Evaluation of Precipitation Network in Snowy Range Observatory

Joongkyu Choi Yeou-Koung Tung

1990
Final Report WWRC-90-34

Final Report

Submitted to

Wyoming Water Research Center
University of Wyoming
Laramie, Wyoming

Submitted by

Joongkyu Choi
Yeou-Koung Tung
Wyoming Water Research Center
University of Wyoming
Laramie, Wyoming

ABSTRACT

This research studies the monthly precipitation information content in the Snowy Range Observatory. The Observatory currently consists of 20 precipitation gages with varying length of record. The study was proceeded through three phases.

Phase I - Estimation of Missing Values Due to various reasons, records of all gages involves missing values. The study was first performed to compare various methods of different complexity to estimate the missing values. Although there was no single method that is universally superior in all circumstances, a simple method of linear inverse distance method was found to be rather accurate. It was then used to fill the missing values in all precipitation gages.

Phase II - Analysis of Monthly Precipitation The time series of monthly total precipitation was first analyzed station by station. The accuracy of different methods of estimating the spatial distribution of the average precipitation were investigated. Then, the spatial structures of the monthly total precipitation were identified using variogram analysis of geostatistics. Information derived from this phase serves as the basis for precipitation network analysis in the next phase.

Phase III - Precipitation Network Analysis The objective of this phase is to examine the effect of having a reduced network, in terms of gage number, on the loss of information content. The present network containing 21 gages was used as the basis for comparison. Two reduced networks each containing 15 and 12 gages, respectively, were subjectively selected on the basis of geographical location, accessibility, and aesthetic considerations. Nonstationary Kriging technique was employed to estimate the spatial distribution of error based on different network configurations. It was found that the two reduced networks have a small increase in error, as compared with 21 stations, during the months of May - September. However, increase in error could be as high as 25%-30% during the months of October - April. High error occurs on the upper third of the watershed resulting from the removal of gages from the area due to accessibility consideration.

ACKNOWLEDGMENTS

The authors wish to thank Steve Gloss, Director of Wyoming Water Research Center, for his support of the study. Gratitude is also extended to Chris Goertler and Tom Wesche for sharing their intimate knowledge on the Snowy Range Observatory. This report is largely based on the M.S. thesis of Mr. Joongkyu Choi for the Department of Statistics, University of Wyoming, Laramie, Wyoming.

CHAPTER I

INTRODUCTION

The Snowy Range Observatory has been maintained since the mid-1960's. Great environmental diversity is found within the relatively small geographical area. The Snowy Range, thus, has been utilized as a study for many research projects. Centered in the Nash Fork Creek drainage of the Medicine Bow National Forest in the upper North Platte River basin (Figure I.1), the Observatory consists of a network of 21 recording precipitation gaging stations and other stations measuring streamflow, humidity, temperature and wind. The list of the stations is presented on Table I.1.

From 1972 to 1988, precipitation data were collected and the various types of the gaging instruments were increased to suit the needs of the research being conducted. During the same period, data collection was interrupted in 1974-1975 which was attributed to the 'energy-boom' in Wyoming resulting in a shift of research focus from high-mountain watershed projects to environmental assessment.

Two primary objectives in this research: (1) to evaluate the existing precipitation gage network in Snowy Range Observatory and (2) to determine the reduced network which retains the maximum amount of precipitation information subject to geographical and strategical constraints.

This research analyzes the monthly precipitation data that have been stored in Water Resources Data System (WRDS) maintained by the Wyoming Water Research Center. Due to unexpected interruption in data collection system, the study first estimates missing values occurring in each station. Before missing values were estimated, various techniques were applied and their performances were examined. Statistical analysis of monthly precipitation data, which is nonstationarity, was performed which served as the basis for formulating the optimal reduced precipitation network model.

The report is organized as the following: Chapter II discusses the estimation of missing values by different methods and compares their performance. In Chapter III, procedures are described and applied to transform a nonstationary monthly precipitation time series to a stationary one. Chapter IV considers the spatial distribution of the monthly precipitation by two types of method: (1) the spatially weighted average using "Inverse Distance" and "Gaussian Smoothing" methods without considering spatial correlation, and (2) 'non-stationary kriging' considering the spatial correlation of monthly precipitation. The performance of these methods were compared on the basis of contour maps for the estimated means and the associated errors. Chapter V presents the model formulation for the optimum precipitation network design using zero-one mixed integer programming. The summary of final results and some recommendations are presented in Chapter VI. The contour maps resulting

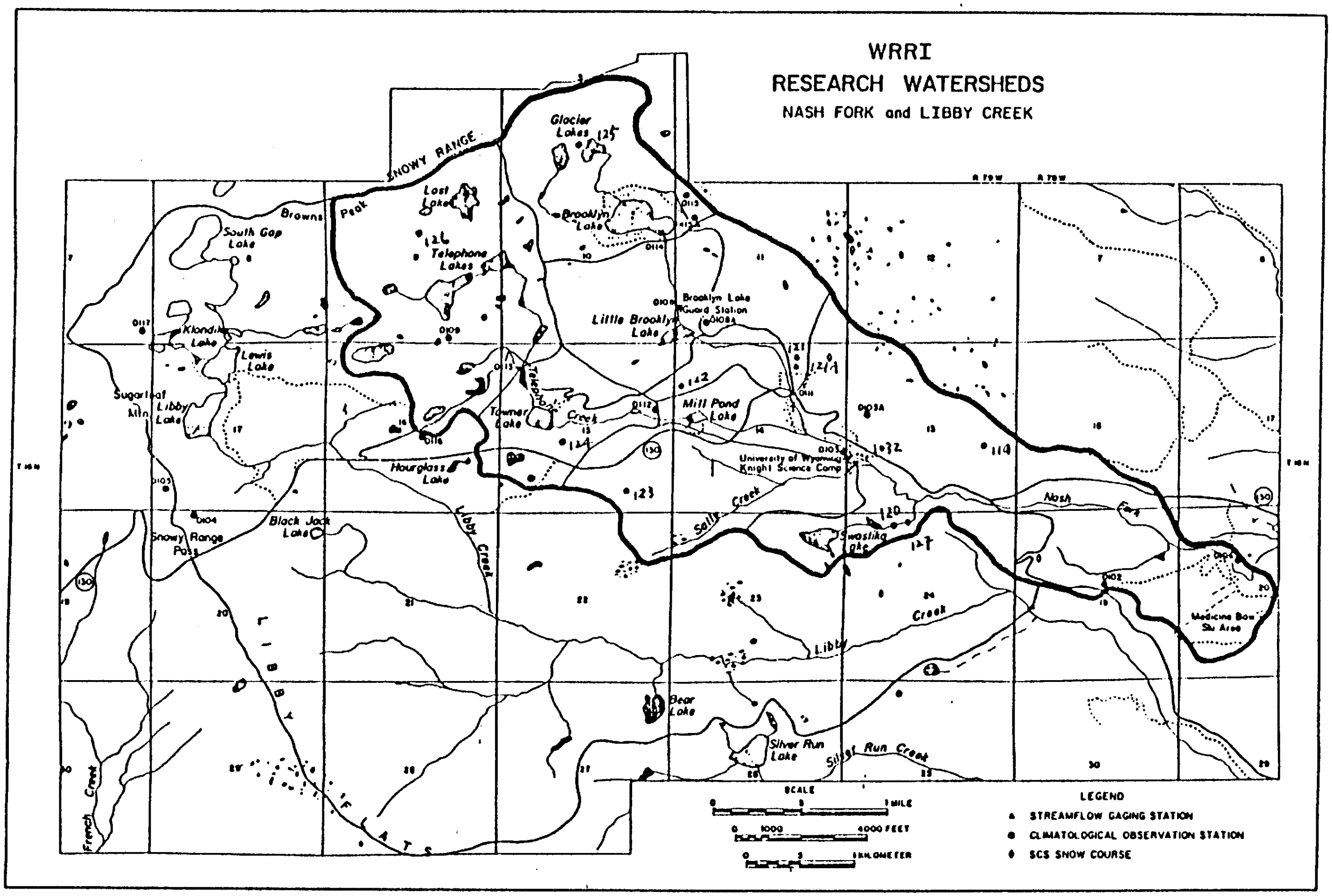


Figure I.1 : Snowy Range Observatory (Wesche,1982)

Station ID	Location			Number of years of record
	Latitude	Longitude	Elevation	
0101	41-18-00	106-09-00	8440.00	12 yrs
0102	41-20-00	106-12-00	9400.00	12 yrs
0103-1	41-21-00	106-14-00	9940.00	14 yrs
0103-2	41-21-00	106-14-00	9940.00	14 yrs
0103-A	41-21-00	106-13-00	10060.00	14 yrs
0106	41-20-00	106-11-00	9100.00	12 yrs
0108-2	41-22-00	106-15-00	10360.00	14 yrs
0108-A	41-22-00	106-15-00	10360.00	14 yrs
0109	41-22-00	106-16-00	10740.00	12 yrs
0115-2	41-22-00	106-15-00	10640.00	14 yrs
0115-A	41-22-00	106-15-00	10560.00	14 yrs
0119	41-21-00	106-13-00	9880.00	12 yrs
0120	41-21-00	106-13-00	9960.00	12 yrs
0121	41-22-00	106-14-00	10320.00	12 yrs
0121-A	41-21-49	106-13-50	10320.00	07 yrs
0122	41-21-00	106-15-00	10380.00	12 yrs
0123	41-21-00	106-15-00	10380.00	12 yrs
0124	41-21-00	106-15-00	10440.00	12 yrs
0125	41-23-00	106-15-00	10800.00	12 yrs
0126	41-22-00	106-16-00	11020.00	12 yrs
0127	41-21-00	106-13-00	9840.00	12 yrs

Table I.1 : List of Precipitation Stations in Snowy Range Observatory (Wesche,1982) from Chapter IV and a FORTRAN program developed for the analysis are given in Appendices A and B, respectively.

CHAPTER II

ESTIMATION OF MISSING DATA

II.1 *Introduction*

The precipitation data from Snowy Range Observatory had been analyzed in this study starting from June 1962 to June 1988 (a total of 169 months). There are, however, observations missing in the data set. When a designed analysis is spoiled by missing data there are basically two ways to perform the analysis. One is to analyze the observed incomplete data set. Alternatively, an approach can be applied to estimate the missing values and then to analyze the 'complete' precipitation data with these estimated values inserted. The first approach is undesirable for this study, in particular, because the missing data occurred very irregularly through the recording period in all existing stations. For example, two monthly precipitation time series containing missing data from stations 106 and 108-2, are shown in Figures II.1 and II.2.

Simultaneous time-space data are required at all stations to characterize the temporal and spatial correlation structures of the precipitations. Only 51 months throughout the entire 169 months were recorded concurrently at the Snowy Range watershed. Thus, ignoring the missing values results in losing about two thirds of the observations at all precipitation gages (on the average, each individual station out of all 21 stations has about 150 observations). For this reason, it is undesirable to conduct this study without estimating the missing values.

In the first attempt, the Box-Jenkins univariate forecasting method (Markridakis et al., 1983 and Vandaele, 1983) was used to analyze temporal correlation of monthly precipitation for each individual station. This is because the missing values occurred irregularly making it difficult to use multivariate time series analysis which requires the observations to occur concurrently.

In univariate time series analysis, the first step is to identify if the time series has a specific ARIMA structure. For this, the autocorrelation function is obtained by SAS/ETS (SAS, 1984) for each individual station. Sample results for stations 102 and 119 are shown in Figures II.3 and II.4. By using the Ljung-Box test (Bowerman et al., 1987 and Markridakis et al., 1983). The auto-correlations were not significant because χ^2 with 30 degrees of freedom is 43.77 at 5% significant level which is large than the statistic value of 39.56 for station 102 and 38.78 for station 119. This implies that the time series of monthly precipitation data for stations 102 and 119 are random with no significant temporal correlation structure. This situation occurs throughout all the remaining 19 precipitation stations in the Observatory. Since no specific ARIMA models could be identified for monthly precipitation data, the use of the Box-Jenkins forecasting method for estimating missing values was abandoned.

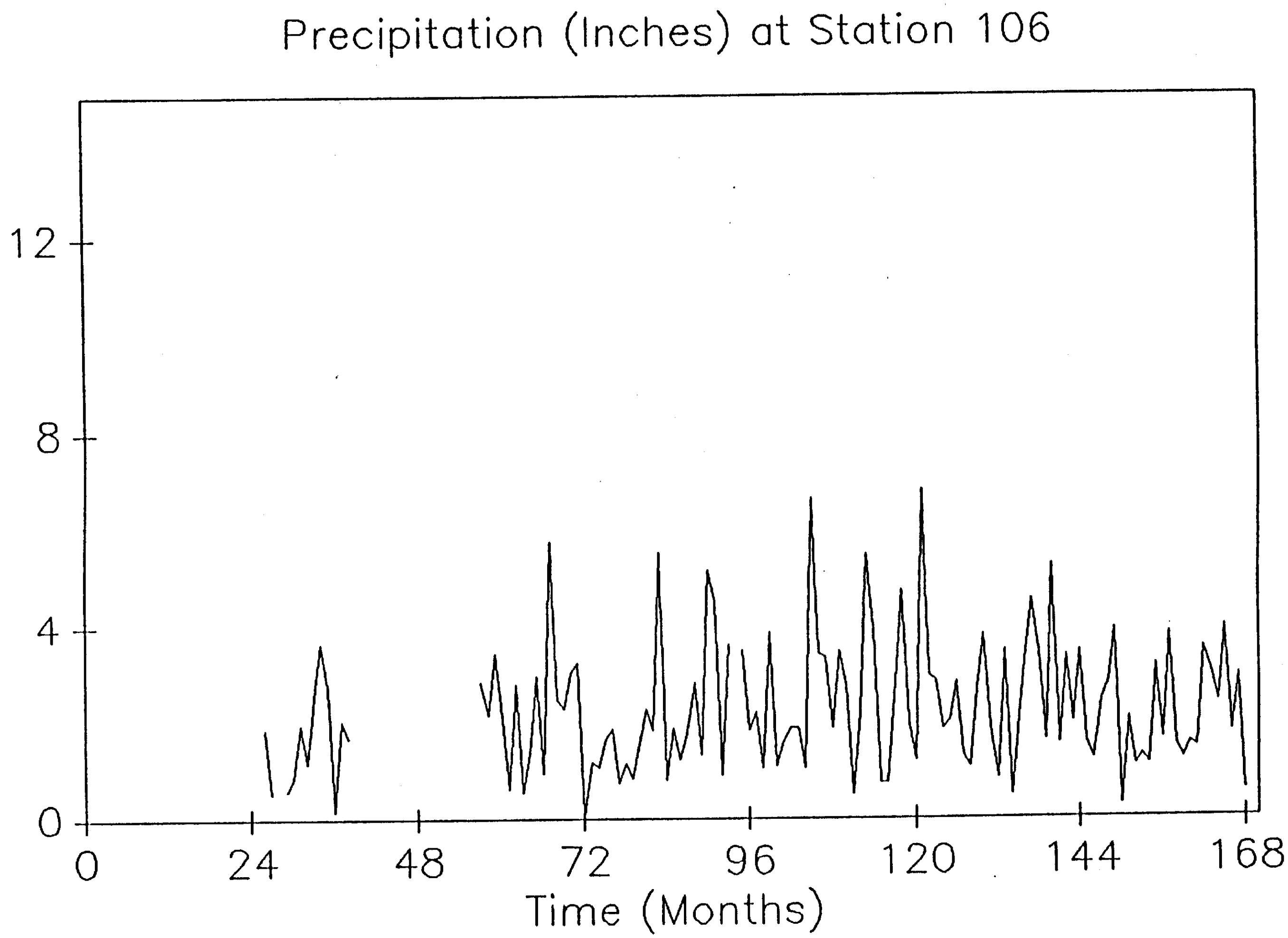


Figure II.1: Time Series of Monthly Precipitation

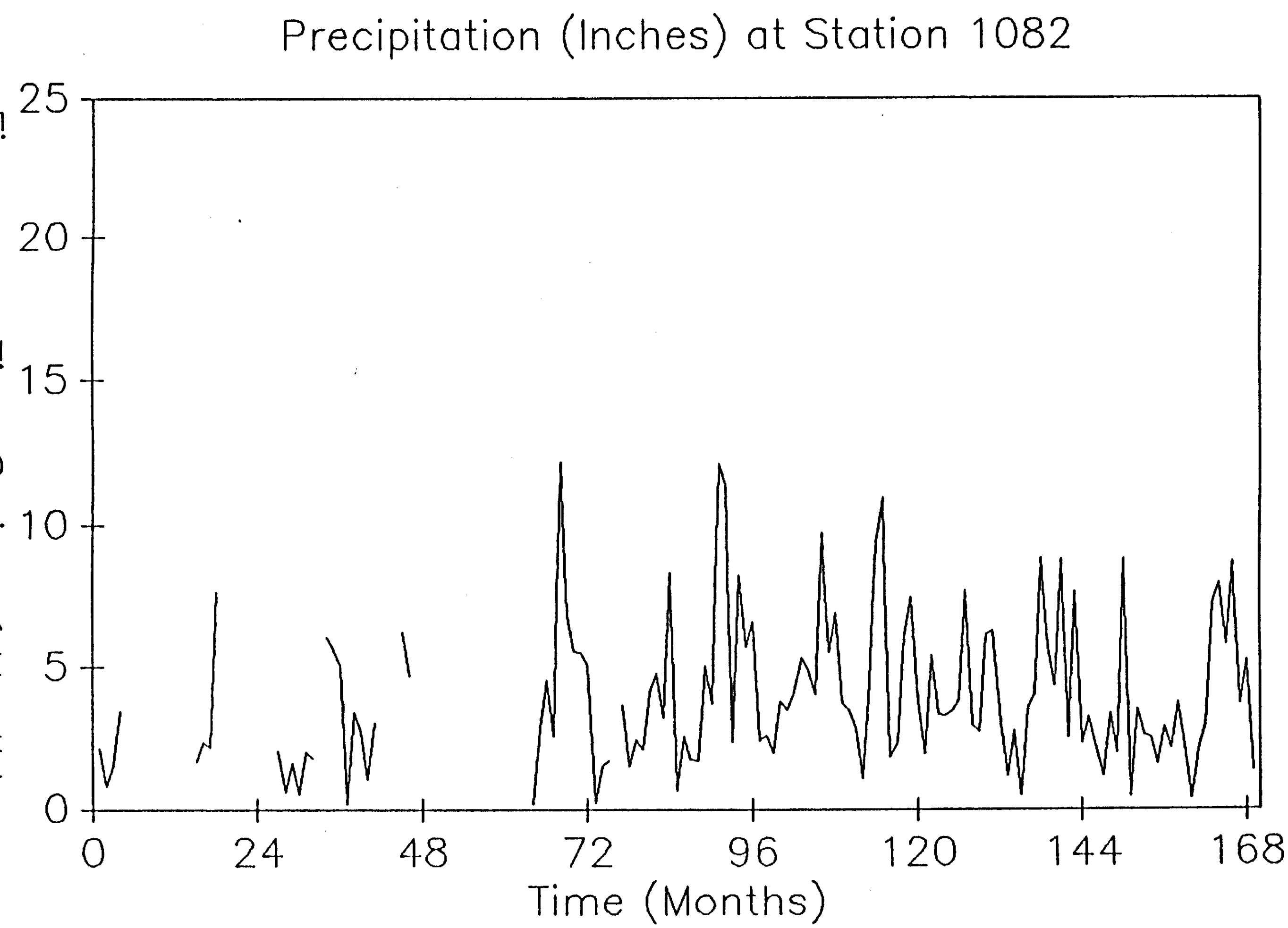


Figure II.2: Time Series of Monthly Precipitation

ARIMA PROCEDURE (STATION 102)

AUTOCORRELATIONS

```
LAG  COVARIANCE  -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
  0     2.57152  |                    |********************|
  1    0.053595  |                .   |   .                |
  2   0.0110287  |                .   |   .                |
  3   -0.083829  |                .  *|   .                |
  4    0.026479  |                .   |   .                |
  5   0.0395429  |                .   |   .                |
  6    -0.42218  |                .***|   .                |
  7    0.168414  |                .   |*  .                |
  8    0.375776  |                .   |***.                |
  9    -0.11497  |                .  *|   .                |
 10    -0.43619  |                .***|   .                |
 11   -0.196167  |                . **|   .                |
 12    0.263037  |                .   |** .                |
 13   -0.246032  |                . **|   .                |
 14   0.0974338  |                .   |*  .                |
 15   -0.162852  |                .  *|   .                |
 16    0.428415  |                .   |***.                |
 17   -0.233829  |                . **|   .                |
 18   -0.102615  |                .  *|   .                |
 19    0.201029  |                .   |** .                |
 20    -0.19482  |                . **|   .                |
 21   -0.396177  |                .***|   .                |
 22    0.188901  |               .    |*   .               |
 23    0.488366  |               .    |****.               |
 24     0.35209  |               .    |*** .               |
 25  -0.0557214  |               .    |    .               |
 26   0.0913949  |               .    |*   .               |
 27   0.0177282  |               .    |    .               |
 28    0.160982  |               .    |*   .               |
 29   -0.490346  |               .****|    .               |
 30    0.249222  |               .    |**  .               |
```

AUTOCORRELATION CHECK FOR WHITE NOISE

TO LAG	CHI SQUARE	DF	PROB	AUTOCORRELATIONS					
6	3.32	6	0.768	0.021	0.004	-0.033	0.010	0.015	-0.164
12	12.03	12	0.444	0.065	0.146	-0.045	-0.170	-0.076	0.102
18	18.69	18	0.411	-0.096	0.038	-0.063	0.167	-0.091	-0.040
24	31.87	24	0.130	0.078	-0.076	-0.154	0.073	0.190	0.137
30	39.56	30	0.114	-0.022	0.036	0.007	0.063	-0.191	0.097

Figure II.3 : Autocorrelations at station 102

```
                    ARIMA PROCEDURE (STATION 119)
                          AUTOCORRELATIONS
LAG COVARIANCE -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
  0    2.73073 |                    |********************|
  1 -0.0133301 |                .   |   .                |
  2  -0.249693 |                . **|   .                |
  3  0.0658215 |                .   |   .                |
  4 -0.0275677 |                .   |   .                |
  5 -0.0601951 |                .   |   .                |
  6  -0.363533 |                .***|   .                |
  7 -0.0297245 |               .    |    .               |
  8   0.433612 |               .    |*** .               |
  9  -0.124588 |               .   *|    .               |
 10  -0.206754 |               .  **|    .               |
 11   -0.18758 |               .   *|    .               |
 12   0.287082 |               .    |**  .               |
 13  -0.186568 |               .   *|    .               |
 14    0.19402 |               .    |*   .               |
 15  -0.121405 |               .   *|    .               |
 16   0.324022 |               .    |**  .               |
 17  -0.155573 |               .   *|    .               |
 18  -0.248737 |               .  **|    .               |
 19   0.100324 |               .    |*   .               |
 20  -0.210515 |               .  **|    .               |
 21  -0.515838 |               .****|    .               |
 22  0.0325981 |               .    |    .               |
 23   0.654807 |               .    |*****               |
 24   0.634183 |               .    |*****               |
 25  -0.308912 |               .  **|    .               |
 26  -0.133387 |              .    *|     .              |
 27  -0.257144 |              .   **|     .              |
 28   0.177734 |              .     |*    .              |
 29  -0.589677 |              . ****|     .              |
 30   0.127712 |              .     |*    .              |
```

AUTOCORRELATION CHECK FOR WHITE NOISE

TO LAG	CHI SQUARE	DF	PROB	AUTOCORRELATIONS					
6	2.43	6	0.876	-0.005	-0.091	0.024	-0.010	-0.022	-0.133
12	7.08	12	0.852	-0.011	0.159	-0.046	-0.076	-0.069	0.105
18	10.97	18	0.896	-0.068	0.071	-0.044	0.119	-0.057	-0.091
24	28.94	24	0.222	0.037	-0.077	-0.189	0.012	0.240	0.232
30	38.78	30	0.131	-0.113	-0.049	-0.094	0.065	-0.216	0.043

Figure II.4 :Autocorrelations at station 119

Alternatively, several estimation methods considering spatial correlation were used to estimate the missing observations. The missing data at a station can be estimated by the weighted average of the observed precipitations from the surrounding stations as

$$\hat{X}(0,t) = \sum_{s=1}^{S} w_s X(s,t) \tag{II.1}$$

where $\hat{X}(0,t)$ is the estimated precipitation amount for a station with missing data at time t, X(s,t) is the observed precipitation from a surrounding station s at time t, w_s is the weight for station s, and S is the total number of surrounding stations used in estimation.

Three types of method are used to in this study estimate the missing values: (1) inverse distance weighing technique, (2) nonlinear programming technique to minimize the variance of the estimates, and (3) regression technique considering the cross-correlations between the station with missing values and the surrounding stations with observations. The resulting regression statistics such as variance, standard error, and R^2 were also used to define the weights for estimating the missing values.

II.2 Estimation Methods

Since the estimates are a weighted linear combination of the observations from the surrounding stations, the results of estimation depend on how the weighing factors are calculated.

II.2.1 Inverse Distance Weighing Method

The inverse distance weighing method considers the premise that weight contributed from a station with observations to the estimated precipitation amount at the station with missing values is inversely proportional to the physical distance between the two stations. The contributing weight for the station s with observation to the station with missing values, w_s, is computed by

$$w_s = \frac{\left(\frac{1}{D_s^n}\right)}{\sum_{s=1}^{S}\left(\frac{1}{D_s^n}\right)} \tag{II.2}$$

where D_s is the distance between the station s with observations and the station with missing values, n is a constant by which the distance is weighted, S is the number of surrounding stations used (Tung, 1983) in the estimation. If the value of α is very large, the weights will be concentrated on a few stations that are very close to the station with missing values. On the other hand, the weight may be dispersed to a large number of the surrounding stations for a small α. For instance, if α equals zero, then every surrounding station under consideration would have an equal weight.

In this study, the values α=1 and 2 are used and they are called the linear inverse distance (IDLIN) method and square inverse distance (IDSQ) method, respectively. For these two methods, the IDSQ technique gives higher weight for the station closer to the point of estimation than the IDLIN method. The amount of monthly precipitation at the station with missing values can be determined by the IDLIN method as

$$\hat{X}(0,t) = \frac{\sum_{s=1}^{S}\left(\frac{X(s,t)}{D_s}\right)}{\sum_{s=1}^{S}\left(\frac{1}{D_s}\right)} \qquad \text{(II.3)}$$

and by the IDSQ method as

$$\hat{X}(0,t) = \frac{\sum_{s=1}^{S}\left(\frac{X(s,t)}{D_s^2}\right)}{\sum_{s=1}^{S}\left(\frac{1}{D_s^2}\right)} \qquad \text{(II.4)}$$

II.2.2 <u>Optimal Weighing Method</u>

Referring to Eq.(II.1), $\hat{X}(0,t)$ is a random variable because it is a linear function of random observations of the surrounding stations. Its degree of uncertainty, represented by the variance, can be computed as

$$\sigma^2_{0,t} = \mathrm{E}\{\,[\hat{X}(0,t)-\mathrm{E}(\hat{X}(0,t))]^2\} = \mathrm{E}\{\,[\hat{X}(0,t)-\sum_{s=1}^{S} w_s\,\mathrm{E}(X(s,t))]^2\}$$

$$= \mathrm{E}\{[\hat{X}(0,t)-\sum_{s=1}^{S} w_s\mu(s,t)]^2\} = \mathrm{E}\{[\hat{X}(0,t)-\mathbf{w}^T\mu(t)]^2\}$$

$$= \mathrm{E}\{\,[\mathbf{w}^T\mathbf{X}(t)-\mathbf{w}^T\mu(t)]^2\,\} = \mathrm{E}\{\,\mathbf{w}^T[\mathbf{X}(t)-\mu(t)]^T[\mathbf{X}(t)-\mu(t)]\,\mathbf{w}\,\}$$

$$= \boldsymbol{w}^T \mathbb{C}_t \boldsymbol{w} \tag{II.5}$$

in which

$$\mathbb{C}_t = \mathrm{E}\{[\boldsymbol{X}(t) - \mu(t)]^T [\boldsymbol{X}(t) - \mu(t)]\} \quad , \tag{II.6}$$

$$\hat{X}(0,t) = \sum_{s=1}^{S} w_s \, X(s,t) \quad , \tag{II.7}$$

$$\mu(s,t) = \mathrm{E}[X(s,t)] \tag{II.8}$$

with C_t being the covariance matrix between stations for month t; X_t and μ_t are Sx1 vectors of the random observation and the mean of monthly precipitation at station s, for s=1,2,..,S, in month t and T is transpose of a vector or a matrix. It is, then, desirable to find the weighing factors that minimizes the variance associated with the estimator X(0,t).

This method is called herein the optimum weighing method (OPTIM) in that it has the object function of minimizing the error variance of the estimator subject to the constraint that the sum of weights is 1. That is,

$$\textit{minimize} \quad \sigma^2_{0t} \tag{II.9}$$

subject to

$$\sum_{s=1}^{S} w_s = 1 \tag{II.10}$$

in which $\mathcal{L}(w,\lambda)$ is the Lagrange function and λ is the Lagrangian multiplier. The above minimization problem is solved by the Lagrange multiplier method which converts the original constrained minimization problem into an unconstrained minimization one as

$$\min \quad \mathcal{L}\,(\boldsymbol{w},\lambda) = \boldsymbol{w}^T \mathbb{C}_t \boldsymbol{w} + 2\lambda(\boldsymbol{w}^T \mathbf{1} - 1) \tag{II.11}$$

Then, the solution to Eq.(II.11) must satisfy the following equations

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{w}} = 2\mathbb{C}_t\, \boldsymbol{w} + 2\lambda\, \mathbf{1} = \mathbf{0} \qquad \text{(II.12)}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = \mathbf{1}^T\, \boldsymbol{w} - 1 = 0 \qquad \text{(II.13)}$$

where 1 is a vector of ones. Solving Eqs.(II.12) and (II.13), the optimal weighing factors can be obtained as

$$\begin{pmatrix} \boldsymbol{w}^* \\ \lambda^* \end{pmatrix} = \begin{pmatrix} \mathbb{C} & \mathbf{1} \\ \mathbf{1}^T & 0 \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} \qquad \text{(II.14)}$$

in which w^* is the vector of optimum weight for the observed stations. Then the missing values can be estimated by

$$\hat{X}(0,t) = \sum_{s=1}^{S} w_s^*\, X(s,t) \qquad \text{(II.15)}$$

in which w_s^* is the optimum weight for the observed station s.

II.2.3 Cross Correlation Weighing (RSQR) Method

Since the monthly precipitation could exhibit spatial correlation, measures of correlation between existing stations can be used to compute the weights. The basic idea is that if a measured station s is strongly correlated to the station with missing values through the whole recording period, then the station s should be given higher weight. Specifically, the weight for station s, w_s, by the RSQR method is computed as

$$w_s = \frac{r_{0s}^2}{\sum_{s=1}^{S} r_{0s}^2} \qquad \text{(II.16)}$$

where r^2_{os} is the coefficient of determination between the station with missing values and the station s with observations.

II.2.4 Variance (or Standard Deviation) Weighing Method

Based on the results of simple regression analysis between two stations, the statistics, such as standard error or its squared value, can also be used to compute the weight. Intuitively, the larger the variance or the standard error for the station, the smaller the weight should be assigned to it. Therefore, the weight should be inversely proportional to the variance and standard error. The inverse variance (INVAR) and inverse standard error (INSTD) techniques are similar to the IDSQ and IDLIN methods, respectively. The only difference is that the IDLIN and the IDSQ techniques require only physical distances between stations to calculate the weights whereas the INVAR and the INSTD methods require the observed monthly precipitation data to compute the statistics to obtain the weights for all stations used. The weights by the INVAR method are calculated as

$$w_s = \frac{\frac{1}{\sigma_{0s}^2}}{\sum_{s=1}^{S} \frac{1}{\sigma_{0s}^2}} \quad , \; s=1,2,\ldots S \qquad \text{(II.17)}$$

and by the INSTD method

$$w_s = \frac{\frac{1}{\sigma_{0s}}}{\sum_{s=1}^{S} \frac{1}{\sigma_{0s}}} \quad , \; s=1,2,\ldots,S \qquad \text{(II.18)}$$

where σ_{0s} is standard error from regression between the station with missing values and station s with observations.

II.2.5 Hybrid Method

In the above methods, the missing monthly precipitation data are estimated from the weighted average of the surrounding observed values using various weighing techniques. In these methods, the actual observed monthly precipitation data from the surrounding stations are used. Alternatively, simple regression equations can be developed for the station to be estimated and all other surrounding stations based on the data concurrently available. These regression equations can be used, as the first attempt, to estimate the missing value from each surrounding station. The resulting estimated missing values from the regression equations are further weighted by the above methods. That is,

$$\tilde{X}(0,t) = \frac{\sum_{s=1}^{S} \hat{X}(0,t)\; w_s}{\sum_{s=1}^{S} w_s} \qquad \text{(II.19)}$$

where $\hat{X}_{(0,t)}$ is the missing monthly total precipitation estimated by the regression equation from station s with observations and $\tilde{X}(0,t)$ is the estimated missing monthly total precipitation computed by the weighted average of regression estimates, X(0,t).

II.3 Comparison Study for the Performance of Estimation Methods

Among the estimation methods, it is desirable to examine the performance of each method and identify the most accurate one. Two criteria are used in this performance evaluation and they are the root-mean-square-error (RMSE)

$$RMSE_{s,k} = \sqrt{\frac{\sum_{t=1}^{T} [X(s,t) - \hat{X}_k(s,t)]^2}{T}} \quad , \; s = 1,2,\ldots S \tag{II.20}$$

and the mean-absolute-error (MAE)

$$MAE_{s,k} = \frac{\sum_{t=1}^{T} |X(s,t) - \hat{X}_k(s,t)|}{T} \quad , \; s = 1,2,\ldots S \tag{II.21}$$

in which T is total months used, and $\hat{X}_k(s,t)$ is the value estimated by weighing method k.

In this comparison, precipitation stations were selected under the condition that there exists, among them, as many concurrently observed values as possible and that all of them are not too close to each other. Nine such stations were chosen and they were 102, 103-1, 108-2, 115-2, 119, 122, 123, 124 and 125. Three different periods data set observed concurrently at all 9 stations were selected to establish regression equations between two stations to compute the corresponding correlation, variance, or standard error for purpose of determining the weights from different weighing techniques. These periods were a 12-month period (7/1982-12/1983), a 24-month period (2/1981-12/1983), and a 48-month period (8/1978-12/1983). Then it was assumed that the observed monthly precipitation (1/1985-7/1988) for these nine stations were missing. The estimated values for the assumed missing period by the different estimation methods on the basis of different record lengths (i.e., 12-, 24-, and 48-month) were compared with the actual observations to calculate the RMSE and MAE.

II.4 Results and Discussion

The results associated with two stations (125, 103-1) out of 9 are presented here. Tables II.1-II.2 show the results of RMSE and MAE including the individual weight and regression equation for each station used. The values of error criteria are also plotted in Figures II.5 and II.6.

Table II.1 presents the comparison of each weighing technique for station 125. Referring to Table II.1(b), the IDLIN method estimates, without using regression, has the smallest values of RMSE (0.592) and MAE (0.436) for all three different periods. The IDSQ method has the second smallest values of errors (RMSE = 0.626, MAE = 0.456). The INVAR technique, considering regression, yields the largest values of RMSE (0.983) and MAE (0.735) for a 12-month period.

Different results for station 103-1 in Table II.2 show that the INSTD method considering regression has the smallest value of RMSE (0.348) but a slightly large MAE (.269) for a 12-month period whereas the IDSQ method, without considering regression, has the smallest value of MAE (0.255) and a rather large RMSE (0.398). Not presented here, the results for station 123 shows that the OPTIM method without using regression has the smallest values of RMSE (0.570) and MAE (0.389). It was also found that when the number of years used to establish regression equations increased, the corresponding errors decreased.

There is no single estimation technique that is uniformly superior in all circumstances. Overall speaking, the linear inverse distance weighing method (IDLIN) is better than all other methods throughout the 9 stations. The IDLIN method is then used to estimate the missing monthly precipitation for further analysis.

