Research Technical Completion Report

MAPPING PRECIPITATION IN IDAHO: A STUDY OF DATA ESTIMATION AND INTERPOLATION PROCEDURES

by

Francis M. Winters, Jr.
Myron Molnau
Department of Agricultural Engineering

Kang-tsung Chang
Department of Geography

Idaho Water Resources Research Institute
University of Idaho
Moscow, Idaho 83843

September, 1989

The research on which this report is based was financed in part by the United States Department of the Interior as authorized by the Water Research and Development Act of 1978 (P.L. 95-467).

Research Technical Completion Report

14-08-0001-G1559-32

MAPPING PRECIPITATION IN IDAHO: A STUDY OF DATA ESTIMATION AND INTERPOLATION PROCEDURES

by

Francis M. Winters, Jr.
Myron Molnau
Department of Agricultural Engineering

Kang-tsung Chang
Department of Geography

Submitted to

U.S. Geological Survey
U.S. Department of Interior
Washington, D.C. 20242

Idaho Water Resources Research Institute
University of Idaho
Moscow, Idaho 83843

September, 1989

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

As the competition for Idaho's available water has increased, increased attention has been paid to the amount and distribution of precipitation in Idaho. To adequately manage this valuable resource, a good understanding of the precipitation pattern is essential. In particular, a new map of mean annual precipitation is needed. The current map was based primarily on data from low-lying weather stations. Since its completion, new data have shown this map to be inaccurate, especially at high elevations where new data from the SNOTEL network are now available.

A new series of mean annual precipitation maps at a scale of 1:250,000 is now being drawn. A computer contouring procedure is used to create the preliminary maps with the final maps being manually drawn.

The database used is the National Weather Service Cooperative Network and the Soil Conservation Service SNOTEL stations. All data records were gathered and standardized to the 1961-1985 base period. Next, multiple regression equations were developed to estimate point precipitation values where no gages exist. Third, point interpolation procedures were investigated as to their appropriateness for use on the data set.

Either the normal ratio method or multiple regression was used to derive missing values. Acceptable estimates were made for all missing data.

Multiple regression equations for point precipitation

data were derived to estimate precipitation in between locations of snow precipitation. The final equations were of the form log(MAP) = f(elevation, location, air mass lifting index). An equation was derived for each of seven regions. Acceptable estimation equations were derived for all but two of these regions.

For the automated production of preliminary isohyets, two point interpolation procedures were examined. The advantages and limitations of the inverse distance method and the punctual Kriging procedure were investigated.

I. INTRODUCTION

GENERAL

In the past several years, increased attention has been paid to the amounts and distribution of precipitation in Idaho, as the competition for Idaho's available water has increased. Agriculture, timber, hydro-electric power, recreation, industries, municipalities, and others have all staked claims to Idaho's water. It is clear that to adequately manage this valuable resource, a good understanding of the precipitation pattern is essential. In particular, an accurate map of mean annual precipitation is needed.

The current map of Idaho's mean annual precipitation was compiled by the National Weather Service (NWS) in 1965. It was based primarily on data from low-lying NWS stations which are near population clusters, because the gauges are manually read and recorded daily. Precipitation values in high elevation areas were extrapolated from these valley (low precipitation) stations. In the years since its completion, changing precipitation patterns and new data have shown that a new map is needed, especially at higher elevations.

A project has been underway for two years under the direction of the State Climatologist, in cooperation with the Soil Conservation Service Snow Surveys, the Bureau of Land Management Idaho State Office, The Forest Service Region Four, and the Bureau of Reclamation Region One, with the primary objective of creating a new set of mean annual precipitation maps for Idaho. These maps correspond to the U.S.G.S. one by two degree quadrangles at a scale of 1:250,000. A number of challenging obstacles have to be overcome in order to produce these maps. Some of these obstacles are addressed in this report including:

1) Data acquisition, adjustment, and standardization.
2) The estimation of point precipitation in areas where no gages have existed.
3) The interpolation of both measured and estimated precipitation values using computer contouring procedures.

The final map production phase of the project will not be included in this report because that phase involves subjective, hand adjustments of isohyets using the computer generated precipitation maps, the topographic base maps, records from manual snow course sites, the old precipitation maps, and the original data set. The knowledge used for the adjustment of isohyets is not easily reduced to computer algorithms.

RATIONALE

The major obstacle to producing these precipitation maps is presented by the inadequate density of precipitation gages to depict the pattern of precipitation with the desired detail. The gage network has improved dramatically in the late 1980's with the introduction of the Soil Conservation Service (SCS) SNOTEL data network. The SNOTEL network was initiated in the early 1980s with the primary goal of quantifying precipitation in the peak water producing zones of the state. These zones are located primarily in remote high elevation areas, which were impossible to monitor until recent technological advancements were made in automated gages. Although the SNOTEL network plays an invaluable role in the quantifying of precipitation in Idaho, the network of gages formed by the combined NWS and SNOTEL data sets was not dense enough to depict the precipitation pattern at the desired scale and detail. To illustrate this, isohyets for the Sandpoint one by two degree quadrangle were generated using the Surfer[1] package with three different sets of data points. Figure 1.1 shows the precipitation pattern using only the existing precipitation stations. Eleven precipitation stations (seven NWS stations, and four SNOTEL stations) are located in the area covered by this quadrangle. Figure 1.2 is the same map with the four SNOTEL stations removed. Clearly, the high elevation data points make a notable difference in the precipitation pattern, but the pattern is not believed to be depicted reasonably, even with the high elevation data points. This map would lead us to believe that there are three major peak precipitation points on this quadrangle.

---

[1]These names are included for the benefit of the reader and do not imply endorsement of performance for the product.

Experience shows that this is not the case. If we had a data point on the top of every major peak and at the bottom of every major valley, the pattern would like that shown in Figure 1.3. This map was made by using the original eleven points and forty-two estimated point values from multiple regression equations. This combined set of fifty-three points was run through a Kriging interpolation procedure. It is clear that in order to produce reasonable computer interpolated maps of precipitation, a method must be derived to estimate point precipitation values in all areas of the state where no gages exists, and appropriate interpolation procedures must be selected.

The three main objectives of this report are:

1) To perform data acquisition and standardization.
2) To formulate equations for the estimation of point mean annual precipitation values where no gages exist.
3) To investigate interpolation procedures for the automated production of isohyets.

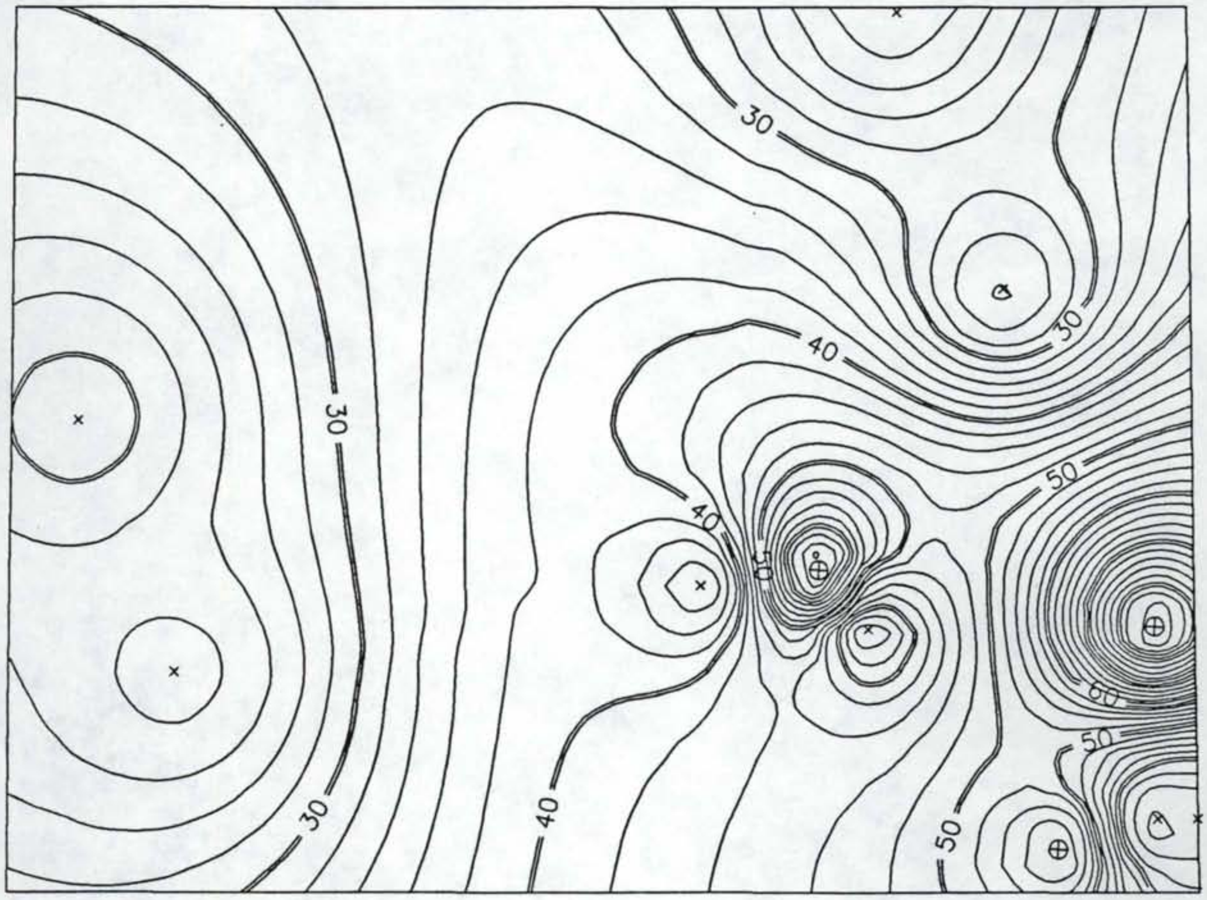


Figure 1.1 Sandpoint One by Two Degree Quadrangle Precipitation Pattern Derived from Three SNOTEL and Eight NWS Data Points

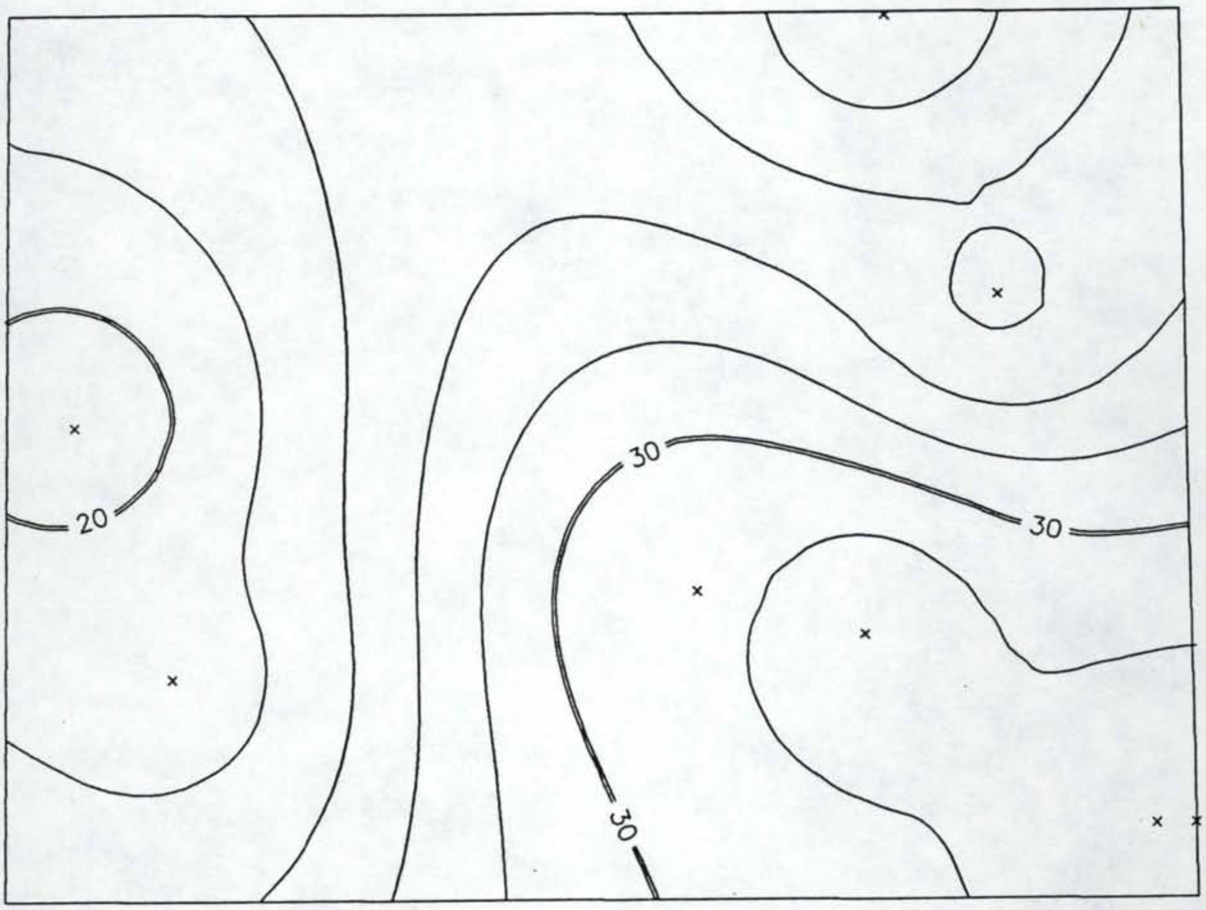


Figure 1.2 Sandpoint One by Two Degree Quadrangle Precipitation Pattern Derived from Eight NWS Data Points

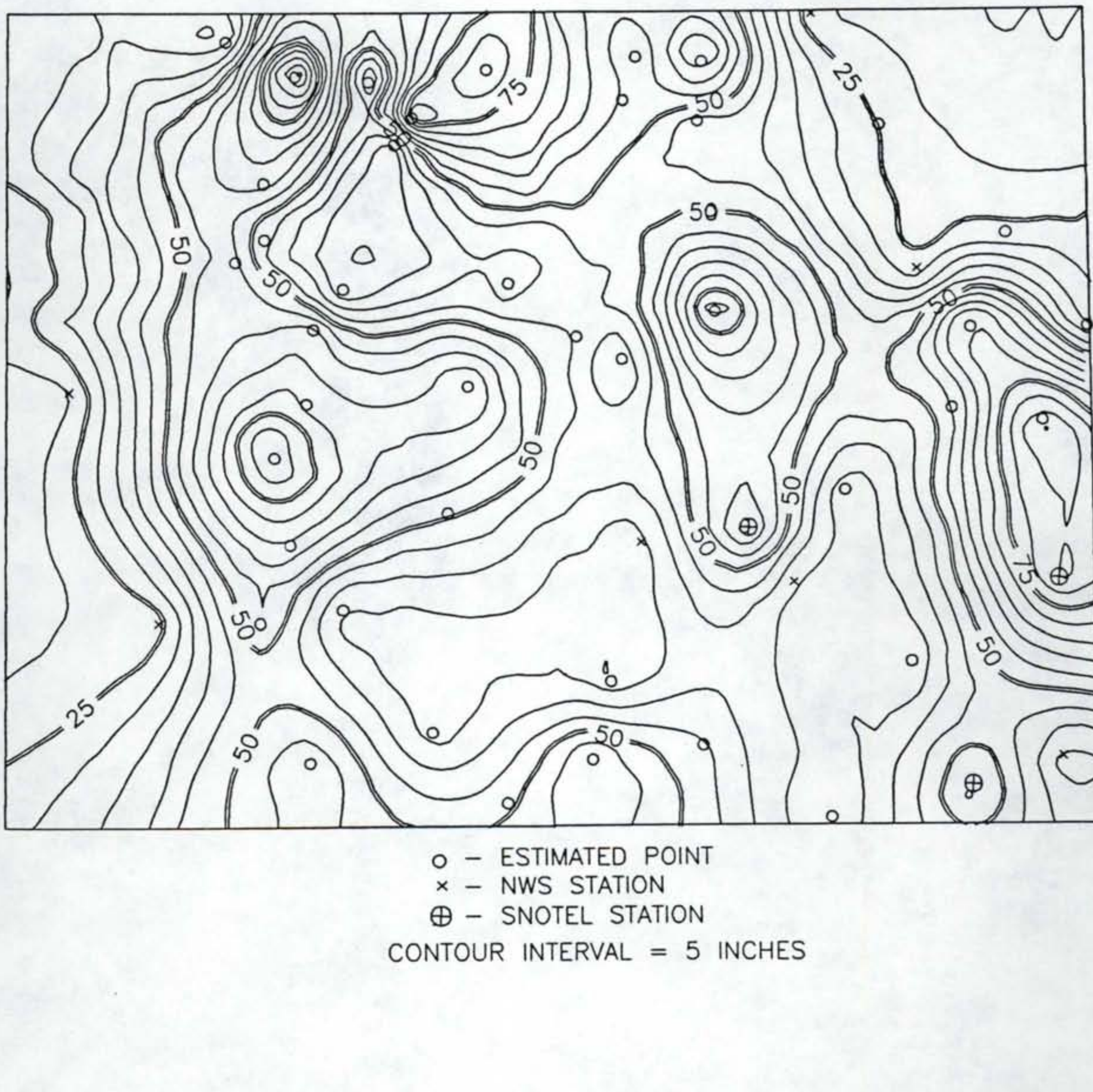


Figure 1.3 Sandpoint One by Two Degree Quadrangle Precipitation Pattern Derived from Three SNOTEL, Eight NWS, and Forty-three Estimated Data Points

ORGANIZATION OF THE STUDY

This study is divided into six sections, each of which will be explained in an individual chapter. Chapter I introduces the topic, and gives some rationale for the study. A literature review will be included in Chapter II. The third chapter deals with the original data acquisition, missing record estimation, and short record extension. Examples of missing data estimation procedures and lists of all of the precipitation stations used are included. A series of regression models for estimation of point precipitation values at sites where no gages exist is described in Chapter IV. The steps used to generate these regression equations include:

1) Regression analysis of mean annual precipitation and the log of mean annual precipitation based on physiographic variables for the entire study area.
2) The same analysis on individual regions of the study area.
3) Discriminate analysis to verify the validity of the regions.
4) Regression runs on mean seasonal and monthly precipitation totals.
5) Final regression runs to formulate the most accurate and easy to use estimation equations.

The fifth chapter will deal with the interpolation procedures which are necessary for automated contouring. The inverse distance method and the punctual Kriging procedure are explained in this chapter. A summary and conclusions are included in the final chapter.

EXPECTED RESULTS

After studying the findings of others who have undertaken similar projects in other states, a statement can be made as to the expected results of this study. It is expected that the data record adjustment phase of the project will work well. Some of the initial stations may have too many missing records to allow for reasonable estimates. These stations will not be used in the initial part of this project but will be used in the production of the final maps. Missing records from the remaining stations will be estimated using different methods depending upon the number of missing records. This process is expected to work well because of the strong cross-correlation between stations and individual station auto-correlations (Foufoula-Georgiou, 1983). It is also expected that the SNOTEL records will play an important role in quantifying precipitation in high elevation areas (Doesken and Schaefer, 1987).

