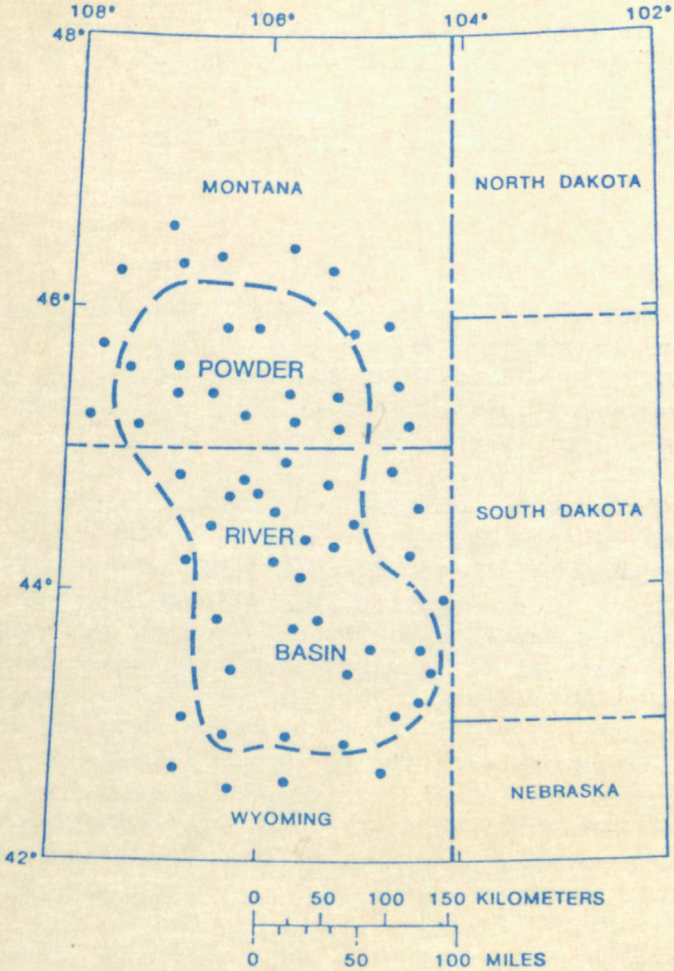
KRIGING ANALYSIS OF MEAN ANNUAL PRECIPITATION, POWDER RIVER BASIN, MONTANA AND WYOMING



U.S. GEOLOGICAL SURVEY

Water-Resources Investigations 80-50

50272-101

REPORT DOCUMENTATION PAGE	1. REPORT NO.	2.	3. Recipient's Accession No.

4. Title and Subtitle
Kriging Analysis of Mean Annual Precipitation, Powder River Basin, Montana and Wyoming

5. Report Date
Approved May 1980

6.

7. Author(s)
M. R. Karlinger and J. A. Skrivan

8. Performing Organization Rept. No.
USGS/WRI 80-50

9. Performing Organization Name and Address
U.S. Geological Survey, Water Resources Division
1201 Pacific Avenue, Suite 600
Tacoma, Washington 98402

10. Project/Task/Work Unit No.

11. Contract(C) or Grant(G) No.
(C)
(G)

12. Sponsoring Organization Name and Address
U.S. Geological Survey, Water Resources Division
1201 Pacific Avenue, Suite 600
Tacoma, Washington 98402

13. Type of Report & Period Covered
Final

14.

15. Supplementary Notes

16. Abstract (Limit: 200 words)

Kriging is a statistical estimation technique for regionalized variables that exhibit an autocorrelation structure. Such structure can be described by a semi-variogram of the observed data. The punctual-kriging estimate at any point is a weighted average of the data, where the weights are determined by using the semi-variogram and an assumed drift, or lack of drift, in the data. Block, or areal, estimates can also be calculated.

The kriging algorithm, based on unbiased and minimum-variance estimates, involves a linear system of equations to calculate the weights. Kriging errors, a by-product of the calculations, can then be used to give confidence intervals of the resulting estimates.

Mean annual precipitation is an important variable when considering restoration of coal strip-mining lands in the Powder River basin, Montana and Wyoming. Two punctual-kriging analyses involving data at 60 stations were made--one assuming no drift in precipitation and one a partial quadratic drift simulating orographic effects. Contour maps of estimates of mean annual precipitation based on a period of 10 years of data were similar for both analyses, as were the corresponding kriging errors. Kriging with data uncertainty, as measured by standard errors of the mean from the 10 years of data, was also performed. Block estimates of mean annual precipitation were made for two subbasins. Runoff estimates were 1-2 percent of the kriged block estimates.

17. Document Analysis a. Descriptors

*statistical methods, Contours, *regional analysis, network design, *precipitation (atmospheric)

b. Identifiers/Open-Ended Terms

*kriging, *universal kriging, *variogram, *semi-variogram, *geostatistics, stochastic interpolation, *mean annual precipitation, *Powder River Basin, Montana, Wyoming

c. COSATI Field/Group

18. Availability Statement	19. Security Class (This Report)	21. No. of Pages
No restriction on distribution	UNCLASSIFIED	29
	20. Security Class (This Page) UNCLASSIFIED	22. Price

(See ANSI-Z39.18) See Instructions on Reverse

OPTIONAL FORM 272 (4-77)
(Formerly NTIS-35)
Department of Commerce

(200)
WRi

KRIGING ANALYSIS OF MEAN ANNUAL PRECIPITATION, POWDER RIVER BASIN, MONTANA AND WYOMING

By M. R. Karlinger and J. A. Skrivan

U.S. GEOLOGICAL SURVEY

Water-Resources Investigations 80-50

1981

UNITED STATES DEPARTMENT OF THE INTERIOR

JAMES G. WATT, Secretary

GEOLOGICAL SURVEY

Doyle G. Frederick, Acting Director

For additional information write to:

U.S. Geological Survey
1201 Pacific Avenue - Suite 600
Tacoma, Washington 98402

CONTENTS

ILLUSTRATIONS

TABLE

CONVERSION FACTORS

Multiply	By	To obtain
inch (in.)	25.4	millimeter (mm)
	2.540	centimeter (cm)
square inch (in^2)	645.16	square millimeter (mm^2)
	6.452	square centimeter (cm^2)
mile (mi)	1.609	kilometer (km)

KRIGING ANALYSIS
OF MEAN ANNUAL PRECIPITATION,
POWDER RIVER BASIN,
MONTANA AND WYOMING

By M. R. Karlinger and J. A. Skrivan

ABSTRACT

Kriging is a statistical estimation technique for regionalized variables that exhibit an autocorrelation structure. Such structure can be described by a semi-variogram of the observed data. The punctual-kriging estimate at any point is a weighted average of the data, where the weights are determined by using the semi-variogram and an assumed drift, or lack of drift, in the data. Block, or areal, estimates can also be calculated.

The kriging algorithm, based on unbiased and minimum-variance estimates, involves a linear system of equations to calculate the weights. Kriging errors, a by-product of the calculations, can then be used to give confidence intervals of the resulting estimates.

Kriging with no data uncertainty is an exact interpolator at the observations. However, such uncertainty can be included in the analysis to take into account data of different reliability.

Mean annual precipitation in the Powder River basin, Montana and Wyoming, is an important variable when considering restoration of coal strip-mining lands of the region. Two punctual-kriging analyses involving data at 60 stations were made -- one assuming no drift in precipitation and one a partial quadratic drift simulating orographic effects. Contour maps of estimates of mean annual precipitation based on a period of 10 years of data were similar for both analyses, as were the corresponding kriging errors. Kriging with data uncertainty, as measured by standard errors of the mean from the 10 years of data, was also performed. Block estimates of mean annual precipitation were made for two subbasins. Runoff estimates were 1-2 percent of the kriged block estimates.

INTRODUCTION

The main purpose of this study is to demonstrate the use of kriging in energy-related problems. The particular application is to describe mean annual precipitation in the Powder River basin in Montana and Wyoming. The distribution of precipitation is an important factor in estimating reclamation success in coal strip-mining areas of the basin.

Kriging is a regionalization technique for estimation that uses the autocorrelation structure of the observed data. If there is no apparent autocorrelation between data points, then kriging has no advantage over fitting a least-squares trend through the data. However, mean annual precipitation exhibited a sufficient autocorrelation structure to satisfy the purpose of the study. All data used in the analysis can be found in Toy and Munson (1978).

Only a cursory review of the theory behind kriging is given. However, sufficient references are included for the interested reader to obtain a more complete development.

THE KRIGING SYSTEM

Until recently, the majority of kriging applications has been in the field of mining for estimating ore reserves. In the past few years, however, kriging has been applied in foreign countries to hydrologic studies dealing with precipitation distribution, ground-water levels, and aquifer-parameter estimation (Delhomme, 1978; and Delfiner and Delhomme, 1975).