Precipitation Data from Snowy Range

Station to be estimated : 125

Stations used : 102,1031,1082,1152,
119,122,123,124

(A). Weights

Station / Methods		Weights for each Stations							
		STA 102	STA 1031	STA1082	STA 1152	STA 119	STA 122	STA 123	STA 124
IDSQ		0.014	0.040	0.181	0.510	0.027	0.098	0.052	0.077
IDLIN		0.049	0.082	0.174	0.292	0.067	0.128	0.093	0.114
INSTD	1 yrs	0.100	0.120	0.131	0.133	0.100	0.165	0.137	0.114
	2 "	0.092	0.112	0.151	0.161	0.088	0.152	0.128	0.117
	4 "	0.097	0.111	0.151	0.168	0.087	0.149	0.118	0.118
RSQR	1 "	0.115	0.127	0.132	0.125	0.118	0.136	0.120	0.128
	2 "	0.100	0.118	0.147	0.136	0.105	0.135	0.126	0.133
	4 "	0.108	0.118	0.142	0.135	0.108	0.135	0.126	0.128
INVAR	1 "	0.132	0.128	0.085	0.094	0.149	0.127	0.163	0.122
	2 "	0.159	0.137	0.079	0.091	0.162	0.121	0.138	0.112
	4 "	0.154	0.138	0.082	0.091	0.164	0.116	0.133	0.122
OPTIM	1 "	0.116	0.132	0.064	0.069	0.156	0.150	0.192	0.121
	2 "	0.136	0.138	0.072	0.082	0.152	0.141	0.160	0.119
	4 "	0.146	0.140	0.071	0.078	0.164	0.128	0.146	0.127

(B). Estimation without using Regression from Individual Station

Years	Error Criterion	Methods					
		IDSQ	IDLIN	INSTD	RSQR	INVAR	OPTIM
1	RMSE	0.626	0.592	0.698	0.719	0.799	0.834
	MAE	0.456	0.436	0.482	0.490	0.530	0.556
2	RMSE	0.626	0.592	0.664	0.692	0.821	0.821
	MAE	0.456	0.436	0.470	0.481	0.541	0.545
4	RMSE	0.626	0.592	0.660	0.699	0.820	0.838
	MAE	0.456	0.436	0.468	0.482	0.540	0.553

Table II.1 : Comparison of Estimation Method for Station 125

(C). Estimation using Regression from Individual Station

Years	Error Criterion	By Individual Station								By Weighted Results					
		STA 102	STA103	STA108	STA 1152	STA119	STA 122	STA123	STA124	IDSQ	IDLIN	INSTG	RSQR	INVAR	OPTIM
1	RMSE	1.293	1.076	.989	.974	1.297	.783	.944	1.138	.933	.938	.948	.972	.983	.971
	MAE	.985	.803	.709	.692	.977	.626	.655	.808	.677	.689	.708	.723	.735	.730
2	RMSE	1.283	1.060	.784	.737	1.347	.781	.923	1.007	.755	.808	.855	.875	.946	.933
	MAE	1.076	.902	.566	.523	1.131	.563	.690	.769	.558	.621	.681	.700	.771	.760
4	RMSE	1.110	.972	.717	.645	1.245	.724	.915	.913	.665	.717	.764	.786	.847	.850
	MAE	.931	.841	.531	.489	1.035	.525	.712	.718	.525	.583	.626	.644	.697	.698

** Note **

$Y_{125,1} = 1.138 + .684\ X_{102}$; SE = 2.454 (0.561)*
$Y_{125,2} = 1.682 + .661\ X_{102}$; SE = 1.642 (0.287)
$Y_{125,4} = 1.261 + .824\ X_{102}$; SE = .982 (0.427)

$Y_{125,1} = .786 + .799\ X_{1031}$; SE = 2.041 (0.675)
$Y_{125,2} = 1.093 + .784\ X_{1031}$; SE = 1.356 (0.396)
$Y_{125,4} = .881 + .884\ X_{1031}$; SE = .860 (0.510)

$Y_{125,1} = .748 + .579\ X_{1082}$; SE = 1.876 (0.733)
$Y_{125,2} = .384 + .708\ X_{1082}$; SE = 1.003 (0.617)
$Y_{125,4} = .374 + .752\ X_{1082}$; SE = .634 (0.730)

$Y_{125,1} = .740 + .569\ X_{1152}$; SE = 1.847 (0.654)
$Y_{125,2} = .419 + .696\ X_{1152}$; SE = .943 (0.530)
$Y_{125,4} = .434 + .749\ X_{1152}$; SE = .571 (0.659)

$Y_{125,1} = 1.139 + .809\ X_{119}$; SE = 2.461 (0.586)*
$Y_{125,2} = 1.640 + .728\ X_{119}$; SE = 1.723 (0.312)
$Y_{125,4} = 1.251 + .871\ X_{119}$; SE = 1.102 (0.425)

$Y_{125,1} = .674 + .914\ X_{122}$; SE = 1.486 (0.778)
$Y_{125,2} = .481 + .904\ X_{122}$; SE = .999 (0.517)
$Y_{125,4} = .370 + .963\ X_{122}$; SE = .641 (0.665)

$Y_{125,1} = .395 + .913\ X_{123}$; SE = 1.791 (0.606)
$Y_{125,2} = .573 + .902\ X_{123}$; SE = 1.181 (0.453)
$Y_{125,4} = .546 + .959\ X_{123}$; SE = .809 (0.577)

$Y_{125,1} = .724 + .770\ X_{124}$; SE = 2.159 (0.685)
$Y_{125,2} = .829 + .825\ X_{124}$; SE = 1.288 (0.508)
$Y_{125,4} = .726 + .907\ X_{124}$; SE = .808 (0.592)

* represents R-Square

Table II.1

Precipitation Data from Snowy Range

Station to be estimated : 1031

Stations used : 102,1082,1152,119,
122,123,124,125

(A). Weights

Station / Methods		Weights for each Stations							
		STA102	STA1082	STA1152	STA119	STA 122	STA123	STA 124	STA 125
IDSQ		0.066	0.120	0.060	0.328	0.192	0.128	0.072	0.035
IDLIN		0.096	0.129	0.092	0.214	0.164	0.134	0.101	0.070
INSTD	1 yrs	0.155	0.112	0.157	0.104	0.100	0.088	0.194	0.090
	2 "	0.158	0.102	0.143	0.095	0.130	0.107	0.186	0.078
	4 "	0.161	0.096	0.137	0.091	0.125	0.116	0.182	0.091
RSQR	1 "	0.128	0.120	0.121	0.134	0.119	0.133	0.134	0.111
	2 "	0.133	0.123	0.125	0.138	0.120	0.136	0.137	0.088
	4 "	0.131	0.123	0.125	0.134	0.123	0.133	0.133	0.098
INVAR	1 "	0.119	0.094	0.097	0.132	0.145	0.147	0.116	0.150
	2 "	0.126	0.099	0.104	0.128	0.142	0.133	0.115	0.154
	4 "	0.133	0.099	0.102	0.138	0.133	0.131	0.122	0.141
OPTIM	1 "	0.116	0.063	0.069	0.155	0.149	0.191	0.120	0.139
	2 "	0.143	0.076	0.086	0.159	0.148	0.168	0.126	0.094
	4 "	0.153	0.075	0.082	0.172	0.134	0.153	0.134	0.096

(B). Estimation without using Regression from Individual Station

Years	Error Criterion	Methods					
		IDSQ	IDLIN	INSTD	RSQR	INVAR	OPTIM
1	RMSE	0.398	0.439	0.481	0.482	0.472	0.415
	MAE	0.255	0.311	0.370	0.361	0.349	0.298
2	RMSE	0.398	0.439	0.462	0.470	0.483	0.401
	MAE	0.255	0.311	0.355	0.355	0.358	.291
4	RMSE	0.398	0.439	0.461	0.479	0.465	0.388
	MAE	0.255	0.311	0.353	0.361	0.345	0.280

Table II.2 : Comparison of Estimation Method for Station 103-1

(C). **Estimation using Regression from Individual Station**

Years	Error Criterion	By Individual Station								By Weighted Results					
		STA102	STA1082	STA1152	STA119	STA122	STA123	STA124	STA125	IDSQ	IDLIN	INSTD	RSQR	INVAR	OPTIM
1	RMSE	.499	.692	.491	.743	.771	.878	.398	.855	.396	.368	.348	.357	.374	.366
	MAE	.394	.547	.396	.419	.632	.606	.258	.599	.256	.267	.269	.274	.280	.264
2	RMSE	.444	.690	.494	.740	.540	.657	.378	.904	.397	.376	.361	.364	.389	.356
	MAE	.357	.540	.387	.438	.367	.472	.274	.723	.279	.284	.284	.285	.304	.275
4	RMSE	.418	.698	.490	.739	.537	.580	.370	.739	.397	.375	.360	.368	.376	.356
	MAE	.324	.529	.388	.427	.367	.413	.259	.613	.277	.283	.281	.286	.293	.274

** Note **

$Y_{1031,1} = .456 + .888\, X_{102}$: SE = .947 (0.894)*
$Y_{1031,2} = .450 + .938\, X_{102}$: SE = .569 (0.897)
$Y_{1031,4} = .323 + .973\, X_{102}$: SE = .370 (0.912)

$Y_{1031,1} = .480 + .618\, X_{1082}$: SE = 1.313 (0.791)
$Y_{1031,2} = .389 + .635\, X_{1082}$: SE = .883 (0.770)
$Y_{1031,4} = .294 + .637\, X_{1082}$: SE = .617 (0.803)

$Y_{1031,1} = .266 + .649\, X_{1152}$: SE = .932 (0.804)
$Y_{1031,2} = .130 + .684\, X_{1152}$: SE = .632 (0.794)
$Y_{1031,4} = .168 + .679\, X_{1152}$: SE = .433 (0.823)

$Y_{1031,1} = .254 + 1.074 X_{119}$: SE = 1.410 (0.977)
$Y_{1031,2} = .403 + 1.028 X_{119}$: SE = .947 (0.966)
$Y_{1031,4} = .244 + 1.056 X_{119}$: SE = .654 (0.956)

$Y_{1031,1} = .533 + .936\, X_{122}$: SE = 1.462 (0.771)*
$Y_{1031,2} = .283 + .863\, X_{122}$: SE = .691 (0.732)
$Y_{1031,4} = .165 + .856\, X_{122}$: SE = .475 (0.804)

$Y_{1031,1} = -.640 + 1.185 X_{123}$: SE = 1.666 (0.965)
$Y_{1031,2} = -.286 + 1.045 X_{123}$: SE = .840 (0.945)
$Y_{1031,4} = -.088 + .988\, X_{123}$: SE = .513 (0.939)

$Y_{1031,1} = -.019 + .948\, X_{124}$: SE = .756 (0.982)
$Y_{1031,2} = .185 + .907\, X_{124}$: SE = .484 (0.953)
$Y_{1031,4} = .124 + .926\, X_{124}$: SE = .327 (0.944)

$Y_{1031,1} = .492 + .845\, X_{125}$: SE = 1.623 (0.675)
$Y_{1031,2} = 1.528 + .505\, X_{125}$: SE = 1.156 (0.396)
$Y_{1031,4} = .906 + .577\, X_{125}$: SE = .654 (0.510)

* represents R-Square

Table II.2

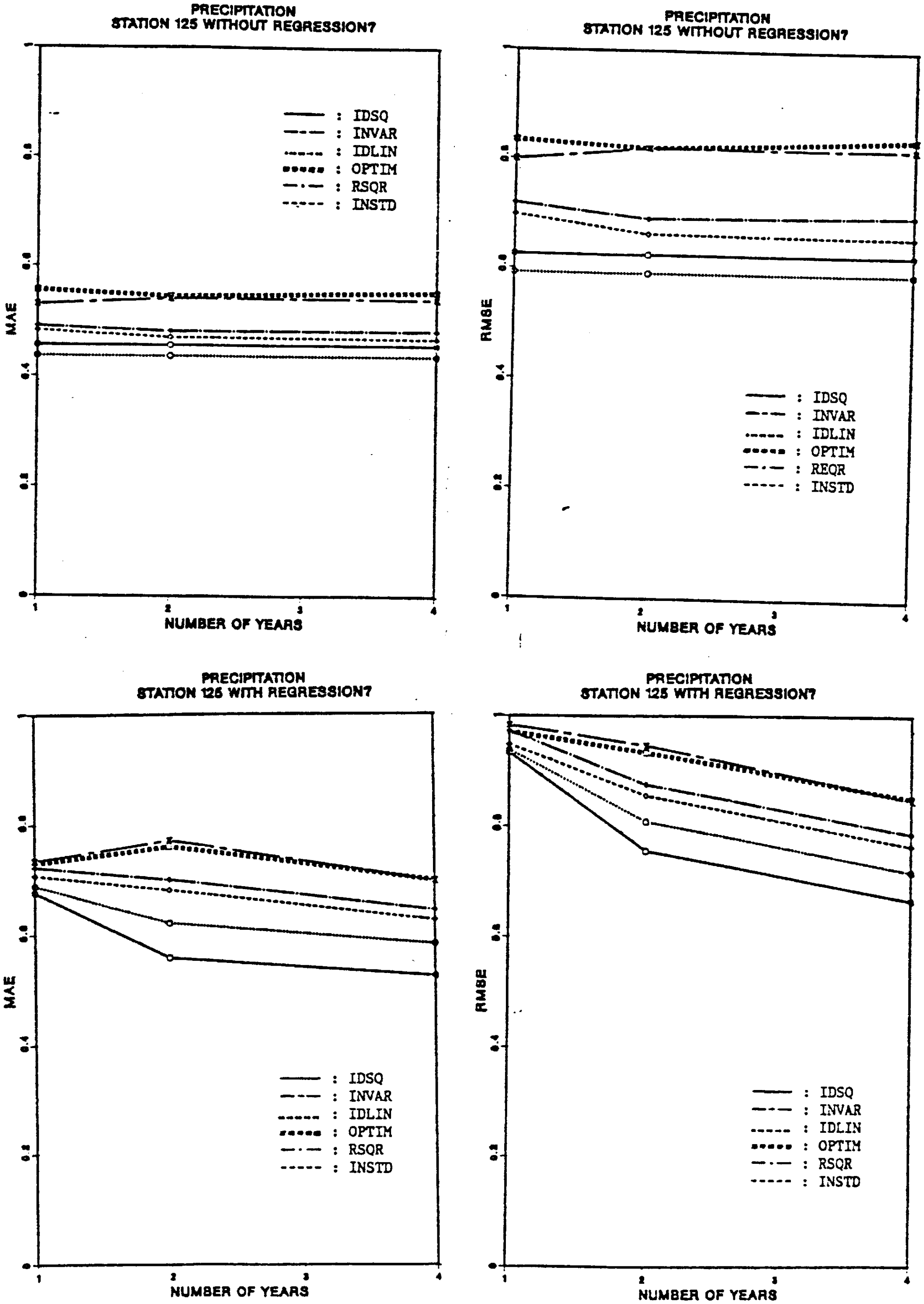


Figure II.5 : Errors of Estimation Methods for Station 125

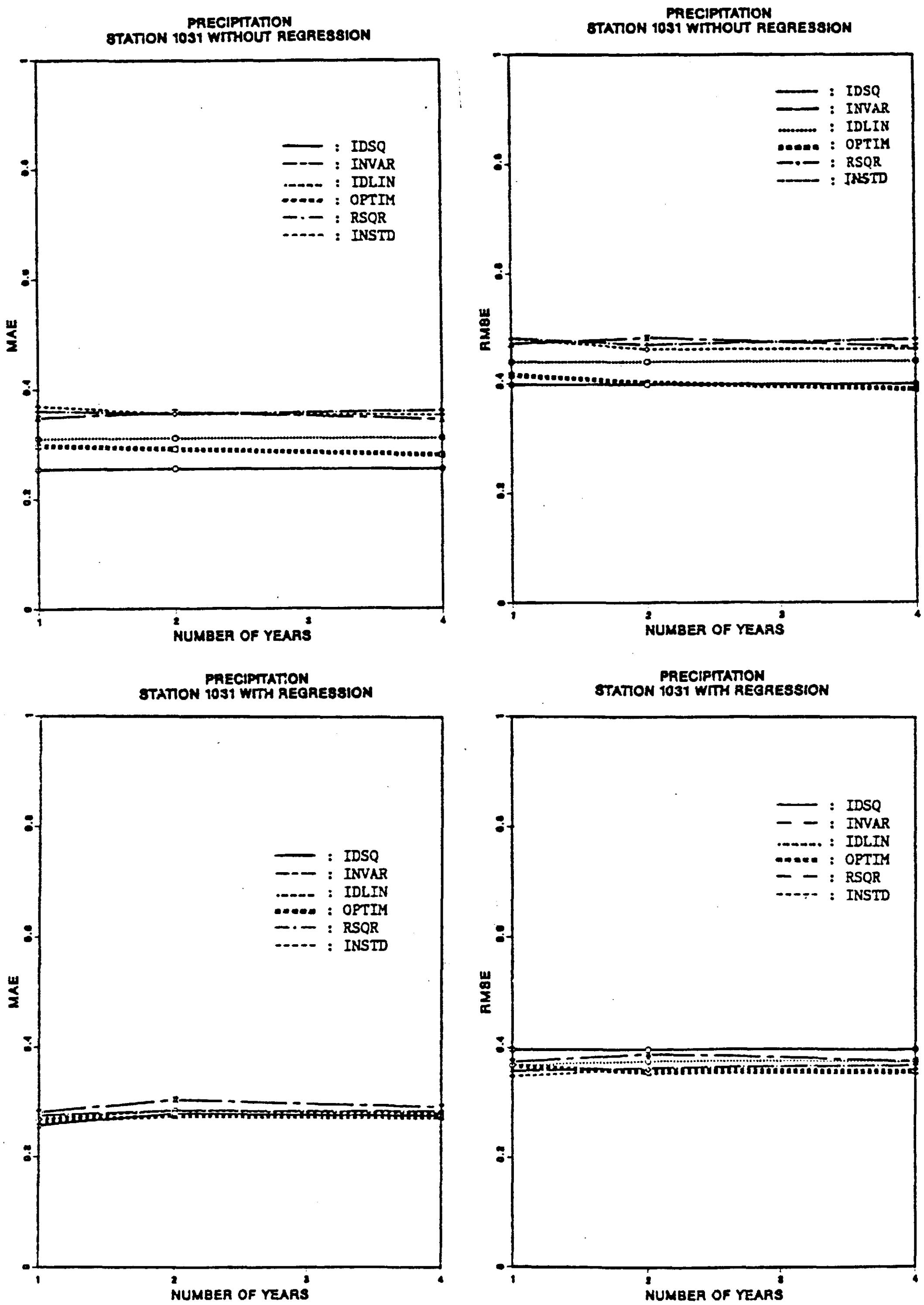


Figure II.6 : Errors of Estimation Methods for Station 103-1

CHAPTER III

ANALYSIS OF MONTHLY PRECIPITATION

When a time series is studied, the observed value of the series at a particular time period should be viewed as a random value. If the series exhibits a systematic change over time, some transformation procedures can be applied to the series to make it stationary. Detrending and deseasonalizing procedures were employed in this study.

III.1 Removal of Long Term Trend

The trend in a time series can be defined as any systematic change in the level with respect to time. When a time series is steadily increasing or decreasing over time, the annual or some periodic mean of the series also intends to change through the whole period. In this case, the analysis of the data might not yield consistent results because of the presence of such trend. Hence it is necessary to remove the trend over time if it exists.

To examine if the monthly precipitation has the over-year trend, the annual mean of each precipitation station is computed as

$$\overline{X}(y) = \frac{1}{12}\sum_{m=1}^{12} X(m,y) \quad , \; y = 1,2,\ldots,Y \qquad \text{(III.1)}$$

in which $\overline{X}(y)$ is the arithematic average of monthly precipitation in year y, X(m,y) is the observed monthly precipitation in month m and year y, and Y is the total number of years in the record. The annual mean precipitations for all 21 stations are plotted in Figures III.1–III.7. Although there are some fluctuations in the annual average precipitations, all stations do not appear to have any visible over-year changes in the annual mean. There is a slightly large fluctuation for the first several years at station 109, but the annual mean seems to stabilize after that. The annual mean at station 121-A has a decreasing trend for the last three years, but there was no more data afterward because the station was removed.

A simple linear regression analysis is performed to fit the annual average precipitation over time. The resulting regression equations are shown in Table III.1 with the p-values for the slope. As can be seen, all the slopes of the fitted regression lines are insignificant at both 10% and 5% levels. Based on these results, the conclusion was made that there were no long-term trend in monthly precipitation. Therefore, it is not necessary to detrend the monthly precipitation data.

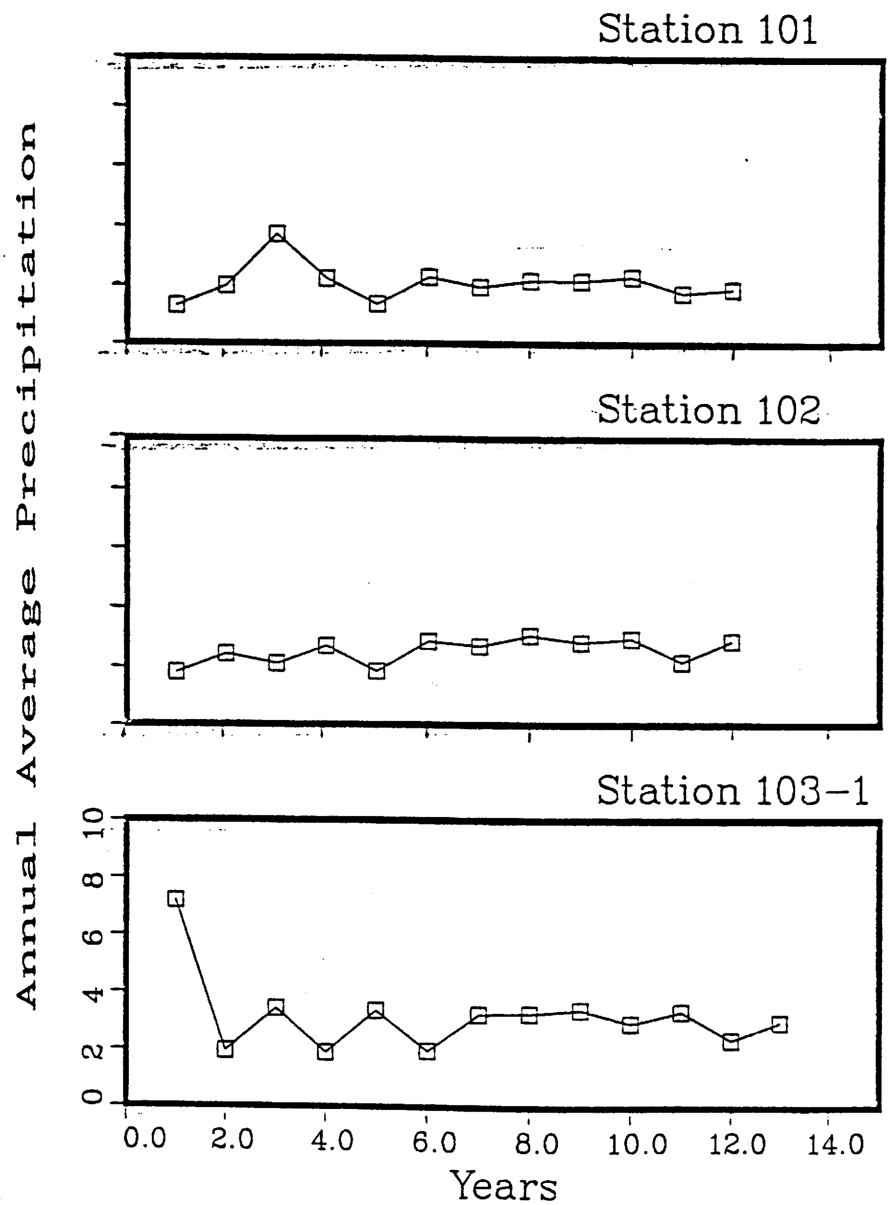


Figure III.1 Long Term Trend of Monthly Precipitation

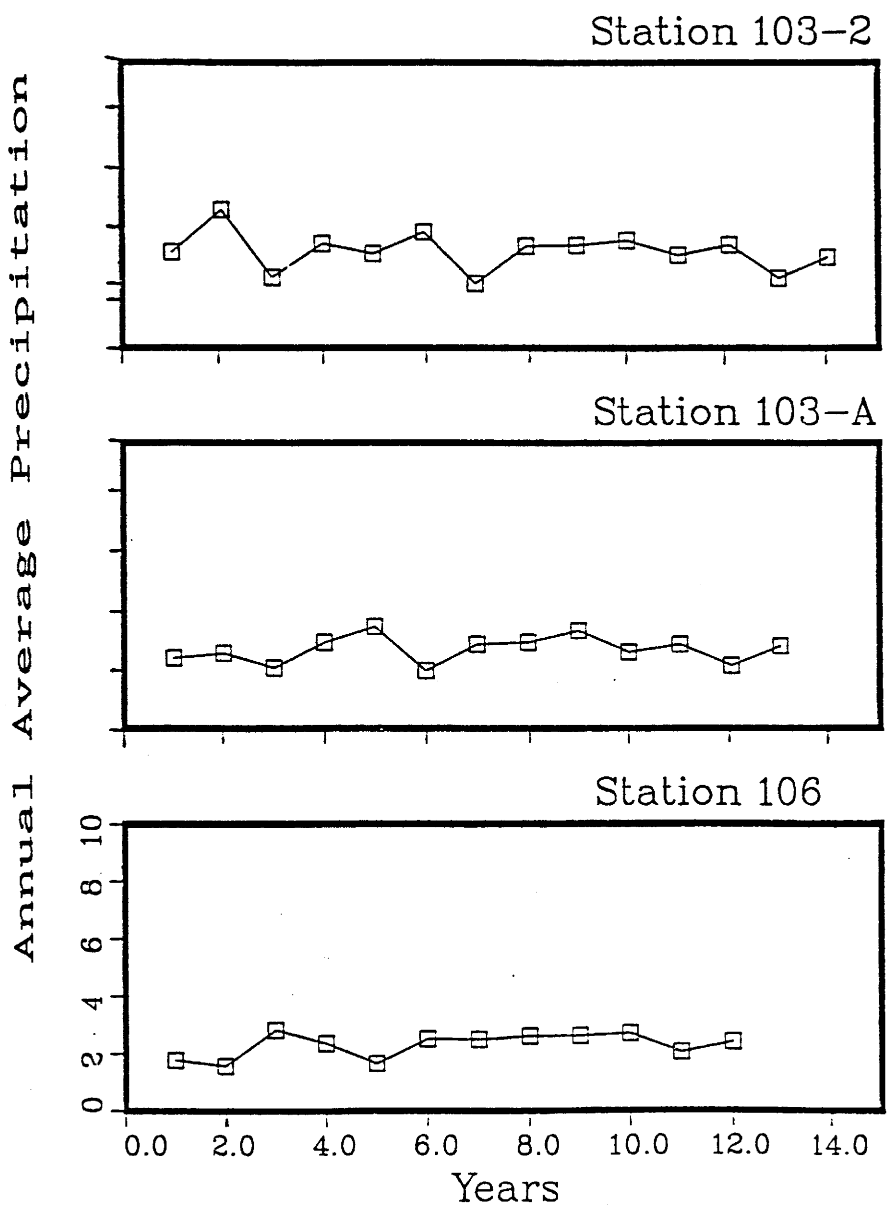


Figure III.2 Long Term Trend of Monthly Precipitation

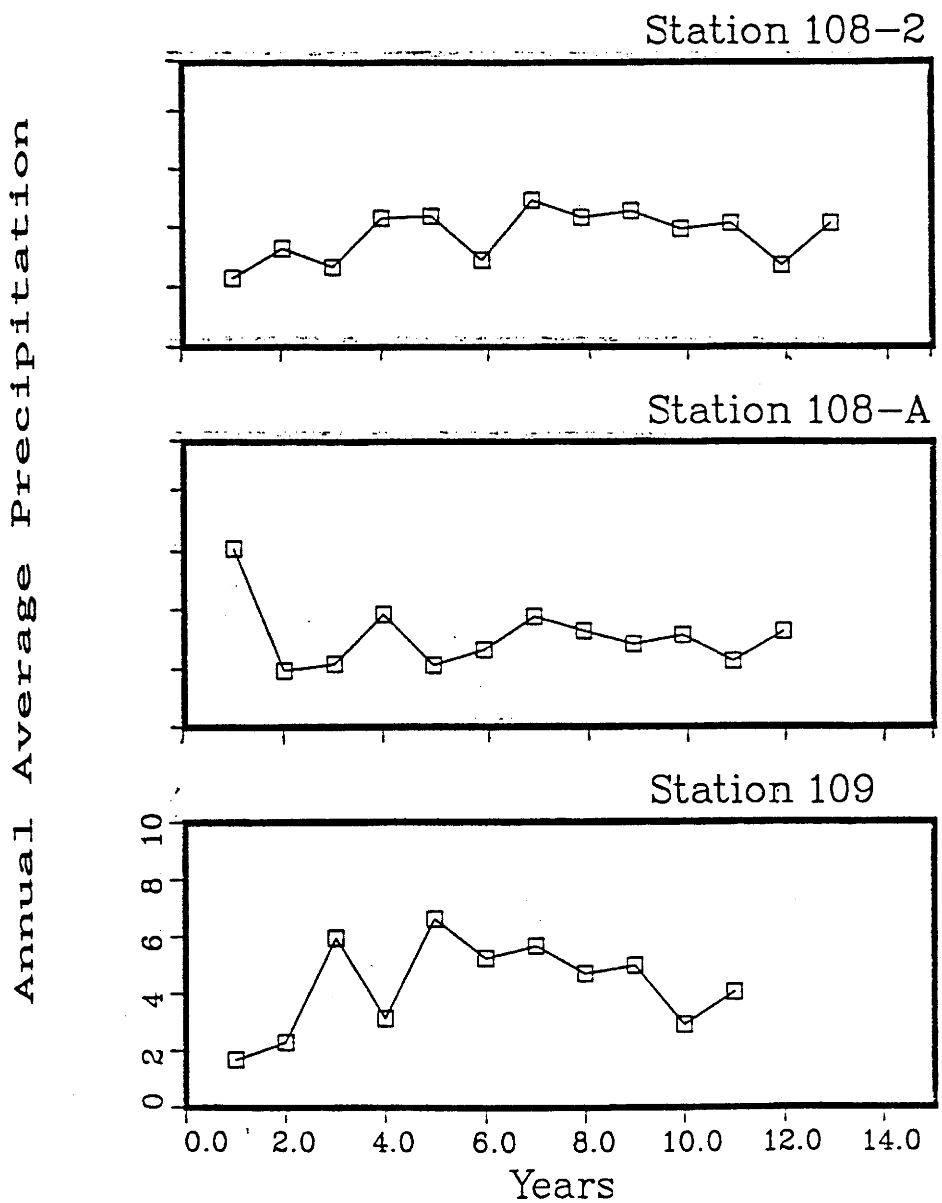


Figure III.3 Long Term Trend of Monthly Precipitation

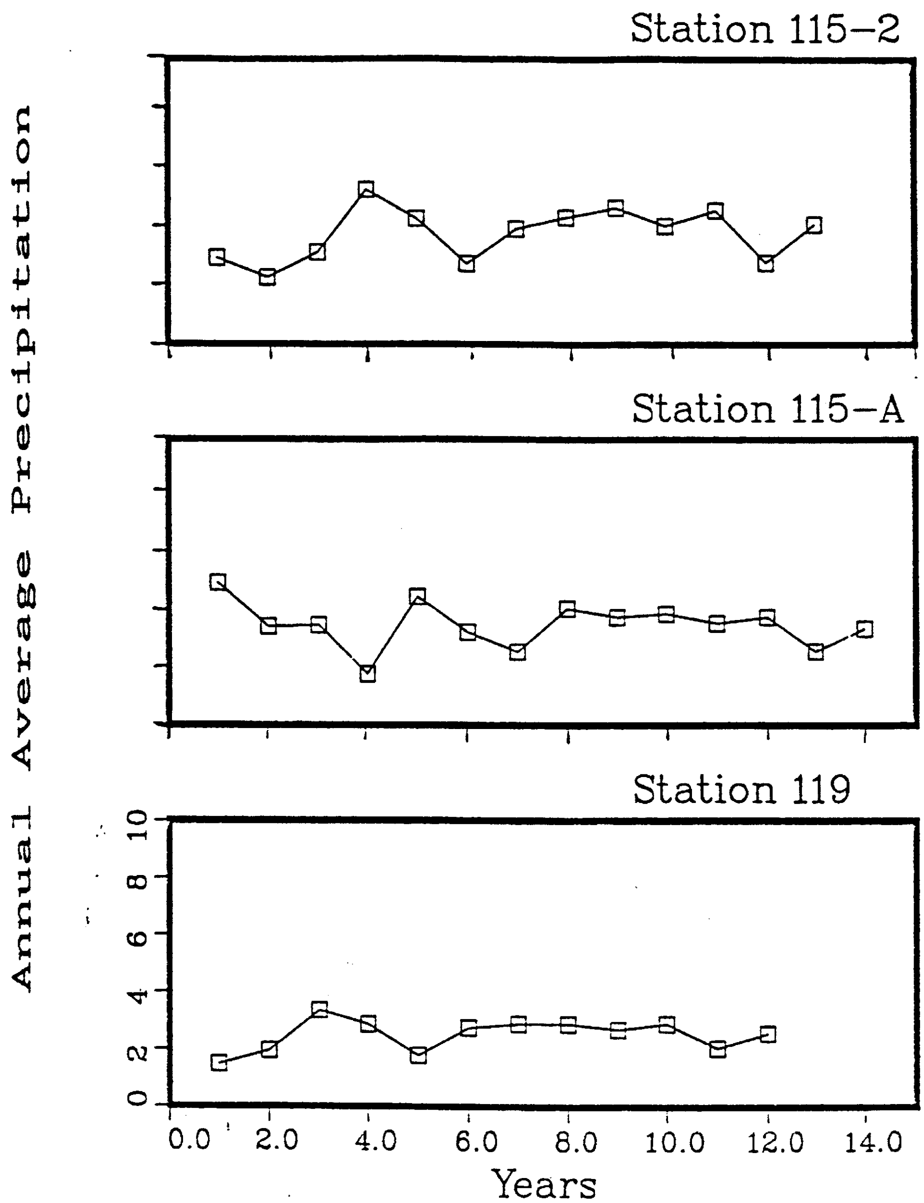


Figure III.4 Long Term Trend of Monthly Precipitation

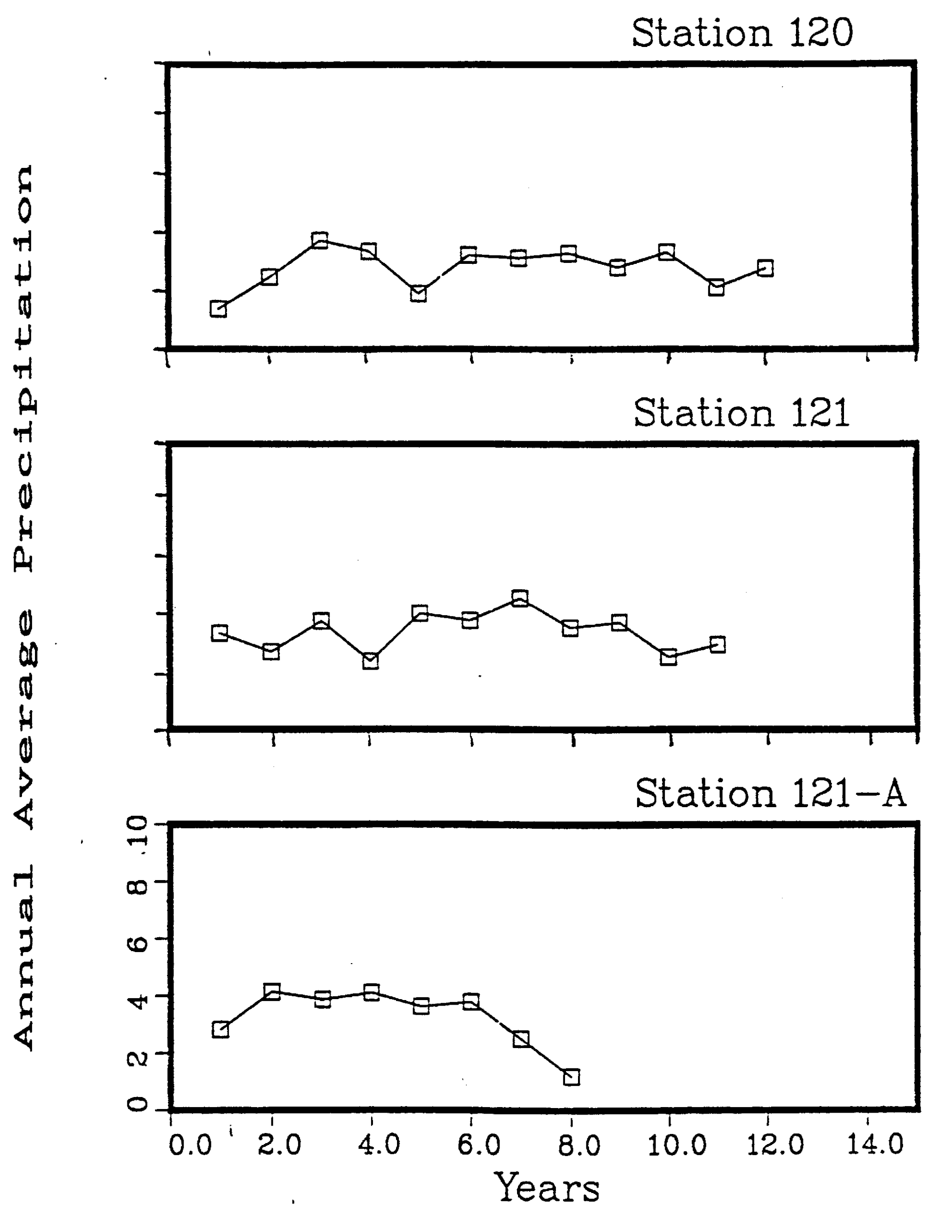


Figure III.5 Long Term Trend of Monthly Precipitation

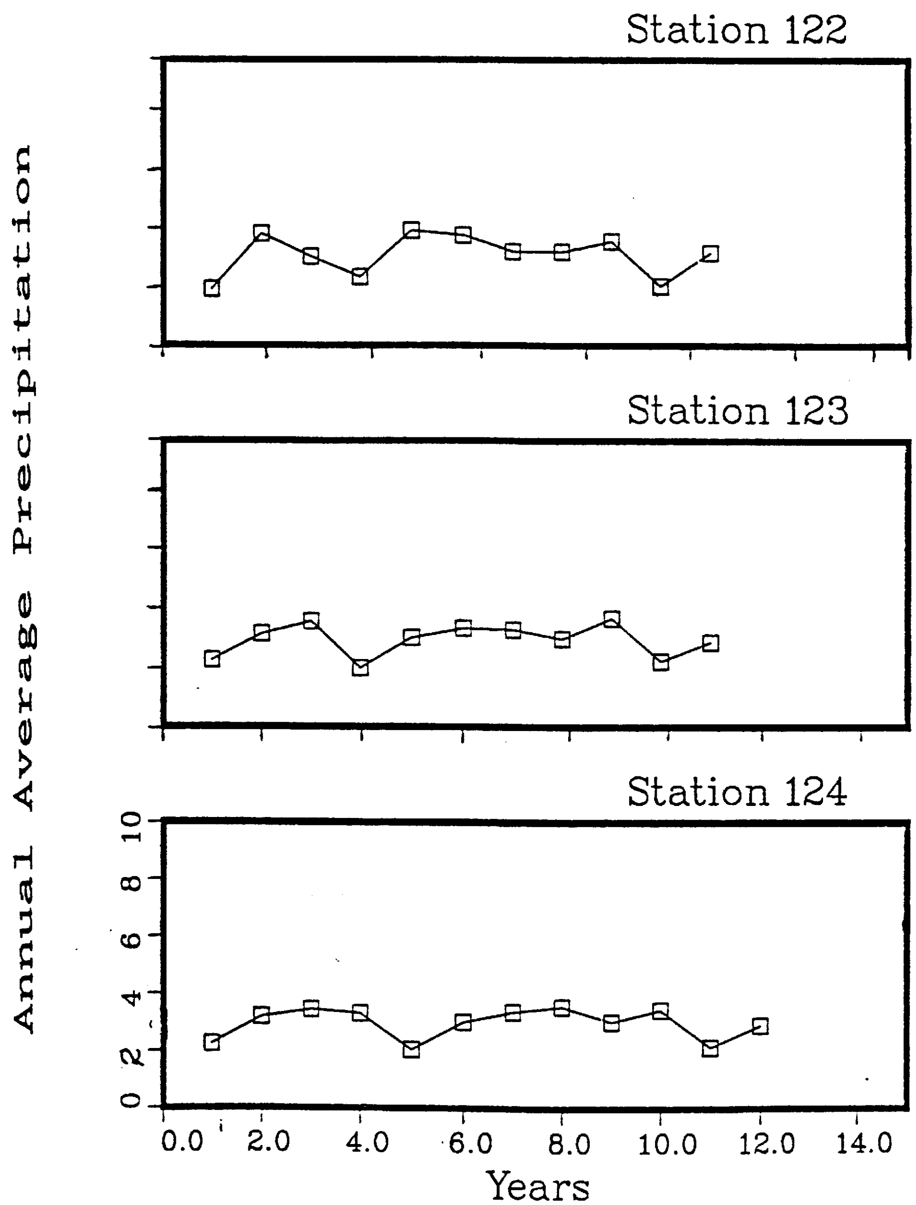


Figure III.6 Long Term Trend of Monthly Precipitation

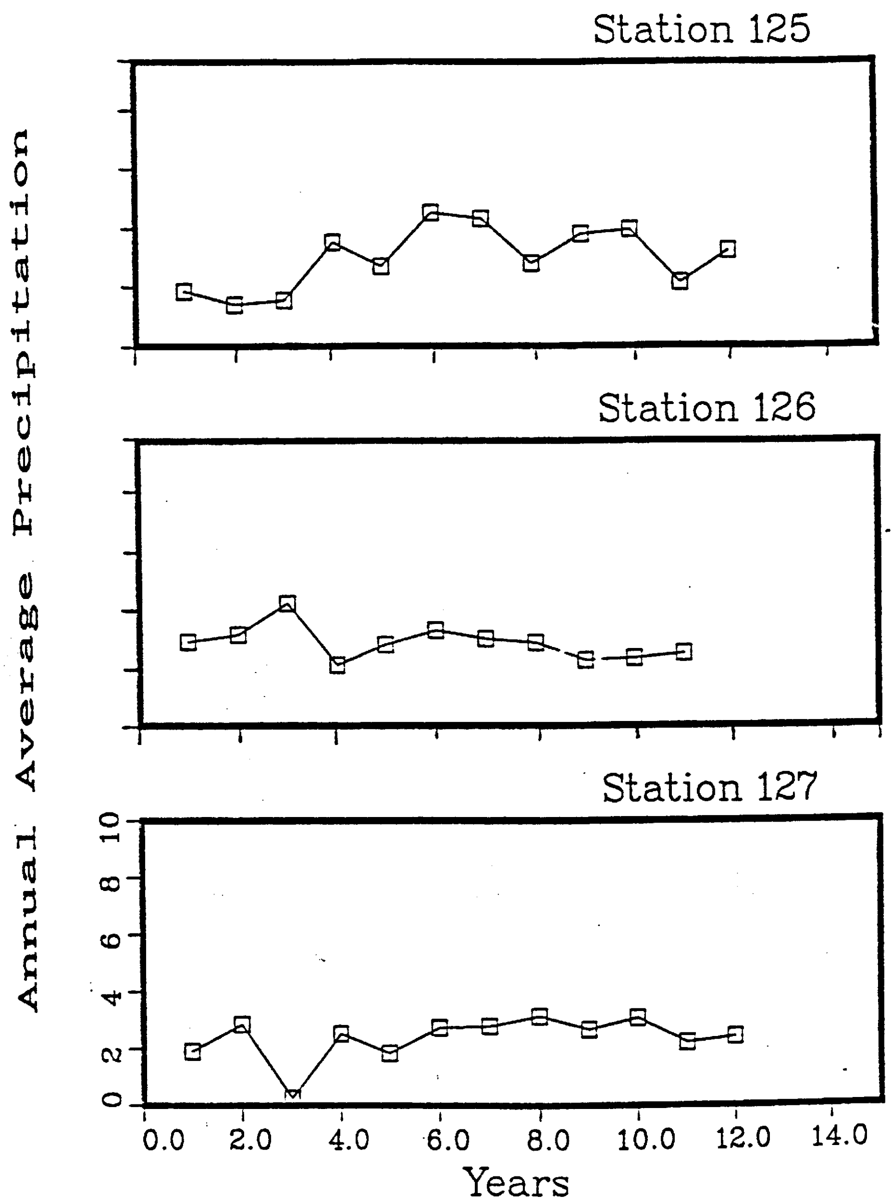


Figure III.7 Long Term Trend of Monthly Precipitation

Station	Regression Equations for Annual Trend
101	$\bar{X}$ = 2.22 - 0.0132 Y, (0.815)*
102	= 2.11 + 0.0708 Y, (0.062)
103-1	= 3.90 - 0.1041 Y, (0.316)
103-2	= 3.68 - 0.0404 Y, (0.387)
103-A	= 2.64 + 0.0182 Y, (0.627)
106	= 1.96 + 0.0506 Y, (0.166)
108-2	= 3.33 + 0.0856 Y, (0.206)
108-A	= 3.71 - 0.0795 Y, (0.443)
109	= 3.55 + 0.1219 Y, (0.453)
115-2	= 3.32 + 0.0793 Y, (0.269)
115-A	= 3.76 - 0.0336 Y, (0.555)
119	= 2.24 + 0.0409 Y, (0.409)
120	= 2.61 + 0.0370 Y, (0.562)
121	= 3.41 + 0.0002 Y, (0.997)
121-A	= 4.34 - 0.2414 Y, (0.133)
122	= 3.05 + 0.0177 Y, (0.813)
123	= 2.93 + 0.0182 Y, (0.762)
124	= 2.96 + 0.0025 Y, (0.959)
125	= 2.17 + 0.1422 Y, (0.130)
126	= 3.42 - 0.0869 Y, (0.135)
127	= 1.81 + 0.0845 Y, (0.211)

$\bar{X}$: Annual mean of monthly precipitation
Y : number of years,
* : P-value of regression coefficient

Table III.1 : Regression Equations of Annual Mean Precipitation

III.2 Removal of Periodicity

In monthly precipitation time series, some types of seasonal patterns would be likely to exist at 12-month intervals. This fluctuation reveals a seasonal cycle or periodicity that recurs about every 12 months. To make the time series stationary, deseasonalization is needed to remove the within-year seasonal fluctuation in the monthly precipitation time series. Two methods (Parzen, 1981) were used in this study to deseasonalize the monthly precipitation data.

III.2.1 Using Monthly Sample Statistics

The monthly averages of the total precipitation for each station on a given month m can be computed as

$$\overline{X}(m) = \frac{1}{Y}\sum_{y=1}^{Y} X(m,y) \quad , \; m = 1,2,\ldots,12 \qquad \text{(III.2)}$$

in which $\overline{X}(y)$ is the average precipitation for month m. Then, the standardization procedure is applied to obtain the deseasonalized data as

$$Z(m,y) = \frac{X(m,y) - \overline{X}(m)}{S(m)} \qquad \text{(III.3)}$$

in which S(m) is the sample standard deviation of monthly precipitation for month m,

$$S(m) = \sqrt{\frac{\sum_{y=1}^{Y} [X(m,y) - \overline{X}(m)]^2}{Y-1}} \quad , m=1,2,\ldots,12 \qquad \text{(III.4)}$$

The sample statistics of the monthly precipitation are shown in Tables III.2 and III.3. The standardized monthly precipitations for stations 106 and 108-2 are presented in Figures III.8 and III.9 to compare with the original time series shown in Chapter I. As can be observed, the seasonal fluctuation that appeared in the original time series data is removed. The standardized monthly precipitation, then, can be considered stationary having zero mean and unit variance.