While formulating the regression equations for estimation of point precipitation values, it is conceivable that the study area should be broken down into regions. The precipitation regimes in various regions of Idaho are dramatically different. One multiple regression equation cannot be expected to accurately quantify the relationship between physiographic parameters and precipitation throughout the entire state.

Multicollinearity will vary from region to region. The cross-correlation between the physiographic variables will be controlled by the overall topographic trends in the regions. This cannot be avoided. For example, the relationship between elevation and longitude is strong in Idaho, because the eastern boundary of the state is the crest of the Bitterroot Mountains.

Favorable results are expected in the northern most portion of the state (the Panhandle). The terrain in this area is by no means gentle, but the precipitation pattern is expected to follow the organized system of ridges and valleys with great regularity. Problems are expected in and around Idaho County. The terrain in this area is complex and somewhat random, and the extremes in precipitation values are great.

When the seasonal and monthly precipitation values are used as dependent variables, the error terms are expected to be smaller than those of the annual models. This should occur as the individual precipitation forming processes at work are quantified separately during different times of the year. It is expected that the wintertime precipitation processes will be more easily quantified than those of the summer, because of the organized nature of the frontal precipitation processes at work in the winter (Arnold, 1989).

II. LITERATURE REVIEW

MEAN PRECIPITATION MAPPING PROGRAMS

The need for long-term mean precipitation maps has been recognized in the United States and throughout the world. In this section, a few projects in which precipitation maps were created or updated are described. Peck & Brown (1962) described an innovative way of producing isohyetal maps for mountainous areas. The anomalies to the general relationship between precipitation and elevation were found to be based on topographic barriers. These anomalies were plotted on the existing precipitation maps and the isohyets were adjusted accordingly. Mean annual precipitation maps for the State of Utah were the major end product of the project.

The old mean annual precipitation map of Colorado was updated only in areas where changes were needed. The following steps were involved: 1) assembly of all available data, 2) calculation of mean monthly, seasonal, and annual values for each station, 3) adjustment of short-term records to the 1951-1980 base period, 4) plotting data points on the old precipitation maps, and 5) adjustment of isohyets to fit the new data. Doesken and others (1984) stated that significant improvements in accuracy were not because of better mapping procedures, but rather due to better, more representative data.

Mean annual precipitation maps were produced for the entire state of Montana for the 1941-1970 base period at a scale of 1:250,000. The Montana project introduced data improvements by estimating mean annual precipitation from snow course data, as well as by utilizing the SCS mountain storage gage network. All short-term records were adjusted to the same thirty-year base period through correlation with surrounding stations. Isohyets were drawn using the known data points with terrain and soil information (U.S.D.A., 1977; Farnes, 1978).

The Vitim River Basin of the Soviet Union was the site for large-scale mapping of long-term mean annual precipitation using graphs of precipitation versus elevation. Isohyets were hand-drawn taking into account orographic anomalies and the direction of moisture-bearing winds. These maps were checked and calibrated using runoff data from the river basin gage sites (Vuglinski, 1972).

DATA RECORD ADJUSTMENT

One of the first steps in any study of long-term mean annual precipitation patterns usually involves the adjustment of data records to a common base period (Peck, 1972). This involves both extending short-term records and filling in missing records. Researchers have treated these tasks in a variety of ways. Foufoula-Georgiou (1983) performed a comparison study of six methods of estimating missing monthly rainfall data records. The normal ratio method generated the most accurate estimates. With as much as twenty percent of the data missing, unbiased estimates were generated with most of the estimation procedures.

Solow and Gorelick (1986) described a process for filling in missing streamflow records in west central Virginia through a co-Kriging procedure. This process took into account spatial dependence as well as cross-correlations and auto-correlations. The resulting streamflow estimates were thought to be more accurate than those derived from regression analysis. They suggested that a similar process could be applied to missing precipitation records to take advantage of the relationship between precipitation and elevation.

In order to produce accurate hand-drawn maps of fifty-year mean and median precipitation for Hawaii, short-term records had to be extended to a fifty-year base period (Giambelluca and Nullet, 1985). If stations with short-term records were discarded, the gage density would not have been sufficient to determine the precipitation pattern. Record extension was accomplished by a process called ridge regression, which correlated precipitation values for the stations with short-term records to those nearby stations with complete fifty-year records. Ridge regression is preferred to standard regression when multicollinearity is a problem between stations used as independent (predictor) variables. This process employs biased estimation to minimize the effects of multicollinearity and optimize the explained variance. The correlation between independent variables is artificially reduced in order to stabilize the estimation equation. A non-negative bias coefficient (k) is used as a scalar. K is added to each of the diagonal elements of the X'X matrix. This results in a more stable X'X matrix from which the parameter estimate can be calculated.

Farnes (1978) introduced an interesting way of estimating long-term mean precipitation values from mountain snow course records measured in both open and forested areas of Montana. A photocanopyometer was used to photographically measure the amount of sheltering over each forest covered snow course site. Adjustments were made to the April First snow water equivalent records based on the amount of canopy cover measured. Snow courses in open areas needed adjustment. These April First snow water equivalent values were then correlated with annual precipitation records from nearby stations to yield mean annual precipitation values at each snow course site. The correlation worked well. Isohyetal maps were produced from these data points and storage gage data. These maps were checked with streamflow runoff records.

Doesken and Schaefer (1987) compared SNOTEL gages with standard National Weather Service gages in Colorado. They determined that improvements in the SNOTEL gages make the catch accuracy superior to NWS gages. They also warn that the two networks form distinct data populations and direct comparisons may be difficult and inappropriate. SNOTEL data are undoubtably invaluable to high elevation hydrologic studies.

PRECIPITATION DISTRIBUTION ANALYSIS

There are numerous studies explaining spatial variation in precipitation amounts and frequencies based on physiographic features. Some of these studies point towards estimating precipitation values where no gages exist, while others are intended to simply offer some insight into the spatial characteristics of the precipitation forming process. In order to produce a series of precipitation frequency maps for the entire western United States, Miller (1972) estimated regression equations relating precipitation frequency to: 1) slope, 2) normal annual precipitation, 3) distance to moisture, 4) elevation, 5) barriers to airflow, 6) location, and 7) roughness. The most important variable was slope. This was measured as vertical change along the mean moisture inflow bearing. Distance to the moisture source and elevation were also important. Eighty-three percent of the variance in precipitation frequency was explained by his final model.

Dingman (1981) explained precipitation in Vermont and New Hampshire in terms of elevation and local physiographic parameters which make up what he called a "precipitation delivery factor". The following is Dingman's equation explaining precipitation at point (i):

$$p_i = a_i + 0.746\ y_i$$

where p_i is precipitation at point (i); a_i is the precipitation delivery factor; and y_i is elevation. The precipitation delivery factor was based on the distance to the water source (Atlantic Ocean), and the rain shadow effect. The change in precipitation with elevation was 0.746 mm per meter after controlling for the other variables.

Ryden (1972) described a project in which the vertical distribution of precipitation was examined at very high elevations (above tree line) in Sweden. Elevation differences of 1500 meters were examined in terms of their effect on precipitation. Ryden found that precipitation increased with elevation in a linear fashion in general, but during high precipitation seasons, the rate of increase with elevation was greater.

The USDA/ARS Northwest Watershed Research Center operates a dense network of precipitation gages in the Reynolds Creek watershed of southwest Idaho. Several microscale precipitation studies have been conducted on this watershed to take advantage of the unusually dense data network. Hanson (1982) studied the spatial distribution of precipitation over the Reynolds Creek watershed through regression analysis. He found a linear relationship between precipitation and elevation. This relationship worked the best when stations were grouped into upwind and downwind categories based on the mean moisture inflow bearing of the watershed. The equation for the downwind group explained ninety-two percent of the variance in precipitation, while the upwind equation explained eighty-six percent. These groupings were not appropriate in the summer due to the sporadic nature of summertime precipitation producing processes (thunderstorms).

Molnau and others, (1980) conducted a two part study on the same network of gages to answer the following questions: 1) Which gauge site characteristics can be used to explain the variation in precipitation?, and 2) What changes in accuracy of precipitation estimation could be expected from various gauge network densities covering the same area? Regression models were run with mean annual precipitation as the dependent variable, and independent variables of elevation, slope, aspect, percent vegetation cover, and soil class. Of these, elevation, vegetation, and slope were significant. The final model used only elevation and vegetation cover, with a r-square of 0.637. The gage network was then stratified into elevation classes to test the effects of various gage densities on the error terms. Samples of thirty, twenty, ten, and five gages were tested. The confidence band widened dramatically when less than twenty gages were used.

An analysis of Colorado's mean December through February precipitation was performed by Spreen (1947) for the 1920-1930 base period. The sample contained twenty-six stations, all with complete data records. A graphic correlation technique was used under the assumption that the independent variables (elevation, slope, exposure, and orientation) had a combined influence on precipitation. Graphs of precipitation versus elevation were produced and adjusted for slope, orientation, and exposure. Equations for estimating precipitation were derived from these graphs. When the estimated values were compared with the actual means, eighty-eight percent of the variance was explained by the four parameters. Thirty percent was explained by elevation alone.

As part of a study to develop techniques for evaluating changes in accuracy of water supply forecasts in Colorado caused by changes in gauge densities and location, Peck and Schaake (1987) needed estimates of winter precipitation values on an evenly spaced grid. The study area consisted of the portion of the Colorado River Basin above Lake Powell. Regression equations were derived to estimate these values. Twenty-five percent of the variance in precipitation was explained when rise, exposure, direction of rise, and azimuth were included. In order to explain more of the variance in precipitation, an atmospheric water balance was used to assign three precipitation indices to each station based on the relative amounts of precipitation expected from storms approaching from three likely directions (west-northwest, west, and west-southwest). After dividing the study area into three regions, they found that the precipitation indices varied smoothly with latitude. The final regression equations used the three precipitation indices and latitude. Approximately two thirds of the variance in precipitation was explained.

Peck (1972) examined a method of predicting the synoptic precipitation pattern from meteorological parameters in Utah without the use of typical storm types. Twelve hour periods were studied. Canonical correlation proved more useful than regression analysis due to the strong multicollinearity among variables. He found that two sets of canonical variables were significant.

Differences between precipitation values measured at seventy-two pairs of stations in the North Central Great Basin were examined by Houghton (1979). Each pair of stations contained one high elevation and one low elevation station. The ratio of precipitation between the two stations was computed for each pair. Orographic influences were expressed by these ratios. Sixty-five percent of the variation in the ratios could be explained using numerous topographic variables in a regression equation. It was found that orographic differences in precipitation and the rain shadow effect were strongest when the airflow was perpendicular to the ridge, and the ridge is of substantial height and length.

Numerous projects have been undertaken in western Canada in an effort to explain the spatial distribution of precipitation (Storr and Ferguson, 1972). In the Marmot Creek watershed of the Kananaskis Valley in Alberta, Canada, a study was conducted using stepwise regression to express rainfall as a function of elevation and slope. The model worked the best when the stations were grouped by aspect. An r value of 0.945

was calculated for the gauges on the south facing slopes, and 0.941 for gauges on the east facing slopes. Also, the increase of snowfall with elevation was greater than that of rainfall. Storr and Ferguson also reported on a study of monthly precipitation in the Okanagan Basin of British Columbia. The first step in this analysis was to develop models to estimate monthly precipitation as a function of elevation, slope, aspect, distance to barriers, barrier height, and shield effect. The shield effect variable was the sum of the barriers along the 236.6 degree (mean 850 mb. flow) bearing. They found that elevation and barrier height were the most useful in explaining the variation of precipitation. They stated, that the slope and aspect variables might have been more effective if they had been combined with the wind direction during precipitation events, into a single "orographic vertical motion parameter."

Bleasdal and Chan (1972) conducted a large scale study on the variation in precipitation throughout the United Kingdom using long-term records from over six thousand rain gauges. A regression equation was derived expressing mean annual rainfall as a function of elevation. Residual values were plotted on contour maps. The line of zero departure (LZD) followed roughly the water divide running north-south along the highest points in the country. Positive departures were found west of the LZD, while negative to the east.

INTERPOLATION PROCEDURES

In order to produce precipitation maps using a contouring program, various interpolation procedures have been utilized. These interpolation procedures are needed to create a continuous, evenly spaced grid of precipitation values through which the contouring algorithm can direct isolines. The quality of the map produced from this grid depends heavily on the appropriateness of the interpolation process used.

Tabios and Salas (1985) reported on a comparison of six interpolation techniques used to estimate point precipitation at ungauged locations. The procedures investigated were the Thiessen polygon method, the classical polynomial by least squares or Lagrange approach, the inverse distance technique, multiquadric interpolation, optimal interpolation, and a Kriging technique. Estimates were checked at five known precipitation gauges which were excluded from the interpolation. Tabiso and Salas determined that the Kriging and optimal interpolation techniques were superior to the others, while the multiquadric method was not far behind in terms of accuracy.

In the United Kingdom, a system is in operation to automatically produce isohyetal maps at a variety of scales. A five kilometer grid of mean annual precipitation values was derived using a cubic spline on a dense network of gauges. The system is the product of the Comprehensive Areal Rainfall Program (Shearman and Salter, 1975).

Hutchinson and Bischof (1983) introduced a new method of estimating the spatial distribution of rainfall in New South Wales. A Laplacian smoothing spline procedure was decided upon. This procedure was proven to be more accurate than the least-squares polynomial method and the weighted interpolation method. Rainfall values were found to vary smoothly in response to changes in latitude, longitude, and elevation.

In central Switzerland, a project was undertaken to produce and evaluate precipitation maps using estimates based on the anomaly method. This two step process involves first, evaluating precipitation as a function of elevation, then, applying a Kriging procedure to the anomalies residual values. The residuals form a regionalized variable, because it is neither totally spatial dependent nor totally random in nature. The anomaly pattern was associated with physiographic features which alter the normal precipitation/elevation relationship (de Montmollin, and others, 1980).

Dingman and others (1988) described two processes of interpolation to produce isohyetal maps in Vermont and New Hampshire. The study focused on the inadequacies in gauge density to sample the area of interest at the desired scale. Kriging procedures were applied in two ways on a ten kilometer grid system. The grid was required for contour generation by the computer mapping program SURF2. First, the actual precipitation values were Kriged without adjustment. This produced the lowest error at seven verification stations. The second procedure involved the derivation of "precipitation delivery factors" which were estimated through regression analysis of precipitation based on elevation. The "precipitation delivery factors" were precipitation values with the effect of elevation removed. This variable was Kriged over the entire two state area. The result was a much smoother pattern in error terms than produced by the earlier method. Although the estimates at individual verification stations were not quite as accurate, the "precipitation delivery factors" method was preferred. The overall confidence level for the precipitation estimates was approximately twenty percent.

III. INITIAL DATA

Data for this study were compiled from the National Weather Service (NWS) Cooperative Network and the Soil Conservation Service (SCS) SNOTEL network. All available records for the state of Idaho and a fifty mile perimeter were assembled for the years 1961-1985. Figure 3.1 shows the study area with all of the NWS and SNOTEL sites.

NWS DATA

All NWS records were organized by month and year. Approximately one percent of the monthly records were missing from the original data base. These were looked up in published NWS records and on microfiche. If found, the values were verified and entered.

Stations with more than five missing years were discarded, assuring that at least eighty percent of the data for each station consisted of actual recorded values (Foufoula-Georgiou, 1985). The remaining missing monthly values were estimated in one of two ways depending upon the number of missing months in each year. In cases of three or fewer missing months in a single year, the normal ratio method was used. When more than three values were missing, the percentage of normal precipitation for that year could not be

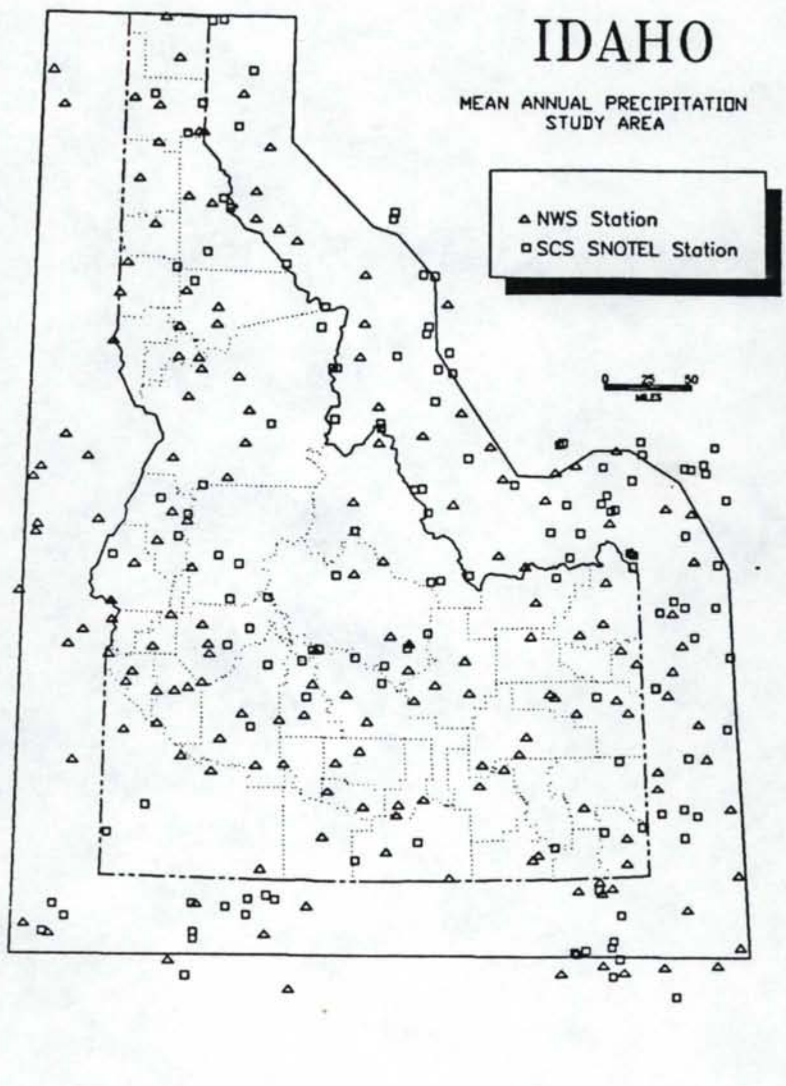


Figure 3.1 Mean Annual Precipitation Study Area and Precipitation Stations

Table 3.1 Normal Ratio Method

Station #163760 Anderson Dam
Missing February and December of 1982

A =	1961-1985 10 month mean precipitation = 16.07 inches for #163760 (excluding Feb. and Dec.)
B =	1982 10 month precipitation total = 18.33 inches for #163760 (excluding Feb. and Dec.)
C =	B/A = ratio of normal = 1.14
D =	1961-1985 mean precipitation value for Feb. = 2.15 inches
E =	1961-1985 mean precipitation value for Dec. = 3.63 inches
F =	D * C = 1982 Feb. estimate = 2.45 inches
G =	E * C = 1982 Dec. estimate = 4.14 inches

accurately estimated. Therefore, the estimate was derived for each month using the correlation method (Peck and Schaake, 1987).

