AN ABSTRACT OF THE THESIS OF

Joseph A. Hevesi for the degree of Master of Science in Geology presented on May 22, 1990.

Title: PRECIPITATION ESTIMATION IN MOUNTAINOUS TERRAIN USING MULTIVARIATE GEOSTATISTICS:

Abstract approved: Redacted for Privacy

Jonathan D. Istok

ABSTRACT

Estimates of average annual precipitation (AAP) are needed for hydrologic modeling at Yucca Mtn., Nevada, site of a proposed, high-level nuclear waste repository. Historical precipitation data and station elevation were obtained for stations in southern Nevada and southeastern California. Elevations for 1,531 additional locations were obtained from topographic maps. The sample direct-variogram for the transformed variable TAAP = ln(AAP) * 1000 was fit with an isotropic, spherical model with a small nugget and a range of 190,000 ft. The sample direct-variogram for elevation was fit with an isotropic model with four nested structures (nugget, Gaussian, spherical, and linear) with ranges between 0 and 270,000 ft. There was a significant ($p = 0.05$, $r = 0.75$) linear correlation between TAAP and station elevation. The sample cross-variogram for TAAP and elevation was fit with two nested structures (Gaussian, spherical) with ranges from 55,000 to 355,000 ft. Alternate model structures and parameters were compared using cross-validation.

Isohyetal maps for average annual precipitation (AAP) were prepared from estimates obtained by kriging and cokriging using the selected models. Isohyets based on the kriging estimates were very smooth, increasing gradually from the southwest to the northeast. Isohyets based on the cokriging estimates and the spatial correlation between AAP and elevation were more irregular and displayed known orographic effects. Indirect confirmation of the cokriging estimates were obtained by comparing isohyets prepared with the cokriging estimates to the boundaries of more densely vegetated and/or forested zones. Estimates for AAP at the repository site were 145 and 165 mm for kriging and cokriging, respectively. Cokriging reduced estimation variances at the repository site by 55% relative

to kriging. The effectiveness of an existing network of stations for measuring AAP is evaluated and recommendations are made for optimal locations for additional stations.

Precipitation Estimation in Mountainous Terrain Using Multivariate Geostatistics

by

Joseph A. Hevesi

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Completed May 22, 1990

Commencement June 1991

APPROVED:

Redacted for Privacy

Professor of Civil Engineering in charge of major

Redacted for Privacy

Head of Geosciences Department

Redacted for Privacy

Dean of Graduate School

Date thesis is presented May 22, 1990

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

PRECIPITATION ESTIMATION IN MOUNTAINOUS TERRAIN USING MULTIVARIATE GEOSTATISTICS

INTRODUCTION

Yucca Mtn., Nevada, is being considered as a possible site for the nation's first high-level nuclear waste repository. Values of average annual precipitation (AAP) are needed for water balance calculations and for defining recharge boundary conditions for groundwater flow models. The length of record for existing precipitation stations at Yucca Mtn. is short (2 to 4 years at the time of this study), and AAP values must be estimated using surrounding southern Nevada and southeastern California stations with longer lengths of record (14 to 53 years). Unfortunately, the areal coverage provided by this network of stations is sparse, due to the rugged mountainous terrain and low population density of the region.

Two techniques commonly used to estimate AAP are interpolation and regression. Interpolation includes the Thiessen polygon method, inverse distance weighting, and geostatistical methods (e.g. kriging and cokriging). Regression can be used when AAP is dependent on one or more easily measured variables, such as elevation or location. The orographic influence of rugged mountainous terrain in the southern Nevada and southeastern California region, where local relief varies from 2,000 to 10,000 ft., results in a significant positive correlation between precipitation and ground elevation. French (1986) defined linear and log-linear relationships between AAP and station elevation for 63 precipitation stations in the southern Nevada and southeastern California. Quiring (1983) used regression to define linear and log-linear relationships between AAP and elevation for stations on the Nevada Test Site (NTS).

Previous studies have utilized univariate geostatistics to estimate AAP at other locations. Tabios and Salas (1985) compared several interpolation methods for estimating AAP and concluded that ordinary and universal kriging were superior to other methods in terms of average estimation accuracy. Dingman et al. (1988) applied kriging to estimate AAP in New Hampshire and Vermont, using an "elevation scale factor" to remove an observed orographic influence on measured values of AAP. However, an application of multivariate geostatistics, utilizing ground elevation as an additional, correlated variable to improve areal estimates of AAP, has not been previously reported.

The two parts of this study present an application of both univariate and multivariate geostatistics to the estimation of AAP at Yucca Mtn. and within a surrounding drainage basin, the Upper Amargosa River Watershed (UARW). Geostatistics is chosen as the method of estimation because: 1) geostatistics theoretically provides the best linear unbiased estimates using available measurements of AAP, where the term "best" refers to a minimization of the estimation variance, 2) a measure of the reliability of each estimated value, in terms of a variance, can be calculated, 3) using the multivariate method, elevation data can be incorporated into the estimation equations, with the potential of increasing estimation accuracy and detail, 4) the developed models of spatial variability can be used to help identify optimum locations for additional precipitation stations, 5) for future studies, the geostatistical model can be expanded to include several additional correlated variables and can also be modified to include temporal variability, 6) geostatistics has not been previously used to estimate AAP at Yucca Mtn.

The overall objective is to provide improved estimates of AAP and to evaluate the practical application of multivariate geostatistics when AAP is "extremely undersampled" relative to elevation. The advantage of utilizing elevation to estimate AAP is that elevation data are much easier, faster, and less expensive to obtain than measurements of AAP. Improvements in estimation accuracy can be measured by relative reductions in calculated estimation variances, provided that assumptions concerning the derived geostatistical models are correct. The reduction in estimation variances should increase as: 1) the number of additional data obtained by using the correlated variable increases, 2) the magnitude of spatial cross-correlation between the two variables increases, 3) the distance of the additional data from the point to be estimated decreases, 4) the arrangement of the additional data around the point to be estimated becomes optimized.

Chapter I of this study concerns the characterization and modeling of the spatial correlations and the spatial cross-correlation of AAP and elevation for the UARW. An important objective is to define model variograms, using structural analysis and the cross-validation technique presented by Delhomme (1978) and Cooper and Istok (1988a). A second objective is to compare the cross-validation statistics obtained with kriging and cokriging using alternate variogram models, and also with results obtained using neighborhood averaging, inverse-distance-weighting, and regression.

Chapter II describes the use of the multivariate geostatistical model defined in Chapter 1 to develop isohyetal maps of AAP for Yucca Mtn. and the UARW. Important objectives include estimating AAP at the location of the proposed repository, determining the global average AAP estimate for the area within the UARW, and identifying areas of maximum and minimum estimates within the UARW. Average, maximum, and minimum estimation variances obtained using multivariate and univariate geostatistics are compared. Calculated estimation variances are utilized for identifying optimum locations for additional precipitation stations and also for predicting the efficiency of future sampling networks. Estimates and calculated estimation variances obtained using several selected geostatistical models, as well as results obtained using inverse-distance and regression methods, are compared in an effort to evaluate the variability of results as a function of both the selected model and the selected estimation technique.

Contribution of Author

Joseph Hevesi was the principal author of this work. Technical guidance and editorial comments were provided by the two coauthors.

PRECIPITATION ESTIMATION IN MOUNTAINOUS TERRAIN USING MULTIVARIATE GEOSTATISTICS: 1. STRUCTURAL ANALYSIS

Joseph A. Hevesi
Fenix & Scisson of Nevada,
P.O.Box 327, MS 721, Mercury, Nevada 89023

Alan L. Flint
U. S. Geological Survey, Water Resources Division,
P.O.Box 327, MS 721, Mercury, Nevada 89023

Jonathan D. Istok
Department of Civil Engineering, Oregon State University,
Corvallis, Oregon 97331

Introduction

Precipitation data are required for nearly all hydrologic analyses and one of the fundamental problems of hydrology is to estimate precipitation at a site using data from surrounding precipitation stations. Obtaining reliable estimates is particularly difficult when the areal coverage provided by the surrounding stations is sparse or when precipitation charactersitics vary rapidly with position. This situation is frequently encountered in mountainous terrain when few stations are available and orographic effects are large. An example of such an area is Yucca Mtn., Nevada, the potential site for the nation's first high-level, nuclear waste repository. Values of average annual precipitation (AAP) are needed for water balance calculations and for defining recharge boundary conditions for groundwater flow models. However, the length of record for precipitation stations at Yucca Mtn. is short (2 to 4 years) and AAP must be estimated from stations in southern Nevada and southeastern California with longer periods of record (14 to 53 years).

Tabios and Salas (1985) compared several methods for estimating AAP and concluded that ordinary and universal kriging were superior to Thiessen polygon, polynomial interpolation, and inverse-distance weighting methods. Kriging provides unbiased, linear estimates with minimized estimation variances and has been widely used to estimate climatological data including AAP (Dingman et al., 1988). An important factor in the choice of estimation method is the observed correlation between AAP and elevation for southern Nevada (Quiring, 1983; French, 1986). Cokriging is a multivariate geostatistical method that utilizes the spatial correlation between two or more variables to reduce estimation variances when one of the variables is undersampled (David, 1977). The magnitude of the reduction depends on the correlation between the variables and on the data configuration.

The two chapters in this thesis present an application of kriging and cokriging to the problem of estimating AAP at Yucca Mtn. The overall objective is to determine if a multivariate geostatistical approach based on correlations between AAP and elevation can be used to improve estimation detail and reduce estimation variances for AAP for Yucca Mtn. An application of multivariate geostatistics, using elevation to estimate AAP has not previously been reported.

The first chapter describes the characterization and modeling of the spatial correlations and cross-correlation for AAP and elevation for an area in southern Nevada and southeastern California that contains Yucca Mtn. The second chapter describes the use of the geostatistical models to develop isohyetal maps for AAP at Yucca Mtn.

Study Area

Geographic and Physiographic Setting

Yucca Mtn. is located in Nye county, Nevada, along the western border of the Nevada Test Site (NTS) (Figure 1). The proposed waste repository is an area of approximately 2 mi^2 beneath the southern extent of Yucca Mtn. The watershed containing Yucca Mtn. is defined, for the purposes of this study, as the Upper Amargosa River Watershed (UARW) (Figures 1 and 2). The area of the watershed is approximately 2,600 mi^2.

Physiography within the UARW varies from the near horizontal depositional surfaces of the Amargosa Desert to the rugged escarpments and steep slopes formed by fault scarps and erosional surfaces in the surrounding mountain ranges. Elevations in the UARW range from 2,000 ft in the Amargosa Desert to 7,640 ft on Pahute Mesa. Local relief (defined here as the change in elevation over a distance of 50,000 ft) within the UARW ranges from near zero in the Amargosa Desert to 4,000 ft in the Spring and Funeral Mtns. Local relief in the vicinity of Yucca Mtn. ranges from 1,000 to 3,000 ft. Chocolate Mtn. is the highest point of the Yucca Mtn. massif at an elevation of 6,703 ft. Elevations in the area of the proposed repository range from 3,980 to 4,940 ft.

Climate

The regional climate in the vicinity of Yucca Mtn. is classified as a midlatitude - desert, with AAP less than 300 mm in most locations. Winter (October to March) precipitation is primarily due to frontal storm systems traveling eastward from the Pacific Ocean. Summer precipitation is primarily due to localized convective-type storms, with possible sources of moisture being the Gulf of Mexico to the southeast and the Gulf of California to the south. After analyzing measurements of AAP from 63 southern Nevada and southeastern California precipitation stations, French (1983a) concluded that the combined effects of the Sierra Nevada rain shadow and the circulation of moisture from the Gulf of Mexico caused a trend of increasing precipitation from west to east across southern Nevada. According to French (1983a), Yucca Mtn. is located within the transition zone between a moisture deficit zone to the west and a moisture excess zone to the east.

Elevation - Precipitation Correlation

Locally, the climate at higher elevations in southern Nevada is more humid due to the orographic influence of surrounding mountainous terrain and to the "virgis" effect, the depletion by evaporation of the falling rain (Jones, 1981). A change in vegetation type and an increase in vegetation density with elevation, offers indirect evidence of increased

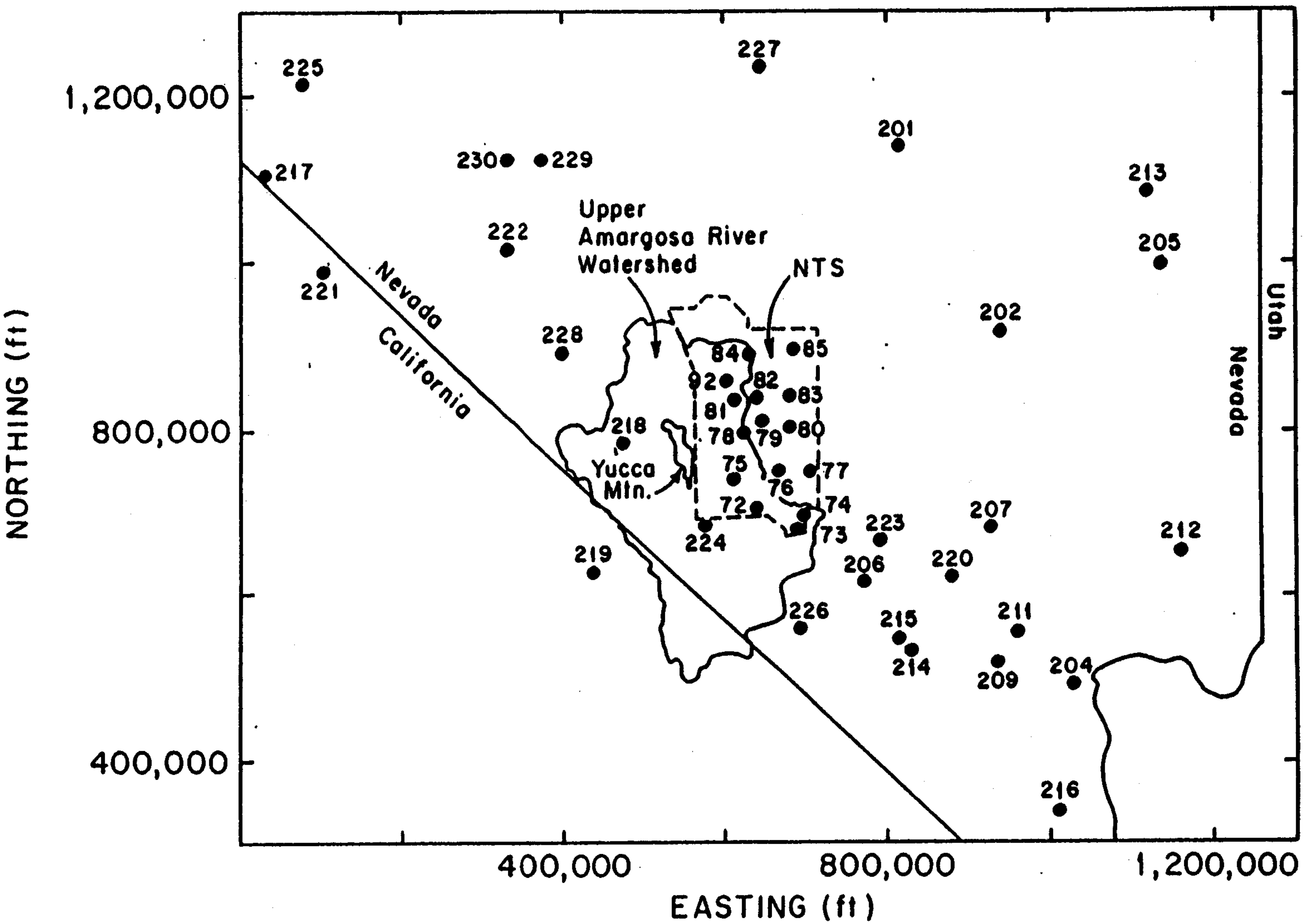

Figure I.1 Location of study area and precipitation stations.

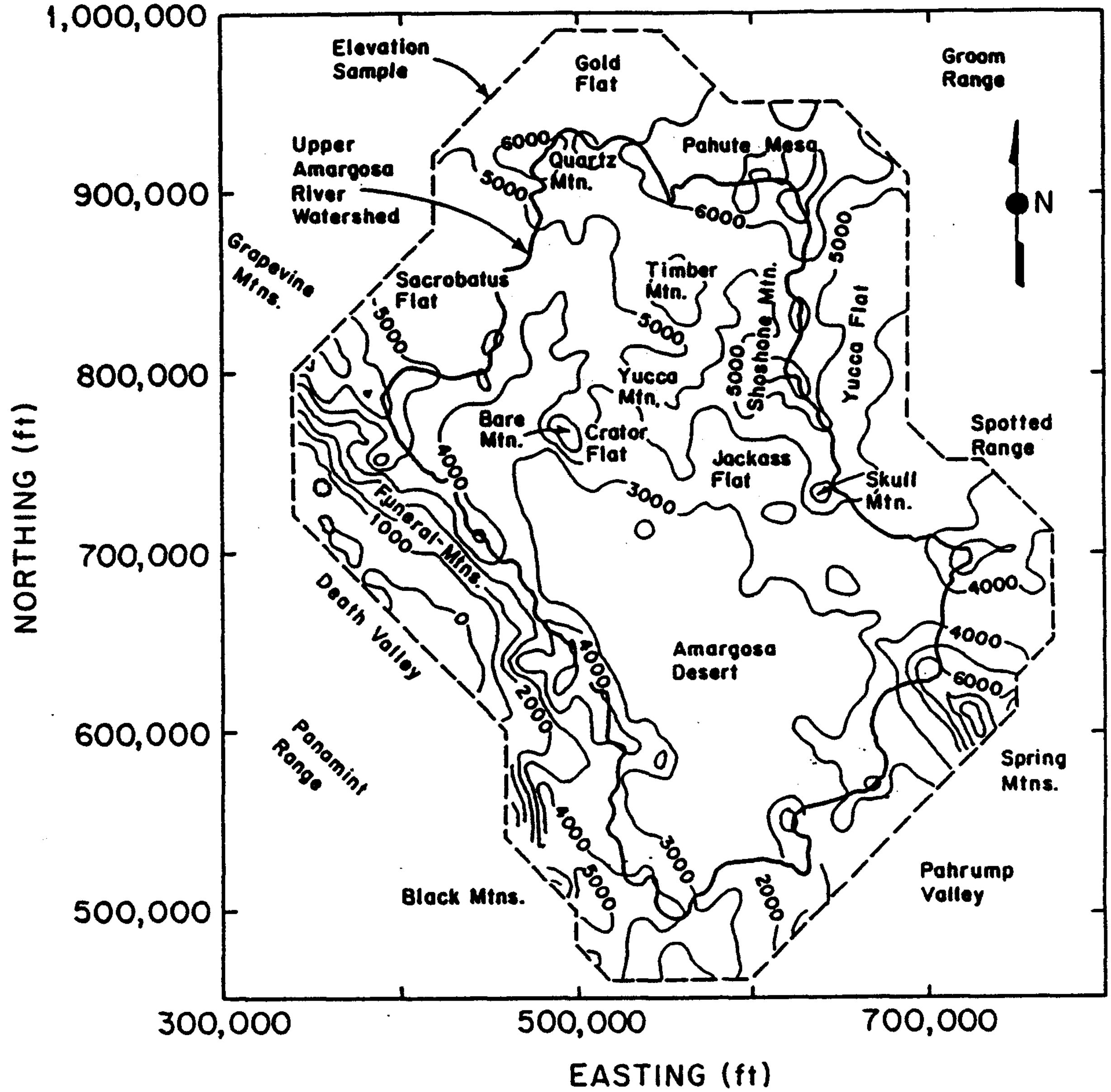

Figure I.2 Topographic map of UARW showing boundary of sampling grid for elevation.

precipitation with increased elevation. Sparce juniper and pinyon pine become prevalent above elevations of roughly 6,000 ft. Forests of ponderosa pine, aspen, and bristlecone pine occur above 8,000 ft in the Spring Mtns. Mountain mahogony, pinyon pine, and juniper occur on Telescope Peak in the Panamint Range. Quiring (1983) calculated a correlation coefficient r = 0.90 between AAP and elevation for 11 stations on the NTS, using both linear and log-linear relationships. French (1986) examined the precipitation - elevation correlation for 12 stations on the NTS with 10 or more years of complete record, and computed r values of 0.92 between AAP and elevation, and 0.87 between log(AAP) and elevation.

Geostatistical Theory

The term geostatistics refers to a collection of statistical methods for describing and modeling the correlation structure of regionalized variables (ReV), i.e., random variables distributed in space and time. The theoretical basis for these methods has been described in detail elsewhere (David, 1977; Journel and Huijbregts, 1978; Isaaks and Srivastava, 1989); only those aspects of theory relevant to the estimation of AAP are presented here.

The set of sample values for a regionalized variable is considered to be a single realization of a random function. In this study the two random functions of interest, $Z_i(x)$ and $Z_j(x)$, are AAP and elevation respectively. The spatial correlation structure for a pair of random functions is defined by the variogram

$$\gamma_{ij}(h) = \frac{1}{2} E \{[Z_i(x+h) - Z_i(x)]\,[Z_j(x+h) - Z_j(x)]\} \qquad (1)$$

where E is the expectation operator and x and x + h are a pair of points separated by the vector h. If i = j, Eq. 1 defines the direct-variogram for Z_i. If $i \neq j$, Eq. 1 defines the cross-variogram for Z_i and Z_j. In this study, Eq. 1 is used to define a direct-variogram for AAP, a direct-variogram for elevation, and a cross-variogram for AAP and elevation.

Because the available data represent only one realization of the random functions, estimating the direct- and cross-variograms is not possible unless certain simplifying assumptions are made. In this work we will assume that the expectations of Z_i and Z_j are stationary, i.e., $E[Z_i(x)] = m_i$ and $E[Z_j(x)] = m_j$ where m_i and m_j are constants, within a circular neighborhood of the point to be estimated. We will further assume that the variograms defined by Eq. 1 are stationary. These two assumptions constitute the intrinsic hypothesis and permit the calculation of sample variograms (see Methods) that can be used to estimate the direct- and cross-variograms for AAP and elevation. Mathematical models can then be selected to represent the spatial structure displayed by the sample variograms.