III.2.2 Using Fourier Series

Another method of deseasonalization is to fit the monthly average mean by a Fourier series (Bloomfield, 1975). Then, the above standardization procedure can be performed using the fitted average, instead of

Station	Month					
	Jan.	Feb.	Mar.	Apr.	May.	Jun.
101	2.03	1.51	2.34	2.47	2.40	1.24
102	2.58	2.15	3.33	3.54	2.85	1.41
103-1	3.04	1.39	3.67	4.05	3.33	1.61
103-2	3.53	2.82	4.26	4.43	3.78	1.69
103-A	2.94	2.29	3.47	3.90	3.29	1.59
106	2.34	1.79	2.85	3.06	2.82	1.27
108-2	6.28	5.09	4.45	5.77	4.97	4.55
108-A	3.12	2.48	3.78	3.99	3.32	1.59
109	6.51	5.02	6.33	6.25	4.48	2.02
115-2	4.85	3.79	5.48	5.55	4.45	1.91
115-A	4.25	3.10	4.87	4.82	3.92	1.67
119	2.63	1.95	3.07	3.55	3.16	1.40
120	3.03	2.35	3.48	3.74	3.19	1.53
121	4.23	3.01	4.73	4.92	4.12	1.75
121-A	4.03	3.01	4.70	4.70	4.09	1.68
122	4.00	2.92	4.13	4.17	3.69	1.52
123	3.19	2.83	3.87	4.37	3.48	1.50
124	3.18	2.49	3.84	4.22	3.37	1.57
125	4.82	3.37	5.06	4.34	3.45	1.55
126	3.05	2.64	3.55	3.64	2.90	1.74
127	2.39	1.94	3.01	3.31	2.70	1.27

Table III.2 : Monthly Average Precipitations in Snowy Range Observatory

(continued)

Station	Month					
	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
101	2.50	1.52	1.47	1.63	2.29	2.49
102	2.51	1.48	1.86	2.18	3.31	3.26
103-1	2.71	1.85	1.82	2.47	3.83	3.56
103-2	2.70	1.90	1.90	2.61	4.19	3.87
103-A	2.61	1.75	1.84	2.29	3.41	3.21
106	2.49	1.56	1.63	1.95	2.69	2.79
108-2	2.27	2.92	2.19	2.32	3.29	5.00
108-A	2.81	2.00	1.96	2.47	3.28	3.62
109	3.02	2.23	2.39	3.23	5.06	5.81
115-2	2.83	2.18	2.09	3.31	4.78	5.47
115-A	2.70	1.96	1.87	2.75	4.01	5.03
119	2.46	1.58	1.52	2.15	2.97	2.84
120	2.58	1.64	1.63	2.24	3.58	3.55
121	2.87	1.97	1.91	2.78	4.46	4.56
121-A	2.57	1.93	1.89	2.76	4.34	4.44
122	2.77	2.00	2.09	2.75	3.63	4.17
123	2.51	1.77	1.77	2.63	3.82	3.72
124	2.49	1.74	1.77	2.53	3.48	3.61
125	2.53	1.86	1.80	2.76	4.25	4.76
126	2.70	2.26	1.94	2.53	3.21	2.85
127	3.30	1.65	1.64	2.11	3.28	2.82

Table III.2 : Monthly Average Precipitations in Snowy Range Observatory

Station	Month					
	Jan.	Feb.	Mar.	Apr.	May.	Jun.
101	1.32	1.18	0.72	1.33	0.94	0.84
102	1.84	1.35	1.33	1.95	1.49	0.88
103-1	2.04	1.61	1.46	2.28	1.46	1.05
103-2	2.28	1.75	1.62	2.58	1.66	1.07
103-A	2.46	1.38	1.24	2.07	1.46	1.06
106	1.57	1.31	1.36	1.61	1.13	0.90
108-2	4.46	3.55	1.96	2.18	1.71	1.74
108-A	1.94	1.35	1.36	1.69	1.39	1.04
109	4.52	3.43	3.24	2.92	1.65	0.99
115-2	3.07	2.20	1.94	2.58	1.75	1.34
115-A	2.48	1.47	2.01	2.23	1.52	1.02
119	1.81	1.33	1.25	1.91	1.46	0.92
120	2.18	1.87	1.63	2.29	1.50	0.98
121	2.67	1.55	1.72	2.47	1.78	0.94
121-A	2.72	1.84	1.78	2.34	1.97	1.02
122	2.18	1.76	1.75	1.92	1.77	1.00
123	2.32	1.96	1.53	2.19	1.48	0.93
124	2.19	1.48	1.64	2.03	1.43	0.98
125	4.11	2.10	2.40	2.11	1.51	1.04
126	1.72	1.59	1.62	1.28	1.15	1.00
127	1.77	1.44	1.42	2.02	1.52	0.97

Table III.3 : Standard Error of Monthly Precipitation in Snowy Range Observatory

(continued)

Station	Month					
	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
101	1.50	0.65	0.94	0.68	1.46	1.63
102	1.83	0.75	1.14	0.81	1.92	1.86
103-1	1.61	0.93	1.05	0.87	2.44	2.44
103-2	1.56	0.97	1.14	0.89	2.51	2.42
103-A	1.53	1.03	0.94	0.68	1.99	1.84
106	1.72	0.78	0.97	0.71	1.52	1.57
108-2	1.43	1.30	0.96	1.43	1.04	2.75
108-A	1.42	0.90	1.05	0.59	1.76	2.13
109	1.63	0.94	1.23	1.40	2.25	3.24
115-2	1.74	1.02	1.18	1.04	2.65	3.28
115-A	1.65	0.90	0.99	0.79	2.43	2.94
119	1.46	0.78	1.12	0.71	2.00	1.67
120	1.57	0.86	1.08	0.89	2.34	2.30
121	1.70	0.89	1.17	0.91	2.42	2.56
121-A	1.82	0.97	1.19	0.89	2.28	2.66
122	1.47	0.95	1.15	0.72	1.74	2.50
123	1.40	0.93	1.18	0.87	2.32	2.32
124	1.36	1.81	1.04	0.76	2.06	2.12
125	1.45	0.98	2.11	0.69	2.12	2.40
126	1.41	1.22	0.92	0.63	1.17	1.61
127	3.10	0.99	1.13	0.95	2.23	1.83

Table III.3 : Standard Error of Monthly Precipitation in Snowy Range Observatory

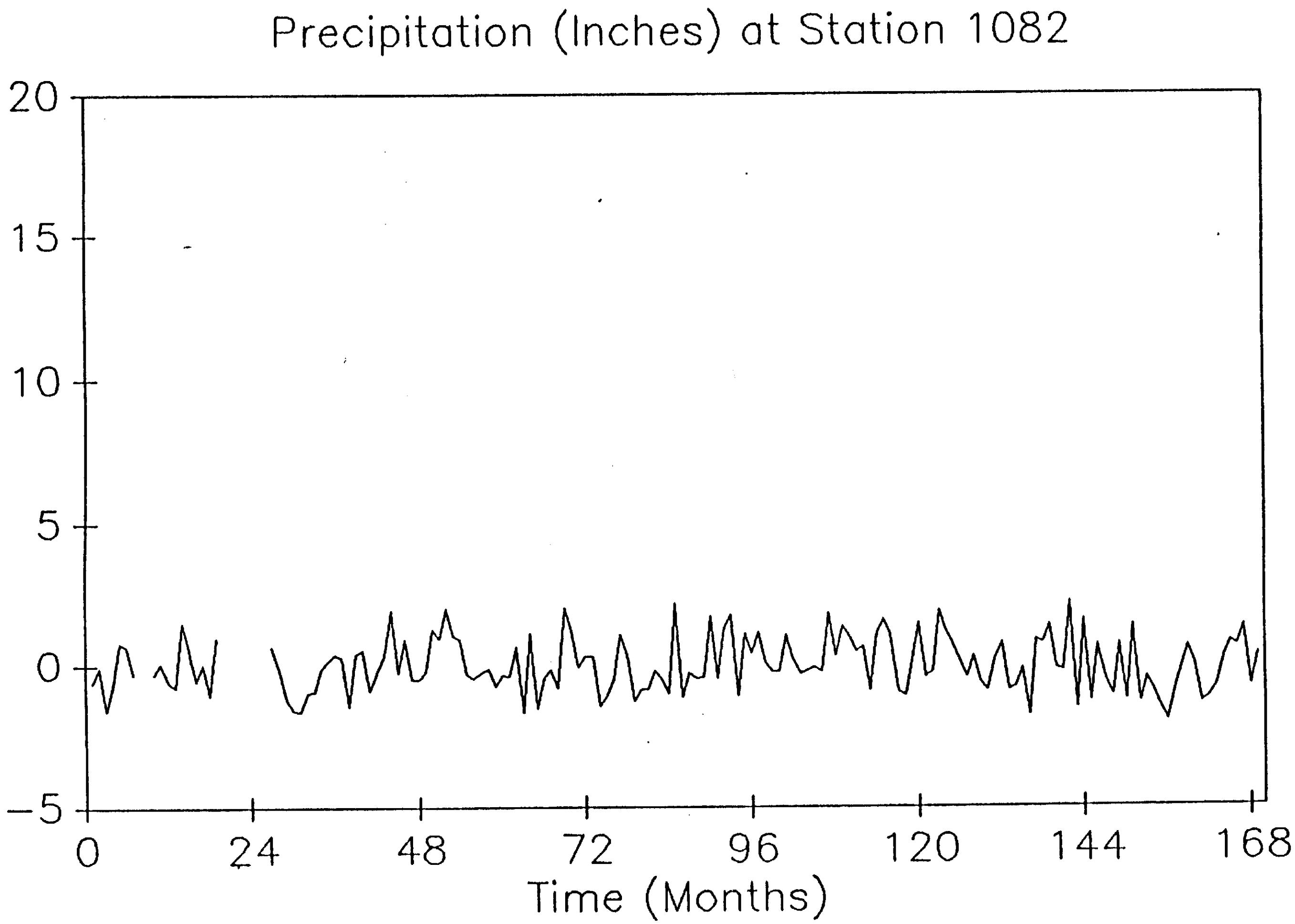


Figure III.8 : Standardized Time Series

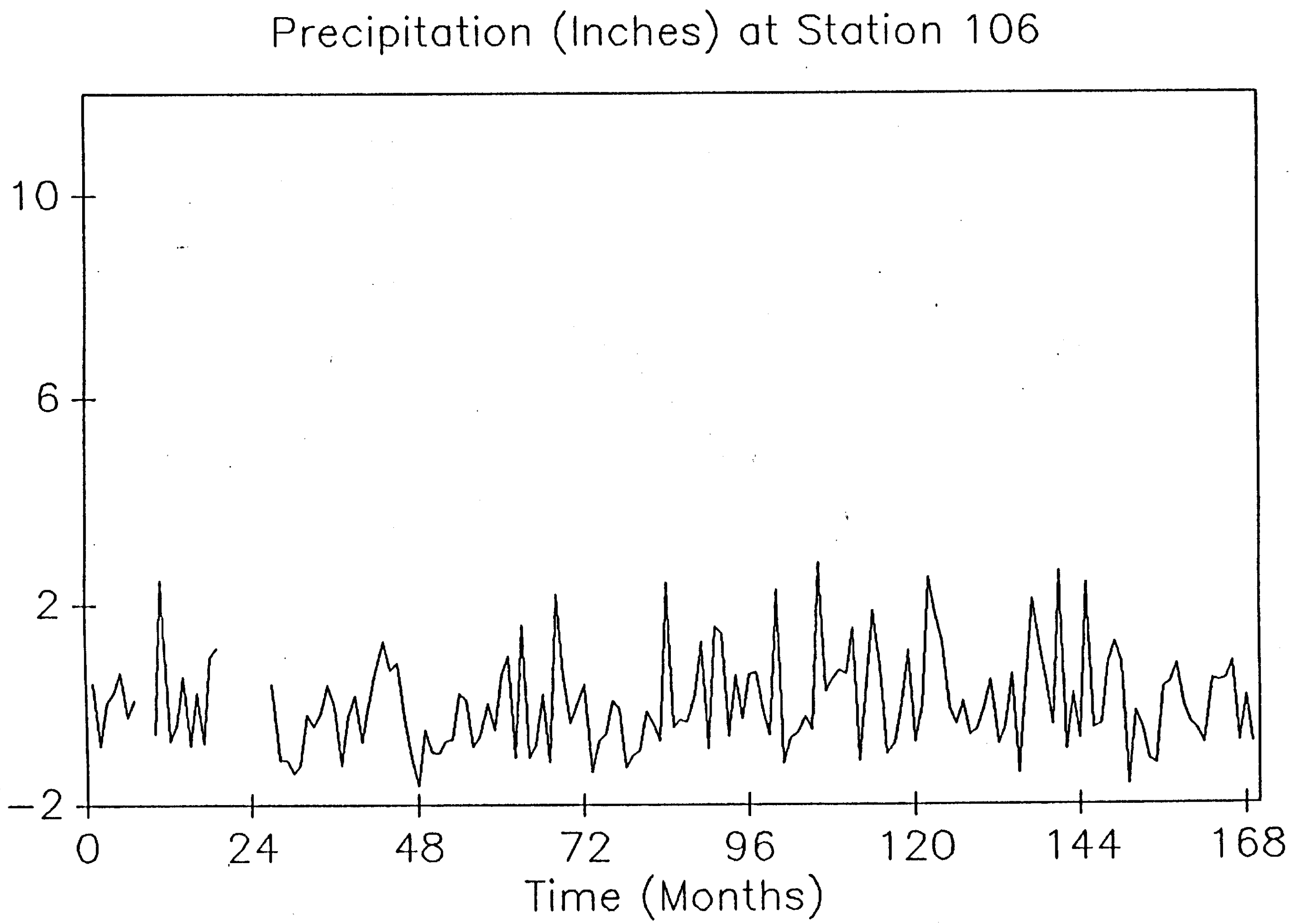


Figure III.9 : Standardized Time Series

sample means. The Fourier series

$$\overline{X}_F(m) = \overline{\overline{X}} + \sum_{f=1}^{6} [\alpha_f \cos(\frac{2\pi fm}{12}) + \beta_f \sin(\frac{2\pi fm}{12})] \qquad \text{(III.5)}$$

is an approach to find the "hidden periodicity" in a time series in which $\overline{X}_F(m)$ is the fitted sample mean by a Fourier series in month m, $\overline{\overline{X}}$ is an overall mean that is computed by

$$\overline{\overline{X}} = \frac{1}{T}\sum_{y=1}^{Y}\sum_{m=1}^{12} X(m,y) \qquad \text{(III.6)}$$

in which T is the total months in the time series used and α_f and β_f are Fourier coefficients that can be computed as

$$\alpha_f = \frac{1}{6}\sum_{m=1}^{12} \overline{X}(m) \cos\left(\frac{2\pi fm}{12}\right) \quad , f = 1,2,\ldots,6 \qquad \text{(III.7)}$$

$$\beta_f = \frac{1}{6}\sum_{m=1}^{6} \overline{X}(m) \sin\left(\frac{2\pi fm}{12}\right) \quad , f = 1,2,\ldots,6 \qquad \text{(III.8)}$$

It should be noted that not all Fourier coefficients are statistically significant. To test their significance for each frequency f, a test statistic v_f (Parzen, 1981)

$$v_f = \frac{T\ (\alpha_f^2 + \beta_f^2)}{\overline{S}} \qquad \text{(III.9)}$$

is used in which $\overline{S}$ is the average of the monthly sample standard deviation of the monthly precipitation that can be
computed by

$$S = \frac{1}{12}\sum_{m=1}^{12} S(m) \qquad \text{(III.10)}$$

The test statistic v_f has a χ^2 distribution with 2 degree of freedom. If the statistic $v_f > 12.0$, for f=1,2,..,5 and $v_f > 3.84$ for f=6, the coefficients at the corresponding frequency can be accepted as significant and thus the cycles of the period 12/f ,for f=1,2,..,6, which correspond to 12, 6, 4, 3, 2.4 and 2 months are also significant. Then the seasonal means can be estimated using those significant coefficients. The estimated coefficients are presented in Tables III.4 and III.5. These tables also show the results of significant tests of the Fourier coefficients. By the results of this analysis, the frequencies f=1, 2, 3 and 6 are mostly significant throughout all 21 stations. The corresponding coefficients are used to estimate the mean precipitation for each month. Using the results from Fourier analysis, the fitted monthly mean, $\overline{X}_F(m)$, is used to transform the detrended time series into a monthly stationary series

$$Z_F(m,y) = \frac{X(m,y) - \overline{X}_F(m)}{S(m)} \qquad \text{(III.11)}$$

Station	α_f					
	f=1	f=2	f=3	f=4	f=5	f=6
101	0.07	0.24	-.20	-.06	0.19	0.36
102	-.03	0.43	-.31	-.09	0.11	0.40
103-1	-.01	0.52	-.34	-.01	0.10	0.41
103-2	0.01	0.56	-.32	0.03	0.11	0.51
103-A	0.09	0.45	-.24	-.03	0.09	0.42
106	0.08	0.34	-.21	-.06	0.19	0.40
108-2	0.57	0.25	0.57	-.02	-.30	-.36
108-A	0.06	0.35	-.25	-.10	0.21	0.35
109	-.17	0.30	-.17	-.17	0.05	0.54
115-2	-.13	0.62	-.28	-.09	0.25	0.38
115-A	-.08	0.52	-.26	-.16	0.29	0.38
119	0.14	0.45	-.20	-.03	0.15	0.39
120	-.02	0.49	-.30	-.02	0.13	0.41
121	-.03	0.62	-.30	-.05	0.16	0.56
121-A	-.05	0.61	-.27	-.00	0.19	0.52
122	-.10	0.37	-.12	-.09	0.25	0.47
123	-.01	0.52	-.29	0.02	0.13	0.31
124	0.05	0.50	-.25	-.09	0.15	0.33
125	-.29	0.34	-.29	-.15	0.18	0.54
126	0.05	0.16	-.26	0.30	0.05	0.28
127	0.04	0.33	-.47	-.06	0.14	0.54

Table III.4 : Estimates of Fourier Coefficients, α_f

Station	β_f					
	f=1	f=2	f=3	f=4	f=5	f=6
101	-.23	0.05	-.35*	-.16	-.12	-.00*
102	-.60*	-.17*	-.42*	-.14	-.16	-.00*
103-1	-.70*	-.16*	-.50*	-.23	-.07	-.00*
103-2	-.95*	-.20*	-.51*	-.22	-.02	-.00*
103-A	-.64*	-.17*	-.45*	-.23	-.08	-.00*
106	-.41*	-.09*	-.40*	-.19	-.10	-.00*
108-2	-1.60*	0.49*	-.29*	-.07	-.22	-.00*
108-A	-.68*	-.15*	-.51*	-.21	-.07	-.00*
109	-2.10*	-.12	-.67*	-.36	0.04	-.00*
115-2	-1.53*	-.20*	-.70*	-.23	0.03	-.00*
115-A	-1.30*	-.19*	-.72*	-.16	0.02	-.00*
119	-.55*	-.14*	-.41*	-.28	-.05	-.00*
120	-.74*	-.06*	-.47*	-.20	-.08	-.00*
121	-1.19*	-.13*	-.65*	-.27	0.02	-.00*
121-A	-1.19*	-.19*	-.60*	-.20	0.04	-.00*
122	-.96*	-.10*	-.52*	-.26	-.02	-.00*
123	-.91*	-.258	-.48*	-.27	-.11	-.00*
124	-.84*	-.22*	-.50*	-.23	-.04	-.00*
125	-.15*	-.04	-.55*	-.19	0.15	-.00*
126	-.53*	-.19	-.31*	-.27	0.03	-.00*
127	-.35	-.04	-.41*	-.34	-.21	-.00*

Table III.5 : Estimates of Fourier Coefficients, β_f
(* represents v_f is significant)

CHAPTER IV

ANALYSIS OF SPATIALLY NONSTATIONARY MONTHLY PRECIPITATION

In Chapter III, the monthly total precipitation data under study shows the nonstationarity over the study area (see Table III.2 and III.3) indicated by that the mean and the variance are not constant. It is, therefore, necessary to have a mechanism to estimate the mean as well as the variance in space so that the standardized stationary random field can be transformed back to preserve the original spatial nonstationary characteristics. Firstly, two different methods that do not consider spatial correlation were used to estimate the monthly average precipitation over the space. The inverse distance weighing technique and the Gaussian smoothing method were used. Secondly, to consider the spatial correlation between stations at Snowy Range Observatory, the nonstationary kriging technique was used to estimate the monthly total precipitation over the space.

IV.1 Estimation without Considering Spatial Correlation

The inverse distance weighing technique and the Gaussian smoothing weighing method use a weighted average of the values from the recording stations to estimate the monthly average precipitation at the specified ungaged location. The weights computed by inverse distance or Gaussian smoothing techniques are only functions of the physical distance between two stations. They do not explicitly account for the spatial variability of monthly precipitation under study. These two techniques also do not allow for computation of the reliability of the estimates.

IV.1.1 Inverse Distance Method

The linear inverse distance weighing technique has the same concept as the IDLIN technique in Chapter II used for estimating the missing values in which the contributing weight of the existing station to ungaged location is determined as

$$w_{s,IV} = \frac{\left(\frac{1}{D_s}\right)}{\sum_{i=1}^{S}\left(\frac{1}{D_i}\right)} \qquad \text{(IV.1)}$$

where D_s is the distance between the specified ungaged location and a recording station s, S is the total number of stations used in estimation, and $w_{s,IV}$ is the weight for the station s by inverse distance method. Then, the monthly average precipitation at the ungaged location, $\overline{X}_{0,IV}(m)$, for month m, is estimated by

$$\overline{X}_{0,IV}(m) = \sum_{s=1}^{S} w_{s,IV} \ \overline{X}_s(m) \tag{IV.2}$$

where $\overline{X}_s(m)$ is the calculated total monthly average precipitation at the recording station s in a given month m. The method does not consider any other factors affecting the estimate except the distances between two locations. This technique also can be employed to estimate the variance of monthly precipitation at the ungaged location in a similar manner as

$$\hat{S}_{0,IV}^2(m) = \sum_{s=1}^{S} w_{s,IV} \ S_s^2(m) \tag{IV.3}$$

in which $\hat{S}_{0,IV}^2(m)$ and $S_s^2(m)$ are the variances at the ungaged location and the recording station s, respectively.

IV.1.2 Gaussian Smoothing Method

The Gaussian smoothing technique (Borgman, 1990; Kallianpul, et al., 1988) uses the Gaussian function

$$f(D_s) = e^{-\frac{1}{2Q} D_s^2} \tag{IV.4}$$

to compute the weights where Q is a constant representing the effective width of smoothing. The weights are determined by

$$w_{s,GS} = \frac{f(D_s)}{\sum_{i=1}^{S} f(D_i)} \quad , \ s = 1,2,\ldots,S \tag{IV.5}$$

The value of Q can be determined subjectively with the idea that the larger the value of Q is, the more smoothing the result will be. A very small value of Q allows only those recording stations in close vicinity of the estimation point to be used to estimate the monthly precipitation at the ungaged location. As can be seen, the weight also depends only on the distance. The estimates of monthly average precipitation can be determined by

$$\overline{X}_{0,GS}(m) = \sum_{s=1}^{S} w_{s,GS} \ \overline{X}_s(m) \tag{IV.6}$$

In this study, Fibonacci search technique (Luenberger, 1984) is applied to find the optimal Q which minimizes the error associated with the estimates using the observed data at all recording stations. The objective function is to minimize

$$\varepsilon^2(m|Q) = \sum_{s=1}^{S} \left[\overline{X}_s(m) - \hat{\overline{X}}_{s,GS}(m|Q) \right]^2 \qquad \text{(IV.7)}$$

in which $\hat{\overline{X}}_{s,GS}(m)$ is the estimates of monthly average precipitation resulting from Gaussian smoothing weights. This process can also be applied to estimate the variance, $S^2_s(m)$. The optimal Q's for estimating $\overline{X}_s(m)$ and $S^2_s(m)$ for each month, m=1,2,...,12, by the Fibonacci search are shown in Table IV.1.

IV.2 Estimation considering Spatial Correlation (Kriging)

Unlike the inverse distance weighing technique and Gaussian smoothing weighing method, a geostatistical kriging method is based on the structure of the spatial variability of monthly total precipitation (Royle et al., 1980; Borgman, 1990; Bras et al., 1985). The measured monthly total precipitation in Snowy Range watershed allows one to analyze the spatial correlation structure in the study area and to incorporate such information to determine the optimum weight which minimizes the variance estimation. The kriging procedure has been developed around the variogram or covariance function and the best linear unbiased estimation.

IV.2.1 Analysis of Spatial Correlation Structure

The computation of the correlation matrix is essential in kriging to find the variogram or covariance structure. Computation of the spatial correlation structure requires that the random field is stationary. Therefore, the standardized monthly total precipitation series described in Chapter III were used. Since the standardization requires estimating the means and variances of the monthly total precipitation, the above Gaussian smoothing method and two other techniques were used. One technique was simply to use sample means and sample variances at the recording stations and the other used the monthly total precipitation statistics calculated by the Fourier Series described in Chapter III.

Using the standardized monthly total precipitation series, the functional relationship between correlations and distances computed by the three different methods are plotted in Figures IV.1, IV.2 and IV.3. All three plots are similar thus indicating that the correlation-distance relationship is not sensitive to the method used to compute the local mean and variance of the monthly total precipitation in Snowy Range watershed. It is interesting to observe that the correlation function is separated into two distinct parts. One, which locates at the lower part, parallel to

Month	Optimal Q	
	For Mean	For Variance
January	0.3819663	0.5280327
February	0.3068254	0.2938475
March	0.2755167	0.7782433
April	0.3068254	0.4530374
May	0.5197246	0.1029390
June	0.5322481	0.3942531
July	0.2003758	0.8325485
August	0.3318724	0.5290438
September	0.2755167	0.5932542
October	0.1565436	0.2842049
November	0.2191610	0.3290381
December	0.4383220	0.2374398

Table IV.1 : Optimal Q for Gaussian Smoothing

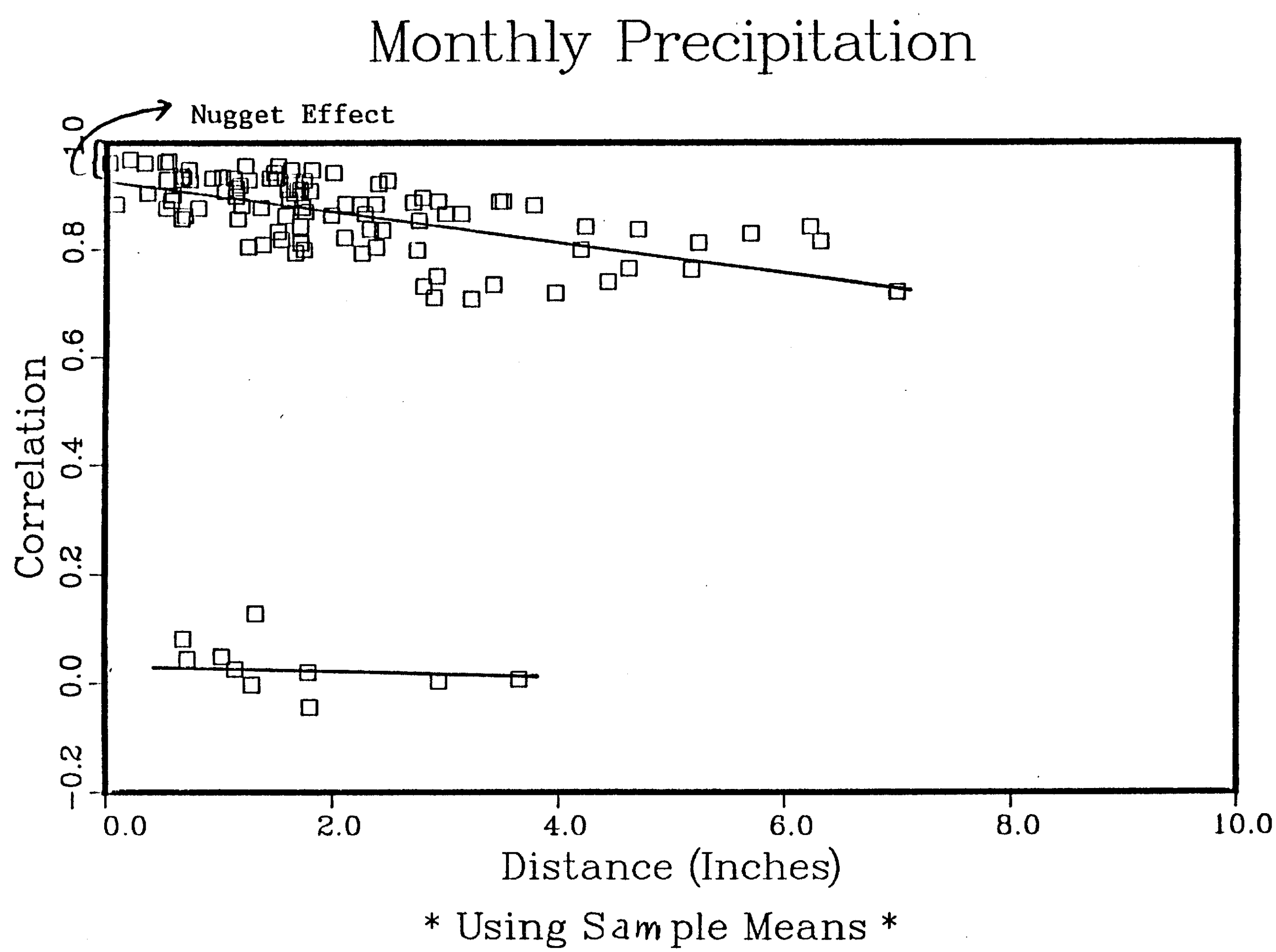


Figure IV.1 : Correlation Structure from using Sample Means

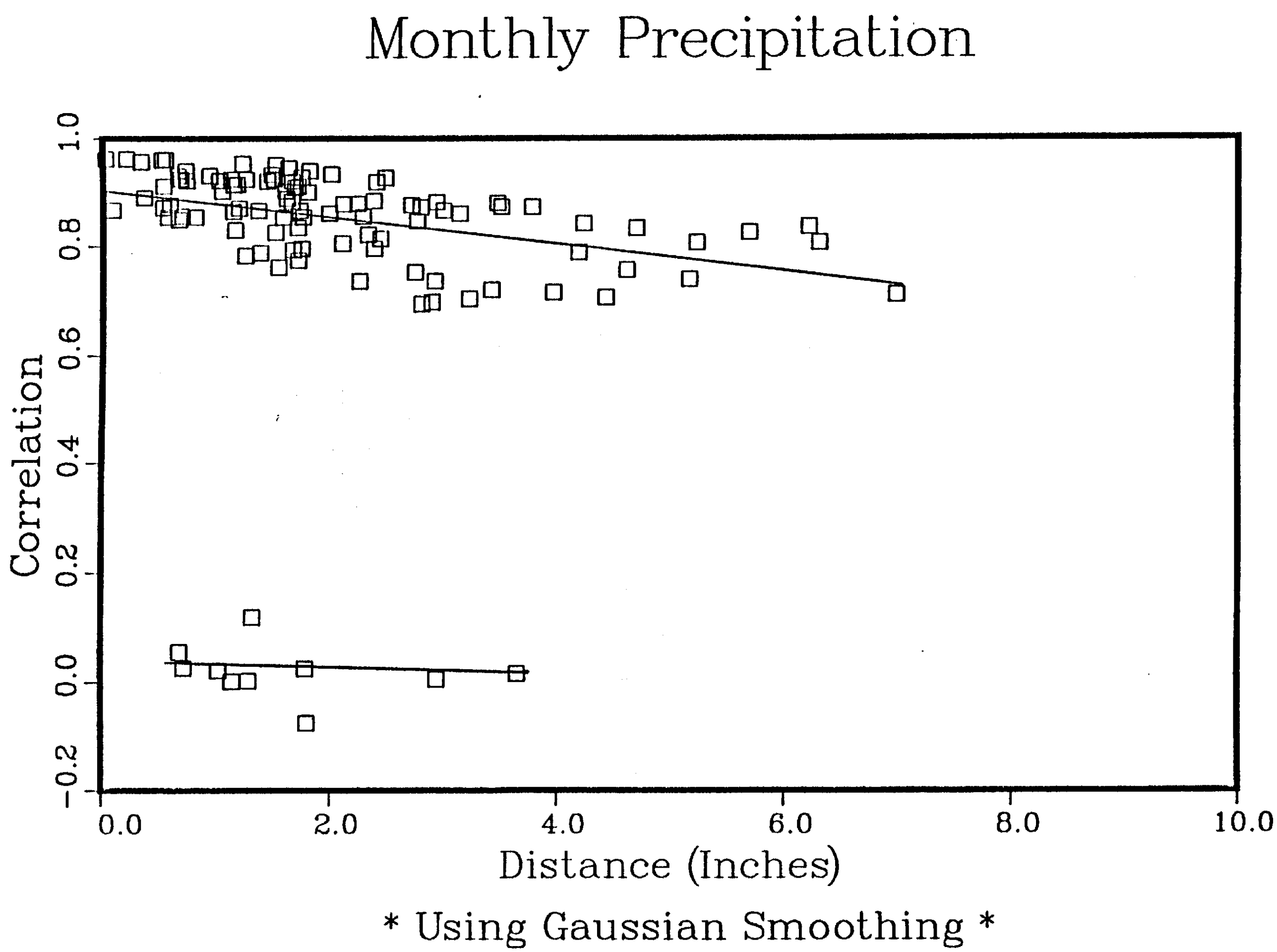


Figure IV.2 : Correlation Structure from using Gaussian Smoothing

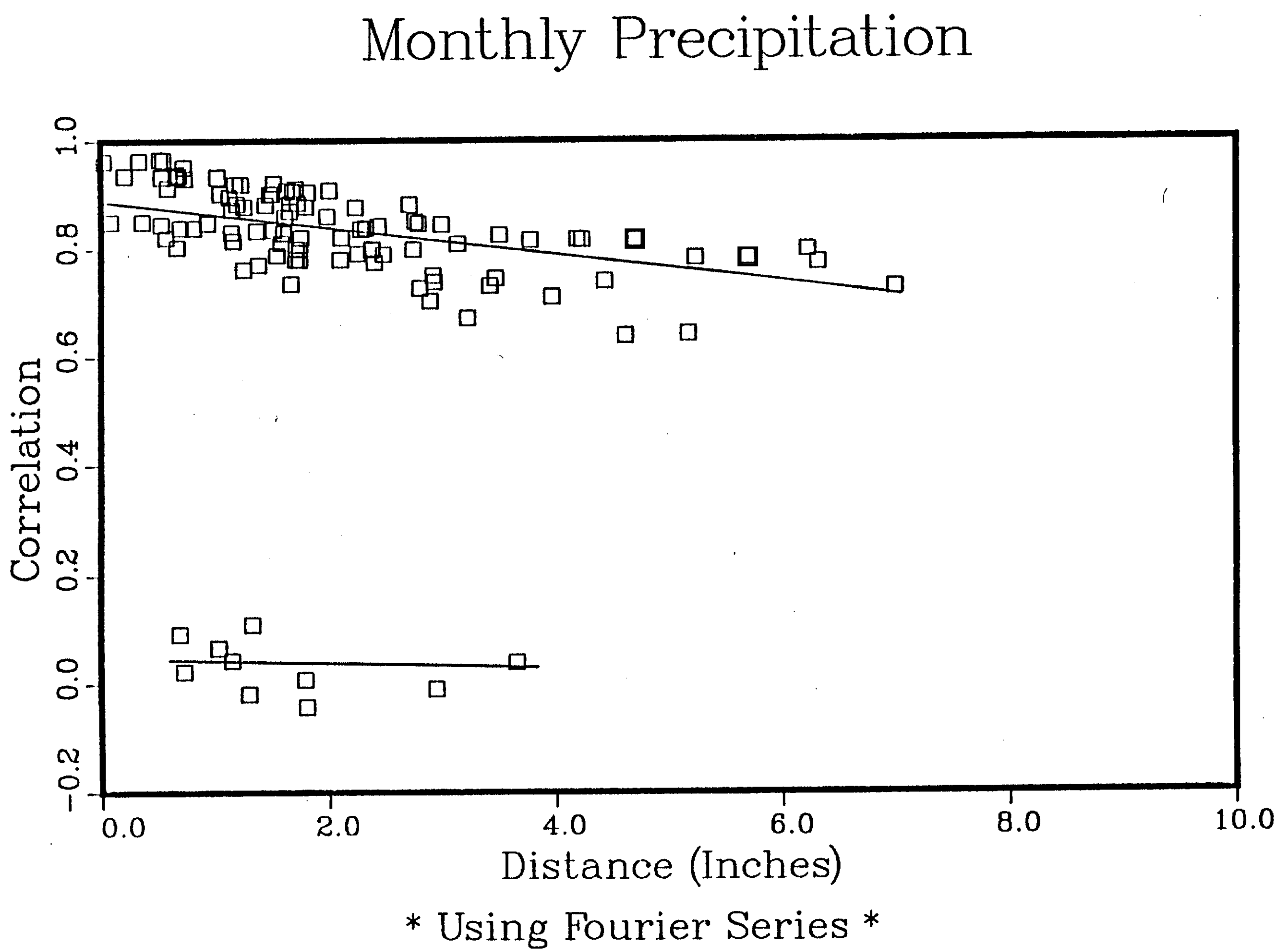


Figure IV.3 : Correlation Structure from using Fourier Series the horizontal axis indicating that the correlation associated these points is a constant against the distances. This part consists of 20 points and all these points are related to station 108-2. This implies that the monthly precipitation collected by station 108-2 is independent of the catch by all other stations. This lower part associated with station 108-2 is assumed to be a constant having the correlation value of the average of the total 20 points.

The upper part of the plot reveals an inverse linear relationship with the distances. This means that the closer the distance between two stations is, the higher the correlation is. These correlations were fitted by regression to find the structure for the three different methods. The resulting regression equations that define the functional relationship for correlation between station s and s', $\rho(d_{s,s'})$ is as follows: when station 108-2 is considered to compute correlation, for all three techniques

$$\hat{\rho}(d_{s,s'}) = constant \qquad \text{(IV.8)}$$

When all other stations are used to calculate the correlation, the resulting regression equation, using sample means and variances, is

$$\hat{\rho}(d_{s,s'}) = 0.928 - 0.0263 d_{s,s'} \quad , R^2 = 36.6\% \qquad \text{(IV.9)}$$

When using the Gaussian smoothing method the regression equation is

$$\hat{\rho}(d_{s,s'}) = 0.914 - 0.0257 d_{s,s'} \quad , R^2 = 30.0\% \qquad \text{(IV.10)}$$

and using the Fourier Series it is

$$\hat{\rho}(d_{s,s'}) = 0.901 - 0.0291 d_{s,s'} \quad , R^2 = 38.4\% \qquad \text{(IV.11)}$$

The resulting regression lines from the three different techniques are not distinguishable. Hence the correlation relation computed by the sample means and the sample variances is adapted in the further study. It should also be noted that there exists a nugget (random) effect (indicated in Figure IV.1) in these plots. The correlation function used in the nonstationary kriging for this study is

$$\hat{\rho}(d_{s,s'}) = \begin{cases} 0.028, \; d_{s,s'} > 0, & \text{(IV.12)} \\ \qquad\qquad \text{for } s' = \text{station 108-2 and} & \\ 0.928 - 0.0263 d_{s,s'}, \; d_{s,s'} > 0, & \text{(IV.13)} \\ \qquad\qquad \text{for } s' \neq \text{station 108-2} & \\ 1, \; d_{s,s'} = 0 \qquad \text{for all } s, s' & \text{(IV.14)} \end{cases}$$

From the above correlation function, the covariance function can be computed as

$$C(d_{s,s'}) = C(s,s') = \sigma(s)\,\sigma(s')\,\rho(d_{s,s'}) \qquad \text{(IV.15)}$$

where $C(d_{s,s'})$ and $\rho(d_{s,s'})$ are the covariance and the correlation between two stations s and s' separated by distance $d_{s,s'}$, respectively.

IV.2.2 Nonstationary Kriging

Nonstationary kriging proposes that the monthly total precipitation value over the specified block or point can be estimated by the known functions, $X_{m,s}$ for s=1,2,..,S, as

$$\hat{X}_{m,s_0} = \sum_{s=1}^{S} w_s\, X_{m,s} \qquad \text{(IV.16)}$$

where S is the total number of recording precipitation stations in the region. The basic function, $X_{m,s}$, is known, but the weighing coefficients, w_s, are to be determined.

IV.2.2.1 Definition

The mean of the time-space random process, $X_{m,s}$, which is assumed to be second-order and nonstationary, is defined as

$$\mu_{m,s} = \mathrm{E}[X_{m,s}] \qquad \text{(IV.17)}$$

where s represents a location in a region and m is a given month (Borgman, 1990). The variance can also be expressed as

$$\sigma_{m,s}^{2} = \mathrm{E}[(X_{m,s} - \mu_{m,s})^2] \qquad \text{(IV.18)}$$

The covariance allows modeling of the spatial dependence of the random process $X_{m,s}$. The covariance function at any two locations, s and s' for month m, is defined as

$$C_m(s,s') = \mathrm{E}\left[(X_{m,s}-\mu_{m,s})^T (X_{m,s'}-\mu_{m,s'})\right] \qquad \text{(IV.19)}$$

where s=1,2,...,S and s'=1,2,...,S. If s=s', $C_m(s,s')$ is called autocovariance and if not, $C_m(s,s')$ is called cross-covariance.