Kriging is an interpolation technique with which values of a regionalized variable can be estimated at a point (punctual kriging) or over an area (block kriging). These estimates are a weighted sum of the observed data in which the weights are determined by using the spatial distribution of the observations. An important by-product of kriging, not found in ordinary regionalization techniques, is a variance estimate of the phenomena at each interpolated point. These variances, or kriging errors, are often used to define confidence intervals about the estimate if errors are assumed to be normally distributed.

There are many descriptions of the kriging system and derivation of the kriging equations, such as Olea (1975) or Matheron (1971). Therefore, only a brief account of the kriging algorithm will be given.

The kriging estimate is a linearly weighted combination of the known data, Z_i, at locations p_i having coordinates (X_i,Y_i). Z_i will be denoted by $Z(p_i)$. Assume there are N such data points in the domain of interest. A punctual-kriging estimate, Z, at any desired location p_o is therefore of the form

$$\hat{Z}(p_o) = \sum_{i=1}^{N} w_i Z(p_i) \tag{1}$$

where $\{w_i\}_{i=1,\ldots,N}$ is a set of weights chosen such that $\hat{Z}(p_o)$ is an unbiased estimate of $Z(p_o)$ and the estimation error for $Z(p_o)$ is minimum. These two criteria can be mathematically stated in the following equation set:

$$\text{minimize } E\,(\hat{Z}(p_o) - Z(p_o))^2 \tag{2}$$

$$\text{subject to } E\,(\hat{Z}(p_o) - Z(p_o)) = 0 \tag{3}$$

where $E(\cdot)$ is the expectation. Substitution of equation 1 into conditions 2 and 3 results in the following equation set:

$$\text{minimize } E\,(\sum_{i=1}^{N} w_i Z(p_i) - Z(p_o))^2$$

$$\text{subject to } E\,(\sum_{i=1}^{N} w_i Z(p_i) - Z(p_o)) = 0\,.$$

The method of solution employed to obtain the weights, w_i, is a Lagrangian optimization (Taylor, 1955). This technique produces the following system of equations (Olea, 1975):

$$\sum_{i=1}^{N} w_i \gamma(h_{ij}) + \mu_o = \gamma(h_{oj}) \quad \text{for all } j \tag{4}$$

$$\sum_{i=1}^{N} w_i = 1 \tag{5}$$

where $\gamma(h)$ is the semi-variogram,
h_{ij} is the distance between observation points p_i and p_j,
μ_o is the Lagrangian multiplier, and
h_{oj} is the distance between the point p_o where an estimate is desired and the point p_j.

The semi-variogram is defined by:

$$\gamma(h) = \tfrac{1}{2}\,\mathrm{Var}\,(Z(p+h) - Z(p)) \tag{6}$$

where Var $(\cdot)$ is the variance. In other words, $\gamma(h)$ is defined by differences, or increments, of the regionalized variable Z over a distance h. $\gamma(h)$ can be approximated from the given set of observations by using the following empirical semi-variogram:

$$\gamma_E(h_\ell) = \frac{1}{2M(h_\ell)} \sum_{i=1}^{M(h_\ell)} (Z(P_{i_1}) - Z(P_{i_2}))^2 \tag{7}$$

where h_ℓ is the distance between observation points p_{i_1} and p_{i_2}, and $M(h_\ell)$ is the number of pairs of points a distance h_ℓ apart. $\gamma_E(h)$ is thus the average square of the difference between Z_i observations a given distance h apart and is calculated for several values of h_ℓ, depending on the spatial distribution of the observations.

In order to solve for the weights w_i in equations 4 and 5, a theoretical form of $\gamma(h)$ (Delhomme, 1979) must be chosen and then evaluated at the required distances. This choice is made by fitting $\gamma_E(h)$ to one of several common forms of $\gamma(h)$ (Skrivan and Karlinger, 1980). Once this form is chosen, then $\gamma(h_{ij})$ and $\gamma(h_{oj})$ are evaluated and equations 4 and 5 are solved for w_i.

Equations 4 through 7 apply for the case in which the expected value of the phenomenon is constant over the domain, that is, $E(Z(p_i)) = C$, a constant, for all i. This situation is termed constant drift or, equivalently, no drift, and the analysis is called kriging, or simple kriging. If the expected value of the phenomenon is not constant but changes in a gradual manner over the domain, that is, has a trend, or drift, and the change can be characterized by a low-order polynomial, then the following equations should replace equations 4 and 5:

$$\sum_{i=1}^{N} w_i\,\gamma(h_{ij}) + \sum_{k=0}^{K} \mu_k f^k(p_i) = \gamma(h_{oj}) \text{ for all } j \tag{8}$$

$$\sum_{i=1}^{N} w_i\, f^k(p_i) = f^k(p_o) \qquad k = 0, \cdots, K \tag{9}$$

where the $f^k(p)$ designate the K+1 monomials used in the drift equation. For example, possible choices for these monomials are $f^0(p) = 1$, $f^1(p) = X$, $f^2(p) = Y$, and $f^3(p) = X^2$. A polynomial of degree three is the maximum needed for most cases. The μ_k's are the corresponding Lagrangian multipliers for this system of equations. The empirical semi-variogram for the system described by equations 8 and 9 is the following:

$$\gamma_E(h_\ell) = \frac{1}{2M(h_\ell)} \sum_{i=1}^{M(h_\ell)} (R(p_{i_1}) - R(p_{i_2}))^2 \tag{10}$$

where the $R(p_i)$ are the residuals of the $Z(p_i)$ off the drift. The coefficients of the drift are not explicitly needed to solve the new system of equations for the w_i. However, the drift has to be known to calculate the $R(p_i)$ in equation 10. Therefore, kriging with this sytem, known as universal kriging, is an iterative process. For suggestions on the use of universal kriging see Skrivan and Karlinger (1979).

Once the system of equations 4 and 5, or 8 and 9, are solved for w_i, the estimates of $Z(p_o)$ can be made by using equation 1. The kriging error, $\hat{\sigma}_K$ (p_o), for the system of equations 4 and 5 (constant-drift case) is

$$\hat{\sigma}_K(p_o) = \sqrt{\sum_{j=1}^{N} w_j \gamma(h_{oj}) + \mu_o} \,. \tag{11}$$

For the system of equations 8 and 9 (drift case) the kriging error is

$$\hat{\sigma}_K(p_o) = \sqrt{\sum_{j=1}^{N} w_j \, \gamma(h_{oj}) + \sum_{k=0}^{K} \mu_k \; f^k(p_o)} \,. \tag{12}$$

It can be shown from both systems of equations that, in the absence of data uncertainty such as sampling errors, estimating a point already in the kriging system, that is, one of the data points, reproduces the observed value at that point. Hence, kriging is an exact interpolator and the kriging error at these observed points is zero.

In most kriging analyses, the final form of the theoretical semi-variogram or of the semi-variogram-drift combination is based upon the results of a validation process. Such validation is the ultimate test for choosing both the theoretical semi-variogram and terms of a drift. The process is a modified split-sample procedure in which the data points are eliminated individually from the data set, then estimated by using the remaining points. There is thus a residual error corresponding to each point in the data set. These errors are then compared to the theoretical kriging errors ($\hat{\sigma}_K$) for each point. The semi-variogram or semi-variogram-drift combination that produces an average residual error near zero and that equates the mean square residual error to the mean kriging error (Gambolati and Volpi, 1979) is chosen for the final solution. Thus the reduced mean-square error, i.e., the ratio of the mean-square residual error to the mean theoretical error, should be close to 1 for validation. Since the kriging error is theoretically equal to the variance of the estimation error for any point, the validation ratio attempts to insure this in a mean-squared criterion.

In the validation process, the average residual error is not quite as sensitive to refinements in the semi-variogram as is the reduced mean-square error. This is fortunate because achieving a reasonable reduced mean-square error gives more confidence to the kriging error at little or no expense to the kriging estimate, and the kriging error is the most important end product of a kriging analysis.

PROJECT AREA

The study area considered in this report is the semiarid Powder River basin in southeastern Montana and northeastern Wyoming. Its east-west extent is approximately 200 mi and its north-south extent is approximately 300 mi. The western edge of the basin is bordered by the Big Horn mountains and the eastern edge encroaches upon the Black Hills area of South Dakota. Within the basin itself is a figure-8 shaped region of coal strip-mine potential (fig. 1).

The precipitation data used in the analysis were taken from Toy and Munson (1978). The mean annual precipitation values were obtained from a 10-year period of data collection, 1965-74 (table 1). Because 10 years of data at each site are available, an estimate of the standard error of the mean annual precipitation can be calculated for each of these sites. These standard errors can also be incorporated into the kriging system (p. 17). The areal coverage of the basin by precipitation stations (fig. 1) appears to be adequate for regional climate appraisal.