The selected models must be positive definite. For a single regionalized variable this condition can be met by selecting "authorized" model functions (e.g. spherical, linear, exponential, and gaussian). For the cross-variogram used in this study this condition can be met by choosing model parameters so that the Cauchy-Schwarz inequality is satisfied for all values of h (Journel and Huijbregts, 1978)

$$\gamma_{ij}(h) \leq \sqrt{\gamma_{ii}(h)\,\gamma_{jj}(h)} \qquad (2)$$

where $\gamma_{ij}(h)$ is the cross-variogram model for AAP and elevation, and $\gamma_{ii}(h)$ and $\gamma_{jj}(h)$ are the direct-variogram models for AAP and elevation, respectively.

The linear estimator for AAP at an unsampled location x_o consists of a weighted combination of the available AAP and elevation data

$$Z_i^*(x_o) = \sum_{k=1}^{n_i} \lambda_{ik}\, z_i(x_k) + \sum_{l=1}^{n_j} \lambda_{jl}\, z_j(x_l) \tag{3}$$

where $Z_i^*(x_o)$ is the estimate for AAP at x_o, λ_{ik} is the weight assigned to the measured value of AAP at location x_k (x_k, $k = 1$, n_i are the locations of the precipitation stations), and λ_{jl} is the weight assigned to the measured value of the elevation at location x_l (x_l, $l = 1$, n_j are the locations of the precipitation stations and the 1531 grid points).

The weights in Eq. 3 are obtained by minimizing the estimation variance

$$\mathrm{Var}\{Z_i^*(x_o) - Z_i(x_o)\} \tag{4}$$

with the constraint that the estimate be unbiased i.e.,

$$E\{Z_i^*(x_o) - Z_i(x_o)\} = 0 \tag{5}$$

The result is a system of linear equations, the cokriging system (Yates and Warrick, 1987)

$$\begin{aligned}
&\sum_{i=1}^{n_i} \lambda_{ik}\, \gamma_{ii}(x_m, x_k) + \sum_{l=1}^{n_j} \lambda_{jl}\, \gamma_{ij}(x_m, x_l) + \mu_i = \gamma_{ii}(x_o, x_m),\ m = 1, n_i \\
&\sum_{i=1}^{n_i} \lambda_{ik}\, \gamma_{ij}(x_m, x_k) + \sum_{l=1}^{n_j} \lambda_{jl}\, \gamma_{jj}(x_m, x_l) + \mu_j = \gamma_{ij}(x_o, x_m),\ m = 1, n_j \\
&\sum_{k=1}^{n_i} \lambda_{ik} = 1 \\
&\sum_{l=1}^{n_j} \lambda_{jk} = 0
\end{aligned} \tag{6}$$

where μ_i and μ_j are Lagrange multipliers. The minimized estimation variance at the unsampled location x_o, $\sigma^2(x_o)$ is

$$\sigma^2(x_o) = \mu_i - \sum_{k=1}^{n_i} \lambda_{ik}\, \gamma_{ii}\, (x_k, x_o) + \sum_{l=1}^{n_j} \lambda_{jl}\, \gamma_{ij}\, (x_l, x_o) \qquad (7)$$

When n_j = 0, Eqs. 6 and 7 define the kriging estimate and estimation variance for regionalized variable i (AAP) otherwise Eqs. 6 and 7 define the cokriging estimate and estimation variance for AAP.

Methods

Precipitation Data

Records of AAP and station elevation for precipitation stations in southern Nevada and southeastern California were obtained from French (1983,1986,1987) and from the unpublished records for two precipitation stations on Yucca Mtn. Eighteen stations were located within the UARW. Many of the stations had been moved during their period of record, and corrections were made by French using a double mass analysis. The length of record was between 2 and 53 years. The maximum and minimum values of AAP are 544 and 69 mm for station 208 in the Spring Mtns. and station 219 in Death Valley, respectively.

Elevation Data

An additional 1,531 ground elevations were manually recorded from 1 : 250,000 scale topographic maps (Caliente, Death Valley, Goldfield, and Las Vegas). The data were recorded at points on a square grid with a spacing of 10,000 ft. The grid boundary extended approximately 50,000 ft beyond the boundary of the UARW to reduce extrapolation and edge effects on estimates of AAP made close to the watershed boundary (Renard and Yancey, 1984; Journel and Huijbregts, 1978). Inspection of contour maps prepared from the elevation sample indicated that the elevation sample accurately represented major physiographic features within the UARW (e.g., the northeast to southwest alignment of basin and range structure).

Ideally, the areal extent of the sampled domain for AAP and elevation should be identical to ensure consistency in the analysis of the spatial correlations and cross-correlations. However, two factors prevented the use of an elevation grid with the same areal extent as the precipitation stations: 1) the sample size for elevation with the selected grid spacing of 10,000 ft. would have been too large, and 2) major changes could be identified in the characteristics of the physiography within southern Nevada and southeastern California, causing the structure of spatial variability for elevation on this scale to be non-stationary (also see the discussion of the sample direct-variogram for elevation).

Preliminary Statistical Analyses

The sample distributions for AAP and elevation were tested for normality by inspection of histograms and normal probability plots. Sample means and variances were calculated for use in fitting models to the direct-variograms and for use in calculating cross-validation statistics and limits of model validity. Sample covariances and correlation coefficients were also calculated and used to fit models to the sample cross-variogram.

Sample Variograms

Sample variograms were computed using (e.g., Journel and Huijbregts, 1978)

$$\gamma^*_{ij}(h) = \frac{1}{2n_i} \sum_{k=1}^{n_i} [z_i(x_k + h) - z_i(x_k)]\,[z_j(x_k + h) - z_j(x_k)] \tag{8}$$

where z_i and z_j are the measured values of AAP and elevation respectively. Calculations were performed using the computer program VARIO (Cooper et al., 1988).

Model Fitting

A set of alternate direct- and cross-variogram models were fitted to the sample variograms with the criteria that each model be positive definite. In this study, nested structures were evident in the sample variograms. Parameters of the variogram models for each structure were selected that satisfied the Cauchy-Schwarz inequality (Eq. 2). Selecting parameters of the cross-variogram model that satisfy Eq. 2 was simplified by plotting the function PDC (for positive definite condition)

$$PDC = \sqrt{\gamma_{ii}(h)\,\gamma_{jj}(h)} \tag{9}$$

on the sample cross-variogram for AAP and elevation. In Eq. 9, $\gamma_{ii}(h)$ and $\gamma_{jj}(h)$ are the model direct-variograms for AAP and elevation, respectively. Using Eq.9, it was easy to determine if parameters of a proposed model cross-variogram were positive-definite, regardless of the number of nested structures. The proposed model was simply plotted with PDC and a check was made that the value of $\gamma_{ij}(h)$ did not exceed the value of PDC for all h.

Cross-Validation Procedure

Cross-validation statistics were used to select model direct- and cross-variograms from the set of alternate models (Delhomme, 1979). The procedure involves deleting a sample from the dataset of ReV, and then using kriging or cokriging to estimate the value of that ReV at the location of the deleted sample, using the remaining samples and the chosen variogram model(s) and parameters. Differences between estimated and measured values were summarized using the cross-validation statistics: Percent Average Estimation

Error (PAEE), Relative Mean Squared Error (RMSE), and Standardized Mean Squared Error (SMSE) (Cooper and Istok, 1988).

Estimates were considered unbiased if the PAEE was close to zero:

$$\mathrm{PAEE} = \frac{100\%}{\bar{z}_i n_i} \sum_{k=1}^{n_i} Z_i^*(x_k) - z_i(x_k) \tag{10}$$

where $Z_i^*(x_k)$ is the estimated value of ReV Z_i at location x_k, $z_i(x_k)$ is the deleted sample value for Z_i at x_k and $\bar{z}_i$ is the sample mean.

Estimates were considered accurate if the RMSE was close to zero:

$$\mathrm{RMSE} = \frac{1}{s^2 n_i} \sum_{k=1}^{n_i} [Z_i^*(x_k) - z_i(x_k)]^2 \tag{11}$$

where s^2 is the sample variance.

The SMSE indicates the consistency of the calculated estimation variances with the observed RMSE. The estimation variances were considered consistent if SMSE was in the range $1 \pm 2\sqrt{2n}$ (Delhomme, 1979).

$$\mathrm{SMSE} = \frac{1}{n_i} \sum_{k=1}^{n_p} \frac{[Z_i^*(x_k) - z_i(x_k)]^2}{\sigma_K^2 (x_k)} \tag{12}$$

Results and Discussion

Precipitation Data Base

In general, standard errors for AAP were greater than 20 mm for stations with lengths of record of less than 25 years (Figure I.3). French (1987), in an analysis of precipitation data for 11 NTS stations, concluded that a standard error less than 20 mm for AAP required a length of record of at least 31 years. Only 12 stations available for this study had a length of record of 31 years or longer which is too small a sample for an analysis of spatial variability. An adequate sample size (42 stations) was obtained by selecting stations with a length of record of at least 8 years (Figure I.3). The locations of the selected stations are in Figure I.1; a summary of the data is in Table 1. Standard errors for these stations ranged from 47.4 mm for station 215, with a length of record of 8 years, to 6.6 mm for station 219, with a length of record of 18 years. Note the minimum value of 69 mm for station 219 in Death Valley and the maximum value of 354 mm for station 215 in the Spring Mtns. Station elevations ranged from 7,549 ft for station 207 in the Sheep Mtns. to -167 ft in Death Valley. The dataset included 10 stations within the boundaries of the UARW. Values of AAP for these stations ranged from 295 mm for station 84 (elevation 7,490 ft) to 85 mm for station 224 (elevation 2,175 ft). Station 75 was closest to the proposed repository site on Yucca Mtn., with an AAP value of 118 mm, and an elevation of 3,422 ft.

Preliminary Statistical Analysis

There was a significant non-linear correlation between AAP and elevation (Figure I.4). The regression equation:

$$AAP = \exp(4.32 + 0.00018\,E), \qquad r^2 = 0.56 \tag{13}$$

was fit to the data where E is the station elevation in feet and AAP is in mm.

French (1986) developed the following regression equation using data from 63 stations in southern Nevada:

$$AAP = \exp(4.31 + 0.00016E), \qquad r^2 = 0.48 \tag{14}$$

Recognizing the poor reliability of AAP values for stations with lengths of record as short as two years, French analyzed the correlation between AAP and elevation for 12 NTS stations with 10 or more years of record, and developed the following equation:

$$AAP = \exp(4.38 + 0.00019E), \qquad r^2 = 0.76 \tag{15}$$

Table I.1 Precipitation data used to estimate AAP.

Station No.	Station Name	Nevada Central Coordinates Easting (ft)	Nevada Central Coordinates Northing (ft)	Elevation (ft)	Average Annual Precipitation (mm)	Length of Record (yrs)	Standard Deviation (mm)
72	ROCK VALLEY	639,000	705,000	3399	157	8	23
73	DESERT ROCK A.P.	688,000	682,000	3297	157	15	17
74	MERCURY	697,000	695,000	3770	159	13	18
75	4JA	611,000	741,000	3422	119	16	15
76	CS	664,000	751,000	4000	208	21	21
77	W5B	705,000	748,000	3081	128	19	12
78	SB	621,000	795,000	5660	225	12	25
79	MID VALLEY	645,000	809,000	4659	226	13	23
80	UCC	681,000	804,000	3921	177	25	17
81	40MN	611,000	837,000	4820	202	18	18
82	TS2	639,000	839,000	4981	245	20	23
83	BJY	679,000	842,000	4072	177	22	17
84	AREA 12	631,000	889,000	7491	295	17	31
85	PHS	683,000	896,000	4564	193	8	22
92	LITTLE F. 2	602,000	859,000	5161	229	8	36
201	ADAVEN	816,000	1,138,000	6398	321	47	16
202	ALAMO	938,000	918,000	3442	128	26	10
204	BOULDER CTY.	1,029,000	490,000	2526	139	50	9
205	CALIENTE	1,134,000	996,000	4406	231	51	11
206	COLD CREEK	770,000	618,000	6001	230	8	33
207	HIDDEN FOREST	927,000	681,000	7550	320	9	27
209	LAS VEGAS A.P.	936,000	520,000	2162	104	33	8
211	SUNRISE MR	961,000	554,000	1821	106	32	8
212	OVERTON	1,157,000	652,000	1221	91	26	13
213	PIOCHE	1,119,000	1,084,000	6119	313	44	16
214	RED ROCK	829,000	535,000	6500	270	8	36
215	ROBERTS RNC.	814,000	544,000	6099	354	8	47
216	SEARCHLIGHT	1,010,000	339,000	3540	185	50	14
217	BASALT	27,000	1,104,000	6349	142	15	17

Table I.1 Continued.

Station No.	Station Name	Nevada Central Coordinates Easting (ft)	Nevada Central Coordinates Northing (ft)	Elevation (ft)	Average Annual Precipitation (mm)	Length of Record (yrs)	Standard Deviation (mm)
218	BEATTY	473,000	785,000	3550	159	47	12
219	DEATH VALLEY	437,000	628,000	-167	69	18	7
220	DESERT G.R.	878,000	623,000	2920	106	42	8
221	DYER	101,000	991,000	4974	125	31	9
222	GOLDFIELD	331,000	1,016,000	5689	136	39	12
223	INDIAN SPRGS	790,000	667,000	3137	116	25	16
224	LATHROP	574,000	686,000	2175	85	21	11
225	MINA	76,000	1,216,000	4551	115	53	7
226	PAHRUMP	692,000	559,000	2700	126	20	16
227	RATTLESNAKE	644,000	1,236,000	5912	121	20	13
228	SACROBAT	400,000	893,000	4019	90	14	13
229	TONOPAH A.P.	375,000	1,123,000	5427	163	29	10
230	TONOPAH	331,000	1,123,000	6093	126	22	11

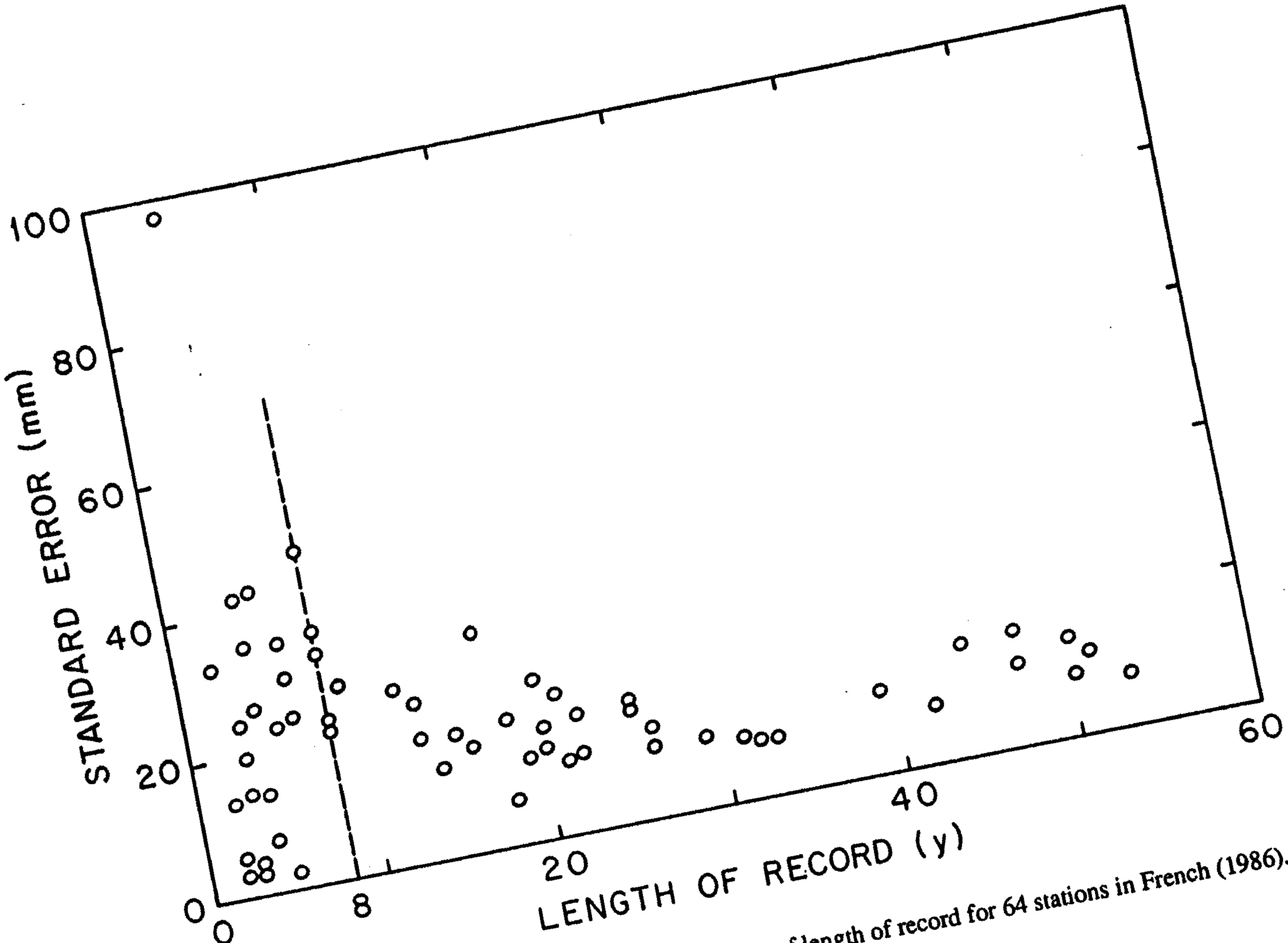

Figure I.3 Standard error for AAP as a function of length of record for 64 stations in French (1986).

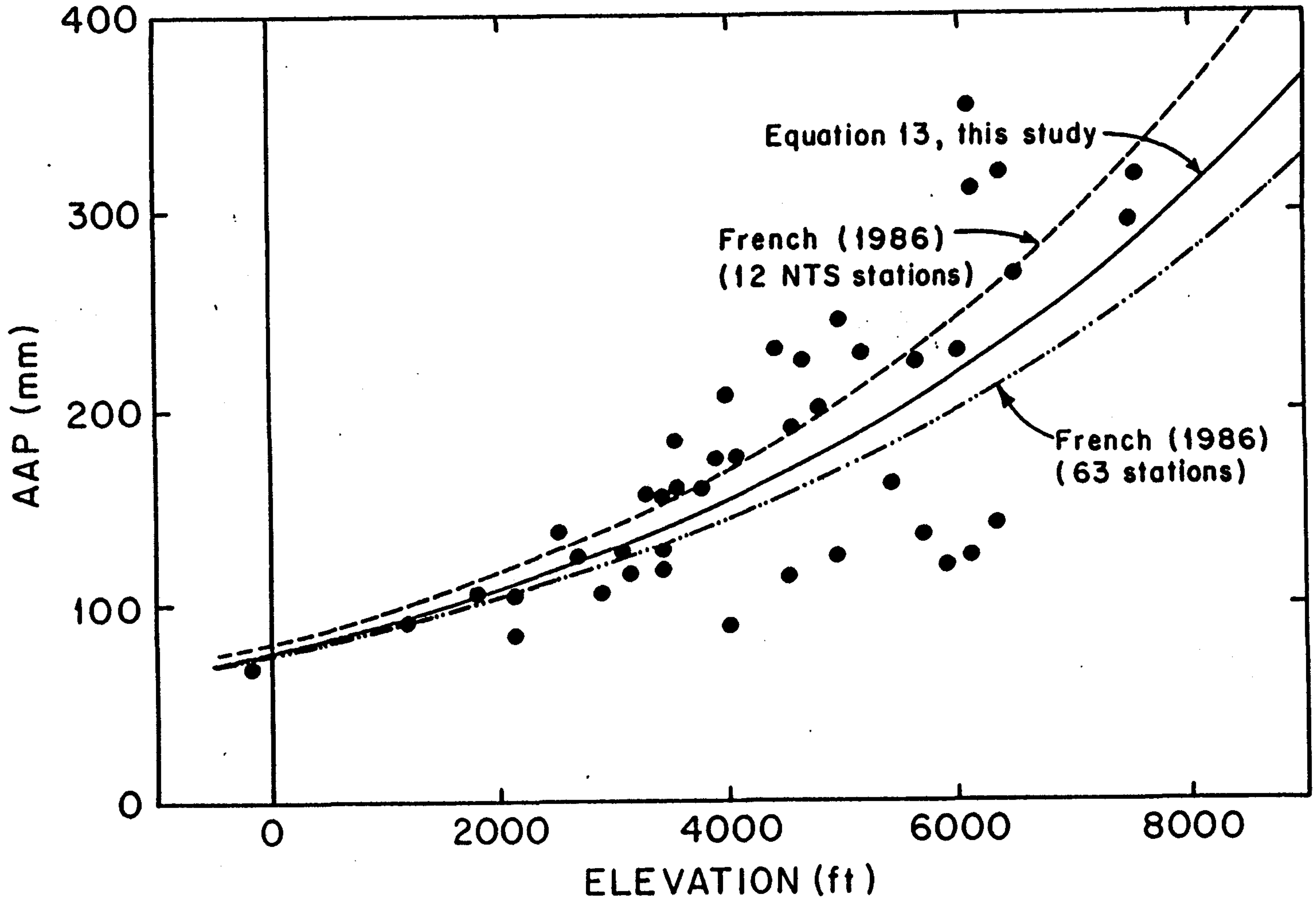

Figure I.4 Regression equations for AAP as a function of station elevation.

Equation 15 should apply in the vicinity of Yucca Mtn., because these NTS stations are much closer to Yucca Mtn. than the stations used to develop Eqs. 13 and 14.