The objective of nonstationary kriging is to estimate the monthly total precipitation over the specified area B in a given month m as

$$\hat{X}_{Bm} = \frac{1}{B}\int_B X_{m,s}\, ds \quad , m = 1,2,..,12 \qquad \text{(IV.20)}$$

When the block B reduces to a point, then

$$\hat{X}_{Bm} = X_{m,s_0} \qquad \text{(IV.21)}$$

To construct the contour maps of the mean and the error, the use of nonstationary point estimation is preferred. The expectation of point kriging is

$$\mu_{m,s_0} = \mathrm{E}[X_{m,s_0}] \quad ,m=1,2,..,12 \qquad \text{(IV.22)}$$

IV.2.2.2 Formulation of Nonstationary Kriging Model

The point monthly total precipitation in month m at the estimation point s_0 can be expressed as a linear combination of the observed data, $X_{m,s}$, for s=1,2,..,S, by

$$\hat{X}_{m,s_0} = \sum_{s=1}^{S} w_s\, X_{m,s} \qquad \text{(IV.23)}$$

The mean square error of monthly total precipitation at a given estimation point s_0, in a given month m, can be rewritten as

$$\sigma_E^2(m) = \mathrm{E}\left[(\hat{X}_{m,s_0} - X_{m,s_0})^2\right] = \mathrm{E}\left[\left(\sum_{s=1}^{S} w_s X_{m,s} - X_{m,s_0}\right)^2\right]$$
$$= \mathrm{E}\left[\left(\sum_{s=1}^{S} w_s (X_{m,s}-\mu_{m,s}) + \sum_{s=1}^{S} w_s \mu_{m,s} - (X_{m,s_0}-\mu_{m,s_0}) - \mu_{m,s_0}\right)^2\right]$$

Assume that the estimator is unbiased, that is,

$$\mathrm{E}[\hat{X}_{m,s_0} - X_{m,s_0}] = 0 \qquad \text{(IV.25)}$$

Then the mean square error, $\sigma_E^2(m)$, is reduced to

$$= \mathrm{E}\left[\left(\sum_{s=1}^{S} w_s (X_{m,s} - \mu_{m,s}) - (X_{m,s_0} - \mu_{m,s_0}) + (\sum_{s=1}^{S} w_s \mu_{m,s} - \mu_{m,s_0})\right)^2\right]$$

$$= \mathrm{E}\left[\left(\sum_{s=1}^{S} w_s (X_{m,s} - \mu_{m,s})\right)^2\right] + \mathrm{E}\left[(X_{m,s_0} - \mu_{m,s_0})^2\right]$$

$$-2\mathrm{E}\left[\sum_{s=1}^{S} w_s (X_{m,s} - \mu_{m,s}) (X_{m,s_0} - \mu_{m,s_0})\right]$$

$$\sigma_E^2(m) = \sum_{s=1}^{S} \sum_{s'=1}^{S} w_s w_{s'} C_m(s, s') + C_m(s_0, s_0) - 2\sum_{s=1}^{S} w_s C_m(s, s_0)$$

$$= \boldsymbol{w}^T \mathbf{C}_m \boldsymbol{w} + \sigma_{m,s_0}^2 - 2\boldsymbol{w}^T \boldsymbol{c}_{m,s_0}$$

The elements on the right-hand side of Eq.(IV.26) can be computed, using information extracted from the data, as

$$C_m(s, s') = s_{m,s}\, s_{m,s'}\, \hat{\rho}_m(d_{s,s'}) \quad , \qquad \text{(IV.27)}$$

$$\hat{\sigma}_{m,s_0}^2 = \hat{s}_{m,s_0}^2 = \sum_{s=1}^{S} w_{s,IV}\, s_{m,s}^2 \quad , \qquad \text{(IV.28)}$$

$$C_m(s, s_0) = s_{m,s}\, \hat{s}_{m,s_0}\, \hat{\rho}_m(d_{s,s_0}) \quad , \qquad \text{(IV.29)}$$

$$\mu_{m,s_0} = \sum_{s=1}^{S} w_{s,IV}\, \mu_{m,s} \qquad \text{(IV.30)}$$

in which $\hat{\rho}_m(s,s')$ and $\hat{\rho}_m(s,s_0)$ are the correlation estimated by the structured correlation function, Eq.(IV.12), (IV.13) or (IV.14), between stations s and s' and between ungaged point s_0 and station s in month m, respectively; $w_{s,IV}$ is the inverse distance weight contributed from the existing station s; $\mu_{m,s}$, $s_{m,s}$ and $\sigma_{m,s}^2$ are the mean, standard deviation, and variance measured at station s in month m, respectively.

After the above terms are computed, the nonstationary kriging model can be expressed as the following

(Model I)

Minimize

$$\sigma_E^2(m) = \mathbf{w}^T \mathbb{C}_m \mathbf{w} + \sigma_{m,s_0}^2 - 2\mathbf{w}^T \mathbf{c}_{m,s_0} \qquad \text{(IV.31a)}$$

subject to

$$\cancel{\text{(IV.31b)}} \quad \mathbf{w}^T \boldsymbol{\mu}_m = \mu_{m,s_0} \quad ,$$

$$\mathbf{w}^T \mathbf{1} = 1 \quad , \qquad \text{(IV.31c)}$$

$$w_s \geq 0 \quad , s=1,2,\ldots,S \qquad \text{(IV.31d)}$$

The commonly used kriging model involves only one constraint Eq.(IV.31b). The solution to such model do not ensure that all weighing factors are nonnegative and add up to one. Therefore, constraint Eqs.(IV.31c) and (IV.31d) are added to the kriging model.

In model I, the local variance, $\sigma_{m,s}^2$, is considered as a constant. However, it is actually a function of the variances measured at every existing station, $\sigma_{m,s}^2$'s, and the corresponding weights which are unknown. Therefore, $\sigma_{m,s}^2$ should be considered as a decision variable and model I is modified as the following

(Model II)

Minimize

$$\sigma_E^2(m) = \mathbf{w}^T \mathbb{C}_m \mathbf{w} + \sigma_{m,s_0}^2 - 2\mathbf{w}^T \mathbf{c}_{m,s_0}$$

$$= \mathbf{w}^T \mathbb{C}_m \mathbf{w} + \sigma_{m,s_0}^2 - 2\mathbf{w}^T [\sigma_{m,s_0} s_{m,s} \hat{\rho}_m(d_{s,s_0})] \qquad \text{(IV.32a)}$$

subject to

$$\mathbf{w}^T \boldsymbol{\mu}_m = \mu_{m,s_0} \quad , \qquad \text{(IV.32b)}$$

$$\mathbf{w}^T \mathbf{1} = 1 \quad , \qquad \text{(IV.32d)}$$

$$\mathbf{w}^T \sigma_m{}^2 = \sigma^2_{m,s_0} \quad , \tag{IV.32c}$$

$$w_s \geq 0 \quad , \; s = 1,2,\ldots,S \tag{IV.32e}$$

in which $\sigma_{m,s}{}^2$ is an additional decision variable representing unknown variance at the estimation point s_0.

IV.2.2.3 Solution Procedure

The Lagrange multiplier method has been used to solve the conventional kriging model subject to only a single unbiased constraint (Borgman, 1990; Journel et al., 1978). However, the Lagrange multiplier method is inefficient for solving models I and II formulated above, which involves several constraints, including non-negativity constraints. In this study, a nonlinear programming code, called GRG2 (Lasdon et al., 1986), is used which is based on the generalized reduced gradient algorithm (Lasdon, 1979; Murtagh, 1978)

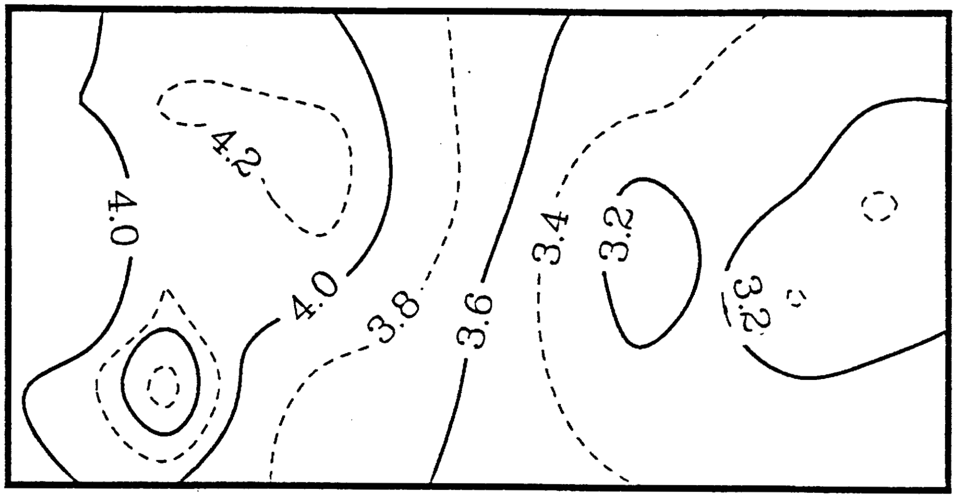
IV.3 Comparison Study

In this section, the results of contour maps for means and errors obtained by the different methods for each month are compared.

IV.3.1 <u>Comparison between Nonstationary Kriging Models I and II</u>

Figure IV.4 shows the mean precipitation contour maps for January derived by nonstationary kriging models I and II. It appears that the two nonstationary kriging models do not yield a significant difference between the mean monthly precipitation with regard to its spatial distribution. The contour maps (see Figure IV.5) for the differences in means are almost flat at zero level. Slightly greater value of the mean precipitations are estimated by model II around stations 109 and 108-2, which are located in the lower-left corner of the study region. Model I, however, has the larger error than model II for every month except August (see Figure IV.6). The error patterns between the two models are distinctly different at the lower-center area (see Figure IV.5) or at the lower-left area (Appendix A). The contour maps for the remaining months are shown in Appendix A.

Model I, Mean



Model II, Mean

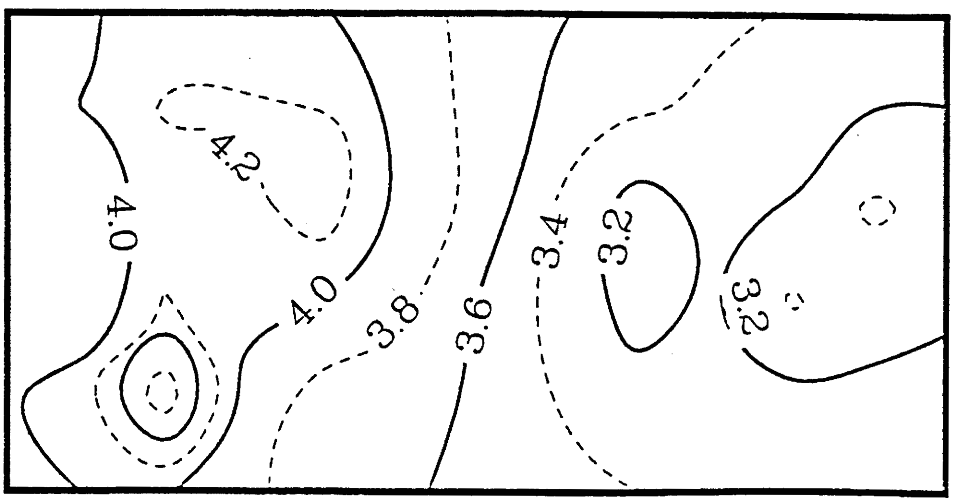


Figure IV.4 Comparison of Mean Contours from Model I & II for January

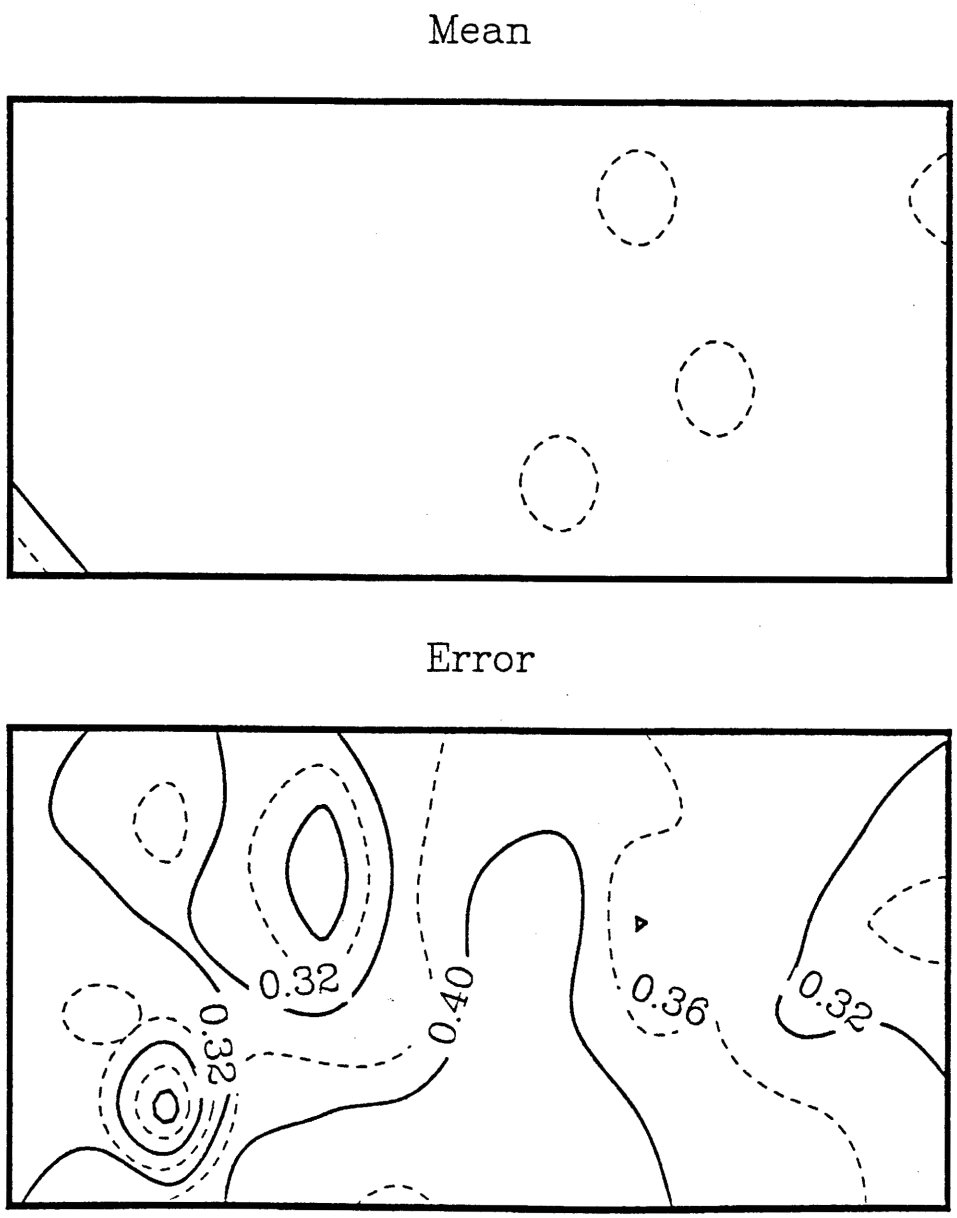
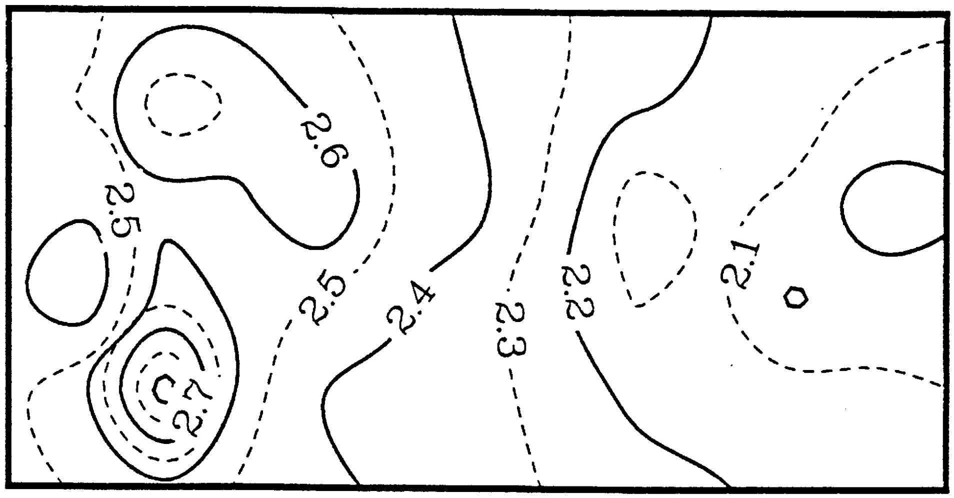


Figure IV.5 Contours of Difference between Model I & II for January

Model I, Error



Model II, Error

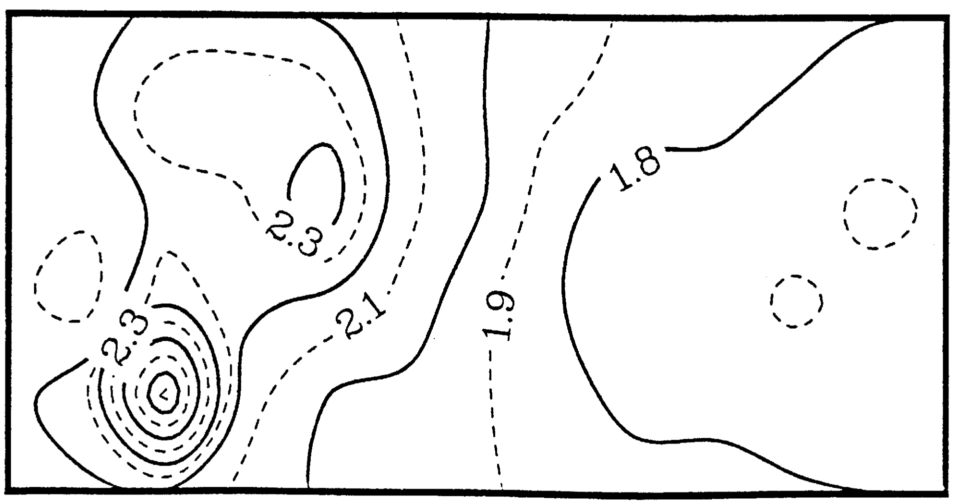


Figure IV.6 Comparison of Error Contours from Model I & II for January

IV.3.2 Comparison of Kriging Models, Inverse Distance and Gaussian Smoothing

It is interesting to compare the contour maps resulting from the Gaussian smoothing weighing method, the inverse distance weighing technique and the kriging method. The contour for the means and the errors estimated by the Gaussian smoothing and the inverse distance methods for January are drawn in Figures IV.7 and IV.8, respectively. The mean precipitation pattern computed by the inverse distance method is quite similar to that computed by the two nonstationary kriging methods. However, the error estimated by the inverse distance technique is always greater over the entire study region than the error estimated by kriging procedures.

The means and the errors obtained by the Gaussian smoothing method vary greatly over the basin compared to those by the other two methods. The Gaussian smoothing technique yields larger values of monthly precipitation at the left area of the watershed than the others and smaller values in the right part.

IV.3.3 Discussion

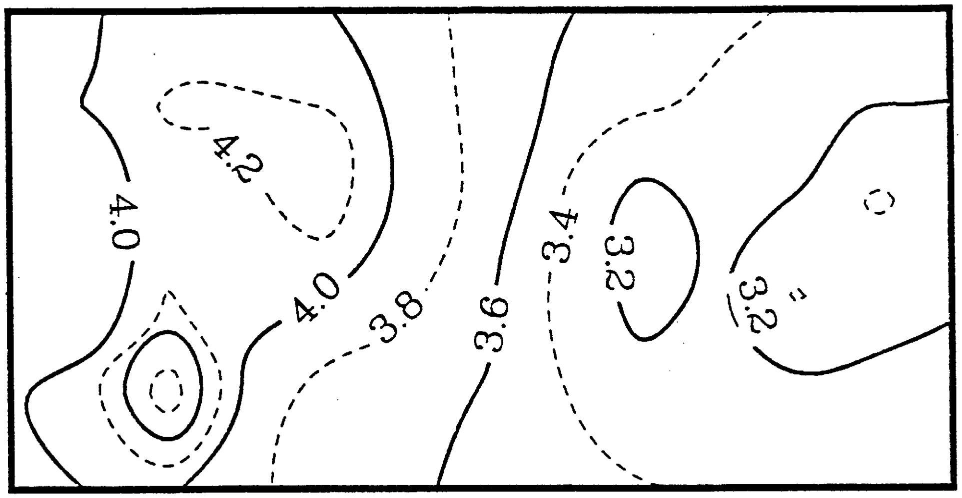
Models I and II yield almost identical spatial patterns of mean monthly total precipitation, but model I has a larger error in estimation. The reason for model II having smaller error is probably due to the fact that model II treats local variance as a decision variable. In model I, the local variance is considered as a constant computed by the initial weights, $w_{s,IV}$. The nonstationary kriging model I does not allow the local variance to use updated optimum weights.

The contour maps for the mean monthly total precipitation obtained by the inverse distance method have a similar pattern to those obtained by the nonstationary kriging. It, however, has larger errors than the kriging results (see Figure IV.7). This might be because the kriging is an optimization procedure which minimizes the estimation error while the inverse distance method does not.

The Gaussian smoothing method yields larger mean monthly total precipitations and errors than all other methods at the left portion of the study area and the lowest values at the right (see Figure IV.8). Referring to Table IV.1, the optimal Q's representing the effective width of smoothing, for the means and the variances have relatively small values less than 1. However, a lot of smoothing is done for estimating the monthly precipitations and the errors (see Figure IV.8).

This comparison study indicates that the kriging method is superior which could be attributed to the fact that the kriging procedure considers the spatial variability while the Gaussian smoothing and the inverse distance methods do not.

Mean



Error

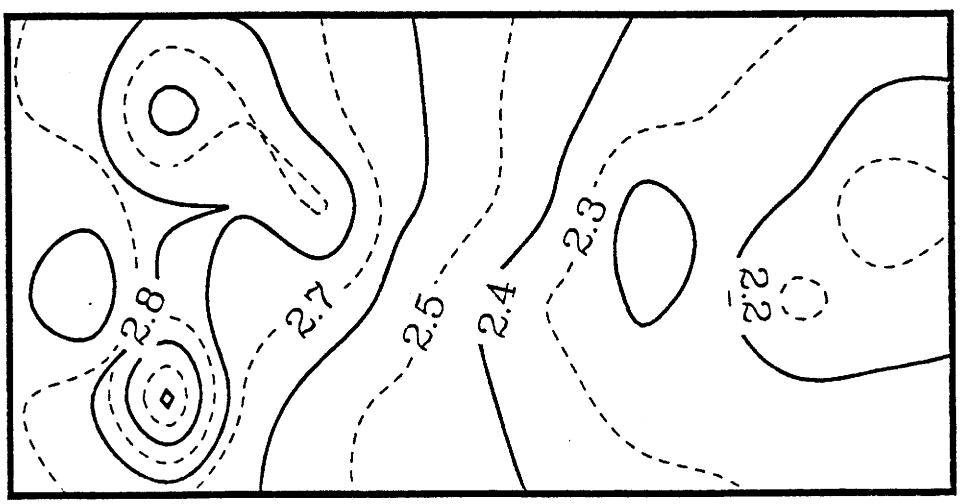
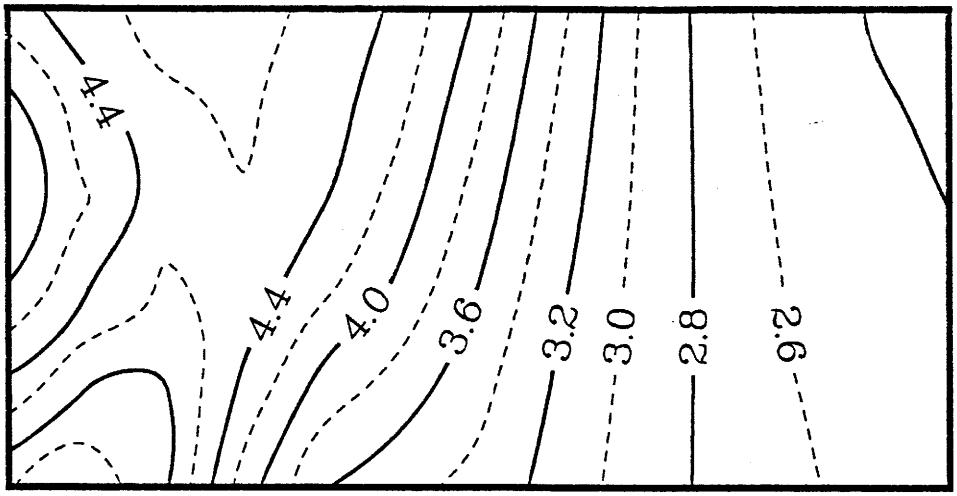


Figure IV.7 Mean & Error Contours from Inverse Distance for January

Mean



Error

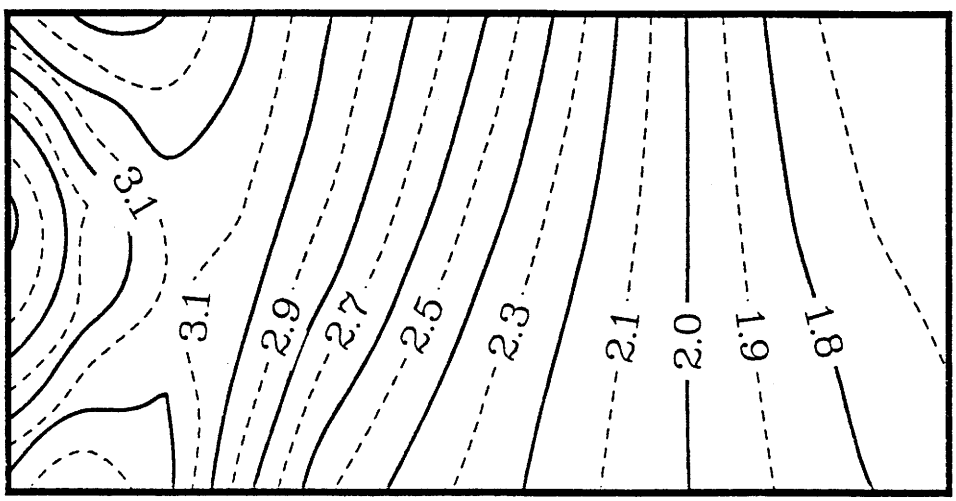


Figure IV.8 Mean & Error Contours from Gaussian Smoothing for January

CHAPTER V

PRECIPITATION NETWORK REDUCTION AND DESIGN

Operation and maintenance of large number of gage station in a remote area requires significant fiscal and manpower commitment. Currently, the Wyoming Water Research Center is contemplating the possibility of future budget reallocation. The task then is set out to examine the possible reduction of network size without sacrificing too much hydrologic information content in the Snowy Range Observatory.

In this Chapter, the formulation of the precipitation network design model that takes into account the spatial characteristics of monthly total precipitation in the study area was considered.

First, the subjective selection of the reduced precipitation network was made based on the geographical locations of the existing stations and other considerations such as the accessibility of the station and the aesthetic aspect. Since precipitation catches are correlated spatially, those stations having similar statistical characteristics for the monthly total precipitation can be reduced without affecting the hydrologic information contents in the study area. Consultations were made with the WRRC researchers and two reduced precipitation networks in the Snowy Range watershed were obtained. Of course, this reduced network so subjectively determined does not guarantee to be the optimum result in statistical information contents.

The model for optimal precipitation network configuration is also considered. The objective of the model is to determine the optimal precipitation network configuration in the Snowy Range watershed that optimizes some types of statistical information measures. In particular, the nonstationary kriging variance is used to construct the object function of the model. In hydrologic network design, there are, generally, two criteria used in the model: (1) accuracy and (2) cost. Some papers that treated this subject considered both (Bras et al., 1976; Loaiciga, 1989) or only accuracy (Shamsi et al., 1988; Bardsley, 1985; Sorman, 1983). This study shall only consider accuracy to find the minimum kriging variance. Mixed integer programming (MIP) is applied to this problem.

V.1 Network Reduction

Two reduced network configurations were considered (see Figures V.1 and V.2). One consisted of 15 stations (101, 102, 103-1, 103-2, 106, 108-2, 109, 115-2, 119, 120, 121, 122, 123, 124 and 127) and the other had 12 stations (101, 103-2, 106, 108-2, 109, 115-2, 119, 120, 121, 122, 123 and 124). The contour maps for the means and the variances of the monthly total precipitation were constructed for each month. The resulting contour maps based on the reduced networks with 15 and 12 stations were compared with those of all 21 precipitation station.

Figure V.3 contains the error maps generated from the nonstationary kriging model I based on 15 and 12 stations, respectively. As can be seen, the errors increase as the number of retained stations decreases. The contour maps of the percentage increment in error associated with the reduced networks related to the full network of 21 stations are plotted in Figure V.4 for January representing the months (from October to April) during which the precipitation is most likely in the form of snow. Figure V.5 is for July representing the months (from May to September) during which the precipitation in general is in the form of rain. Contour maps of error percentage increment for the remaining months are shown in Appendix B. It was observed that, during June - September, the percentage increase in error associated with the reduced network is rather small. However, during the snow months (October-April), the error increases up to about 27% in the upper part of the watershed with a reduced network of 15 stations. Further reduce the network to 12 stations resulted in a slightly higher percentage increment in error to about 30%. The error increment associated with the reduced network in the lower part of the watershed is insignificant all year around. This comparison is also made using the nonstationary kriging model II (see Figures V.6 and V.7). The results from model II are similar to those from model I.

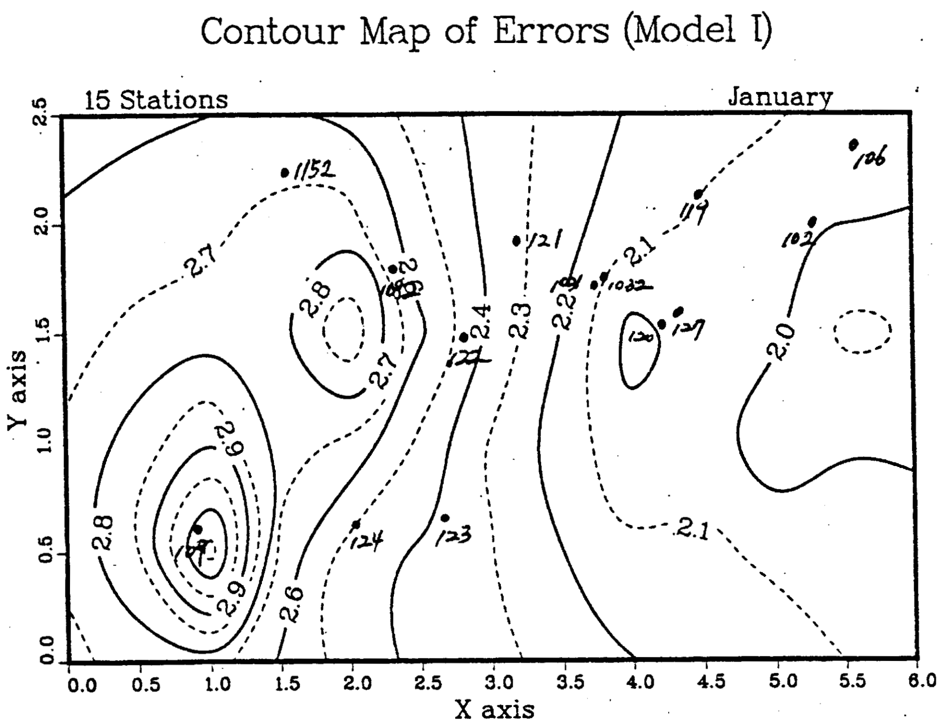
The comparison is also performed for the mean monthly precipitation (see Figure V.8). The results show that the differences in means between 15 and 21 or 12 and 21 stations are not distinctly different.

V.2 Optimal Network Design

V.2.1 Model

Consider that there are K estimation points and S existing precipitation stations in the study area. The objective is to identify the optimal subset of existing stations to be retained to minimize the estimation error. Based on Eq.(IV.31a), the error, measured by the MSE, at the k-th estimation point within the study area for the month m is

$$\hat{\sigma}_k^2(m) = \boldsymbol{w}_k^T \mathbf{C}_m \boldsymbol{w}_k - 2\, \boldsymbol{w}_k^T \boldsymbol{c}_{m,k} + \sigma_{m,k}^2 \qquad \text{(V.1)}$$



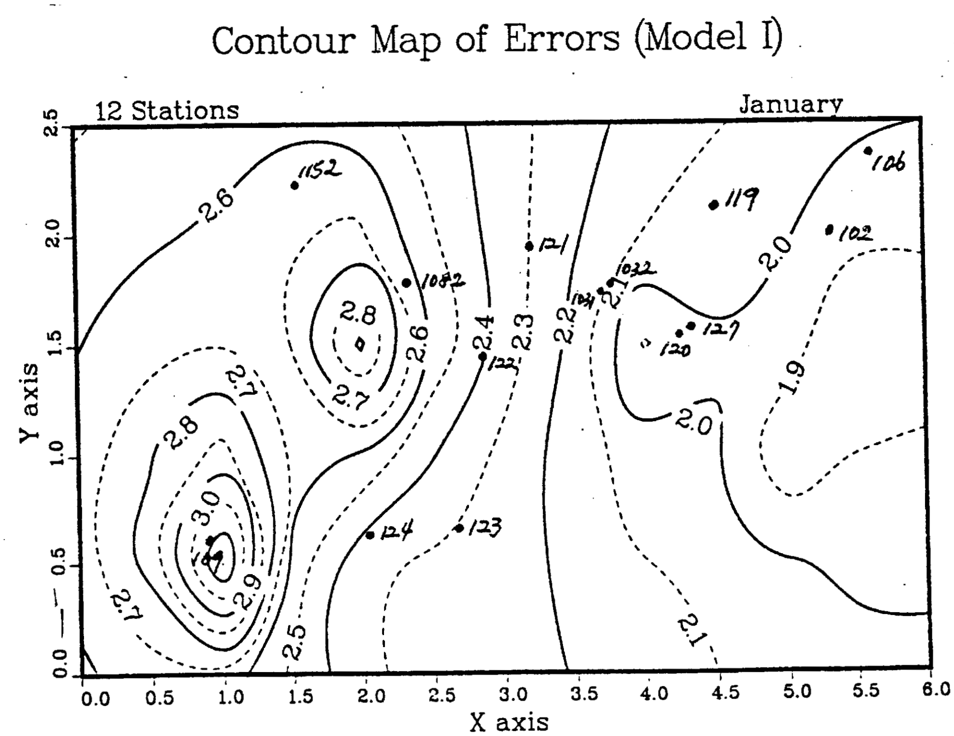
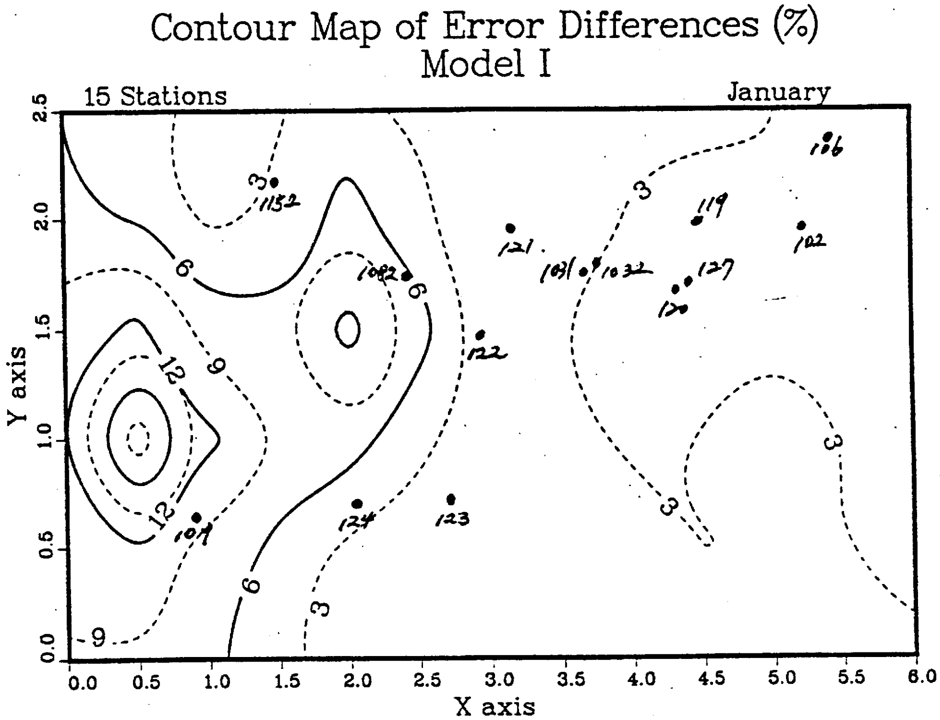


Figure V.1 : Contour Map of Error for Reduced Network (Model I)



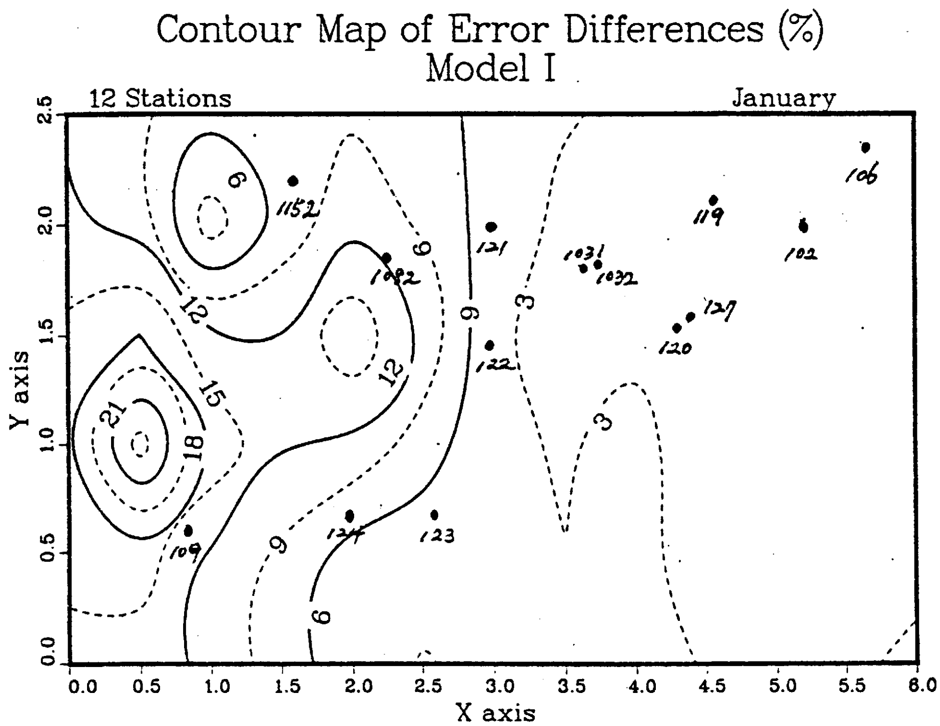
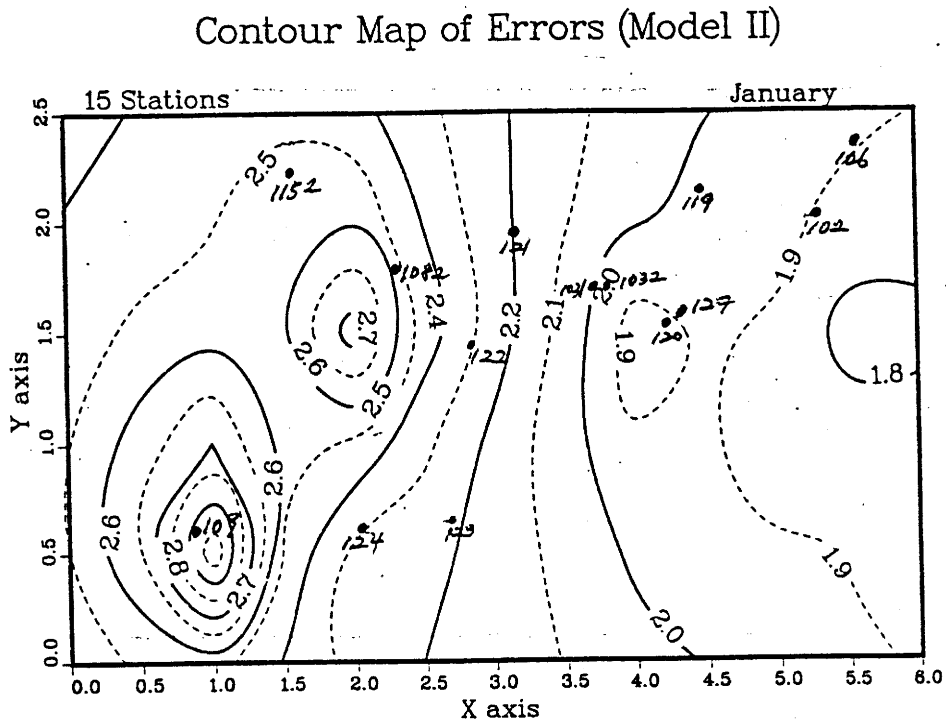


Figure V.2 : Contour Map of Percentage Difference in Error by Reduced Network (Model I)



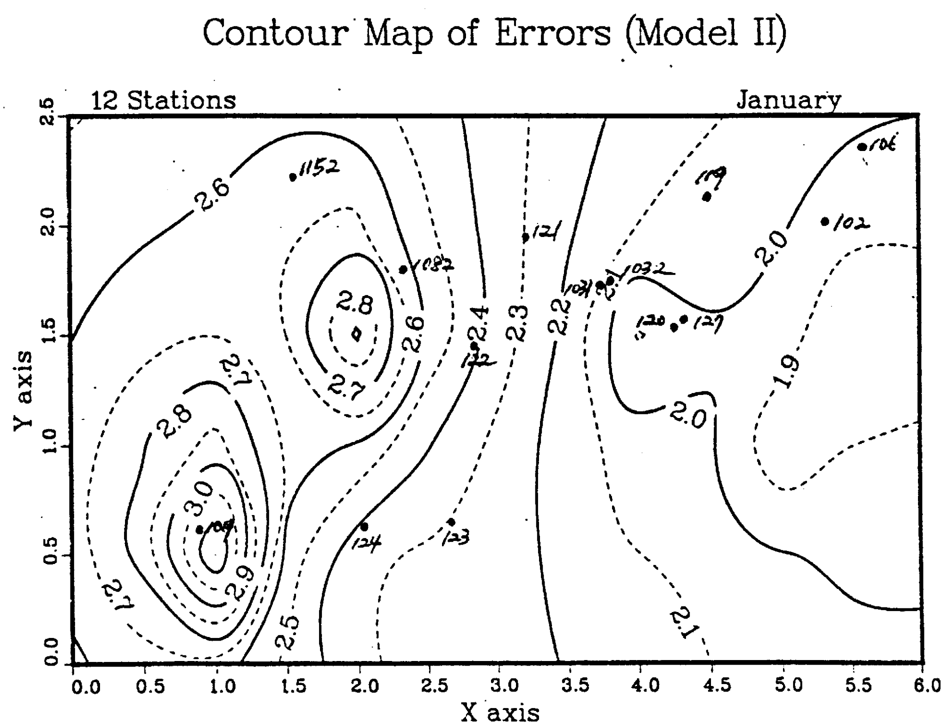
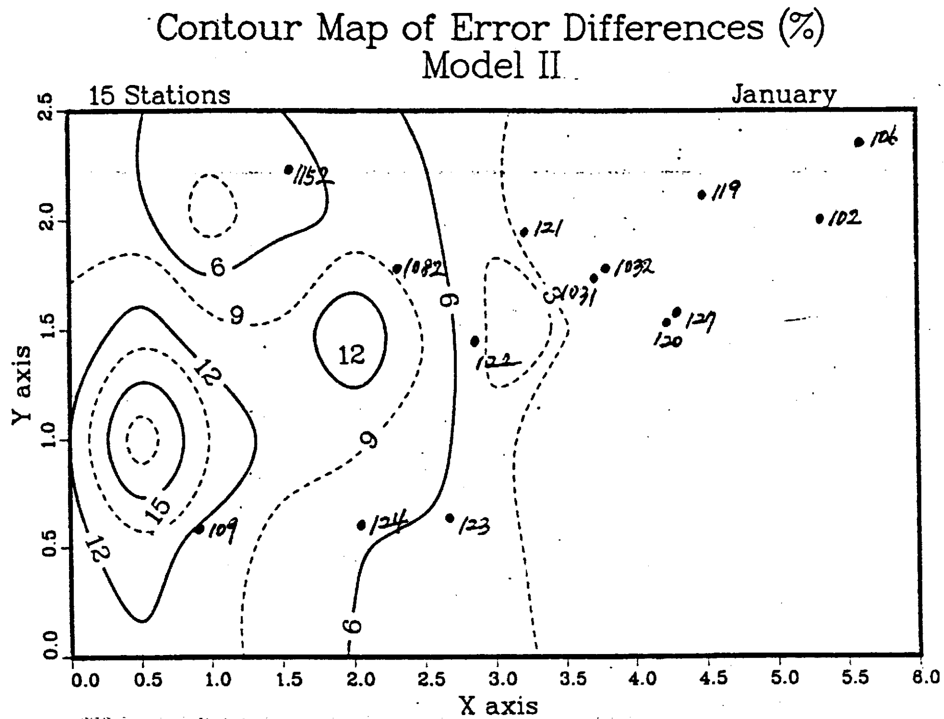


