NASA-CR-170457

Combining Remotely Sensed and Other Measurements for Hydrologic Areal Averages

(E83-10204) COMBINING REMOTELY SENSED AND OTHER MEASUREMENTS FOR HYDROLOGIC AREAL AVERAGES Interim Report, May - Oct. 1982 (Hydex Corp.) 164 p HC A08/MF A01 CSCL 08H

N83-21448

Unclas
G3/43 00204

OCTOBER 31, 1982

National Aeronautics and Space Administration

Goddard Space Flight Center
Greenbelt, Maryland 20771

COMBINING REMOTELY SENSED AND OTHER MEASUREMENTS FOR HYDROLOGIC AREAL AVERAGES

OCTOBER 31, 1982

INTERIM REPORT FOR PERIOD MAY - OCTOBER 1982

Prepared by:

HYDEX CORPORATION
SUITE 502
11150 MAIN STREET
FAIRFAX, VIRGINIA 22030

E. R. Johnson, E. L. Peck and T. N. Keefer

Prepared for:

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
GODDARD SPACE FLIGHT CENTER
GREENBELT ROAD
GREENBELT, MARYLAND 20771

PREFACE

This report is a part of a three-phase study by the Hydex Corporation for the National Aeronautics and Space Administration (NASA). The study is funded under the Conservation and Pollution section of the AgRISTARS program, a joint program for Agricultural and Resources Inventory Surveys Through Aerospace Remote Sensing.

The objective of the Hydex study is to develop and test technology for improving the usefulness of hydrologic modeling by incorporating remotely sensed measurements in operational procedures. The short-term objective of this phase of the study is to develop methods for using remotely sensed measurements with hydrologic models to ensure that simulated results are consistent with observed conditions.

This interim report describe a method (correlation area) for combining remotely sensed measurements with standard measurements to obtain improved areal average values of hydrologic variables.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF ILLUSTRATIONS

IN APPENDIXES

LIST OF ILLUSTRATIONS (Continued)

IN APPENDIXES

LIST OF TABLES

IN APPENDIXES

CHAPTER 1

INTRODUCTION

1. STUDY BACKGROUND

The study described in this report is part of continuing, long-term effort to maximize use of remotely sensed hydrologic data. Specifically, the report addresses the data weighting aspect of incorporating remotely sensed information in hydrologic models.

For several reasons, remotely sensed information has not been of much value for improving the ability to model the land phase of the hydrologic cycle. One reason is the dissimilarity in the time and space averages as envisioned by the hydrologic model, as exist in the real world, and as measured by remote sensing systems. A second reason is the problem of discriminating one type of measurement from another. One type of sensor usually responds to several variables at one time.

Many remote sensing techniques provide indirect measurement of land characteristics, vegetative cover, and the states of water in the hydrologic cycle. Such measurements should provide valuable information for improving the ability to model the hydrologic cycle. To date, however, this use of remote sensing techniques has been of limited value. In fact, federal agencies responsible for forecasting the flow of rivers and predicting water supplies are not using remote sensing techniques to provide a primary data base in their operational hydrologic forecasting programs.

The National Aeronautics and Space Administration (NASA) has played a strong role in working with other federal agencies to promote the use of remote sensing in hydrologic models. NASA and four other federal agencies cooperate in the AgRISTARS (Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing) program. The AgRISTARS program is directed toward developing the technology and testing the capability to use remotely sensed data in more economical ways in seven agriculturally related groups, one of which is conservation and pollution.

In the conservation and pollution group, three tasks have been defined:

TASK 1. Conservation Inventory

TASK 2. Water Resources Management

TASK 3. Snowpack Assessment

As part of the Task 2 (Water Resources Management) program, NASA contracted (No. NAS5-26446) with the Hydex Corporation for "Hydrological Modeling Survey Studies." The objective was to determine the suitability of present and planned remote sensing capabilities for commonly used hydrologic models.

The contract resulted in two major reports. In an interim report, "Review of Hydrologic Models for Evaluating Use of Remote Sensing Capabilities" (NASA Contractor Report CR 166674 dated 31 March 1981), Hydex presented information on the structure,

parameters, states, and required inputs of seven hydrologic models. The seven models studied were:

- Antecedent Precipitation Index (API)
- Chemicals, Runoff and Erosion from Agricultural Management Systems (CREAMS),
- National Weather Service River Forecast System (NWSRFS),
- Storage, Treatment, Overflow, Runoff Model (STORM),
- Stanford Watershed Model IV (SWM),
- Streamflow Synthesis and Reservoir Regulation (SSARR), and
- NWSRFS Snow Accumulation and Ablation Model.

In the final report, "Strategies for Using Remotely Sensed Data in Hydrologic Models" (NASA Contractor Report CR 66729 dated 31 July 1981), Hydex summarized the current relationship of remote sensing to models. It was determined that a major reason for lack of use of remote sensing was the lack of a one-to-one correspondence between any remotely sensed variables and model parameter or states. Several of the models examined could, however, be used with remotely sensed data if limited model modifications were made and appropriate data preprocessing was done.

In the current investigation, Hydex is undertaking a project to develop a suitable interface between remote sensing and current hydrologic models. Specifically, interfaces are to be

developed or modifications made to allow remotely sensed soil moisture data to be used in the NWSRFS rainfall/runoff model and remotely sensed snow covered area and water equivalent data to be used in the NWSRFS snow model.

2. NEED FOR STUDY

A recent panel on Water Resources of the Space Applications Board, Assembly of Engineering, National Research Council, stressed the importance of remote sensing techniques for predicting water resources. Among the panel's conclusions was that to be useful for prediction, remotely sensed data must be compatible with mathematical modeling of hydrologic systems.

The importance of remote sensing for improving the usefulness of hydrologic modeling for water resources prediction has been well stated and supported by the National Research Council Panel on Water Resources. Some other related factors, however, have not been stressed by the panel. Hydrologic modeling currently depends on the data base of ground measurements collected by national networks such as those maintained by the National Weather Service (NWS) of the National Oceanic and Atmospheric Administration (NOAA) and the U.S. Geological Survey (USGS) of the U.S. Department of the Interior. The quantity and quality of these networks have been steadily declining as a result of decreased resources for supporting the existing hydrometeorological networks. Remote sensing capabilities provide a viable method to offset this loss of information.

Another factor not adequately covered by the panel is the problem of applying conceptual hydrologic models in the drier areas of the United States. In those areas, precipitation is primarily convective (thunderstorms) and the ground-based data-collection network is not adequate to provide accurate information on the average precipitation. The NWS forecasters, therefore, have found it more practical to use simpler (black box) hydrologic models that are easier to adjust. However, with these black box models, it is not possible to predict, for example, low flows during drought periods and associated probabilities of occurrence. If remote sensing techniques could provide enough additional data to warrant the use of improved models, the predictive ability could be increased considerably.

3. PURPOSE AND SCOPE

The long-range purpose of the current Hydex studies is to produce operational models capable of accepting remotely sensed data and methods for preprocessing the data for use by the models. The short-term objective is to develop a sound theoretical base for making needed model modifications. The purpose of this report is to describe a method for combining measurements with widely divergent spatial scales and accuracies: it is part of the short-term objective.

The scope of this study is limited to the problem of assigning weights to point data, line data, and distributed data for entry into a model. The methods and techniques discussed are not model specific although emphasis is placed on the needs of the NWSRFS models.

In this report, the statistical weighting techniques that are developed will ultimately allow entry of widely differing types of measurements into models. For example, point soil moisture values can be combined with line-integral values from aircraft and areal-average values from satellites. The technique described uses the statistical properties of the measurements in such a way that any meaningful measurement can be used. Ten percent accurate areal averages can be successfully used along with 98 percent accurate point values. Proper account is made of the spatial resolution of each.

The remainder of the report is divided into two chapters. Chapter 2 describes a technique whereby the statistical properties of the measurements may be used to combine them for use in modeling. The mathematical development of the technique is provided in the Appendixes. Chapter 3 describes the commonly used techniques for measuring soil moisture, snow covered area, and snow water equivalent. The general properties and statistical characteristics of each measurement are summarized.

CHAPTER 2

COMPUTING AREAL AVERAGES

1. INTRODUCTION

The states of a lumped parameter (lumped input) conceptual hydrologic model represent indices to basin-wide average conditions of one or more components of the hydrologic system. To update the state or states of a hydrologic model, it is necessary to derive areal averages of one or more variables based on measurement technology for the area of the basin modeled.

Measurements may be obtained for a point, a line, or an area by many techniques. Each of these measurements contains information that should be of value in determining the best or most representative average over the basin.

Objective analysis techniques are available to determine weights to apply to each measurement in order to obtain an unbiased estimate of a random field. Such analyses have been done for fields of point measurements (1,2).* The mathematics required to handle line and areal samples becomes very cumbersome

---

*References are presented on Page 88

(a discussion of obtaining basin averages by objective analysis is presented in Appendix A). To avoid the need to use the objective analysis technique, a heuristic approach has been developed.

The purpose of this chapter is to describe the technique developed for combining measurements from various technologies to produce a reliable estimate of the mean areal average values that could be used operationally for hydrologic forecasting and prediction.

2. CRITERIA

The technique for combining measurements to produce an areal average should meet the following criteria:

1. It should not be dependent on particular measurement technologies, i.e., it should allow new types of measurements to be incorporated as they are developed.

2. It should be an objective technique that can be automated. Although it may have parameters that must be estimated subjectively, it should not require day-to-day human intervention.

3. It should work regardless of the mix of data available in any one time step. Operational data of any particular type may not be available under certain conditions; therefore, the algorithm must not depend on the availability of any particular data type.

4. It should explicitly recognize the sampling geometry of the data. This means that areal average samples from satellite platforms, point samples on the ground, and flight lines from airborne instruments will be treated differently.

5. It should explicitly recognize differences in measurement accuracy of different technologies.

6. It must produce some measure of the accuracy of the areal estimate. Assuming that the data base available at any time will vary from day to day, the accuracy of the estimated areal average will also vary from time to time.

3. OVERVIEW-CORRELATION AREA METHOD

This section presents an overview of an algorithm that meets the criteria listed above for estimating areal average from data of various sampling characteristics. Although the algorithm meets the criteria and performs in a desirable fashion, giving greater weight to measurements that are more accurate and that cover a large portion of the basin, the algorithm is a heuristic procedure. In essence, it is an "engineering approach" that takes liberties at certain points with a more theoretically correct approach in the interest of simplicity and operational capability.

The technique is called the correlation area method. It is based on assigning a weight to each sample proportional to the degree to which that sample is correlated with the variable of interest inside the basin area. Only that portion of the basin with which the individual sample is more highly correlated than any other sample is considered; this portion is called the "sample area" of each sample.

To illustrate the basic principles, a simple case is discussed. There are three measurements (a point, a line, and an areal sample) for the basin as shown in Figure 2.1. (A more detailed discussion of computing basin averages by the correlation area method may be found in Appendix B.)

For the purpose of this illustration, the following assumptions are made regarding the correlation of each measurement with the random field of a hydrologic variable (Details on the sampling correlation functions are given in Appendix C.):

(a) The areal measurement has a correlation of 0.7 with the true areal average over the sample domain.

(b) The line measurement has a correlation of 0.85 with a random point along the flight line; that correlation decreases on either side of the line as shown by the values of the parallel dashed lines on Figure 2.1.

(c) The point measurement has a correlation of 0.94 with the variable at the point, decreasing around the point as shown by the values of the circles on Figure 2.1.

The area of the basin not encompassed within the 0.7 correlation lines for the line measurement and the 0.7 circle for the point measurement indicates that the areal sample provides the best estimate for this area. The areal sample measurements have a 0.7 correlation with the random field of the variable for this area. For the rest of the basin, the point measurement (within the 0.7 correlation circle) and the line measurement (between the 0.7 dashed lines) provide a better estimate than does the areal sample measurement. The areas assigned to each of the three measurements are shown by the shaded areas in Figure 2.2.