The Normal Ratio Method

For estimating one, two, or three missing month values, the Normal Ratio Method first computed the ratio of the precipitation totals for the nine to eleven months containing data to the 1961-1985 means for the same months. This ratio was then multiplied by the 1961-1985 mean monthly value to yield an estimate for each missing month. In other words, an adjustment for the year's percentage of normal was applied to each mean monthly value. The 1982 record for the Anderson Dam Station exemplifies this process (Table 3.1)

The Correlation Method

In the event of more than three missing monthly values in one year, regression equations were derived using the entire 1961-1985 record from nearby stations. The precipitation for the individual month in question was estimated as a function of precipitation for that month at several surrounding stations. In some cases, as many as eight stations were considered, because of their close proximity to the station in question. For example, station 165038, Kuna 2 NNE, is surrounded by eight stations within thirty miles. In other cases, fewer stations were available. Station 161663, Challis, has only two stations for comparison within fifty miles. Most of the stations had four comparison stations. Only NWS stations were used. The process by which SNOTEL records were adjusted (explained in the next section) render them unsuitable for this comparison with NWS stations.

A stepwise regression procedure was used to estimate an equation based on the comparison stations. This was done for each missing monthly value one station at a time. A "no intercept" option was used so only the parameter estimates and the independent variables would be needed to formulate the final estimates. Independent variables which were not significant at the 0.15 level were automatically dropped by the stepwise procedure. Because of strong cross-correlation between comparison stations, only one or two stations were included by the stepwise procedure in most cases. Most of the equations explained between eighty-five and ninety-five percent of the variance in the monthly means of the station in question. Figure 3.1 and Table 3.2 illustrate this procedure.

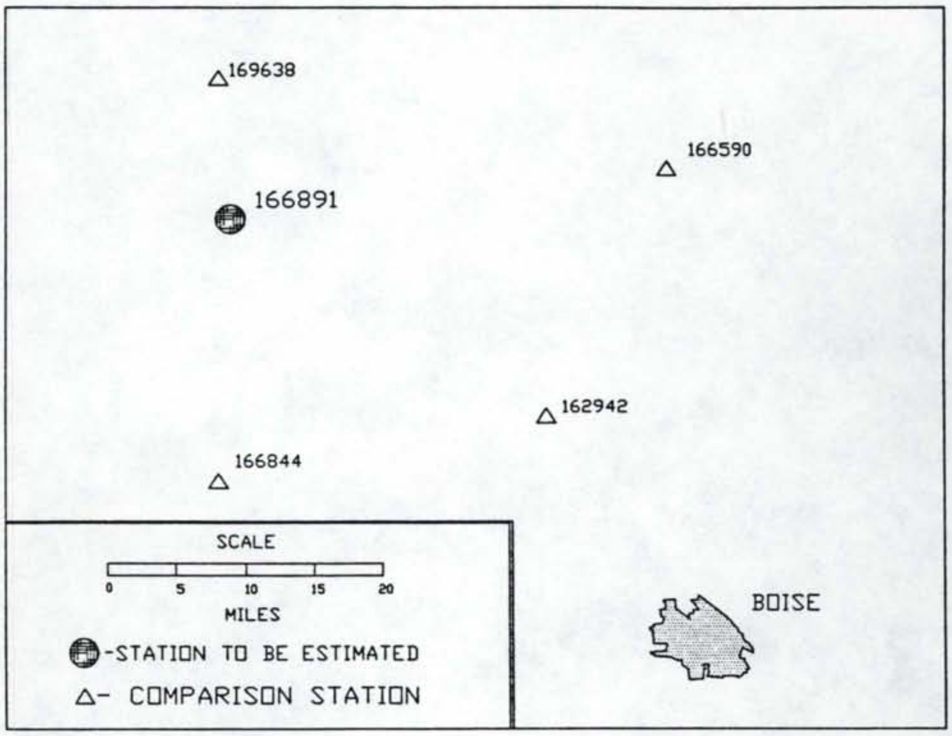


Figure 3.2 Comparison Stations for Station #166891

Table 3.2 Results of Stepwise Regression When More Than Three Months Missing in a Single Year

Station: #1606891 Payette missing Jan-Dec of 1980
Comparison Stations:
162942- Emmett 2E
166590- Ola 4S
166844- Parma Exp Station
169638- Weiser 2SE

Month	R-square	Independent Variables	Prob > $\lvert T \rvert$	Parameter Estimate
Jan	0.96	102942	0.0028	0.58193563
		106844	0.1489	0.27168522
Feb	0.96	102942	0.0020	0.44872011
		106590	0.0105	0.26378949
Mar	0.91	102942	0.0001	0.85377225
Apr	0.96	106844	0.0001	0.65787434
May	0.95	102942	0.0165	0.25746774
		106844	0.1504	0.25336399
		109638	0.0299	0.32661945
Jun	0.87	106590	0.0301	0.23835589
		106844	0.0011	0.46472220
Jul	0.88	102942	0.0098	0.34357031
		106590	0.0001	0.32985120
Aug	0.96	106844	0.0001	0.38223354
		109638	0.0001	0.77343045
Sep	0.91	106844	0.0001	0.38717049
		109638	0.0027	0.35435839
Oct	0.96	102942	0.0947	0.19684191
		106590	0.0001	0.42791017
Nov	0.97	106590	0.0030	0.27014455
		109638	0.0006	0.51188003
Dec	0.96	102942	0.0005	0.98889507
		106590	0.0774	0.21486911
		109638	0.0856	-0.26120517

Estimates for 1980:

January	2.2	July	0.1
February	1.5	August	0.0
March	1.5	September	0.3
April	0.5	October	1.3
May	1.1	November	1.3
June	0.7	December	1.8

When one of the comparison stations also had the same month missing, the resulting equation could not be solved for an estimate. This occurred in approximately thirty cases. If an earlier step in the stepwise procedure produced an acceptable equation, it was used. Otherwise, the regression was rerun without the unusable comparison station.

All estimates were calculated on an individual monthly basis rather than by seasons or years to minimize the effect of error on the twenty-five year mean values, and to take advantage of all partial year records. Acceptable estimates were calculated for all missing data values. All NWS data used in the study are listed in Table 3.3. Mean annual precipitation values include estimates.

Table 3.3 National Weather Service Data
1961-1985

INDEX NUMBER	STATE	STATION NAME	MAP	ELEV
160010	ID	ABERDEEN EXP STA	9.90	4441
160227	ID	AMERICAN FALLS 1 SW	12.40	4320
160282	ID	ANDERSON DAM	21.60	3880
160375	ID	ARCO 3 SW	11.60	5330
160448	ID	ARROWROCK DAM	19.40	3280
160470	ID	ASHTON	21.30	5260
160667	ID	BAYVIEW MODEL BASIN	25.00	2080
160915	ID	BLACKFOOT 2 SSW	11.80	4490
161002	ID	BLISS 4 NW	10.80	3280
161018	ID	BOISE LUCKY PEAK DAM	14.60	2840
161022	ID	BOISE WSFO	12.20	2838
161079	ID	BONNERS FERRY 1 SW	23.70	1860
161195	ID	BRUNEAU	8.00	2530
161303	ID	BURLEY FAA AP	10.00	4160
161363	ID	CABINET GORGE	32.30	2260
161380	ID	CALDWELL	11.10	2370
161408	ID	CAMBRIDGE	20.30	2650
161514	ID	CASCADE 1 NW	22.70	4900
161636	ID	CENTERVILLE ARB RCH	27.40	4300
161663	ID	CHALLIS	7.90	5175
161671	ID	CHILLY BARTON FLAT	8.30	6260
161932	ID	COBALT	17.90	5010
161956	ID	COEUR D' ALENE R S	26.10	2160
162187	ID	COUNCIL	26.90	2950
162260	ID	CRATERS OF THE MOON	17.00	5900
162444	ID	DEER FLAT DAM	10.30	2510
162575	ID	DIXIE	31.10	5620
162604	ID	DOLLARHIDE SUMMIT	38.40	8650
162676	ID	DRIGGS	17.30	6120
162707	ID	DUBOIS EXP STA	13.10	5450
162875	ID	ELK CITY R S	30.00	4060
162892	ID	ELK RIVER 1 S	37.30	2920
162942	ID	EMMETT 2 E	13.40	2370
163108	ID	FAIRFIELD R S	16.60	5070
163143	ID	FENN R S	39.10	1590
163297	ID	FORT HALL IND AGENCY	12.50	4460
163417	ID	GALENA	30.00	7300
163448	ID	GARDEN VALLEY R S	23.90	3134
163554	ID	GIBBONSVILLE	15.90	4480
163631	ID	GLENNS FERRY	10.60	2510
163732	ID	GRACE	15.90	5550
163760	ID	GRAND VIEW 2 W	7.30	2400
163771	ID	GRANGEVILLE	23.70	3360
163882	ID	GROUSE	14.10	6100

Table 3.3 continued

INDEX NUMBER	STATE	STATION NAME	MAP	ELEV
163942	ID	HAILEY AP	17.50	5310
163964	ID	HAMER 4 NW	9.60	4790
164140	ID	HAZELTON	10.30	4060
164150	ID	HEADQUARTERS	39.60	3140
164268	ID	HILL CITY 1 W	15.10	5090
164295	ID	HOLLISTER	10.80	4550
164384	ID	HOWE	10.00	4820
164442	ID	IDAHO CITY	25.60	3970
164456	ID	IDAHO FALLS 16 SE	16.50	5850
164455	ID	IDAHO FALLS 2 ESE	12.70	4770
164460	ID	IDAHO FALLS 46 W	9.40	4940
164457	ID	IDAHO FALLS FAA AP	11.30	4730
164598	ID	ISLAND PARK	32.00	6300
164670	ID	JEROME	10.90	3740
164793	ID	KAMIAH	24.00	1210
164831	ID	KELLOGG	30.00	2320
165011	ID	KOOSKIA	24.70	1260
165038	ID	KUNA 2 NNE	10.20	2680
165241	ID	LEWISTON WSO AP	12.70	1413
165275	ID	LIFTON PUMPING STA	11.10	5930
165462	ID	MACKAY R S	10.00	5900
165544	ID	MALAD	16.30	4552
165559	ID	MALAD CITY	15.00	4470
165685	ID	MAY	8.30	5110
165708	ID	MCCALL	28.40	5025
165980	ID	MINIDOKA DAM	10.30	4210
166053	ID	MONTPELIER R S	15.00	5943
166152	ID	MOSCOW UNIV OF IDAHO	25.40	2660
166174	ID	MOUNTAIN HOME AFB	11.30	3190
166388	ID	NEW MEADOWS R S	25.50	3870
166424	ID	NEZ PERCE	21.60	3150
166542	ID	OAKLEY	11.90	4600
166590	ID	OLA 4 S	20.00	2990
166681	ID	OROFINO	25.00	1030
166764	ID	PALISADES	20.70	5385
166844	ID	PARMA EXP STA	12.20	2220
166877	ID	PAUL 1 ENE	9.80	4210
166891	ID	PAYETTE	11.50	2150
167040	ID	PICABO	13.90	4880
167046	ID	PIERCE	42.90	3190
167211	ID	POCATELLO WSO AP	12.50	4450
167264	ID	PORTHILL	20.70	1775
167301	ID	POTLATCH 3 NNE	25.40	2600

Table 3.3 continued

INDEX NUMBER	STATE	STATION NAME	MAP	ELEV
167386	ID	PRIEST RIVER EXP STA	31.90	2380
167648	ID	REYNOLDS	11.40	3930
167673	ID	RICHFIELD	12.20	4310
167706	ID	RIGGINS	17.00	1800
168062	ID	SAINT MARIES	29.80	2220
168137	ID	SANDPOINT EXP STA	33.80	2120
168380	ID	SHOSHONE 1 WNW	10.30	3950
168022	ID	ST ANTHONY 1 WNW	14.50	4950
168786	ID	STREVELL	11.70	5280
168928	ID	SWAN FALLS PWR HOUSE	8.40	2330
168937	ID	SWAN VALLEY 2 E	17.30	5270
169065	ID	TETONIA EXP STA	17.30	6170
169119	ID	THREE CREEK	12.70	5460
169498	ID	WALLACE WOODLAND PRK	39.40	2940
169560	ID	WARREN	28.70	5899
169638	ID	WEISER 2 SE	11.80	2103
300110	MT	ALDER 17S	5.40	5850
302404	MT	DILLON AIRPORT	10.10	5216
302500	MT	DRUMMUND AVIATION	13.30	3943
302793	MT	ENNIS	13.30	4953
303366	MT	GALLATIN GATEWAY 10SSW	23.10	5480
303570	MT	GLEN 4N	9.20	5050
303707	MT	GRANT 4NE	9.70	5840
303885	MT	HAMILTON	13.60	3529
303984	MT	HAUGAN 3E	28.90	3124
304038	MT	HEBGEN DAM	31.60	6489
304084	MT	HERON 2NW	34.10	2240
305020	MT	LIBBY 32SSE	25.90	3600
305015	MT	LIBBY RS 1NE	18.20	2080
305030	MT	LIMA	12.00	6273
305745	MT	MISSOULA WSO AP	13.60	3190
305811	MT	MONIDA	14.40	6785
307318	MT	SAINT REGIS R.S.	21.50	2680
307894	MT	STEVENSVILLE	12.50	3375
307964	MT	SULA 3ENE	16.30	4475
308043	MT	SUPERIOR	16.80	2710
308211	MT	THOMPSON FALLS PH	23.10	2380
308597	MT	VIRGINIA CITY	17.00	5758
308857	MT	WEST YELLOWSTONE	22.40	6657
309067	MT	WISDOM	11.50	6060
309082	MT	WISE RIVER 3WNW	11.60	5730
321905	NV	CONTACT	10.80	5365
322189	NV	DEETH	12.10	5338

Table 3.3 continued

INDEX NUMBER	STATE	STATION NAME	MAP	ELEV
323114	NV	GIBBS RANCH,NV	10.40	6000
325392	NV	MOUNTAIN CITY RS,NV	13.40	5620
325818	NV	OROVADA,NV	10.90	4310
326005	NV	PARADISE VALLEY 1NW,NV	9.80	4675
328346	NV	TUSCARORA,NV	12.70	6180
328988	NV	WELLS	10.80	5650
410412	OR	BAKER FAA AP, ORE.	11.10	3368
410417	OR	BAKER KBKR, OR	12.00	3444
411924	OR	COVE 1ENE,OR	21.50	2920
412135	OR	DANNER, ORE.	12.60	4225
412672	OR	ENTERPRISE, 2S, OR	13.40	3880
413604	OR	HALFWAY, ORE.	22.00	2670
414175	OR	IRONSIDE 2W, OR	12.00	3915
418746	OR	UNION ES, OR	14.00	2765
418797	OR	VALE, ORE.	10.00	2240
418997	OR	WALLOWA, ORE.	17.60	2923
491244	UT	CAUSEY DAM	23.50	5500
493087	UT	GARDEN CITY SUMMIT	31.30	7600
494727	UT	KLONDIKE NARROWS	43.60	7400
495207	UT	LOST CREEK RESERVOIR	26.00	6125
496414	UT	OGDEN SUGAR FACTORY	17.00	4280
497271	UT	RICHMOND	17.70	4680
498760	UT	TONY GROVE RANGER STATIO	34.90	6250
531395	WA	CHEWELAH 2S	21.30	1635
531650	WA	COLVILLE AP	18.90	1862
560027	WY	AFTON	18.70	6210
560603	WY	BEDFORD 3SE	22.10	6425
560695	WY	BIG PINEY	8.90	6820
560865	WY	BONDURANT 3NW	20.90	6504
561736	WY	CHURCH BUTTES G.	9.00	7075
563100	WY	EVANSTON 1 E	11.50	6810
563396	WY	FONTENELLE DAM	7.00	6480
564910	WY	JACKSON	16.90	6230
565105	WY	KEMMERER	10.40	6958
565345	WY	LAKE YELLOWSTONE	20.50	7770
566165	WY	MERNA	14.80	7700
566428	WY	MOOSE	21.70	6470
566440	WY	MORAN	24.60	6750
566555	WY	MOUNTAIN VIEW	8.60	6800
568315	WY	SNAKE RIVER STATION	32.70	6920
569025	WY	TOWER FALLS	17.10	6266
569905	WY	YELLOWSTONE PARK	15.70	6230

173 Stations

SCS SNOTEL DATA

The SCS SNOTEL network was designed to quantify precipitation and snow water equivalent at high elevations. Gages were placed strategically to capture precipitation in the major precipitation production zones and to maximize the ability to correlate SNOTEL records with existing long-term snow course records (Palmer, 1989). The system is fully automated. Precipitation records are transmitted to a central computer without manual reading. The SNOTEL data play a large role in this study. For the first time, a good network of high elevation precipitation stations is available in Idaho. The last mean annual precipitation map for Idaho was produced in 1965 by the NWS. At that time, high elevation precipitation values were extrapolated from low-lying National Weather Service Stations using models developed from a few storage gages. This procedure was shown to be inappropriate in Colorado in the absence of good high elevation data (Crow, 1982). The SNOTEL network gives us a much better starting point for the modeling of precipitation throughout the state (Doesken and others, 1984).

All available SNOTEL data were gathered. It is important that all data records in the study be adjusted to the same base period (Peck, 1972). Twenty-five year averages for 1961-1985 were calculated by the Soil Conservation Service Snow Surveys staff through a series of comparisons to nearby NWS storage gauges with long periods of record (Palmer, 1988).

For each year that a particular SNOTEL site was in place, the percentage of normal precipitation was calculated at all nearby gauge sites. This percentage of normal was based on the 1961-1985 base period. It was assumed that even though precipitation totals are quite different between stations at various elevations, the percentage of normal precipitation for a given year should be fairly constant over relatively small areas. Mean annual values were estimated for each year of SNOTEL record based on the percentage of normal observed at the nearby stations. These five to eight yearly estimates were compared and averaged to yield final estimates of 1961-1985 mean annual precipitation.

The mean monthly precipitation values were also calculated. This was done by dividing the mean annual value into monthly values based on the monthly distribution experienced at the surrounding NWS gauge sites. Some high elevation stations were expected to have monthly distributions which were quite different from NWS stations in the valleys. The seasonal distribution of precipitation observed at some of the valley stations, shows summer maximums, which are dominated by intense summertime convection storms. These valley stations often receive little or no precipitation from the passing of wintertime fronts which account for the winter maximums at high elevations. The monthly distribution was checked against snow course data and adjusted where needed. It is apparent that the strong winter maximum observed at high elevations does not always apply to the monthly distribution in the valleys (Palmer, 1989).

One record was downloaded for each station. Each record contained the station name, October through September monthly estimates, and a mean annual precipitation estimate. These records were then combined with the index information (elevation, latitude, longitude, etc). No adjustments were made on the SCS estimates. All SNOTEL stations used are listed in Table 3.4.