Figure V.3 : Contour Map of Error for Reduced Network (Model II)



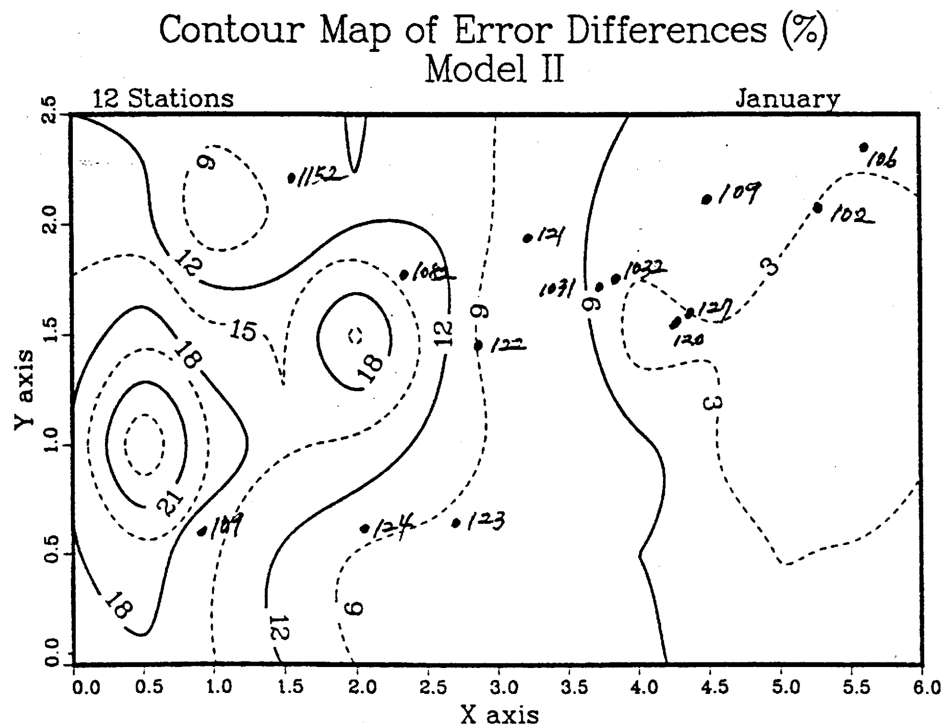
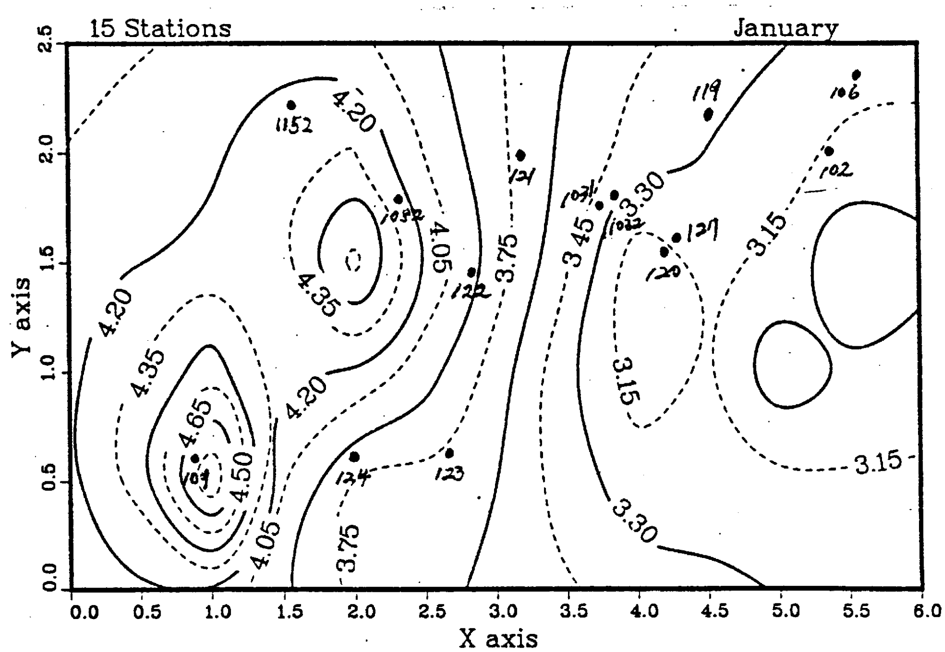


Figure V.4 : Contour Map of Percentage Difference in Error by Reduced Network (Model II)

Contour Map of Means (Model I)



Contour Map of Means (Model II)

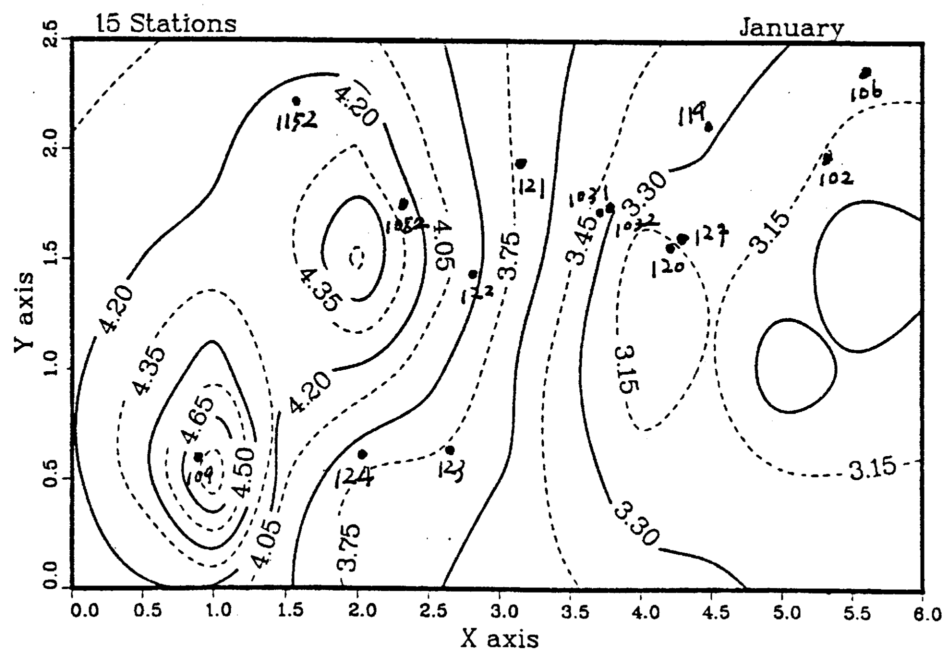


Figure V.5 : Comparison of Contour Map of Mean for Reduced Network by Model II where w_k is a Sx1 vector of weights for the k-th estimation point, $c_{m,k}$ is a Sx1 vector of the estimated covariance between the estimation point k and the existing stations S, $\sigma_{m,k}^2$ represents the local variance at the estimation point k for month m estimated by the contributing weights and the corresponding variances at the retained stations. A representative measure of the estimation error for a study area is the aerial averaged error or, equivalently, the total aerial error

$$\sum_{k=1}^{K} \sigma_k^2(m) = \sum_{k=1}^{K} [\boldsymbol{w}_k^T \mathbb{C}_m \boldsymbol{w}_k - 2\boldsymbol{w}_k^T \boldsymbol{c}_{m,k} + \sigma_{m,k}^2] \tag{V.2}$$

In the network design context, the status of each existing station, to be retained or removed, is unknown. Therefore, zero-one integer variables, z_s, for s=1,2,..,S, are introduced to the model. If the station s is to be retained, z_s has the value of one, otherwise, z_s=0.

The objective of this precipitation network design problem is to minimize the total MSE over the specified area. The proposed mixed integer programming (MIP) model to optimally select the existing precipitation stations for month m is the following:

$$\text{Minimize} \quad \sum_{k=1}^{K} \sigma_k^2(m) = \sum_{k=1}^{K} [\boldsymbol{w}_k^T \mathbb{C}_m \boldsymbol{w}_k - 2\, \boldsymbol{w}_k^T \boldsymbol{c}_{m,k} + \sigma_{m,k}^2] \tag{V.3}$$

The constraints in the model include the following:

(a) The sum of contributing weight from the retained stations to individual estimation point is unity.

$$\boldsymbol{w}_k^T \mathbf{1} = 1 \quad , \; k = 1,2,\ldots,K \tag{V.4}$$

(b) The mean monthly total precipitation at the estimation point k is a linear weighted average of the monthly mean precipitation measured from the retained stations

$$\boldsymbol{w}_k^T \boldsymbol{\mu}_m = \mu_{m,k} \quad , \; k = 1,2,\ldots,K \tag{V.5}$$

(c) Because any existing station s, if retained, can possibly contribute to compute monthly precipitation at all K estimation points, the upper bound value for the sum of contributing weights from the existing station s will be K. If not retained (z_s=0), the contribution from station s will be zero.

$$\sum_{k=1}^{K} w_{s,k} - z_s K \le 0 \quad , \; s = 1,2,\ldots,S \tag{V.6}$$

(d) The number of stations to be retained cannot exceed the specified number N determined by fiscal or geographical consideration.

$$\sum_{s=1}^{S} z_s \le N \quad , \tag{V.7}$$

(e) The contributing weights are non-negative.

$$\begin{aligned} w_k &\ge 0 \;, & k &= 1,2,\ldots,K, \\ z_s &= 0 \;\; or \;\; 1 \;, & s &= 1,2,\ldots,S \end{aligned} \tag{V.8}$$

The first and second constraints are the same as those of model I for nonstationary kriging described in Chapter IV.

The proposed optimal network design model is nonlinear. For each month, the MIP problem includes the zero-one integer variables, z_s, and non-negative real decision variables, w_k, $k=1,2,..,K$. The number of decision variables is S+SxK real-valued decision variables for the weighing factors and S zero-one integer decision variables. The number of constraints is 2K+S+1 (not including nonnegativity constraints).

V.2.2 Solution Algorithm

Since the optimal network design model is a nonlinear MIP problem, there is no solution software immediately available. It is proposed to linearize the objective function by retaining the first order Taylor expansion term. The first-order Taylor expansion of the MSE for the given estimation point k about $w_k = w_k^0$ is

$$f(\boldsymbol{w}_k) = \boldsymbol{w}_k^T \mathbb{C}_m \boldsymbol{w}_k - 2\, \boldsymbol{w}_k^T \boldsymbol{c}_{m,k} + \sigma_{m,k}^2 \tag{V.9}$$

$$= f(\boldsymbol{w}_k^0) + \nabla^T f(\boldsymbol{w}_k^0)\,(\boldsymbol{w}_k - \boldsymbol{w}_k^0) + \varepsilon \tag{V.10}$$

$$\begin{aligned} \approx\; &[f(\boldsymbol{w}_k^0) - 2\, \boldsymbol{w}_k^0\, \mathbb{C}_m\, \boldsymbol{w}_k^0 + \sigma_{m,k}^2(m)] \\ &+ 2\, \boldsymbol{w}_k^T \mathbb{C}_m \boldsymbol{w}_k^0 - 2\, \boldsymbol{w}_k^T \boldsymbol{c}_{m,k} \end{aligned} \tag{V.11}$$

$$= constant_k^0 + 2\,[\boldsymbol{w}_k^T \mathbb{C}_m \boldsymbol{w}_k^0 - \boldsymbol{w}_k^T \boldsymbol{c}_{m,k}] \tag{V.12}$$

$$= constant_k^0 + 2\, \mathbf{w}_k^T\, [\mathbb{C}_m\, \mathbf{w}_k^0 - \mathbf{c}_{m,k}] \qquad (V.13)$$

Substituting Eq.(V.13) into Eq.(V.3), the linearized objective function of the network design model is

Minimize

$$\sum_{k=1}^{K} f(\mathbf{w}_k) = \sum_{k=1}^{K} \left[\mathbf{w}_k^T\, \mathbb{C}_m\, \mathbf{w}_k - 2\, \mathbf{w}_k^T\, \mathbf{c}_{m,k} + \sigma_{m,k}^2\right]$$

$$\approx \sum_{k=1}^{K} constant_k^0 + 2 \sum_{k=1}^{K} \mathbf{w}_k^T \left[\mathbb{C}_m\, \mathbf{w}_k^0 - \mathbf{c}_{m,k}\right] \qquad (V.14)$$

The above objective function is equivalent to

minimize

$$\sum_{k=1}^{K} \mathbf{w}_k^T \left[\mathbb{C}_m \mathbf{w}_k^0 - \mathbf{c}_{m,k}\right] \qquad (V.15)$$

subject to

$$\mathbf{w}_k^T\, \mathbf{1} = 1 \quad , k = 1,2,\ldots,K \qquad (V.16)$$

$$\mathbf{w}_k^T \boldsymbol{\mu}_m = \mu_{m,k} \quad , k = 1,2,\ldots,K \qquad (V.17)$$

$$\sum_{k=1}^{K} w_{sk} - z_s K \le 0 \quad , s = 1,2,\ldots,S \qquad (V.18)$$

$$\sum_{s=1}^{S} z_s \le N \; , \qquad (V.19)$$

$$\begin{array}{ll} \mathbf{w}_k \ge 0 \, , & k = 1,2,\ldots,K \, , \\ z_s = 0 \;\; or \;\; 1 \, , & s = 1,2,\ldots,S \end{array} \qquad (V.20)$$

The linearized network design model is then solved by the computer program ZOOM (Zero-One Optimization Methods) developed by Marsten (1988).

Note that the linearized network design model requires initial guess on the weighing factors, which may or may not be the optimal one. Therefore, the linearized model must be solved iteratively each time the weighing factors are revised and updated if the current solutions are different from the previous solutions. The procedure then is repeated until the solutions converge.

V.2.3 Numerical Results and Discussion

Knowing that station 121-A had already been removed from the Snowy Range Observatory, the remaining twenty stations were then used in the model to find the optimum subset of precipitation stations to be retained. The network model was solved to select six stations (N=6) out of 20. Six stations (103-1, 106, 108-2, 119, 121, 124) which were not too close to each other were chosen as the initial solution. After numerous iterations, the best six stations did not converge; they varied from one iteration to another even though the minimum error was smaller or stayed the same. Different initial solutions for the weighing factors such as those computed by inverse distance or from previous kriging results were used and the algorithm still failed to converge.

One possible explanation for the algorithm failure was that the solutions to the optimal network design model was very sensitive to the initial stations chosen in linearization procedure. To select the best 6 out of 20 existing stations, there are $\binom{20}{6} = 38760$ possibilities. The solution selected may be far away from the optimal one.

CHAPTER VI

SUMMARY AND RECOMMENDATION

In Chapter II, several estimation methods were used to estimate missing values and the accuracy of each method was compared with RMSE and MAE. There was no single estimation method that is uniformly superior in all circumstances. However, in the majority of the cases considered, the linear inverse distance weighing method (IDLIN) was better than all other methods through all 9 stations.

In Chapter III, detrending and deseasonalizing procedures were employed to make the time series data stationary. It was found that there was no discernable long-term trend in annual average precipitation. Both monthly sample statistics and the fitted monthly sample mean by a Fourier series were used to standardize the monthly precipitation for removing the within-year seasonal pattern in monthly precipitation time series.

In Chapter IV, the inverse distance and Gaussian smoothing weighing methods, which do not consider spatial correlation, were used to estimate the monthly average precipitation over the space. The nonstationary kriging, which considers spatial correlation, was also used to estimate the monthly precipitation. Two models were formulated by the nonstationary kriging and they were compared with the contour maps for the means and the errors. The models I and II yielded almost identical spatial patterns of monthly precipitation, but the model I had large errors in estimation. The kriging models, the inverse distance method, and Gaussian smoothing technique were also compared with the contour maps for the means and the errors. It was observed that the kriging method was superior to other two methods.

In Chapter V, two reduced precipitation networks, each retains 15 and 12 gages, are subjectively selected on the basis of spatial locations of gages, record length of gage, and aesthetic reason. The results indicated that error increases as the number of retained stations decreases. The maximum increase in error during May-September associated with the reduced networks of 15 (28% reduction in network size) and 12 (43% reduction in network size) stations, from an original 21 stations, are about 8% and 10%, respectively. However, during October-April when precipitation is most likely to be in the form of snow, the percentage error increases to a maximum of 28% for the 15-station network and 30% for the 12-station network. The area within the watershed where larger error occurs is above the Brooklyn Lake.

A model for optimal precipitation network configuration was developed. The resulting formulation was a nonlinear integer programming model to which there is no commercial software available for use. A procedure was attempted to linearize the nonlinear function and so that the well developed linear integer progamming technique can be applied. Unfortunately, the solutions from the linearization procedure were not stable. Two possible alternative methods can be applied to solve the problem. One is to use dynamic programming approach which might be applied to find improved solutions. By dynamic programming approach, the optimization problem is divided into stages representing the number of existing stations to be retained or removed. In each stage, the state can be the list of all existing stations in the network to be retained or removed. The problem can be solved backward or forward using appropriate recursive procedure. Alternatively, recognizing the network design model has a quadratic objective function, it is possible to solve the model by embedding the quadratic programming algorithm into the branch and bound algorithm.

From the study, the following conclusions can be made:

(1) During the months of May-September the precipitation gage number can be significantly reduced without loosing much precipitation information content in the Snowy Range watershed.

(2) During months of snow (October-April), larger error could occur in the upper part of the watershed above Brooklyn Lake with reduced network. The accuracy on the lower portion of the watershed is not significantly affected by the reduced network.

(3) From the above observation, it places quite a dilemma in the present operation of the precipitation network. The upper portion of the watershed area where high error occurs is generally difficult to access especially during the winter season. Currently, the gages are maintained and checked by Water Center staff on the weekly basis. To reduce labor intensiveness of maintaining the gages and to retain the precipitation accuracy one or two gages that can operate over longer period with high reliability can be installed in the remote area.

Selected References

Bardsley, W.E. and Manly, B.F.J., Notes on Selecting and Optimum Raingauge Subset, Journal of Hydrology, 76, 197-201, 1985

Bloomfield, P., Fourier Analysis of Time Series: An Introduction, New York: Wiley,1975

Borgman, L.E., Review of Kriging Procedure, submit for publication, 1990

Bras, R.L. and Iturbe, I.R., Random Functions and Hydrology, Addison-Wesley, 1985

Bras, R.L., Network Design for the Estimation of Areal Mean of Rainfall Events, Water Resources Research, V12, No.6, 1185-1195, Dec., 1976

Bowerman, B.L. and O'connell, R. T., Time Series Forecasting, 2nd Ed., Duxbury Press, 1987

Kallianpur, G. and Karandikar, R.L., White Noise Theory of Prediction, Filtering and Smoothing, Gorden and Breach Science Publishers, 1988

Lasdon, L.S. and Warren, A.D., GRG2 User's Guide, 1986

Loaiciga, H.A., An Optimization Approach for Groundwater Quality Monitoring Network Design, Water Resources Research, V25, No.8, 1771-1782, Aug., 1989

Luenberger, D.G., Linear and Nonlinear Programming, 2nd Ed., Addison-Wesley, 1984

Markridakis, S., Wheelwright, S.C. and McGee, V.E., Forecasting Methods and Applications, 2nd Ed., N.Y., Wiley, 1983

Marsten, R., User's Manual for : ZOOM/XMP, 1988

Parzen, E. and Maidment, D.R., "A Cascade Model of Monthly Municipal Water Use", Texas Engineering Experiment Station, p157, 1981

Royle, A., Geostatistics, Mcgraw Hill, 1980

Shamsi, U.M. and Quimpo, R.G., An Application of Kriging to Rainfall Network Design, Nordic Hydrology, 19, 137-152, 1988

Sorman, U. and Balkan, G., An Application of Network Design Procedures for Redesigning Kizilirmak River Basin Rainguage Network, Turkey, Hydrological Sciences Journal, 28, 2, 6/1983

Taha, H.A., Operations Research, 3rd Ed., 1983

Tung, Y.K., Point Rainfall Estimation for a Mountainous Region, Journal of Hydraulic Engineering, V109, No.10, Oct., 1983

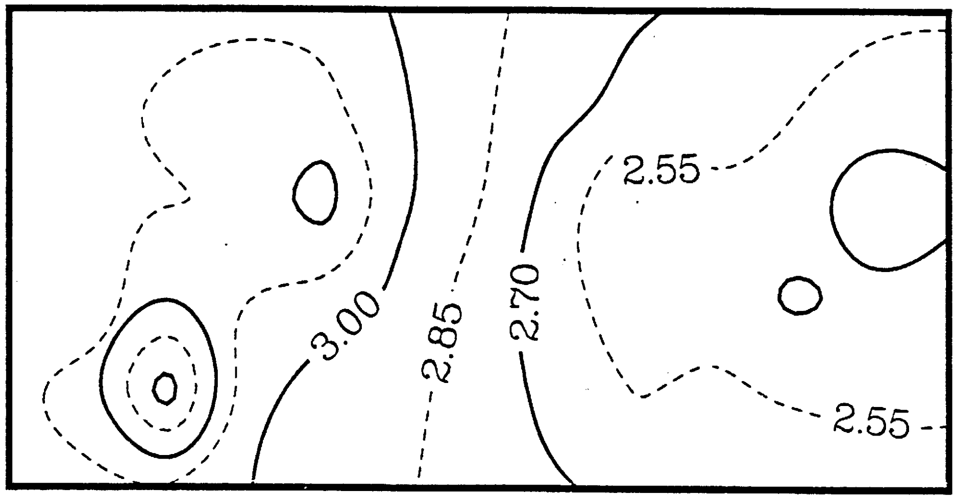
Vandaele, W., Applied Time Series and Box-Jenkins Models, Academic Press Inc., 1983

Wesche, T.A., The Snowy Range Observatory: An Update and Review, Water Resources Series, No. 81, 1982

APPENDICES

Appendix A : Contour Maps for Means and Errors

Model I, Mean



Model II, Mean

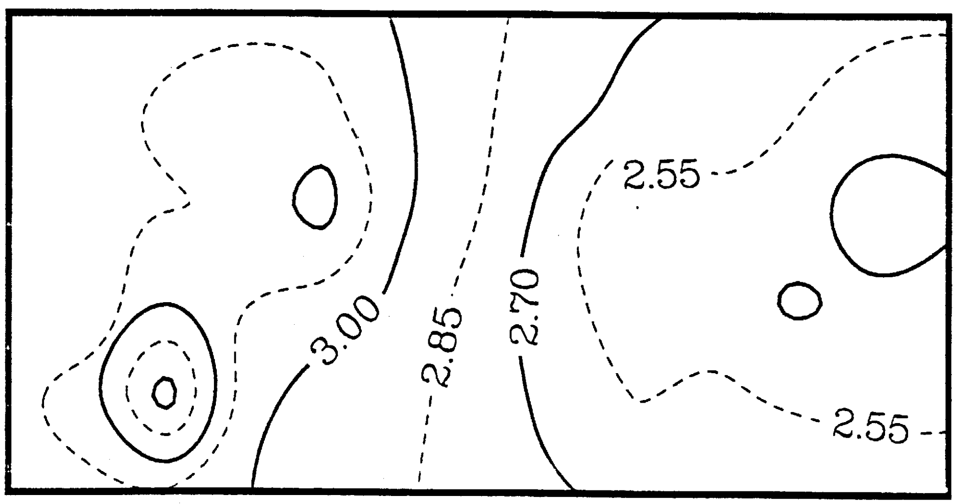


Figure IV.4-1 Comparison of Mean Contours from Model I & II for February

Mean

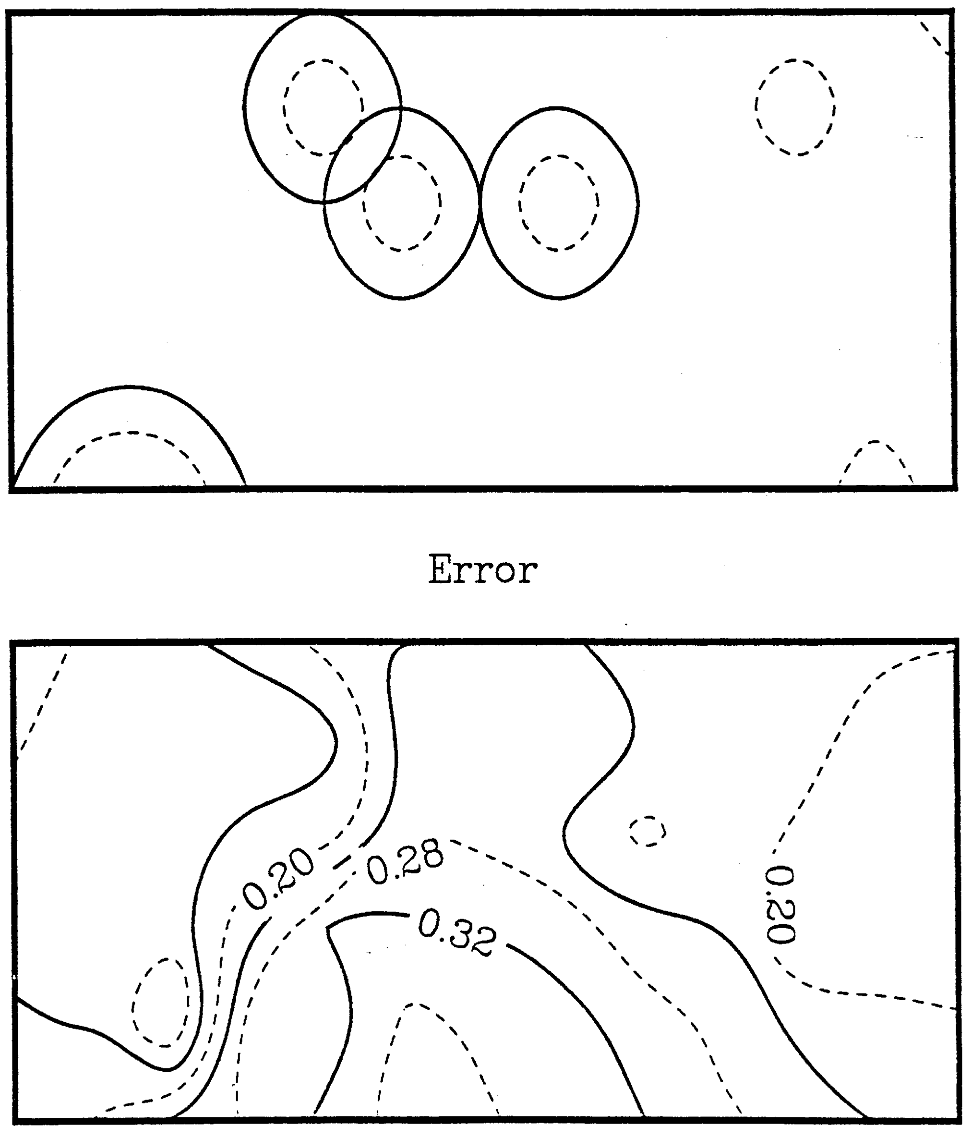
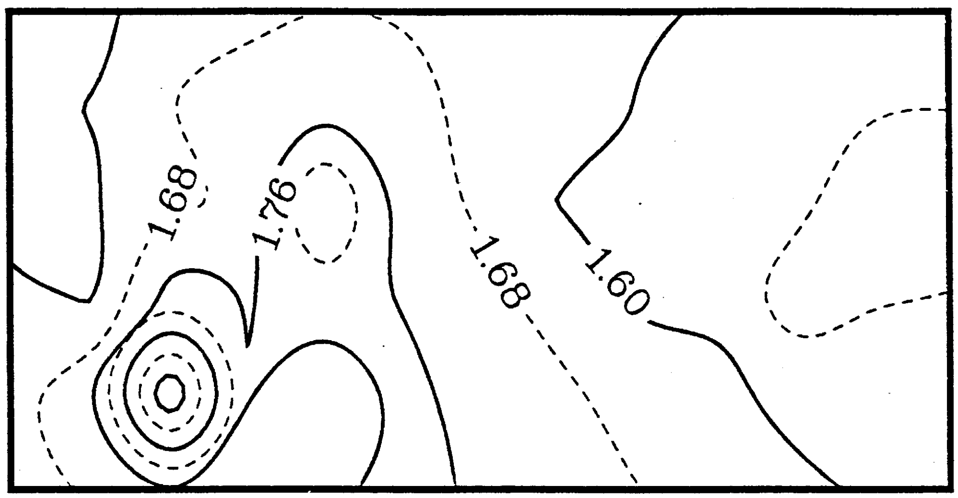


Figure IV.5-1 Contours of Difference between Model I & II for February

Model I, Error



Model II, Error

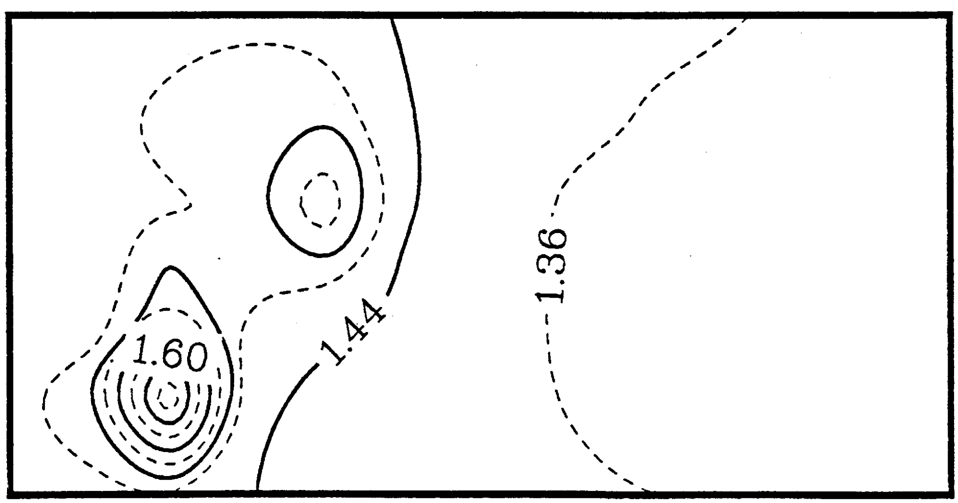
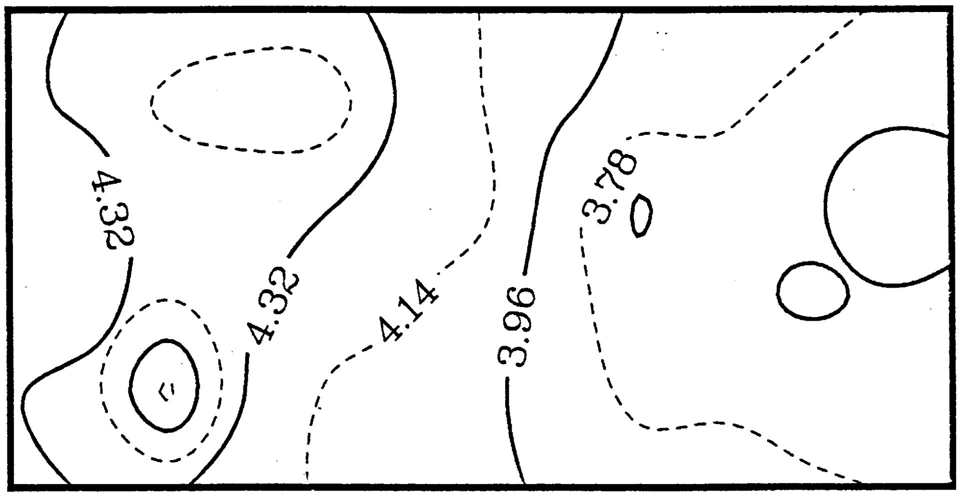


Figure IV.6-1 Comparison of Error Contours from Model I & II for February

Model I, Mean



Model II, Mean

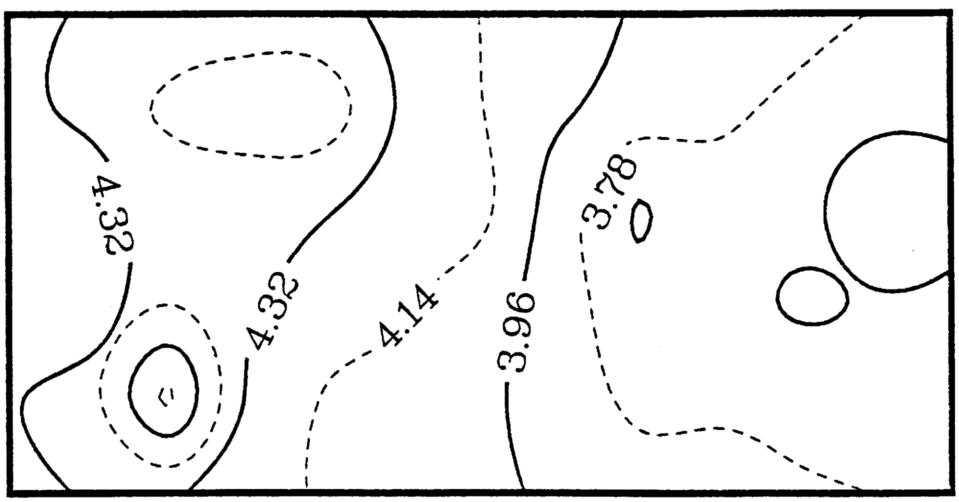
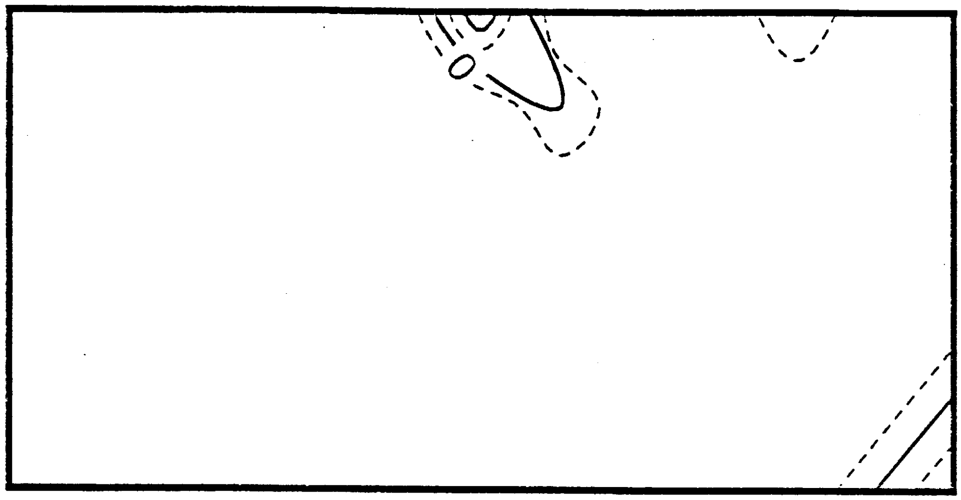


Figure IV.4-2 Comparison of Mean Contours from Model I & II for March

Mean



Error

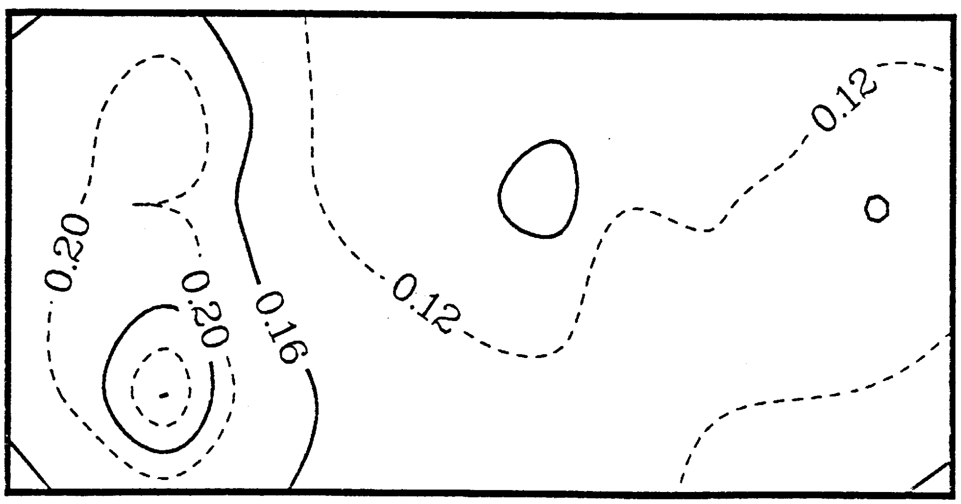
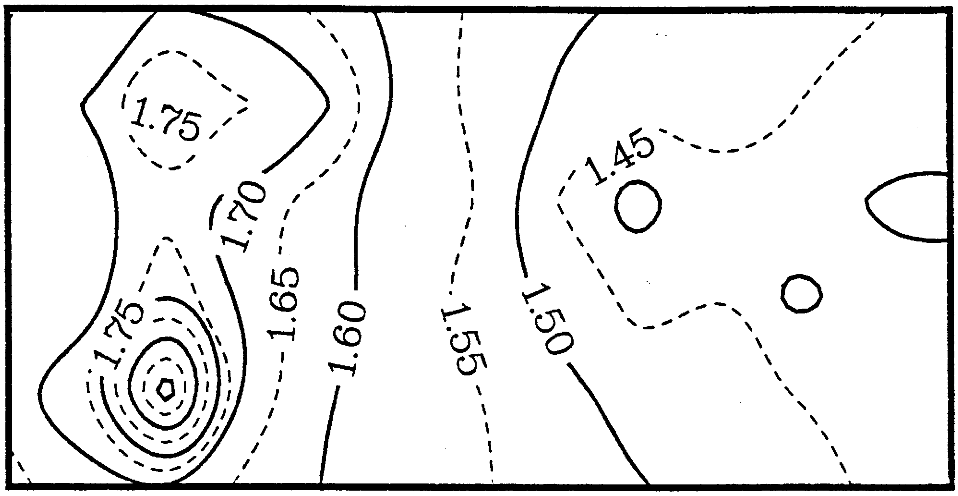


Figure IV.5-2 Contours of Difference between Model I & II for March

Model I, Error



Model II, Error

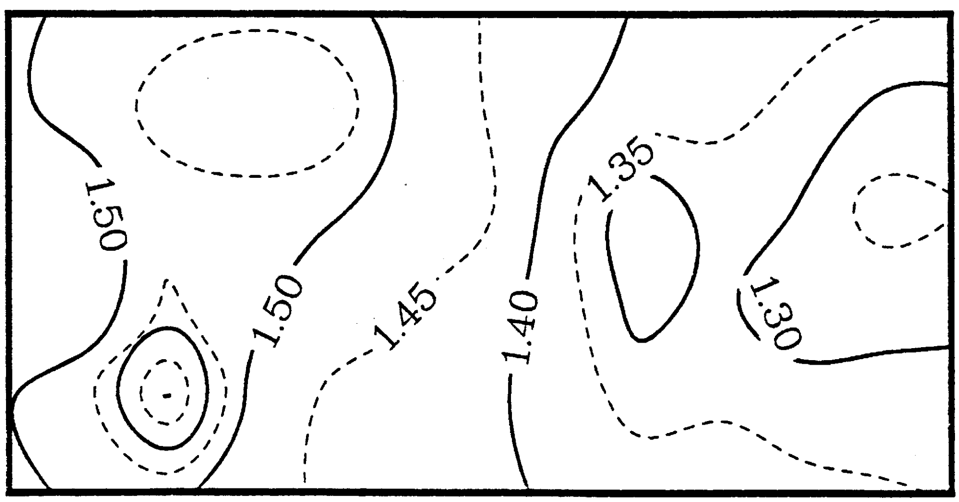
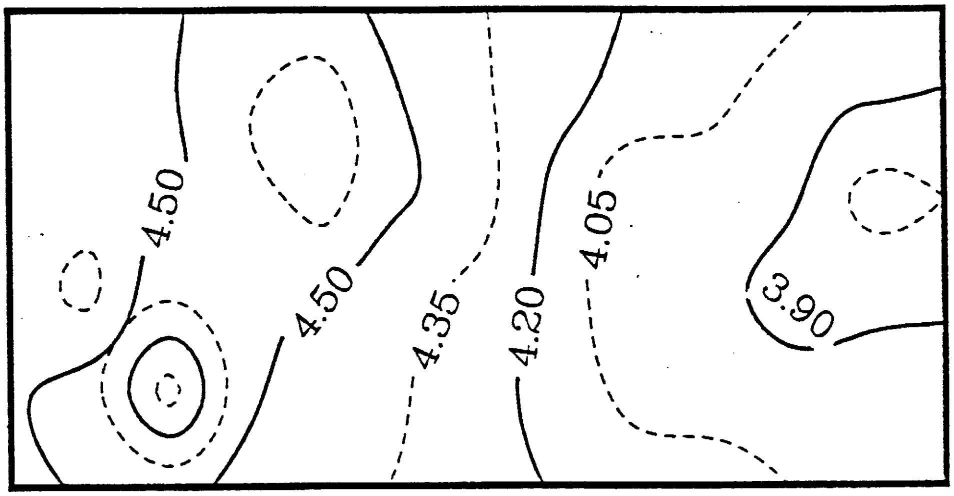