ORIGINAL PAGE IS
OF POOR QUALITY

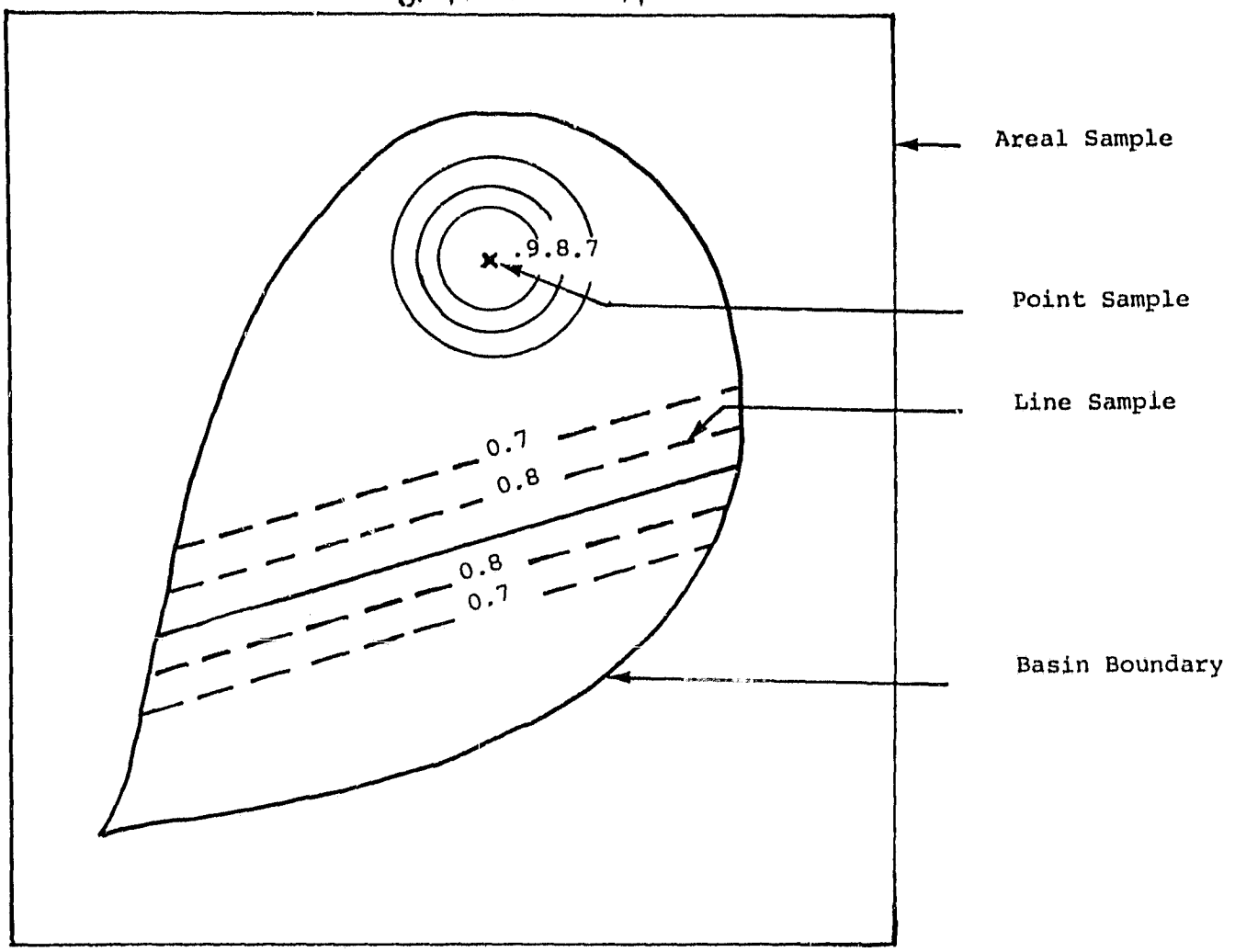


FIGURE 2-1. EXAMPLE OF CORRELATION AREA METHOD

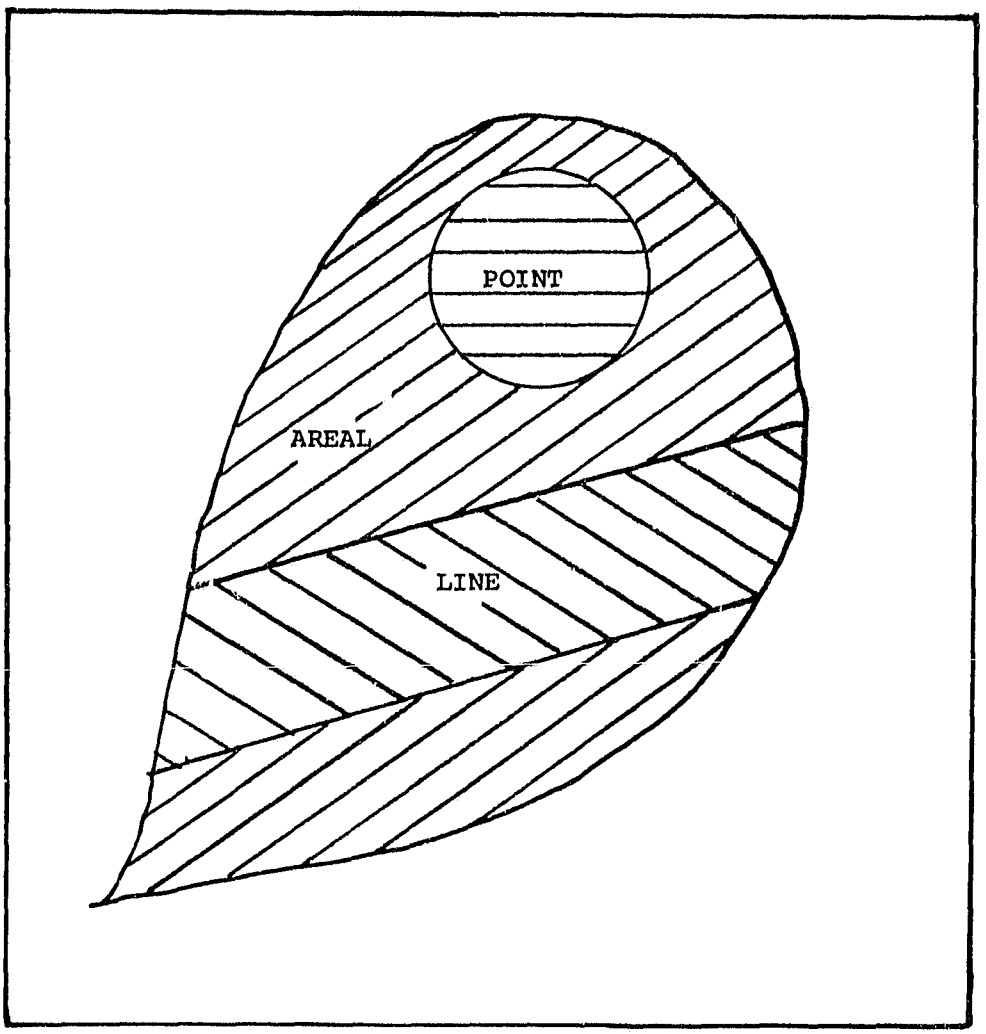


FIGURE 2-2. AREAS ASSIGNED TO POINT, LINE AND AREAL MEASUREMENTS (EXAMPLE)

The informational value of each measurement is dependent upon how much area is assigned to that measurement and on how well that information correlates with the random variable in that area. Since the correlation of the point and line measurements varies for different portions of these areas, the informational values are not consistent over these area.

The approach to determining the best average values for the basin includes a weighting for each area. The correlation area is equal to the sums of the products of the correlation multiplied by the portion of the area assigned to that measurement covered by that correlation.

For example, the correlation area for the point sample, A_p, in the case illustrated in Figure 2.1 and 2.2 is approximately:

A_p=0.92 X (area inside the 0.9 circle)

+0.85 X (area between the 0.9 and 0.8 circles)

+0.75 X (area between the 0.8 and 0.7 circles)

The correlation area for the line measurement, A_ℓ, is determined by the sum of the products of the average correlation multiplied by the area between each correlation line on either side of the flight line as shown in Figure 2.1.

In this example, the correlation area for the areal sample, A_a is 0.7 times the area of the basin assigned to the areal measurement as shown on Figure 2.2.

The final estimated areal average value, $\hat{P}$, of the variable for the basin is simply a weighted average of the measurements. The weight (λ_i) for a measurement is equal to its correlation area (A_i) divided by the sum of all the correlation areas for the basin.

$$\lambda_i = \frac{A_i}{\Sigma A_i} \qquad (2.1)$$

A measure, ω_m, of the overall accuracy of the final estimate ($\hat{P}$) is obtained by dividing the sum of the correlation areas (A_i) by actual area of the basin (B).

$$\omega_m = \frac{\Sigma A_i}{B} \qquad (2.2)$$

By its definition ω_m must be between 0 and 1 and can be loosely referred to as the correlation of the estimated value ($\hat{P}$) with the actual average (P) of the parameter over the basin.

The algorithm has an intuitively reasonable behavior; more accurate measurements get a larger sample area, a larger correlation area, and, thus, a larger weight than less accurate samples. The location of the samples also affect the shape of the sample area.

When multiple line, point, and areal samples are considered, the shape of the sample areas becomes considerably more complex than that for the simple case shown in Figure 2.1, but the basic concept is unchanged.

Details on the complete algorithm to handle complex cases are given in the following appendixes:

- Appendix D - Correlation Area of Point Samples;
- Appendix E - Correlation Area of Line Samples;
- Appendix F - Correlation Area of Areal Samples;
- Appendix G - Testing a Point for Inclusion in the Basin; and
- Appendix H - Subdividing the Basin.

4. DATA REQUIREMENTS

The sampling correlation functions discussed above require that several covariance parameters be estimated to implement the correlation area technique. These covariance parameters are related both to the accuracy of particular measurement technologies and to the correlation decay in space of the random field under study.

The purpose of this section is to introduce these data requirements. (Appendix C presents a more detailed discussion.)

4.1 Decay Parameter

The variable under study is assumed to be a random field. This random field, whether soil moisture, snow water equivalent, or whatever, is assumed to have a homogeneous mean value and variance and isotropic covariance in space described by a simple exponential covariance function.*

*A discussion of the theoretical and practical implications of these assumptions is deferred to the final report of the study.

The decay parameters α describes the value of information at one location in estimating the random field at another site. Thus, if the correlation parameter α is 0.1 kilometer e^{-1}, the value of the random field at a point has a correlation of e^{-1} with values 10 kilometers distant regardless of direction.

The three possible approaches for estimating the value of α are the use of historical data, real-time data, and a conceptual model. Of these three approaches, the historical approach and the conceptual model approach offer the most promise for operational estimates of α.

4.1.1 Historical Data (Multiple-Realization Approach)

Historical data on soil moisture or snow water equivalent can be analyzed to estimate α. Then the value of α can be assumed to apply to current conditions. The difficulties with this approach are: (a) procuring a historical data base, and (b) developing a procedure to "stratify" the data for different values of α related to some easy-to-identify property of current conditions.

4.1.2 Real Time Data (Single-Realization Approach)

If enough point data are available in real-time, a value of α can be estimated for the current condition of the random field. Using remotely sensed data for this approach is rather difficult because of the sample averaging properties of this type of data.

4.1.3 Conceptual Model

The conceptual model approach can be illustrated by considering soil moisture as the product of two random fields: the field capacity and the fraction of field capacity that is filled. The variability of the field capacity can be related to a soil map. If a conceptual model is developed to relate the statistics of the "fraction of field capacity" to some easy to estimate parameters (e.g., the antecedent precipitation index), it may be possible to estimate α for the soil moisture without actual measurements of soil moisture.

4.2 Point Sample, C_p

The value of C_p is the correlation of a point measurement with the true value at that point. Based on investigations of the accuracy of ground sampling technologies, it should be relatively easy to estimate C_p. Furthermore, it is expected that C_p will be generally "large", say, approximately 0.9 or better, for most technologies.

4.3 Line Sample, C_ℓ

The value of C_ℓ is the correlation of a flight-line sample with the value at a point randomly located along the flight line. Presumably, historical ground truth experiments will be sufficient to estimate C_ℓ. It should be understood that C_ℓ actually mixes two effects. The first is the accuracy of the technology as it measures the true flight-line average, and the second is the correlation of a point along the flight line with

the flight-line average. A more accurate technology will increase C_ℓ, but even a perfect measurement technology would not guarantee a C_ℓ of 1.0 because of the variability of point values along the flight line. Thus, C_ℓ should be expected to be less than C_p.

4.4 Areal Sample, C_a

The value of C_a is the correlation of an areal sample with the true areal average over the sampling domain. Presumably ground truth experiments will provide information to estimate C_a, but some degree of subjectivity may be needed as well.

More sophisicated techniques than those outlined above can be developed to use data from all available technologies simultaneously to estimate the statistical properties of the measurement error in each technology and the statistical properties of the random field under study. However, the required research is beyond the scope of this study.

5. SUMMARY

The correlation area method provides a means for combining measurements of various sampling geometries with measurement accuracy to produce an estimated mean areal value over a watershed and a measure of the accuracy of the mean areal value. The method is summarized below.

5.1 Data Requirements

First, the geometry of the samples and the basin must be completely described. The basin outline is described by connected straight line sections between vertex points on the basin boundary; a counterclockwise convention is adopted to order the basin boundary points. The location of all measurements is required; the location of point samples, the end points of any flight-line samples, and the corners of any areal samples must be known. A local x-y coordinate system can be produced from latitude-longitude data.

A single value is used to represent the accuracy of each measurement technology. As described above, these values are

- C_p - the correlation of a point sample with the true value at that point;
- C_ℓ - the correlation of a flight-line sample with the true value at a randomly located point along the flight line; and
- C_a - the correlation of an areal sample with the true areal average over the sampling domain.

In addition to the accuracy measures, C_p, C_ℓ, C_a, a correlation parameter, α, is required. This parameter describes how quickly the variable under study decorrelates with distance.

5.2 Procedure

Having provided the data listed above, the correlation area method proceeds in several steps as summarized below.

5.2.1 Step 1. Combine Areal Samples

If more than one areal sampling technology is available, the various areal samples must be combined into a single set of areal samples. The method for combining areal samples is described in Appendix F and involves knowledge of C_a for each technology in order to weight two samples over the region of their overlap.

5.2.2 Step 2. Divide Into Subbasins For Each Areal Sample

In this step, the entire basin is subdivided into the portion of the basin that is in the domain of each areal sample. Processing proceeds from this point by considering a single subbasin at a time. In the end, all subbasins are recombined to obtain an areal average over the entire basin (Appendix H).

5.2.3 Step 3. Point Samples

Point samples are considered one at a time. For each sample being considered, the subbasins developed in Step 2, are considered one at a time. For each subbasin, the sample area for the point sample will be the entire subbasin at first, but the effect of every other measurement, considered one at a time, may be to reduce the sample area to a region smaller than the subbasin. After the geometry of the sample area region for each subbasin is defined, the areas of the sample area and the correlation area are determined (Appendix D).

5.2.4 Step 4. Line Samples

In a manner analogous to that employed for point samples, the sample area and correlation area are found for each line sample in each subbasin. Some special processing is required for line samples that terminate within the basin boundary (Appendix E).

5.2.5 Step 5. Areal Sample

Because of the subbasin division in Step 2, each areal sample is limited to a single subbasin. The sample area of the areal sample is simply the subbasin area reduced by the sample areas of all point and line samples in the subbasin. The correlation area is the correlation multiplied by the sample area (Appendix F).

5.2.6 Step 6. Weight Computation

The total correlation area, A_i, of each measurement, m_i, is the sum of correlation areas in each subbasin. The weight assigned to each measurement is λ_i, which is computed by Equation 2.1 as

$$\lambda_i = \frac{A_i}{\Sigma A_i} \qquad (2.1)$$

Thus, the weights sum to 1.0 over all measurements.