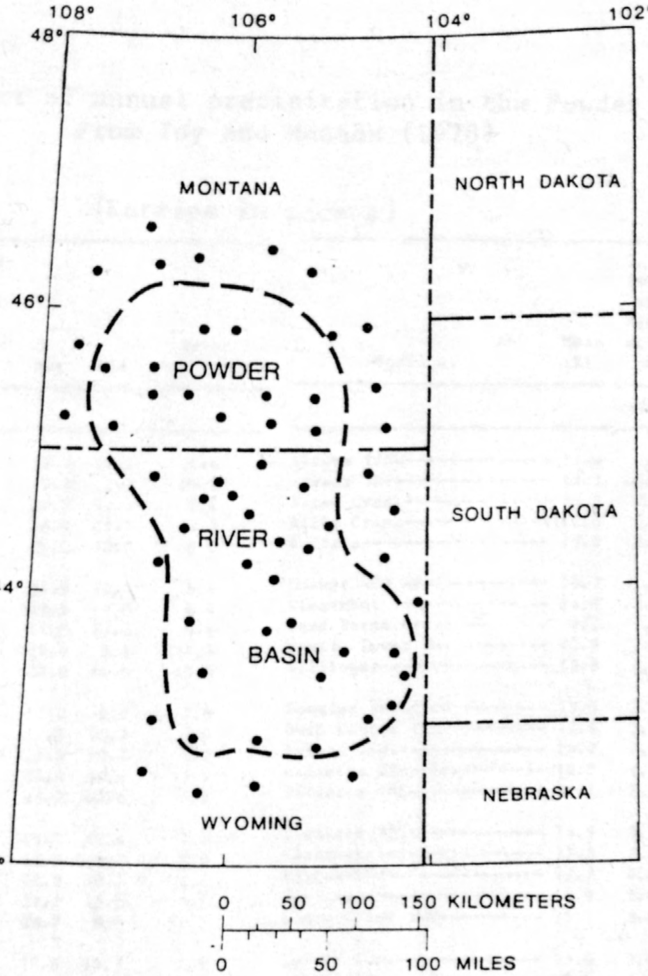
ESTIMATION OF MEAN ANNUAL PRECIPITATION

Semi-Variogram Determination

The first step in the kriging analysis is to estimate a semi-variogram. The maximum distance between any two points in the precipitation network is approximately 300 mi. The h axis of the empirical semi-variogram was subdivided into 13 intervals of 15-mile lengths to provide each point on the semi-variogram sufficient definition ($M(h_{ij}) \geq 50$). The last interval included all distances greater than 195 mi. Equation 7 was then used to calculate the isotropic empirical semi-variogram for a constant-drift case (fig. 2). Examination of the graph of the semi-variogram indicated that a linear equation was appropriate, and the following initial theoretical semi-variogram (fig. 2) was chosen:

$$\begin{aligned} (h_{ij}) &= 0 && h_{ij} = 0 \\ (h_{ij}) &= 3.0 + 0.01667\, h_{ij} && 0 < h_{ij} \leq 300 \text{ mi} \qquad (13) \\ (h_{ij}) &= 8.0 && h_{ij} > 300 \text{ mi} \end{aligned}$$

The value of 3.0 is denoted as the nugget effect which is sometimes the result of an insufficient gage density at small values of h. A microstructure may exist, but the smallest distance between any two gages is still too coarse to sufficiently define the spatial correlation at this smaller scale. Other causes of the nugget effect in the case of precipitation data include instrument error. This results from the gage being located such that it is affected by wind or other factors during precipitation events. Although this would be more prominent in the case of semi-variograms of storm-related precipitation, it can still be a factor in a semi-variogram of mean annual precipitation.



EXPLANATION

Boundary of region of potential mining

• Precipitation station used in the analysis

FIGURE 1.--Location of Powder River basin, region of potential coal-strip mining and precipitation stations. Modified from Toy and Munson (1978).

TABLE 1.--Statistics of annual precipitation in the Powder River basin. From Toy and Munson (1978)

[Entries in inches]

Station	Mean ($\bar{x}$)	Standard deviation (σ)	Standard error of mean (S.E.)	Max	Min	Range (Max - Min)
MONTANA						
Albion 1N	14.4	1.9	0.6	18.8	12.2	6.6
Biddle 8SW	14.6	3.2	1.0	22.0	11.1	10.9
Birney 2SW	14.0	1.7	.5	16.7	11.3	5.4
Brandenburg	15.5	2.5	.8	18.9	11.7	7.3
Broadus	14.4	2.7	.9	18.2	10.2	8.0
Busby	17.2	2.7	.8	21.0	12.7	8.4
Colstrip	16.9	2.1	.7	19.8	13.0	6.9
Crow Agency	17.0	2.8	.9	21.0	11.2	9.8
Custer	13.6	3.0	.9	19.9	8.6	11.3
Ekalaka	18.2	4.7	1.5	27.0	14.0	13.0
Forsyth 2E	13.2	2.5	.8	17.2	9.9	7.3
Hardin	13.9	1.9	.6	17.2	10.2	7.0
Hysham	13.3	1.9	.6	15.9	10.7	5.2
Hysham 19SSE	15.1	2.4	.8	18.4	10.9	7.5
Kirby 1S	20.2	2.4	.8	23.2	15.6	7.6
Miles City FAA AP	16.0	2.2	.7	18.7	11.6	7.1
Mizpah 4NNW	13.5	3.2	1.0	17.7	8.7	9.0
Moorhead 9NE	14.1	2.5	.8	18.8	10.7	8.1
Otter 9SSW	18.9	2.2	.7	21.7	15.0	6.7
Powderville 8NNE	14.8	4.0	1.3	20.7	8.5	12.2
Ridgeway 1S	13.7	2.3	.7	17.8	10.5	7.3
Sonnette 2WNW	15.6	2.2	.7	17.6	12.1	5.5
Wyola	17.5	3.0	1.0	20.4	11.3	9.1
Yellowtail Dam	20.0	2.8	.9	22.7	13.8	8.9
WYOMING						
Alcova 17NW	11.4	3.9	1.2	19.0	6.4	12.6
Arvada 3N	13.3	2.9	.9	16.7	7.0	9.7
Bates Creek	13.0	2.0	.6	16.2	10.8	5.4
Billy Creek	12.0	3.9	1.2	19.5	7.3	12.2
Buffalo	13.6	3.6	1.1	20.7	9.1	11.6
Casper WSO AP	12.3	2.3	.7	15.3	8.1	7.2
Clearmont 5SW	14.8	2.6	.8	18.7	10.7	8.0
Dead Horse Creek	9.2	1.6	.5	12.9	7.2	5.7
Devils Tower No. 2	17.9	2.3	.7	23.2	15.8	7.4
Dillinger	13.8	2.1	.7	16.7	10.1	6.6
Douglas Aviation	13.0	3.8	1.2	17.2	6.6	10.6
Dull Center 1SE	12.9	3.0	1.0	17.0	8.0	9.0
Echeta 2NW	14.7	2.6	.8	17.4	8.8	8.6
Gillette 2E	16.7	1.7	.5	19.1	14.6	4.5
Gillette 18SW	16.1	2.4	.7	20.1	13.2	6.9
Glenrock 5ESE	14.9	5.1	1.6	21.9	7.0	14.9
Glenrock 14SSE	17.6	2.8	.9	21.9	13.1	8.8
Kaycee	12.7	2.6	.8	16.8	9.0	7.8
Keeline	16.4	5.0	1.6	23.4	8.3	15.1
Lance Creek 3WNW	15.1	3.6	1.1	20.8	8.6	12.2
Lawver 10SW	13.0	2.0	.6	15.0	9.4	5.6
Leiter 9N	15.1	2.6	.8	20.7	11.2	9.5
Lost Cabin	8.1	2.5	.8	12.4	4.8	7.6
Midwest 1SW	14.7	3.4	1.1	19.8	8.5	11.3
Moorcroft	11.1	1.7	.6	13.6	9.0	4.6
Morrisey	12.0	2.4	.7	15.5	9.1	6.4
Newcastle	14.7	2.5	.8	18.2	10.8	7.4
Powder River 2SW	11.8	3.0	1.0	16.3	6.8	9.5
Recluse 14NNW	15.2	2.7	.9	18.9	10.4	8.5
Redbird 1NW	14.0	4.1	1.3	20.4	8.8	11.6
Reno	11.8	3.0	.9	15.0	6.3	8.7
Rochelle 3E	13.0	3.2	1.0	17.1	6.8	10.3
Sheridan Field Station	15.2	3.0	.9	20.9	11.5	9.4
Spencer 10NE	13.3	3.2	1.0	20.2	8.3	11.9
Sundance	17.4	3.1	1.0	21.3	11.7	9.6
Ten Sleep 4NE	12.9	3.2	1.0	18.8	8.6	10.2
Ten Sleep 16SSE	12.5	2.7	.9	18.6	8.7	9.9
Upton	15.1	2.4	.8	17.6	9.3	8.3
Weston 1E	13.0	2.9	.9	19.6	9.6	10.0