Figure IV.6-2 Comparison of Error Contours from Model I & II for March

Model I, Mean



Model II, Mean

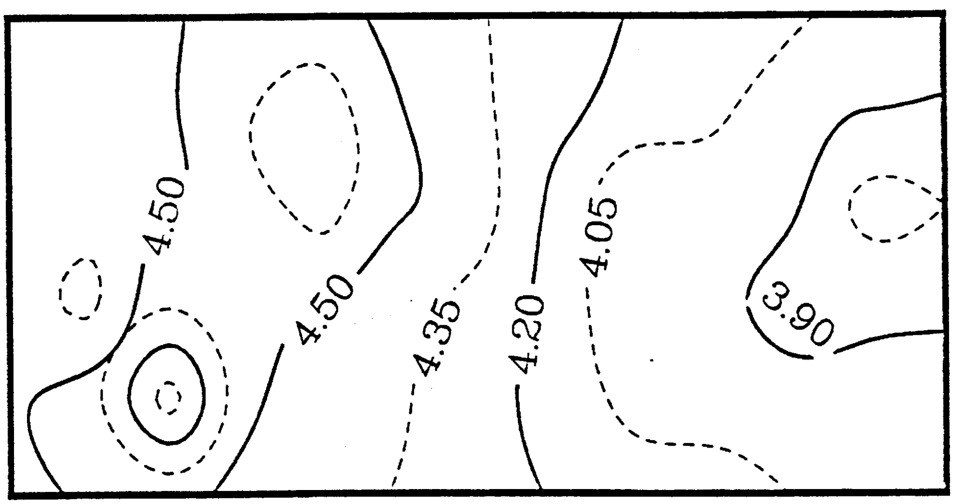


Figure IV.4-3 Comparison of Mean Contours from Model I & II for April

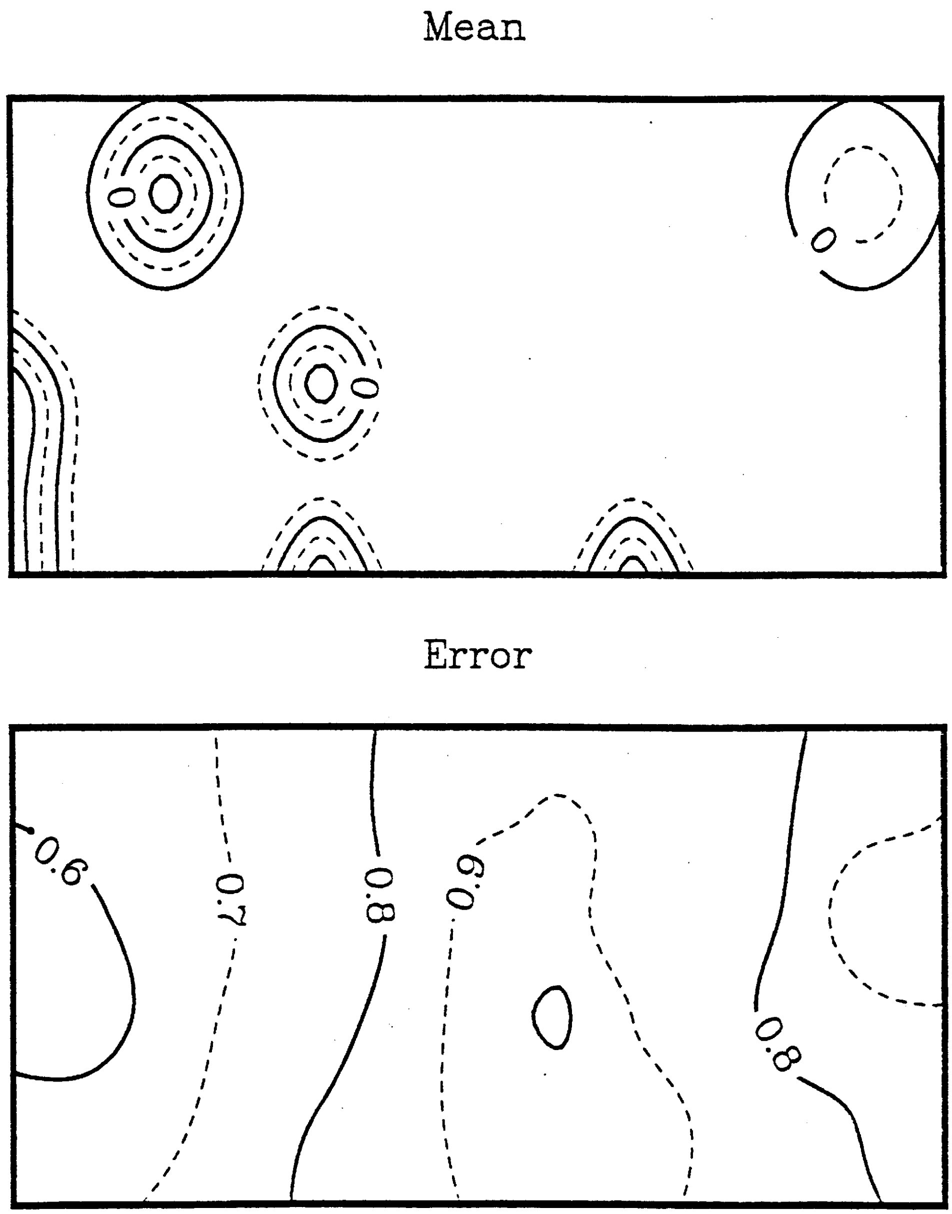
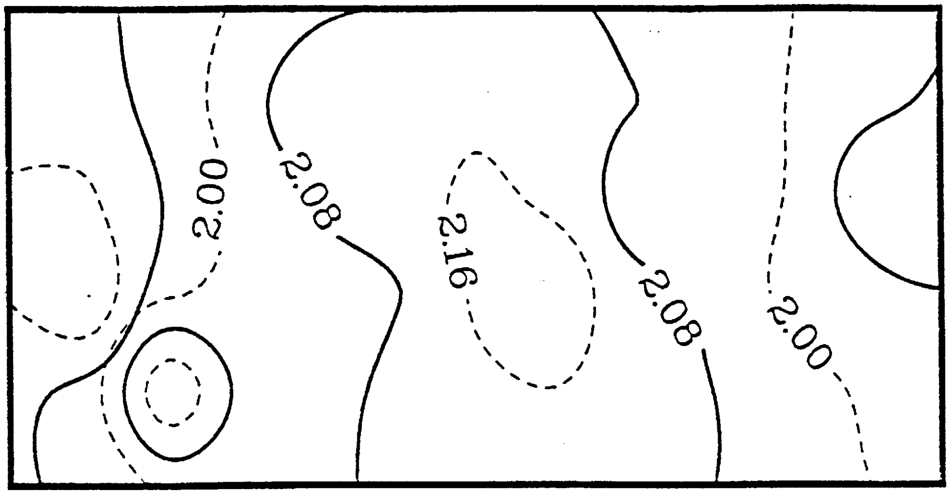


Figure IV.5-3 Contours of Difference between Model I & II for April

Model I, Error



Model II, Error

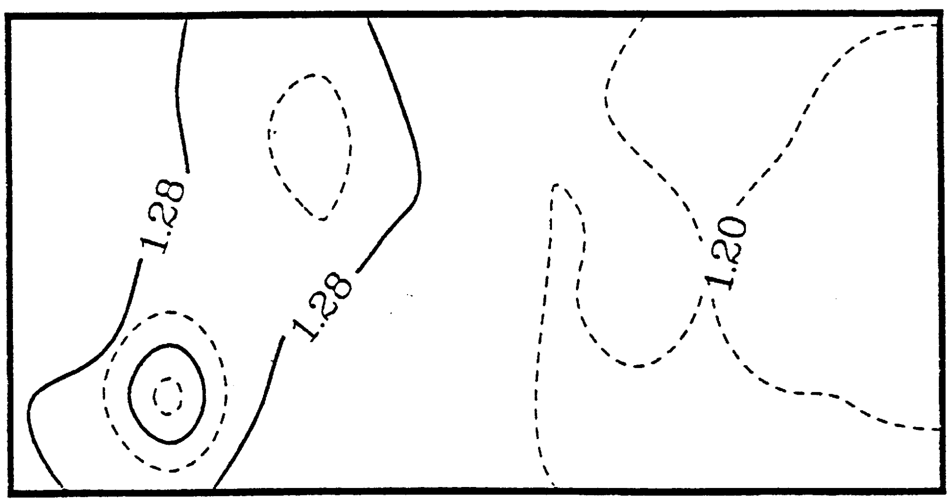
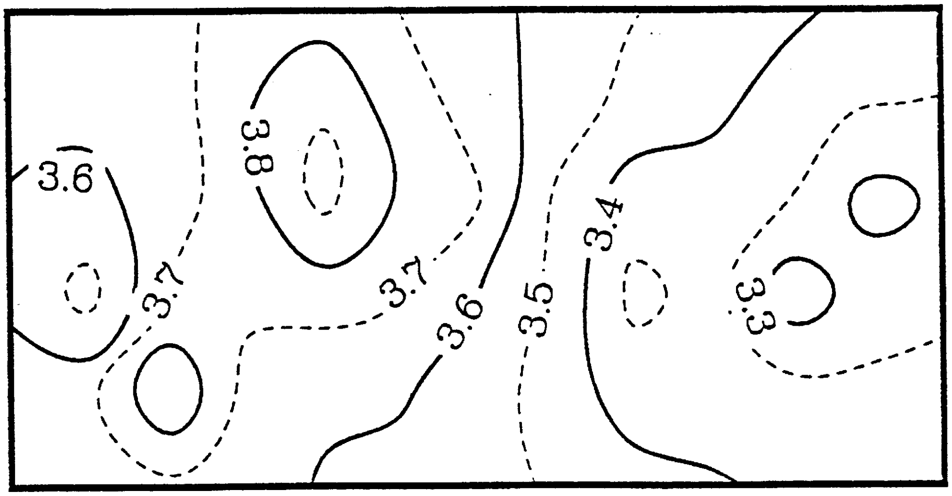


Figure IV.6-3 Comparison of Error Contours from Model I & II for April

Model I, Mean



Model II, Mean

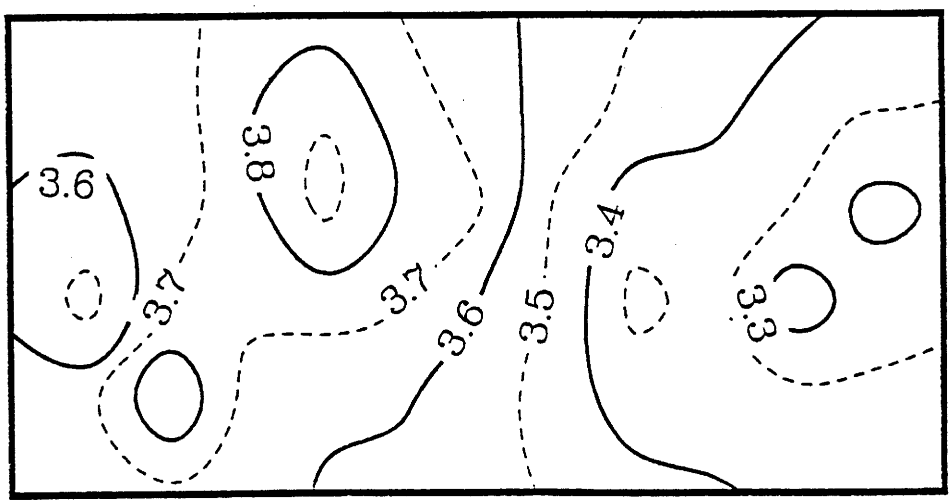
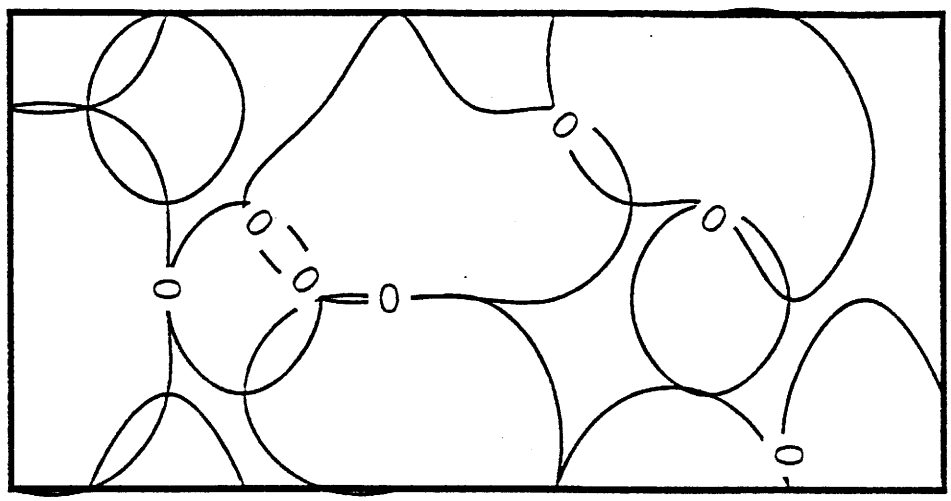


Figure IV.4-4 Comparison of Mean Contours from Model I & II for May

Mean



Error

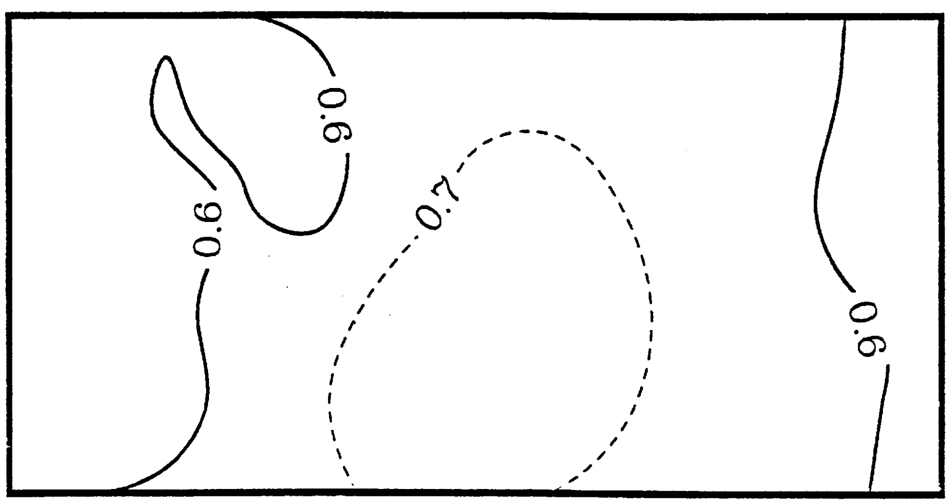


Figure IV.5-4 Contours of Difference between Model I & II for May

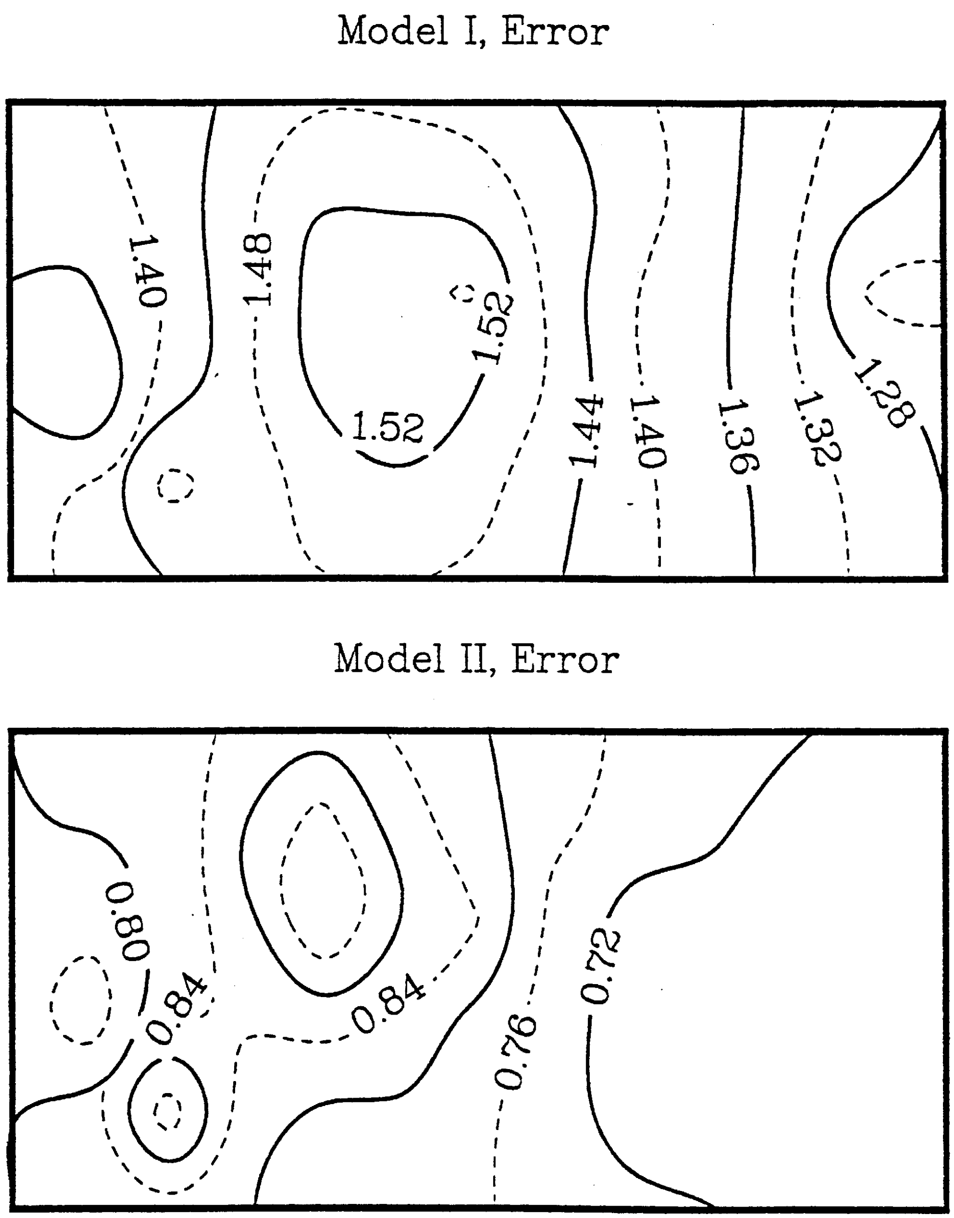


Figure IV.6-4 Comparison of Error Contours from Model I & II for May

Model I, Mean

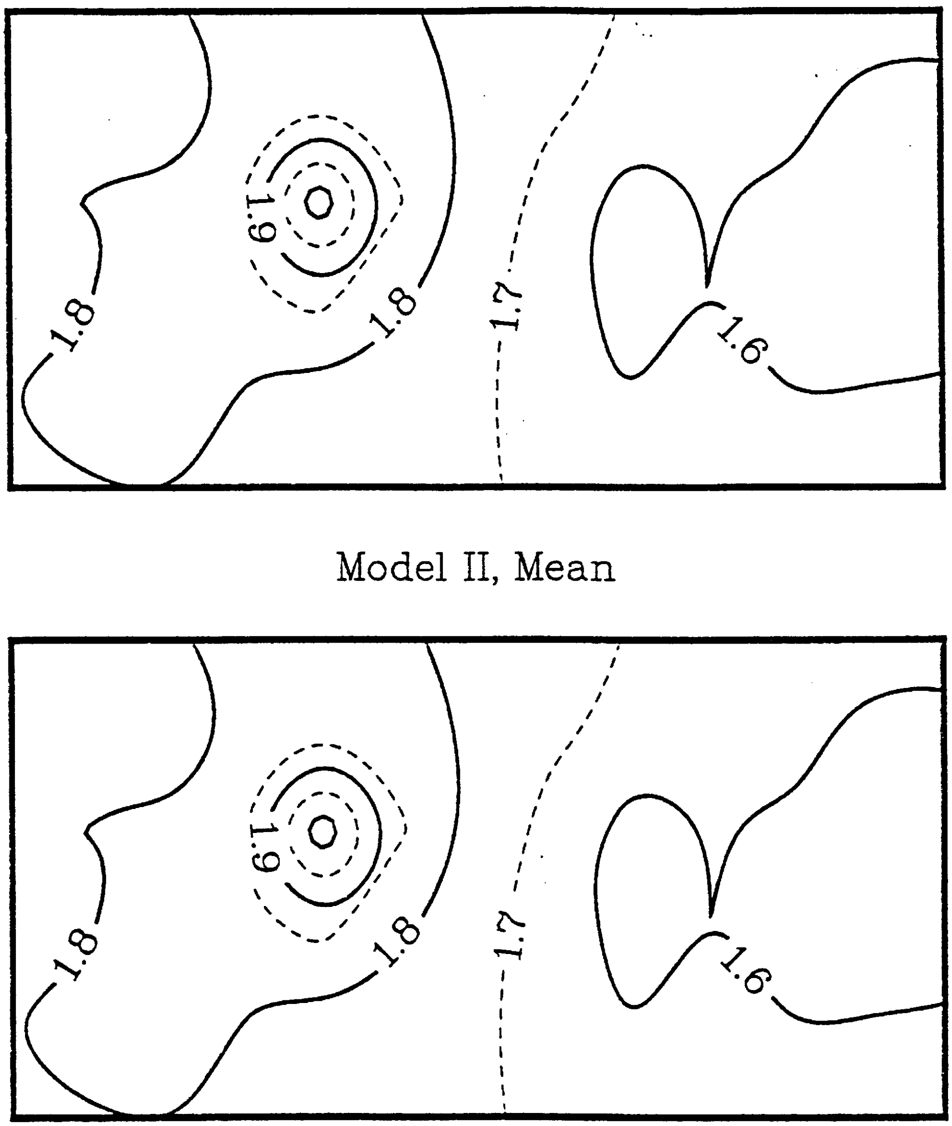
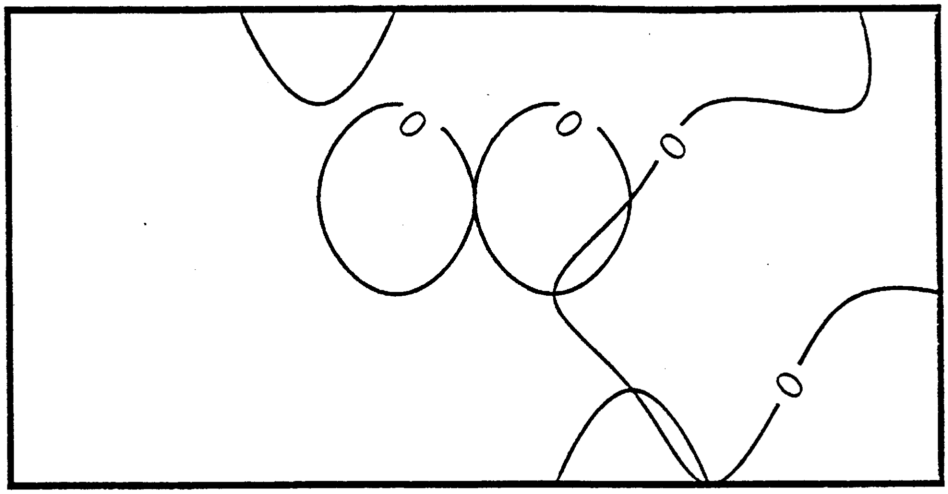


Figure IV.4-5 Comparison of Mean Contours from Model I & II for June

Mean



Error

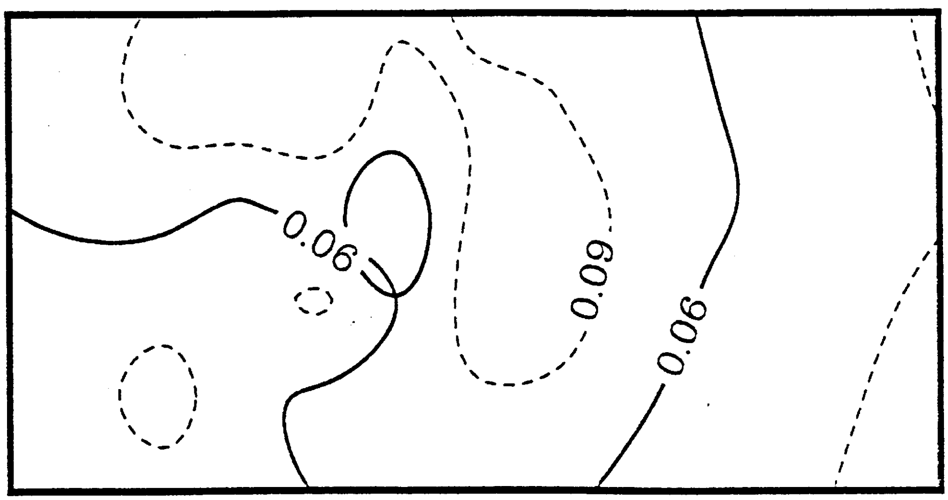
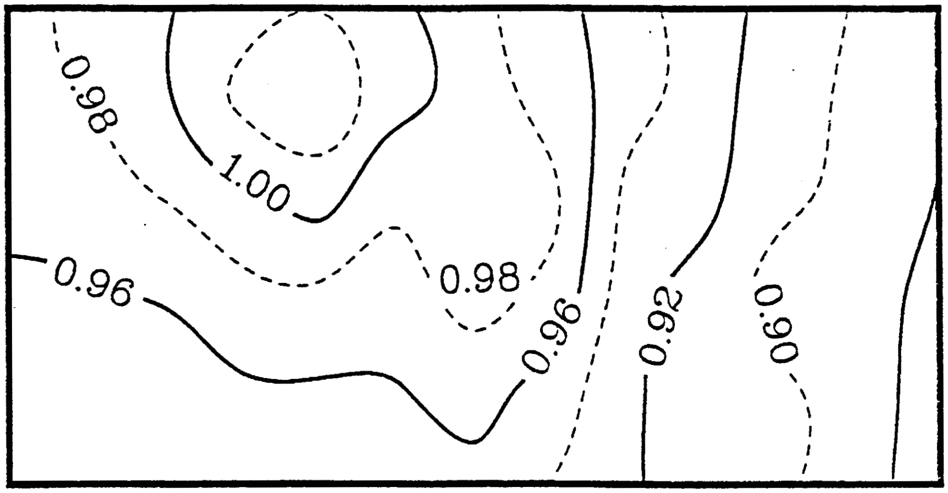


Figure IV.5-5 Contours of Difference between Model I & II for June

Model I, Error



Model II, Error

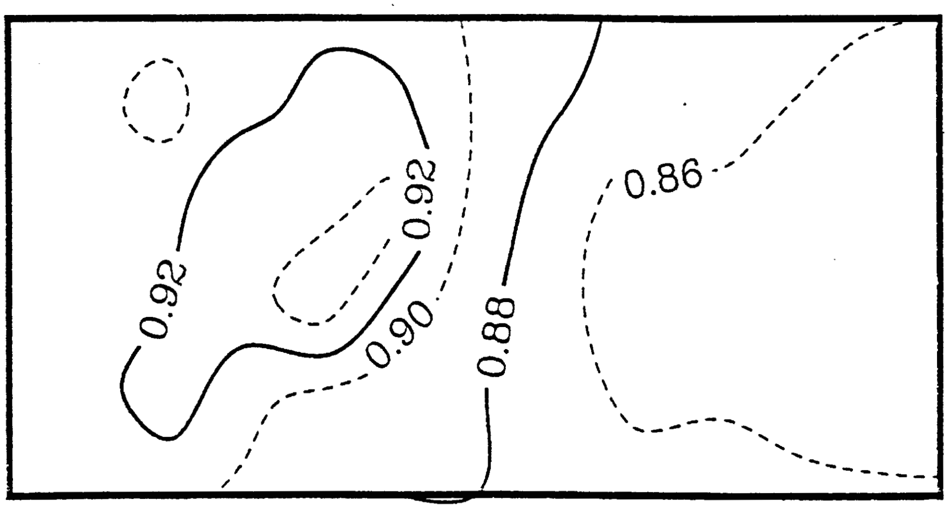
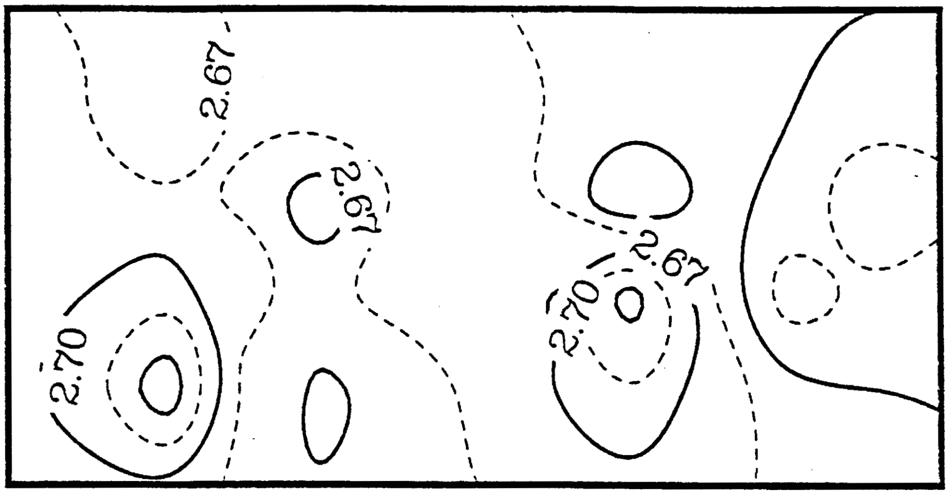


Figure IV.6-5 Comparison of Error Contours from Model I & II for June

Model I, Mean



Model II, Mean

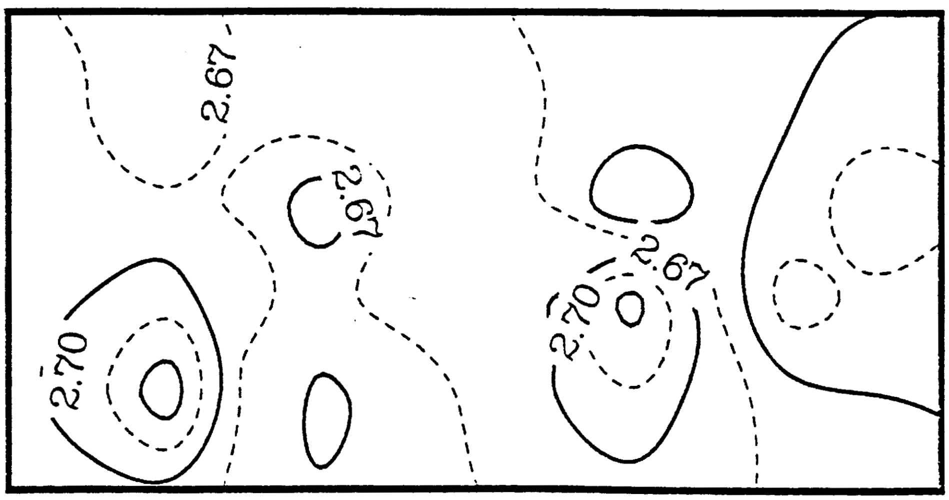
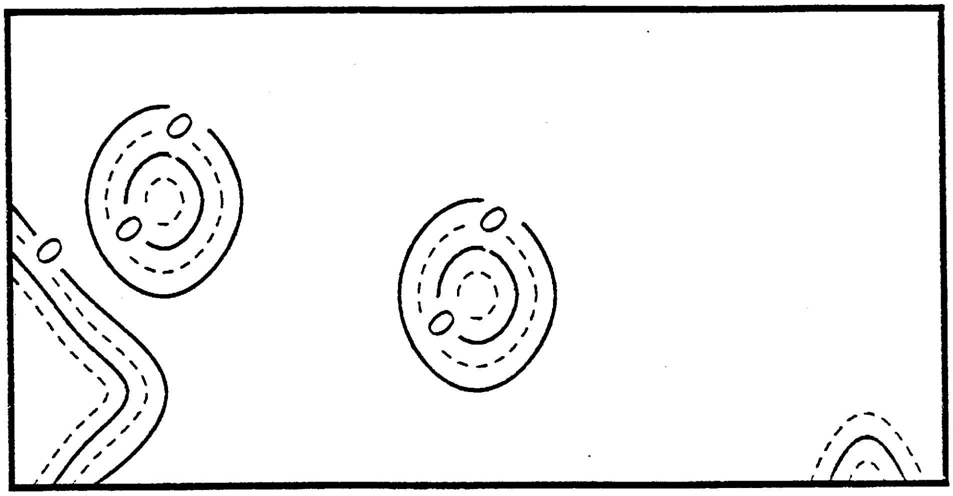


Figure IV.4-6 Comparison of Mean Contours from Model I & II for July

Mean



Error

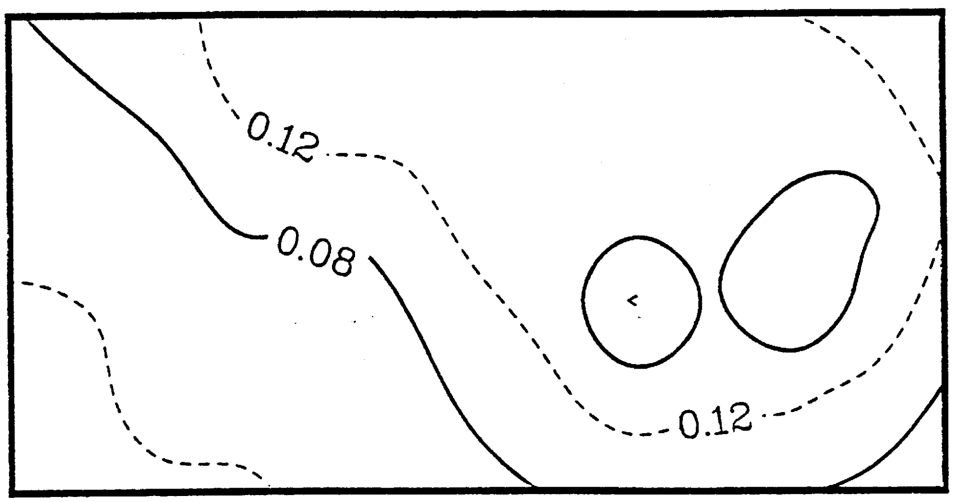
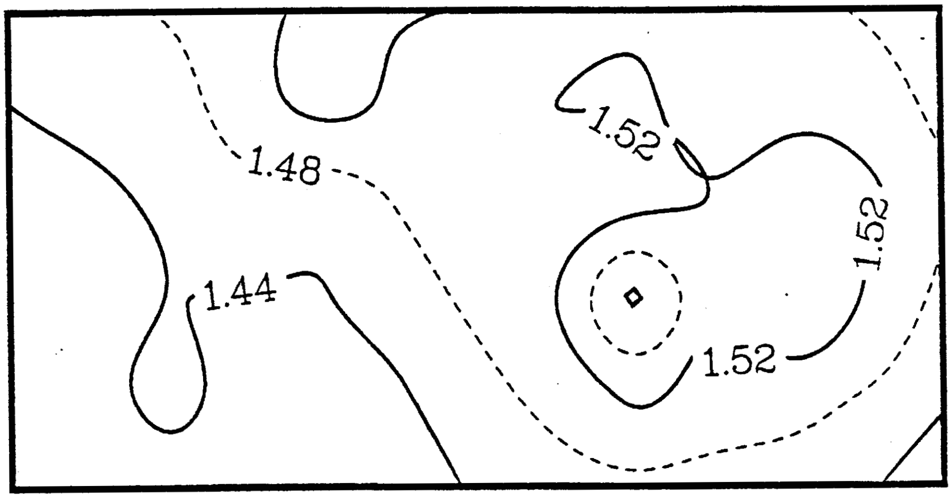


Figure IV.5-6 Contours of Difference between Model I & II for July

Model I, Error



Model II, Error

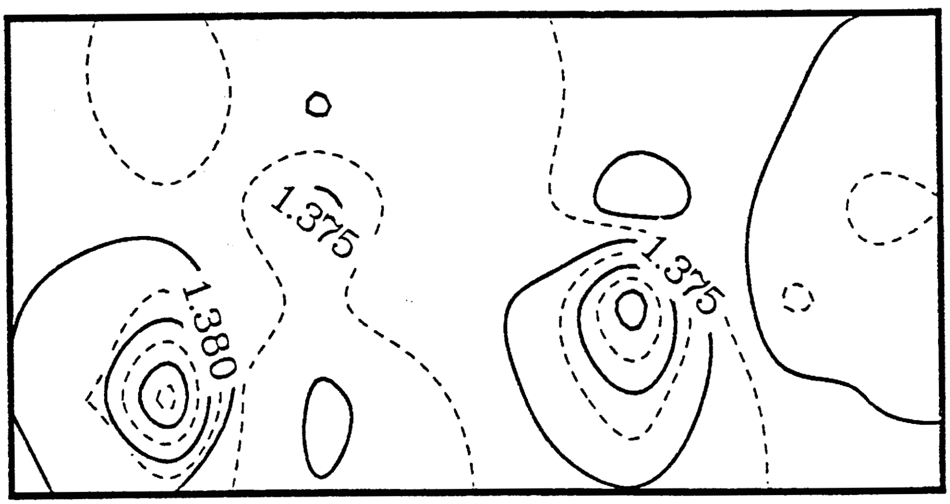
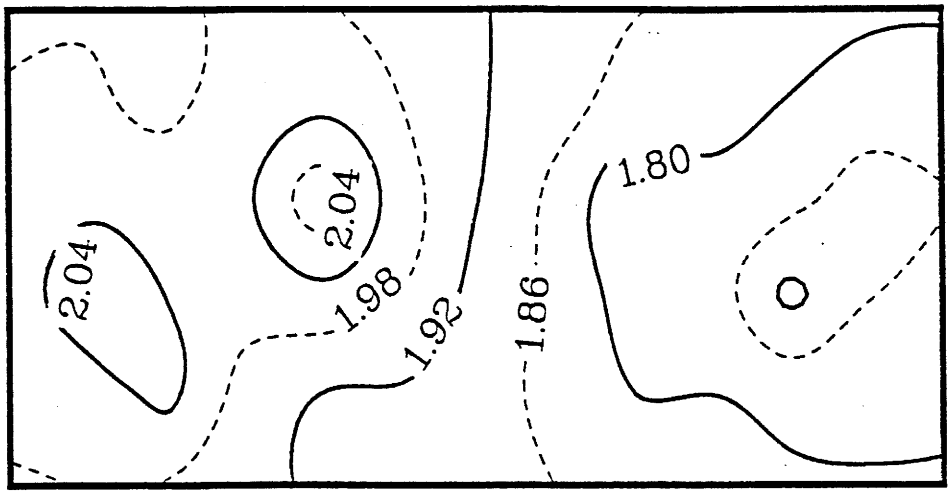


Figure IV.6-6 Comparison of Error Contours from Model I & II for July

Model I, Mean



Model II, Mean

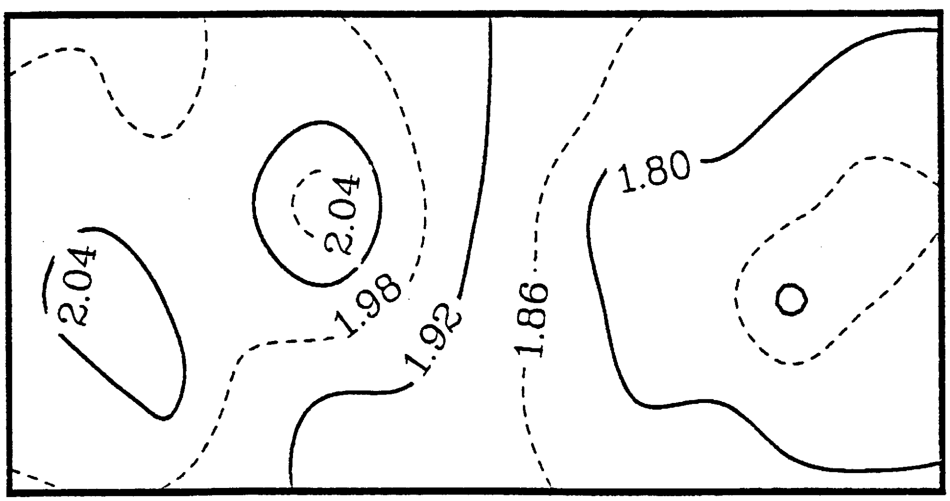
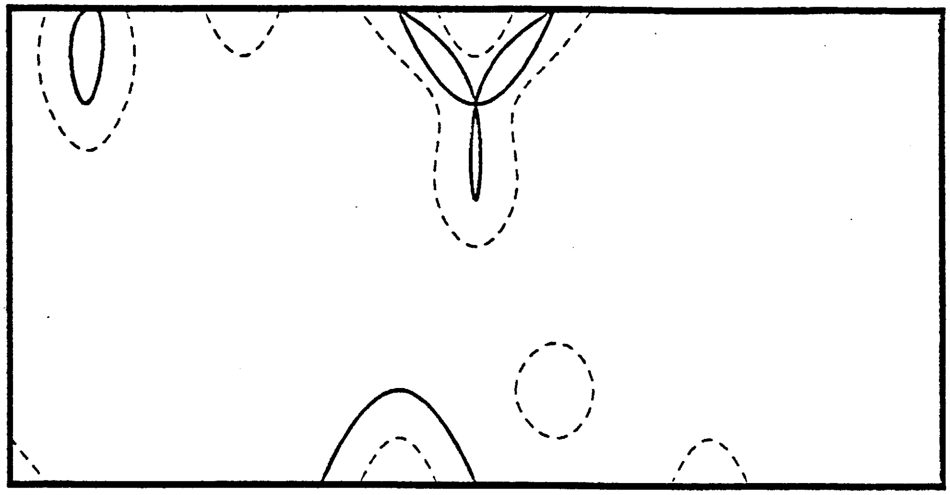