5.2.7 Step 7. Areal Estimate

The estimated areal average over the basin is $\hat{P}$ defined by

$$\hat{P}=\Sigma\lambda_i m_i \qquad (2.3)$$

5.2.8 Step 8. Accuracy Measure

The accuracy measure of $\hat{P}$ is ω_m given in Equation 2.2 as

$$\omega_m = \frac{\Sigma A_i}{B} \qquad (2.2)$$

where B is the total basin area.

This implies that

$$0 \leq \omega_m \leq 1.0$$

where ω_m=1.0 is a perfect estimate.

CHAPTER 3

MEASUREMENTS CHARACTERISTICS

1. INTRODUCTION

One of the long-term purposes of this study is to develop methods for updating in an operational mode the NWSRFS snow accumulation and ablation (snow melt) model and the NWSRFS soil moisture accounting model. Updating is a means of keeping the simulated values (state variables) of a model consistent with the real world (i.e., with measurements of hydrologic variables in the basin).

The areal extent of the snow cover is one hydrologic variable that can be measured directly by remote sensing techniques. In addition, the snow melt model has a state variable representing the percent of the basin covered by snow.

The model variables that relate to the water equivalent of the snow cover and the soil moisture represent average values for the entire basin (or subbasin). There are no known techniques for directly measuring the average values of these two variables or any combinations of them over a basin. They must be estimated from point samples, from remotely sensed data, or from both. The commonly used approach is to derive estimates of areal averages from ground based point measurements.

Statistical information is required for application of the correlation area method (described in Chapter 2) to compute basin averages of the water equivalent of the snow cover and of the soil moisture.

The purpose of this chapter is to review the physical and statistical properties of the three hydrological measured variables for which updating procedures are being developed.

The statistical properties of the snow cover and of soil moisture vary considerably with differenct climatic regions. For this reason it was felt it would be best to select an area in which the climate was fairly homogenous and one in which the procedures would be most useful. For purposes of discussion, the North Central Plains area was selected. It is an area that is fairly homogeneous in its vegetation, soils, and climate, is a major agricultural area, and is subject to considerable flooding. Data and results from other areas of the United States are included where appropriate or when they are the only data that exist.

2. MODEL CONSIDERATIONS

From the previous studies (3,4) it is evident that there is not a simple one-to-one relationship between a state variable, such as soil moisture, of the model and the measured field values from the basin. The updating process must take into account how the models perceive the real world and must provide measurements to fit into this perception. The state variables of concern will now be discussed.

The NWSRFS (Sacramento) soil moisture accounting model divides the area of the basin that receives precipitation into three parts. These parts are shown in the schematic diagram in Figure 3.1. They are the impervious area (parameters PCTIM), the

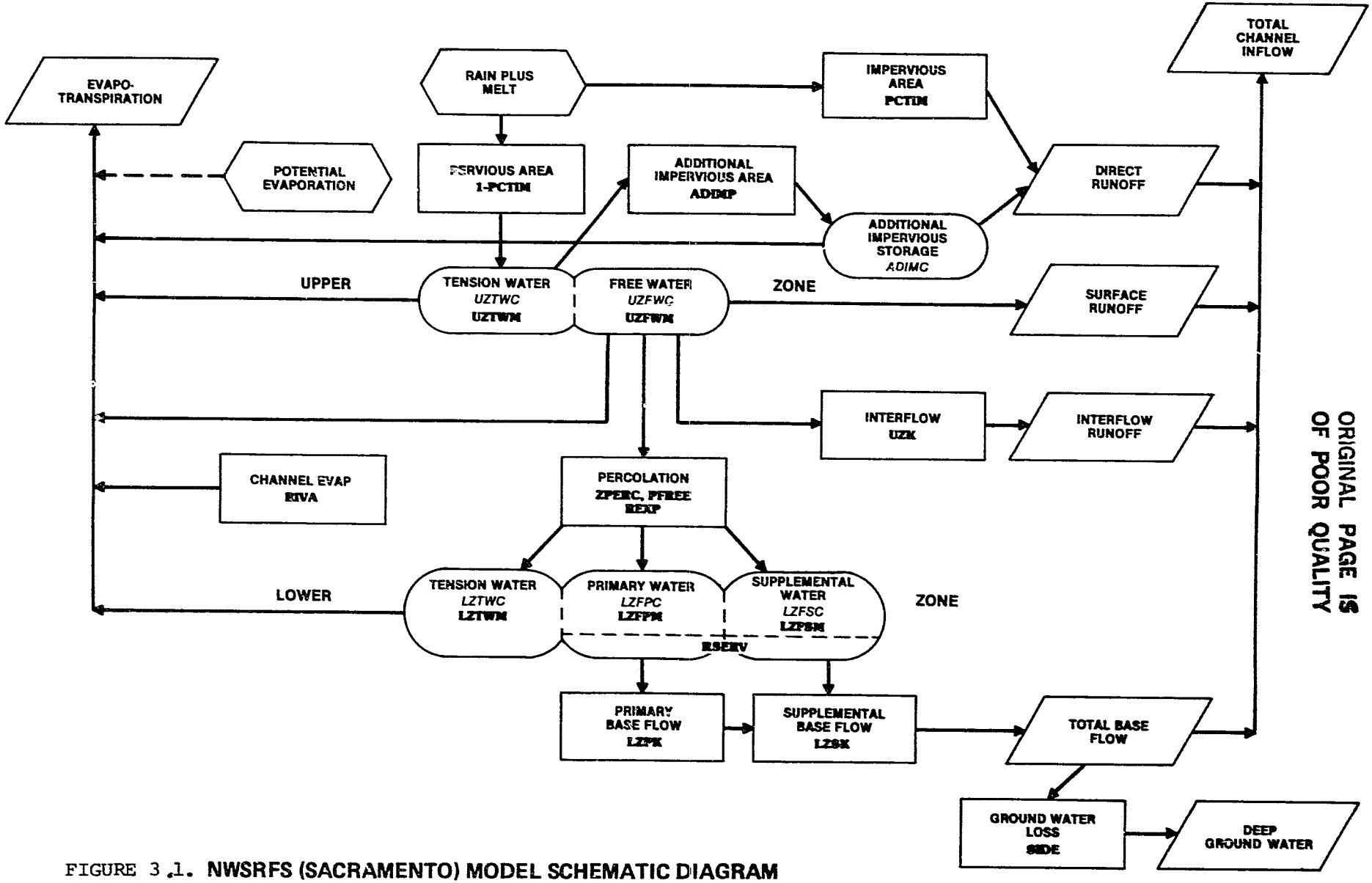


FIGURE 3.1. NWSRFS (SACRAMENTO) MODEL SCHEMATIC DIAGRAM

pervious area (1-PCTIM), and the part of the pervious area that becomes impervious as the wetness of the basin increases (the additional impervious area, parameter ADIMP).

The PCTIM is that portion (fraction) of the basin that is impervious and contiguous with the stream channels. It includes all water-covered areas and other areas such as solid rock from which there is a direct stream response when any precipitation occurs. The ADIMP is that portion (fraction) of the basin that becomes impervious as all tension water storage becomes full (upper zone tension water, UZTW, and lower zone tension water, LZTW on Figure 3.1) becomes full. In this study, the ADIMP is referred to as the wet area of the basin. The dry portion of the basin is the pervious portion of the basin (1-PCTIM) less the portion that is assigned as the wet portion (ADIMP).

The snow melt model is shown schematically in Figure 3.2. Although there is no simple state variable that directly relates to the water equivalent of the snow cover as measured (total mass of water above the ground), a value of the water equivalent may be determined by summing several states of the model. The areal extent of snow cover is also of concern.

3. INFORMATION REQUIREMENTS

Five measurements of the hydrologic conditions of the basin are required in order to use the correlation area method and to update the two selected models. These measurements are the

(1) fraction of the basin that is snow covered,

(2) fraction of the basin that is covered by water,

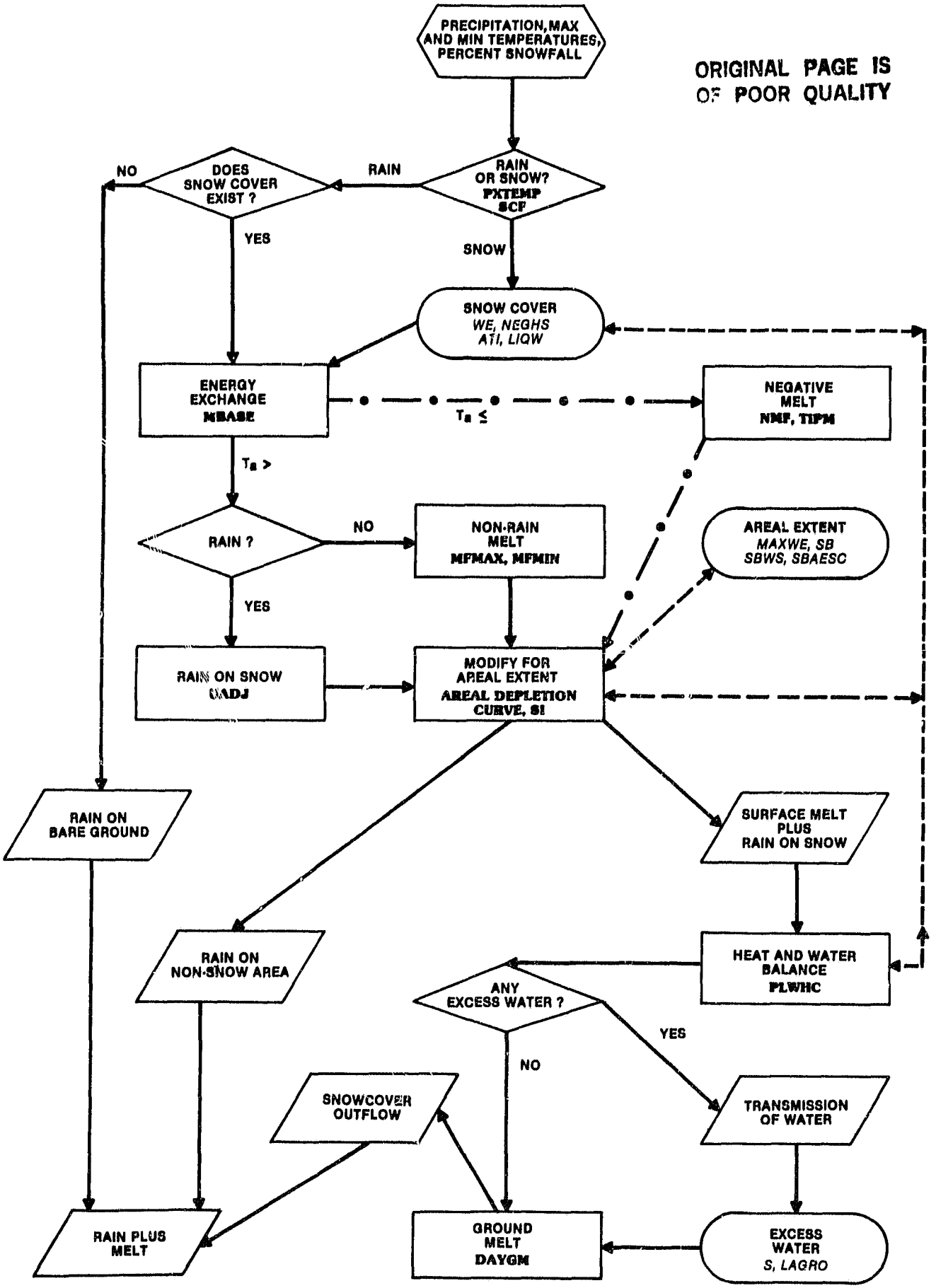


FIGURE 3.2. NWSRFS (ANDERSON) SNOWMELT MODEL SCHEMATIC DIAGRAM

(3) fraction of wet area of basin (ADIMP),

(4) average water equivalent of the snow cover (mm) over the entire basin, and

(5) Average soil moisture (percent by dry weight; mm) on the dry portion of the basin to a specific depth, Dm.

In addition to the above measurements, statistical information on the accuracy of the measurements (water equivalent and soil moisture) and on the variability of the measured variables in the basin is required. This information include a measure of how well the point measurements correlate with other measurements of the same variable in the basin,the decay parameter α, and the correlation values (C_p,C_ℓ, and C_a) discussed in Chapter 2.

A section of this chapter is devoted to a discussion of the measurement technology and to the determination of the statistical information required for each of the five measurements required.

4. AREAL EXTENT OF SNOW COVER

The determination of the areal extent of snow cover is the most operational applied remote sensing measurement. The techniques for determining snow covered area and the problems associated with the measurements are well documented.

The discussion of areal extent of snow cover is divided into three sections. First, the available data sources from aircraft

to satellite are discussed; second, the techniques available for interpreting the data are discussed; and third, the statistical properties of the measurements relative to the correlation area method are discussed.

4.1 Data Sources

A variety of remote sensing systems report data suitable for use in determining the areal extent of snow cover. The most popular system is the multispectral scanner (MSS) imagery from the LANDSAT satellites. The MSS records spectral energy data from four bands (referred to as Bands 4,5,6 and 7). The data from the two LANDSAT satellites are available every nine days and have a resolution of 80 m (approximately 1 acre). LANDSAT MSS Band 5 permits rapid snow mapping; Band 7 is also used for snow mapping under light cloud cover and hazy conditions. False-color composite images from several bands are used in difficult conditions such as forested areas, under clouds, or in steep terrain.

A second source of satellite data is the TIROS (NOAA) series of satellites. Very high resolution radiometers (VHRR) provide once-a-day coverage in the visible light range and twice-a-day coverage in the thermal infrared range. A resolution of 1 km is available directly below the satellite track.

A third data source is the Geostationary Operational Environmental Satellite (GOES) system. The GOES acquires imagery

in the visible and thermal infrared ranges by means of visible infrared spin scan radiometers (VISSRs). VISSR imagery is available as frequently as every 30 minutes, and the data resolution is 1 km.