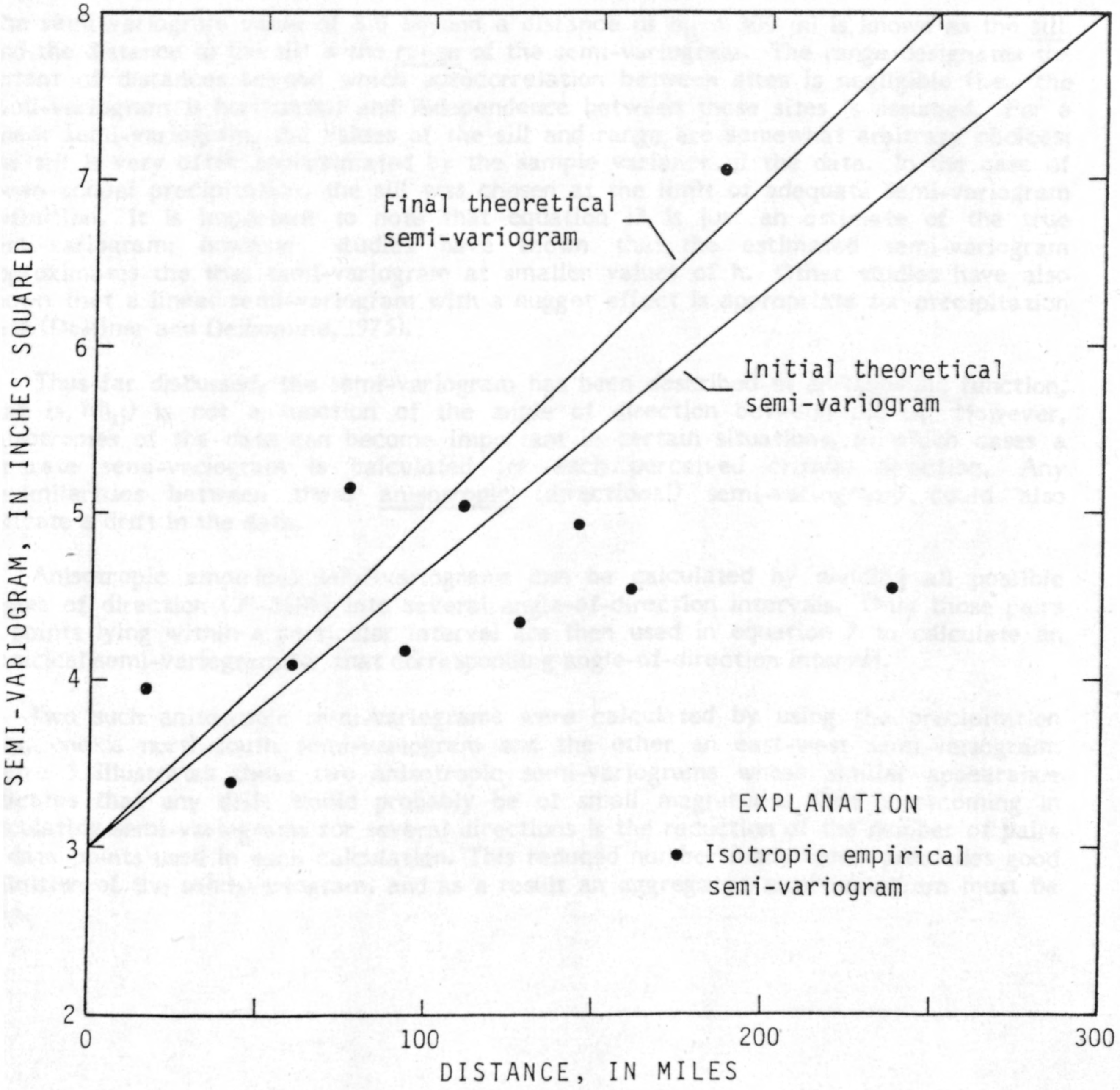


FIGURE 2.--Isotropic empirical semi-variogram of mean annual precipitation and theoretical semi-variograms used in Kriging analyses.

The semi-variogram value of 8.0 beyond a distance of h_{ij} = 300 mi is known as the sill, and the distance to the sill is the range of the semi-variogram. The range designates the extent of distances beyond which autocorrelation between sites is negligible (i.e., the semi-variogram is horizontal) and independence between these sites is assumed. For a linear semi-variogram, the values of the sill and range are somewhat arbitrary choices; the sill is very often approximated by the sample variance of the data. In the case of mean annual precipitation, the sill was chosen at the limit of adequate semi-variogram definition. It is important to note that equation 13 is just an estimate of the true semi-variogram; however, studies have shown that the estimated semi-variogram approximates the true semi-variogram at smaller values of h. Other studies have also shown that a linear semi-variogram with a nugget effect is appropriate for precipitation data (Delfiner and Delhomme, 1975).

Thus far discussed, the semi-variogram has been described as an isotropic function, that is, $\gamma(h_{ij})$ is not a function of the angle of direction between points. However, anisotropies of the data can become important in certain situations, in which cases a separate semi-variogram is calculated for each perceived critical direction. Any dissimilarities between these anisotropic (directional) semi-variograms could also indicate a drift in the data.

Anisotropic empirical semi-variograms can be calculated by dividing all possible angles of direction (0°-360°) into several angle-of-direction intervals. Only those pairs of points lying within a particular interval are then used in equation 7 to calculate an empirical semi-variogram for that corresponding angle-of-direction interval.

Two such anisotropic semi-variograms were calculated by using the precipitation data, one a north-south semi-variogram and the other an east-west semi-variogram. Figure 3 illustrates these two anisotropic semi-variograms whose similar appearance indicates that any drift would probably be of small magnitude. One shortcoming in calculating semi-variograms for several directions is the reduction of the number of pairs of data points used in each calculation. This reduced number many times precludes good definition of the semi-variogram, and as a result an aggregated semi-variogram must be used.

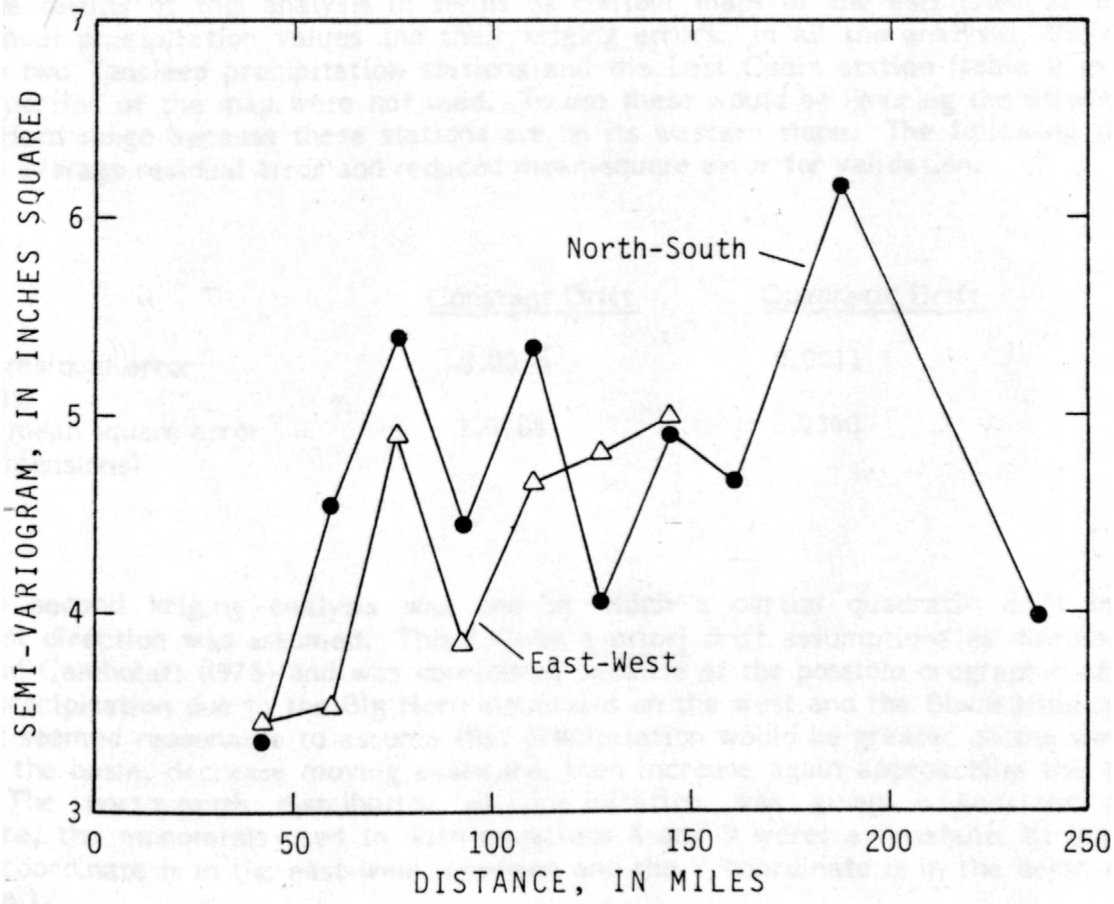


FIGURE 3.--North-South and East-West anisotropic empirical semi-variograms of mean annual precipitation.