Table 3.4 SCS SNOTEL Data

INDEX NUMBER	STATE	STATION NAME	MAP	ELEV
1611F11	ID	ASPEN GROVE	27.80	6500
1615F04	ID	ATLANTA SUMMIT	46.50	7580
1615E11	ID	BANNER SUMMIT	41.20	7040
1616E11	ID	BEAR BASIN	37.90	5350
1613F03	ID	BEAR CANYON	28.40	7900
1616A08	ID	BEAR MOUNTAIN	86.70	5400
1616E10	ID	BEAR SADDLE	37.30	6180
1615F07	ID	BENNETT MOUNTAIN	31.60	6560
1615E02	ID	BIG CREEK SUMMIT	48.60	6580
1616G01	ID	BULL BASIN	35.00	5460
1614F11	ID	COUCH SUMMIT	26.80	6840
1615E08	ID	COZY COVE	33.00	5380
1611E37	ID	CRAB CREEK	28.80	6860
1615E04	ID	DEADWOOD SUMMIT	62.00	6860
1616C15	ID	ELK BUTTE	67.10	5550
1611G06	ID	EMIGRANT SUMMIT	41.30	7390
1614F01	ID	GALENA	21.20	7440
1614F12	ID	GALENA SUMMIT	34.00	8780
1611G16	ID	GIVEOUT	21.00	6840
1613E27	ID	HILTS CREEK	25.80	8000
1613G01	ID	HOWELL CANYON	37.60	7980
1615B21	ID	HUMBOLDT GULCH	50.00	4250
1615E09	ID	JACKSON PEAK	47.00	7070
1614C05	ID	LOLO PASS	53.10	5240
1615B02	ID	LOOKOUT	54.50	5140
1615B14	ID	LOST LAKE	88.00	6110
1614F03	ID	LOST-WOOD DIVIDE	36.40	7900
1614G02	ID	MAGIC MOUNTAIN	31.30	6880
1613E18	ID	MEADOW LAKE	34.90	9150
1614E01	ID	MILL CR SUMMIT	31.00	8800
1613E06	ID	MOONSHINE	24.80	7440
1613D16	ID	MOOSE CREEK	30.40	6200
1615F01	ID	MORES CR SUMMIT	52.40	6100
1614E04	ID	MORGAN CREEK	30.50	7600
1616A04	ID	MOSQUITO RIDGE	58.80	5200
1615D06	ID	MOUNTAIN MEADOWS	51.60	6360
1616G07	ID	MUD FLAT	21.60	5730
1612G18	ID	OXFORD SPRING	28.80	6740
1614C04	ID	SAVAGE PASS	49.30	6170
1616A10	ID	SCHWEITZER BASIN	68.50	6090
1615D01	ID	SECESH SUMMIT	52.20	6520
1616C01	ID	SHERWIN	43.50	3200
1611G01	ID	SOMSEN RANCH	29.20	6800
1616E05	ID	SQUAW FLAT	46.10	6240
1613F09	ID	SWEDE PEAK	28.60	7640

Table 3.4 continued

INDEX NUMBER	STATE	STATION NAME	MAP	ELEV
1614F04	ID	VIENNA MINE	48.80	8960
1616D08	ID	WEST BRANCH	44.20	5560
1613F0	ID	WHITE KNOB	23.00	7700
30MX08	MT	BANFIELD MTN	38.40	5600
30MQ44	MT	BARKER LAKES	36.30	8250
30MN08	MT	BEAGLE SPRINGS	25.90	8850
30MI38	MT	BEAVER CREEK	38.00	7850
30MH17	MT	BIG SKY SCS	33.20	7700
30MI35	MT	BLACK BEAR	63.70	7950
30MQ13	MT	BLACK PINE	29.40	7100
30MR10	MT	BLOODY DICK	30.00	7550
30MD31	MT	BOX CANYON	26.90	6700
30MR26	MT	CALVERT CREEK	19.30	6430
30MI29	MT	CARROT BASIN	54.10	9000
30MH23	MT	CASHE CREEK	26.50	7800
30MM06	MT	CHRISTENSEN RANCH	14.50	6000
30MH08	MT	CLOVER MEADOW	37.20	8800
30MD30	MT	COLLEY CREEK	23.60	6300
30MQ33	MT	COMBINATION	21.90	5600
30MR19	MT	DARKHORSE LAKE	49.20	8700
30MN07	MT	DIVIDE	28.10	7800
30MH12	MT	FOUR MILE	28.60	6900
30MX05	MT	GARVER CREEK	26.50	4250
30MX03	MT	HAWKINS LAKE	49.20	6450
30MZ10	MT	HOODOO BASIN	72.80	6050
30MP22	MT	KRAFT CREEK	41.80	4750
30MI03	MT	LAKEVIEW RIDGE	33.50	7400
30MS23	MT	LEMHI RIDGE	28.60	8100
30MD13	MT	LICK CREEK	35.40	6860
30MH11	MT	LOWER TWIN	44.60	7900
30MQ38	MT	LUBRECHT FLUME	22.90	4680
30MQ14	MT	LUBRECHT FOREST HQ	18.20	4100
30MI31	MT	MADISON PLATEAU	44.10	7750
30MD19	MT	MILL CREEK	27.50	7500
30MD12	MT	MONUMENT PEAK	40.00	8850
30MM11	MT	MULE CREEK	29.80	8300
30MW02	MT	NEZ PERCE CAMP	31.80	5650
30MD07	MT	NORTHEAST ENTRANCE	25.90	7350
30MQ36	MT	PETERSON MEADOW	27.90	7200
30MD24	MT	PLACER BASIN	41.60	8830
30MX12	MT	POORMAN CREEK	65.80	5100
30MH21	MT	ROCK CREEK MEADOWS	32.40	8160
30MR22	MT	SADDLE MOUNTAIN	44.40	7900
30MD16	MT	SHOWER FALLS	53.00	8100
30MQ03	MT	SKALKAHO SUMMIT	42.60	7250
30MP23	MT	SKYLARK TRAIL	53.60	6200

Table 3.4 continued

INDEX NUMBER	STATE	STATION NAME	MAP	ELEV
30MH13	MT	TAYLOR PEAKS	42.60	8500
30MI24	MT	TEPEE CREEK	30.20	8000
30MV13	MT	TWELVEMILE	46.60	5600
30MV12	MT	TWIN LAKES	66.20	6400
30MQ43	MT	WARM SPRINGS	48.50	7800
30MI30	MT	WHISKEY CREEK	38.60	6800
32S049	NV	BEAR CREEK,NV	38.50	7800
32S231	NV	BIG BEND,NV	19.30	6700
32S345	NV	BUCKSKIN LOWER,NV	23.90	6700
32S229	NV	GOAT CREEK,NV	37.70	8800
32S239	NV	GRANITE PEAK,NV	32.00	7800
32S232	NV	JACK CREEK UPPER,NV	31.90	7250
32S222	NV	JACKS PEAK,NV	44.80	8420
32S240	NV	LAMANCE CREEK,NV	23.00	6000
32S233	NV	LAUREL DRAW,NV	27.90	6700
32S227	NV	POLE CREEK RS,NV	24.40	8330
32S050	NV	SEVENTYSIX CREEK,NV	23.50	7100
32S354	NV	TAYLOR CANYON,NV	13.00	6200
49S105	UT	BEN LOMOND PEAK	75.60	8000
49S262	UT	BEN LOMOND TRAIL	48.80	6000
49S264	UT	BUG LAKE	28.40	7950
49X006	UT	BURTS-MILLER RANCH	19.70	7900
49S270	UT	DRY BREAD POND	32.70	8350
49S076	UT	HORSE RIDGE	42.80	8260
49S276	UT	LITTLE BEAR (UPPER)	44.30	6550
49S082	UT	MONTE CRISTO RANGER STAT	43.50	8960
49X850	UT	PINE CANYON	40.20	8000
49S114	UT	TONY GROVE LAKE	54.00	8400
56S292	WY	BASE CAMP	33.40	7030
56S294	WY	BLIND BULL SUMMIT	38.20	8750
56S299	WY	CANYON	29.40	7940
56S404	WY	COULTER CREEK	42.70	7020
56S302	WY	GRASSY LAKE	58.20	7265
56X044	WY	GROVER PARK DIVIDE	31.90	7000
56S461	WY	INDIAN CREEK	34.70	8240
56S097	WY	LEWIS LAKE DIVIDE	56.30	7850
56S306	WY	LOOMIS PARK	35.70	8240
56S100	WY	PHILLIPS BENCH	41.90	8200
56S099	WY	SALT RIVER SUMMIT	26.30	7700
56S467	WY	SNIDER BASIN R.S.	20.90	8060
56S089	WY	SPRING CREEK DIVIDE	31.20	9000
56S310	WY	SYLVAN LAKE	39.90	8420
56S311	WY	TOGWOTEE PASS	40.70	9580
56S410	WY	TWO OCEAN PLATEAU	48.60	9160
56S473	WY	WILLOW CREEK	48.00	8450

136 Stations

function of elev, 2) regression analysis expressing MAP as a function of elev, qx, qy, ALI5, ALI10, ALI20, and ALI30, and 3) using a Log (base 10) transformation on mean annual precipitation to form a new dependent variable.

After the preliminary investigation was completed on the entire study area, the study area was divided into seven regions to further refine the models. First, these regions were tested using two types of discriminant analysis. After the regions were adjusted, the remaining six steps were performed to arrive at the most appropriate final regression equations. These steps consisted of a set of regression runs (one run for each region). These regression runs were set up as follows:

1) Log (10) MAP as a function of elev, qx, qy, ALI5, ALI10, ALI20, and ALI30.
2) Log (10) MAP as a function of elev.
3) Log (10) mean seasonal precipitation as a function of elev, qx, qy, ALI5, ALI10, ALI20, and ALI30.
4) Log (10) mean monthly precipitation as a function of elev, qx, qy, ALI5, ALI10, ALI20, and ALI30. (This was done for Region One only.)
5) Log (10) MAP as a function of elev, qx, qy, and each ALI measure separately.
6) Final regression estimation equations - log (10) MAP as a function of elev, qx, qy, and the best ALI measure for each region.

The PC-SAS package was used to run each of these statistical analyses. This progression of steps used to refine the precipitation estimation models will be explained further in this chapter.

IV. ESTIMATION OF MEAN ANNUAL PRECIPITATION AT UNGAGED POINTS

In order to map precipitation at the desired scale (1:250,000), we need to have a method of estimating point precipitation values at locations throughout the study area where no gages exist. A series of regression equations were derived in an effort to come up with the best procedure for making these estimates. The goal was to make the equations as accurate as possible while also keeping them easy to use. We must keep in mind that these equations will be used to estimate hundreds of ungaged locations during the map production phase of the project, so the number and complexity of variables must be kept to a minimum. The independent variables used were elev (elevation), qx (easting), qy (northing), and various versions of an ALI (air mass lifting index). These variables are explained in the following section. Of the 309 original data points, 256 were used in the regression runs. The fifty-three remaining were all on the perimeter of the study area, either off the available quadrangles, or too close to the edge of the study area to measure the ALI variables. Therefore, they were discarded for this portion of the study.

The regression models were refined through a progression of eleven steps which are described in this chapter. The first three steps consisted of general exploratory regression analysis on the entire study area. They are: 1) regression analysis expressing mean annual precipitation (MAP) as a

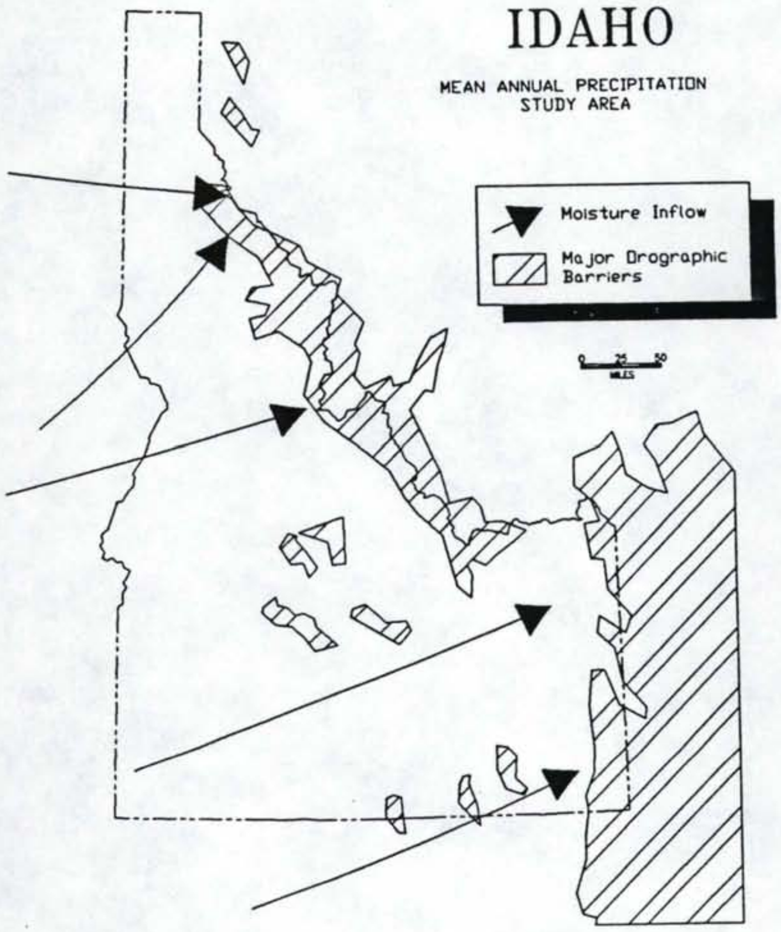


Figure 4.1 Principal Moisture Inflow Paths and Major Orographic Barriers (Miller, 1972)

THE INDEPENDENT VARIABLES

Elev

The elev variable represents the elevation above mean sea level as recorded in the NWS and SCS SNOTEL indices. Elevations of points to be estimated are measured on the one by two degree quadrangles to the nearest 200 feet.

Qx and Qy

Qx and qy are surrogates for longitude and latitude, respectively. These are cartesian coordinates based on a polyconic map projection with the origin at forty degrees north, 120 degrees west. The coordinates were scaled to one inch units at 1:250,000. This transformation was done for three reasons: 1) to make sure that the maps would overlay the U.S.G.S. quadrangles, 2) to facilitate convenient coordinate measurements, and 3) to make the units of qx and qy equal and consistent throughout the study area (unlike latitude and longitude).

ALI

ALI stands for air mass lifting index. This is an experimental variable, which assigns a relative vertical lift or decent to each station encountered by moisture bearing winds (Miller, 1972). In a regression analysis of precipitation based on elevation, slope, aspect, distance to barriers, barrier height, and a shield effect, Storr and Ferguson (1972, p. 251) stated:

> It is recognized that the land slope and aspect might be more effective in the regression analysis if they were combined with wind direction during precipitation into a single orographic vertical motion parameter.

Their statement sums up the idea behind the ALI variable.

Four ALI measures were tested at each station. The elevations measured at each of five, ten, twenty, and thirty miles upwind of the station along the mean wintertime 700 mb bearing were subtracted from the elevation of the station. This resulted in four different measures of rise (positive ALI) or decent (negative ALI) to each station. These measures will be called ALI5, ALI10, ALI20, and ALI30.

The 700 mb mean wintertime wind bearing was used for two reasons. First, the majority of the precipitation in the area falls during the winter months. Second, the variable was designed to explain orographically induced differences in the precipitation pattern. In the winter, the flow of moisture is much more organized and stronger than in the summer. Orographic influences are greater when the flow is steady and of greater intensity (Houghton, 1979).

A mean 700 mb wind bearing of 261 degrees (nine degrees south of west) was used (Arnold, 1989). Although this parameter varies slightly from place to place in Idaho, the flow bearing is constant enough to warrant the use of one bearing for the entire study area. Figure 4.1 shows the principal paths of moisture inflow and major orographic barriers as depicted by Miller (1972). These moisture inflow bearings confirm the findings of Arnold (1989) and the use of 261 degrees as the mean moisture movement bearing.

THE REGRESSION ANALYSIS

The Entire Study Area

The first step involved a regression analysis of mean annual precipitation for the entire study area based on elevation alone. Although elevation was significant at the 0.0001 level, the r-square was 0.16, and the root mean square error was 14.32. This yields a 95 percent probability of estimating within plus or minus twenty-eight inches at the mean. This preliminary analysis suggests that, while the precipitation pattern is not determined by elevation exclusively, elevation does contribute significantly to the explanation of the variation in precipitation throughout Idaho. It is clear that other variables and further refinements are needed.

The next run set mean annual precipitation as a function of elev, qx, qy, ALI5, ALI10, ALI20, and ALI30. The resulting r-square was 0.62. The root mean square error was now 9.82. This yields a 95 percent confidence band of plus or minus 19.64 inches. Elev, qx, qy, ALI5, ALI20, and ALI30 were all significant at the 0.05 level. Although further refinement is needed for accurate estimation models, it is apparent that elevation, location, and the airmass lifting measures all help to explain the precipitation regime.

One of the problems which became apparent during this analysis is the combination of the NWS and SCS SNOTEL data populations. Figure 4.2 shows the frequency distribution of mean annual precipitation from the combined data set. The distribution is far from the normal as the SNOTEL data (high precipitation, high elevation stations) skew to the right. Doesken and Schaefer (1987) concluded that the SNOTEL data set for Colorado has unique characteristics which make direct comparisons to NWS records difficult and sometimes inappropriate.

The problems presented by the non-normal distribution resulting from the combination of the two gauge populations were rectified by taking the log (base 10) of mean annual precipitation (Figure 4.3). The resulting variable will be referred to as logMAP hereafter.

The same regressions were recalibrated with logMAP as the dependent variable. The resulting r-square was still 0.62, but the root mean square error was 0.159 measured in log form. It should be noted that the equation was much more sensitive in this form, as slight biases in fitting logMAP could translate to large biases in mean annual precipitation (Afifi and Clark, 1984).

REGIONALIZATION

Because of the increased sensitivity of the logMAP model, the study area was divided into seven regions (Figure 4.4), which were expected to have fairly homogeneous precipitation regimes. The tradeoff between the homogeneity of small regions and the degrees of freedom allowed by the sample sizes was considered. A minimum of fifteen stations was deemed acceptable for any one region.