Figure IV.4-7 Comparison of Mean Contours from Model I & II for August

Mean



Error

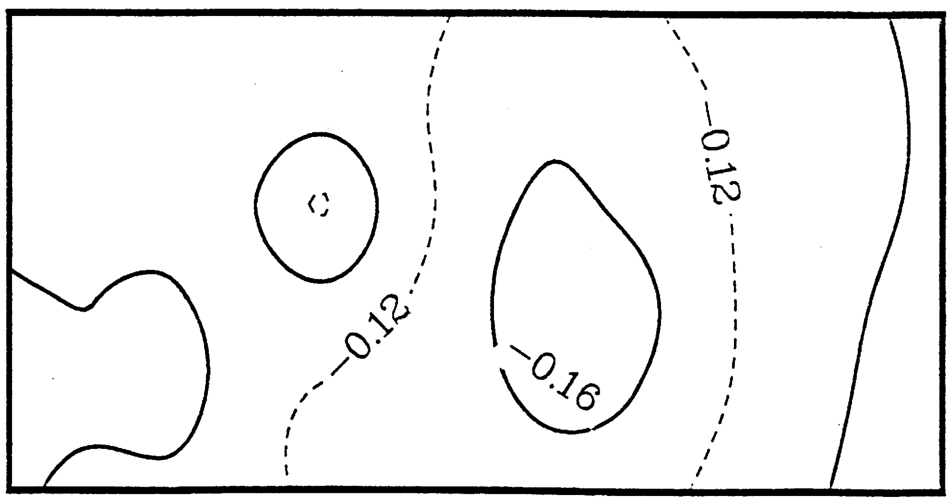
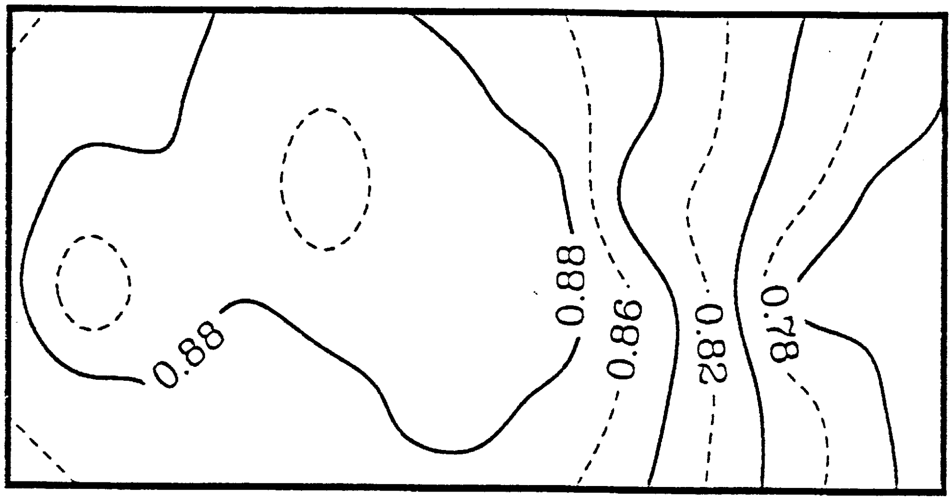


Figure IV.5-7 Contours of Difference between Model I & II for August

Model I, Error



Model II, Error

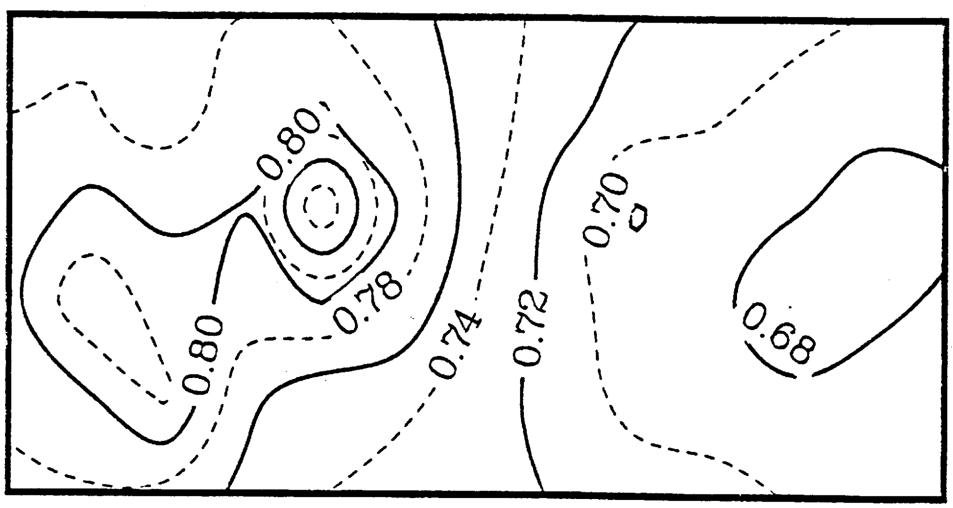
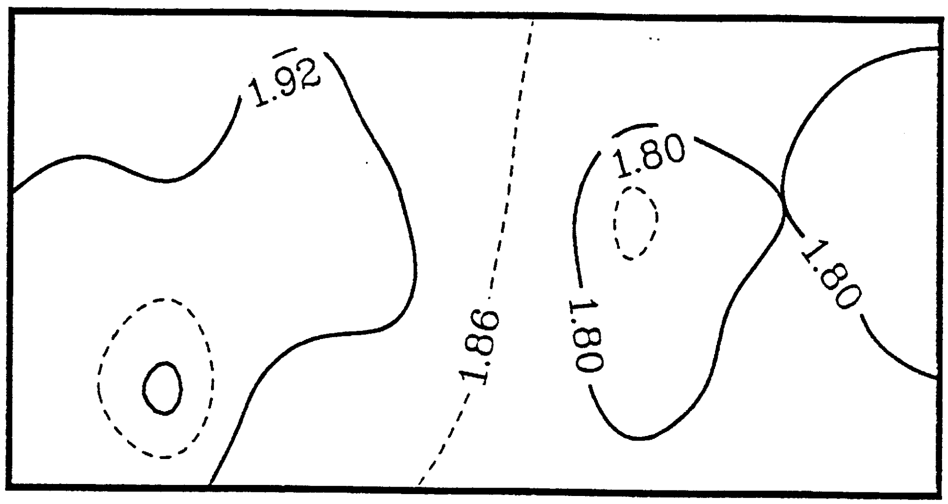


Figure IV.6-7 Comparison of Error Contours from Model I & II for August

Model I, Mean



Model II, Mean

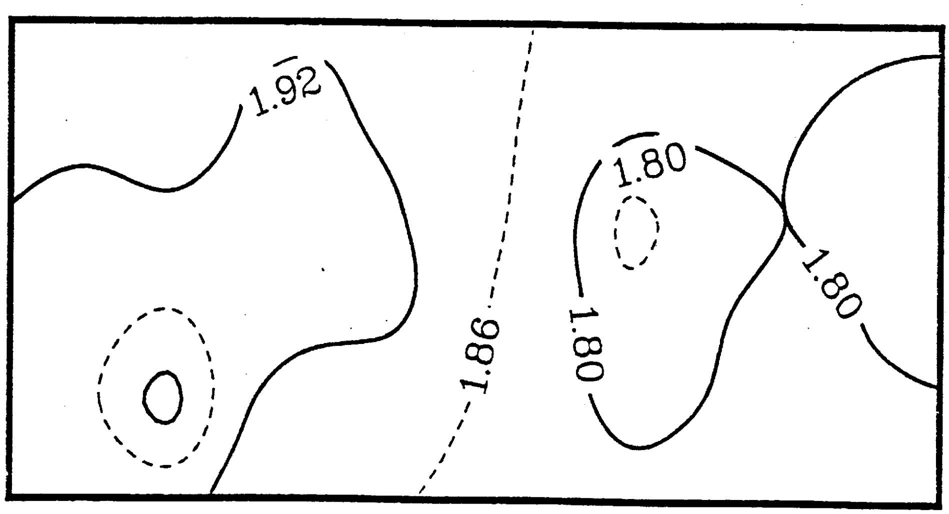
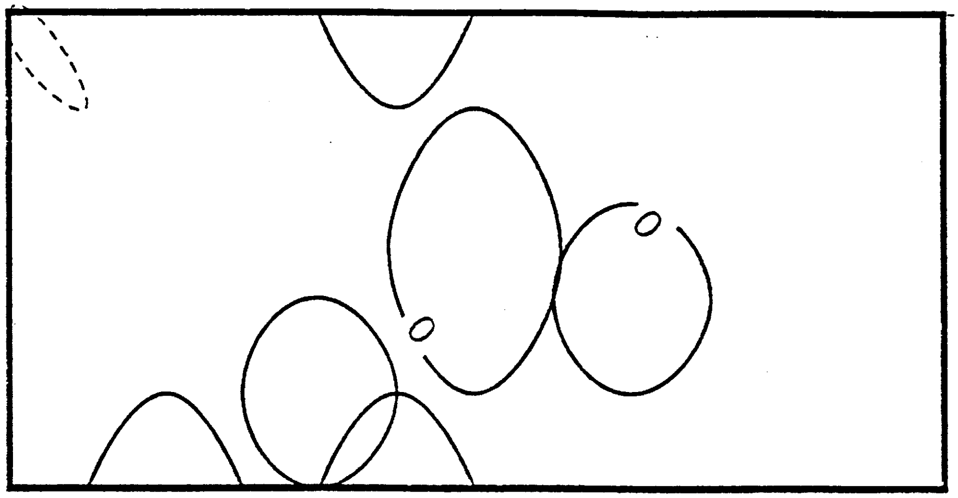


Figure IV.4-8 Comparison of Mean Contours from Model I & II for September

Mean



Error

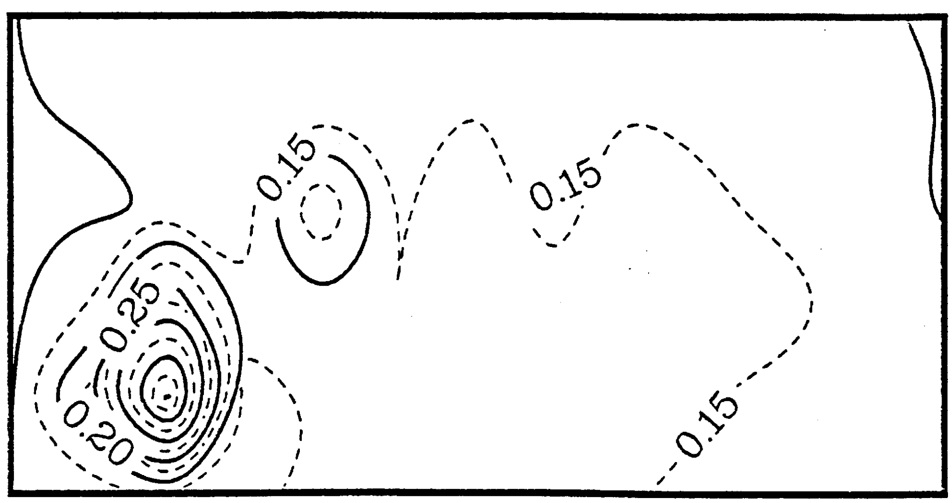
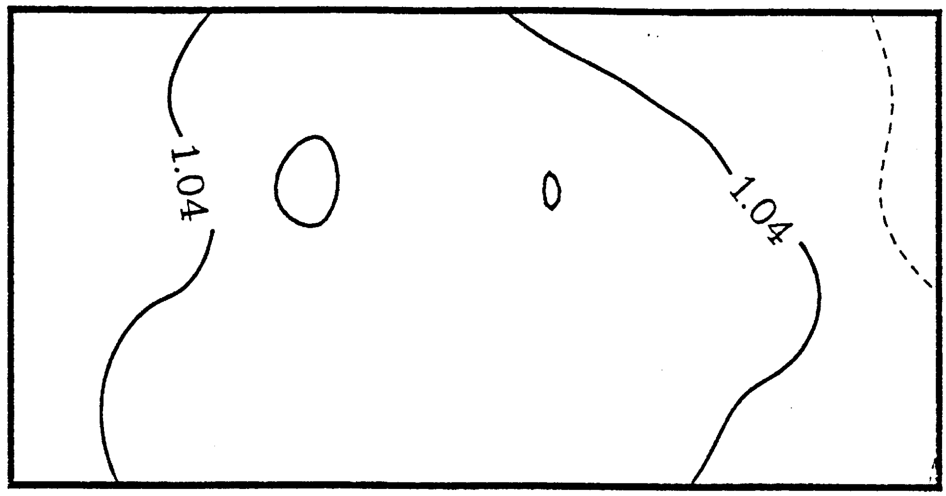


Figure IV.5-8 Contours of Difference between Model I & II for September

Model I, Error



Model II, Error

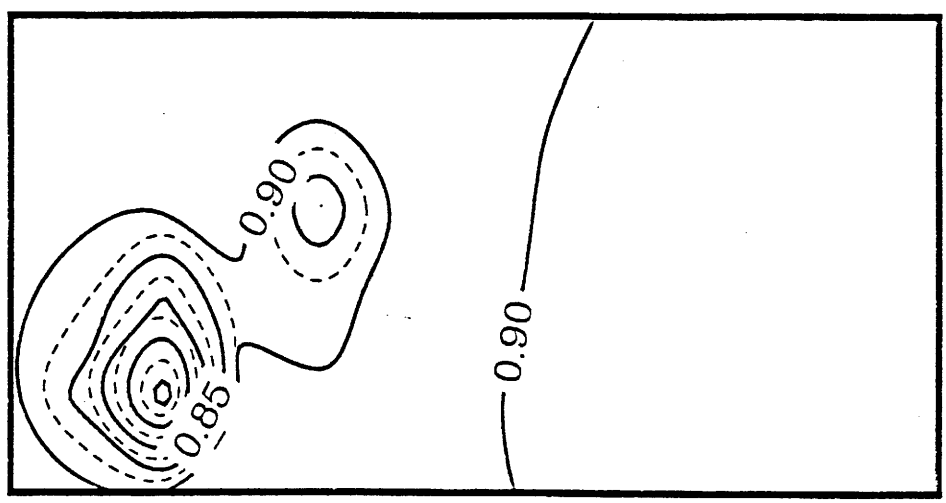
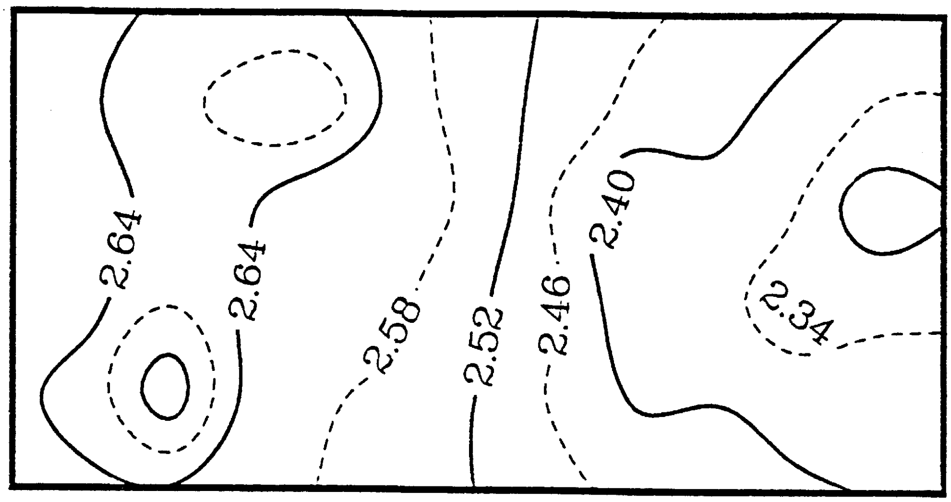


Figure IV.6-8 Comparison of Error Contours from Model I & II for September

Model I, Mean



Model II, Mean

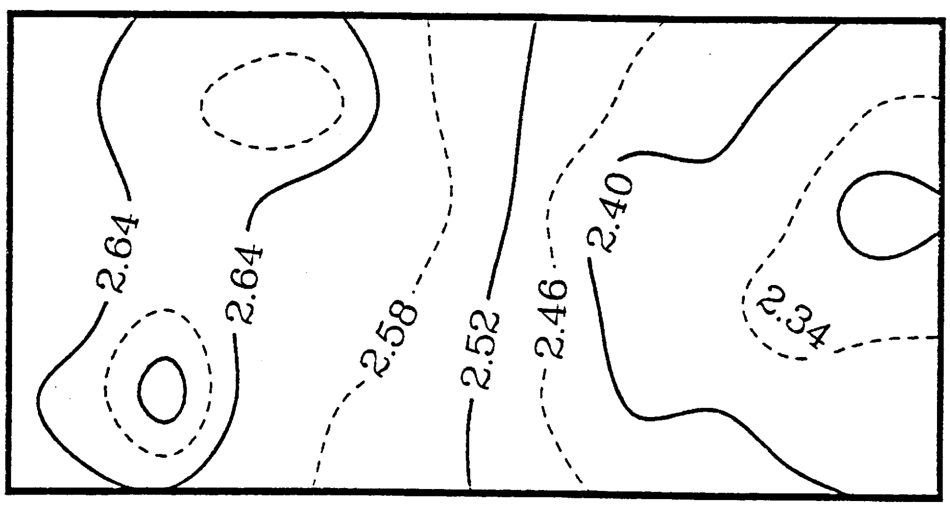


Figure IV.4-9 Comparison of Mean Contours from Model I & II for October

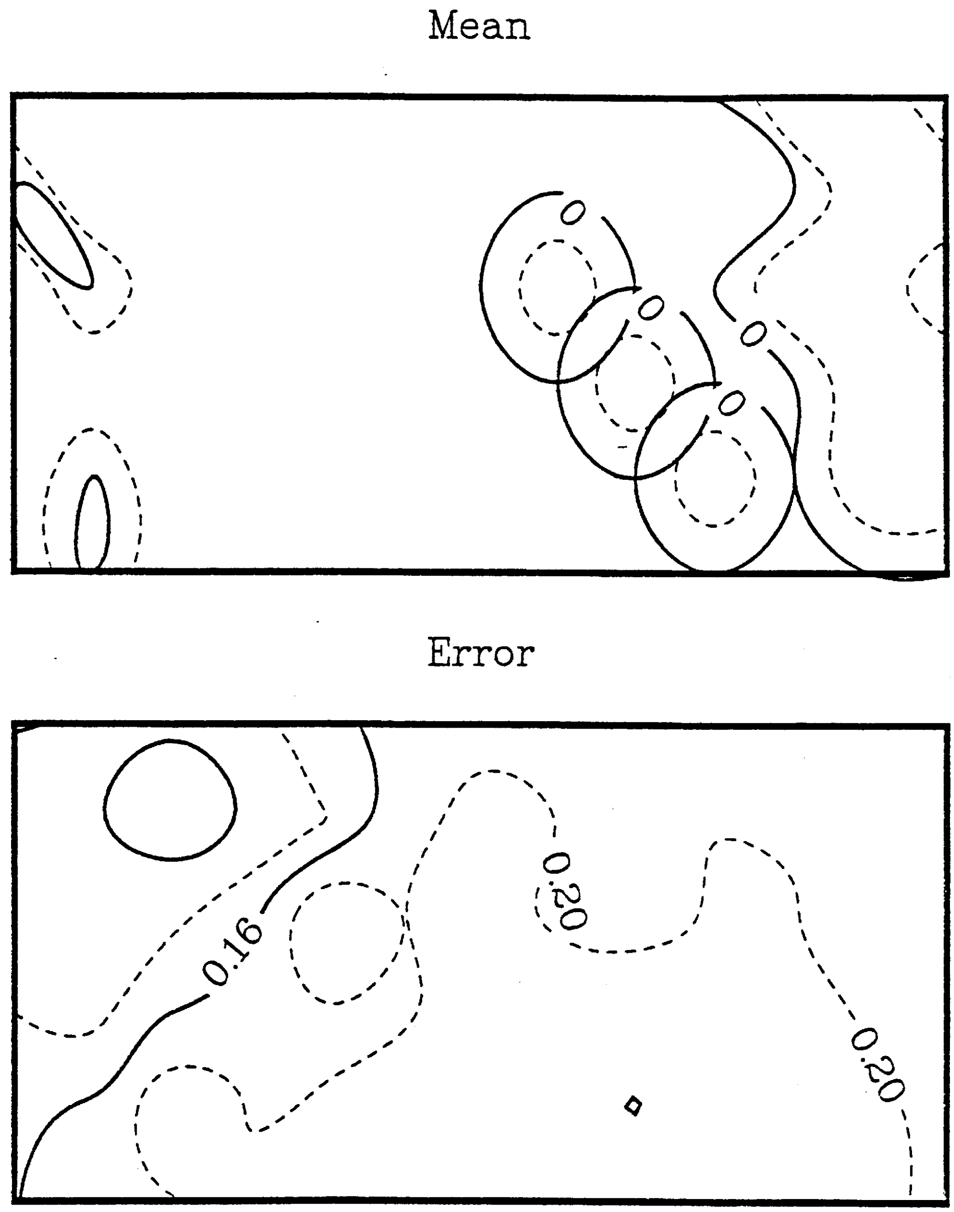
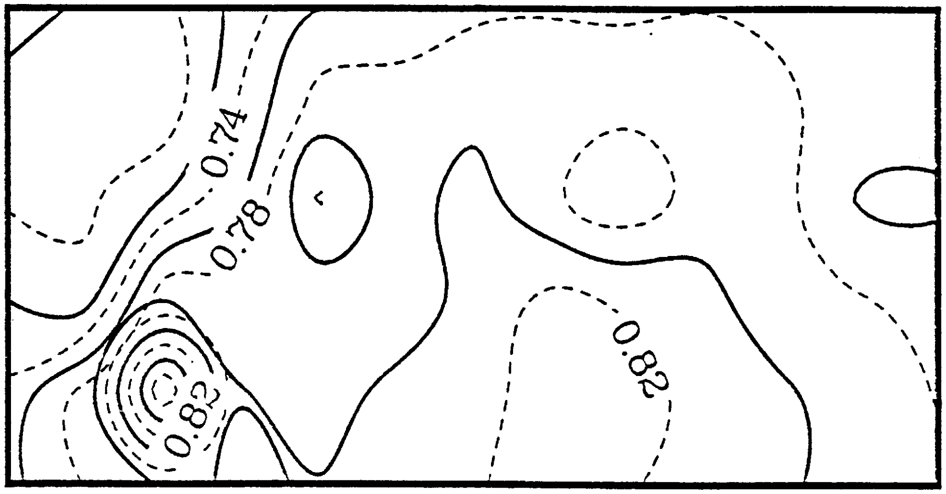


Figure IV.5-9 Contours of Difference between Model I & II for October

Model I, Error



Model II, Error

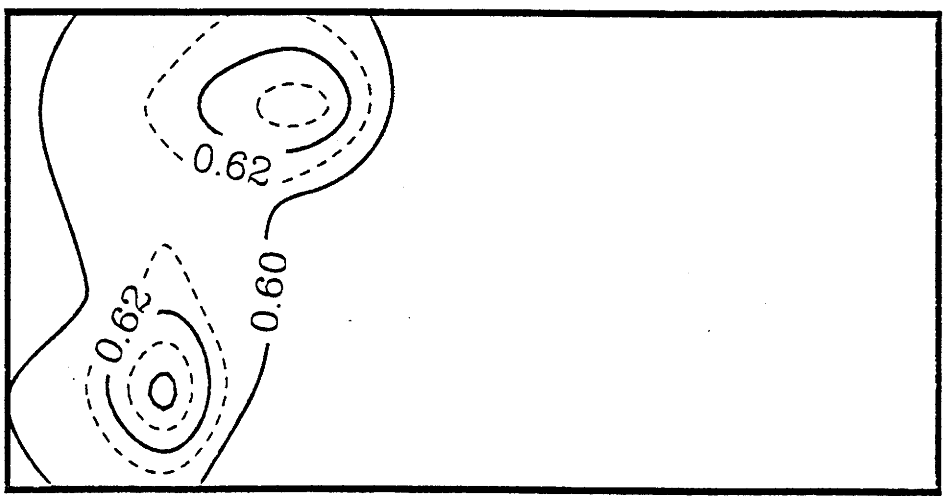
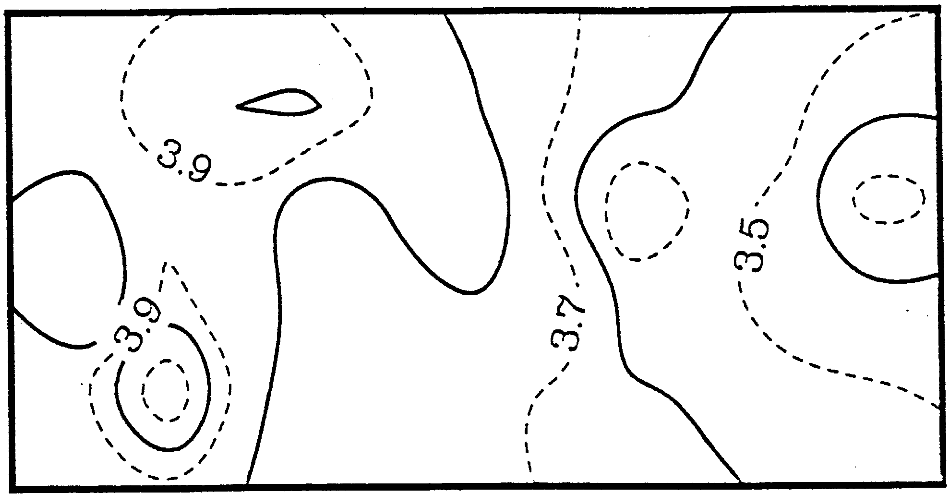


Figure IV.6-9 Comparison of Error Contours from Model I & II for October

Model I, Mean



Model II, Mean

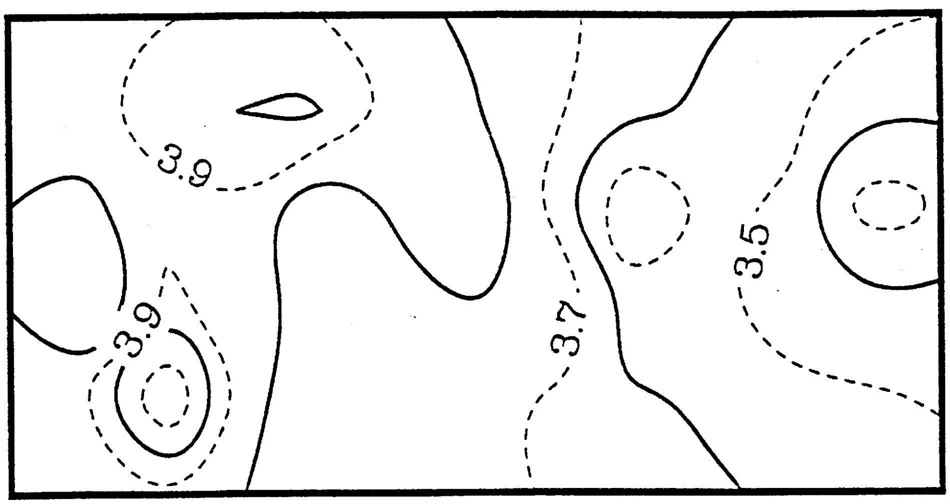
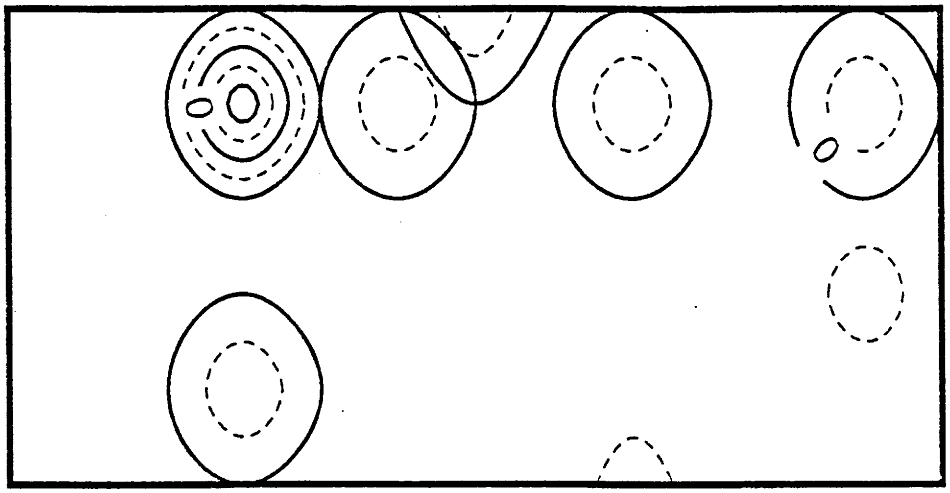


Figure IV.4-10 Comparison of Mean Contours from Model I & II for November

Mean



Error

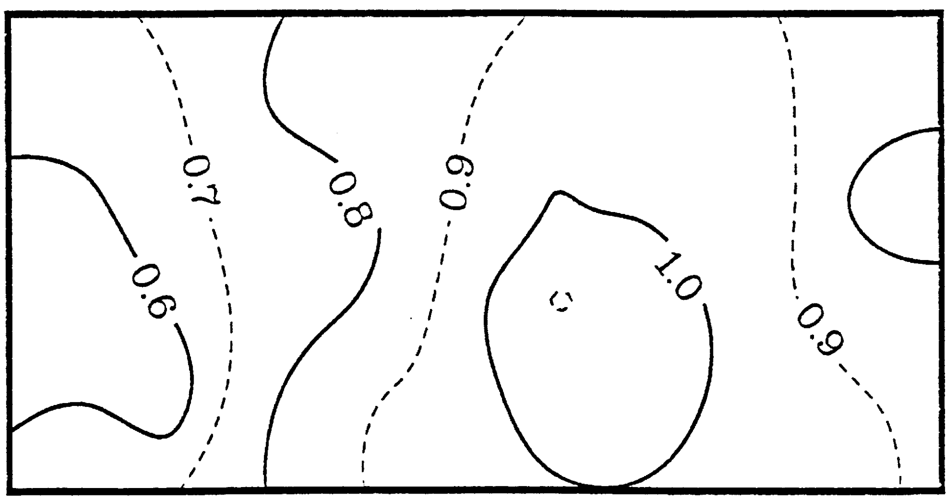
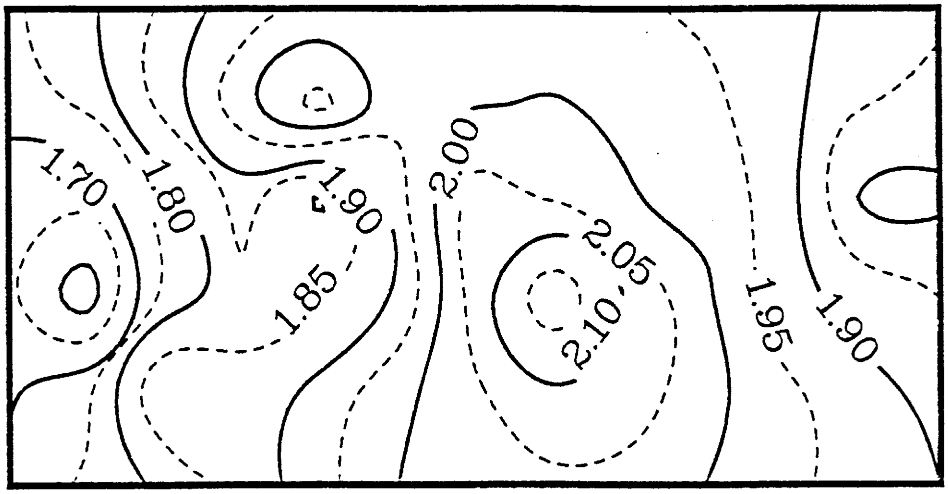


Figure IV.5-10 Contours of Difference between Model I & II for November

Model I, Error



Model II, Error

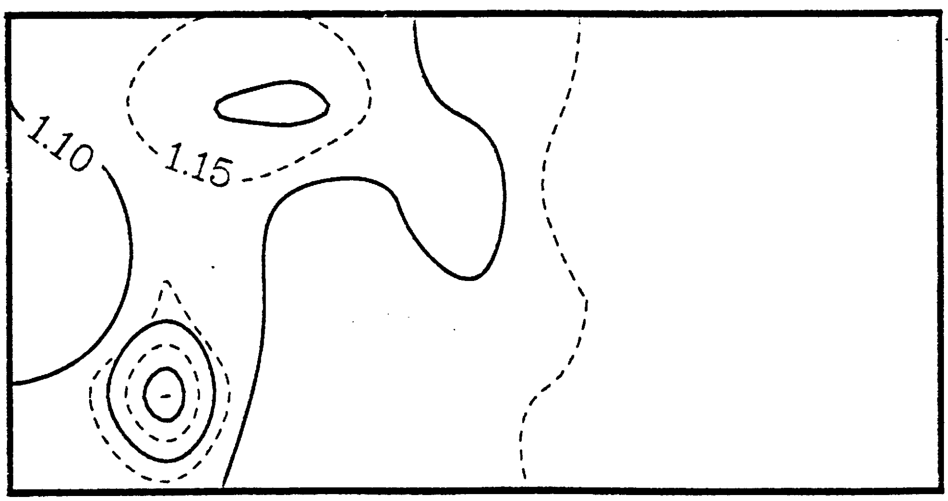
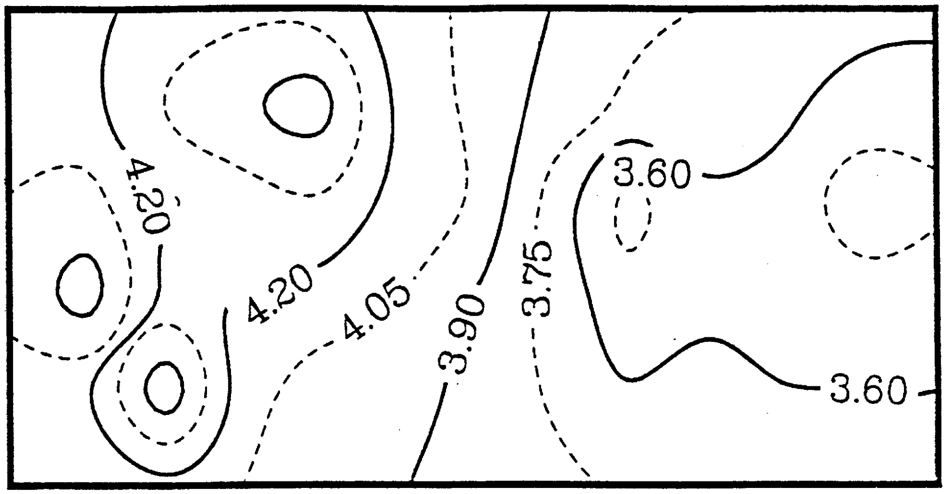