Punctual Kriging

Two punctual-kriging analyses were performed in this study. The first was an analysis assuming no drift in a regional 10-year mean annual precipitation. Figure 4 shows the results of this analysis in terms of contour maps of the estimates of these mean annual precipitation values and their kriging errors. In all the analyses, the data from the two Tensleep precipitation stations and the Lost Cabin station (table 1) in the western portion of the map were not used. To use these would be ignoring the effects of the Big Horn range because these stations are on its western slope. The following table gives the average residual error and reduced mean-square error for validation.

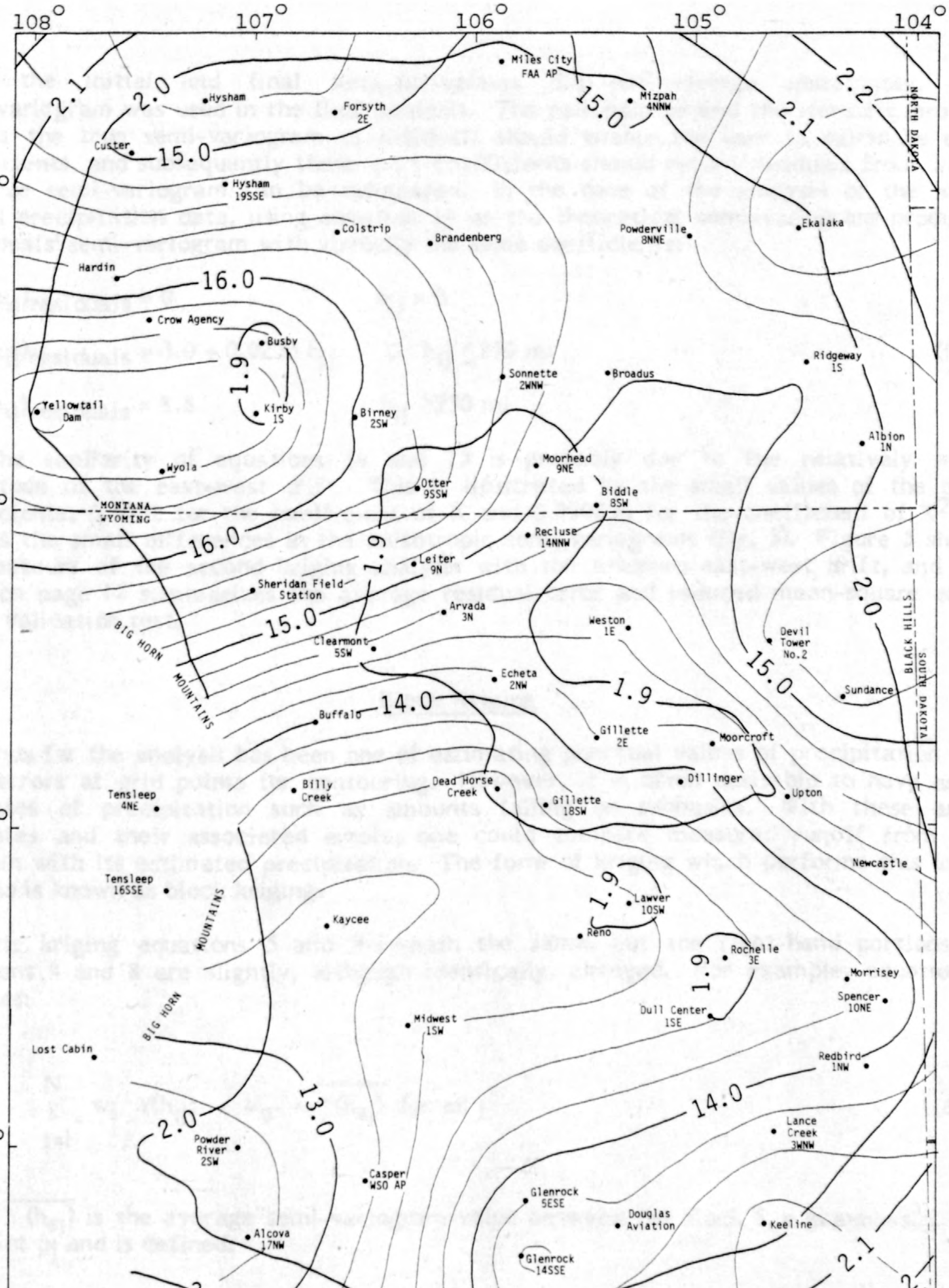
	Constant Drift	Quadratic Drift
Average residual error (inches)	-0.0015	0.0011
Reduced mean square error (no dimensions)	1.0488	1.0340

The second kriging analysis was one in which a partial quadratic drift in the east-west direction was assumed. This follows a priori drift assumptions as discussed in Volpi and Gambolati (1978) and was considered because of the possible orographic effects on the precipitation due to the Big Horn mountains on the west and the Black Hills on the east. It seemed reasonable to assume that precipitation would be greater on the western edge of the basin, decrease moving eastward, then increase again approaching the Black Hills. The north-south distribution of precipitation was given a constant drift. Therefore, the monomials used in with equations 8 and 9 were: a constant; X; and X^2. (The X coordinate is in the east-west direction and the Y coordinate is in the north-south direction.)

The final theoretical semi-variogram (fig. 2) chosen for both kriging analyses, as a result of the validation process, took the following form:

$$
\begin{aligned}
&(h_{ij}) = 0 && h_{ij} = 0 \\
&(h_{ij}) = 3.0 + 0.0200\, h_{ij} && 0 < h_{ij} \leq 250 \text{ mi} \\
&(h_{ij}) = 8.0 && h_{ij} > 250 \text{ mi}
\end{aligned}
\qquad (14)
$$

An assumed drift in the analysis required the use of equation 10 rather than 7 in the semi-variogram estimation. The iteration process consisted of assuming an initial semi-variogram, calculating drift coefficients by using the kriging system (Skrivan and Karlinger, 1979), then determining the semi-variogram from the calculated residuals.



EXPLANATION

•Reno Precipitation station

—— 14.0 —— Line of equal mean annual precictation. Interval 0.2 inch

—— 2.2 —— Line of equal kriging error. Interval 0.1 inch

FIGURE 4.--Kriging mean annual precipitation and associated kriging errors using a constant drift.

When the initial and final semi-variograms did not change appreciably, that semi-variogram was used in the final analysis. The rationale behind the iterative process is that the true semi-variogram of residuals should enable the user to calculate drift coefficients, and subsequently these drift coefficients should return residuals from which the true semi-variogram can be calculated. In the case of the analysis of the mean annual precipitation data, using equation 14 as the theoretical semi-variogram produced a residuals' semi-variogram with virtually the same coefficients:

$$
\begin{aligned}
(h_{ij})_{residuals} &= 0 && h_{ij} = 0 \\
(h_{ij})_{residuals} &= 3.0 + 0.0233\, h_{ij} && 0 < h_{ij} \leq 250 \text{ mi} \\
(h_{ij})_{residuals} &= 8.8 && h_{ij} > 250 \text{ mi}
\end{aligned}
\qquad (15)
$$

The similarity of equations 14 and 15 is probably due to the relatively small magnitude of the east-west drift. This is illustrated by the small values of the drift coefficients, 0.0032 for the coefficient of X and 0.000036 for the coefficient of X^2, as well as the small differences in the anisotropic semi-variograms (fig. 3). Figure 5 shows the contours of the second kriging analysis with the assumed east-west drift, and the table on page 12 summarizes the average residual error and reduced mean-square error of the validation test.