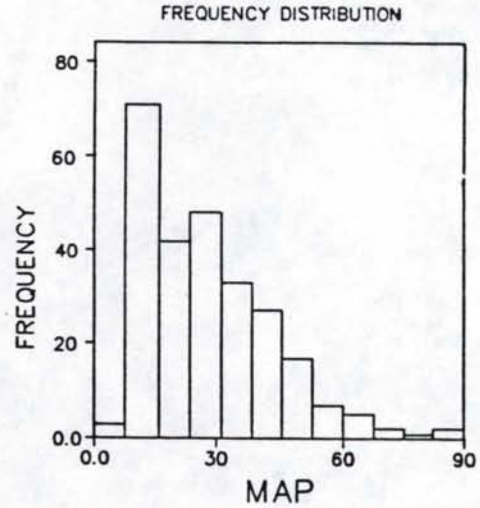


Figure 4.2 Frequency Distribution of MAP

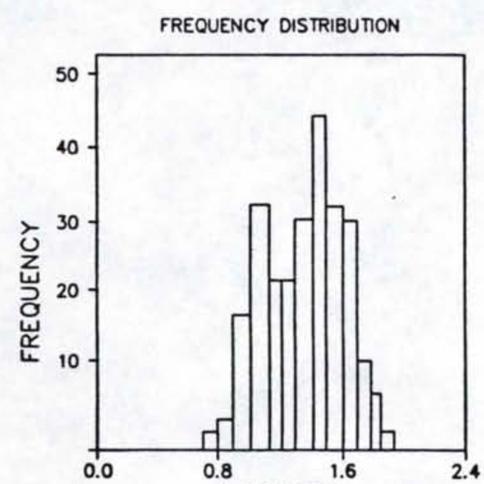


Figure 4.3 Frequency Distribution of LogMAP

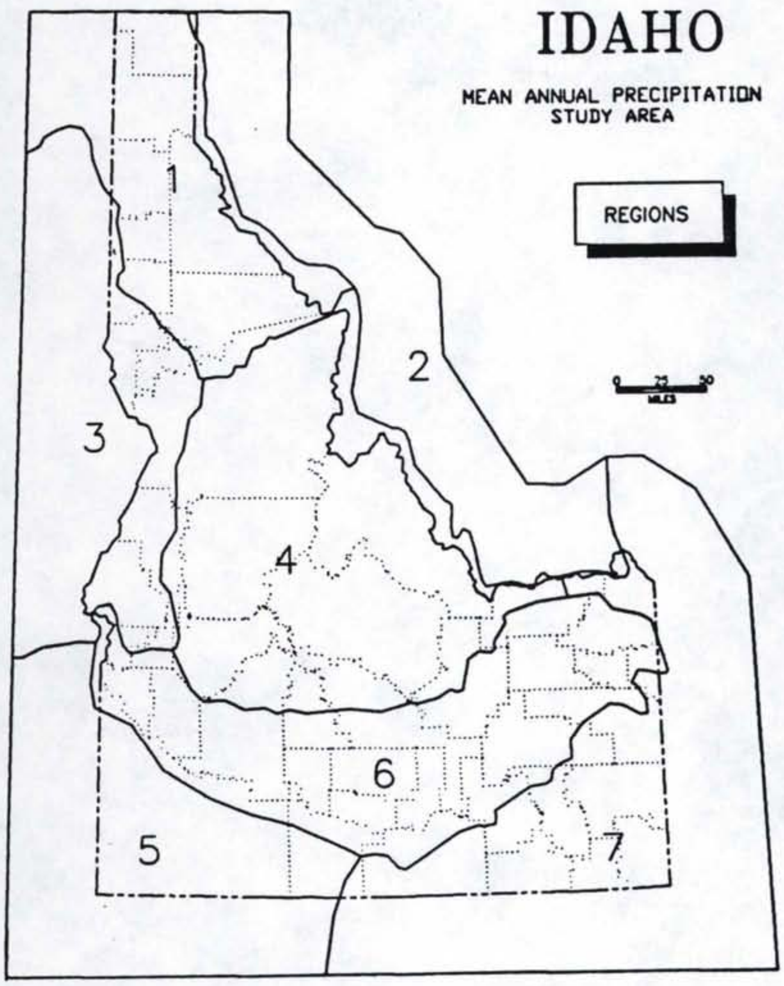


Figure 4.4 Study Areas and Regions

Regional boundaries were based on physiographic features delineated on the topographic maps. It was assumed that major changes in landform would result in changes in the precipitation formation processes. For example, it was expected that the precipitation regime on the west side of the Bitterroot Mountains should be very different from that on the east. Likewise, the precipitation formation processes at work in the Snake River Plain, east of Boise, and the Snake-Columbia Canyons to the west should be very different from those at work in the nearby mountains. The climatic division as defined by The National Climatic Data Center, NOAA, were also used as guidelines. Some of these intra-regional boundaries are fairly obvious on the topographic maps, such as the crest of the Bitterroots and the border formed by the Sawtooth Mountains abutting the Snake River Plain. Other regional breaks were needed to separate large areas where no abrupt landform changes were apparent, such as the border between Region One and Region Four (the Middle Fork of the Clearwater River and the Lochsa River) and the border between Regions Four and Seven (Monida Pass). Although no dramatic landform changes are associated with some of these boundaries, they were all needed to differentiate the various general precipitation formation zones of Idaho. These boundaries should represent statistical breaks of varying significance, depending upon the amount of landform change associated with them.

Two types of discriminant analysis were used to test the statistical homogeneity of the regions and to make minor adjustments in questionable areas. First, the entire study area was tested in a seven-class discriminant analysis to check the overall statistical individuality and homogeneity of the regions. This was done twice, once before and once after minor changes were made. Then, each pair of adjacent regions was tested in a two class discriminant analysis to verify each intraregional boundary. The variables used in the discriminant test were the same ones used in the regression analysis:

1) LogMAP	5) ALI5
2) Qx	6) ALI10
3) Qy	7) ALI20
4) Elev	8) ALI30

Seven Class Discriminant Analysis

The initial seven-class discriminant test produced an average percentage of correctly classified stations of 81.3 percent. Each misclassified observation was examined on the topographic maps. The fact that the discriminant function placed a station in a different region was not grounds for the regions to be changed, but it did, however, warrant further investigation. The following changes were made after careful examination of the discriminant results and the landforms on the maps:

1) Station 30MI22 was moved from Region Two to Region Four.

2) Stations 30MR22, 30MI03 and 1611E37 were moved to

Region Seven.

3) Station 161514 was moved from Region Three to region Four.

All of these moves were accomplished by slightly adjusting the boundaries.

The final seven-class discriminant test was rerun after these changes were made. Appendix A.1 shows the results of this test. The average percentage of correctly classified stations was 82.9 percent. This means that the regions can be separated by their unique statistical properties fairly well. This discriminant test gives justification for the separation of the study area into regions. The discriminant function derived by the computer was able to correctly classify most of the stations. The goal of the discriminant analysis was simply to verify whether or not the regional separations were appropriate and statistically recognizable.

Two Class Discriminant Tests

Each boundary between regions was tested to verify if it formed a recognizable statistical separation. Each pair of adjacent regions was tested in a two class analysis. Again, all of the variables involved in the regression analysis were used. Appendices A.2 through A.14 are the results of these two class discriminant runs. Only misclassified stations were listed.

The combined average of the percentages of correctly classified observations for all of the two region pairs was

96.75 percent. All of the discriminant functions classified over 90 percent of the stations correctly with the exception of three. Seven stations were misclassified from Region Four when tested with Region Three (87 percent correctly classified). The probability of the classification was less than 0.68 in all of these cases. This result was not surprising because the long border between these regions is not based on any severe landform changes.

Thirteen of the fifteen stations in Region Five were correctly classified when tested against Region Six (87 percent). This too was not surprising because of the relatively few data points in Region Five.

Five stations were misclassified from Region Six when tested against Region Seven (89 percent). These regions also share a long border and the topographic change between them is not abrupt or well defined.

In eight cases, 100 percent of the stations were correctly classified. All of the discriminant results from both the seven-class test and each of the two class tests, show that the regional boundaries form recognizable statistical breaks and that the pattern of data values was neither totally random, nor completely spatially dependent.

The ability of each two class discriminant function to correctly clarify the stations was controlled by more than just the homogeneity of the regions. Other contributing factors were the degree of landform change at the boundary, and the relative length of the boundary. For example, it would be expected that the boundary between Region Two and Region Four should discriminant very well because of the severe topographic change associated with the crest of the Bitterroot Mountains. It would also be expected that Regions Three and Six be statistically separate because of the short boundary they share.

PRELIMINARY REGRESSION ANALYSIS BY REGION

A regression analysis was run for each region using logMAP as the dependent variable, and elev, qx, qy, ALI5, ALI10, ALI20, and ALI30 as the independent variables. This was not an attempt to formulate final estimation equations, but rather to see how each variable would work in each region. Tables 4.16 through 4.22 show the results of each test. Only variables significant at the 10 percent level are shown. The simple correlation matrices can be found in Appendix B.1 through B.7. Table 4.23 show the results of the same test using only elev as the independent variable.

As expected, each region produced different results in the regression analysis. The inter-correlations between the independent variables is somewhat controlled by the topography of the regions. This changes dramatically from one region to another. The inter-correlation between elevation, qx, and qy is determined by the general topographic trends. For example, as we go from north to south in Region Five, the elevation increases steadily. This is why elev and qy have a strong, negative correlation. By definition, all of the ALI measures should be somewhat inter-correlated, but the degree to which they are correlated depends upon the relative relief of the region.

Region One

The simple correlation matrix for Region One (Appendix B.1) showed the highest correlation (r = 0.95) between elev and logMAP of any of the regions. The overwhelming control of elevation in this region was evident in the multiple regression analysis, as it accounted for 91 percent of the variance in precipitation, and was significant at the one percent level (Table 4.16).

Table 4.1 Preliminary Regression Results for Region One

F Value	Prob>F
53.623	0.0001

Root MSE	0.04599	R-square	0.9542
C.V.	2.86788		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob >|T|
ELEV	1	0.000076480	0.00002182	3.505	0.0025

Region Two

In Region Two, we find an interesting problem with multicollinearity. When elev was used alone, the r-square was 0.005. This means that it explains less than one percent of the variance. The correlation coefficient for elevation and logMAP was lower in Region Two than in any other region (0.0737) (Appendix B.2). Qx explained most of the variation when all of the variables were included. This means that easting plays the most important role in explaining precipitation. The further a station is from the crest of the Bitterroots, the lower its precipitation. Following qx, elev was significant at the 10 percent level. This seems strange because of its small effect alone. One possible explanation is that elev is important after controlling for qx. The multicollinearity in this region renders the equation unstable. The expected accuracy of estimation is low.

Table 4.2 Preliminary Regression Results for Region Two

F Value	Prob>F
6.198	0.0007

Root MSE	0.15708	R-square	0.6955
C.V.	12.45576		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob >|T|
ELEV	1	0.000148	0.00008192	1.806	0.0867
QX	1	-0.016316	0.00809523	-2.016	0.0582

Region Three

In Region Three, elevation alone explained 43 percent of the variance in logMAP. Yet, when the other variables were included, elev was not significant. Qx, ALI5, ALI10, and ALI20 were all significant at the 10 percent level, and they combined to explain 88 percent of the variance in logMAP. Again, elev was correlated with the other independent variables.

Table 4.3 Preliminary Regression Results for Region Three

F Value	Prob>F
7.414	0.0085

Root MSE	0.06916	R-square	0.8811
C.V.	5.09431		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob >|T|
QX	1	0.011589	0.00416316	2.784	0.0272
ALI5	1	0.000090863	0.00004089	2.222	0.0617
ALI10	1	-0.000084421	0.00004270	-1.977	0.0885
ALI20	1	0.000100	0.00002871	3.496	0.0100

Region Four

Region Four possesses by far the most complex terrain of any region in the study area. For this reason, none of the independent variables are inter-correlated to a high degree. This was one reason that elev, qx, qy, ALI20, and ALI30 all are significant at the 10 percent level. Only eight percent of the variances in logMAP was explained by elev alone, while 77 percent was explained by all of the variables combined. Because of the complexity of the terrain in this area, it was expected that ALI5 would work the best, as it would introduce the least "noise". The longer measures of ALI often covered more than one peak or valley. This expected outcome did not come forth in the results. ALI30 appears it will contribute the most, even in this complex topography.

Table 4.4 Preliminary Regression Results for Region Four

F Value	Prob>F
22.538	0.0001

Root MSE	0.11001	R-square	0.7742
C.V.	7.51105		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob >|T|
ELEV	1	0.000050663	0.00001659	3.055	0.0037
QX	1	-0.009163	0.00191260	-4.791	0.0001
QY	1	0.005625	0.00109245	5.149	0.0001
ALI20	1	0.000031620	0.00001311	2.412	0.0199
ALI30	1	0.000030358	0.00001037	2.929	0.0053

Region Five

Region Five has a fairly sparse network of gages (only fifteen stations). Elev explained 18 percent of the variation in logMAP. Eighty-nine percent was explained by all of the variables combined. ALI5, ALI20, and ALI30 were all significant at the 10 percent level. The sign on the ALI20 parameter estimate was negative. This was not at all what we would expect. ALI20 seemed to work backwards in this region.

Table 4.5 Preliminary Regression Results for Region Five

F Value	Prob>F
5.697	0.0250

Root MSE	0.13254	R-square	0.8692
C.V.	10.20765		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob >|T|
ALI5	1	0.000239	0.00010652	2.240	0.0664
ALI20	1	-0.000244	0.00007352	-3.315	0.0161
ALI30	1	0.000228	0.00009724	2.343	0.0576

Region Six

Region Six consists of the flattest, low-lying area of the state (the Snake River Plain), and is the driest of all the regions. This region rises slowly from west to east. Elev alone explained 40 percent of the variation in precipitation, while all of the other variables combined explained 76 percent. Elev, qx, qy, and ALI30 were all significant at the five percent level. It is understandable that ALI30 had the highest correlation with logMAP, because of the gentle rise along the mean moisture movement bearing experienced throughout this region. The longest measure of ALI expressed this rise the best. ALI measures at forty or fifty miles may have worked even better, but these long distances are hard to cover at the scale used.

Table 4.6 Preliminary Regression Results for Region Six

F Value	Prob>F
170426	0.0001

Root MSE	0.06816	R-square	0.7625
C.V.	6.22390		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob >|T|
ELEV	1	0.000070750	0.00001412	5.009	0.0001
QX	1	-0.002909	0.00073311	-3.968	0.0003
QY	1	0.004821	0.00121769	3.959	0.0003
ALI30	1	0.000040737	0.00001253	3.251	0.0024

Region Seven

Elev explained one-third the variation in logMAP when used alone. The other variables combined with elevation explained 69 percent of the variation. Qy, ALI10, ALI20, and ALI30 were all significant at the five percent level.

Table 4.7 Preliminary Regression Results for Region Seven

F Value	Prob>F
21.909	0.0001

Root MSE	0.13190	R-square	0.6928
C.V.	9.24318		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob >|T|
QY	1	0.005746	0.00099675	5.765	0.0001
ALI10	1	0.000027135	0.00001296	2.093	0.0400
ALI20	1	0.000049332	0.00002056	2.399	0.0192
ALI30	1	0.000047642	0.00001809	2.634	0.0104

Table 4.8 Regression Results from LogMAP Based on Elev

Region	R Square	Root MSE	Elevation Prob>¦T¦	Parameter Estimate (Intercept)	Parameter Estimate (Elev)
1	0.9124	0.05510	0.0001	1.244360	0.0000106
2	0.0054	0.24746	0.7147	1.210117	0.000010124
3	0.4264	0.11150	0.0083	1.143983	0.000098276
4	0.0811	0.20875	0.0369	1.224544	0.000037536
5	0.1836	0.23415	0.1264	0.978169	0.0000058420
6	0.3972	0.10091	0.0001	0.843567	0.00061007
7	0.3274	0.18709	0.0001	0.623107	0.000114

Analysis of Mean Seasonal Precipitation

Two sets of regression equations were run for each region with the log of mean summertime (April - September) precipitation and the log of mean wintertime (October - March) precipitation as dependent variables. These variables will be called logsum and logwin respectively. Elev, qx, qy, ALI5, ALI10, ALI20, and ALI30 made up the independent variables. It was considered possible that by separating the year into seasons, the equations might estimate values more accurately. Table 4.24 shows the results of each seasonal test. No overall improvement in the r-square, the root mean square error, or the 95 percent confidence bands occurred when the seasonal precipitation values were used.

Table 4.9 Mean Seasonal Regression Results

* = Significant at .1 level

Dependant Variable	R Square	Root MSE	Elev	Qx	Qy	ALI5	ALI10	ALI20	ALI30
			Independent variables						
Region One									
Logsum	0.88	0.058	*						
Logwin	0.97	0.045	*				*	*	*
Region Two									
Logsum	0.57	0.124				*			
Logwin	0.85	0.110							
Region Three									
Logsum	0.91	0.053	*	*	*				
Logwin	0.85	0.110				*	*	*	
Region Four									
Logsum	0.74	0.086	*		*			*	*
Logwin	0.74	0.163	*	*	*			*	*
Region Five									
Logsum	0.80	0.135				*		*	*
Logwin	0.88	0.146		*		*		*	*
Region Six									
Logsum	0.79	0.069	*		*				
Logwin	0.75	0.087	*	*	*				*
Region Seven									
Logsum	0.55	0.121			*			*	
Logwin	0.67	0.180		*	*		*	*	*

Analysis of Mean Monthly Precipitation

Mean monthly precipitation values were used as dependent variables in another set of regression runs. The January through December mean monthly values in log form will be called logm1 - logm12. The number of runs needed to study all seven regions by month made it unfeasible, so Region One was used as a test case. Table 4.25 shows the results of the monthly analysis. No improvement in overall accuracy was found when each month was run separately. The explained variance ranged from 98 percent in January, to 81 percent in August. It was not surprising that the equations worked better during the winter months because of the organized, consistent nature of precipitation forming processes at work during the winter. The 95 percent confidence band for the entire year equation was plus or minus 2.22 inches. The average 95 percent confidence band from all of the twelve month equations combined was plus or minus 2.29 inches. No overall improvement in accuracy was brought about by the monthly equations.

Table 4.10 Mean Monthly Regression Results for Region One

* = Significant at the .1 level

Dependant Variable	R Square	Root MSE	Independent variables Elev	Qx	Qy	ALI5	ALI10	ALI20	ALI30
Logm1	0.98	0.040	*						
Logm2	0.97	0.015	*			*	*	*	
Logm3	0.93	0.070	*		*				
Logm4	0.89	0.065	*		*				
Logm5	0.86	0.065	*						
Logm6	0.84	0.066							
Logm7	0.84	0.068	*					*	
Logm8	0.81	0.074							
Logm9	0.89	0.061	*			*			
Logm10	0.87	0.070							
Logm11	0.92	0.061	*						*
Logm12	0.97	0.044	*				*	*	*

THE FINAL EQUATIONS

It was determined through examination of each of the sets of output, that the entire year averages should be used as the dependent variables. Although the r-square values were higher for some of the monthly and seasonal equations than those produced by the annual models, no improvements in the root mean square error terms were found. An advantage of using the mean annual precipitation models is that the error terms form definite confidence bands at the mean. When the estimates from two or more equations are summed, the resulting confidence band cannot be explicitly expressed.

In order to produce equations which were both accurate and easy to use, only one ALI variable was used in each final equation. This greatly reduce the number of measurements needed for each point to be estimated, without sacrificing much accuracy. Elev, qx, and qy were used in each equation regardless of their significance. These three variables were measured for each point in order to locate the point on the quadrangle, and to calculate the appropriate ALI measure.