The regression lines are compared in Figure 4 and indicate a poorer fit with increased station elevation. For example, at an elevation of 7,000 ft, estimates of AAP range from 301 mm (Eq. 15) to 228 mm (Eq. 14). This may be due to: 1) the small number of stations below 2,000 ft., 2) a "rain shadow" effect caused by the location of stations on the lee slope of higher topography or adjacent to blocking ranges or mountains, 3) regional trends in AAP, such as the west to east increase hypothesized by French (1983a), 4) increased measurement error with increased elevation, due to increasing amounts of snow, and 5) increased sampling variability at higher elevations.

Histogram and normal probability plots indicated that the values of AAP were not normally distributed. The transformation Y = log (AAP) improved the fit of the frequency histogram to a normal distribution. The standardized skewness was reduced from 2.19 to 0.34 by the transformation (Table 2). A log-normal distribution is typical for AAP data in arid climates (Jones, 1981). The linear transformation: TAAP (for transformed AAP) = ln(AAP)*1000 was performed to reduce numerical errors during cross-validation, kriging, and cokriging.

The elevation data were approximately normally distributed, with the exception of a small bi-modal characteristic caused by a few negative elevations in the eastern half of Death Valley resulting in a standardized skewness of -3.89 for the elevation sample (Table 2). Inspection of the normal probability plot, however, showed an approximate overall fit to a normal distribution and no attempt was made to improve this fit by a transformation.

Sample Direct-Variogram for TAAP

Because of the small sample size for TAAP, directional variograms could only be computed using very wide angle and distance classes (Cooper et al., 1988) and this resulted in a loss of directional resolution. For this reason, isotropism was assumed and an omnidirectional sample variogram was computed. This variogram indicated a range of spatial correlation for TAAP of about 200,000 ft (Figure 5) and there was no evidence of drift for distances up to 641,000 ft. This supports the assumption of second-order stationarity for TAAP.

Sample Direct-Variogram for Elevation

The large sample size for elevation permitted the calculation of valid directional sample variograms. Sample variograms were calculated for directions, from 0° (east) to 180° (west) in steps of 30°. An angle class of 45° and a distance class of 10,000 ft were

Table I.2 Descriptive statistics for AAP, TAAP, and elevation.

Variable	n	min	$\bar{x}$	max	s^2	standardized skewness
AAP (mm)	42	69	175	354	5,278	2.19
TAAP	42	4239	5087	5870	163,151	0.34
Elevation(ft)	1595	-250	3932	8400	2,466,980	-3.89

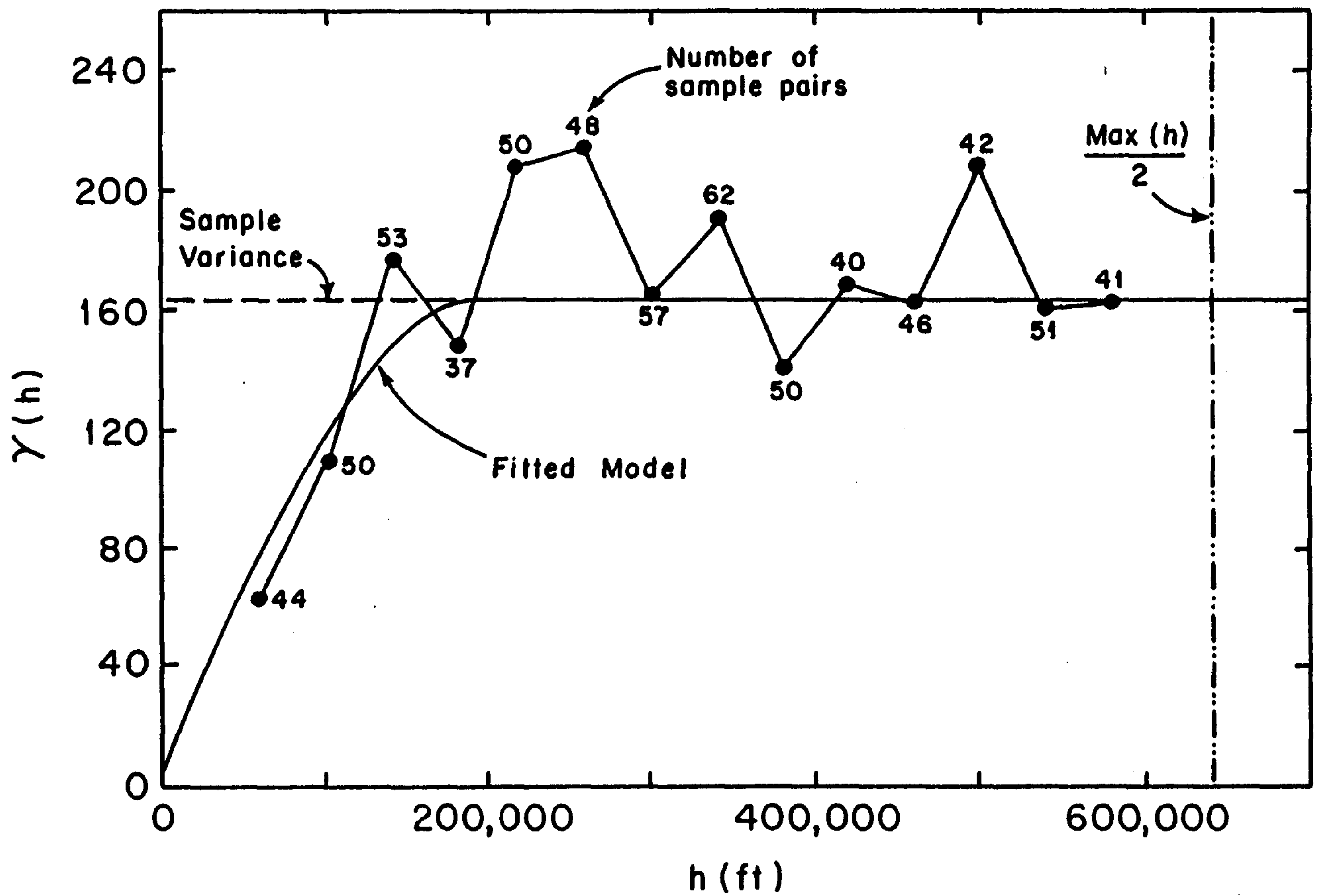

Figure I.5 Sample direct-variogram and fitted model for TAAP.

used. Two sample variograms were selected to represent the observed spatial variability: the ENE - WSW sample variogram (direction 30°) corresponding to the direction of maximum variability, and the NNW - SSE variogram (direction 120°) corresponding to the direction of minimum variability. The maximum distance between sample pairs in the ENE - WSW direction was smaller than for the NNW - SSE direction (Figure 2). The number of sample pairs, however, was very large; many points were calculated with more than 10,000 sample pairs. However, only 21 sample pairs were used to calculate $\gamma(h)$ for the minimum value of h (3,000 ft).

A comparison of the directional variograms indicated zonal anisotropism (Journel and Huijbregts, 1978), with the ENE-WSW variogram increasing to more than 4 times the NNW-SSE variogram at distances of 300,000 ft (Figure 6). Maximum spatial variability occurs in directions transverse to Basin and Range structure, as defined by Death Valley, the Grapevine Range, the Funeral and Black Mtns., the Amargosa Desert, and the northern Spring Mtns. (Figure 2). Minimum spatial variability occurs in directions parallel to the alignment of Basin and Range physiography. Drift is apparent in both sample variograms for distances greater than 400,000 ft (Figure 6). Analysis of the physiography showed the drift to be caused by a gradual increase in the average elevation of basins from south to north in the southern Nevada region. This trend defines the transition from the Mohave to the Great Basin Deserts. Drift is also apparent in the omnidirectional variogram for distances between 300,000 and 600,000 ft. The apparent decrease in the omnidirectional variogram between 320,000 and 500,000 ft is caused by the differences in areal extents of the data for AAP and elevation.

For distances less than 200,000 ft, the sample variograms were very similar and the assumption of isotropism was considered valid (Figure 7). Inspection of topographic maps confirms that spatial structure is isotropic on a local scale for about 70% of the elevation sample (Figure 2), greatly dampening the effect of Basin and Range anisotropism on the sample variograms. Inspection of elevation contours within the UARW indicated the assumption of isotropism to be justifiable for distances less than 200,000 ft because the areas of maximum local anisotropism, such as the western slopes of the Grapevine Range and the Funeral, Black and Spring Mtns. are not included (Figure 2).

There are several reasons for maintaining the assumption of stationarity for distances less than 200,000 ft. The first reason is to simplify model fitting. Model simplification becomes important during cross-validation (see below), where it can become difficult to identify the large number of parameters for more complicated models. It is also important during estimation where the use of universal kriging/cokriging techniques assumes that the form of drift is known (Journel and Huijbregts, 1978; Clark, 1979;

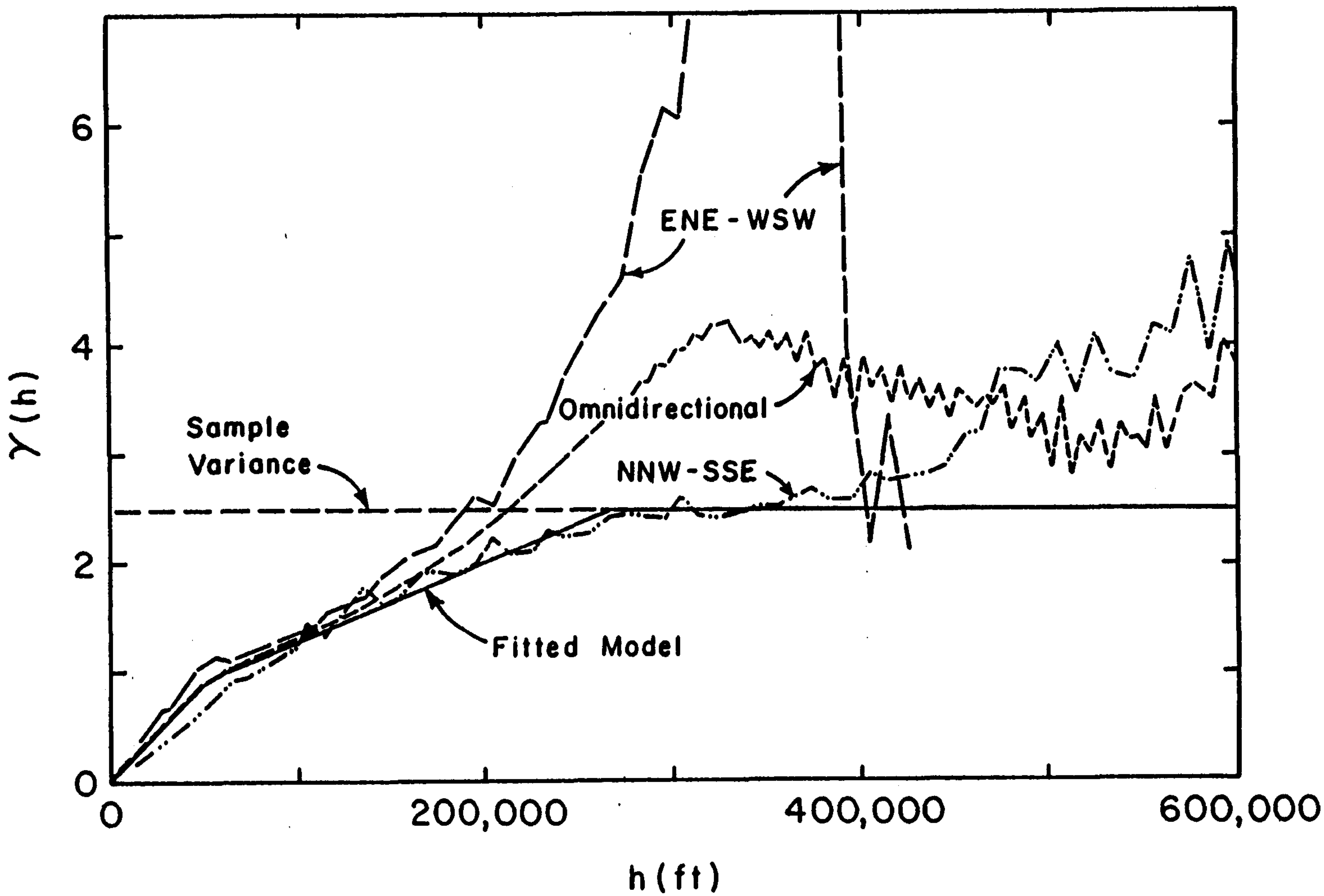

Figure I.6 Sample direct-variogram and fitted model for elevation.

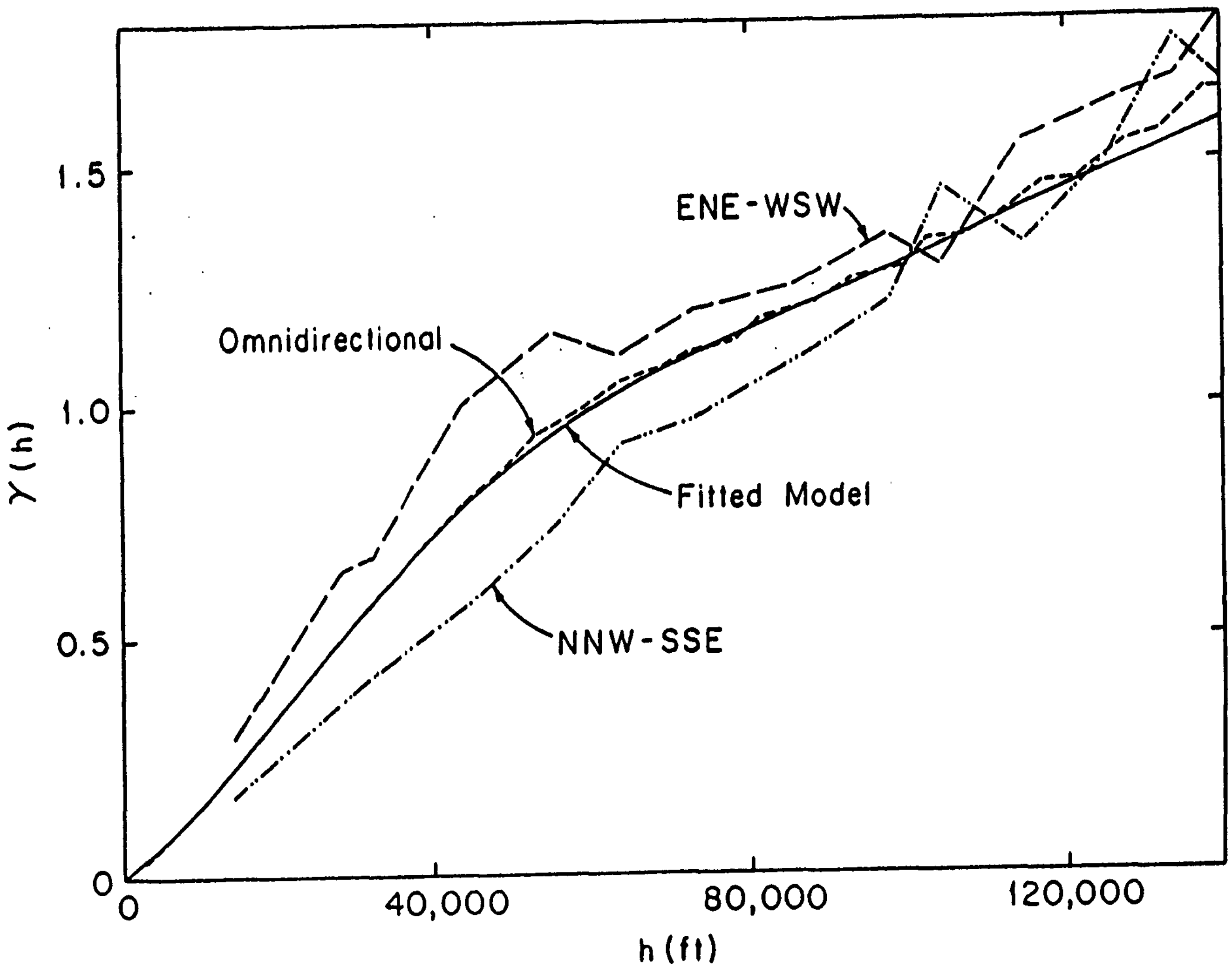

Figure I.7 Sample direct-variogram and fitted model for elevation (expanded scale).

David, 1977). An additional reason is to maintain consistency between the direct-variogram models for TAAP and elevation; if two ReVs are spatially cross-correlated, the presence of an anisotropic structure in the direct-variogram for one ReV implies the existence of an anisotropic structure in the direct-variogram for the second ReV. Because the small sample size for TAAP did not permit the calculation of informative directional variograms, the use of an isotropic variogram model for elevation for distances less than 200,000 ft, was considered preferable to proposing an unobservable anisotropic structure for TAAP.

Sample Cross-Variogram for Elevation and TAAP

The sample cross-variogram indicated a range of spatial cross-correlation between 150,000 and 220,000 ft (Figure 8). For distances between 450,000 and 640,000 ft the cross-variogram fluctuates about the sample variance, suggesting stationarity. However, a negative trend was observed for distances between 220,000 and 450,000 ft. This may be due to the majority of stations being aligned in a northwest to southeast direction (Figure 1). This direction transgresses both the trend of increasing elevation from south to north, which is known to exist, and the trend of increasing AAP from west to east, hypothesized by French (1983). If more stations were available to the southwest and the northeast of the NTS, the negative trend might not be observed in the isotropic sample cross-variogram. A more detailed structural analysis of the TAAP - Elevation spatial cross-correlation was not possible with the available data. Anisotropism probably exists in the true spatial structure of TAAP because of the observed anisotropism in the elevation direct-variograms, and the observed correlation between elevation and TAAP.

Cross-Validation: TAAP Direct-Variogram Model

A large number of alternate models were investigated for the direct-variogram for TAAP and many of these models had favorable cross-validation statistics (Table 3). The selected model consisted of a nugget (5,000) and a spherical structure with a range of 190,000 ft (36 mi.) (Table 4, Figure 5). The existence of a nugget is physically reasonable for AAP at this scale because of the short length of record of the precipitation data used to compute AAP at each station and because of precipitation measurement errors. If additional closely spaced stations were available or if longer lengths of record were available or if longer lengths of record were available sample variograms for AAP would be expected to have very small nugget, due to the expected continuous spatial variation in AAP.

Particular attention was paid to the selection of a search radius (for use in the method of sliding neighborhoods) because of the small sample size and large areal extent of

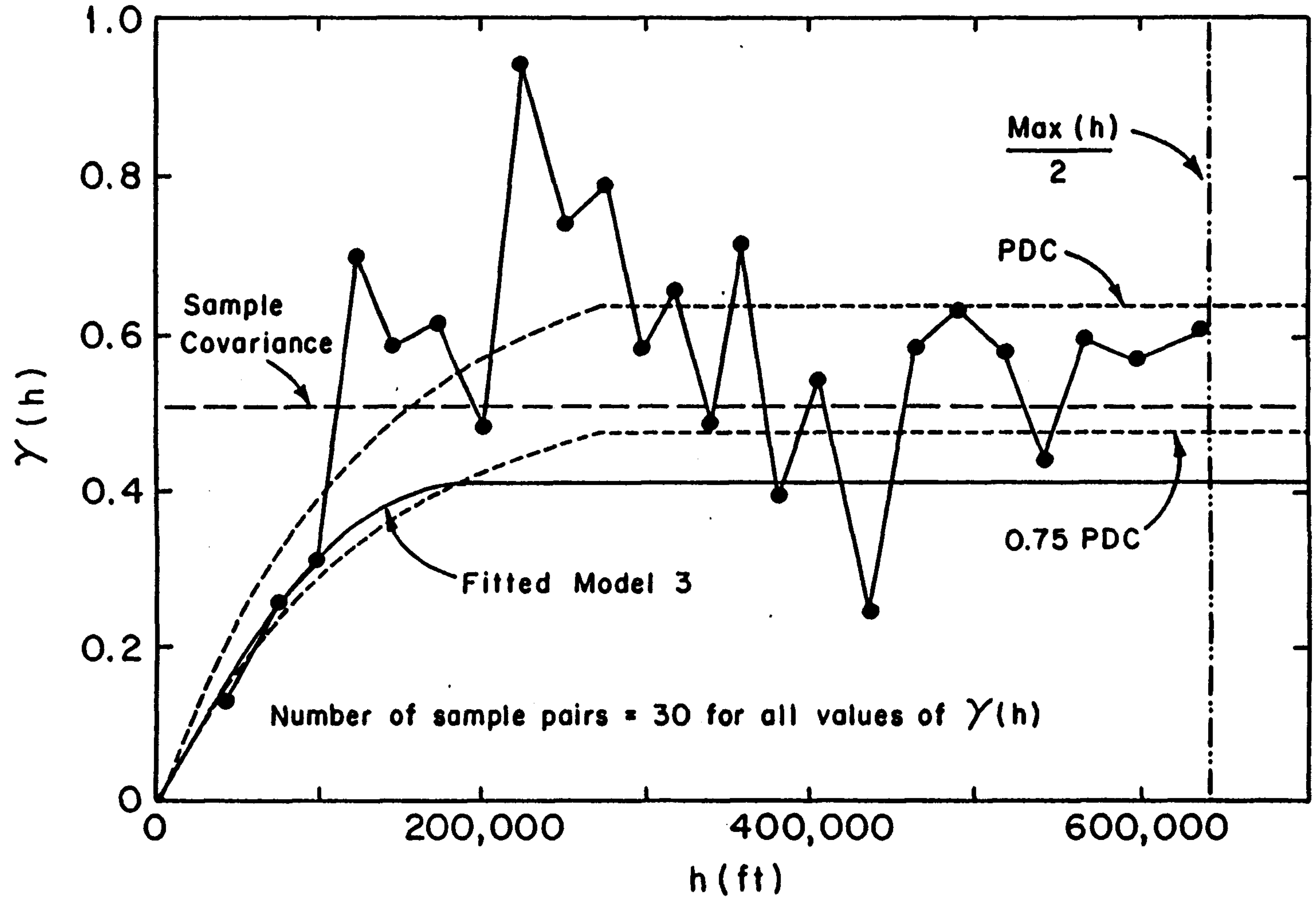

Figure I.8 Sample cross-variogram and fitted model for TAAP and elevation.