A final source of data on snow covered area is aircraft flights. Aerial photography of various kinds (high altitude, low altitude, oblique) has been used for many years and continues to be used. Aircraft flights still offer the best resolution (less than 1 m) when flown at low altitude.

4.2 Interpretation Techniques

Most remote sensing data are available in formats that range from photographic prints and transparencies to digital tapes. Snow covered area is determined from the various formats by several techniques, the most common of which are described here briefly.

Most methods of estimating snow covered area involve photo interpretation techniques or computer analysis. The photo interpretation methods are generally viewed as more timely and less expensive.

The most common (and recommended) photo interpretation tool is the zoom transfer scope, which the zoom transfer scope allows

the user to stretch, rotate, and zoom in or out on images recorded on film transparencies. The altered image is projected on a topographic map of known scale wherein the snow line can be traced. The zoom transfer scope can be used with images from GOES, TIROS (NOAA), and LANDSAT, as well as with aerial photography.

A second mapping technique is density slicing, a method by which a positive transparency is laid on a light table with an opaque mask covering areas not of interest. A camera records the transparency image in shades of gray and breaks the available light range into 12 discrete levels. The 12 levels can be displayed on a monitor in 12 false colors. An operator selects the colors associated with snow and the system electronically planimeters the snow covered areas.

A third mapping method, called color additive viewing, makes use of several images at once. The technique is effectively used with multispectral scanner data. In this method up to four images are registered with one another to produce a false color or natural color composite picture. The resultant image is projected on a base map and interpreted manually, much like the zoom transfer scope method.

A final analysis method involves computer-assisted analysis of digital imagery tapes. Each LANDSAT scene, for example, is available as a computer compatible tape (CCT), wherein each pixel

is recorded as digital bits describing the reflectance level in each spectral band. The digital information can be used to project false color pictures on a graphics terminal. Available software allows the user to manipulate the data and "teach" the computer to recognize snow (or any number of other features). The computer can easily perform electronic planimetering and other useful computations.

4.3 Statistical Properties

Normally the correlation area method requires two coefficients in order to incorporate a particular type of remotely sensed data into a model (see Chapter 2). The first coefficient is a measure of the accuracy of the method, and the second is a measure of the spatial area over which the measurement applies.

The snow covered, water covered, and wet area estimates are different from other measurements in that no areal extent coefficient is involved when applying the correlation area method. It is implicitly assumed here that a given measurement applies to the entire basin. Thus, only a measure of accuracy is required.

In Chapter 2 the measure of accuracy chosen for areal average measurements was C_a. The value of C_a is the correlation between the given measurement (in this case the estimated percent of snow covered area) and the true value. In the case of snow

covered area, C_a will depend on the imagery type (primarily on the resolution), the operator, interpretation technique, terrain, vegetation cover, and weather (cloud cover). At this time it is not possible to assign precise values of C_a under all circumstances. Some data are, however, available, and two of the factors upon which C_a depends--imagery type and interpretation technique--will be discussed to the extent possible.

4.3.1 Imagery Type (Resolution)

If large, relatively flat, plowed fields with little vegetation are photographed from an aircraft, high resolution is possible. Under such conditions, the snow covered area could be determined with an accuracy of 95 percent or better with LANDSAT data in basins greater than 25 sq km and with NOAA (1 km resolution) for basins greater than 200 sq km; that is, C_a=0.95. Conversely, if a GOES VISSR image (1 km resolution) is used to determinate snow covered area in a 40 sq km basin in forested mountain terrain, the accuracy might be 50 percent or less; that is, C_a=0.5.

The problem of accuracy as determined by imagery resolution has been addressed by Jackson and McCuen (5) in relation to impervious area determination from LANDSAT data. Their analysis is equally valid for snow covered area.

In Jackson and McCuen's terminology, the minimum area that can be discriminated in an image is a "sensor resolution element" (often called a pixel). The area over which the measurement is desired is referred to as a "hydrologic model unit." They

develop a model for the standard error of estimate of a collection of sensor resolution elements in relation to the standard error for a single resolution element.

Comparisons were made between the LANDSAT-based estimates and the "true" parameter values for each of the 576 individual sensor resolution elements. Extensive sampling of low level aerial photos was used to develop the "true" values. The parameter values for groups of sensor resolution elements were then averaged to obtain a parameter value for each group. This procedure was performed for the "true" data set and each of the LANDSAT-based data sets, and the average parameter values were compared. Two methods--parallelpiped analysis and radiometrically corrected mixture model analysis--were used to determine whether a particular sensor resolution element contained impervious area.

As a result of the analysis it was determined that a log-transformation equation was a good model of the standard error versus unit size [number of sensor resolution elements per hydrologic model unit (basin)]. Figure 3.3 illustrates the variation of standard error as determined by Jackson and McCuen.

The figure can be used on a guide in selecting the portion of C_a that is the result of sensor resolution. The user should compare the size of the basin being analyzed with the size of the sensor resolution elements. If 50 resolution elements (pixels) or more cover the basin, the standard error is less than 10 percent and C_a is greater than 90 percent. Conversely, if only one resolution element covers the basin, $C_a \cong 1-0.25=0.75$. If the

ability to determine whether snow actually covers a sensor resolution element is less than the ability to detect impervious area, then the C_a values would be reduced accordingly.

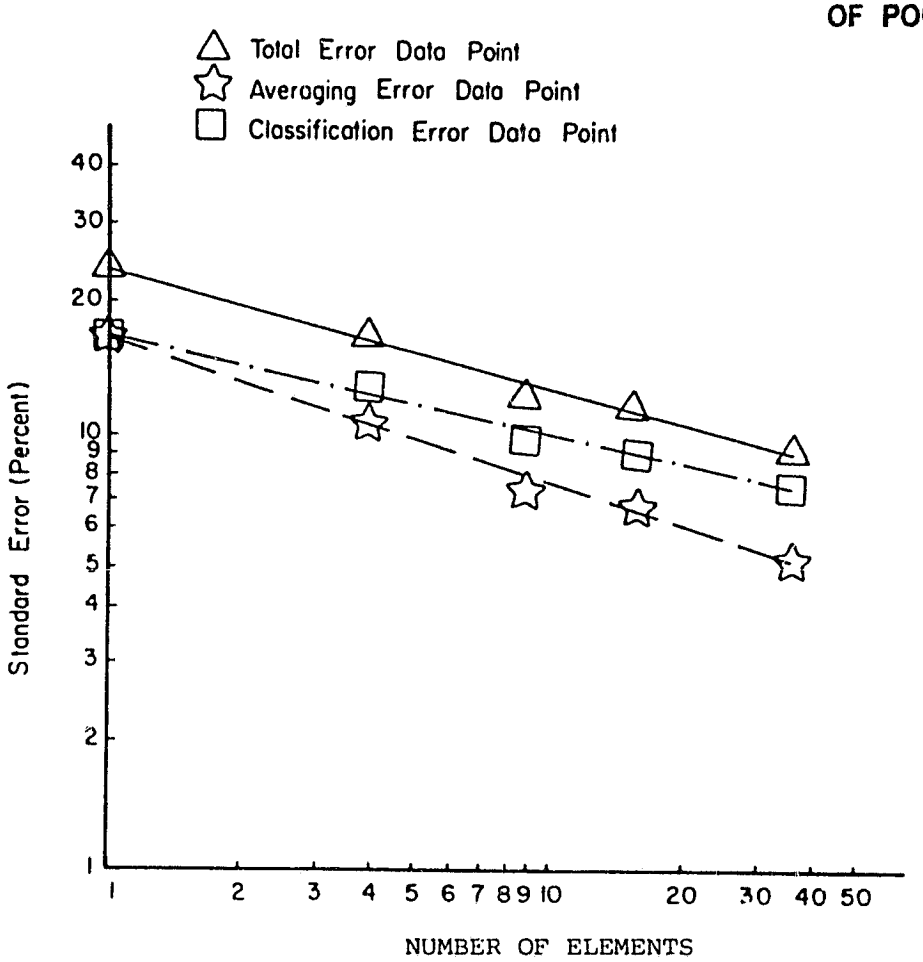


FIGURE 3.3 NUMBER OF SENSOR RESOLUTION ELEMENTS COVERING A SINGLE BASIN

4.3.2 Interpretation methods

Some indication of the variation in accuracy between interpretation methods can be obtained from NASA's Application Systems Verification and Transfer (ASVT) program for snow mapping (6), Volume IV, "Operational Applications of Satellite Snow-Cover Observations-Colorado Field Test Center," presents data on snow covered area determination in various basins by several interpretation techniques. The techniques include areal photography and LANDSAT image analysis by zoom transfer scope, color additive viewer, density slicer, and computer methods (G.E. Image 100 System, Colorado State University Computer). Table 3.1 compares the data. (The grid method was considered too poor to be useful.)

TABLE 3.1 COMPARISON OF SIX METHODS OF SNOW MAPPING PERFORMED IN THE COLORADO ASVT STUDY

Image Date	Drainage	Aerial Photography	Percent Basin Snowcover					
			Zoom Trans Scope	Color Additive	Density Slicer	Grid	Image 100	CSU Comp
May 12, 1974	Conejos		42	37	38			
	Alamosa		51	39	35			
	South Fork		28	30	31			
	Rio Grande		27	8	6			
May 30, 1974	Conejos		16	14	15			
	Alamosa		19	19	17			
	South Fork		6	12	10			
	Rio Grande		7	3	2			
June 3, 1975	Conejos		47	43	31	22	28	12
	Alamosa		63	44	28	38		
	South Fork		30	40	31	28		
	Rio Grande		25	20	9	11		
April 3, 1978	Conejos	87	89					
April 13, 1976	Conejos	81	84					

The authors of the ASVT report believe that the zoom transfer scope data analysis is the most accurate snow mapping method. By comparing the area values in the remaining columns with those in the zoom transfer scope column, an idea of relative accuracy can be determined. It is evident that some analysis techniques introduce bias as well as random error.

The implications for estimating C_a corrections required because of interpretation technique are somewhat subjective. By averaging the percent difference between the color additive viewer determined areas and the zoom transfer determined areas, a value of -7.9 percent is obtained. If the 100 percent positive error in the South Fork Basin is neglected, a value of -17 percent is obtained. This value implies a strong bias as well as a decrease in accuracy. Thus, it would be reasonable to reduce C_a by 0.1 to 0.2 if the color additive viewer is used. Similar analysis of the density slicer results indicates an average error of -23.3 percent from the zoom transfer values, making a reduction of C_a on the order of 0.2 to 0.3 appropriate. The data from other analysis techniques are too limited to allow any estimate of their effects on C_a.

5. AREAL EXTENT OF WATER COVERED AREA

The subject of this section, areal extent of water covered area, and the subject of the subsequent section, areal extent of

wet area, are directly related to the NWSRFS (Sacramento) model. The two areal extent estimates will be used in estimating and updating the percent impervious area (PCTIM) and additional impervious area (ADIMP) discussed in Chapter 2.

The discussion in this section is divided into two parts. First, the data sources and interpretation techniques, are discussed, and second, the statistical properties of the measurements related to the correlation area method are discussed.

5.1 Data Sources and Interpretation Methods

The data sources for determining water covered area do not differ substantially from those for determining snow covered area. However, some of the sources are more likely to be used than others. The most practical data sources are:

- LANDSAT MSS,
- NOAA-2 VHRR, and
- aerial photography, both conventional and infrared.

Similarly, the interpretation methods are nearly the same. The zoom transfer scope, color additive viewer, and computer processing methods still apply. A short description of the factors that distinguish water covered areas in the various types of data is now presen ed.

The data available from the LANDSAT MSS will be discussed first. Recall from the previous section that the MSS records spectral energy data from four bands (referred to as Bands 4,5,6, and 7). The data are available every 9 days and have a resolution of 80 m (approximately 1 acre). Two of the bands, 6 and 7, are highly effective at distinguishing open water areas. Band 6 measures energy in the upper and lower infrared region and is effective in detecting land use and giving maximum land-water contrast. Band 7 measures energy in the infrared region and provides the best land-water descrimination.

The NOAA-2 VHRR on the TIROS satellites detects both visible and thermal infrared radiation. The data are available twice daily and have a resolution of 1 km. Water covered area can be determined from either visible or infrared data. Determinations made with visible data depend on the contrast between the water and the surrounding land; determinations made with the thermal infrared data depend on temperature differences. Wiesnet, et al.(7), describe the use of the VHRR data for flood mapping and report reasonable results.

Moore and North (8) describe the use of a variety of methods to detect water covered areas, including aerial photography. They indicate that in open agricultural areas and urban areas water boundaries are easily detected by all types of photography and imagery. Water covered areas are more difficult to define in vegetated areas. More and North report that in March, even with trees dormant, black-and-white and color photographs show only treetops. Color infrared photographs, however, show a distinct

tone in inundated woods. After April, both color infrared photography and thermal infrared imagery detect only treetops. The December to March time frame is best for mapping water covered areas under vegetation.

5.2 Statistical Properties

It is assumed that the determination of areal extent of water covered area will be made for an entire drainage basin at once. Thus, just as with snow covered area, only one statistical characteristic is required for the correlation area method. The characteristics is again C_a, the correlation between the given measurement (percent water covered area) and the true value. The value of C_a depends on the imagery type (resolution), vegetation cover, and other factors. Guidelines for estimating C_a under various conditions will now be presented.

5.2.1 Imagery Type (Resolution)

The effects of sensor resolution or accuracy have already been discussed in relationship to snow covered area. The work of Jackson and McCuen (5) relating standard error to the relationship between hydrologic model units (basin size) and sensor resolution elements (pixel size) was presented. There is no reason to believe that their analysis is not also applicable for determining water covered area (with appropriate adjustments for the ability to discriminate water as opposed to impervious area).

Using the same principles as before, if 50 or more sensor resolution elements cover the basin, C_a will be 0.90 or higher. A C_a value higher than 0.9 would apply, for example, if low-altitude aerial photography is used in an area with little vegetation cover. If the basin size approaches the resolution element size, the C_a should be 0.75 or less.