Regression equations were generated for all seven regions using elev, qx, qy, and each of the ALI variables separately. The resulting root mean square error values were compared to determine which ALI measure should be used for each region. The ALI measures producing the lowest root mean square error for each region are listed below:

Region One - ALI20
Region Two - ALI10
Region Three - ALI20
Region Four - ALI20 and ALI30
Region Five - ALI20
Region Six - ALI30
Region Seven - ALI20

In Region Four, both ALI20 and ALI30 produced equal error terms. In this case, ALI20 was selected because it would be easier to measure for point estimation. It is not surprising that ALI20 worked the best in five of the seven regions. The only two regions in which another ALI variable worked better were Regions Two and Six. In these two regions, the landforms and precipitation forming processes are unlike any others found in the rest of the study area. Spreen (1947) stated that the exposure of each gage site to moisture inflow was used along with other physiographic variables by W.T. Wilson to explain the variance in precipitation in northwest Wyoming. This exposure was measured at a number of different radii around the stations in order to discern which radius was the most effective in explaining the variation in precipitation. Exposure worked the best when measured at a twenty mile radius. Tables 4.26 through 4.32 show the final parameter estimates and their resulting statistics.

Region One

The final point estimation model for Region One explained more of the variance in precipitation than any of the other region's models. This can be contributed primarily to the overwhelming effect of elevation on precipitation. Both elev and ALI20 were significant at the five percent level. All of the variables combined explained 94 percent of the variance in logMAP. In Tables 4.11-4.17, the upper and lower 95 percent CL refer to the upper and lower limits of the 95 percent confidence band when the estimated value is equal to the mean. As the estimates depart from the mean, the expected confidence bands increase.

Table 4.11 Final Regression Results for Region One

Dependent Variable: LOGMAP

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	4	0.78360	0.19590	84.988	0.0001
Error	21	0.04841	0.00231		
C Total	25	0.83200			

Root MSE	0.04801	R-square	0.9418
Dep Mean	1.60364	Adj R-sq	0.9307
C.V.	2.99385		

Dep Mean (inches) 40.15
Upper 95% C.L. (inches) 49.86
Lower 95% C.L. (inches) 32.33

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob >|T|
INTERCEP	1	1.413373	0.19270423	7.334	0.0001
ELEV	1	0.000081213	0.00001204	6.745	0.0001
QX	1	0.000465	0.00230995	0.201	0.8423
QY	1	-0.000834	0.00094373	-0.883	0.3870
ALI20	1	0.000026855	0.00000987	2.722	0.0128

Region Two

Region Two was by far the most problematic. The precipitation forming processes in this region are unlike those of any other region in the study area. The mean moisture inflow to the portion of the region east of the dump zone of the Bitterroot Mountains is from the east (Miller, 1972). The use of 261 degrees as the prevailing airmass movement bearing for the ALI measures is the least appropriate in this region. The strong inter-correlation between qx and elev also caused some confusing results. The final equation explained 69 percent of the variation in precipitation, while elev and qx were significant at the 10 percent level.

Table 4.12 Final Regression Results for Region Two

Dependent Variable: LOGMAP

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	4	1.06518	0.26629	12.355	0.0001
Error	22	0.47416	0.02155		
C Total	26	1.53934			

Root MSE	0.14681	R-square	0.6920
Dep Mean	1.26109	Adj R-sq	0.6360
C.V.	11.64140		

Dep Mean (inches) 18.27
Upper 95% C.L. (inches) 35.39
Lower 95% C.L. (inches) 9.40

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob >|T|
INTERCEP	1	1.793856	1.05299021	1.704	0.1025
ELEV	1	0.000115	0.00002705	4.242	0.0003
QX	1	-0.014919	0.00691986	-2.156	0.0423
QY	1	-0.000033725	0.00493499	-0.007	0.9946
ALI10	1	-0.000003212	0.00001972	-0.163	0.8721

Region Three

In Region Three, qx and ALI20 appeared to be most important in explaining the variance in logMAP. They were both significant at the five percent level. The final model explained 74 percent of the variance.

Table 4.13 Final Regression Results for Region Three

Dependent Variable: LOGMAP

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	4	0.20768	0.05192	7.010	0.0059
Error	10	0.07406	0.00741		
C Total	14	0.28174			

Root MSE	0.08606	R-square	0.7371
Dep Mean	1.35765	Adj R-sq	0.6320
C.V.	6.33881		

Dep Mean (inches) 22.79
Upper 95% C.L. (inches) 33.60
Lower 95% C.L. (inches) 15.45

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	0.631284	0.33163131	1.904	0.0861
ELEV	1	0.000020094	0.00003175	0.633	0.5410
QX	1	0.012730	0.00483632	2.632	0.0251
QY	1	0.001555	0.00175355	0.887	0.3959
ALI20	1	0.000056113	0.00002244	2.501	0.0314

Region Four

The most pleasing results came from Region Four. All of the variables were significant at the one percent level. It was expected that the complex terrain and precipitation regime in this region would cause some problems in the regression analysis, but the variables worked surprisingly well. Seventy-two percent of the variance in logMAP was explained by the model.

Table 4.14 Final Regression Results for Region Four

Dependent Variable: LOGMAP

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	4	1.77995	0.44499	31.782	0.0001
Error	49	0.68606	0.01400		
C Total	53	2.46601			

Root MSE	0.11833	R-square	0.7218
Dep Mean	1.46465	Adj R-sq	0.6991
C.V.	8.07885		

Dep Mean (inches) 29.15
Upper 95% C.L. (inches) 49.73
Lower 95% C.L. (inches) 17.09

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob >|T|
INTERCEP	1	1.274728	0.14291445	8.920	0.0001
ELEV	1	0.000070416	0.00001503	4.684	0.0001
QX	1	-0.011332	0.00186676	-6.070	0.0001
QY	1	0.005717	0.00116260	4.918	0.0001
ALI20	1	0.000044411	0.00001322	3.359	0.0015

Region Five

The lowest percentage of explained variance of any of the models was produced in Region Five (57 percent). This is probably due to the sparse network of gages in the area, and the large non-uniform land area of the region. Only elev was significant at the 10 percent level. The coefficient of variation in this area was considerably higher than that of the other regions. This is most likely because of an inconsistent source region for precipitation with major flow patterns alternating from south to west.

Table 4.15 Final Regression Results for Region Five

Dependent Variable: LOGMAP

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	4	0.45934	0.11484	2.982	0.0799
Error	9	0.34655	0.03851		
C Total	13	0.80589			

Root MSE	0.19623	R-square	0.5700
Dep Mean	1.29842	Adj R-sq	0.3789
C.V.	15.11289		

Dep Mean (inches) 19.88
Upper 95% C.L. (inches) 48.20
Lower 95% C.L. (inches) 8.20

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob >|T|
INTERCEP	1	-0.426817	0.94535228	-0.451	0.6623
ELEV	1	0.000275	0.00010122	2.721	0.0236
QX	1	-0.009066	0.00677147	-1.339	0.2134
QY	1	0.016594	0.00941504	1.762	0.1118
ALI20	1	-0.000103	0.00008717	-1.187	0.2657

Region Six

All of the variables were significant at the one percent level, while 75 percent of the variance was explained. Although precipitation forming processes in this area are different from those of other parts of the state (largely convectional), the model worked remarkably well. The region is by far the most homogeneous, and contains the least complex terrain of all the regions.

Table 4.16 Final Regression Results for Region Six

Dependent Variable: LOGMAP

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	4	0.55593	0.13898	30.415	0.0001
Error	41	0.18735	0.00457		
C Total	45	0.74329			

Root MSE	0.06760	R-square	0.7479
Dep Mean	1.09516	Adj R-sq	0.7233
C.V.	6.17253		

Dep Mean (inches) 12.45
Upper 95% C.L. (inches) 16.38
Lower 95% C.L. (inches) 9.46

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob >¦T¦
INTERCEP	1	0.720287	0.07377563	9.763	0.0001
ELEV	1	0.000070689	0.00001298	5.448	0.0001
QX	1	-0.002722	0.00067384	-4.040	0.0002
QY	1	0.004927	0.00119934	4.108	0.0002
ALI30	1	0.000044245	0.00000939	4.714	0.0001

Region Seven

All of the variables included in the final model for Region Seven were significant at the five percent level. Sixty-four percent of the variance was explained.

Table 4.17 Final Regression Results for Region Seven

Dependent Variable: LOGMAP

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	4	2.45882	0.61471	31.347	0.0001
Error	71	1.39229	0.01961		
C Total	75	3.85111			

Root MSE	0.14003	R-square	0.6385
Dep Mean	1.42699	Adj R-sq	0.6181
C.V.	9.81328		

Dep Mean (inches) 26.73
Upper 95% C.L. (inches) 50.29
Lower 95% C.L. (inches) 14.21

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	1.292302	0.20205125	6.396	0.0001
ELEV	1	0.000060562	0.00002113	2.866	0.0055
QX	1	-0.005188	0.00205723	-2.522	0.0139
QY	1	0.004329	0.00091986	4.706	0.0001
ALI20	1	0.000086382	0.00001737	4.974	0.0001

Summary

The goal of this section of the study was to derive regression equations to estimate precipitation values at points where no gages exist. These equations were to be both as accurate and as simple to use as possible. Through a series of exploratory regression runs, the most efficient variables were decided upon. Acceptable models were derived for each of the seven regions with the possible exception of Region Two. In Region Two, multicollinearity had some peculiar effect on the regression results as the variables were run in various combinations. The final equations should prove most useful in Region Four, where the topography and the precipitation pattern are the most complex.

The scope of this study covers the entire state, and roughly a fifty mile perimeter. The level of detail at which the precipitation pattern is being examined is appropriate for this scope. The equations derived in this section are not necessarily appropriate for the delineation of the precipitation pattern at a larger scale. In other words, to accurately estimate precipitation values for a study encompassing a small area, such as a small drainage basin, these equations should not be used. The one by two degree U.S.G.S. topographic maps used for all of the variable measurements have a contour interval of 200 feet and a scale of 1:250,000. None of the variables have a finer resolution than these maps, and consequently, neither do the estimation equations. The equations which were estimated in this section, are useful only in the region from which they were generated. The quantitative relationships between precipitation and the variables used here cannot be expected to hold true in other parts of the country.

V. INTERPOLATION PROCEDURES

Interpolation procedures for automated mapping can be grouped into a number of categories depending upon the scope of the study, spatial characteristics of the data, and the desired output. The two general classes of interpolation procedures are differentiated by the type of data for which they are intended. Point interpolation uses point data such as spot elevations or precipitation data. Areal interpolation procedures are appropriately used on areal data such as population densities or per capita income by state. Point methods are generally used for contour mapping (isometric), while areal methods are adapted to isopleth mapping (Lam, 1983).

The following discussion concentrates on point interpolation. Most computer contouring procedures use a two step process for generating contours from point data. First, an evenly spaced grid, or matrix, of data values is generated. The density of this grid will determine the resolution of the map produced. The intersections of grid lines define points which must have estimated values associated with them. Once the matrix of points is in place, the second step of the process consists of simply tracing the contour lines in and out of the appropriate squares defined by each set of four grid points. This can be done very quickly by the computer.

The main difference between various point interpolation procedures lies in the estimation of grid point values. Two maps produced from the same data set can have quite different appearances depending upon the characteristics of the grid algorithms such as the search shape, the search distance, the search orientation, and the spatial dependance, or distance decay function. The spatial dependance model is expressed by the point estimation algorithm, and is at the heart of the interpolation process.

SEARCH METHODS

Each grid point is assigned an estimated value based on its proximity to a number of neighboring data points. The search shape and distance control the number of points to be considered. The typical default shape is circular, with a user defined radius. When a grid point is to be estimated, all of the data points within the radius are used in the interpolation. When the grid procedure used all of the data points, the search method is called global. The global method is usually time consuming because of the large number of data points which must be considered. Often, the extra time spent in calculating the weights of distant points does not increase the accuracy of the estimate. As the distance between points increases, the relative importance, or weight of the points decreases. The search radius should be set to include only points which are close enough to have an effect on the estimate. The selective search is very important when contouring a variable which has a spatial dependance that breaks down rapidly with distance. When the search shape is other than circular, more points can be included in one direction than another. This can be useful when the user has some notion of the general orientation of the phenomenon to be contoured. For instance, the orientation of the ridges on the Sandpoint one by two degree quadrangle is primarily north-south. The precipitation pattern should roughly follow this orientation. When creating the grid for this quadrangle, an oval shaped search oriented north-south is most appropriate. When each grid point is estimated, the probability is greater for more points to be considered along this orientation than to the east or west. When no previous knowledge of the pattern is available, the circular search is most appropriate. The search can also be constrained by specifying the number of points to be used from each quadrant or octant around the point. This equalizes the number of points to be considered in any one direction, even if some quadrants contain numerous nearby points (Surfer Version 4 Reference Manual 1989; Davis, 1986).

POINT INTERPOLATION METHODS

Two point interpolation procedures were considered for this project. They are the inverse distance method, and the punctual, or simple Kriging method. This portion of the study is limited to these two methods, because they are believed to be the most appropriate of the readily available procedures.

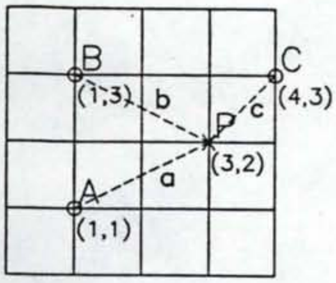
Inverse Distance

The inverse distance method of point estimation is one of the fastest and most straightforward ways of estimating points based on a number of neighboring known points. Each estimated point is derived as a weighted average of the neighboring data points. The weights, or contribution coefficient applied to each known point is a function of its distance to the point to be estimated. The most common distance function is the inverse of distance squared, although the power of the distance decay function can be varied. This method assumes a predefined spatial dependence between points as expressed by the power of the distance function. The general formula for the inverse distance model is:

$$P(x,y) = \frac{\sum_{i=1}^{n} Z_i / (d_i)^e}{\sum_{i=1}^{n} 1 / (d_i)^e}$$

where P(x,y) is the value of the point to be estimated, Z_i is the value of the neighboring point, d is the distance between the points Z_i and P(x,y), n is the number of neighboring points, and e is the distance weighting power (Surfer Version 4 Manual 1989; Monmonier, 1982).

As an example of this method, the point "P" (Figure 5.1) will be estimated using the inverse distance method with a distance weighting power of two. For simplicity, three points (A, B, and C) will be used. The distance between A, B, and C, and the point P are expressed as a, b, and c, respectively.



Z VALUES

A = 20
B = 25
C = 30

DISTANCES

a = 2.236
b = 2.236
c = 1.414

$$P = \frac{20 / (2.236)^2 + 25 /(2.236)^2 + 30 / (1.414)^2}{1 / (2.236)^2 + 1 / (2.236)^2 + 1 / (1.414)^2}$$

$$P = \frac{4.0 + 5.0 + 15.0}{.20 + .20 + .50}$$

$$P = \frac{24.0}{.90}$$

$$P = 26.7$$

Figure 5.1 Point P Estimated Using the Inverse Distance Method

The major disadvantage of this method is that the spatial relationship of the variable over distance is predefined. This may or may not be appropriate for the data set. Another limitation is that no estimate of the relative uncertainty (error term) can be calculated. One would assume that points which are closely surrounded by numerous neighbors would be more reliable than those estimated from only a few distant neighbors. A measure of this reliability would be useful, but is unavailable with this algorithm.

Two advantages of this method are its speed and its ease of use. This method does one set of calculations for each point. This one step process is easily encoded into an expedient computer algorithm. The user interface needed to run this procedure is limited. Once the search criteria and the grid density have been established, the user needs only to specify the distance weighting power.

Kriging

The Kriging process was developed by, and named for, D.G. Krige. It was originally intended for mine elevation, but is well suited to the production of a variety of contour maps (Davis, 1986; Lam, 1983). The Kriging procedure, like the inverse distance procedure, assigns a numeric weight to each known point in the search. These weights are then multiplied by the point values, to yield an estimate. Three attributes of the Kriging process set it apart from the other interpolation methods. First, the weights of all the points sum to one, and produce the minimum possible standard error of estimate. Second, the weights are derived from a model of the spatial dependence characteristics of the specific data set in use. Third, the standard error of estimate is explicitly expressed for each grid point. This standard error is also derived from the spatial dependence characteristics of the data set.

Many different variations exist on the Kriging algorithm. The one considered for this project is the punctual or simple Kriging algorithm. The underlying assumptions to the punctual Kriging procedures requires stationarity of the regionalized variable. A regionalized variable is one whose spatial characteristics are neither totally spatially deterministic nor completely random, but somewhere in between. Stationarity means that the surface portrayed by the data set has similar and somewhat continuous statistical characteristics.

Semivariograms

The spatial dependence of the sample is determined through the derivation of a semivariogram. This is a graph on which spatial variance versus lag distance is shown. The assumption is that when two measurements are made at the same location, the difference between the measured values is small. As the distance between the points increases, the difference in the measured values (spatial variance) also increases, until the points are far enough apart that there is no relationship between them. The semivariance for each lag distance can be calculated by the following formula:

$$\hat{\gamma}(h) = 1/2 * 1/N(h) \sum_{i=1}^{n} [X(\mu_i) - X(\mu_i + h)]^2$$

where γ is the semivariance, h is the lag distance, N is the number of lags at that lag distance, and x is the measured values. Figure 5.2 shows a set of known data points and the calculations needed to produce the semivariogram. The data points are all on the same line for illustration simplicity. The process for generating a semivariogram for randomly distributed points is similar to the one shown here. The nugget, or intercept, is the expected variation when sampling the same point (lag distance of zero). If the nugget is large, the sampling process is relatively inaccurate. The sill is the semivariance after which the points are no longer related, whereas the range is the usable portion of the semivariogram.