Table I.3 Combinations of structures and parameters, and range of resulting cross-validation statistics for alternate direct-variogram models for TAAP.

TAAP Direct-Variogram Model

Structures+	C	a (ft)	PAEE min	PAEE max	RMSE min	RMSE max	SMSE min	SMSE max
Nugget	0 - 20,000	0	0.29	0.80	0.57	1.08	0.743	1.736
Linear	0.815 - 1.63	100,000 - 200,000						
Nugget	0 - 30,000	0	0.28	0.83	0.60	0.99	0.680	1.739
Spherical	133,151 - 163,151	100,000 - 300,000						
Nugget	0 - 30,000	0	0.24	0.78	0.58	0.96	0.911	3.824
Gaussian	133,151 - 163,151	120,000 - 210,000						
Nugget	0 - 30,000	0	0.32	0.80	0.63	0.89	0.869	1.452
Gaussian	65,000 - 120,000	155,000 - 190,000						
Spherical	33,151 - 93,151	155,000 - 200,000						

+Structures — Equation

Nugget: $\gamma(h) = C$

Spherical: $\gamma(h) = C\left[\frac{1.5h}{a} - \frac{0.5h^3}{a^3}\right], h < a$

$= C, h \geq a$

Gaussian: $\gamma(h) = C\left[1 - \exp\left(-\frac{h^2}{b^2}\right)\right], a = b\sqrt{3}$

= practical range (Journel and Huijbregts, 1978)

Linear: $\gamma(h) = Ch, h < a$

$= C, h \geq a$

Table I.4 Structures, parameters, and cross-validation statistics for fitted variogram models.

Variogram	PAEE	RMSE	SMSE	Structures	C	a (ft)
TAAP Direct-variogram	0.45	0.71	0.99	Nugget	5,000	
				Spherical	158,151	190,000
				Search Radius = 480,000 ft		
Elevation Direct-variogram	0.02	0.04	1.01	Nugget	5,500	
				Gaussian	415,000	61,000
				Spherical	185,000	70,000
				Linear	7	270,000
				Search Radius = 120,000 ft		
Elevation TAAP Cross-variogram						
Model 1	-0.58	0.363	1.61	Gaussian	100,000	87,000
				Spherical	309,446	190,000
				Search Radius = 120,000 ft		
Model 2	0.17	0.326	1.00	Nugget	3,880	
				Gaussian	55,000	67,000
				Spherical	350,000	200,000
				Search Radius = 120,000 ft		
Model 3	0.06	0.321	0.99	Gaussian	55,000	61,000
				Spherical	355,000	190,000
				Search Radius = 120,000 ft		
Model 4	-0.21	0.337	1.12	Nugget	3,880	
				Gaussian	50,000	87,000
				Spherical	361,960	190,000
				Search Radius = 120,000 ft		
Model 5	-0.52	0.370	1.45	Gaussian	55,000	87,000
				Spherical	366,960	190,000
				Search Radius = 120,000 ft		

the sampled domain for TAAP (Figure 1). RMSE and SMSE results were not very sensitive to changes in the neighborhood search radius for radii between 300,000 and 600,000 ft. For radii less than 300,000 ft, the results were more sensitive to changes in the search radius than to changes in the model parameters. This occurred because not all of the 42 TAAP values could be estimated for radii less than 300,000 ft due to an insufficient number of nearby stations (Figure 1). For example, for a search radius of 100,000 ft, only 30 values of TAAP could be estimated.

Based on plots like Figure 9, a search radius of 480,000 ft was selected for TAAP.

Cross-Validation: Elevation Direct-Variogram

Although a large number of alternate models were investigated for the direct-variogram for elevation the cross-validation statistics for these models were not sensitive to the choice of structures or parameters. To obtain good cross-validation statistics, a closely fitting model was necessary only for distances less than 140,000 ft for which the assumption of isotropism and stationarity were considered valid.

The three cross-validation statistics for the fixed model were not very sensitive to changes in the search radius. This is due to the fact that almost all of the elevation data are located on a regular grid, and the values closest to the estimated point receive the most weight in the kriging system.

Cross-validation: TAAP-Elevation Cross-variogram

Several values of the sample cross-variogram lie above the PDC curve, suggesting the possibility that models fitted to the sample cross-variogram may fail the positive definite condition (Figure 8). The inconsistency between the areal extent of the two sample domains, combined with the known non-stationarity of elevation for distances greater than 200,000 ft made it difficult to fit models to the sample cross-variogram. In an effort to formulate a valid model, several assumptions were made. First, stationarity of the cross-variogram was assumed for distances up to 140,000 ft. Second, it was assumed that the maximum range of the cross-variogram does not exceed the range for the direct-variogram for TAAP. Third, because the smallest distances between precipitation stations occur, on average, between stations within the NTS (Figure 1), and because the NTS is mostly within the elevation grid, the first few values of the sample cross-variogram (distances less than approximately 120,000 ft) were assumed to be more representative of the spatial cross-correlation structure than the sample cross-variogram as a whole. The reasoning used in this assumption is that, on a local scale, the TAAP - elevation correlation is less likely to be obscured by regional effects. Evidence for this is given by the improved

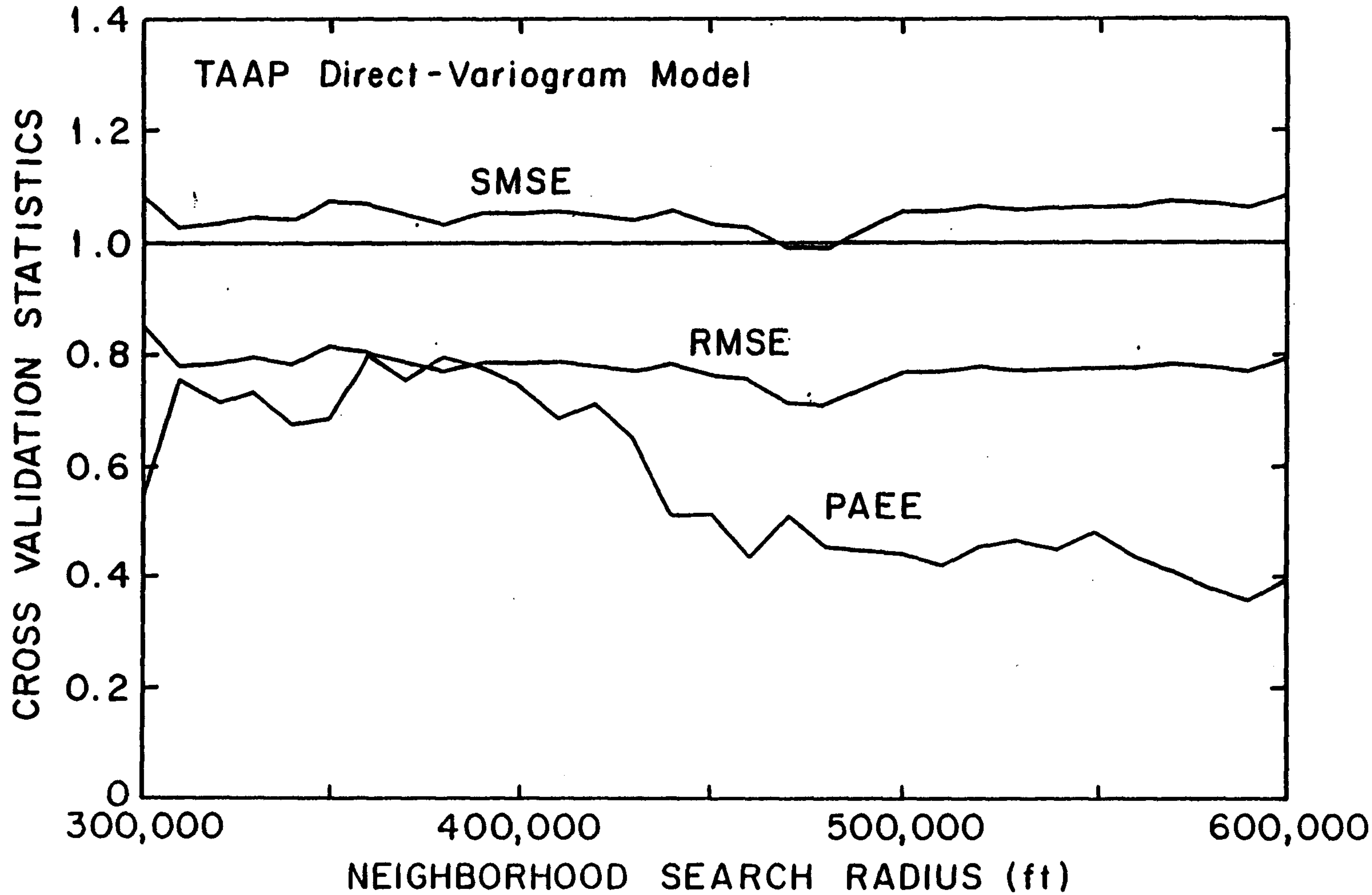

Figure I.9 Effect of neighborhood search radius on cross-validation statistics for fitted, direct-variogram model for TAAP.

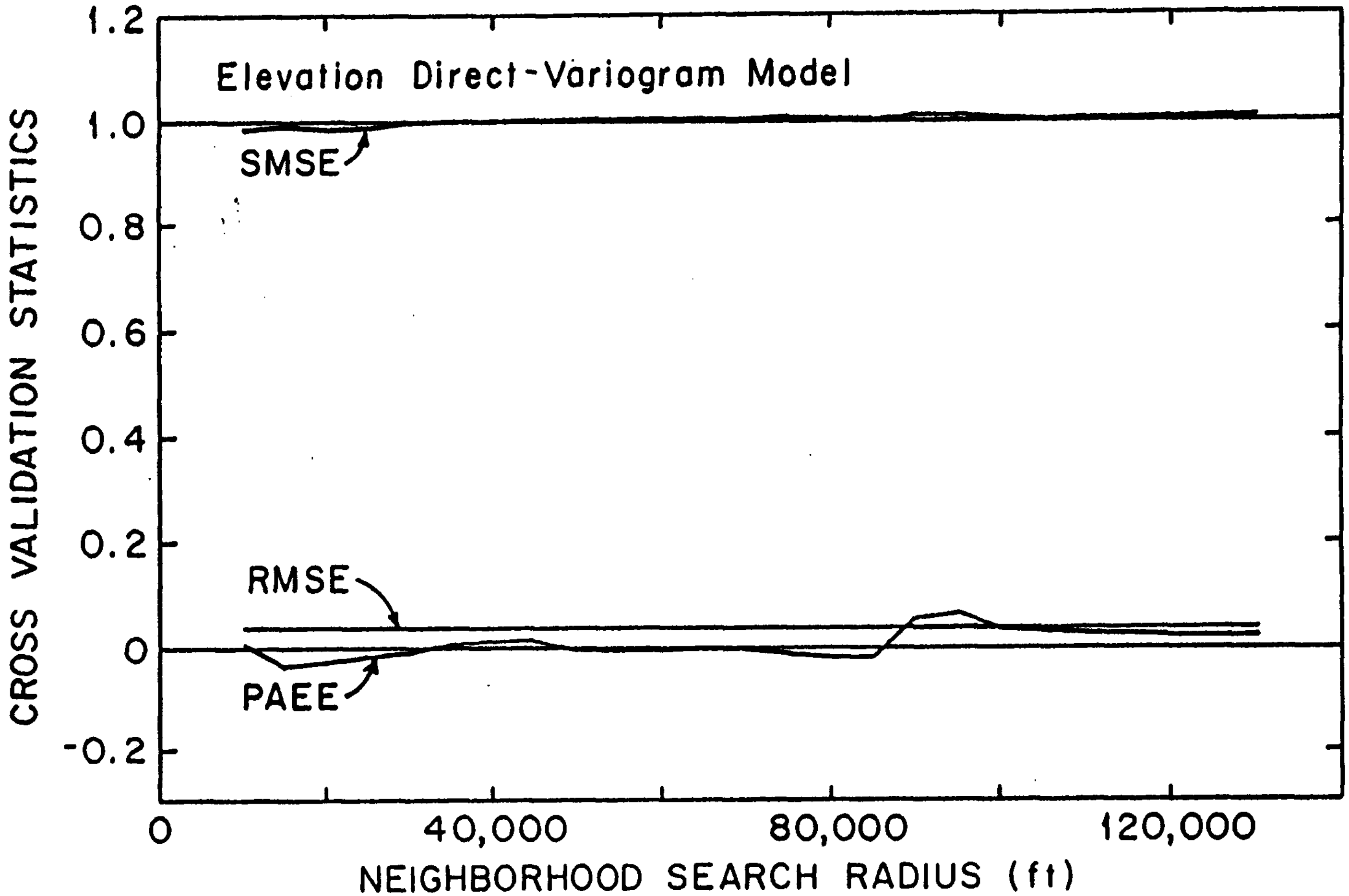

Figure I.10 Effect of neighborhood search radius on cross-validation statistics for fitted, direct-variogram model for elevation.

precipitation - elevation correlation observed by French (1986, 1983) and Quiring (1983) for the NTS stations relative to the correlation observed for all southern Nevada and southeastern California stations.

As an aid to model fitting a 0.75 PDC curve was defined to represent the positive definite condition and the sample correlation coefficient of 0.75:

$$0.75 \text{ PDC} = 0.75 \sqrt{\gamma_{ii}(h)\, \gamma_{jj}(h)} \tag{16}$$

where $\gamma_{ii}(h)$ and $\gamma_{jj}(h)$ are the fitted direct-variogram models for TAAP and elevation. The intersection of the 0.75 PDC curve with the range for the TAAP direct-variogram at a distance of 190,000 ft defines a "psuedo-sill" for the model cross-variogram. The psuedo-sill can be considered a derived covariance that accounts for the inconsistency between the TAAP and elevation sampled domains. By substituting the values of the model direct-variograms into Eq. 16 for h = 190,000 ft, the psuedo-sill could be calculated:

$$\begin{aligned} \text{Psuedo–sill} &= 0.75 \sqrt{\gamma_1(190{,}000)\, \gamma_2(190{,}000)} \\ &= 410{,}000 \text{ ft} \end{aligned} \tag{17}$$

An important assumption implicit in this reasoning is that the spatial variability of TAAP is stationary. If the variance within the UARW is actually less than that observed for the southern Nevada and southeastern California region (i.e., if TAAP is actually non-stationary), the psuedo-sill value defined by Eq. 17 could represent an exaggerated correlation between TAAP and elevation. This assumption is somewhat contradictory to observations concerning the non-stationarity of elevation and the correlation between elevation and precipitation. However, the assumptions were considered necessary to provide a starting point for modeling the cross-variogram. It is important to note that the 0.75 PDC curve provided a fairly good fit to the sample cross-variogram for distances less than 120,000 ft (Figure 8).

Figure 11 shows the cross-validation statistics for alternate models plotted against the elevation neighborhood search radius. Inspection of Figure 11b indicates a general improvement in the RMSE result for cokriging relative to kriging with an increase in elevation search radius. At a search radius of 105,000 ft, the RMSE for several multivariate models is reduced by approximately 55% relative to the RMSE of the TAAP direct-variogram model at a search radius of 480,000 ft. However, the pronounced trend in the RMSE for the multivariate models is not continuous.

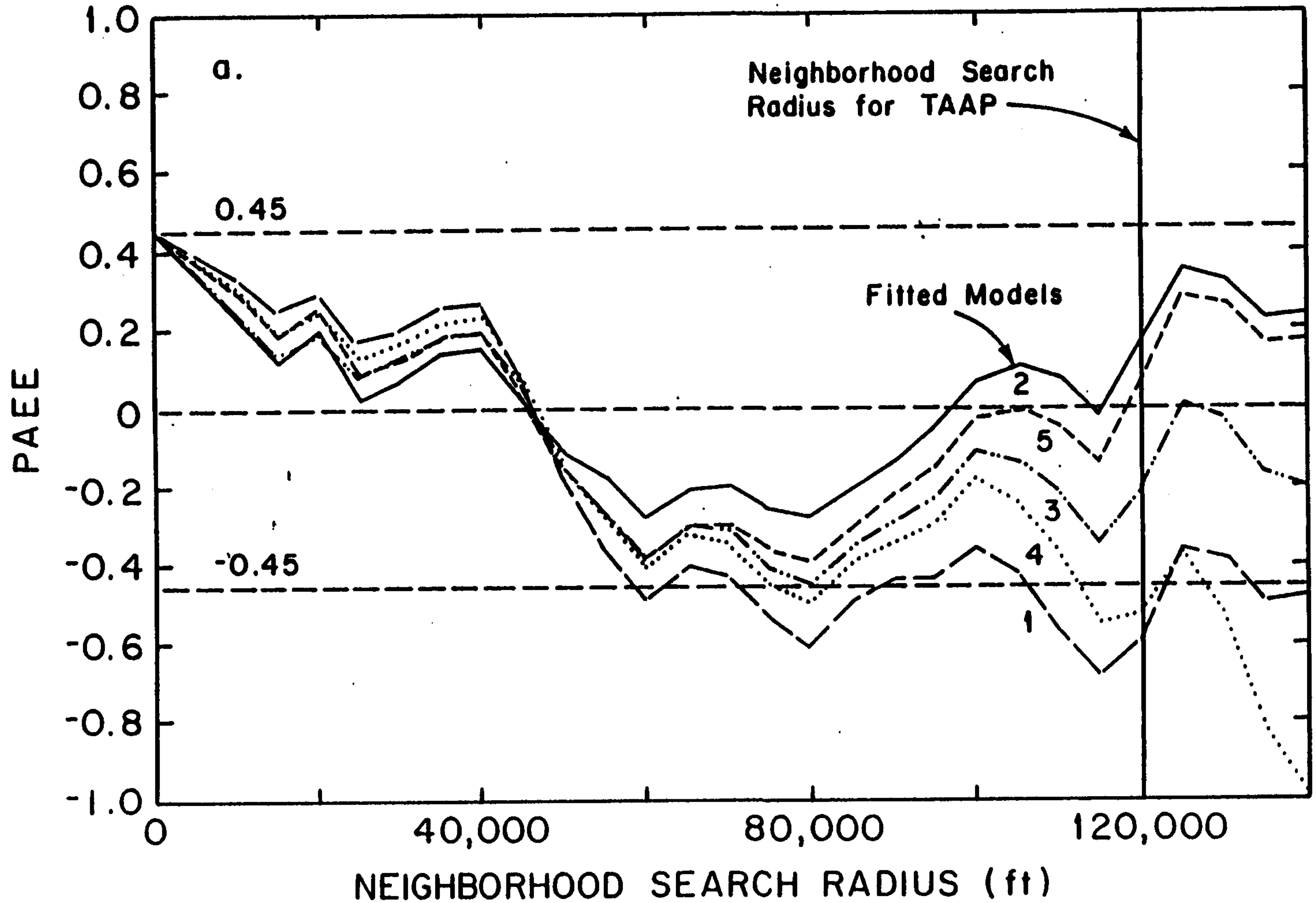

Figure I.11 Effect of neighborhood search radius for elevation on cross-validation statistics for alternate cross-variogram model for TAAP and elevation.

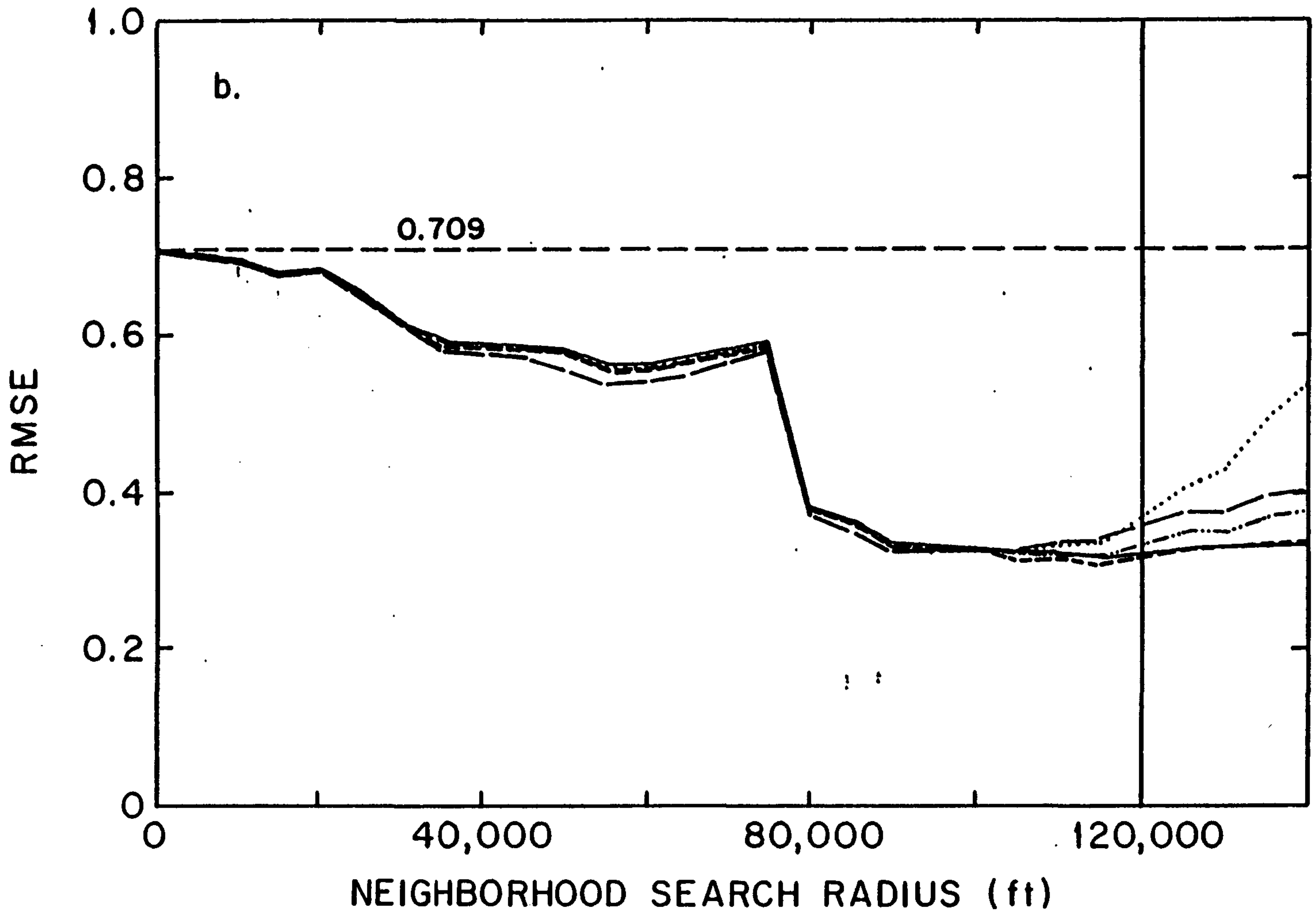

Figure I.11 (Continued).