Figure IV.6-10 Comparison of Error Contours from Model I & II for November

Model I, Mean



Model II, Mean

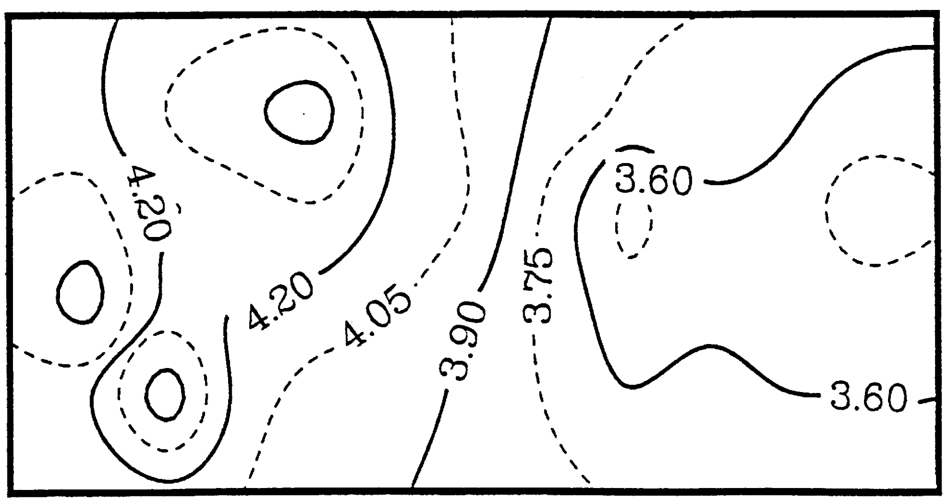


Figure IV.4-11 Comparison of Mean Contours from Model I & II for December

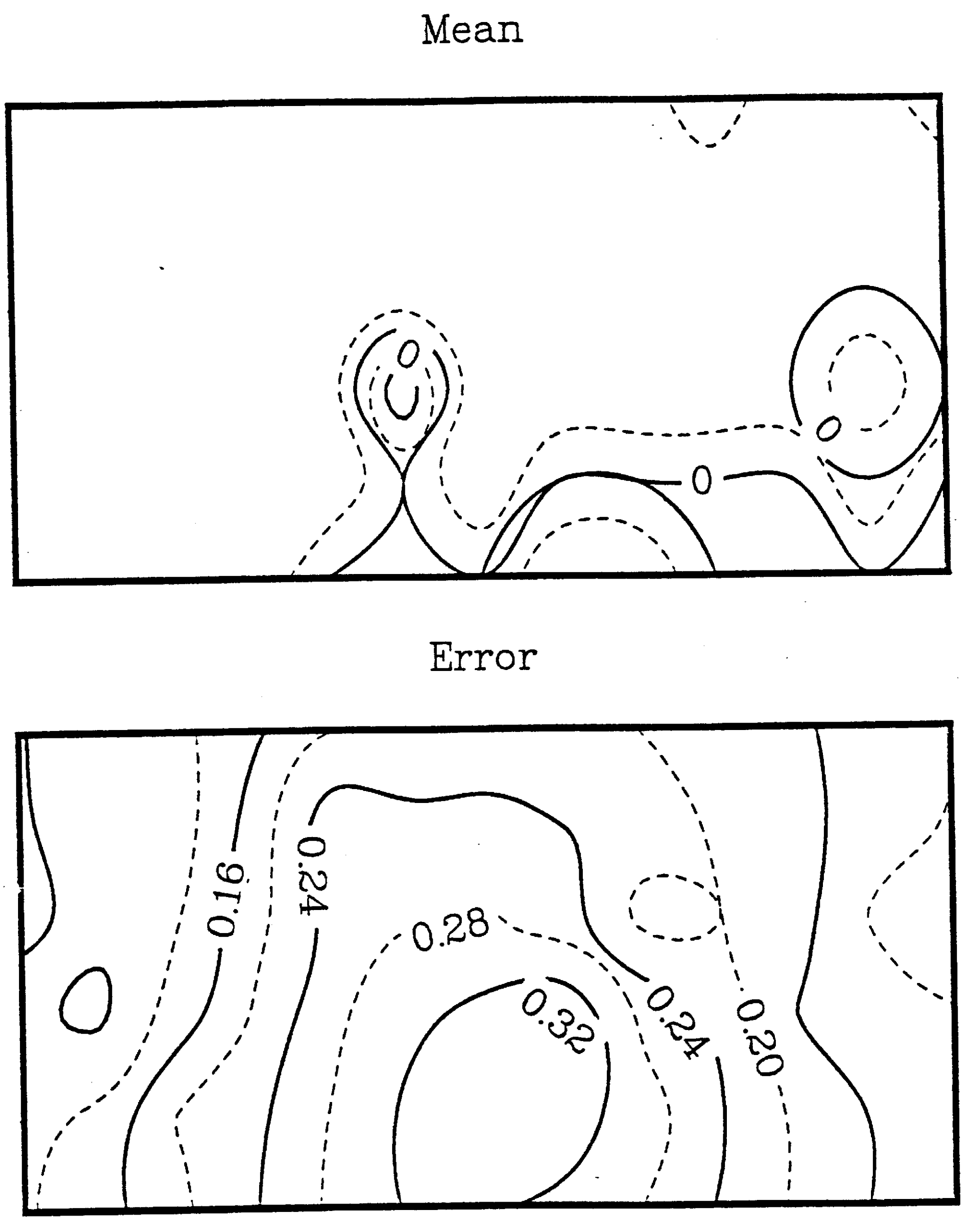


Figure IV.5-11 Contours of Difference between Model I & II for December

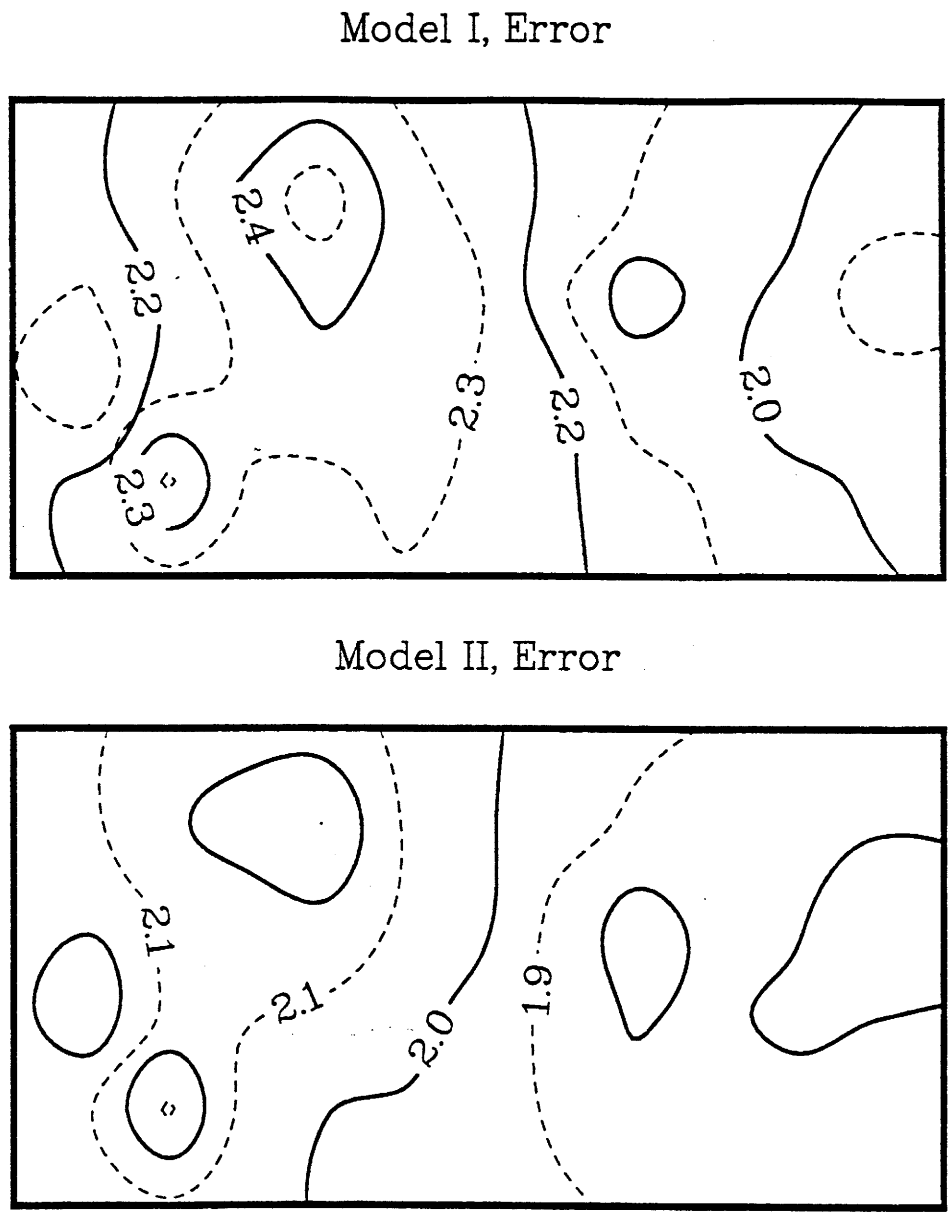


Figure IV.6-11 Comparison of Error Contours from Model I & II for December

Appendix B : Computer Program for Kriging

```
      PROGRAM KRIGING
C
C     THIS PROGRAM IS DOING AS (1) CALCULATES
C     DISTANCES BETWEEN STATIONS WITH THE ALTITUDE,
C     LATITUDE AND LONGITUDE (2)COMPUTES MEANS AND
C     VARIANCES OF INPUT VARIABLES (3) ESTIMATES THE
C     MEANS, VARIANCES AND STANDARD DEVIATIONS USING
C     GAUSSIAN SMOOTHING (4) STANDARDIZE THE VARIABLES
C     TO BE STATIONARY (5) COMPUTES THE COVARIANCES
C     BETWEEN VARIABLES (6) AT EACH ESTIMATION (GRID)
C     POINT, ESTIMATES THE MEANS AND THE STANDARD
C     DEVIATIONS USING INVERSE DISTANCE WEIGHTS AND
C     GAUSSIAN SMOOTHING WEIGHTS FOR CONTOUR MAPS (7)
C     AFTER FINDING THE COVARIANCE FUNCTION FROM THE
C     CORRELATION MATRIX, OBTAIN THE OPTIMUM WEIGHTS
C     FROM NON-STATIONARY KRIGING (GRG2 PROGRAM IS
C     APPLIED TO THE SOLUTION PROCEDURE). THEN MONTHLY
C     PRECIPITATION IS ESTIMATED WITH OPTIMUM WEIGHTS
C     AND THE MEANS AND THE STANDARD DEVIATIONS ARE
C     OBTAINED AGAIN.
C
C     INPUT DATA:
C
C       X(S,T)   = PRECIPITATION VARIABLE AT STATION S AT
C                  MONTH T WHERE S=1,2,..,21 AND T=1,2,..,169
C       RAT(S)   = LATITUDE OF STATION S
C       RONG(S) = LONGITUDE OF STATION S
C       ALT(S)   = ALTITUDE OF STATION S
C       GRIDP(K,3) = CHOSEN K NUMBER OF ESTIMATION (GRID)
C                    POINTS WITH X:Y:Z COORDINATES
C
C     OUTPUT DATA :
C
C       AVCONT(K)   = MONTHLY AVERAGE ESTIMATED BY INVERSE
C                     DISTANCE, GAUSSIAN SMOOTHING OR OPTIMUM
C                     WEIGHTS AT ESTIMATION POINT k,
C                     k=1,2,..,K
C       STDCONT(K) = STANDARD DEVIATION ESTIMATED AT
C                     ESTIMATION POINT k
C
C     VARIABLE :
C
C       DIST(S,S')     = DISTANCE BETWEEN STATION S AND S'
C       DISTN(S)       = DISTANCE BETWEEN STATION S AND
```

```
C                        ESTIMATION POINT AT EACH ITERATION
C       AVERMON(S,M)   = MONTHLY AVERAGE OF VARIABLE AT
C                        STATION S AND A GIVEN MONTH M
C       STDERR(S,M)    = STD_DEV OF      "              "
C       VARIANCE(S,M)  = VARIANCE OF     "              "
C       AVHAT(S,M)     = MONTHLY AVERAGE ESTIMATED BY
C                        GAUSSIAN SMOOTHING AT STATION S AND
C                        A GIVEN MONTH M
C       STDHAT(S,M)    = STD_DEV ESTIMATED     "        "
C       AC(M)          = OPTIMAL C FOR THE MEAN COMPUTED BY
C                        FIBONACCI SEARCH AT A GIVEN MONTH
C       SC(M)          = OPTIMAL C FOR THE STD_DEV
C       VC(M)          = OPTIMAL C FOR THE VARIANCE
C       COV(S,S')      = CORRELATION BETWEEN STATION S AND S'
C       SW(S)          = INVERSE DISTANCE WEIGHT
C       GW(S)          = GAUSSIAN SMOOTHING WEIGHT
C       GRID(3)        = TEMPORARY VARIABLE OF ESTIMATIO
C                        POINT
C       GRIDP(K,3)     = LOCATION VARIABLE OF     "       "
C       VCONTOUR(K,T)  = ESTIMATED DATA USING OPTIMUM WEIGHT
C       RHO1(S)        = CORRELATION BETWEEN STATION S AND
C                        EACH ESTIMATION POINT FITTED BY THE
C                        VARIOGRAM
C       RHO2(S,S')     = CORRELATION BETWEEN STATION S AND S'
C                        FITTED BY THE VARIOGRAM
C       C(S,S')        = ESTIMATED LHS COVARIANCE ON KRIGING
C                        SYSTEM
C       COI(S)         =     "     RHS        "             "
C       AMU(S)         = MEAN VALUE ON NONSTATIONARY KRIGING
C       AMU0           = ESTIMATED MEAN OF ESTIMATION POINT
C       TOTVAR         = ESTIMATED VARIANCE OVER THE BASIN
C       AVCONT(K)      = MONTHLY AVERAGE COMPUTED BY THE DATA
C                        ESTIMATED FROM EACH WEIGHTING METHOD
C       STDCONT(K)     = STD_DEV          "                "
C       VARCONT(K)     = VARIANCE         "                "
C
      implicit double precision(a-h,o-z), integer(i-n)
      PARAMETER(itime=169,MONTH=12,istat=21,igrid=32)
C
      common /coma/DIST(istat,istat)
      common /comb1/c(istat,istat)
      common /comb2/coi(istat)
      common /comb3/amu(istat)
      common /comb4/amu0
      common /comc/imon
      common /comd/distn(istat)
      common /comg/totvar
C
      DIMENSION RAT(istat), RONG(istat), ALT(istat)
```

```
      DIMENSION X(istat,itime), Z(istat,itime)
      DIMENSION AVERMON(istat,MONTH), STDERR(istat,MONTH),
     &          VARIANCE(istat,MONTH)
      DIMENSION avhat(istat,month), stdhat(istat,month),
     &          varhat(istat,month)
      DIMENSION ac(MONTH), sc(MONTH), vc(MONTH)
      DIMENSION COV(istat,istat)
      DIMENSION sw(istat), gw(istat)
      DIMENSION GRID(3), GRIDP(igrid,3),
     1          VCONTOUR(igrid,itime)
      DIMENSION RHO1(istat), rho2(istat,istat)
      DIMENSION avcont(igrid), stdcont(igrid),
     1          varcont(igrid)
      DIMENSION skip(igrid)
      CHARACTER FILNAME*80, NAMES(istat)*4, NAMEM(MONTH)*3
C
      OPEN(UNIT=5,FILE='PREcLOC.DAT',STATUS='OLD')
      OPEN(uNIT=50,FILE='GRIDPNT.DAT',STATUS='OLD')
C      OPEN(uNIT=50,FILE='WGRIDPNT.DAT',STATUS='OLD')
C
      OPEN(UNIT=51,FILE='PM0101.DAT',STATUS='OLD')
      OPEN(UNIT=52,FILE='PM0102.DAT',STATUS='OLD')
      OPEN(UNIT=53,FILE='PM01031.DAT',STATUS='OLD')
      OPEN(UNIT=54,FILE='PM01032.DAT',STATUS='OLD')
      OPEN(UNIT=55,FILE='PM0103A.DAT',STATUS='OLD')
      OPEN(UNIT=56,FILE='PM0106.DAT',STATUS='OLD')
      OPEN(UNIT=57,FILE='PM01082.DAT',STATUS='OLD')
      OPEN(UNIT=58,FILE='PM0108A.DAT',STATUS='OLD')
      OPEN(UNIT=59,FILE='PM0109.DAT',STATUS='OLD')
      OPEN(UNIT=60,FILE='PM01152.DAT',STATUS='OLD')
      OPEN(UNIT=61,FILE='PM0115A.DAT',STATUS='OLD')
      OPEN(UNIT=62,FILE='PM0119.DAT',STATUS='OLD')
      OPEN(UNIT=63,FILE='PM0120.DAT',STATUS='OLD')
      OPEN(UNIT=64,FILE='PM0121.DAT',STATUS='OLD')
      OPEN(UNIT=65,FILE='PM0121A.DAT',STATUS='OLD')
      OPEN(UNIT=66,FILE='PM0122.DAT',STATUS='OLD')
      OPEN(UNIT=67,FILE='PM0123.DAT',STATUS='OLD')
      OPEN(UNIT=68,FILE='PM0124.DAT',STATUS='OLD')
      OPEN(UNIT=69,FILE='PM0125.DAT',STATUS='OLD')
      OPEN(UNIT=70,FILE='PM0126.DAT',STATUS='OLD')
      OPEN(UNIT=71,FILE='PM0127.DAT',STATUS='OLD')
      OPEN(UNIT=6,FILE='krige.OUT',STATUS='UNKNOWN')
C
      READ(5,*)((Rong(ISTA),Rat(ISTA),ALT(ISTA)),
     1         ISTA=1,istat)
      READ(50,*)((GRIDP(I,J),J=1,3),I=1,igrid)
      READ(51,*)(X(1,K),K=1,itime)
      READ(52,*)(X(2,K),K=1,itime)
      READ(53,*)(X(3,K),K=1,itime)
```

```
      READ(54,*)(X(4,K),K=1,itime)
      READ(55,*)(X(5,K),K=1,itime)
      READ(56,*)(X(6,K),K=1,itime)
      READ(57,*)(X(7,K),K=1,itime)
      READ(58,*)(X(8,K),K=1,itime)
      READ(59,*)(X(9,K),K=1,itime)
      READ(60,*)(X(10,K),K=1,itime)
      READ(61,*)(X(11,K),K=1,itime)
      READ(62,*)(X(12,K),K=1,itime)
      READ(63,*)(X(13,K),K=1,itime)
      READ(64,*)(X(14,K),K=1,itime)
      READ(65,*)(X(15,K),K=1,itime)
      READ(66,*)(X(16,K),K=1,itime)
      READ(67,*)(X(17,K),K=1,itime)
      READ(68,*)(X(18,K),K=1,itime)
      READ(69,*)(X(19,K),K=1,itime)
      READ(70,*)(X(20,K),K=1,itime)
      READ(71,*)(X(21,K),K=1,itime)
C
      DATA NAMES/'101','102','1031','1032','103A','106',
     &            '1082','108A','109','1152','115A','119',
     &            '120','121','121A','122','123','124',
     &            '125','126','127'/
      DATA NAMEM /'JUN','JUL','AUG','SEP','OCT','NOV','DEC',
     &             'JAN','FEB','MAR','APR','MAY'/
C
      FILNAME = 'KRIGE00'
C                                GET THE DISTANCE BETWEEN
C                                STATIONS
      CALL DISTANT (RAT,RONG,ALT)
C                                GET THE MONTHLY AVERAGE,
C                                THE STD_DEV AND THE
C                                VARIANCE AT THE GIVEN
C                                STATION
      CALL AVERAGE (X,AVERMON,STDERR,VARIANCE)
c                                ESTIMATE MEAN, STD_DEV AND
C                                VARIANCE WITH GAUSSIAN
C                                SMOOTHING AND OBTAIN
C                                THE OPTIMUM C
      call GAUSSSMO (avermon,ac,avhat)
      call GAUSSSMO (stderr,sc,stdhat)
      call GAUSSSMO (variance,vc,varhat)
C                                STANDARDIZE THE VARIABLES
      CALL STANDARD (x,avermon,stderr,Z)
C                                GET THE COVARIANCE OF
C                                STANDARDIZED VARIABLES
C                                BETWEEN TWO DIFFERENT
C                                STATIONS
      CALL CORRCOV (z,cov,const)
```

```
c                                  ITERATION FOR EACH MONTH
      DO 10 IMON = 1, MONTH
        IF (IMON.LE.9) THEN
         WRITE(FILNAME(7:7),'(I1)') IMON
        ELSEIF (IMON.LE.99) THEN
         WRITE(FILNAME(6:7),'(I2)') IMON
        ENDIF
        CLOSE(UNIT=10)
        OPEN(10,FILE=FILNAME,STATUS='NEW')
C                                  Decide the grid points
        DO 30 iter =1,igrid
         DO 31 I = 1, 3
   31     grid(i) = gridp(iter,i)
c                                  Get the new distances
C                                  between each grid point
C                                  and station S
          call DISTGRID (grid,rat,rong,alt)
c                                  Choose the weighting
C                                  method
          method=3
          if(method.eq.1) then
c                                  Get the 'IVDIS' weights
           CALL IDWEIGHT (sw)
c                                  Estimate the means and
C                                  std_dev for contour line
C                                  using the weights
           call ESTIMATE
C                   (sw,iter,avermon,stderr,avcont,stdcont)
c
          elseif(method.eq.2) then
c                                  Get the weights from
C                                  Gaussian Smoothing using
C                                  the optimum  C, then
C                                  estimate the means
c                                  and std_dev using the
C                                  weights
           call GAWEIGHT (ac,gw)
           call ESTIMATE
                (gw,iter,avermon,stderr,avcont,skip)
           call GAWEIGHT (sc,gw)
           call ESTIMATE
                (gw,iter,avermon,stderr,skip,stdcont)
c
          elseif(method.eq.3) then
           do 32 i=1,istat
   32       amu(i) = avermon(i,imon)
c                                  Get the initial weights
C                                  for GRG2
           CALL IDWEIGHT (const,sw)
```

```
c                                  Get the estimated RHS and C
                                LHS correlation matrices
          call RLHSRHO(const,rhol,rho2)
c                                  Get the estimated LHS and C
                                RHS Covariance matrices
C                                  for Kriging
          CALL RLHSCOV (sw,cov,stderr,avermon,rhol,rho2)
c                                  Get the optimum weights
C                                  from GRG2
          call OPTWEIHT(sw)
c                                  Get the estimated monthly
c                                  PRECIPITATION using the
C                                  optimum weights FOR given
C                                  month
          do 33 jj=imon,itime,12
           VSUM = 0.
           DO 34 II = 1,istat
            if(x(ii,jj).ne.9999.) then
             VSUM = VSUM + sw(II)*x(II,jj)
            else
             vsum=9999.
            endif
   34      continue
   33      VCONTOUR(iter,jj) = VSUM
c                                  Get the mean and variance C
                                of estimated monthly TOTAL
c                                  precipitation from kriging
          call contaver
C              (vcontour,iter,avcont,stdcont,varcont)
         endif
   30   continue
c
        do 40 i=1,igrid
   40    write(10,110) gridp(i,1),gridp(i,2),
C    &                 avcont(i),stdcont(i)
   10 continue
  110 format (32(4x,f5.3,3x,f5.3,3x,f11.7,3x,f11.7))
      STOP
      END
C
C*******************************************************
C*******************************************************
C
C   COMPUTE THE AVERAGE AND STANDARD DEVIATION OF DATA
C
      SUBROUTINE AVERAGE (X,AVERMON,STDERR,VAR)
C
C     INPUT DATA :
C       X(S,T)       = MONTHLY PRECIPITATION DATA AT STATION
```

```
C                        S AND AT MONTH T
C
C      OUTPUT DATA :
C        AVERMON(S,M) = MONTHLY MEAN OF DATA AT STATION S
C                       AT A GIVEN M
C        STDERR(S,M)  = STD_DEV OF DATA "          "
C        VAR(S,M)     = VARIANCE OF DATA "          "
C
      implicit double precision(a-h,o-z), integer(i-n)
      PARAMETER(itime=169,istat=21,MONTH=12)
C
      dimension X(istat,itime), AVERMON(istat,MONTH),
     &     STDERR(istat,MONTH), VAR(istat,MONTH)
C
      DO 10 ISTA = 1, istat
       DO 20 MON = 1, MONTH
        SSUM = 0.
        ACOUNT = 0.
        DO 40 itim = MON, itime, MONTH
         IF (X(ISTA,itim).NE.9999.) THEN
          SSUM = SSUM + X(ISTA,itim)
          ACOUNT = ACOUNT + 1.
         ELSE
         ENDIF
   40   CONTINUE
        AVERMON(ISTA,MON) = SSUM / ACOUNT
        SUM = 0.
        DO 50 itim = MON, itime, MONTH
         IF (X(ISTA,itim).NE.9999.) THEN
          SUM = SUM + (X(ISTA,itim)-AVERMON(ISTA,MON))**2
         ENDIF
   50   CONTINUE
        VAR(ISTA,MON)=SUM/(ACOUNT-1.)
        STDERR(ISTA,MON) = DSQRT(VAR(ISTA,MON))
   20  CONTINUE
   10 CONTINUE
      RETURN
      END
C
C ************************************************************
C ************************************************************
C
C   COMPUTE THE CORRELATION BETWEEN TWO STATIONS
C
      SUBROUTINE CORRCOV (SX,SCOV,const)
C
C     INPUT DATA :
C       SX(S,T)      = THE STANDARDIZED MONTHLY PRECIPITATION
C                      DATA AT STATION S AND IN MONTH T
```

```
C
C     OUTPUT DATA :
C       SCOV(S,S')   = CORRELATION BETWEEN STATION S AND S'
C       CONST        = AVERAGE OF CORRELATIONS RELATED TO
C                      STATION 108-2
C
      implicit double precision(a-h,o-z), integer(i-n)
      PARAMETER (istat=21,itime=169)
      dimension SX(istat,itime),SCORR(istat,istat),
     &     SCOV(istat,istat)
C
      DO 10 ISTA = 1, istat
       DO 20 JSTA = 1, istat
        XSUM = 0.
        YSUM = 0.
        DF = 0.
         DO 30 I = 1,itime
          IF (SX(ISTA,I).NE.9999.) THEN
           DF = DF + 1.
           XSUM = XSUM + SX(ISTA,I)
           YSUM = YSUM + SX(JSTA,I)
          ENDIF
   30    CONTINUE
C
         XMEAN = XSUM / DF
         YMEAN = YSUM / DF
C
         RNUM = 0.
         SSX = 0.
         SSY = 0.
         DO 40 I = 1,itime
          IF (SX(ISTA,I).NE.9999.) THEN
           RNUM = RNUM + (SX(ISTA,I)-XMEAN)*
     1             (SX(JSTA,I)-YMEAN)
           SSX = SSX + (SX(ISTA,I)-XMEAN)**2
           SSY = SSY + (SX(JSTA,I)-YMEAN)**2
          ENDIF
   40    CONTINUE
         DENUM = DSQRT(SSX*SSY)
         SCORR(ISTA,JSTA) = RNUM / DENUM
         SCOV(ISTA,JSTA) = SCORR(ISTA,JSTA)
   20  CONTINUE
   10 CONTINUE
      sum0=0.
      ccount=0.
      do 80 i=1,istat
       if(i.ne.7) then
        ccount=ccount+1.
        sum0=sum0+scov(i,7)
```

```
         endif
   80 continue
        const=sum0/ccount
        do 81 i=1,istat
         if(i.ne.7) then
          scov(i,7)=const
          scov(7,i)=const
         endif
   81 continue
        RETURN
        END
C
C*********************************************************
C*********************************************************
C
C    STANDARDIZE THE DATA
C
      SUBROUTINE STANDARD (x,aver,stder,SZ)
C
C     INPUT DATA :
C       X(S,T)      = MONTHLY PRECIPITATION DATA
C       AVER(S,M)   = MONTHLY AVERAGE OF DATA AT STATION
C                     S AND IN A GIVEN MONTH M
C       STDER(S,M)  = STANDARD DEVIATION OF DATA AT STATION S
C                     AND IN A GIVEN MONTH M
C
C     OUTPUT DATA :
C       SZ(S,T)     = STANDARDIZED DATA AT STATION S AND IN
C                     MONTH T
C
      implicit double precision(a-h,o-z),integer(i-n)
      PARAMETER(istat=21,MONTH=12,itime=169)
      dimension x(istat,itime),SZ(istat,itime),
     &     aver(istat,MONTH),stder(istat,MONTH)
C
      DO 10 ISTA = 1, istat
       DO 20 MON = 1, MONTH
        DO 30 I = MON, itime, 12
         IF (x(ISTA,I).NE.9999.) THEN
         sz(ista,i)=(x(ista,i)-aver(ista,mon))/
     7               stder(ista,mon)
         else
          sz(ista,i)=9999.
         ENDIF
   30   CONTINUE
   20  CONTINUE
   10 CONTINUE
      RETURN
      END
```

```
C
C ***********************************************************
C ***********************************************************
C
C   GET THE DISTANCE BETWEEN TWO STATIONS USING
C   THE LATITUDE(MILE),LONGITUDE(MILE) AND ALTITUDE(MILE)
C   OF EACH STATION.
C
      SUBROUTINE DISTANT(SLAT,SLONG,SALT)
C
C     INPUT DATA :
C       SLAT(S)    = LATITUDE OF STATION S
C       SLONG(S)   = LOGITUDE OF STATION S
C       SALT(S)    = ALTITUDE OF STATION S
C
C     OUTPUT DATA :
C       DIST(S,S') = DISTANCE BETWEEN STATION S AND S'
C
       implicit double precision(a-h,o-z),integer(i-n)
       PARAMETER (istat=21)
C
       common /coma/disT(istat,istat)
       dimension SLAT(istat), SLONG(istat), SALT(istat)
C
       DO 44 I = 1, istat
        DO 33 J = 1, istat
         RATIT = ABS(SLAT(I)-SLAT(J))
         RONGIT = ABS(SLONG(I)-SLONG(J))
         ALTIT = ABS(SALT(I)-SALT(J))/5280.
         DIST(I,J) = DSQRT(RATIT**2+RONGIT**2+ALTIT**2)
   33   CONTINUE
   44  CONTINUE
C
      RETURN
      END
C
C *************************************************************
C *************************************************************
C
C   GET THE DISNCE BETWEEN EXISTING STATION AND THE
C   ESTIMATION POINT
C
      subroutine distgrid (grid,rat,rong,alt)
c
C     INPUT DATA :
C      GRID(X,Y,Z)  = THE COORDINATES OF EACH ESTIMATION
C                     POINT
C      RAT(S)       = LATITUDE OF STATION S
C      RONG(S)      = LOGITUDE OF STATION S
```

```
C      ALT(S)       = ALTITUDE OF STATION S
C
C     OUTPUT DATA :
C       DISTN(S)      = DISTANCE BETWEEN A GIVEN ESTIMATION
C                       POINT AND STATION S
C
      implicit double precision(a-h,o-z),integer(i-n)
      parameter(istat=21)
      common /comd/distn(istat)
      dimension grid(3),rat(istat),rong(istat),alt(istat)
c
      do 10 ista=1,istat
       a=grid(1)-rong(ista)
       b=grid(2)-rat(ista)
       c=(grid(3)-alt(ista))/5280.
   10  distn(ista)=dsqrt(a**2+b**2+c**2)
      return
      end
c
c**********************************************************
c**********************************************************
c
C   CALCULATE THE INVERSE DISTANCE WEIGHT
C
      subroutine idweight (w)
c
C     INPUT DATA :
C       DISTN(S)    = DISTANCE BETWEEN STATION S AND THE
C                     ESTIMATION POINT
C
C     OUTPUT DATA :
C       W(S)        = INVERSE DISTANCE WEIGHT OF STATION S
C
      implicit double precision(a-h,o-z),integer(i-n)
      parameter(istat=21)
      common /comd/distn(istat)
      dimension w(istat)
c
      dsum=0.
      do 11 ista=1,istat
   11  dsum=dsum+1./distn(ista)
      do 12 ista=1,istat
        w(ista)=(1./distn(ista))/dsum
   12 continue
      return
      end
c
c**********************************************************
c**********************************************************
```

```
c
C   COMPUTE THE CORRELATION USING THE CORRELATION FUNCTION
C
      subroutine RLHSRHO(const,rho1,rho2)
c
C     INPUT DATA :
C       DIST(S,S')   = DISTANCE BETWEEN STATION S AND S'
C       DISTN(S)     = DISTANCE BETWEEN STATION S AND
C                      ESTIMATION POINT
C
C     OUTPUT DATA :
C       RHO1(S)      = ESTIMATED CORRELATION BETWEEN STATION
C                      S AND ESTIMATION POINT
C       RHO2(S,S')   = ESTIMATED CORRELATION BETWEEN STATION
C                      S AND S'
C
      implicit double precision(a-h,o-z),integer(i-n)
      parameter(istat=21)
      common /coma/dist(istat,istat)
      common /comd/distn(istat)
      dimension rho1(istat),rho2(istat,istat)
c
      do 10 ista=1,istat
       if(distn(ista).eq.0.) then
        rho1(ista)=1.
       else
        if(ista.ne.7) then
c
        rho1(ista)=.927988-.026262*distn(ista)
        else
C                                  FOR STATION 108-2
         rho1(ista)=const
        endif
       endif
c
       do 11 jsta=1,istat
        if(dist(ista,jsta).eq.0.) then
         rho2(ista,jsta)=1.
        else
         if(ista.ne.7.or.jsta.ne.7) then
c
          rho2(ista,jsta)=.927988-.026262*dist(ista,jsta)
         else
C                                  FOR STAITON 108-2
          rho2(ista,jsta)=const
         endif
        endif
   11  continue
   10 continue
```

```
      return
      end
c
c*********************************************************
C*********************************************************
c
C   COMPUTE THE ELEMENTS IN RIGHT HAND SIDE OF THE OBJECT
C   FUNCTION OF KRIGING
C
      subroutine rlhscov(w,stder,aver,rho1,rho2)
c
C     INPUT DATA :
C       W(S)        = INVERSE DISTANCE WEIGHT
C       STDER(S,M) = STANDARD ERROR OF DATA AT STATION S
C                    AND IN A GIVEN MONTH M
C       AVER(S,M)  = AVERAGE OF DATA AT STATION S AND IN
C                    A GIVEN MONTH M
C       RHO1(S)    = CORRELATION BETWEEN STATION S AND
C                    ESTIMATION POINT
C       RHO2(S,S') = CORRELATION BETWEEN STATION S AND S'
C       IMON       = A GIVEN MONTH
C
C     OUTPUT DATA :
C       C(S,S')    = COVARIANCE BETWEEN STATION S AND S'
C       COI(s)     = COVARIANCE BETWEEN STATION S AND
C                    ESTIMATION POINT
C       AMUO       = ESTIMATED MEAN OF ESTIMATION POINT
C       TOTVAR     = ESTIMATED VARIANCE ESTIMATION POINT
C
      implicit double precision(a-h,o-z),integer(i-n)
      parameter(istat=21)
      common /comb1/c(istat,istat)
      common /comb2/coi(istat)
      common /comb4/amu0
      common /comc/imon
      common /comg/totvar
      dimension w(istat),stder(istat,12),aver(istat,12),
     1          rho1(istat),rho2(istat,istat)
c
      sum1=0.
      sum2=0.
      sum3=0.
      do 10 ista=1,istat
       sum1=sum1+w(ista)*aver(ista,imon)
       sum3=sum3+w(ista)*stder(ista,imon)**2
   10  sum2=sum2+w(ista)*stder(ista,imon)
      do 11 ista=1,istat
       do 12 jsta=1,istat
   12   c(ista,jsta)=stder(ista,imon)*stder(jsta,imon)*
```

```
     1                rho2(ista,jsta)
   11  coi(ista)=stder(ista,imon)*sum2*rho1(ista)
       amu0=sum1
       totvar=sum3
      return
      end
c
c *******************************************************
c *******************************************************
c
C   CALCULATE THE OPTIMUM WEIGHT BY KRIGING
C
      subroutine optweiht(xx)
c
C     INPUT DATA :
C       XX(S)      = INVERSE DISTANCE WEIGHT OF STATION S
C
C     OUPUT DATA :
C       XX(S)      = OPTIMUM WEIGHT OF STATION S
C
C     SUBROUTINE 'GRGSUB' IS DEVELOPED BY LASDEN (1979)
C     ALL VARIABLES USED FOR 'GRGSUB' CAN BE REFERRED
C     FROM GRG2 USER'S GUIDE
C
      implicit double precision(a-h,o-z), integer(I-n)
      parameter(istat=21)
      integer*4 ncore,nnvars,nfun,maxbas,maxhes
     ~common /comb1/c(istat,istat)
      common /comb2/coi(istat)
      common /comb3/amu(istat)
      common /comb4/amu0
      common /comc/imon
      common /comg/totvar
      logical maxim,inprnt,otprnt
      dimension defaul(19),z(15000),ramcon(3),ramvar(istat),
     1          nonbas(istat),redgr(istat),fcns(3),
     1          inbind(3),blvar(istat),buvar(istat),
     1          blcon(3),bucon(3),rmults(3),
     1          ttitle(19),xx(istat)
      data ncore /15000/
      data nnvars,nfun,maxbas,maxhes /21,3,3,21/
      data ramcon /8hOBJFUNCN,8hUNBIASED,8hSUMOFWGT/
      data ramvar /8hSTAT0101,8hSTAT0102,8hSTAT1031,
     1             8hSTAT1032,8hSTAT103A,8hSTAT0106,
     1             8hSTAT1082,8hSTAT108A,8hSTAT0109,
     1             8hSTAT1152,8hSTAT115A,8hSTAT0119,
     1             8hSTAT0120,8hSTAT0121,8hSTAT121A,
     1             8hSTAT0122,8hSTAT0123,8hSTAT0124,
     1             8hSTAT0125,8hSTAT0126,8hSTAT0127/
```

```
      data blvar/istat*0.0/,buvar/istat*1.0/
      data blcon /0.0,0.0,0.0/, bucon /1.0E3,0.0,0.0/
      data ttitle /19*4h    /
c
      inprnt=.true.
      otprnt=.true.
c
      do 10 i=1,19
       defaul(i)=1.0
   10 continue
c
      nnobj=1
      call GRGSUB(inprnt,otprnt,ncore,nnvars,nfun,maxbas,
     1            maxhes,nnobj,ttitle,blvar,buvar,blcon,
     1            bucon,defaul,fpnewt,fpinit,fpstop,fpspiv,
     1            pphlep,nnstop,iitllm,llmser,iipr,iipn4,
     1            iipn5,iipn6,iiper,iidump,iiquad,lderiv,
     1            mmodcg,ramcon,ramvar,xx,fcns,inbind,
     1            rmults,nonbas,redgr,nbind,nnonb,inform,z)
c
c      write(10,200) (xx(j),j=1,nnvars)
c      write(10,210) (fcns(i),i=1,nfun)
c      write(10,220) (nonbas(i),redgr(i),i=1,nnonb)
c      write(10,230) (inbind(i),rmults(i),i=1,nbind)
c      write(10,240) nnonb,nbind
c      write(10,250) inform
  200 format(///1h1,1x,'FINAL VALUES OF DECISION
     1       VARIABLES:'/5(5(5X,E12.5)/))
  210 format(/1x,'FINAL VALUES OF CONSTRAINTS AND
     1       OBJECTIVE:'/5(5X,E12.5))
  220 format(/1x,'REDUCED GRADIENT:'/5(5(5X,I3,2X,E12.5)/))
  230 format(/1x,'VALUES OF LAGRANGE MULTIPLIERS:'/
     1       5(5X,I3,2X,E12.5))
  240 format(/1x,'NUMBER OF STRUCTURAL NONBASIC VARIABLES:',
     1       5X,I5/1X,'NUMBER OF BINDING CONSTRAINTS:'5X,I5)
  250 format(/1x,'REASON FOR TERMINATION:',5X,I5)
      return
      end
c
c**********************************************************
c**********************************************************
c
C   COMPUTE THE OBJECT AND CONSTRAINTS FUNCTION
C
      subroutine GCOMP(g,x)
c
C     INPUT DATA :
C       G(3)       = TOTAL FUNCTIONS USED IN THE PROBLEM
C       X(S)       = UPDATED INPUT SOLUTION
```

```
C
C     OUTPUT DATA :
C       X(S)        - UPDATED OUTPUT SOLUTION
C
C     SUBROUTINE GCOMP IS DEVELOPED BY LASDEN (1979)
C
      implicit double precision(a-h,o-z), integer(i-n)
      parameter(istat=21)
      common /comb1/c(istat,istat)
      common /comb2/coi(istat)
      common /comb3/amu(istat)
      common /comb4/amu0
      common /comg/totvar
      dimension g(3),x(istat)
c
      f1=0.
      f2=0.
      f3=0.
      f4=0.
      do 10 i=1,istat
       do 11 j=1,istat
   11   f1=f1+x(i)*x(j)*c(i,j)
       f3=f3+x(i)*amu(i)
       f4=f4+x(i)
   10  f2=f2+x(i)*coi(i)
      g(1)=f1-2.*f2+totvar
      g(2)=f3-amu0
      g(3)=f4-1.
      return
      end
c
C*********************************************************
C*********************************************************
C
C   ESTIMATE THE MEANS AND THE VARIANCE USING GAUSSIAN
C   SMOOTHING
C
      SUBROUTINE GAUSSSMO (X,GC,GX)
C
C     INPUT DATA :
C       X(S,M)      - THE MEAN OR THE VARIANCE OF DATA AT
C                     STATION S AND IN MONTH M
C
C     OUTPUT DATA :
C       GC(M)       - OPTIMAL SMOOTHING FACTOR IN MONTH M
C       GX(S,M)     - ESTIMATED MEANS OR VARIANCE BY
C                     GAUSSIAN SMOOTHING
C
      implicit double precision(a-h,o-z),integer(i-n)
```

```
      PARAMETER (istat=21,MONTH=12)
C
      common /coma/DIST(iSTAT,iSTAT)
      common /comf/PX(iSTAT)
      dimension X(iSTAT,MONTH),GX(iSTAT,MONTH),
     &     XHAT(iSTAT),GC(MONTH)
C
      DO 10 MON = 1, MONTH
        DO 20 ISTA = 1, iSTAT
   20     PX(ISTA) = X(ISTA,MON)
C
        CALL FIBON (PC)
        CALL GAUSSHAT (PC,XHAT)
        GC(MON) = PC
c         print*,pc
C
       DO 30 ISTA = 1,iSTAT
   30   GX(ISTA,MON) = XHAT(ISTA)
   10 CONTINUE
      RETURN
      END
C
C ***********************************************************
C ***********************************************************
C
C    CALCULATE THE GAUSSIAN SMOOTHING WEIGHT
C
      SUBROUTINE GAUSSHAT (CC,RESULT)
C
C     INPUT DATA :
C       CC          = UPDATED OPTIMAL SMOOTHING FACTOR
C       DIST(S,S') = DISTANCE BETWEEN STATION S AND S'
C
C     OUTPUT DATA :
C       RESULT(S)  = UPDATED ESTIMATED DATA OF STATION S
C
      implicit double precision(a-h,o-z),integer(i-n)
      PARAMETER (istat=21)
      common /coma/DIST(iSTAT,iSTAT)
      common /comf/PX(iSTAT)
      dimension RESULT(iSTAT),F(iSTAT),W(iSTAT)
C
      DO 10 ISTA = 1, iSTAT
       DO 11 J = 1, iSTAT
        F(J) = DEXP((-DIST(ISTA,J)**2)/(2.*CC))
   11 CONTINUE
      FSUM = 0.
      DO 20 I = 1, iSTAT
   20 FSUM = FSUM + F(I)
```

```
      DO 30 I = 1, iSTAT
   30  W(I) = F(I) / FSUM
      WSUM = 0.
      DO 40 I = 1, iSTAT
   40  WSUM = WSUM + W(I)*PX(I)
      RESULT(ISTA) = WSUM
   10 CONTINUE
      RETURN
      END
C**********************************************************
C**********************************************************
C
C     SEARCH THE OPTIMAL SMOOTHING FACTOR BY FIBONACCI
C     SEARCH METHOD
C
      SUBROUTINE FIBON (GC)
C
C      INPUT DATA :
C        PX(S)      = UPDATED ESTIMATED DATA
C
C      OUTPUT DATA :
C        GC         = UPDATED SMOOTHING FACTOR
C
      implicit double precision(a-h,o-z),integer(i-n)
      common /comf/px(21)
      DIMENSION F(500)
      DATA A,B,ALPHA/0.,10.,0.01/
C
      BB=B
      AA=A
      UNCIV=B-A
      FNPR=UNCIV/ALPHA
      F(1)=1.0
      F(2)=1.0
      DO 5 I=3,500
      F(I)=F(I-1)+F(I-2)
      IF(FNPR-F(I))6,6,5
    5 CONTINUE
    6 IFN=I
      IFPRT=I
   15 P1=AA+(F(IFN-2)/F(IFN))*UNCIV
      P2=AA+(F(IFN-1)/F(IFN))*UNCIV
      CALL FUNC(P1,ERR1)
      CALL FUNC(P2,ERR2)
      IF(ERR1-ERR2)10,11,12
   12 AA=P1
      GOTO 13
   10 BB=P2
      GOTO 13
```

```
   11 BB=P2
      AA=P1
   13 UNCIV=BB-AA
      IFN=IFN-1
      IF(IFN-3)20,15,15
   20 GC=BB
       RETURN
      END
C
C**********************************************************
C**********************************************************
C
C    CALCULATE THE ERROR CONCERNING TO FIND THE
C    OPTIMAL SMOOTHING FACTOR
C
      SUBROUTINE FUNC (CC,ERR)
C
C     INPUT DATA :
C       CC          = UPDATED SMOOTHING FACTOR
C       PX(S)       = OBSERVED DATA AT STATION S
C       PXH(S)      = ESTIMATED DATA AT STATION S
C
C     OUTPUT DATA :
C       ERR         = THE ERROR CONCERNING THE ESTIMATES
C
      implicit double precision(a-h,o-z),integer(i-n)
      PARAMETER (istat=21)
      common /comf/PX(iSTAT)
      dimension PXH(iSTAT)
C
      CALL GAUSSHAT (CC,PXH)
      ESUM = 0.
      DO 20 I = 1, iSTAT
       ESUM = ESUM + (PX(I)-PXH(I))**2
   20 CONTINUE
      ERR = ESUM
      RETURN
      END
c
c **********************************************************
C **********************************************************
c
C   CALCULATE THE MEAN AND THE ERROR OF THE ESTIMATED
C   DATA WITH OPTIMUM WEIGHT
C
      SUBROUTINE CONTAVER (X,ITER,AVERMON,STDERR,VAR)
C
C     INPUT DATA :
C       X(K,T)      = THE ESTIMATED DATA AT ESTIMATION
```

```
C                      POINT K AND IN MONTH T
C        ITER        = REPRESENTS ITER-TH ESTIMATION POINT
C
C      OUTPUT DATA :
C        AVERMON(K) = MEAN OF ESTIMATED DATA AT ESTIMATION
C                      POINT K
C        STDERR(K)  = STANDARD ERROR OF DATA AT ESTIMATION
C                      POINT K
C        VAR(K)     = VARIANCE OF DATA AT ESTIMATION POINT K
C
      implicit double precision(a-h,o-z), integer(i-n)
      common /comc/imon
      PARAMETER(itime=169,igrid=32)
C
      dimension X(igrid,itime),AVERMON(igrid),
     &     STDERR(igrid),VAR(igrid)
C
        SSUM = 0.
        ACOUNT = 0.
        DO 40 ITIM = IMON, ITIME, 12
         IF (X(ITER,ITIM).NE.9999.) THEN
          SSUM = SSUM + X(ITER,ITIM)
          ACOUNT = ACOUNT + 1.
         ELSE
         ENDIF
   40   CONTINUE
        AVERMON(ITER) = SSUM / ACOUNT
        SUM = 0.
        DO 50 ITIM = IMON, ITIME, 12
         IF (X(ITER,ITIM).NE.9999.) THEN
          SUM = SUM + (X(ITER,ITIM)-AVERMON(ITER))**2
         ENDIF
   50   CONTINUE
        VAR(ITER)=SUM/(ACOUNT-1.)
        STDERR(ITER) = DSQRT(VAR(ITER))
      RETURN
      END
c
c *********************************************************
C *********************************************************
c
C   CALCULATE THE GAUSSIAN SMOOTHING WEIGHT BY THE OPTIMAL
C   SMOOTHING FACTOR
C
      subroutine GAWEIGHT (gc,gw)
c
C     INPUT DATA :
C       GC(M)      = OPTIMAL SMOOTHING FACTOR IN MONTH M
C
```

```
C       OUTPUT DATA :
C         GW(S)       = GAUSSIAN SMOOTHING WEIGHT AT STATION S
C
        implicit double precision(a-h,o-z),integer(i-n)
        parameter(istat=21)
        common /comc/imon
        common /comd/distn(istat)
        dimension gw(istat),f(istat),gc(12)
c
        cc=gc(imon)
        do 10 ista=1,istat
         f(ista)=dexp((-distn(ista)**2)/(2.*cc))
   10   print*,f(ista),distn(ista),cc
        fsum=0.
        do 20 ista=1,istat
   20    fsum=fsum+f(ista)
        do 30 ista=1,istat
   30    gw(ista)=f(ista)/fsum
        return
        end
c
c ********************************************************
C ********************************************************
c
C       ESTIMATE THE MEANS AND STANDARD ERROR USING THE GIVEN
C       WEIGHT
C
        subroutine ESTIMATE
     &           (w,iter,avermon,stderr,avhat,stdhat)
c
C       INPUT DATA :
C         W(S)          = THE WEIGHT AT STATION S
C         AVERMON(S,M)  = THE OBSERVED MEAN AT STATION S AND
C                         IN MONTH M
C         STDERR(s,M)   THE OBSERVED STANDARD ERROR
C
C       OUTPUT DATA :
C         AVHAT(K)      = ESTIMATED MEAN AT ESTIMATION POINT K
C         STDHAT(K)     = ESTIMATED STANDARD ERROR AT
C                         ESTIMATION POINT K
C
        implicit double precision(a-h,o-z),integer(i-n)
        parameter(istat=21,igrid=32,month=12)
        common /comc/imon
        dimension w(istat),avermon(istat,month),
     1          stderr(istat,month),
     1          avhat(igrid),stdhat(igrid)
c
        avsum=0.
```

```
      stdsum=0.
      do 10 ista=1,21
       avsum=avsum+w(ista)*avermon(ista,imon)
   10 stdsum=stdsum+w(ista)*stderr(ista,imon)
c
      avhat(iter)=avsum
      stdhat(iter)=stdsum
      return
      end
```