Resources, University of Arizona, 1990.

Gupta, V. K., Preface for special section on Investigation of mesoscale precipitation fields, *Water Resour. Res., 21*, 1223-1225, 1985.

Hershfield, D. M., On the spacing of raingages, *Symposium on the the Design of Hydrological Networks, quebec 1965, Int. Assoc. scient. Hydrol. Publ. 67*, 72-79, 1965.

Hershfield, D. M., A note on areal rainfall determination, *Water Resources Bulletin, 5 (4)*, 49-55, 1969.

Huff, F. A., Spatial distribution of rainfall rates, *Water Resource Research, 6(1), 254-260*, 1970.

Hutchinson, P., A contribution to the problem of spacing raingages in rugged terrain, *J. of Hydrology, 12*, 1-14, 1970.

Isaaks, H. E., R. M. Srivastava, Applied geostatistics, Oxford University Press, New York, 1989.

Journel, A. G., and C. J. Huijbregts, Mining geostatistics, *Academic Press*, New York, 1978.

Kalma, J. D., J. Lomas, M. Thaller, and Y. Shashoua, An accurate small orifice raingage, *Water Resources Res. 5 (5)*, 300-305, 1969.

Kassim, A. H. M., and N. T. KOTTEGODA, Rainfall network design through comparative kriging methods, *Hydrological Sciences, 36, 3 (6)*, 223-240, 1991.

Lenton, R. L. and I. Rodriguez-Itube, A multidimensional model for the systems of processes of areal rainfall average, *Water Resources Research 13 (3)*, 605-612, 1977a.

Lenton, R. L. and I. Rodriguez-Itube, Rainfall network systems analysis: The optimation of total areal storm depth, *Water Resources Research 13(5)*, 825-836, 1977b.

Luo, J., and Y. Xing, Economic statistics analysis method and forecast (in Chinese), Qing Hua University Press, 270-299, 1987.

Matalas, N. C., Optimum gaging station location, *Pro. IBM Scientific Computing Symp. on Water and Air Resources Management*, 1967.

Matheron, G., The theory of regionalized variables and its application, *Cahiers du Centre de Morphologie Mathenatique Ecole des Mines*, Fonntainbleau, France, 211pp, 1971.

McConkey, B. G., W. Nicholaichuk and H. W. Cutforth, Small Area Variability of warm-season precipitation in a semiarid climate, *Agricultural and Forest Meteorology, 49*, 225-242, 1990.

Mejia, J. M. and I. Rodriquez-Iturbe, Multidimentional characterization of the rainfall process, *Ralph M. Parsons Laboratory Report No. 177*, MIT, Department of Civil engineering, 1973.

Mejia, J. M. and I. Rodriquez-Iturbe, On the synthesis of random field sampling from the spectrum: An application of generation of hydrologic spatial processes, *Water Resources Research 10 (4)*, 705-711, 1974.

Merva, G. E., N. D. Strommen, and E. H. Kidder, Rainfall variaton as influenced by wind and topography, *J. Appl. Meteor, 15 (7)*, 728-732, 1971.

Michaud, J.D. and S. Sorooshian,Rainfall-runoff modeling of flash floods in semi-arid watersheds, *Technical Report HWR No. 92-030*, Department of Hydrology and Water Resources, University of Arizona, 319pp., 1992, .

Neff, E. L., Principles of precipitaiton network disign for intensive hydrologic investigation, *IAHS, 67*, 49-55, 1965.

Neyman, J., and E. L. Scott, Some current problems of rain simulation research, *in Proceedings of the International Symposium on Uncertainties in Hydrologic and Water Resources Systems, 3*, 1167-1221, University of Arizona, Tucson, Ariz, 1972.

Niemczynowicz, J. and Jonsson, O., Extreme rainfall events in Lund 1979-1980, *Nord. Hydrol., 12 129-142, 1981.*

Orlanski, I., 1975, A rational subdivision of scale for atmospheric processes, *Bull. Am. Met. Soc. 56(5)*, 527-530, 1975.

Osborn, H. B., and L. J. Lane, Optimum gaging of thurderstorm rainfall in Southeastern Arizona, *Water Resources Research*, 8 (1), 259-265, 1972.

Osborn, H. B., and E. M. Laursen, Thunderstorm runoff in sourtheastern Arizona, *Proc. A. S. C. E., J. Hydraulic division, 99 (HY7)*, 1129-1145, 1973.

Osborn, H. B., K. G. Renard, and J. R. Simanton, Dense networks to measure convective rainfall in Southwestern United State, *Water Resources Research, 15 (6)*, 1701-1711, 1979.

Peck, E. L., Discussion of problems in measuring precipitation in mountainous areas, *WMO Publ. 326 (1)*, 5-16, 1973.

Rhodes, H. R., and D. E. Myers, Correspondence analysis used in the evaluation of Lakewater chemistry in the Adirondacks, *Journal of Chemometrics, 5*, 273-290, 1991.

Roche, M., Hydrologie de Surface, *Ganthier-Villars*, Paris, 430pp., 1963.

Rodriguez-Iturbe, I. and V. M. Mejia, On the transformation a point rainfall to areal rainfall, *Water Resources Research 10 (4)*, 729-735, 1974.

Rouhani, S. and D.E. Myers, 1990, Problems in space-time Kriging of geohydrological data, *Mathematical Geology, 22 (5),* 611-623, 1990.

Rycroft, H. B., Random sampling of rainfall, *J. S. Afr. For. Ass., 18,* 71-81, 1947.

Sharon, D., Spatial analysis of rainfall data from dense networks, *Bulletin of the International Association of Hydrological Sciences, XVII, 3 (10),* 291-299, 1972.

Sellers, W. D., and R. H. Hill, Arizona climate, the University of Arizona Press, Tucson, Arizona, 6-19, 1974.

Shilling, W. and Harms, R. W., Raumliche Variabilitat von Niederschlag und Abfluszbildung, *Auswirkungen auf den Alofuszprozesz, Dtsche, Gewasserk, Mitt.,* 1983.

Shilling, W., A quantitative assessment of uncertainties in stormwater modeling, *Proc. Inc. Cont. Urban Storm Drain., Chalmers Univ. Technol., Gothenburg,* 625-633, 1984a.

Shilling, W., Short-term flow forecasting for real time stormwater control, *Proc. Inc. Cont. Urban Storm Drain., Chalmers Univ. Technol., Gothenburg,* 1111-1119, 1984b.

Shih, S. F., Rainfall variation analysis and optimization of gaging system, *Water Resources Research, 18 (4),* 1269-1277, 1982.

Sun, H., Geostatistics and its application (in Chinese), China Mining University Press, 282pp, 1990.

Tabios III, Q. G., and J. d. Salas, a comparative analysis of techniques for spatial interpolation of precipitation, *Water Resources Bulletin, 21 (3),* 365-380, 1985.

Thiessen, A H., Precipitation averages for large areas, Monthly Weather Review, 39(7), 1082-1084, 1911

Tian, D., S. Sorooshian, and D. E. Myers, Correspondence Analysis with Matlab, *Computers and Geosciences,* in press.

Troutman, B. M., Runoff prediction errors and bias in parameter estimation induced by spatial variability of precipitation, *Water Resources Research, 19 (3),* 791-810, 1983.