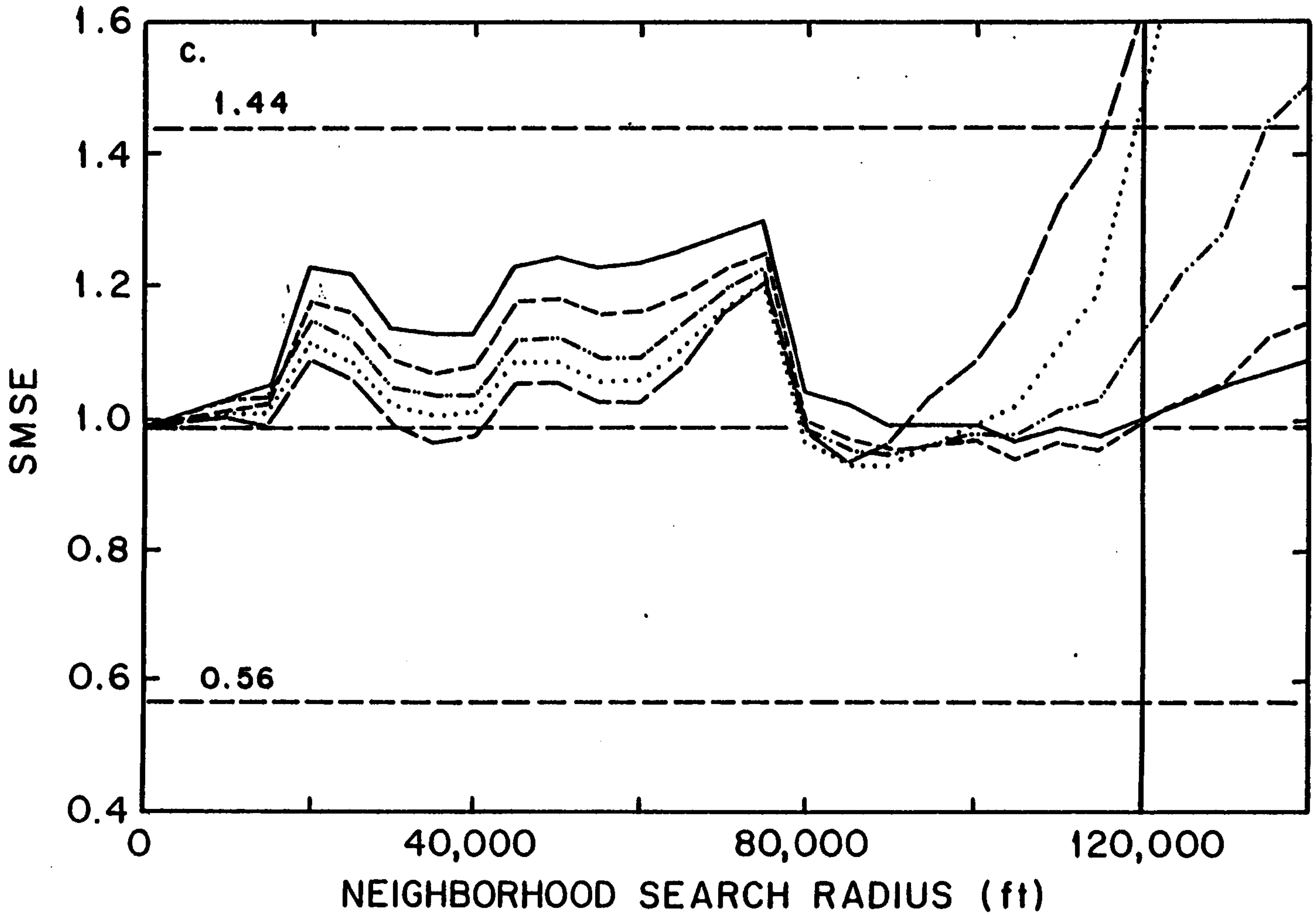

Figure I.11 (Continued).

The sharp decrease in the RMSE beyond 75,000 ft is due to the change from kriging to cokriging in the estimation of a deleted TAAP value as the elevation search radius is increased. Cokriging was not performed if the only elevation data available were the elevation of the deleted precipitation station. This behavior of the RMSE indirectly indicates the sensitivity of the statistic to possible outliers in the TAAP sample.

In general, it was observed that the sensitivity of the cross-validation statistics rapidly increased, with a reduction in model performance and consistency, prior to model failure in terms of the positive definite condition. This is shown by a comparison of Figures 11a, 11b, and 11c, and noting the complete deterioration of SMSE for models 1, 4, and 5 with increasing and divergent RMSE trends. The PAEE result for the model with the most rapidly increasing RMSE and SMSE results (model 5) also becomes divergent and approaches the limits of validity at 140,000 ft. It is interesting to note that model 2, which is a close approximation of the 0.75 PDC curve, displays favorable model performance and consistency throughout the interval of 80,000 to 140,000 ft. Relative stability of the cross-validation statistics was considered in addition to model performance at some specific elevation search radius for the selection of a cross-variogram.

PAEE values for models 2, 3, and 4 decreased throughout the entire range of elevation search radii, and all models indicate an approximate 78% reduction in PAEE for a search radius of 45,000 ft (Figure 11a). Between 15,000 and 75,000 ft, SMSE values for most models were considerably greater than 1.00, though still well within the range of 1.44 to 0.56, which define the upper and lower limits of model validity (Figure 11c). SMSE values showed a similar trend up to 85,000 ft, beyond which SMSE results become large. There is a sharp reduction in SMSE for all models between 75,000 and 80,000 ft, following the same discontinuity observed with the RMSE results. This is again a direct result of the switch from kriging to cokriging in the estimation of one of the TAAP sample values.

For search radii greater than 100,000 ft, SMSE is more sensitive to changes in the parameters of the cross-variogram model than to changes in the elevation search radius. PAEE values follow a trend of increasing model sensitivity with increasing elevation search radius, and changes in model parameters become more important than changes in the search radius beyond approximately 100,000 ft, suggesting that a careful fitting of the cross-variogram is important. Only slight changes in model parameters were likely to cause large changes in validation results, particularly with large elevation search radii (Table 4 and Figure 11). The use of a large elevation search radius, however, was desirable because the large reduction in RMSE indicates large potential reductions in estimation variances, and because of edge effects that occur when small search neighborhoods are used (Renard and Yancey, 1984).

Model 3 (Table 4) was selected because of the relative stability and favorable performance of the combined cross-validation results (Figure 11). The selected elevation search radius of 120,000 ft incorporates up to 449 elevation values into the cokriging system for TAAP. Model 3 was selected over model 2, which had similar cross-validation statistics, because model 3 provided a greater potential reduction in calculated estimation variances. Nested structures were used in model 3 to improve cross-validation results. Although evidence of nested structures is lacking in the sample cross-variogram, the influence of the nested structures in the elevation direct-model on the 0.75 PDC curve is evident.

Cross-Validation: Alternate Estimation Methods

The cross-validation results for the fitted direct- and cross-variogram models are compared with alternate estimation methods in Table 5. The order of performance in terms of RMSE, starting with the least favorable estimation method, was: 1) neighborhood-averaging, 2) inverse-distance weighting, 3) inverse-distance-cubed weighting, 4) inverse-distance-squared weighting, 5) kriging, 6) log-linear regression, 7) linear regression, 8) cokriging. Methods 2, 3, and 4 failed the test for validity based on the PAEE statistic. The performance of kriging is approximately the same as the performance of the inverse-distance-squared method in terms of RMSE.

Log-linear and linear regression provided a 53% improvement in RMSE relative to neighborhood averaging and a 34% reduction in RMSE relative to kriging. Interestingly, log-linear regression does not pass the test for validity in terms of PAEE, even though the value of r (and RMSE) is slightly higher than for linear regression. In general, the performance of linear regression was better than interpolation methods. This suggests the correlation of AAP with elevation to be more important in estimating of AAP than the spatial correlation of available AAP measurements.

Cokriging with an elevation search radius of 120,000 ft provided a 68% improvement in RMSE relative to neighborhood averaging, and a 55% improvement in RMSE relative to kriging. The values of PAEE for cokriging and linear regression were similar but the value of RMSE for cokriging was improved by 31% relative to linear regression. SMSE for cokriging and kriging were nearly identical.

Table I.5. Cross-validation results for alternate estimation methods.

Method	Search Radius (ft)	PAEE	RMSE	SMSE
Neighborhood Average	400,000	2.32	0.90	-
	480,000	0.62	0.93	-
	560,000	0.31	0.97	-
Inverse Distance	400,000	4.04	0.81	-
	480,000	3.05	0.81	-
	560,000	2.64	0.85	-
Inverse Distance Squared	400,000	6.29	0.76	-
	480,000	5.72	0.74	-
	560,000	5.36	0.75	-
Inverse Distance Cubed	400,000	6.51	0.83	-
	480,000	6.17	0.81	-
	560,000	5.96	0.81	-
Kriging	400,000	0.75	0.78	1.05
	480,000	0.45	0.71	0.99
	560,000	0.44	0.78	1.07
Log-linear Regression	-	-3.93	0.46	-
Linear Regression	-	-0.04	0.47	-
Cokriging	55,000+	-0.26	0.56	1.16
	80,000	-0.39	0.38	1.00
	120,000	0.05	0.32	0.99

+ Search radius for elevation. Search radius for TAAP = 480,000 ft.

Summary and Conclusions

Measured values of average annual precipitation (AAP) for 42 precipitation stations in the southern Nevada - southeastern California region were found to approximate a log-normal distribution, and a transformed variable, TAAP = ln(AAP) * 1000, was defined for the purpose of estimating AAP, using multivariate geostatistics, for a proposed high-level nuclear waste repository site on Yucca Mtn. The use of stations with as few as 8 years of record, was required in order to provide a sample size large enough for an adequate geostatistical analysis.

An isotropic spherical model variogram with a relatively small nugget (5000), a sill equal to the sample variance (163,151) and a range of 190,000 ft (36 mi) was selected as the best representation of the spatial structure for TAAP, given the data currently available. The assumption of isotropism was necessary because the small sample size did not allow for the calculation of informative directional sample variograms. Model selection was based on a visual fit of the model to the sample variogram, cross-validation statistics, and knowledge of regional climatology.

Computed directional sample variograms for ground elevation indicated a pronounced anisotropic spatial structure for elevation within the southern Nevada region. The axis of maximum spatial variability is orientated in the ENE - WSW direction, transverse to the dominant alignment of Basin and Range structure. Modeling the spatial structure for elevation under the assumptions of isotropism and stationarity was considered desirable given the isotropic, spherical model used for TAAP. This was possible using the assumption of quasi-stationarity and limiting the applicability of the model elevation direct-variogram to distances less than 140,000 ft (26 mi) for which the directional variograms coincided fairly well with the omnidirectional sample variogram, and for which drift was not observed.

Two nested transition structures were observed in the isotropic sample variogram for elevation: a large-scale structure representing the spatial variability between basins and ranges, and a small-scale structure representing the variability within individual basins and ranges. The larger structure was modeled by a linear model, while the smaller structure was modeled using a combination of spherical and gaussian models. A nugget was used to improve model performance in terms of the cross-validation results, and to account for known errors in the elevation data.

A correlation coefficient of 0.75 was calculated between TAAP and elevation. Sample cross-variograms indicated positive spatial cross-correlation between TAAP and elevation with a range of approximately 190,000 ft. Evidence for a non-stationary cross-correlation structure was observed as a drift in the sample cross-variogram between distances of 200,000 and 450,000 ft (40 and 85 mi). The anomalous structure was concluded to be a caused by the arrangement of precipitation stations in a northwest to southeast alignment, transgressing both the observed trend of increasing elevation from south to north and (possibly) the hypothesized trend of increasing AAP from west to east. The elevation and AAP trends cause values of TAAP to be lower for a given elevation northwest of the UARW in the "deficit" zone of French (1983), relative to AAP values for the same elevation southeast of the UARW, in the "excess zone". The clustering of stations within the NTS, which are within or directly adjacent to the UARW boundary, are within the "transition zone". Distances between 200,000 and 450,000 ft represent the approximate range of distances between the grouping of stations in each zone. The addition of AAP values to the southwest and northeast of the NTS cluster would probably eliminate the negative trend observed in the sample cross-variogram. The true spatial structure of AAP is likely to be anisotropic, with significant drift components, because TAAP shows significant correlation with elevation, and the sample variograms for elevation show strong anisotropism with at least two drift components. However, it is impossible to identify anisotropism without the aid of directional sample variograms, and these could not be computed given the small sample size for TAAP.

Cokriging provided the best cross-validation statistics relative to regression methods and other interpolation methods, including kriging. The RMSE result of the selected multivariate model, using an elevation search radius of 120,000 ft., was reduced from 0.71 for kriging 0.32 for cokriging, a 55% improvement. The PAEE result was reduced from 0.45 to 0.05 using cokriging, an 89% improvement. The SMSE result changed only slightly from kriging to cokriging, indicating similar model performance in terms of the consistency of calculated estimation variances. The RMSE result for the selected multivariate model indicated approximately a 31% improvement relative to regression. It was concluded that the utilization of the spatial correlation of TAAP and elevation, in addition to the sample correlation for the two variables, was important.

References

Clark, I., 1979: Practical Geostatistics. Applied Science Publishers, London, U.K.

Cooper, R. M., J. D. Istok, and A. L. Flint, 1988: Three-dimensional, cross-semivariogram calculations for hydrogeological data. Ground Water, 26, 638-646.

David, M., 1977: Geostatistical Ore Reserve Estimation. Elsevier Scientific Publishing Company, New York, N.Y.

Delhomme, J. P., 1979: Spatial Variability and Uncertainty in Groundwater Flow Parameters: A Geostatistical Approach. Water Resources Research, 15, 269-280.

Dingman, S. L., D. M. Seely-Reynolds, and R. C. Reynolds III, 1988: Application of Kriging to Estimating Mean Annual Precipitation in a Region of Orographic Influence. Water Resources Bulletin, 24(2), 329-339.

French, R. H., 1987: Effects of the Length of Record on Estimates of Annual and Seasonal Precipitation at the Nevada Test Site. U.S. Dept. of Energy, Nevada Operations Office, #DE-AC08-85NV10384.

French, R. H., 1986: Daily, Seasonal, and Annual Precipitation at the Nevada Test Site, Nevada. U.S. Dept. of Energy, Nevada Operations Office, #DE-AC08-85NV10384, Publication No. 45042.

French, R. H., 1983: A Preliminary Analysis of Precipitation in Southern Nevada. DOE/NV/10162-10, Water Resources Center, Desert Research Institute, Las Vegas, Nevada, 39.

Isaaks, E. H. and R. M. Srivastava. 1989: An Introduction to Applied Geostatistics. Oxford University Press, New York.

Jones, K.R., 1981: Arid Zone Hydrology for Agricultural Development, FAO Irrigation and Drainage Paper No. 37, Food and Agriculture Organization of the United Nations, Rome.

Journel, A. G. and C. J. Huijbregts, 1978: Mining Geostatistics. Academic Press, New York, N.Y.

Quiring, R. F., 1983: Precipitation Climatology of the Nevada Test Site. Nuclear Support Office, Las Vegas, NV.

Renard, D., and J. D. Yancey, 1984: Smoothing Discontinuities when Extrapolating Using Moving Neighborhood. Geostatistics for Natural Resources Characterization, Part II, 679-690, D. Reidel Publishing Co.

Tabios, G. Q. III., and J. D. Salas, 1985: A Comparative Analysis of Techniques for Spatial Interpolation of Precipitation. Water Resources Bulletin, 21(3), 365-380.

Yates, S.R., and A.W. Warrick, 1987: Estimating Soil Water Content Using Cokriging. Soil Sci. Soc. of Am. J., 51, 23-30.

PRECIPITATION ESTIMATION IN MOUNTAINOUS TERRAIN USING MULTIVARIATE GEOSTATISTICS: 2. ISOHYETAL MAPS

Joseph A. Hevesi
Fenix & Scisson of Nevada
P.O. Box 327, M.S. 721, Mercury, Nevada 89023

Alan L. Flint
U.S. Geological Survey, Water Resources Division
P.O. Box 327, M.S. 721, Mercury, Nevada 89023

Jonathan D. Istok
Department of Civil Engineering, Oregon State University
Corvallis, Oregon 97331

Introduction

This chapter presents the first application of multivariate geostatistical methods to the problem of estimating average annual precipitation (AAP) at Yucca Mtn., Nevada, site of a proposed, high-level nuclear waste repository. Orographic effects in the rugged mountainous terrain at the site result in a significant positive correlation between AAP and station elevation (Quiring, 1983; French, 1986, Chapter I). French (1986) and Quiring (1983) fit linear and nonlinear regression equations to AAP and station elevation for 63 precipitation stations in southern Nevada and southeastern California. Dingman et al. (1988) used an "elevation scale factor" to remove an observed orographic influence of elevation on AAP in Vermont and New Hampshire. In Chapter I precipitation data from 42 stations in the region were used to compute an r value of 0.75 for the correlation between the transformed variable TAAP = 1000 ln (AAP) and station elevation. Sample direct- and cross-variograms were fit with isotropic, variogram models containing several nested structures with ranges from 55,000 to 355,000 ft.

In this chapter, the model variograms presented in Chapter I are used to prepare isohyetal maps for AAP using estimates obtained from kriging and cokriging. One goal was to determine the effectiveness of cokriging in providing increased detail in isohyets and reducing estimation variances at the site. Additional goals were to evaluate the effectiveness of an existing precipitation network for measuring AAP and selecting optimal locations for additional precipitation stations.

Methodology

Study Area

The proposed high-level nuclear waste repository is located within a 2 mi^2 area beneath the southern extent of Yucca Mtn., Nevada (Figure I.1, Chapter I). The Upper Amargosa River Watershed (UARW), was defined for use in estimating AAP in the vicinity of Yucca Mtn. A description of the physiography and climate of the UARW are in Chapter I.

Precipitation and Elevation Data

Measured values of AAP from a network of 42 precipitation stations in the southern Nevada - southeastern California region with lengths of record from 8 to 53 years were utilized for the purpose of estimating AAP (Table I.1, Chapter I). The nearest station to the repository (station 75, AAP = 118.6 mm) is located approximately 10 miles ESE and the farthest station (station 225, AAP = 114.8 mm) is located approximately 140 miles WNW. Ten of the 42 selected stations are within the boundaries of the UARW. Station elevations were combined with 1531 elevations recorded from topographic maps using a square grid with a spacing of 10,000 ft. A total of 747 grid points were located within the UARW. This density was considered sufficient for preparing isohyetal maps of AAP for the UARW.

Calculating Estimates

Fitted model direct- and cross-variograms for the transformed variable, TAAP = ln(AAP) * 1000 and for elevation are in Chapter I. Estimates and estimation variances for TAAP were obtained using ordinary kriging and cokriging (Journel and Huijbregts, 1978). The method of sliding neighborhoods was used with a nearest-neighbor, Quadrant octant search. Estimates were obtained for a grid of points identical to that used to obtain the elevation sample to maximize the reduction in estimation variances obtained using cokriging. This configuration also permitted direct comparisons between estimates obtained using kriging and cokriging and estimates obtained using regression and inverse-distance weighting. Separate search radii were used for TAAP and elevation; the radii were selected using cross-validation (Chapter I). Isohyetal maps were prepared using the SURFER software package (Golden Graphics, 1989).

To compute estimates for AAP, inverse transformations were made using the following equations (Cooper and Istok, 1988):

$$AAP^*(x_o) = K_o \exp[TAAP^*(x_o)/1000 + \sigma_k^2(x_o)/1{,}000{,}000] \quad (1)$$

$$\sigma_{AAP}^2(x_o) = m^2[\exp(\sigma^2)\ (1 - \exp\ (\sigma_k^2(x_o)1{,}000{,}000)] \quad (2)$$

where

$$K_o = \left(\frac{1}{n} \sum_{i=1}^{n} TAAP^*(x_o)/1000 \right) / m \quad (3)$$

$AAP^*(x_o)$ is the estimate for AAP at the point x_o, $TAAP^*(x_o)$ and σ_k^2 are the estimate and estimation variance for TAAP obtained using kriging or cokriging, m is the sample mean for AAP, σ_{AAP}^2 is the estimation variance for AAP, and σ_k^2 and is the sample variance for TAAP. The parameter K_o is used to correct for bias in the inverse transformation.

Effectiveness of Cokriging

Estimates and estimation variances obtained from kriging and cokriging were compared visually and by plotting the percent reduction defined as: percent reduction = 100% * (kriging variance - cokriging variance) / kriging variance. Estimates and estimation variances were also compared in an effort to investigate the sensitivity of results to alternate variogram models.