5.2.2 Vegetation Cover and Other Factors

The effects of vegetation cover on C_a values will depend on the type of data available, the time of year (hence density of vetetation), and the percent of the drainage area that is vegetated. Little or no data exist to specifically identify the effect of vetetation on the accuracy of water covered area estimates. However, general effects of vegetation on interpretation results have been documented.

The results of Moore and North (8) have been discussed earlier. Conventional aerial photography is largely ineffective in determining water covered area where tree cover exists. Color infrared photographs or infrared imagery is effective under some tree cover in the period of December through March. No remote sensing technique is particularly effective in the period of heavy leaf cover or under crop cover.

In applying the correlation area method it is suggested that be adjusted as follows to account for vegetative cover and seasonal effects. First, the user must approximate the boundaries of the water covered area. This should be done even if the entire basin is covered with dense forest. Next, the

percent of the water covered area overlaid by vegetation should be estimated. Next, the user should estimate the impact of the vegetative cover on the estimate of areal extent of water covered area by outlining the mimimum area likely to be covered by water and the maximum area likely to be covered by water in the areas where visability is poor. In the late spring through fall for all type of photographs and imagery the combined estimate of C_a is

$$C_a = C_{a_r}(1-\%_{wv})\left[(1-\frac{A_{max}-A_{min}}{A_L})\right] \qquad (3.1)$$

where

C_{a_r} is the C_a value based strictly on sensor resolution,

$\%_{wv}$ is the approximate percentage of water covered area also covered by vegetation,

A_{max} is the user's estimate of the maximum area covered by water,

A_{min} is the user's estimate of the minimum area covered by water, and

A_L is the user's best estimate of the water covered area.

In the period from early fall to mid-spring when using infrared photographs that can successfully detect water through the vegetation cover,

$$C_a = C_{a_r}\left[1-\frac{A_{max}-A_{min}}{A_L}.\right] \qquad (3.2)$$

When using conventional photography with limited ability to see through vegetation at all times

$$C_a = C_{a_r}\left(1-\%_{wv}\right) \qquad (3.3)$$

Although the procedure for determining C_a is subjective, complete accuracy in determining C_a is not critical. The power of the correlation area method allows the use of measurements with widely differing accuracies. No harm comes to the model from small errors in the weighting of the different measurements.

6. AREAL EXTENT OF WET AREAS

Before discussing methods for determining wet areas, a short review of the definition is in order. The NWSRFS (Sacramento) model has state variables discribing the impervious area of the drainage basin and the additional impervious area covered by wetting and drying of areas along streams, lakes, and other watercourses. This additonal impervious area (ADIMP in the model) is the wet area that must be determined. Determining the wet area is related both to determing the water covered area and to detecting soil moisture.

In this section the data sources and interpretation techniques for determining wet areas are treated and then the determination of the accuracy measure for application of the correlation area method is discussed.

6.1 Data Sources and Interpretation Techniques

The data sources for determining wet areas are limited by the nature of the problem. That is, the wet areas will often be a fairly small percentage of the drainage area in relationship to the dry areas. Thus, a relatively high degree of resolution will be required in order to make meaningful measurements. It is unlikely, for example, that sensors with resolutions of 1 km (such as those aboard the NOAA-2 satellite) could make meaningful determinations of wet area.

At this time, the most likely sources of data for wet area determination are LANDSAT imagery and color infrared photography. Satellite microwave sensors at this time lack adequate resolution. Aircraft-borne sensors do not offer full-basin coverage in most cases (a stated assumption of the areal measurements in the correlation area method).

Procedurally, interpretation of photography or imagery for determination of wet areas will differ little from determination of water covered areas. The zoom transfer scope, color additive viewer, density slicer, and computer processing (LANDSAT) can be used.

The available data sources are not ideal for determining wet areas. Until high resolution, large-area coverage microwave or thermal infrared data are available, it will be difficult to discriminate between open water and saturated on very wet ground. Wet area will most likely be estimated by computing the change over a short time span in the water covered area. For example, the water covered area in a basin under flood conditions minus

the water covered area under dry conditions provides an estimate of the maximum wet area. LANDSAT data can provide this type of information. More accurate estimates of wet area might be made by compiling color infrared photography of a basin on several successive days and at different times. Differences in the surface soil temperature and water temperature over the course of a day can be used to distinguish wet areas. (The section on Soil Moisture in this chapter provides details.)

6.2 Statistical Properties

Data on the accuracy of wet area determinations are virtually nonexistent. However, it is reasonable to assume that the accuracy of wet area data is about the same as the accuracy of water covered area data that comes from the same sources. It is thus recommended that the areal extent of wet areas be determined by using the same formulas that were used to determine water covered area. The C_a values from current remote-sensing techniques will probably be quite low (perhaps 0.5 or less). It should again be emphasized that poor quality measurements may have value and are handled by the correlation area method.

7. WATER EQUIVALENT OF THE SNOW COVER

The water equivalent of the snow cover is the depth of water that would result from snow melting. Water equivalent is usually

determined by sampling with a snow tube. Most samples of the snow cover in the United States are made using a Federal Snow Sampler, a tube with an inside diameter of 37.8 mm (1.48 in.) and designed for use on deeper snow than commonly occurs in the North Central Plains region.

The snow measurement program in the western United States is coordinated and supervised by the Soil Conservation Service (SCS) of the Department of Agriculture. No single agency is responsible for snow sampling outside the western states. The only routine measurements are collected at synoptic meteorological stations of the NWS. Similar measurements are made once or twice a week at selected NWS climatological stations operated by cooperative, unpaid, observers. The measurements by the NWS stations are usually made by inverting an eight-inch rain gage to collect a core and then weighing the core to obtain the water equivalent. Some of these gages are fitted with a set of cutting teeth to facilitate core sampling when ice lenses are.

The observed water equivalents of the snow cover from some of the NWS stations are published in the monthly and annual Climatological Data publications of the Environmental Data Information Service (EDIS) of NOAA.

During periods of heavy snow cover, special surveys are operated by a number of federal and state agencies (e.g., NWS, SCS, and the U.S. Army Corps of Engineers) to obtain additional information on the water equivalent of the snow cover. Some of these measurements are made using the Adirondack Sampler, a fiberglass tube approximately 10 cm (4 in.) in diameter. These

measurements are not published, but maps showing the measured water equivalents near the time of maximum snow cover for years during which heavy snow cover occured in the North Central Plains area are available at the NWS Hydrologic Research Laboratory, in Silver Spring, Maryland.

Even a casual review of these maps of the water equivalent near the time of the maximum snow cover provides an insight into the need for an improved measurement methodology. A look at the individual measurement shows an extremely high variability among adjacent points. However, a general pattern exists and an analysis can be performed. From this analysis the NWS River Forecast Centers select basin averages for individual basins. The accuracy of these values for an individual basin (1300 to 2600 sq km or 500 to 1000 sq miles) is highly questionable, and the need for improvement in estimating the areal average of the water equivalent of the snow cover becomes apparant.

7.1 Measurement Accuracy

Snowfall or snow on the ground is the most difficult form of precipitation for which to determine representative measurement (9). The Federal Snow Sampler is the snow measuring equipment that has been most extensively evaluated. Evaluation of the Adirondack sampler has shown that it is very accurate when used by an experienced and trained observer.

Most of the observers in the North Central Plains area receive very little training on either the actual measurement technique or the selection of a measuring site. The evaluation

of the accuracy of the measurements is complicated because of this lack of trained personnel and because of the variability of the snow cover. The representativeness of any point sample varies from one measurement to another because of the redistribution of the snow by wind movement. In addition, the distribution of the snow is greatly dependent upon the variability of the roughness of the surface. The agricultural cropping patterns (height of stubble, type of crops, cultivation of fields, etc.) makes an objective analsis of the statistical properties of random snow samples extremely difficult and uncertain.

The recently introduced aerial gamma radiation program for obtaining line measurements of the water equivalent of snow cover provides a measurement technology with an associated evaluation of the accuracy of the measurements (6). Many of the statistics required for the correlation area method may be determined from data associated with the gamma radiation program. The statistical information required for the correlation area method is discussed in the following sections.

7.2 Point Measurements

The statistical value required for point measurements of snow water equivalent is the correlation (C_p) of a point sample with the true value at that point. This value can not be computed directly since no data are available.

An estimate of the accuracy of a sample (for example, the standard deviation of the measurements from the true mean) and a

measure of the variability of the measurements over a random field can be combined to obtain an estimate of C_p.

Data for use in this approach are not available. However, some knowledge is available on the accuracy of a point measurement if we assume no bias. The variability (σ) of the measurements in a random field (in this case, the basin) is dependent upon the variability of the snow cover. Snow is redistributed by wind and the resulting distribution is largely a function of the variability of the roughness of the ground cover. Since the areas being considered are primarily agricultural, a measure of this roughness can be estimated from a knowledge of the cropping practices in the basin.

Because of interactions that will be discussed later in this report, the actual value of C_p for a first estimate is not critical. Because of the uncertainty in the measurements of point data in the North Central Plains region, the accuracy of measurements is assumed to be ±15 percent with no bias. A corresponding value of 0.85 is suggested as an initial value for C_p for the North Central Plains region. This value is based on engineering judgement and field experience in making point measurements under adverse winter conditons in the region.

7.3 Line Measurements

Although Line measurements of hydrologic variables in a basin can be obtained in a number of ways, the most common is the aerial survey.

The NWS, is conducting an airborne gamma radiation measurement program in the North Central Plains area, and measurements of the average water equivalent of snow cover for specific flight lines are available for operational hydrologic purposes. At the present time, approximately 240 lines have been established and are surveyed each season when snow occurs.

During the research and development phase of the NWS aerial gamma radiation measurement program, extensive ground surveys of the snow cover were conducted to provide ground truth for test and evaluation of the gamma radiation technique. A large portion of the research was conducted at an established line (8.45 miles long) near Luverne, Minnesota (10). These unpublished data are an aid in determining the statistical parameters required for applying the correlation area method.

As noted in the previous section on point measurements, the variability of the snow cover is highly dependent upon the crop cover (roughness) of the area. For example, field experience has shown that the snow depth varies from 111.8 cm (44 in.) covering a quarter section field of high corn stubble to 30.5 cm (12 in.) covering an adjacent quarter section on which the stubble was cut low. Even with such extreme differencies, the average snow depth over a basin for the same cover conditions has been found to be very consistent.

Table 3.2 is a summary of the variability of the snow cover over the southern three miles (Line W) of the Luverne, Minnesota,

TABLE 3.2 VARIATION IN SNOW COVER WITH CROP TYPE
Luverne, Minnesota

SURVEY
17 February 1970

TYPE OF COVER	DEPTH MEASUREMENTS (NUMBER)	AVERAGE DEPTH (inches)	AVERAGE WATER EQUIVALENT (inches)	DENSITY (Percent)
Corn Stubble	191	10.0	2.40	24
Soy Bean Stubble	48	8.2	2.26	28
Pasture	44	7.0	2.13	30
Hay Stubble	54	7.3	1.91	26
Alfalfa Stubble	128	6.6	1.90	29

SURVEY
12 February 1972

TYPE OF COVER	DEPTH MEASUREMENTS (Number)	AVERAGE DEPTH (inches)
Corn Stubble (east-west)	213	12.2
Corn Stubble (north-south)	112	15.0
Soy Beans Stubble	56	8.4
Pasture	55	9.9
Hay Stubble	79	9.0

research line (Line A) in February 1970 and February 1972. Ground surveys were made along the entire three miles with depth readings every 10 paces and water equivalent measurements each 100 paces.

For the 17 February 1970 data, average densities were established for the different fields and applied to the snow depths to obtain the water equivalent information. The data for 1970 shows the greatest depths and water equivalent for fields of corn stubble and least for field of alfalfa stubble. This finding is consistent with the height of the stubble of these crops.

The lower part of the table shows information on snow depths by crop type for 12 February 1972. In that survey, the corn stubble was subdivided into direction of planting (east-west and north-south). The average snow depth was greater (38.1 cm or 15.0 in.) on the fields planted north-south than on those planted east-west (31 cm or 12.2 in.). An assumption could be made that the north-south rows provide more of a wind break to the generally westerly winds and would result in a higher deposition of snow.

The primary purpose of presenting the information in Table 3.2 is to demonstrate the high variability and at the same time the consistency in an area.

Table 3.3 presents statistical information on the water equivalent and snow depths for Line W as measured over a three year period. The variability of the snow cover results not only from wind action but is highly dependent upon the air temperature

TABLE 3.3 SNOW COVER MEASUREMENT STATISTICS
Luverne, Minnesota

LOCATION	DATES OF SURVEY				
	6Jan1970	17Feb1970	6Jan1971	11Feb1971	11Feb1972
LINE W (3 Mile)					
WATER EQUIVALENT					
Number of Samples	10	35	66	42	31
Mean (inches)	1.60	2.38	1.05	1.19	2.57
Standard Deviation (inches)	0.34	0.80	0.38	0.43	0.60
Coefficient of Variation (percent)	21	34	36	36	24
SNOW COVER (DEPTH)					
Number of Samples	140	510	329	243	544
Mean (inches)	8.7	8.5	5.5	4.0	11.5
Standard Deviation (inches)	2.0	2.9	1.5	116	3.8
Coefficient of Variation (percent)	23	34	27	40	33
DENSITY					
Percent	18	25	19	39	22
LINE A (8.45 Miles)					
WATER EQUIVALENT					
Mean (inches)	1.68	2.32	1.06	1.35	2.32

during and after snowfall. Although there is a correlation (positive) between the water equivalent and the standard deviation of the water equivalent measurements, the coefficients of variability ($Cv=\sigma/\bar{f}$) are not consistent.

The water equivalents for the 3-mile portion of the research line (Line W) are well correlated with the average water equivalents for the entire line (Line A).

The airborne gamma radiation measurement technique is used by several countries (11). Evaluation of that technique is described in the literature. For the North Central Plains area, the standard error of measurement has been given as ±1 cm (10).