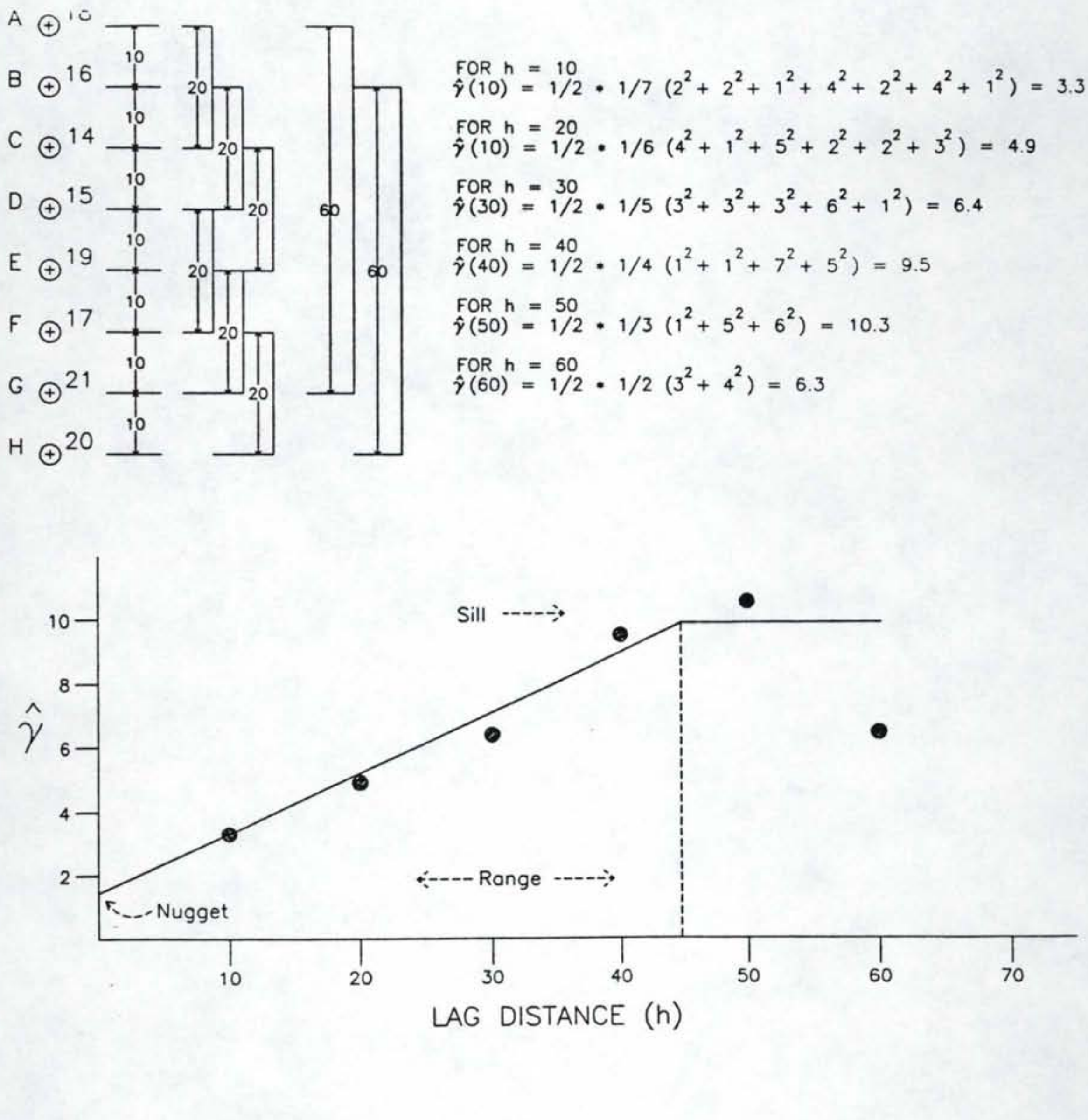


Figure 5.2 The Generation of a Sample Semiovariogram for a Linear Set of Data Points

has been derived, the semivariance for new lag distances to be estimated can be derived from either the semivariogram or an equation which represents it. The Kriging process uses the semivariance to assign the proper weights to known points when deriving an estimate for a new grid point.

Kriging Algorithm

The Kriging procedure uses a set of simultaneous equations to estimate each grid point. The number of equations (n+1) and the number of elements (n+1) in each equation is a function of the number of points in the search (n). The goals in determining the weights for each known point are: the weights must sum to one, and the standard error of estimation must be the minimum possible. The following is the set of simultaneous equations used for estimating a point based on three known points:

$$
\begin{aligned}
W_1\gamma(h_{11}) + W_2\gamma(h_{12}) + W_3\gamma(h_{13}) + \lambda &= \gamma(h_{1p}) \\
W_1\gamma(h_{12}) + W_2\gamma(h_{22}) + W_3\gamma(h_{23}) + \lambda &= \gamma(h_{2p}) \\
W_1\gamma(h_{13}) + W_2\gamma(h_{23}) + W_3\gamma(h_{33}) + \lambda &= \gamma(h_{3p}) \\
W_1 + W_2 + W_3 + 0 &= 1
\end{aligned}
$$

where W_1, W_2, and W_3 are the weights to be applied to known points one, two, and three respectively, $\gamma(h_{ij})$ is the semivariance over the distance h (between points i and j), and λ is the slack, or Lagrange multiplier.

The Lagrange multiplier is used as the fourth variable to insure that the estimate will have the minimum possible error.

The weights are then used in the following formula to arrive at the point estimate, $\hat{Y}_p$:

$$\hat{Y}_p = W_1 Y_1 + W_2 Y_2 + W_3 Y_3$$

The standard error of estimation, S_ϵ^2 , is calculated as follows (Davis, 1986):

$$S_\epsilon^2 = W_1 \gamma (h_{1p}) + W_2 \gamma (h_{2p}) + W_3 \gamma (h_{3p}) + \lambda$$

The major advantages to the punctual Kriging process over the other interpolation methods are threefold. First, the spatial dependence of the data set in use determines the weighting procedure. Second, the weights are calculated to generate the estimate with the lowest possible error. This error will be distributed randomly as long as there is no overall trend in the database. Third, the relative confidence level can be automatically derived from the standard error output. Contour lines showing both the distribution of the data and the distribution of the error can be produced from one run of the Kriging process.

The Kriging process does have some disadvantages. The most obvious limitation is the time and computation cost involved in generating a grid of point estimates. The second disadvantage is that it requires a greater degree of user knowledge and input than needed by other interpolation methods. In many cases, the function used to simulate the semivariogram must be fit to the points by the user (Englund and Sparks, 1988). A third disadvantage is that if the original data set is not appropriate for the method chosen(the underlying assumptions are violated). The resulting error terms may be misleading.

While there are advantages and limitations to the Kriging process, the use of Kriging has grown dramatically in the past several years. The process has spanned disciplines, and is a useful tool in the automated generation of numerous types of contour lines.

A Comparative Example

Two maps of the Spokane one by two degree quadrangle were produced for comparison. The first map was made with the inverse distance squared method (Figure 5.3), while the second map utilized the punctual Kriging procedure (Figure 5.4).

While the differences between these two maps are slight, the Kriging procedure produces a more continuous pattern, while the inverse distance method formed more closed isohyets. The map derived by the Kriging process more closely resembles the precipitation pattern expected for this area based upon experience.

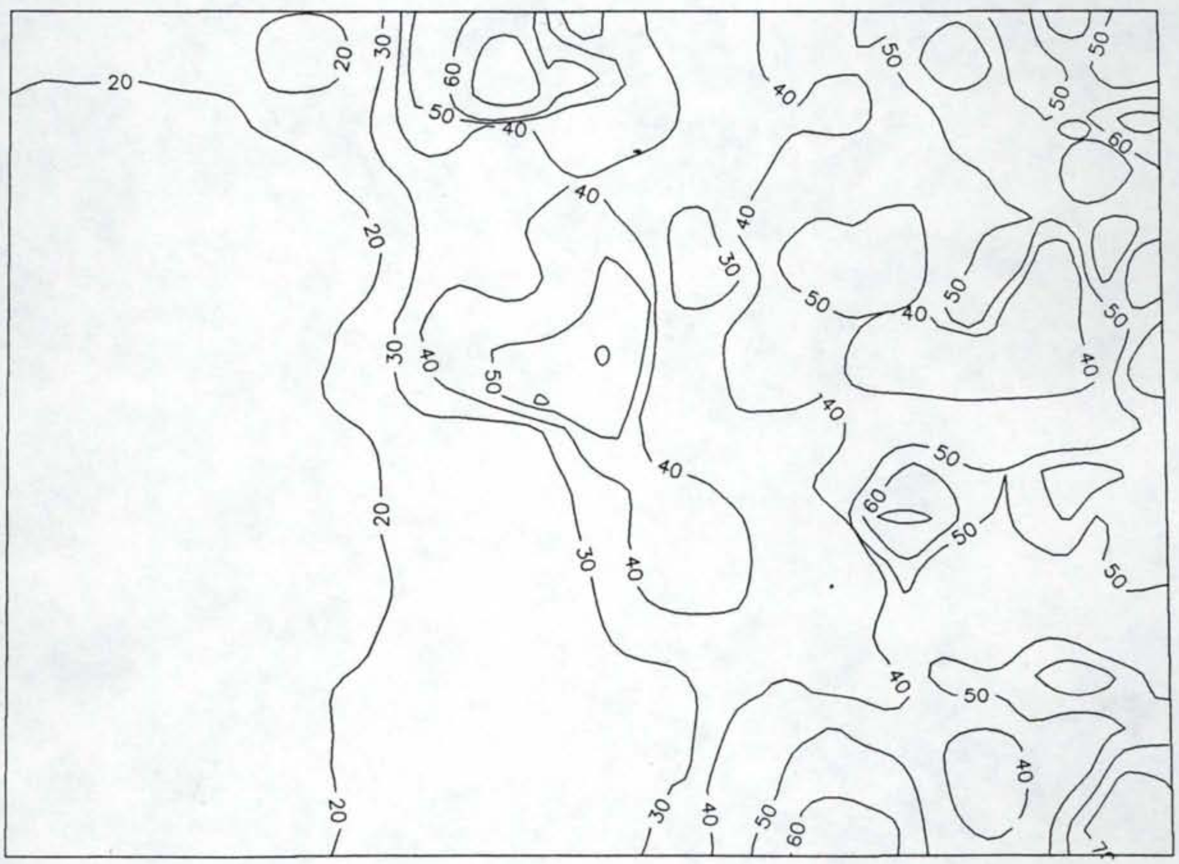


Figure 5.3 Spokane One by Two Degree Quadrangle Precipitation Pattern Using the Inverse Distance Squared Method

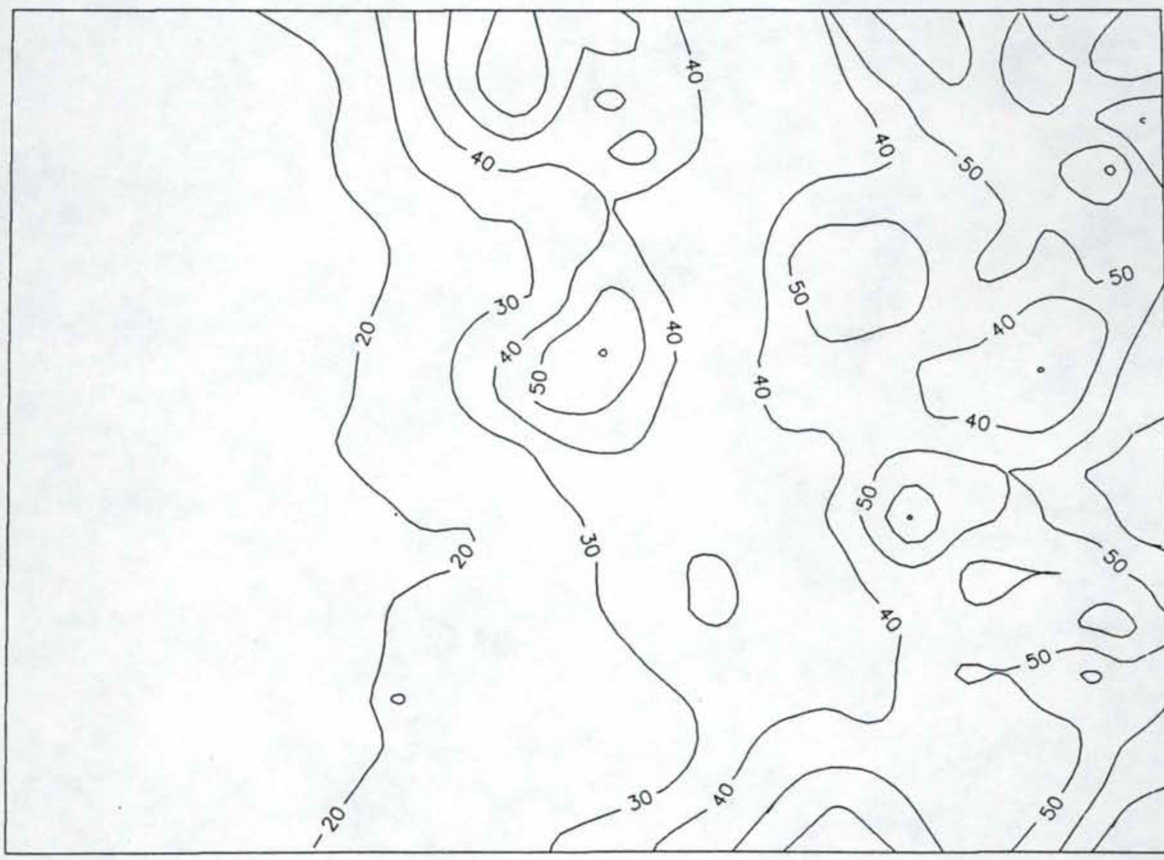


Figure 5.4 Spokane One by Two Degree Quadrangle Precipiation Pattern Using the Punctual Kriging Method

VI. CONCLUSIONS AND DISCUSSION

DATA ESTIMATIONS

Although some of the individual monthly estimates could have been slightly more accurately estimated if they had been derived from regression equations, the method used was acceptable, considering time constraints and the accuracy of the final map output. Given that twenty-five years of record were used, the effect of estimation error on the final mean annual precipitation values is negligible. The record estimation procedures used for the NWS records are believed to be at least as accurate as those applied to the SNOTEL data. If a project were undertaken to look at individual yearly data, a more accurate missing record estimation procedure may be required. However, for the study at hand, the data estimation procedures were more than adequate.

REGRESSION ANALYSIS

Regression equations were used to derive mean annual precipitation values at sites throughout the study area where no gages exist. Acceptable regression equations were derived for five of the seven regions studied.

In Region Two, multicollinearity caused some confusing results. We cannot assume that values estimated by this final equation are reliable, because of the instability imposed by the cross-correlations between the independent variables.

In Region Five, the sign on the ALI20 parameter estimate is negative. This is not what was expected. This may have been caused by the small number of stations in this area which spread out over a large, somewhat heterogeneous area. The estimates derived from the final equation for region five are also believed to be unreliable.

Where the point estimation equations are needed the most (region four), they are most reliable. No one variable in region four exerts a disproportionate control over precipitation. All of the variables used are highly significant.

The ninety-five percent confidence bands are given for each of the equations in Chapter Four. These confidence bands are only valid when the estimated value is equal to the mean. As the estimated values depart from the mean, the expected confidence bands widen considerably. The ninety-five percent confidence band, should be considered for comparison only. When the final equations were tested on a number of points, and plotted on the topographic maps, the estimated values seemed reasonable. When these estimated values were checked with graphs of precipitation versus elevation, all of the estimates were in a reasonably expected range.

When points were estimated at the bottom of some narrow canyons, the values were unreasonably low. This makes sense, because these narrow river canyons are too small to have a large effect on the airflow, but the elevations at their bottoms are low. The estimates generated are appropriate for relatively broad valleys, but not necessarily for narrow canyons. Caution must be used when estimating points in such cases.

The equations generated by the regression analysis cannot be used at another scale or in areas other than the ones they were intended for. The equations simply quantify the relationship between precipitation observations and the physiographic parameters at the observation sites. Localized changes in landform alter this relationship dramatically.

INTERPOLATION

Three interpolation procedures were considered for the production of isohyets using both the actual and estimated data points. The procedures considered were:

1) The Inverse Distance Method of the Golden Software, Inc. Surfer Package

2) The Kriging Method of Surfer

3) The Punctual Kriging Method of the Environmental Protection Agency Geo-Ease Package

Surfer

The Surfer package has somewhat limited search criteria options. The search shape is controlled by the relative grid spacing in the x and y directions. If the grid has a uniform density in both directions, the search shape is circular. This constraint affects the operation of both of the interpolation procedures of the Surfer package. The Kriging algorithm is also limited in Surfer. It assumes a linear semivariogram with no nugget. This may or may not be appropriate for production of isohyets.

The Surfer package excels in the actual plotting of isolines and production of the final computer maps. This is facilitated easily with high quality pen plotter output. All of the maps can be produced repetitively with the same layout.

Geo-Ease

The Kriging process of the Geo-Ease package is flexible and complete. The search criteria can be easily altered. The search shape can be altered to an oval of any width and height. The axis of the search can be oriented to any angle. This allows flexibility in tailoring the search to the known landform patterns. The semivariogram is produced from the pairs of points in the data set, which are selected by the search criteria. The function describing the semivariogram, the nugget, and the sill, are all user defined in an interactive (trial and error) mode. When the user is pleased with the fit of the function to the semivariogram, Kriging takes place on either groups of four points or individual points, as specified by the user. The functions available for the definition of the semivariogram are linear, spherical, Gaussian, and experimental. The Geo-Ease package is limited in the production of the final isolines and map plotting.

FUTURE STUDY AND ALTERNATIVES

The methods used in this study were appropriate, considering the availability of hardware, software, and data. If a Geographic Information System (GIS) with digital terrain data were available at the onset of this project, the methodology would have been much different. Three aspects of the project could be changed in order to improve the mapping of precipitation with the use of a GIS. They are the measurement of the ALI variable, automated use of the regression equations to produce the final grid of estimates, and the use of co-Kriging. The last two of these options would result in the elimination of one of the two major steps (regression analysis, or interpolation).

The measurement of four ALI variables to each station and one ALI variable to each estimated point was the most time consuming part of the project. With the use of a GIS and digital terrain data, this task could be accomplished automatically. This would allow for experimentation with various combinations of ALI measures. One possibility would be a cumulative sum of the relative elevations taken every mile up to twenty miles.

With the use of a GIS, each point in a grid could be estimated directly using the regression equations, thus eliminating the interpolation step. However, it is uncertain whether this would produce a more realistic map. In areas where the regression equations estimates are unreliable, the map output would probably be less accurate than the one produced by the current method.

Another possible alternative would be to use co-Kriging. This is somewhat similar to the punctual Kriging procedure, but the data input is not constrained to one variable. Co-

Kriging accepts a number of independent variables, which are used to adjust the weighting of the dependent variable. A series of correlations are formulated, and used. In this case, the precipitation data would be input along with the elevation data, and the ALI data for every point on the desired grid. First co-Kriging would quantify the relationship between elevation, ALI, and precipitation. Then, the weighting process would proceed as in punctual Kriging, but taking into account and adjusting for the independent variables. This is probably the best method for this project. The maps produced would be finely tuned to the changes in terrain.

The steps described in this thesis lead to the automated production of isohyets. Given the available hardware, software, and data, the methods are appropriate. The maps produced, while knowingly not completely accurate, are by far a more realistic depiction of the precipitation pattern, if not the amounts, than available through previous methods. The confidence level on the final output can not be exactly quantified because of the combined sources of possible error. These maps should be a good basis upon which to begin the process of producing a set of final maps suitable for use in a GIS.

REFERENCES CITED

Afifi, A.A., and V. Clark, 1984. Computer-Aided Multivariate Analysis. Lifetime Learning Publications, Belmont, California.