Effectiveness of Existing Stations with Short Periods of Record

A network of stations has recently been established to measure precipitation in the vicinity of the repository site. However, the length of record for these stations is too short at present to be useful in estimating AAP. The effectiveness of these stations for estimating AAP in the future as more data are collected was evaluated by computing the reduction in estimation variance that would occur if a value of AAP was available at each station location. This was done by adding the coordinates of these stations to an existing network of 28 active stations with longer lengths of record and computing reductions in estimation variances using kriging and cokriging. Of the 42 stations used for obtaining estimates of AAP, 14 are no longer active and thus were not included in the analysis of the existing network.

Identifying Locations for Additional Precipitation Stations

During the course of this study it was necessary to identify optimal locations for additional precipitation stations for this network. The method of fictitious points (Journel and Huijbregts, 1978) was used to select these locations. In this approach a station location is selected by identifying the grid point which has the maximum estimation variance with the existing network. The coordinates of that point are then added to the available data, and the process repeated to identify the location for the next station.

Results and Discussion

Comparison of Kriging and Cokriging Estimates

The kriging estimates defined a smooth surface that increased gradually from the southwest to northeast (Figure II.1a). Orographic effects are apparent only as a general trend of increasing AAP towards the northeastern edge of the UARW, corresponding to precipitation gauges located at higher elevations on the NTS. Estimates for AAP at Yucca Mtn. ranged from 180 mm in the north to 115 mm in the south with an estimate of 145 mm at the proposed repository site. Within the UARW, the maximum estimate of 287 mm occurred at location (630,000E, 890,000N) at an elevation of 7,500 ft. Estimates exceeding 200 were obtained for about 15% of the UARW. Estimates of 85 to 100 mm occurred throughout a large area containing Death Valley and the Amargosa Desert. A minimum estimate of 86 mm was obtained for location (500,000E, 650,000N) at an elevation of 5,000 ft. in the Funeral Mtns. Estimates of less than 150 mm were obtained for more than 50% of the points within the UARW.

The kriging estimates were positively correlated with grid point elevations. A regression equation was fit to the kriging estimates and elevations (Figure II.2)

$$AAP^* = \exp(4.51 + 0.00012E) \qquad (4)$$
$$r^2 = 0.46$$

where AAP^* is the kriging estimate for AAP at a grid point (mm) and E is elevation of the grid point (ft). Equation 4 indicates a smaller orographic influence (i.e., the regression parameter multiplying E in Eq. 4 is smaller) for AAP than that defined by precipitation stations in the southern Nevada and southeastern California region (Equation 13, Chapter I).

The cokriging estimates defined a more detailed and irregular surface and conformed more closely to the topography relative to the kriging estimates (Figure II.3a). A maximum estimate of 335 mm was obtained at location (600,000E, 890,000N) on Pahute Mesa, close to the northeastern boundary of the UARW. Estimates exceeding 250 mm were obtained for a relatively small area encompassing Pahute Mesa, Timber and Shoshone Mtns., and the northern portion of Yucca and Dome Mtns. Estimates of AAP at the repository site ranged from 155 to 175 mm, depending on elevation, and were approximately 20 mm greater than kriging estimates. Cokriging estimates at Yucca Mtn. ranged from minimum values of 130 mm along the southern bare of the Mtn. to maximum

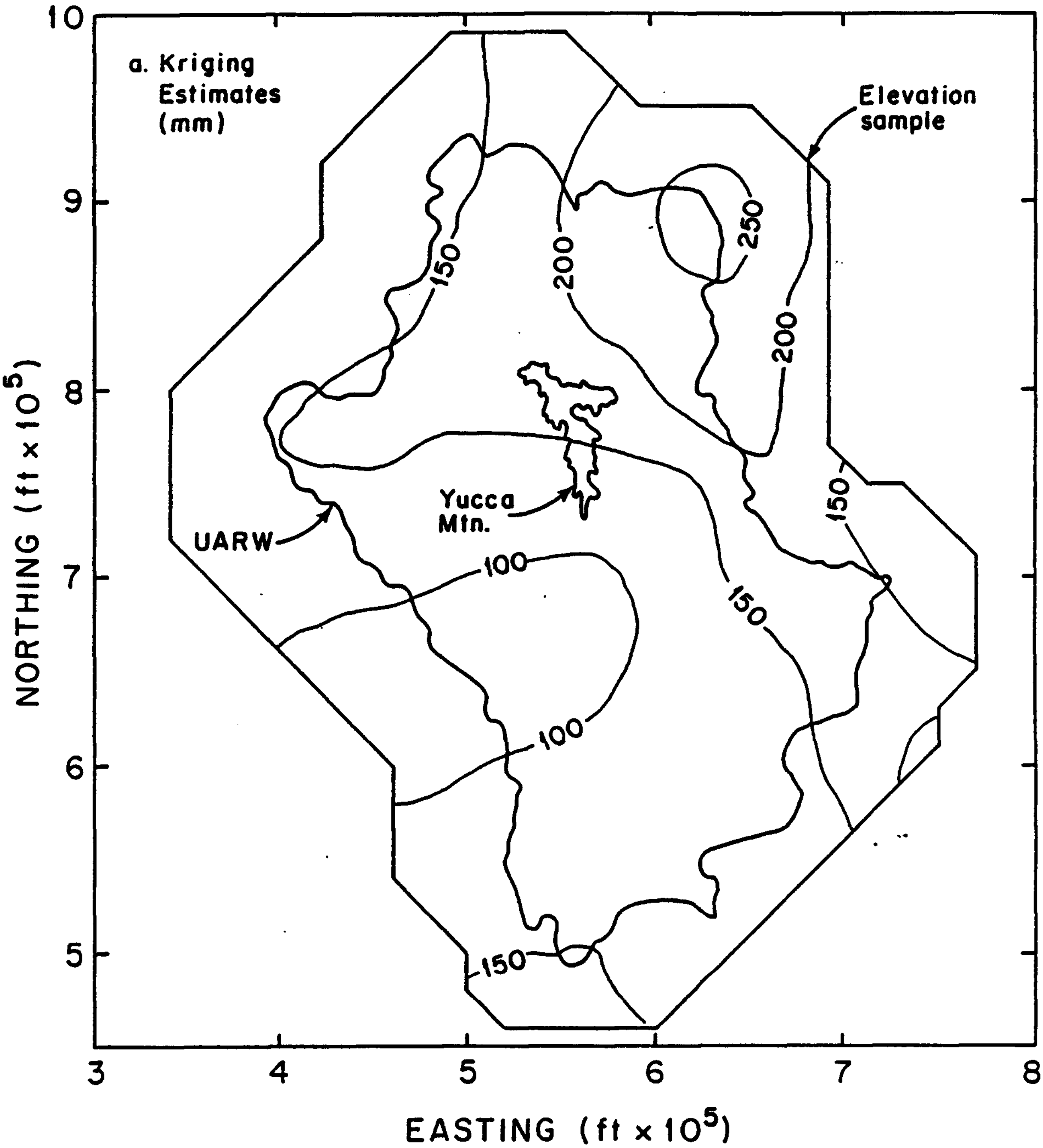

Figure II.1 Kriging estimates and estimation variances for AAP.

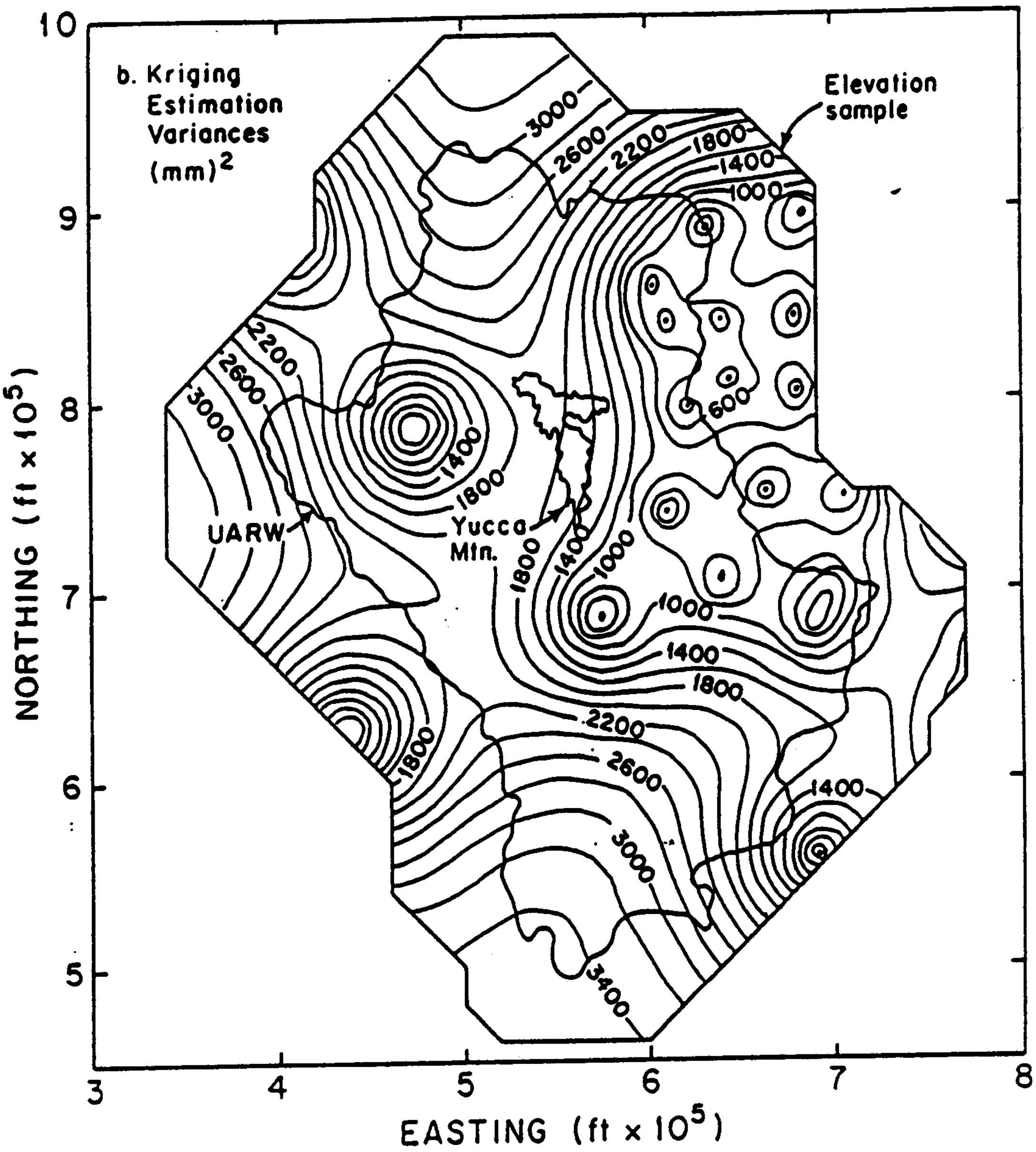

Figure II.1 (Continued).

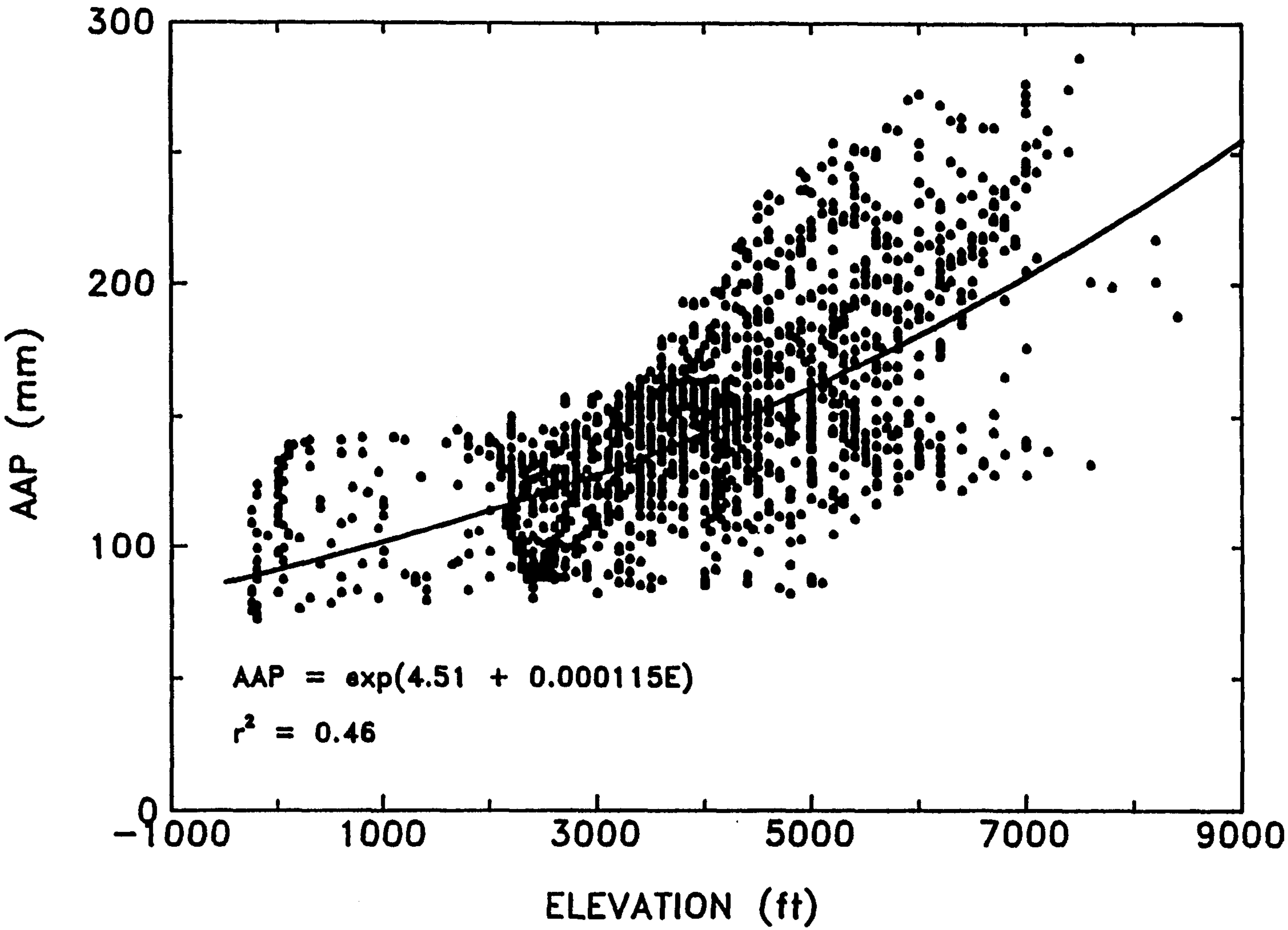

Figure II.2 Kriging estimates for individual grid points.

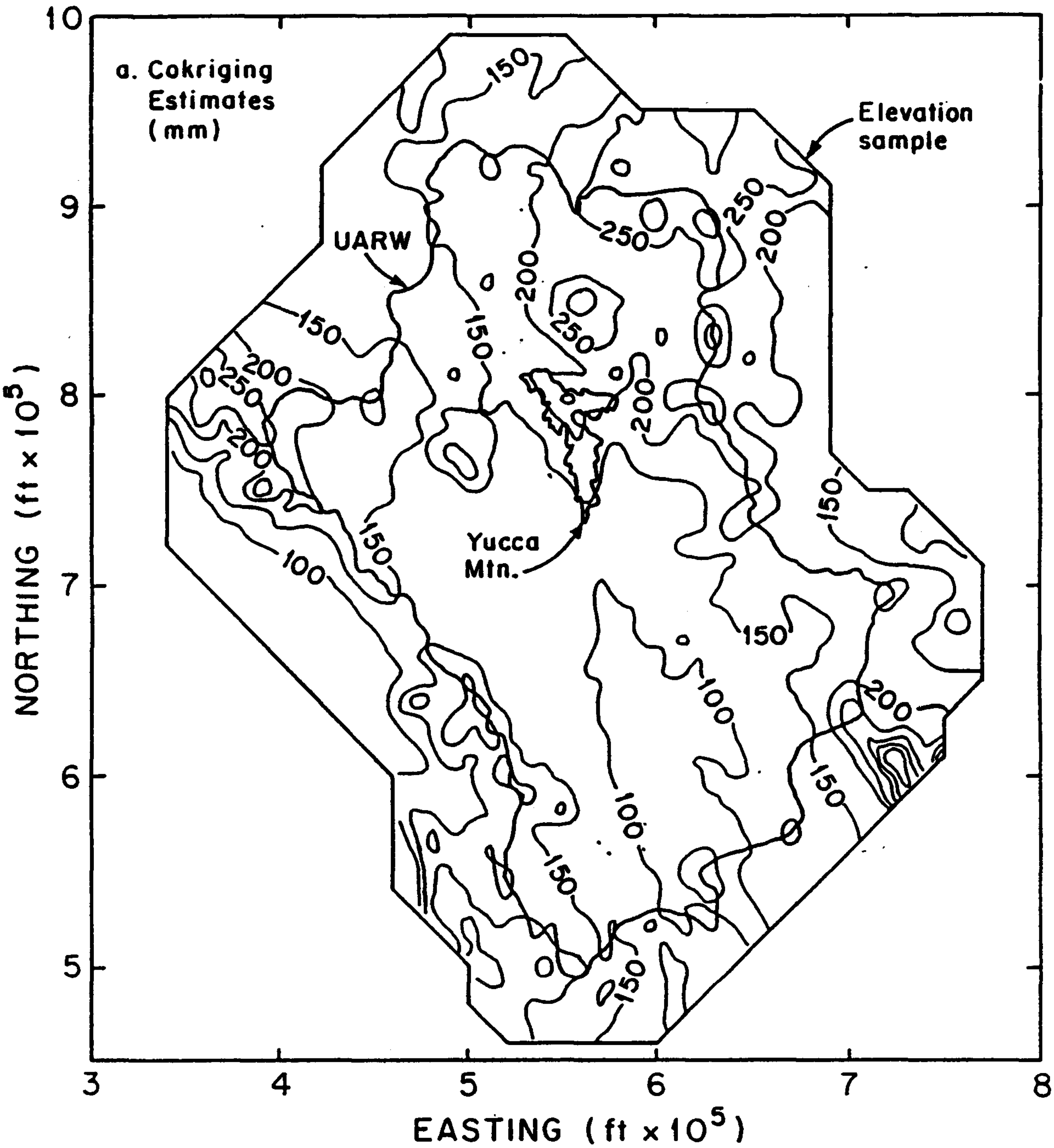

Figure II.3 Cokriging estimates and estimation variances for AAP.

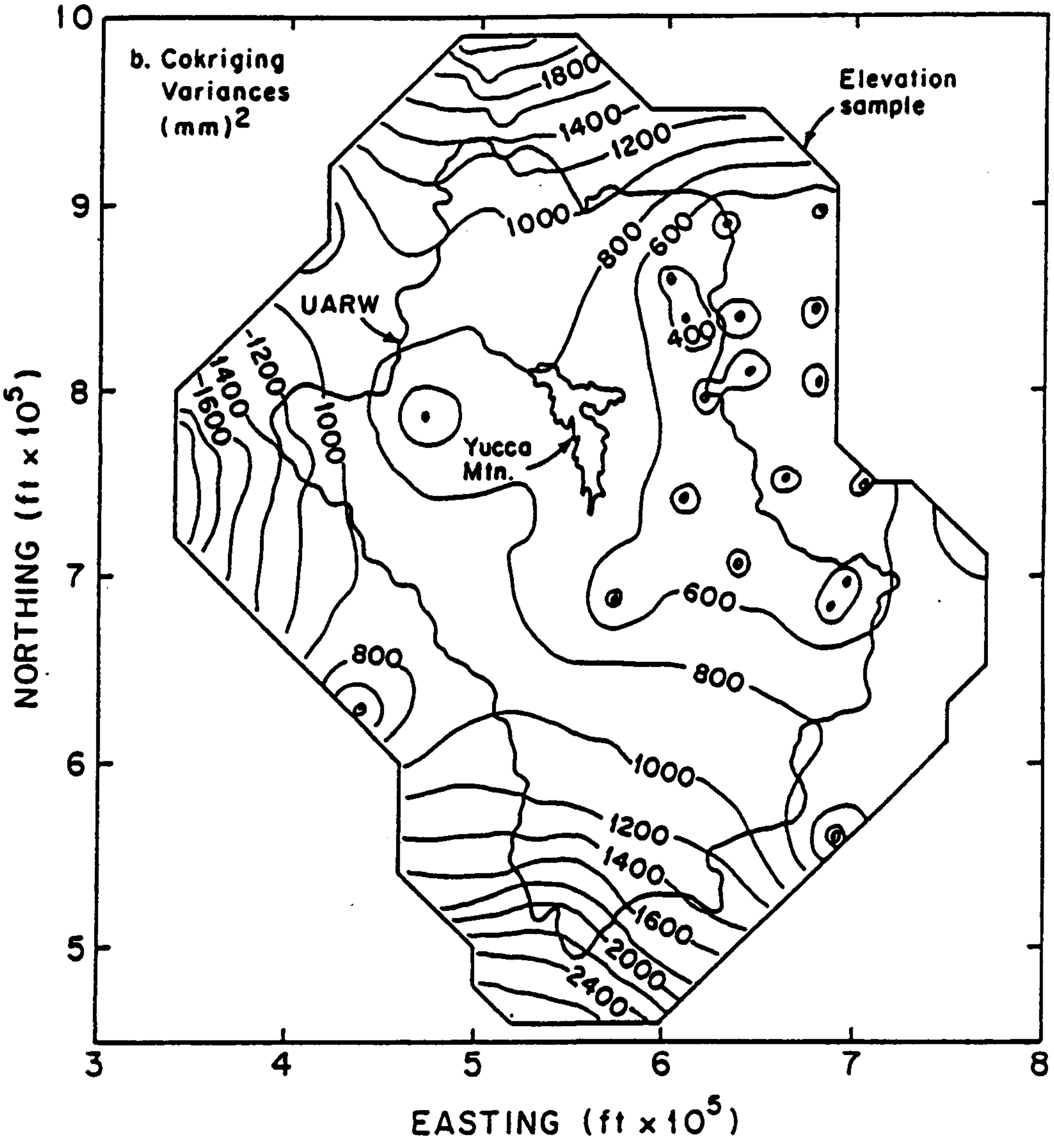

Figure II.3 (Continued).