Block Kriging

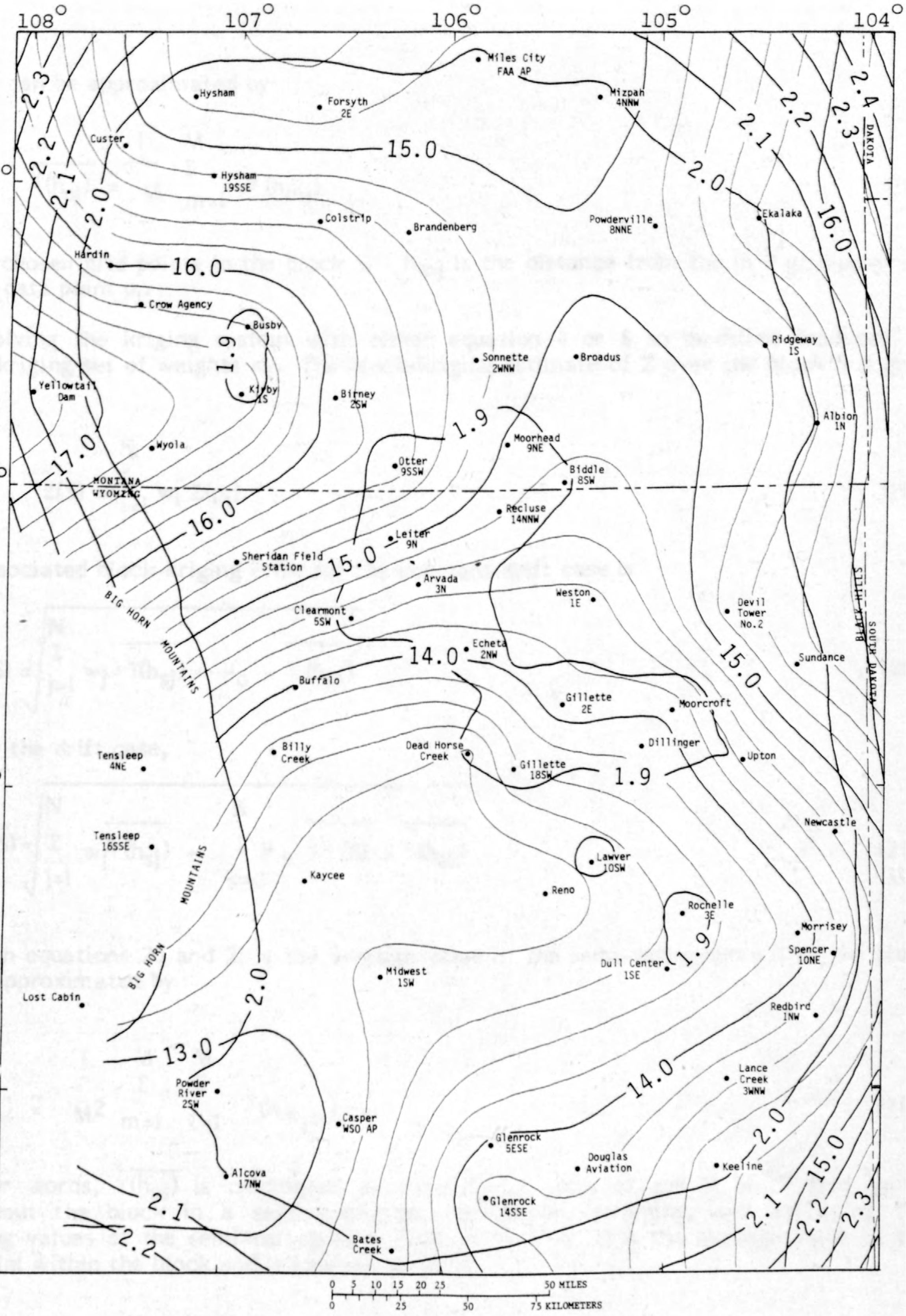
Thus far the analysis has been one of estimating punctual values of precipitation and their errors at grid points for contouring. However, it is often desirable to have areal estimates of precipitation such as amounts falling on subbasins. With these areal estimates and their associated errors, one could compare measured runoff from the subbasin with its estimated precipitation. The form of kriging which performs this areal analysis is known as block kriging.

The kriging equations 5 and 9 remain the same, but the right-hand portions of equations 4 and 8 are slightly, although identically, changed. For example, equation 4 becomes:

$$\sum_{i=1}^{N} w_i \ \gamma(h_{ij}) + \mu_o = \overline{\gamma(h_{sj})} \ \text{ for all } j \qquad (16)$$

where $\overline{\gamma(h_{sj})}$ is the average semi-variogram value between the block S, e.g. subbasin, and the point p_j and is defined:

$$\overline{\gamma(h_{sj})} = \frac{1}{S}\int_S \gamma(h_{sj})dp_s \quad \text{for all points } p_s \text{ in S.} \qquad (17)$$



EXPLANATION

• Reno Precipitation station

—— 14.0 —— Line of equal mean annual precictation. Interval 0.2 inch

—— 2.2 —— Line of equal kriging error. Interval 0.1 inch

FIGURE 5.--Kriged mean annual precipitation and associated kriging errors using a quadratic drift.

$\overline{\gamma(h_{sj})}$ can be approximated by

$$\overline{\gamma(h_{sj})} \cong \frac{1}{M} \sum_{m=i}^{M} \gamma(h_{mj}) \tag{18}$$

for M chosen grid points in the block S. h_{mj} is the distance from the m^{th} grid point in S to the data point p_j.

Solving the kriging system with either equation 4 or 8 so modified produces the block-kriging set of weights w_i. The block-kriging estimate of Z over the block S is given by

$$\hat{Z}(S) = \sum_{i=1}^{N} w_i \, Z(p_i). \tag{19}$$

The associated block-kriging error for the constant-drift case is

$$\hat{\sigma}_k(S) = \sqrt{\sum_{j=1}^{N} w_j \, \overline{\gamma(h_{sj})} + \mu_o - \overline{\gamma(h_{ss})}} \tag{20}$$

and for the drift case,

$$\hat{\sigma}_k(S) = \sqrt{\sum_{j=1}^{N} w_j \overline{\gamma(h_{sj})} + \sum_{k=0}^{K} \mu_k \, \overline{f^k(S)} - \overline{\gamma(h_{ss})}} \tag{21}$$

$\overline{\gamma(h_{ss})}$ in equations 20 and 21 is the average value of the semi-variogram within the block and is approximated by

$$\overline{\gamma(h_{ss})} \cong \frac{1}{M^2} \sum_{m=1}^{M} \sum_{\ell=1}^{M} \gamma(h_{m,\ell}) \ . \tag{22}$$

In other words, $\overline{\gamma(h_{ss})}$ is calculated by rotating all pairs of the M block grid points throughout the block in a semi-variogram calculation, summing, and averaging the resulting values of the semi-variogram. $\overline{f^k(S)}$ in equation 21 is the average value of the monomial within the block and is approximated by

$$\overline{f^k(S)} \cong \frac{1}{M} \sum_{S=1}^{M} f^k(p_s)$$

for grid points p_s in block S.

Consistent with intuition and as a mathematical result is the fact that the block-kriging error has a smaller value than the punctual-kriging error of any corresponding point within the block.

In the analysis of the Powder River basin precipitation, two such block or areal estimates were made. One was for the Belle Fourche River subbasin and another was for the Lance Creek subbasin (fig. 6). The following table lists the results as they compare to subbasin river-discharge estimates taken from Hodson, Pearl, and Druse (1973).

Subbasin	Runoff $(ft^3/s)/mi^2/yr$	Runoff in./yr	Kriged mean annual precipitation in./yr
Belle Fourche	0.013	0.18	13.56
Lance Creek	0.013	0.18	14.20

The comparison shows that only between 1 and 2 percent of the precipitation results as direct runoff, while the remainder is lost as evapotranspiration or infiltration to ground water. This small percentage is reasonable from evapotranspiration considerations of a semiarid climate. No analysis was performed to determine if the soil types in these subbasins are conducive to high infiltration.

Kriging with Data Uncertainty

Kriging with the assumption of no uncertainty of the data assigns the weighting effects of each data point to the desired point or block estimate solely through the influence of the semi-variogram. However, oftentimes one has data of different qualities or reliabilities, and it could be useful to take these differences in data types into account. The accounting for these uncertainties requires only a slight modification to the kriging system.

In this study, the values of the mean annual precipitation have associated standard errors of the mean calculated from 10 years of data at each point. These time-sampling errors can be viewed as the indicators of reliability, or uncertainty, of the spatial data to represent the true values at each site.

These errors, $\varepsilon(p_i)$, are assumed to have a variance σ_i^2 and obey the following requirements (Delhomme, 1978):

$$E(\varepsilon(p_i)) = 0 \qquad \text{for all } i$$

$$\text{Cov}(\varepsilon(p_i), \varepsilon(p_j)) = 0 \quad \text{for } i \neq j$$

$$\text{Cov}(\varepsilon(p_i), Z(p_j)) = 0 \quad \text{for all } i \text{ and } j$$

where Cov $(\cdot)$ is the covariance.