The correlation of an airborne line measurement with a random point along the line (C_ℓ) is dependent to a large extent on the average water equivalent for the line. For example, when the average water equivalent is 2.54 cm (1 in.), C_ℓ would be much less than when the average water equivalent is 7.6 cm (3 in.).

The value of C_ℓ also depends upon melting conditions of the snow. During the snow accumulation period, the consistency in the snow cover is as discussed above. When melting begins, the variability becomes much greater (and, thus, the value for C_ℓ becomes lower). Areas in which there is slight shade may have little or no melting while at the same time, other areas at slightly lower elevations may have no snow or standing water. The measurement of the areal extent of the snow cover becomes more important as the snow covered area begins to decrease.

The relationship of the average water equivalent to the value of C_ℓ for use with line measurements from the NWS aerial

surveys would have the form indicated in Figure 3.4. For periods of light snow cover, the relative error of the measured line average of the water equivalent of the snow cover increases. The result should be a decrease in C_ℓ as shown on Figure 3.4. For depths greater than 4 cm (1.6 in.) (up to a depth were the accuracy decreases because of large attenuation of the radiation, approximately 30 cm or 11.7 in.) a value of 0.8 is indicated for C_ℓ.

It is recommended that Figure 3-4 be used to obtain an initial value for C_ℓ (using an estimate value of water equivalent) until field testing of the technique is conducted and an improved relationship is defined.

7.4 Areal Measurement

Currently, no operational or near-operational means is available for estimating snow water equivalent over large areas. The greatest potential for such measurements appears to be in microwave sensors. Schmugge (12) reports a variety of experiments that indicate that microwave brightness temperatures are a function of snow water equivalent. However, microwave response is a function of snow depth, grain size, water content, type and condition of underlying soil, and wavelength of observation. Considerable research will be needed before any reasonable areal estimates are feasible.

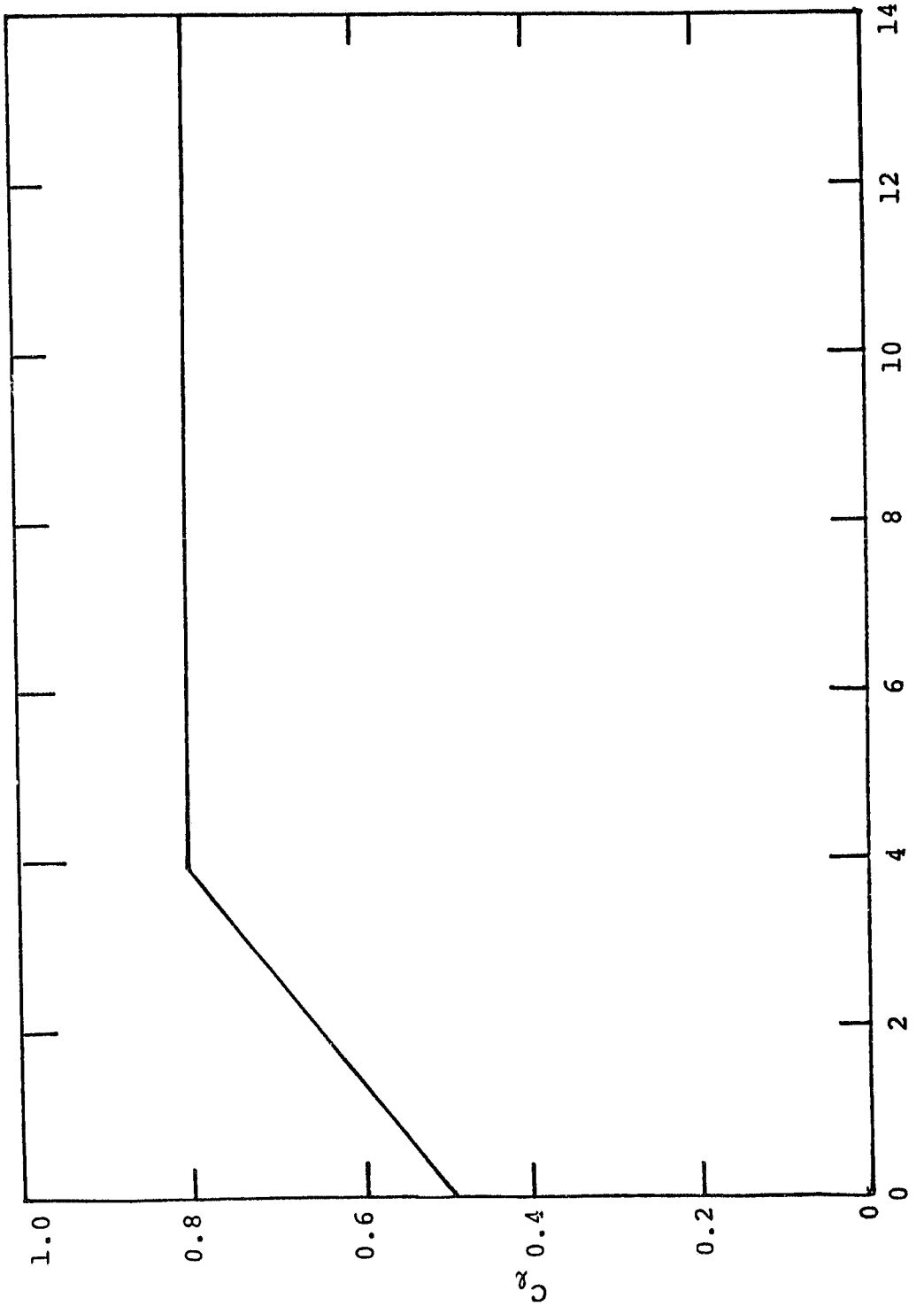


FIGURE 3.4 HYPOTHETICAL RELATION OF AVERAGE BASIN WATER EQUIVALENT WITH LINE CORRELATION FACTOR, C_ℓ.

7.5 Correlation Decay

The correlation decay factor (α) represents how well a point measurement correlates with the surrounding area. Figure 3-5 depicts the decay of the correlation with increasing distance from the measured point for which a correlation of 1.0 is assumed. The correlation decays to e^{-1} at a distance of $1/\alpha$. When distances are expressed in kilometers, if α=0.5 a correlation of 0.37 would exist with a point 2 km away but the correlation would be essentially zero with a point measured 10 km distant. If α=0.1, the correlation decays to 0.82 at 2 km and 0.37 at 10 km.

As shown in Table 3-2, the water equivalents of Line W (3 miles) are generally very well correlated with those for line A (8.45 miles long). During the snow accumulation period, the correlation field is also a function of the storm (precipitation) pattern .

A high variability greatly increases the value of α (a reduced correlation with surrounding points). Since the variability generally increases when melting begins, the value of during snow melting periods should be greater than during the snow accumulation period.

Three basic approaches are described in Chapter 2 to estimate the value of α: from historical data, from real-time data, and with a conceptual model. There are not sufficient data to apply the historical approach and the conceptual approach is difficult to apply since there is no simple relationship between water equivalent values and any other measurable parameters.

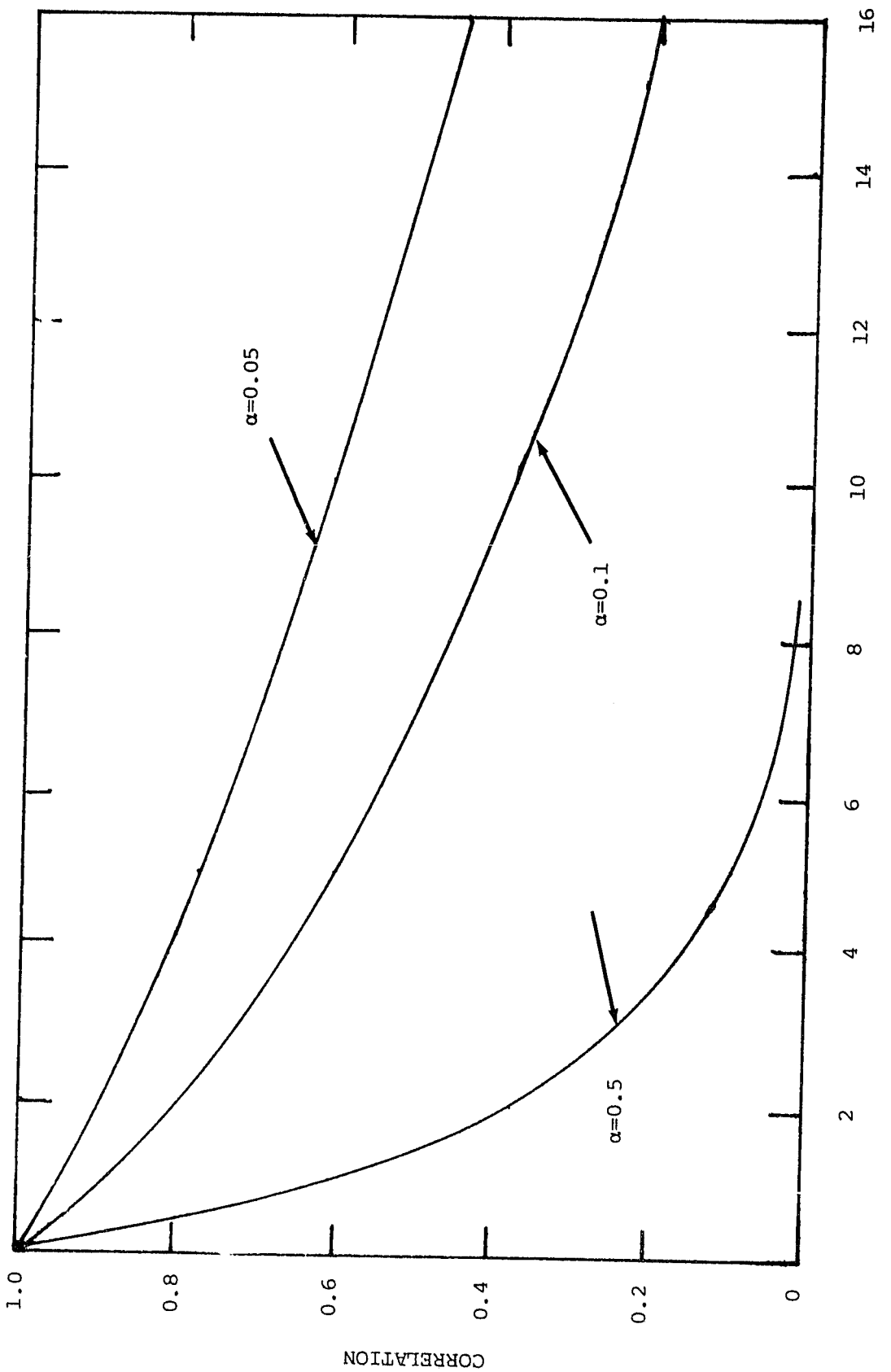


FIGURE 3.5. CORRELATION DECAY WITH DISTANCE FOR VARIOUS VALUES OF CORRELATION DECAY FACTOR, α.

The use of real-time data to estimate the value of for water equivalent measurements offers the best approach. Even this approach has little hope of success until remote sensing can provide a map of water equivalents from which α can be computed.

Until real-time data are available to compute α, an estimate based on field experience will be necessary. For distances measured in kilometers, a value of $\alpha=0.05$ is suggested for the North Central Plains region during the snow accumulation period. During the snow melting period, α should be much larger, increasing to 0.5 or greater as the snow melting nears its end.

8. SOIL MOISTURE

Soil moisture is generally considered one of the most variable of all hydrologic parameters. It is also the least measured of the variables of importance for hydrologic modeling. There is no national network of established measuring stations as there is for precipitation, evaporation, or even water equivalent of the snow cover.

A large part of the variability of soil moisture is due to the large variability of soil types (13).

Most measurements of soil moisture are those made for agricultural research studies. The aerial gamma radiation measuring program of the NWS has provided line measurements of soil moisture for the same survey lines as used for snow cover measurements.

The soil moisture measurement that most closely represents the soil moisture state(s) of a hydrologic model is a measure of the absolute amount of moisture in soil at a particular depth. For such a measurement, it is necessary to know the depth of the sample precisely as well as the weight of the moisture. Such measurements are nearly impossible for the top few centimeters of the soil under actual field conditions because of agricultural debris.

The most common measurement is the percentage of moisture by weight, SM_w, with respect to the weight of the dry soil. The bulk density ρ_b, of the soil [dry mass of a known volume of soil expressed as weight (in grams) per unit volume (in cubic cm)] can be used to relate the percentage of soil moisture by weight, SM_w, to the percentage of the volume, SM_v, by the equation:

$$SM_v = \frac{\rho_b}{\rho_w} SM_w \qquad (3.4)$$

where ρ_w is the density of water (considered equal to unity) and ρ_b is the bulk density of the soil. The equivalent depth of soil moisture for a specific depth of soil h(cm), may be calculated by the equation:

$$\text{Soil Moisture (cm)} = \frac{\rho_b}{\rho_w} = h\,(SM_w) \qquad (3.5)$$

The SCS publishes information, including bulk density by specific depths, for soil types.

Almost all available data on soil moisture are expressed as percentage by weight, SM_w . All further references to

percentage of soil moisture in this report will be moisture percentage by weight. The actual use of the information for relating the soil moisture to the hydrologic models should account for the variation in bulk density of the soils involved.

Changes in the bulk density of the surface layer of the soil occur, especially under frozen conditions. Use of the soil moisture measurements without correcting for changes in bulk density to determine the amount of water in the layer of soil under such conditions can result in a significant error (14).

8.1 Dry Portion of Basin

The dry portion of the basin has been defined for this study as the pervious area of the basin minus the additional impervious area. It represents those areas in which the soil allows water to infiltrate sufficiently to generally avoid ponding on the surface and or those areas in which the slope is sufficient for excess water to drain to other areas (landscape integration).

Areal averages that are based on random sampling of soil moisture have a high variability. A large portion of the variability is often the result of the inclusion of measurements from drainage pathways that persistently maintain near saturated conditions close to the surface of the ground. During periods of rainfall, these drainage pathways can quickly become saturated and become the additional impervious area of the basin .