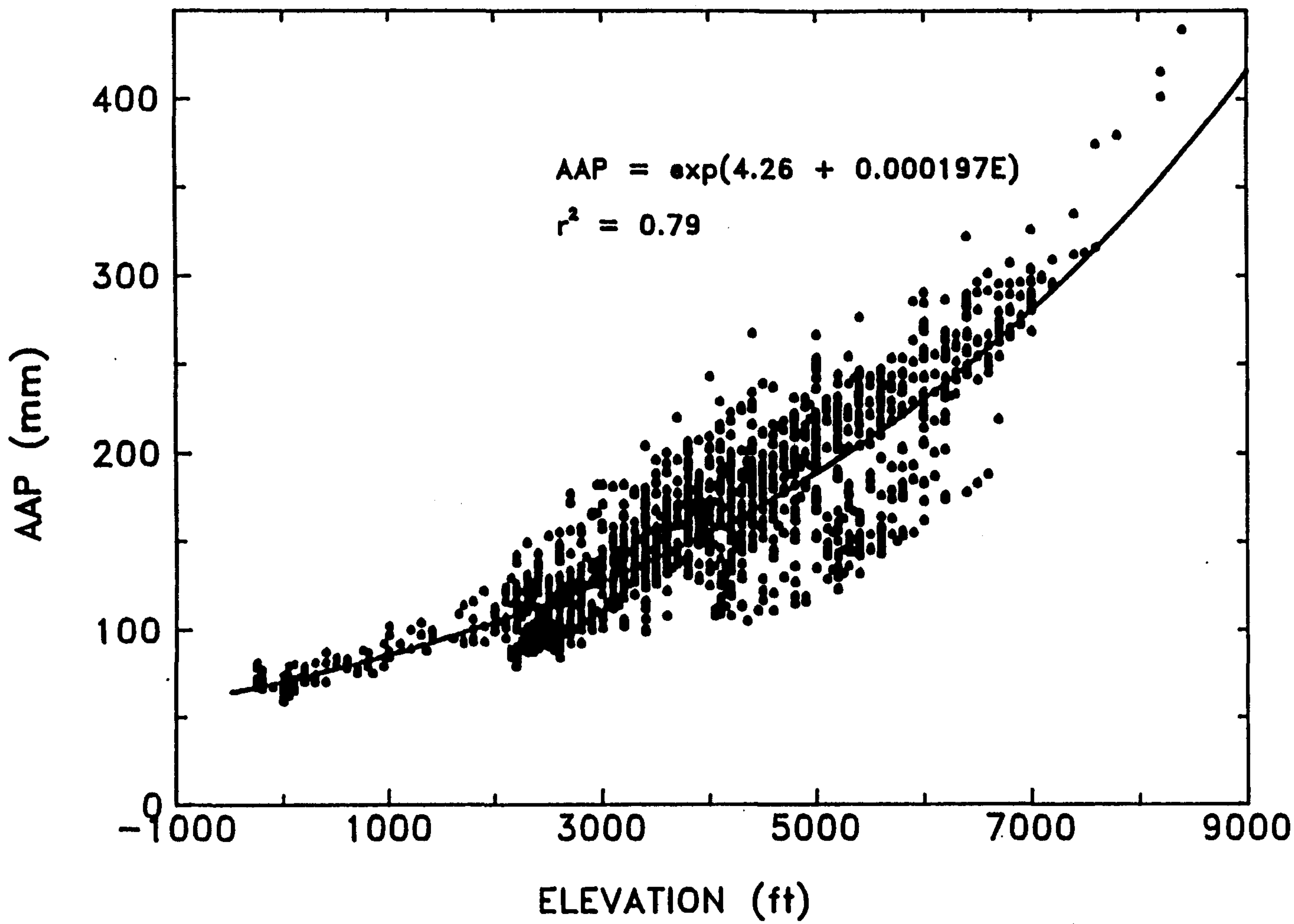

Figure II.4 Cokriging estimates for individual grid points.

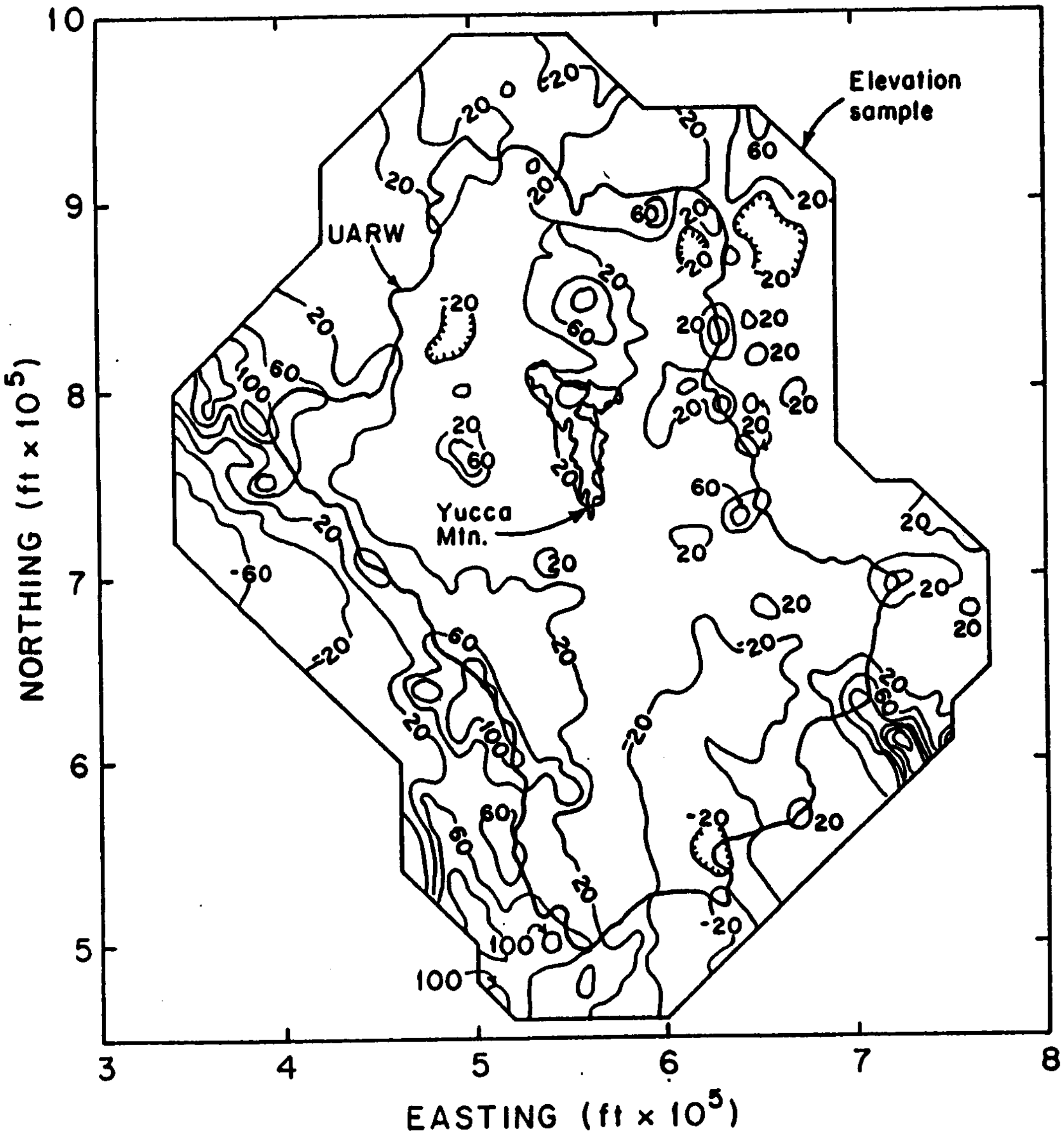

Figure II.5 Difference between kriging and cokriging estimates.

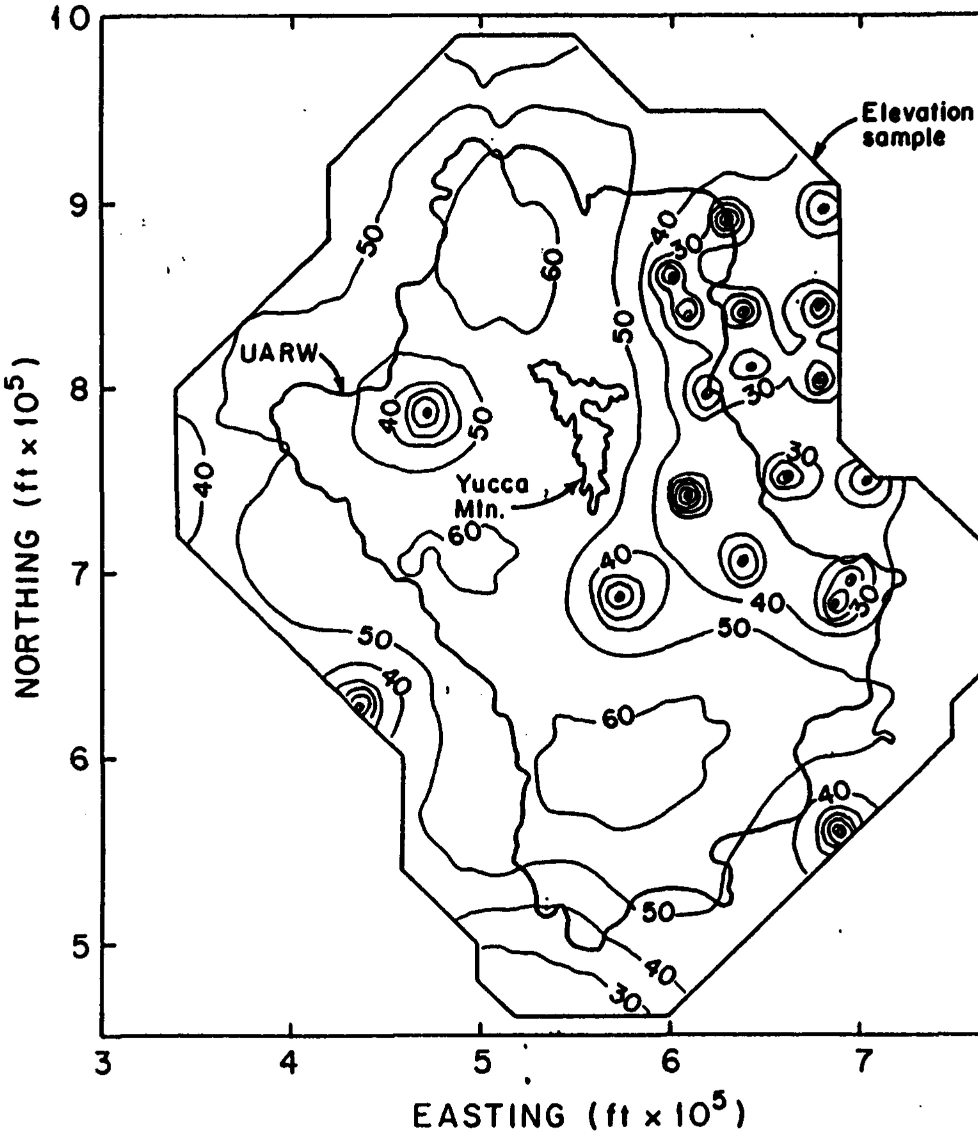

Figure II.6 Percent reduction in estimation variances obtained by cokriging.

values of 250 mm for elevation exceeding 6,000 ft. at the northern extent of Yucca Mtn. This represents a maximum gradient from south to north of approximately 36 mm/100 ft.

A scatterplot indicates that cokriging preserved the log-linear relationship between AAP and elevation (Figure II.4a). A regression equation was fit to the cokriging estimates and elevations:

$$AAP^{*} = \exp(4.26 + 0.00020E) \qquad (5)$$
$$r^2 = 0.79$$

For the range of elevations representative of Yucca Mtn. (3,000 to 6,000 ft), equation 5 defines a gradient for AAP of +31 mm/1000 ft., and this is in close agreement with the observed gradient defined by the samples used to develop the multivariate model (Chapter I). Local relief within the area of the proposed repository site is approximately 900 ft., therefore, values of AAP are predicted to vary 28 mm across the site. As represented by the elevation data used in this study, relief within the repository site is only 500 ft.

Contours of the quantity: cokriging estimate - kriging estimate conformed closely to contours of elevation; large positive values (cokriging estimates exceed kriging estimates) occurred along the crests of mountain ranges while negative values occurred in the basins and valleys (Figure II.5). The largest values occurred in areas of high local relief and in areas farthest from the precipitation stations. At Yucca Mtn., cokriging estimates were up to 60 mm larger than the kriging estimates. Cokriging estimates were up to 40 mm smaller than the kriging estimates at Yucca Flat and in the southern portion of the Amargora Desert.

Comparison of Kriging and Cokriging Estimation Variances

Both kriging and cokriging estimation variances showed the same trends with small values near the cluster of stations on the NTS and large values at the boundaries of the UARW (Figures II.1b and II.3b). However, estimation variances were reduced 30% by cokriging for approximately 90% of the UARW (Figure II.6). Reductions in estimation variances of up to 40% occurred at distances of 20,000 to 30,000 ft (3.8 to 5.7 miles) from precipitation stations. The largest reductions in estimation variances (65%) occurred in the northern portion of the UARW and reductions of more than 60% occurred in the south-central and western portions of the UARW. Reductions in estimation variance of 55% occurred at the proposed repository site on Yucca Mtn. The magnitude of reductions in estimation variances decreased toward the boundaries of the UARW, as a result of edge effects caused by the boundary of the elevation sample. This is most pronounced in the southern, northern, and western extents of the UARW, where reductions are less than

40%. For these locations, the elevation search neighborhood does not include more than one or two precipitation stations, and information regarding the cross-correlation for TAAP and elevation, as defined by the model cross-variograms greatly diminished relative to locations closer to the clustering of stations within the NTS.

In addition, the use of the relatively small elevation search radius results in a discontinuity in estimation variances in areas where precipitation stations are sparse; when only a few precipitation stations are included in the sliding neighborhood, the loss or addition of a single station resulted in an abrupt change in estimation variance.

In general, the areas of maximum estimation variance were similar for kriging and cokriging. The reason for this is the available elevation data are located on a regular grid resulting in a relatively uniform reduction in estimation variances throughout the UARW. Thus, the location (but not the magnitude) of kriging and cokriging estimation variances is determined primarily by the spatial distribution of the precipitation stations.

Vegetated Zone Analysis

Kriging and cokriging estimates were compared with "vegetated zones", which were areas identified on topographic maps, areal photographs, and in the field as being more densely vegetated woodlands. In general, these zones define the approximate boundaries of fully developed Pinyon Pine - Juniper communities, and are representative of higher elevations (about 6,000 ft and above) in the UARW and vicinity which receive relatively more precipitation and have lower air temperatures. The zones did not include lower elevation areas that were more heavily vegetated due to the occurrence of springs or a shallow water table.

The 250 mm isohyet was selected for comparison because of the location of the 250 mm cokriged isohyet relative to the vegetated zones, and also because of the locations of precipitation stations with values of AAP of 250 mm and greater relative to the vegetated zones. The objective was to compare the shape of the kriged and cokriged surfaces with the locations of the vegetated zone boundary. The actual value of the selected isohyet was considered as being somewhat arbitrary. However, Houghton et al. (1975) investigated generalized relationships among elevation, annual precipitation, and the distribution of plant communities in the Great Basin region of Nevada and observed a correlation between the lower limit of the Pinyon - Juniper community, and elevation of approximately 5,900 ft., and AAP of 250 mm. Unpublished correlations between elevation, annual precipitation, and plant communities in the southern Nevada region helped substantiate 250 mm as a reasonable value to be correlated with the lower boundary of the Pinyon - Juniper community (John Emerick, Colorado School of Mines, personal communication).

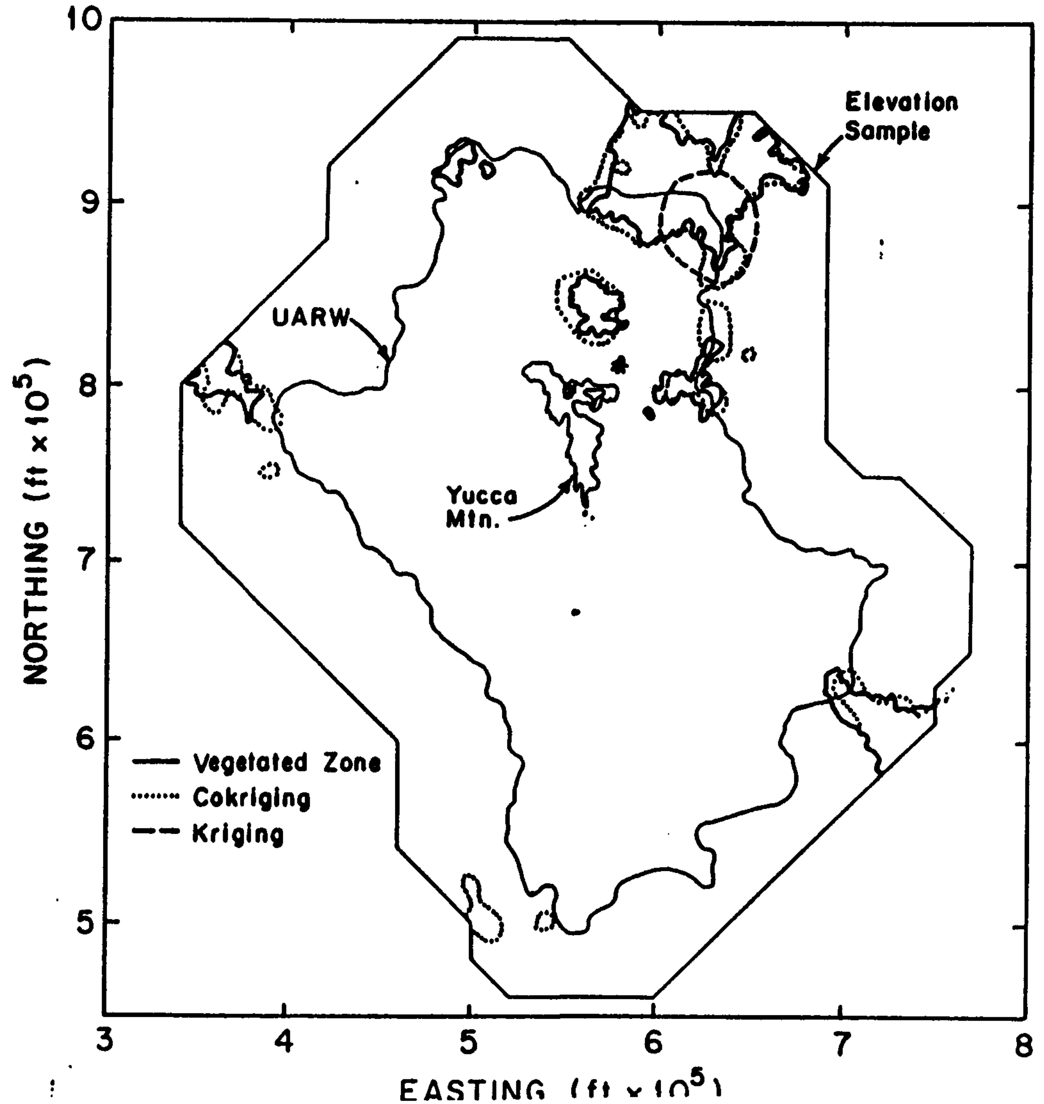

Figure II.7 Comparison of vegetated zone boundaries with 250 mm isohyet obtained by kriging and cokriging.

In general, the 250 mm kriging isohyet did not coincide with vegetated zone boundaries (Figure II.7), while the 250 mm cokriging isohyet coincided very closely with the vegetated zone boundaries, specifically in the areas of Shoshone and Timber Mtns., Yucca Mtn., and the summit of Dome Mtn. (Figure II.7). The 250 mm cokriging isohyet is also present at elevations above 6,000 ft in the Black Mtns. but there is no vegetated zone in this area. Cokriged estimates for this areal may be erroneously high due to a rain-shadow effect from Telescope Peak and because there are no precipitation stations within the elevation search radius for grid points in this area. The 250 mm cokriging isohyet also fails to conform to the vegetated zone boundary identified on the summit of Quartz Mtn. and Black Mtn. (directly adjacent to Quartz Mtn.), along the northwest corner of the UARW boundary. This value is lower than expected for this elevation assuming a stationary correlation for AAP and elevation in the southern Nevada - southeastern California region, and causes estimates for the summits of Quartz Mtn. (180 mm) and Black Mtn. (210 mm) also to be lower than expected for their elevations. According to the hypothesis of French (1983), station 228 is located in the "deficit" zone, and a stationary correlation for AAP and elevation should not be assumed for the entire southern Nevada - southeastern California region. The location of vegetated zones on Quartz Mtn. and Black Mtn., however, suggests that a significant increase in AAP does exist in this area. Note that the cokriged estimates are 40 to 60 mm greater than the kriged estimates for the summit locations.

Global Estimates for AAP

Global (area averaged) estimates for AAP were obtained by computing the arithmetic average of AAP estimates for points within the UARW (Istok and Cooper, 1988). The global estimate obtained using the cokriging estimates was 157 mm, 12 mm greater than that obtaining using the kriging estimates. The value of 157 mm is within the range of 141 to 170 mm obtained using regression, but is not within the range of 159 to 178 mm obtained using alternate interpolation methods. The wide range in the regression results can be expected because of considerable differences in the datasets used to develop each equation (Chapter I).