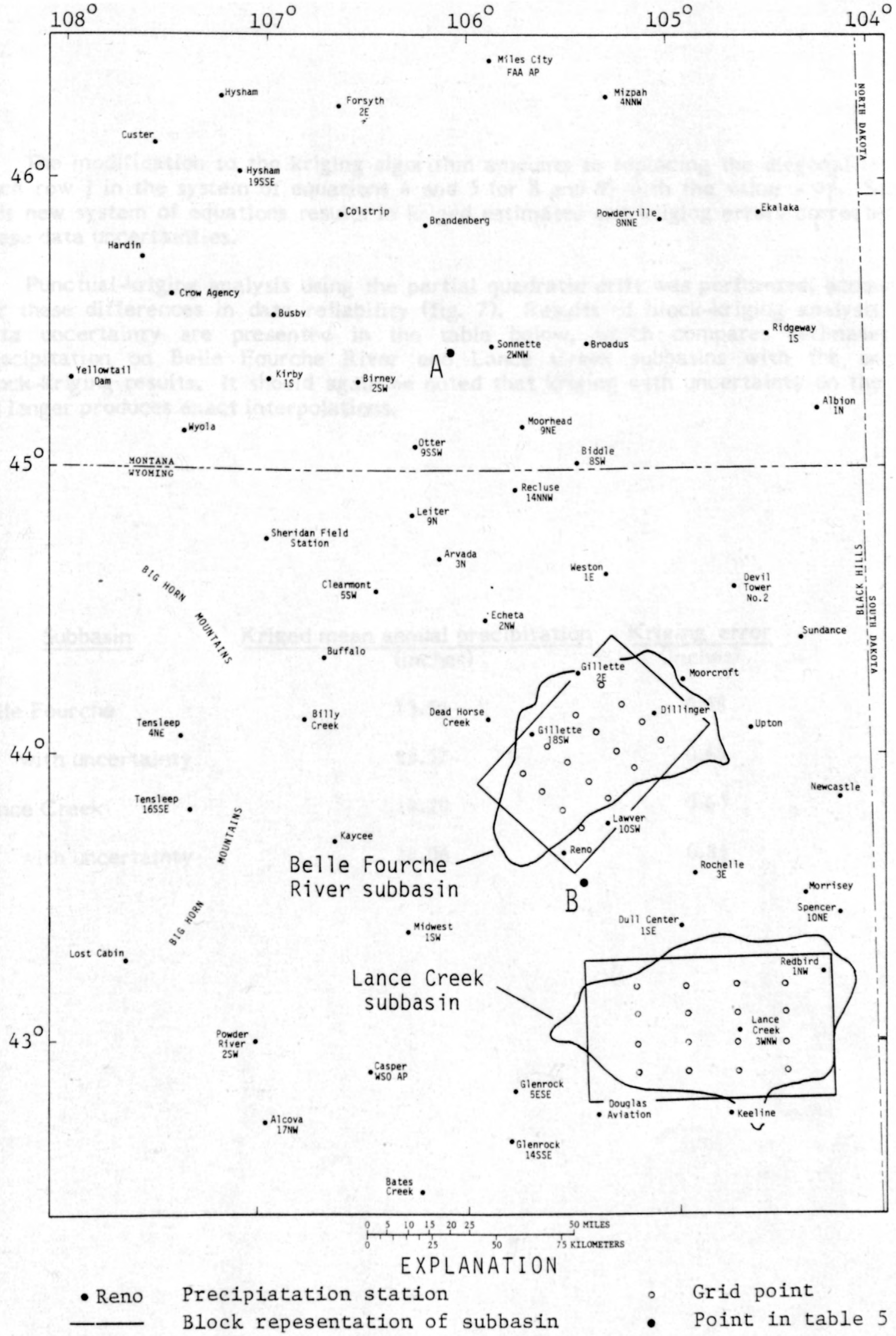
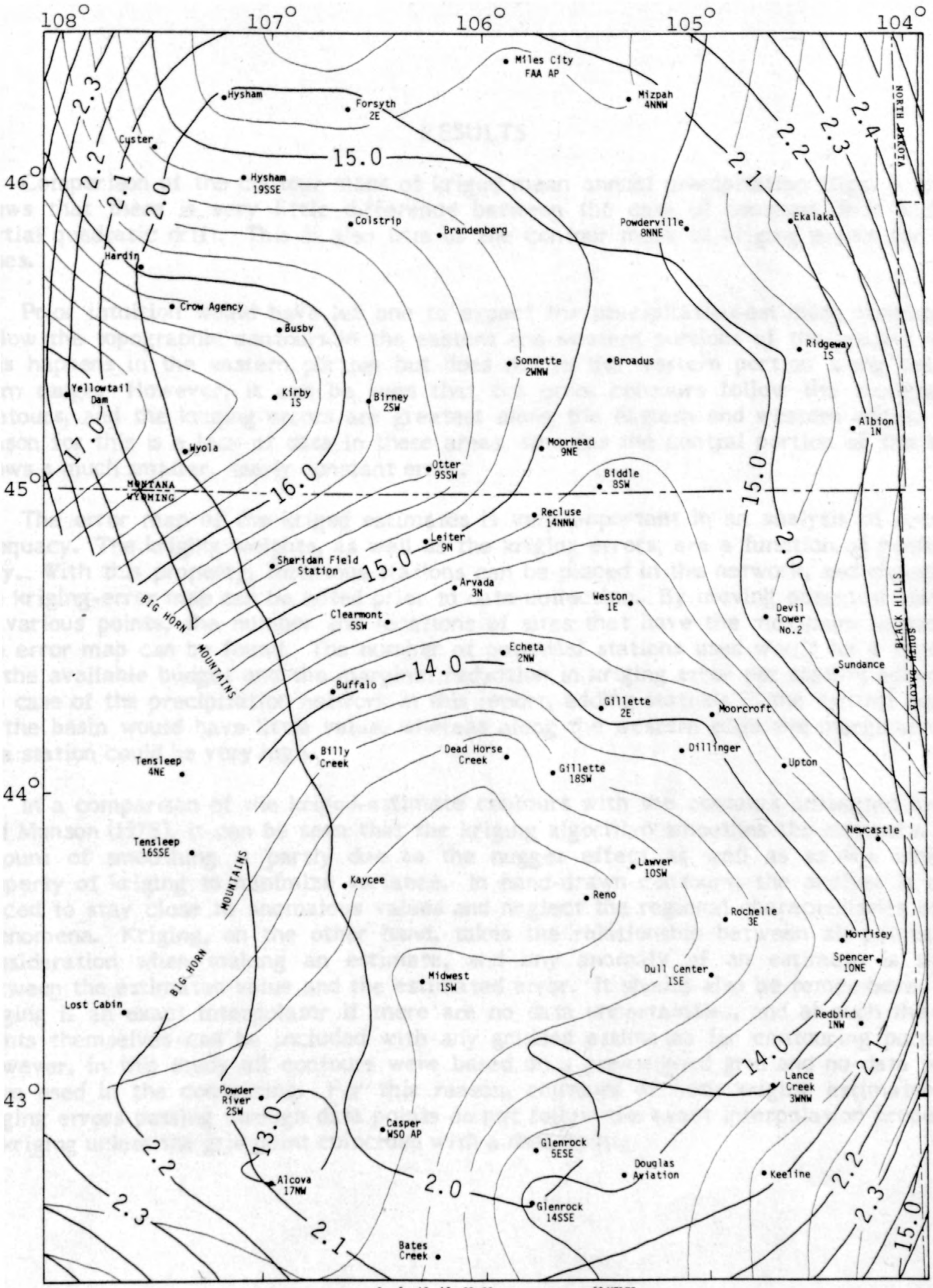


FIGURE 6.--Belle Fourche River and Lance Creek subbasins, with block representatation and location of grid points.

The modification to the kriging algorithm amounts to replacing the diagonal term in each row j in the system of equations 4 and 5 (or 8 and 9) with the value $-\sigma_j^2$. Solving this new system of equations results in kriged estimates and kriging errors corrected for these data uncertainties.

Punctual-kriging analysis using the partial quadratic drift was performed, accounting for these differences in data reliability (fig. 7). Results of block-kriging analysis with data uncertainty are presented in the table below, which compares estimates for precipitation on Belle Fourche River and Lance Creek subbasins with the original block-kriging results. It should again be noted that kriging with uncertainty on the data no longer produces exact interpolations.