Arnold, D.L., 1989. "Specification of Winter Precipitation in Idaho from the Mean Monthly 700-mb Circulation."Thesis, University of Idaho, Moscow.

Bleasdale, A., and Y.K. Chan, 1972. "Orographic Influences on the Distribution of Precipitation." World Meterological Organization: Symposium on Distribution of Precipitation in Mountainous Areas, Norway, 322-333.

Crow, L.W., 1982. "Report on Methods of Estimating Precipitation in Mountainous Areas in Colorado." Colorado State University, Fort Collins, Colorado.

Davis, J.C., 1986. Statistics and Data Analysis in Geology (2nd ed.). John Wiley & Sons, New York.

de Montmollin, F.A., R.J. Oliver, R.G. Simard and F. Z. Wahlen. 1980. "Evalution of a Precipitation Map Using a Smoothed Elevation Precipitation Relationship and Optimal Estimates (Kriging)." Nordic Hydrology, 11:113-120.

Dingman, S.L., 1981. "Elevation. A Major Influence on the Hydrology of New Hampshire and Vermont, USA." Hydrological Sciences Bulletin, 26:399-413.

Dingman, S.L., et al., 1988. "Application of Kriging to Estimating Mean Annual Precipitation in a Region of Orographic Influence." Water Resources Bulletin 24:329-339.

Doesken, N.J., and G.L. Schaefer, 1987. "The Contribution of SNOTEL Precipitation Measurements to Climate Analysis, Monitoring, and Research in Colorado." Western Snow Conference, Vancouver, British Columbia.

Doesken, N.J., T.B. McKee and B.D. Richter 1984. Analysis of Colorado Average Mean Annual Precipitation for the 1951-1980 Period. Atmospheric Sciences Dept., Climatology Report 84-1, Colorado State University, Fort Collins, Colorado.

Englund, E., and A. Sparks, 1988. Geo-Eas (Geostatistical Environmental Assessment Software) User's Guide. U.S. Environmental Protection Agency: Las Vegas, Nevada.

Farnes, P.E., 1978. Development and Use of Mountain Precipitation Map. U.S.D.A.-Soil Conservation Service, Bozeman, Montana.

Foufoula-Georgiou, E., 1983. "Estimation of Missing Values in Monthly Rainfall Series." Proceedings of Stormwater and Water Quality Model User Group Meeting, Gainesville, Florida, 177-190.

Giambelluca, T.W., and M.A. Nullet, 1985. Ridge Regression to Adjust Rainfall Normals to a Common Base Period. University of Hawaii at Monoa, Honolulu, Hawaii.

Hanson, C.L., 1982. "Distribution and Stochastic Generation of Annual and Monthly Precipitation on a Mountainous Watershed in Southwest Idaho." Water Resources Bulletin 18:875-883.

Houghton, J.G., 1979. "A Model for Orographic Precipitation in the North-Central Great Basin." Monthly Weather Review 107:1462-1475.

Hutchinson, M.F., and R.J. Bischof, 1983. "A New Method for Estimating the Spatial Distribution of Mean Seasonal and Annual Rainfall Applied to Hunter Valley, New South Wales." Australian Meteorlogical Magazine. 31:179-184.

Lam, N.S.-N., 1983. "Spatial Interpolation Methods: A Review." The American Cartographer. 10:129-149.

Miller, J.F., 1972. "Physiographically Adjusted Precipitation-Frequency Maps." World Meteorological Organization: Symposium on Distribution of Precipitation in Mountainous Areas, Norway, 264-277.

Molnau, Myron, Walter J. Rawls, David L. Curtis and C.C. Warnick , 1980. "Gauge Density and Location for Estimating Mean Annual Precipitation in Mountainous Areas." Water Resources Bulletin. 16:428-432.

Monmonier, M.S., 1982. Computer-Assisted Cartography. Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Palmer, P., 4 August 1988. Interview, Boise, Idaho.

Palmer, P., 23 March 1989. Interview, Moscow, Idaho.

Peck, E.L., 1972. "Relations of Orographic Winter Precipitation Patterns to Meterological Parameters." World Meterological Organization: Symposium on Distribution of Precipitation in Mountainous Areas, Norway, 234-242.

Peck, E.L., and J.C. Schaake, 1987. "Network Design for Water Supply Forecasting in the West." Twenty-Third Annual Conference & Symposium, American Water Resources Association, Salt Lake City, Utah.

Peck, E.L., and M.J. Brown, 1962. "An Approach to the Development of Isohyetal Maps for Mountainous Areas." Journal of Geophysical Research. 67:681-694.

Ryden, B.E., 1972. "On the Problem of Vertical Distribution of Precipitation, Especially in Areas with Great Height Differences." World Meterological Organization: Symposium on Distribution of Precipitation in Mountainous Areas, Norway, 362-372.

Shearman, R.J., and P.M. Salter, 1975. "An Objective Rainfall Interpolation and Mapping Technique. "Hydrological Sciences Bulletin. 31:179-184.

Solow, A.R., and S.M. Gorelick, 1986. "Estimating Monthly Streamflow Values by Co-Kriging." Mathematical Geology. 18:785-809.

Spreen, W.C., 1947. "A Determination of the Effect of Topography Upon Precipitation." Transactions, American Geophysical Union 28:285-290.

Storr, D., and H.L. Ferguson, 1972. "The Distribution of Precipitation in Some Mountainous Canadian Watersheds." World Meterological Organization: Symposium on Distribution of Precipitation in Mountainous Areas, Norway, 243-263.

Surfer Version 4 Reference Manual, 1989. Golden, Colorado.

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

U.S.D.A., 1977. "Average Annual Precipitation Montana." Soil Conservation Service, Bozeman, Montana.

Vuglinski, V.S., 1972. "Methods for the Study of Laws for the Distribution of Precipitation in Medium-High Mountains (Illustrated by the Vitim River Basin)." World Meterological Organization: Symposium on Distribution of Precipitation in Mountainous Areas, Norway, 212-221.

APPENDIX A

DISCRIMINATE ANALYSIS RESULTS

Appendix A.1 Seven Class Discriminate Analysis Results

Number of Observations and Percent Classified into Regions

	1	2	3	4	5	6	7	Total
1	26	0	0	0	0	0	0	26
	100.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00
2	8	17	0	3	0	0	0	28
	28.57	60.71	0.00	10.71	0.00	0.00	0.00	100.00
3	2	0	13	0	0	0	0	15
	13.33	0.00	86.67	0.00	0.00	0.00	0.00	100.00
4	2	4	4	41	0	2	2	55
	3.64	7.27	7.27	74.55	0.00	3.64	3.64	100.00
5	0	0	1	0	13	1	0	15
	0.00	0.00	6.67	0.00	86.67	6.67	0.00	100.00
6	0	0	3	2	1	34	5	45
	0.00	0.00	6.67	4.44	2.22	75.56	11.11	100.00
7	0	0	0	1	1	1	71	74
	0.00	0.00	0.00	1.35	1.35	1.35	95.95	100.00
Total	38	21	21	47	15	38	78	258
Percent	14.7	8.1	8.1	18.2	5.8	14.7	30.2	100
Priors	0.143	0.143	0.143	0.143	0.143	0.143	0.143	

Error Count Estimates for Regions

	1	2	3	4	5	6	7
Rate	0.0000	0.3929	0.1333	0.2545	0.1333	0.2444	0.0405
Priors	0.1429	0.1429	0.1429	0.1429	0.1429	0.1429	0.1429

Appendix A.2 Two Class Discriminate Results for Regions 1 and 2

	1	2
1	25 96.15%	1 3.85%
2	1 3.75%	27 96.43%

Misclassified Stations	Probability
167264	0.9107
30MX12	0.6826

Appendix A.3 Two Class Discriminate Results for Regions 1 and 3

	1	3
1	26 100%	0 0
3	1 6.67%	14 93.33%

Misclassified Stations	Probability
164793	0.8761

Appendix A.4 Two Class Discriminate Results for Regions 1 and 4

	1	4
1	26 100%	0 0%
4	3 5.45%	52 94.55%

Misclassified Stations	Probability
163143	0.9026
1614C04	0.6382
30MV12	0.6593

Appendix A.5 Two Class Discriminate Results for Regions 2 and 4

	2	4
2	27 96.43%	1 3.57%
4	5 9.09%	50 90.91%

Misclassified Stations	Probability
163554	0.7350
165685	0.5787
1614C04	0.72624
30MS23	0.7606
30MR22	0.8288
30MV12	0.6351

Appendix A.6 Two Class Discriminate Results for Regions 2 and 7

	2	7
2	27 96.43%	1 3.57%
7	1 1.37%	72 98.63%

Misclassified Stations	Probability
1614G02	0.7599
30MI24	0.8914

Appendix A.7 Two Class Discriminate Results for Regions 3 and 4

	3	4
3	15 100%	0 0%
4	7 12.73%	48 87.27%

Misclassified Stations	Probability
160448	0.6716
161514	0.6220
162875	0.5220
163143	0.7476
163448	0.5081
165685	0.6275
165708	0.6300

Appendix A.8 Two Class Discriminate Results for Regions 3 and 5

	3	5
3	15 100%	0 0%
5	0 0%	15 100%

Appendix A.9 Two Class Discriminate Results for Regions 3 and 6

	3	6
3	15 100%	0 0%
6	0 0%	45 100%

Appendix A.10 Two Class Discriminate Results for Regions 4 and 6

	4	6
4	51 92.73%	4 7.27%
6	1 2.22%	44 97.78%

Misclassified Stations	Probability
160448	0.9684
1613F03	0.9956
163448	0.7036
1613E27	0.6228
165685	0.6855

Appendix A.11 Two Class Discriminate Results for Regions 4 and 7

	4	7
4	53 96.36%	2 3.64%
7	2 2.74%	71 97.26%

Misclassified Stations	Probability
1611E37	0.9823
1613G01	0.6583
1614G02	0.9881
30MI03	0.9892

Appendix A.12 Two Class Discriminate Results for Regions 5 and 6

	5	6
5	13 86.67%	2 13.3%
6	1 2.22%	44 97.78%

Misclassified Stations	Probability
164295	0.5616
166542	0.5287
418797	0.9170

Appendix A.13 Two Class Discriminate Results for Regions 5 and 7

	5	7
5	15 100%	0 0%
7	1 1.37%	72 98.63%

Misclassified Stations	Probability
1614G02	0.8879

Appendix A.14 Two Class Discriminate Results for Regions 6 and 7

	6	7
6	40 88.89%	5 11.11%
7	2 2.74%	71 97.26%

Misclassified Stations	Probability
160470	0.6296
1613F03	0.8356
162676	0.7602
164456	0.7723
166542	0.6321
168786	0.5624

APPENDIX B

SIMPLE CORRELATION MATRICES

Appendix B.1 Simple Correlation Matrix for Region 1

Correlation

CORR	ELEV	QX	QY	ALI5
ELEV	1.0000	0.4450	-0.0741	0.7728
QX	0.4450	1.0000	-0.5635	0.2230
QY	-0.0741	-0.5635	1.0000	-0.1309
ALI5	0.7728	0.2230	-0.1309	1.0000
ALI10	0.6284	0.0132	-0.2795	0.6263
ALI20	0.8051	0.2147	-0.1021	0.6461
ALI30	0.8072	0.1646	-0.0654	0.5755
LOGMAP	0.9552	0.4324	-0.1508	0.7660

Correlation

CORR	ALI10	ALI20	ALI30	LOGMAP
ELEV	0.6284	0.8051	0.8072	0.9552
QX	0.0132	0.2147	0.1646	0.4324
QY	-0.2795	-0.1021	-0.0654	-0.1508
ALI5	0.6263	0.6461	0.5755	0.7660
ALI10	1.0000	0.5993	0.6235	0.5970
ALI20	0.5993	1.0000	0.8752	0.8619
ALI30	0.6235	0.8752	1.0000	0.8544
LOGMAP	0.5970	0.8619	0.8544	1.0000

Appendix B.2 Simple Correlation Matrix for Region 2

Correlation

CORR	ELEV	QX	QY	ALI5
ELEV	1.0000	0.6672	-0.5750	0.5000
QX	0.6672	1.0000	-0.9624	0.0202
QY	-0.5750	-0.9624	1.0000	0.0687
ALI5	0.5000	0.0202	0.0687	1.0000
ALI10	0.6219	0.3159	-0.1988	0.5022
ALI20	0.5847	-0.1302	0.2421	0.7131
ALI30	0.6225	0.0085	0.0771	0.4130
LOGMAP	0.0737	-0.5678	0.5996	0.3740

Correlation

CORR	ALI10	ALI20	ALI30	LOGMAP
ELEV	0.6219	0.5847	0.6225	0.0737
QX	0.3159	-0.1302	0.0085	-0.5678
QY	-0.1988	0.2421	0.0771	0.5996
ALI5	0.5022	0.7131	0.4130	0.3740
ALI10	1.0000	0.5657	0.4155	0.1404
ALI20	0.5657	1.0000	0.6686	0.6117
ALI30	0.4155	0.6686	1.0000	0.4754
LOGMAP	0.1404	0.6117	0.4754	1.0000

Appendix B.3 Simple Correlation Matrix For Region 3

Correlation

CORR	ELEV	QX	QY	ALI5
ELEV	1.0000	-0.0982	-0.4396	0.5860
QX	-0.0982	1.0000	-0.1894	-0.0969
QY	-0.4396	-0.1894	1.0000	0.0522
ALI5	0.5860	-0.0969	0.0522	1.0000
ALI10	0.5553	-0.1311	0.1178	0.8864
ALI20	0.8098	-0.2014	-0.2524	0.6362
ALI30	0.7444	-0.3175	-0.1481	0.5397
LOGMAP	0.6530	0.2558	-0.1824	0.6214

Correlation

CORR	ALI10	ALI20	ALI30	LOGMAP
ELEV	0.5553	0.8098	0.7444	0.6530
QX	-0.1311	-0.2014	-0.3175	0.2558
QY	0.1178	-0.2524	-0.1481	-0.1824
ALI5	0.8864	0.6362	0.5397	0.6214
ALI10	1.0000	0.7729	0.5790	0.5960
ALI20	0.7729	1.0000	0.7272	0.7373
ALI30	0.5790	0.7272	1.0000	0.3519
LOGMAP	0.5960	0.7373	0.3519	1.0000

Appendix B.4 Simple Correlation Matrix for Region 4

Correlation

CORR	ELEV	QX	QY	ALI5
ELEV	1.0000	0.5215	-0.2025	0.4849
QX	0.5215	1.0000	-0.0783	0.0379
QY	-0.2025	-0.0783	1.0000	-0.0116
ALI5	0.4849	0.0379	-0.0116	1.0000
ALI10	0.3728	-0.0195	-0.0012	0.6317
ALI20	0.4507	-0.1251	-0.0211	0.3263
ALI30	0.3764	-0.1916	-0.0632	0.3690
LOGMAP	0.2848	-0.4020	0.3136	0.3967

Correlation

CORR	ALI10	ALI20	ALI30	LOGMAP
ELEV	0.3728	0.4507	0.3764	0.2848
QX	-0.0195	-0.1251	-0.1916	-0.4020
QY	-0.0012	-0.0211	-0.0632	0.3136
ALI5	0.6317	0.3263	0.3690	0.3967
ALI10	1.0000	0.4140	0.4436	0.4110
ALI20	0.4140	1.0000	0.5456	0.6328
ALI30	0.4436	0.5456	1.0000	0.6431
LOGMAP	0.4110	0.6328	0.6431	1.0000

Appendix B.5 Simple Correlation Marix for Region 5

Correlation

CORR	ELEV	QX	QY	ALI5
ELEV	1.0000	0.3343	-0.9063	0.6098
QX	0.3343	1.0000	-0.4963	0.0929
QY	-0.9063	-0.4963	1.0000	-0.3558
ALI5	0.6098	0.0929	-0.3558	1.0000
ALI10	0.6252	0.0595	-0.3567	0.9005
ALI20	0.8513	-0.0068	-0.7037	0.6630
ALI30	0.8700	-0.0502	-0.6498	0.6454
LOGMAP	0.4285	-0.2559	-0.1633	0.6738

Correlation

CORR	ALI10	ALI20	ALI30	LOGMAP
ELEV	0.6252	0.8513	0.8700	0.4285
QX	0.0595	-0.0068	-0.0502	-0.2559
QY	-0.3567	-0.7037	-0.6498	-0.1633
ALI5	0.9005	0.6630	0.6454	0.6738
ALI10	1.0000	0.6135	0.6757	0.6983
ALI20	0.6135	1.0000	0.8984	0.3845
ALI30	0.6757	0.8984	1.0000	0.6267
LOGMAP	0.6983	0.3845	0.6267	1.0000

Appendix B.6 Simple Correlation Matrix for Region 6

Correlation

CORR	ELEV	QX	QY	ALI5
ELEV	1.0000	0.7444	0.2012	-0.0867
QX	0.7444	1.0000	0.1696	-0.0864
QY	0.2012	0.1696	1.0000	-0.0988
ALI5	-0.0867	-0.0864	-0.0988	1.0000
ALI10	0.0635	0.1118	-0.0576	0.7934
ALI20	0.3811	0.1008	-0.0201	0.3079
ALI30	0.4446	0.2072	-0.1173	0.3324
LOGMAP	0.6302	0.2059	0.3523	0.0490

Correlation

CORR	ALI10	ALI20	ALI30	LOGMAP
ELEV	0.0635	0.3811	0.4446	0.6302
QX	0.1118	0.1008	0.2072	0.2059
QY	-0.0576	-0.0201	-0.1173	0.3523
ALI5	0.7934	0.3079	0.3324	0.0490
ALI10	1.0000	0.4045	0.4333	0.1924
ALI20	0.4045	1.0000	0.6922	0.5586
ALI30	0.4333	0.6922	1.0000	0.6184
LOGMAP	0.1924	0.5586	0.6184	1.0000

Appendix B.7 Simple Correlation Matrix for Region 7

Correlation

CORR	ELEV	QX	QY	ALI5
ELEV	1.0000	0.3103	0.1957	0.4494
QX	0.3103	1.0000	0.1674	-0.1650
QY	0.1957	0.1674	1.0000	-0.3872
ALI5	0.4494	-0.1650	-0.3872	1.0000
ALI10	0.3327	-0.3068	-0.2007	0.3964
ALI20	0.4629	-0.2610	-0.3795	0.6209
ALI30	0.4743	-0.0703	-0.3888	0.5262
LOGMAP	0.5722	-0.1997	0.2293	0.4013

Correlation

CORR	ALI10	ALI20	ALI30	LOGMAP
ELEV	0.3327	0.4629	0.4743	0.5722
QX	-0.3068	-0.2610	-0.0703	-0.1997
QY	-0.2007	-0.3795	-0.3888	0.2293
ALI5	0.3964	0.6209	0.5262	0.4013
ALI10	1.0000	0.4708	0.3699	0.4536
ALI20	0.4708	1.0000	0.7879	0.5909
ALI30	0.3699	0.7879	1.0000	0.5288
LOGMAP	0.4536	0.5909	0.5288	1.0000