During the research phase of the NWS aerial gamma radiation survey technique program, a large number of ground surveys of soil moisture were made for calibration purposes (10). These

studies demonstrated a consistency (temporal and spatial) among soil moisture measurements obtained from the dry portion of a basin.

Figure 3.6 is a plot of average soil moisture measurements for Line A near Luverne, Minnesota, for three dates during the winter of 1969 - 1970. A fairly good relationship is shown for the soil moisture along the line and among the three surveys. Filippova (15) proposed excluding data from drainage pathways in computing average soil moisture on experimental catchments to obtain a more accurate average of the dry portions of the basin.

The statistical characteristics (C_p, C_ℓ, C_a, α and σ) discussed in the following sections are for the dry area of the basin as defined previously in this report.

8.2 Point Measurements

The accuracy of a point measurement of soil moisture for the upper few centimeters of the soil depends upon the measurement technology used. Several type of sensors are in use, i.e., neutron tubes and those based on electrical resistance and heat conductivity; most are difficult to maintain and operate.

The gravimetric method of observing the weight of a soil sample before and after heating to a temperature of 105°C for several hours is considered as a standard. The accuracy of a gravimetric measurement under laboratory conditions is well within ±1 percent.

The data available for operational use would come from a variety of sensors and observers. Thus, a value for C_p to use in

ORIGINAL PAGE IS
OF POOR QUALITY

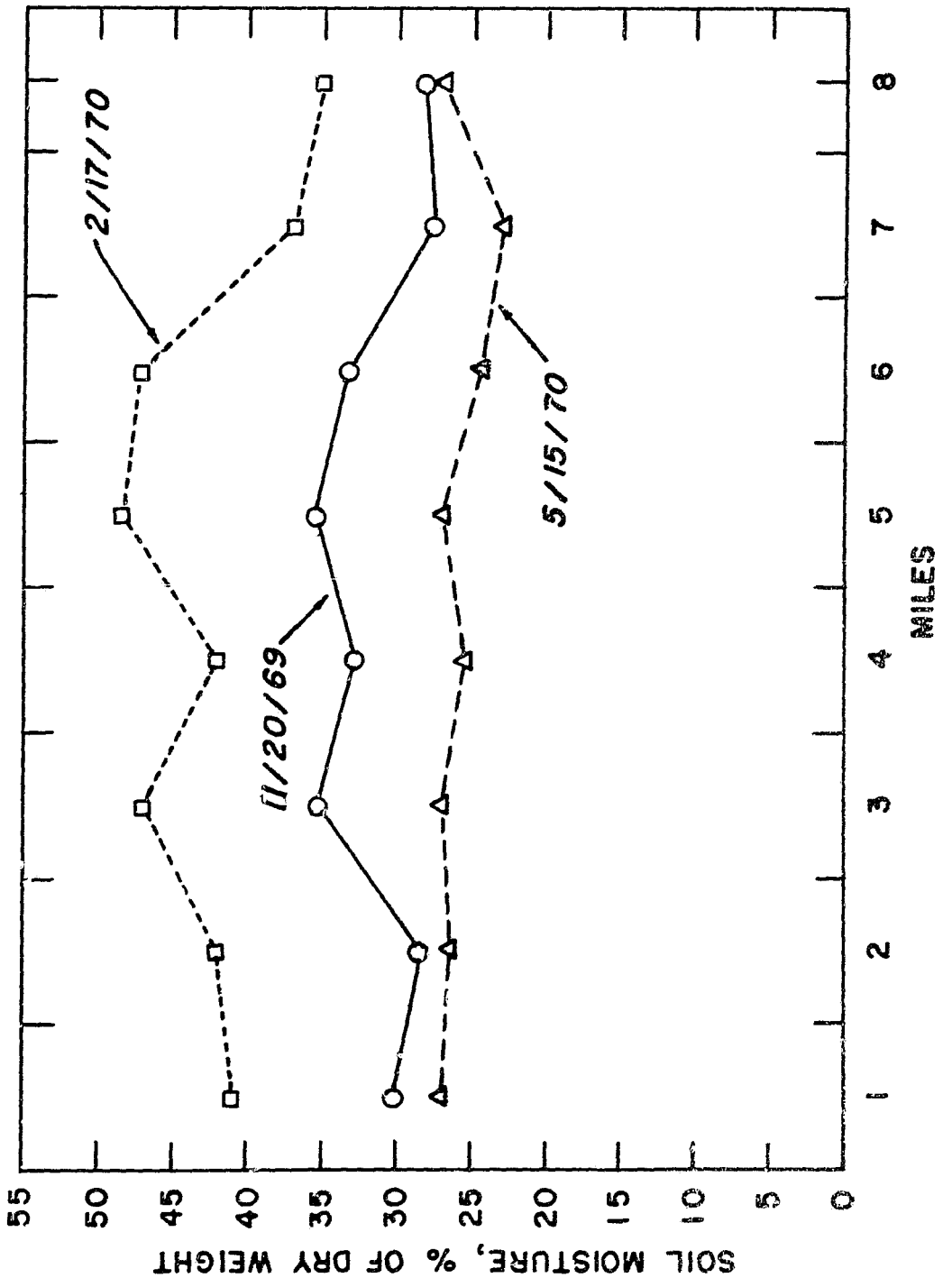


FIGURE 3.6. Average Soil Moisture for Eight Mile Line Near Luverne, Minnesota, 1969-1970 Season

the correlation area method is impossible to determine explicitly. A value of 0.90 is suggested for C_p until actual field operations can provide an improved value.

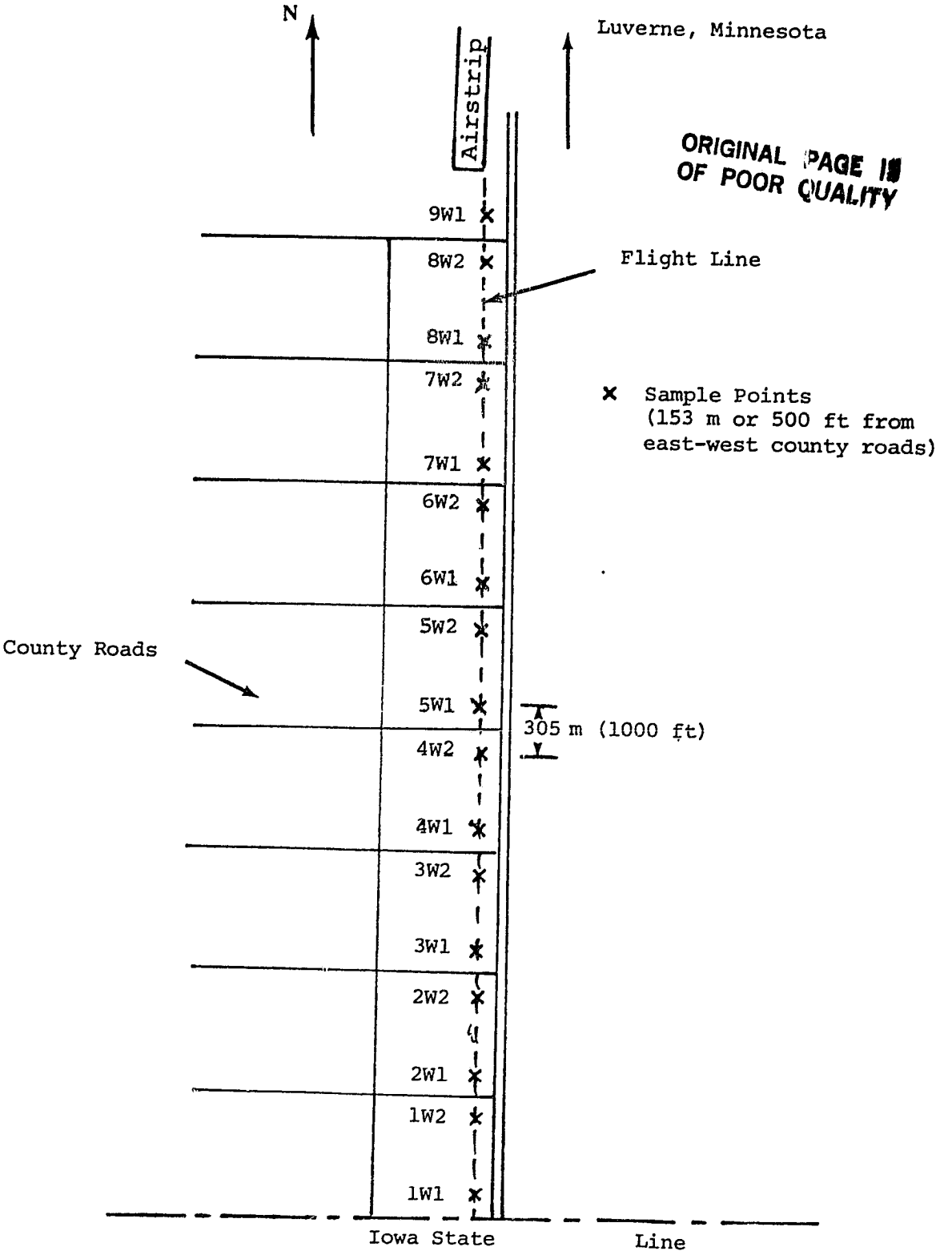
8.3 Line Measurements

The NWS gamma radiation program has provided a unique set of soil moisture measurements for determining a value for C . A considerable amount of the data available are for the 8.45 mile research line (Line A) near Luverne, Minnesota. The line is parallel (305 m or 1000 ft west) to State Highway 75 between the Luverne airport and the Iowa state line. A network of 17 sampling points was established in 1969 as shown on Figure 3.7.

The area covered by the flight line is agricultural and is typical of much of the North Central Plains region. Most of the land, except for the extreme northern portion of the flight line, would be considered to be part of the dry portion of a basin. Small ephemeral stream beds cross the flight line and the area of these can be considered as part of the wet (additional impervious) portion of the basin.

Soil moisture measurements are available for 23 days between November 1969 and February 1976 when research surveys were conducted at the originally established sample points [152 m (500 ft) north and 152 m 500 ft south of each crossroad as shown on Figure 3-7)].

Six additional soil moisture surveys are available since the NWS began an operational snow cover and soil moisture measurement program in 1979 (10).



LOCATION OF SAMPLING POINTS

FIGURE 3-7. RESEARCH FLIGHT LINE NEAR LUVERNE, MINNESOTA.

The data from the Luverne research line have been used to investigate the spatial correlation of soil moisture. The lower 7 miles of the flight line is typical of the slightly rolling farm land of the general region. The terrain of mile 8 is fairly flat but is underlined by gravel and is fairly well drained.

The land in mile 9 is typical of the areas that are considered wet land in this report. The soil moisture remains very high except during drought conditions. During periods of rainfall this area quickly become part of the area that supplies runoff directly to streamflow. The soil moisture samples were collected during the research periods by NWS and by soil scientists and other personnel of the SCS. They were analyzed by a private soil testing and engineering laboratory. Soil samples taken during the operational stage since 1979 have primarily been collected and analyzed by personnel of the SCS.

A summary of the soil type and other soil parameters is shown in Table 3.4 for each of the 17 sampling points along the Luverne, Minnesota, flight line.

The field capacity of a soil is the amount of water held in the soil after excess gravitational water has drained away and the downward movement has materially decreased. The value of the field capacity for the silty, loamed soils in the Luverne, Minnesota, area averages about 32 percent.

The measured soil moisture values for the 17 sampling points during the period 1969 to 1976 are shown in Table 3-5. The

TABLE 3.4 SOIL CHARACTERISTICS RESEARCH FLIGHT LINE
Luverne, Minnesota

SAMPLING POINT	SOIL SERIES*	MOIST BULK DENSITY (G/CM3)	SOIL DRAINAGE
1W1	SAC	1.25 - 1.30	Well
1W2	TRENT	1.15 - 1.25	Well
2W1	SAC	1.25 - 1.30	Well
2W2	MOODY	1.25 - 1.30	Well
3W1	WHITEWOOD	- -	Somewhat Poorly
3W2	MOODY	1.25 - 1.30	Well
4W1	TRENT	1.15 - 1.25	Well
4W2	MOODY	1.25 - 1.30	Well
5W1	MOODY	1.25 - 1.30	Well
5W2	MOODY	1.25 - 1.30	Well
6W1	MOODY	1.25 - 1.30	Well
6W2	GRACEVILLE	1.10 - 1.25	Well & Moderately Well
7W1	GRACEVILLE	1.10 - 1.25	Well & Moderately Well
7W2	FAIRHAVEN	1.25 - 1.40	Well
8W1	BISCAY	1.20 - 1.30	Poorly
8W2	COMFREY	1.20 - 1.40	Poorly & Very Poorly
9W1	COMFREY	1.20 - 1.40	Poorly & Very Poorly

*Descriptions of soil series are given in Appendix J.