Subbasin	Kriged mean annual precipitation (inches)	Kriging error (inches)
Belle Fourche	13.56	0.58
with uncertainty	13.57	0.64
Lance Creek	14.20	0.65
with uncertainty	14.06	0.81



EXPLANATION

• Reno   Precipitation station

—— 14.0 —— Line of equal mean annual precictation. Interval 0.2 inch

—— 2.2 —— Line of equal kriging error. Interval 0.1 inch

FIGURE 7.--Kriged mean annual precipitation and associated kriging errors using a quadratic drift and data uncertainty.

RESULTS

Comparison of the contour maps of kriged mean annual precipitation (figs. 4 and 5) shows that there is very little difference between the case of constant drift and the partial quadratic drift. This is also true of the contour maps of kriging errors for both cases.

Prior intuition would have led one to expect the precipitation-estimate contours to follow the topographic contours in the eastern and western portions of the project area. This happens in the eastern portion but does not in the western portion along the Big Horn range. However, it can be seen that the error contours follow the topographic contours, and the kriging errors are greatest along the eastern and western edges. One reason for this is a lack of data in these areas, whereas the central portion of the basin shows a much smaller, nearly constant error.

The error map of the kriged estimates is very important in an analysis of network adequacy. The kriging weights, as well as the kriging errors, are a function of geometry only. With this property, fictitious stations can be placed in the network, and changes in the kriging-error map can be noted prior to data collection. By moving potential stations to various points, the number and locations of sites that have the maximum impact on the error map can be found. The number of potential stations used would be a function of the available budget and the marginal reduction in kriging error per station added. In the case of the precipitation network in this report, adding stations in the central portion of the basin would have little value, whereas along the western edge the marginal value of a station could be very high.

In a comparison of the kriged-estimate contours with the contours presented by Toy and Munson (1978), it can be seen that the kriging algorithm smoothes the contours. The amount of smoothing is partly due to the nugget effect as well as to the inherent property of kriging to minimize variance. In hand-drawn contours, the analyst is often forced to stay close to anomalous values and neglect the regional characteristics of the phenomena. Kriging, on the other hand, takes the relationship between all points into consideration when making an estimate, and any anomaly of an estimate is shared between the estimated value and the estimated error. It should also be remembered that kriging is an exact interpolator if there are no data uncertainties, and as such the data points themselves can be included with any gridded estimates for contouring purposes. However, in this study all contours were based on a preassigned grid and no data points were used in the contouring. For this reason, contours of both kriging estimates and kriging errors passing through data points do not follow the exact interpolation properties of kriging unless the grid point coincided with a data point.

One consideration which has not yet been discussed is the number of data points one should use in kriging. The dimensions of the matrix resulting from the kriging system are directly dependent on the number of data values used, and, therefore, in addition to mathematical and structural considerations, there is the computational-cost consideration. Perhaps the best way to choose the number of points is to consider the characteristics of the semi-variogram, for example, the range as compared to the scatter of the data points and the purpose of performing the analysis. In this study, all the data points were used because the semi-variogram did not warrant reducing the number nor did computer costs prohibit using all the points. As a check, however, mean annual precipitation at two points (see fig. 6 for their locations) were estimated by using the 15 nearest stations and compared to those estimated by utilizing all 60 stations. These comparisons, which show small differences, are given in the following table.

	Kriged mean annual precipitation (inches)		Kriging error (inches)	
	Observations		Observations	
point*	60	15	60	15
A	15.51	15.51	1.92	1.98
B	13.15	13.35	1.92	1.97

* See fig. 6 for point locations.

RECENT DEVELOPMENTS IN KRIGING

Kriging and universal kriging have been used in estimation problems for over 15 years and have been fairly successful in their applications. However, there has always been a difficulty in reducing bias of the empirical semi-variogram. This has been especially true of universal kriging where the semi-variogram of the residuals off the drift had to be estimated, as discussed previously.

In recent years, however, Matheron (1971) and others from the School of Geostatistics in Fontainebleau, France, have developed a theory of intrinsic random functions of order k. A description of the theory is beyond the scope of this paper, but the reader is referred to Delfiner (1976) for a practical development. The main thrust of the theory is, however, that the kriging system can be modified to generate generalized increments that filter out polynomial drifts in the estimation procedure. These increments allow the estimation procedure to be completed by using a generalized covariance in place of a semi-variogram. The validation procedure of reducing the average residual error as much as possible and bringing the reduced mean-square error close to 1 is the same as already discussed.

The advantage of such a system when a drift is present in the data is that it is conducive to automatically fitting the drift and generalized covariance to a particular problem. A disadvantage of the system as it is now presented in the literature is that it can be used only for complete polynomials describing the drift (the concept of drift is not used in the new theory but is acceptable for purposes of this section). Hence, the case of an a priori-drift polynomial presented in this report would not be acceptable in the new theory.

An analysis was made, however, by applying this new theory to the precipitation data with no a priori assumptions on the drift. It was found that the best results were obtained when a constant drift was used rather than a complete linear-polynomial or complete quadratic-polynomial drift. However, in the case of a constant drift, the new theory reduces to the case of using a constant drift and linear semi-variogram, which was the situation originally presented. The semi-variogram of the new theory agreed with the linear one chosen.

SUMMARY

Kriging can be a valuable tool in many regionalization problems and interpolations of spatial data that are correlated. Other methods of estimation that neglect the correlation between the data are sometimes poor choices because they do not utilize all the information contained in the data. Kriging attempts to do this through the properties of the semi-variogram or generalized covariance. The semi-variogram estimated from the data is sometimes strongly biased, but usually a satisfactory one can be found based on the validation process. The validation process is the ultimate method of choosing a correct semi-variogram-drift set in any kriging system, even for the automatically fitted models of the recent theory of generalized covariance.

A great advantage that kriging has over other regionalization techniques, such as least-squares regression, is that it provides an independent measure of error for each estimate. The kriging-error map is very important for evaluating the adequacy of a data network, and the kriging system enables one to determine the effects of additional measurements on the error map prior to any actual changes.

The analyses performed in this study on the precipitation network in the Powder River basin are a demonstration of the flexibility of the kriging system. The size of the area and extent of the data define the limitations to which the results are applicable. Thus, the present study is useful primarily as a macroscopic regional analysis.

Two kriging analyses, based on data at 60 sites, were used to calculate contour maps of mean annual precipitation — one analysis assumed a constant drift of precipitation in the basin and the other assumed a partial quadratic drift to simulate orographic effects. The calculated contours of precipitation and their associated kriging errors for these two analyses were nearly identical, as were the validation results.

Data uncertainty, as measured by standard errors of the mean and based on a period of 10 years of data, was also included in the kriging analysis. These estimates were similar to those from the two previous analyses.

Kriging can also be used to estimate block, or areal, values of the regionalized variable. Mean annual precipitation for two subbasins, Belle Fourche River and Lance Creek, were so estimated. The estimated runoff was 1-2 percent of the block-kriging estimates.

REFERENCES

Delfiner, P., 1976, Linear estimation of nonstationary spatial phenomena, in NATO Advanced Study Institute on advanced geostatistics in the mining industry, Rome, Italy, 1975: Boston, D. Reidel Publishing Co., p. 49-68.

Delfiner, P., and Delhomme, J. P., 1975, Optimum interpolation by kriging, in NATO Advanced Study Institute on display and analysis of spatial data, Nottingham, England, 1973: New York, John Wiley and Sons, p. 96-114.

Delhomme, J. P., 1978, Kriging in the hydrosciences: Advances in Water Resources, v. 1, no. 5, p. 251-266.

Gambolati, G., and Volpi, G., 1979, Groundwater contour mapping in Venice by stochastic interpolators 1. theory: Water Resources Research, v. 15, no. 2, p. 281-290.

Hodson, W. G., Pearl, R. H., and Druse, S. A., 1973, Water resources of the Powder River basin and adjacent areas, northeastern Wyoming: U.S. Geological Survey Atlas HA-465.

Matheron, G., 1971, The theory of regionalized variables and its application, Les Cahiers du Centre de morphologie mathematique No. 5: Fontainebleau, France, 211 p.

Olea, R. A., 1975, Optimum mapping techniques using regionalized variable theory: Kansas Geological Survey, 137 p.

Skrivan, J. A., and Karlinger, M. R., 1980, Semi-variogram estimation and universal kriging program: U.S. Geological Survey Computer Contribution, 98 p.

Taylor, A. E., 1955, Advanced Calculus: Boston, Ginn and Company, 786 p.

Toy, T. J., and Munson, B. E., 1978, Climate appraisal maps of the rehabilitation potential of strippable coal lands in the Powder River basin, Wyoming and Montana: U.S. Geological Survey Map MF-932, 2 sheets.

Volpi, G., and Gambolati, G., 1978, On the use of a main trend for the kriging ing technique in hydrology: Advances in Water Resources, v. 1, no. 6, p. 345-349.

☆GPO 1981 797-044