Effectiveness of Currently Active Stations for Estimating AAP

Because of the importance of AAP to water balance calculations, ten additional precipitation stations have recently been established in the vicinity of Yucca Mtn. The data from these stations and from the 28 stations used in this study that are currently active will be used in the future to estimate AAP at the site. Therefore it is desirable to estimate the magnitude of estimation variances that can be obtained using this network. The cokriging estimation variances will be reduced substantially at Yucca Mtn. by this network but will remain fairly large in the western half of the UARW (Figure II.8). With the present network, estimation variances at the proposed repository site are reduced to approximately 400 mm^2. Maximum estimation variances within the UARW still occur at the southern extent, the northern extent, and the western extent of the boundary, where variances are 4,250 mm^2, 3,500 mm^2, and 3,250 mm^2, respectively. Variances are less than 1,250 mm^2 for a large area comprising the northeast portion of the UARW due to the existing precipitation network.

Effectiveness of Additional Fictitious Stations

The method of fictitious points (Journel and Huijbreghts, 1978) was used to identify locations for additional stations in an attempt to reduce estimation variances for AAP in the western half of the UARW. Twenty additional locations were identified (Figure II.9). The numbers indicate the order in which the locations were identified. The procedure used the point of maximum estimation variance to identify each successive station location. Once a location was identified, that location remained fixed according to the procedure used in this study, and this influenced all successive locations. This resulted in 12 of the 20 identified locations occurring on the UARW boundary (Figure II.9). Such a configuration would probably not have been the case if locations had been identified using reductions in the average estimation variance within the UARW, as opposed to maximum point reductions in estimation variance within the UARW.

The map of estimation variances obtained with the addition of the 20 fictitious stations predicts significant improvements in estimation reliability for the UARW (Figure II.9). Variances for approximately 90% of the UARW are less than 1,250 mm^2, with a maximum estimation variance of 1,300 mm^2 occurring along the northwest boundary. In general, estimation variances throughout the UARW are much more uniform than those obtained from the presently operating precipitation network. Steep gradients in estimation variances occur only near precipitation stations, and beyond the boundaries of the UARW, where variances steadily increase to over 2,500 mm^2.

No reduction in estimation variances occurred in the vicinity of the proposed repository. Estimation variances were reduced by only 400 mm^2 for a small portion along the northwestern extent of the Yucca Mtn. massif. In addition, no reductions occurred throughout the entire northeast portion of the UARW due to the existing network of NTS stations. From this analysis, it can be concluded that the spatial structure of AAP over the

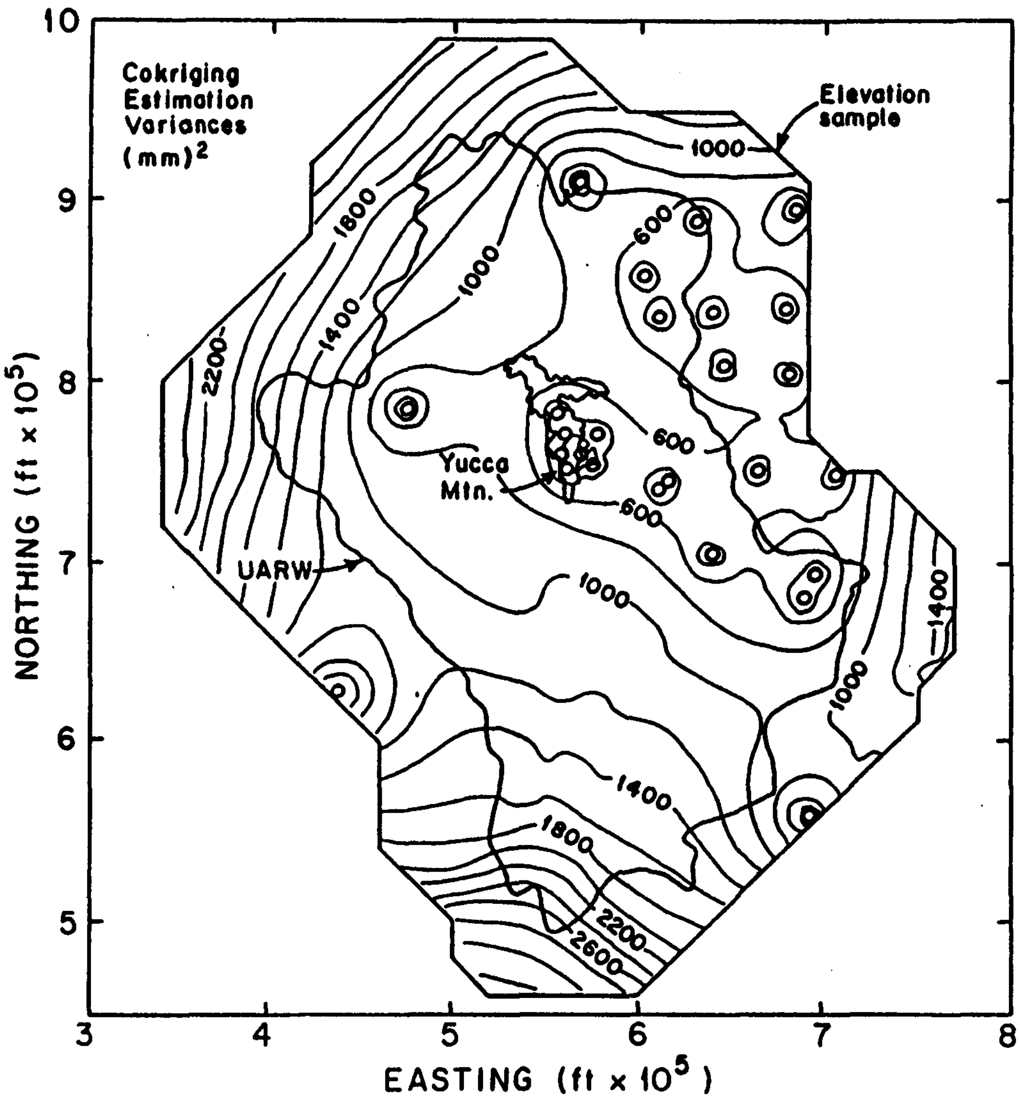

Figure II.8 Cokriging variances for 38 active precipitation stations.

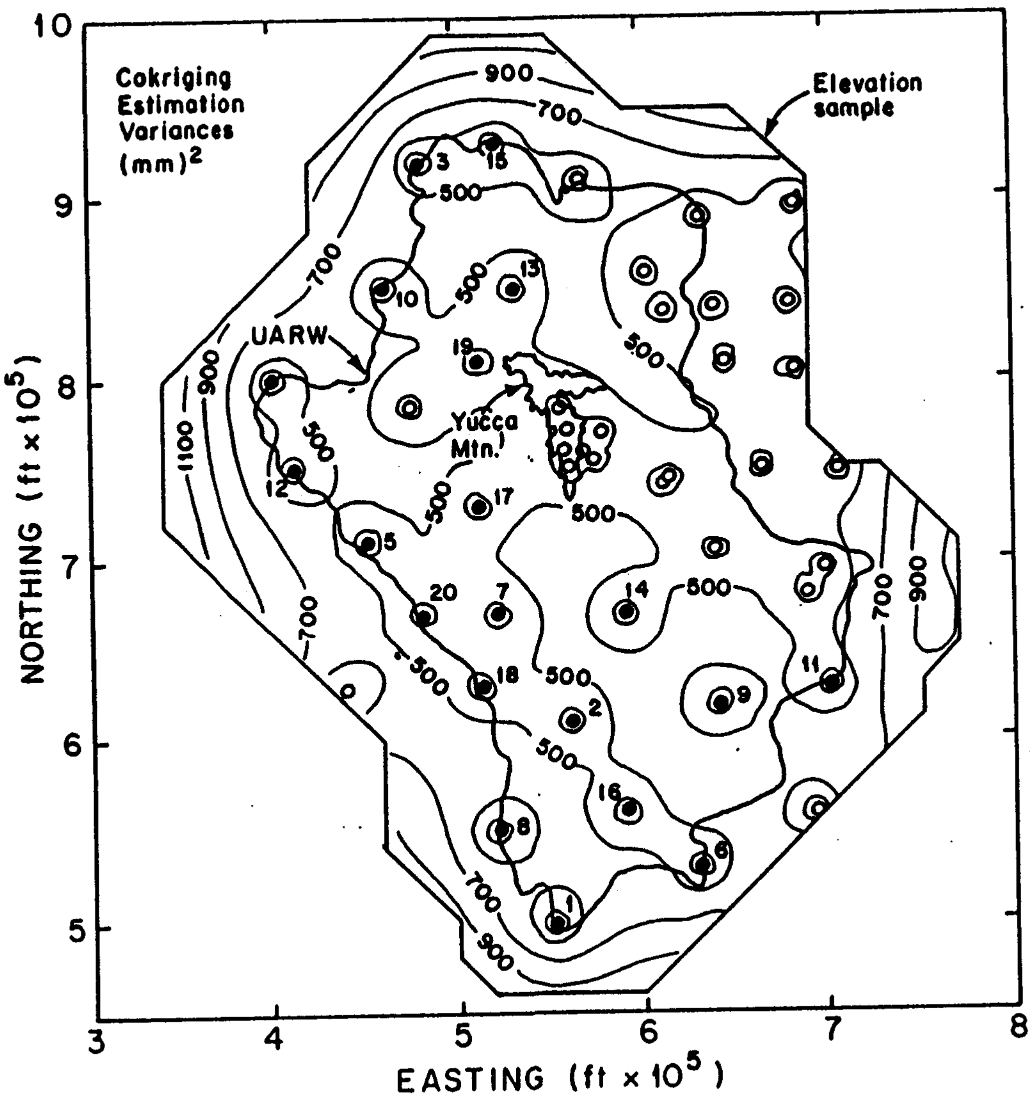

Figure II.9 Cokriging variances for 38 active precipitation and 20 fictitious stations.

UARW is such that estimates made in the northern portion would not be influenced by potential measurements available in the southern portion of the watershed.

A plot of the observed maximum and average cokriging variance was used to investigate the relative influence of each successive fictitious station in improving the efficiency of the network for providing accurate estimates within the UARW (Figure II.10). As expected, the maximum variance decreases much more rapidly than the average variance, and the addition of the first 3 to 4 stations provides the most significant reductions. Addition of the 20th station does not provide a significant improvement in overall estimation accuracy. Predictions concerning potential improvements in sample efficiency were made by fitting empirical formulas to the observed values. To reduce extrapolation error, the fact that both the maximum and average variance would be zero if fictitious stations were located at all of the 747 grid points within the UARW was used during regression to find the best fit curves. This assumes that the selected grid adequately models AAP on the scale of UARW, and that the variability of AAP between stations separated by a distance of 10,000 ft (approximately 2 mi) can be neglected.

In general, the curves cannot be used as exact predictors of estimation accuracy, but are somewhat useful in predicting the relative efficiency of additional stations in improving estimation accuracy. The formulas predict that the maximum and average cokriging variance would be 300 mm^2 for a total of 40 precipitation stations added to the existing network. If assumptions concerning the multivariate log-normal distribution of AAP are correct, an average estimation error of +/- 34 mm is calculated for a 95% confidence interval. For 50 additional stations, the formulas predict that the maximum variance would be less than the average variance, which is an impossible result caused by extrapolation error.

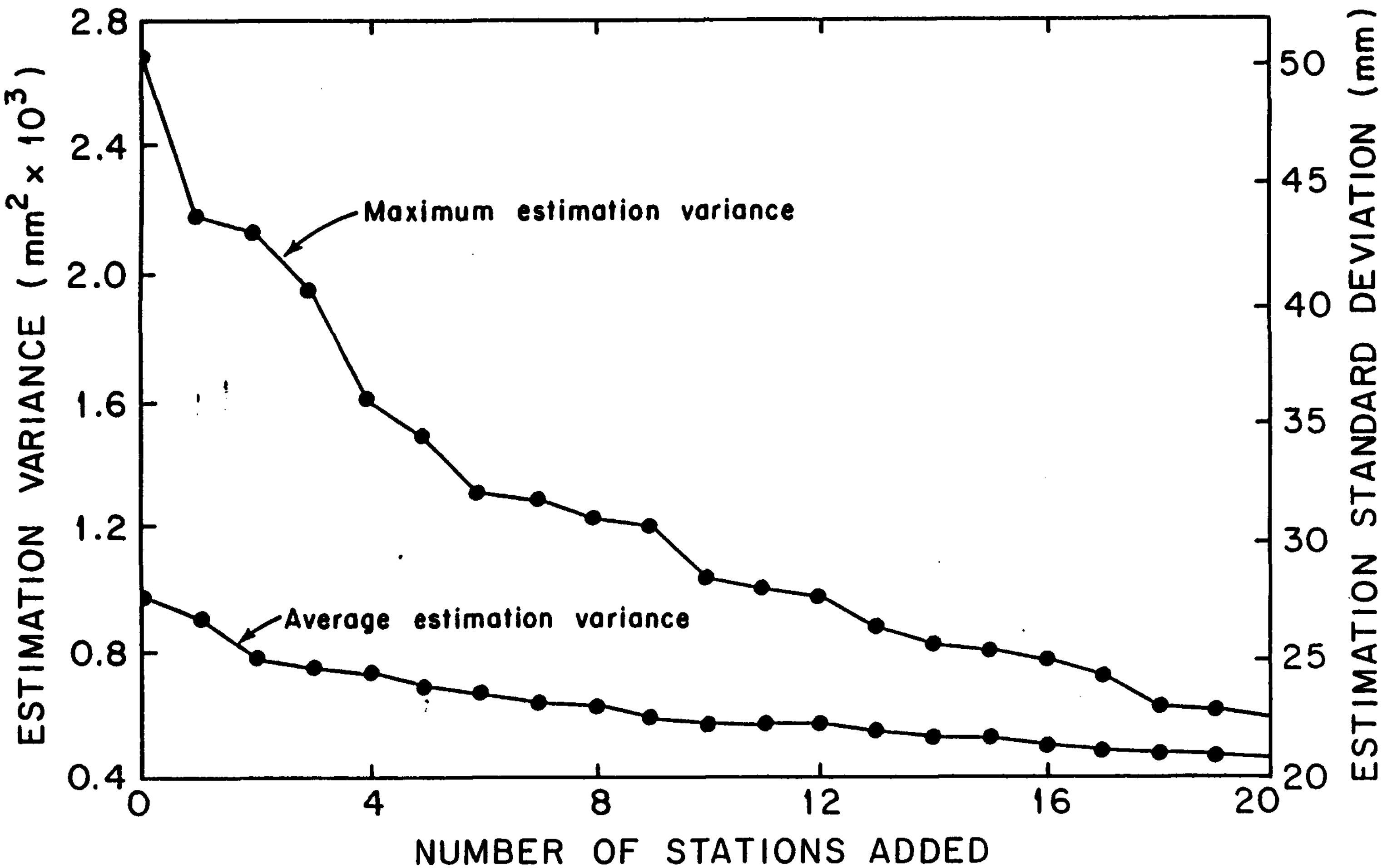

Figure II.10 Reduction in estimation variance with the addition of fictitious stations.

Summary and Conclusions

Cokriging was used to estimate average annual precipitation (AAP) at Yucca Mtn., site of a proposed, high-level nuclear waste repository. The estimation procedure utilized the positive correlation between AAP and elevation reported in Chapter I, precipitation data for 42 stations in the southern Nevada and southeastern California region surrounding the site, and elevation data for 1,531 points located on a grid within the Upper Amargosa River Watershed (UARW). The cokriging estimate for AAP was 165 mm for the approximate center of the proposed repository site. A similar estimate was obtained using regression equations developed in Chapter I and French (1986). The estimate obtained using kriging and the precipitation data only was 145 mm, while estimates obtained using other interpolation methods were about 175 mm.

Previous estimates for AAP at Yucca Mtn. range from 100 to 150 mm/year (Winograd and Thordarson 1975, Quiring 1983). A value of 150 mm was used to help estimate net infiltration at Yucca Mtn. by Montazer and Wilson (1984). An estimate of net infiltration based on an AAP value of 150 mm was also used by Czarnecki (1985) to calculate groundwater recharge for use in a numerical flow model for the site.

Cokriging reduced the estimation variances by 55% at the proposed repository site. The cokriging estimates conformed closely to the topography and displayed known orographic influences of mountain ranges and basins. An indirect confirmation of the cokriging estimates was obtained by comparing contours of AAP estimates with the boundaries of more highly vegetated or forested zones. In most cases the 250 mm contour corresponded closely with the boundaries of these zones.

The effectiveness of a network of currently active stations for estimating AAP was evaluated. Based on these results we concluded that: 1) Precipitation stations are needed along the southern and western boundaries of the UARW, because stations to the south and to the west of the UARW are very sparse or completely lacking. This should be taken into consideration even if reductions in the average estimation variance within the UARW is desired. 2) Precipitation stations are also needed within the Amargosa Desert. This is a relatively low elevation area (2,000 to 2,500 ft), and stations here would be important in helping to model the elevation - precipitation correlation within the UARW. 3) Additional precipitation stations are needed to the northwest and to the southwest of Yucca Mtn. to help reduce high estimation variances close to Yucca Mtn. Currently, only one station (Beatty) exists to the west of Yucca Mtn. within the UARW. This station is located at a distance of almost 20 miles from the proposed repository site.

References

Cooper, R.M. and J.D. Istok, 1988: Geostatistics Applied to Groundwater Contamination. I: Methodology - J. Environ. Engrg., ASCE, 114(2), 270-286.

French, R. H., 1986: Daily, Seasonal, and Annual Precipitation at the Nevada Test Site, Nevada. U.S. Dept. of Energy, Nevada Operations Office, #DE-AC08-85NV10384, Publication No. 45042.

Golden Software Inc. 1989. Surfer. P.O. Box 281, Golden, Colorado 80402.

Hevesi, J.A., A.L. Flint, and J.D. Istok. 19XX: Precipitation Estimation in Mountainous Terrain using Multivariate Geostatistics: 1. Structural Analysis. Submitted to Journal of Applied Meteorology.

Istok, J.D. and R.M. Cooper, 1988: Geostatistics Applied to Groundwater Contamination. III. Global Estimates - J. Environ. Engrg., ASCE, 114(4), 915-928.

Journel, A.G. and C.J. Huijbregts, 1978: Mining Geostatistics. Academic Press, New York, N.Y.

Montazer, P. and W. E. Wilson, 1984: Conceptual Hydrologic Model of Flow in the Unsaturated Zone, Yucca Mountain, Nevada. U.S. Geological Survey Water-Resources Investigations Report 84-4345.

Quiring, R. F., 1983: Precipitation Climatology of the Nevada Test Site. Nuclear Support Office, Las Vegas, NV.

Bibliography

Clark, I., 1979: Practical Geostatistics. Applied Science Publishers, London, U.K.

Cooper, R.M. and J.D. Istok, 1988: Geostatistics Applied to Groundwater Contamination. I: Methodology - J. Environ. Engrg., ASCE, 114(2), 270-286.

Cooper, R. M., J. D. Istok, and A. L. Flint, 1988: Three-dimensional, cross-semivariogram calculations for hydrogeological data. Ground Water, 26, 638-646.

David, M., 1977: Geostatistical Ore Reserve Estimation. Elsevier Scientific Publishing Company, New York, N.Y.

Delhomme, J. P., 1979: Spatial Variability and Uncertainty in Groundwater Flow Parameters: A Geostatistical Approach. Water Resources Research, 15, 269-280.

Dingman, S. L., D. M. Seely-Reynolds, and R. C. Reynolds III, 1988: Application of Kriging to Estimating Mean Annual Precipitation in a Region of Orographic Influence. Water Resources Bulletin, 24(2), 329-339.

French, R. H., 1987: Effects of the Length of Record on Estimates of Annual and Seasonal Precipitation at the Nevada Test Site. U.S. Dept. of Energy, Nevada Operations Office, #DE-AC08-85NV10384.

French, R. H., 1986: Daily, Seasonal, and Annual Precipitation at the Nevada Test Site, Nevada. U.S. Dept. of Energy, Nevada Operations Office, #DE-AC08-85NV10384, Publication No. 45042.

French, R. H., 1983: A Preliminary Analysis of Precipitation in Southern Nevada. DOE/NV/10162-10, Water Resources Center, Desert Research Institute, Las Vegas, Nevada, 39.

Golden Software Inc. 1989. Surfer. P.O. Box 281, Golden, Colorado 80402.

Hevesi, J.A., A.L. Flint, and J.D. Istok. 19XX: Precipitation Estimation in Mountainous Terrain using Multivariate Geostatistics: 1. Structural Analysis. Submitted to Journal of Applied Meteorology.

Isaaks, E. H. and R. M. Srivastava. 1989: An Introduction to Applied Geostatistics. Oxford University Press, New York.

Istok, J.D. and R.M. Cooper, 1988: Geostatistics Applied to Groundwater Contamination. III. Global Estimates - J. Environ. Engrg., ASCE, 114(4), 915-928.

Jones, K.R., 1981: Arid Zone Hydrology for Agricultural Development, FAO Irrigation and Drainage Paper No. 37, Food and Agriculture Organization of the United Nations, Rome.

Journel, A. G. and C. J. Huijbregts, 1978: Mining Geostatistics. Academic Press, New York, N.Y.

Montazer, P. and W. E. Wilson, 1984: Conceptual Hydrologic Model of Flow in the Unsaturated Zone, Yucca Mountain, Nevada. U.S. Geological Survey Water-Resources Investigations Report 84-4345.

Quiring, R. F., 1983: Precipitation Climatology of the Nevada Test Site. Nuclear Support Office, Las Vegas, NV.

Renard, D., and J. D. Yancey, 1984: Smoothing Discontinuities when Extrapolating Using Moving Neighborhood. Geostatistics for Natural Resources Characterization, Part II, 679-690, D. Reidel Publishing Co.

Tabios, G. Q. III., and J. D. Salas, 1985: A Comparative Analysis of Techniques for Spatial Interpolation of Precipitation. Water Resources Bulletin, 21(3), 365-380.

Yates, S.R., and A.W. Warrick, 1987: Estimating Soil Water Content Using Cokriging. Soil Sci. Soc. of Am. J., 51, 23-30.