TABLE 3.5 SOIL MOISTURE
PERCENT BY WEIGHT
Luverne, Minnesota

RESEARCH LINE 1969-1970

DATE	FLIGHT LINE MEASURING POINT																	LINE A AVERAGE
	1W1	1W2	2W1	2W2	3W1	3W2	4W1	4W2	5W1	5W2	6W1	6W2	7W1	7W2	8W1	8W2	9W1	
Nov 20, 1969	33.6	26.9	29.4	28.7	36.5	34.2	31.3	33.5	29.6	41.6	35.8	30.8	26.6	28.9	26.3	24.0		32.1
Jan 6, 1970	-	-	-	-	-	41.7	-	-	-	-	-	-	-		-	-		
Jan 17, 1970	41.3	39.6	-	-	62.1	-	44.7	-	39.4	57.5	47.2	-	-	37.0	30.7		42.4	
Mar 5, 1970	47.6	34.4	-	-	54.5	39.8	39.2	-	37.3	50.6	40.1	-	-	36.7	41.0			
Mar 30, 1970	38.5	35.1	40.1	38.1	47.0	40.5	39.5	-	34.4	43.3	37.5	32.7	49.0	30.5				
Apr 28, 1970	27.8	19.7	25.7	11.6	12.7	24.6	27.6	-	25.3	22.7	27.4	26.5	-	19.1	21.5			
May 15, 1970	26	28	26	27	28	26	26	25	26	28	22	27	26	20	23	19		26.0
Aug 20, 1970	19.1	17.1	18.3	14.7	21.5	15.7	17.0	16.3	15.7	18.6	18.8	17.9		22.0	10.9			17.5
Nov 5, 1970	34	26	31	30	44	-	37	31	30	44	31	33	29	30		35	43	33.4
Jan 6, 1971	36.9	28.3	34.2	36.2	38.6	29.1	31.1	34.1	30.4	33.7	28.8	33.8	33.8	28.7	31.4	33.3	55.4	33.4
Feb 12, 1971	30.1	24.2	28.0	33.7	32.8	26.4	33.1	30.3	30.4	36.8	23.1	37.5	30.3	24.8	36.7	31.2	76.0	30.7
Mar 23, 1971	41.9	33.2	37.2	44.2	42.4	36.4	36.8	37.6	33.6	41.5	38.4	34.5	36.4	40.1	33.5	36.9	63.8	41.1
Nov 19, 1971	32.2	26.1	32.7	33.8	41.2	29.9	32.4	38.2	31.7	34.4	33.8	32.4	32.3	28.3	33.2	29.9	36.5	32.9
Jan 13, 1972	32.2	27.4	32.4	36.4	34.0	30.0	32.2	32.0	32.8	41.5	34.5	33.8	33.0	31.2	52.2	32.3	65.2	36.0
Feb 12, 1972	29.9	26.8	23.6	31.6	35.6	31.2	25.2	26.6	22.7	40.3	-	33.0	24.3	18.8	33.3	29.3	28.8	29.1
Mar 13, 1972	52.4	43.8	41.0	-	54.4	43.5	63.7	42.8	41.3	43.1	52.7	36.5	47.3	44.5	50.1	41.0		46.7
Mar 30, 1972	28.5	30.6	22.9	33.9	35.1	31.0	30.8	31.2	30.2	35.6	28.7	31.5	32.5	26.2	30.7	24.1	41.0	30.9
Jun 12, 1972	31.0	26.8	30.5	30.6	40.2	31.9	26.6	32.8	26.1	33.0	27.5	27.6	21.3	28.5	25.0	27.0	38.4	30.0
Jun 19, 1972	33.5	29.4	31.8	34.6	48.8	29.4	29.9	28.2	31.2	35.5	28.8	23.1	27.1	28.9	29.8	28.2	43.4	31.9
Mar 6, 1973	57.3	35.9	43.2	50.6	52.0	39.7	56.6	45.4	38.0	44.2	43.3	38.4	40.3	44.6	35.6	35.2	55.5	44.5
Mar 7, 1974	35.9	30.6	32.2	40.4	40.4	33.0	28.7	34.7	33.6	35.2	31.4	33.6	32.9	36.8	25.8	37.2	53.4	35.1
May 12, 1975	29.0	24.1	28.3	-	34.1	25.4	22.5	24.3	26.8	29.0	24.0		26.5	26.1	25.7	24.6	34.7	27.1
Jan 13, 1976	50.7	30.8	35.0	51.0	35.2	34.2	40.8	35.8	27.4	32.9	33.9	57.6	32.6	31.8	42.7	25.7	40.2	37.7
AVERAGE	35.9	29.3	31.2	33.7	39.6	32.1	34.2	32.2	30.6	37.9	32.8	32.7	31.3	31.0	31.9	30.2	47.9	33.1

average value of 47.9 for point 9W1 is a higher than the field capacity value of 32 percent, which is a good indication of the wetness of that location.

Many of the measured values for the other points, especially those taken during February and March, are also higher because of increased soil moisture either from lower levels or melting snow. The additional water above field capacity is held in place (above field capacity) by either frozen ground and or by the temperature gradient in the soil. The temperature gradient results from the cold temperature (at least $0^{o}C$) at the ground surface when snow is present and the mean annual temperature lower in the soil. High values of soil moisture are common in the North Central Plains area during February and March when snow has covered the ground for some time (16).

The correlation among the measured values for the period 1969 to 1976 is also shown in Table 3.6. The values are high indicating that the variability of soil moisture for the dry portion of the basin may not be as great as generally considered.

The cross correlation among the sampling points (except with 9W1) is much greater if the winter months of January through March are not included in the calculations.

The value of C_{ℓ} is dependent upon, the accuracy of the measurement technology and the correlation of a point along the flight line with the true flight line average. The second factor is related to the variability along the flight line and thus will include both the dry and wet areas along the survey line.

TABLE 3.6 CROSS CORRELATIONS, SOIL MOISTURE MEASURING SITES
LUVERNE, MINNESOTA, Nov 1969 - Feb 1976

	1W1	1W2	2W1	2W2	3W1	3W2	4W1	4W2	5W1	5W2	6W1	6W2	7W1	7W2	8W1	8W2	9W1	LINE AVG
1W1	1.0	.80	89	.86	.70	.85	.89	.66	.76	.61	.83	.71	.77	.76	.64	.56	.17	.92
1W2		1.0	.80	.86	.85	.92	.85	.80	.87	.72	.88	.45	.84	.81	.59	.56	.06	.89
2W1			1.0	.77	.77	.85	.82	.90	.86	.66	.84	.51	.76	.89	.59	.68	.37	.92
2W2				1.0	.75	.78	.75	.84	.71	.53	.70	.75	.76	.70	.68	.45	.24	.92
3W1					1.0	.82	.70	.79	.82	.77	.77	.28	.60	.75	.47	.69	-.10	.82
3W2						1.0	.82	.90	.85	.86	.91	.52	.73	.83	.56	.64	.09	.91
4W1							1.0	.84	.83	.65	.89	.54	.87	.77	.66	.61	.27	.90
4W2								1.0	.87	.71	.87	.58	.80	.84	.63	.65	.31	.89
5W1									1.0	.80	.88	.39	.90	.80	.64	.74	.45	.91
5W2										1.0	.81	.39	.48	.59	.55	.65	.14	.78
6W1											1.0	.45	.83	.82	.63	.66	.02	.92
6W2												1.0	.39	.36	.66	.15	.08	.61
7W1													1.0	.82	.64	.71	.54	.98
7W2														1.0	.45	.75	.38	.89
8W1															1.0	.51	.43	.76
8W2																1.0	.58	.76
9W1																	1.0	.45
LINE AVG																		1.0

The accuracy of the aerial gamma radiation technology measurements of the average of the soil moisture along the flightline is reported to be 1.2 cm (0.5 in.) (10). The variability of soil moisture along the flightline is dependent upon the relative portion (and distribution) that is wet land. The variability also is greater during January through March when higher soil moisture values may occur.

Since 1979 the NWS has established additional ground truth soil moisture measurement surveys at other flight lines similar to that at Luverne, Minnesota. In the North Central Plains area, 64 flight lines have had one or more soil surveys conducted, with an average of 16 points per line. An analysis of these data gives the results shown in Figure 3.8. The average monthly standard deviations of the soil moisture measurements for the average flight-line are shown on Figure 3.8. The average is higher (7.5 percent) during February and March than for April, June, and November (4.9 percent).

An independent set of soil moisture measurements was made on July 20-21, 1982, at an AgRISTAR project in Minnesota (private communication from Dr. Thomas Carroll, NWS). A large number of gravimetric measurements (5 cm or 2 in. in depth) were made in a 7.8 km (3.5 sq miles) area of the Dry Creek Basin of the Cottonwood River Basin. The area is agricultural and similar to that near Luverne, Minnesota, with corn and soybeans the primary crops. The average soil moisture for the gravimetric measurements was 16.6 percent with a standard deviation of 4.86 percent,

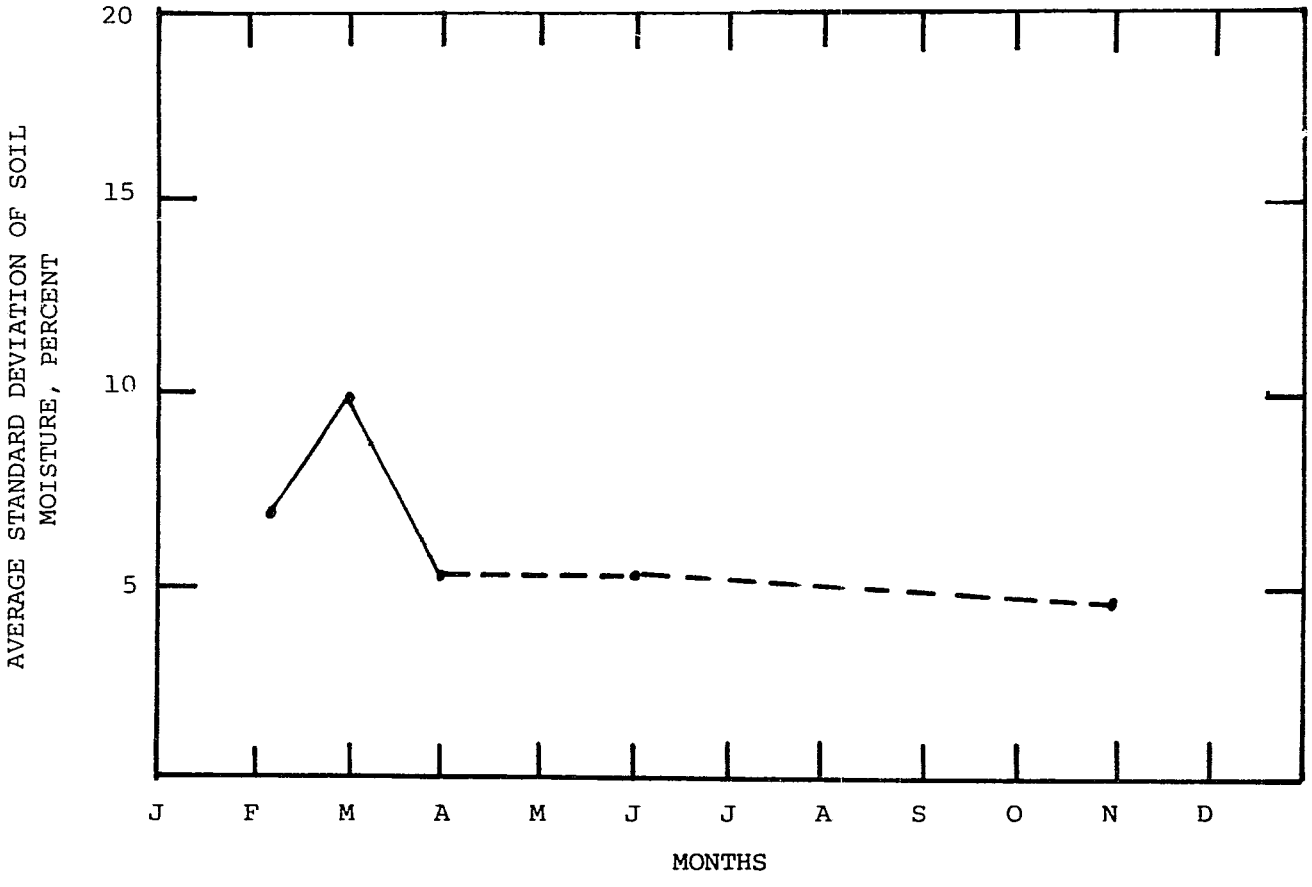


FIGURE 3.8 MONTHLY VALUES OF THE STANDARD DEVIATION OF SOIL MOISTURE FOR CALIBRATION FLIGHT LINES.

supportive of the 4.9 percent standard deviation indicated from the ground surveys of the aerial gamma radiation program.

Although no objectives analysis for C_ℓ is possible, the above data on soil moisture variability and correlation serve as guidelines for an estimate. The average cross-correlation values among the sampling points of the research line at Luverne is on the order of 0.65.

The rightmost column of Table 3.6 shows the correlation of each sampling point with the line average for those years during which a line average was computed. These are high values, with an average of 0.84. The 17 point along the line are considered fairly representative of most agricultural areas in the North Central Plains region with about 6 percent of the area being wet land (1 point in 17).

The average error for the gamma radiation technology in measuring the average soil moisture of a flight line has been reported as ±3.9 percent (17). Thus, a value for C_ℓ of approximately 0.75 is recommended. As additional field data become available from the NWS operational program, the value of C_ℓ could be refined for use with the correlation area method.

The use of microwave equipment for remote sensing of soil moisture is in research and rapidly approaching an operational state. Peck, Larson, et al. (18), present data comparing gamma and passive microwave soil moisture measurements. Their results indicate that the microwave measurements were less correlated to soil moisture variations than were the gamma radiation measurements. However, a positive correlation of 0.82 to 0.84

exists between the two measurement methods. This high correlation suggests that C_ℓ for microwave should be reduced by 20 to 25 percent over values for the gamma measurements.

8.4 Correlation Decay

The intercorrelations among the sampling points along the Luverne, Minnesota, flight line (Table 3.6) provide a means to obtain a value for the correlation decay, α, for soil moisture.

Figure 3.9 is a plot of correlations between points with distance except for correlations with point 9W1 (wet land). The plot represents the distribution of correlation with distance for the dry portion of the basin if the assumption as to the percentage of wet land (about 6 percent) is correct.

A review of the plotted values on Figure 3.9 indicate that a single value of α may not be appropriate to represent the correlation decay for soil moisture for the dry portions of a basin. The appropriate value appears to be inversley proportional to area size. A value of 0.3 to 0.5 for α may be required for very small areas (less than 1 sq km) and decreasing to as low as 0.06 as shown by the curve on Figure 3.9 for estimating areal averages for large areas such as 2,600 sq km (1,000 sq mi).

The variability of the soil moisture (higher correlation for same distance) has been shown to be less when soil moisture values are lower than field capacity (no frozen ground, no water standing on the soil surface, and no large sustained temperature gradient under the snow cover). As data become more available from the NWS program, a procedure to adjust α based on the average value of the soil moisture could be developed.

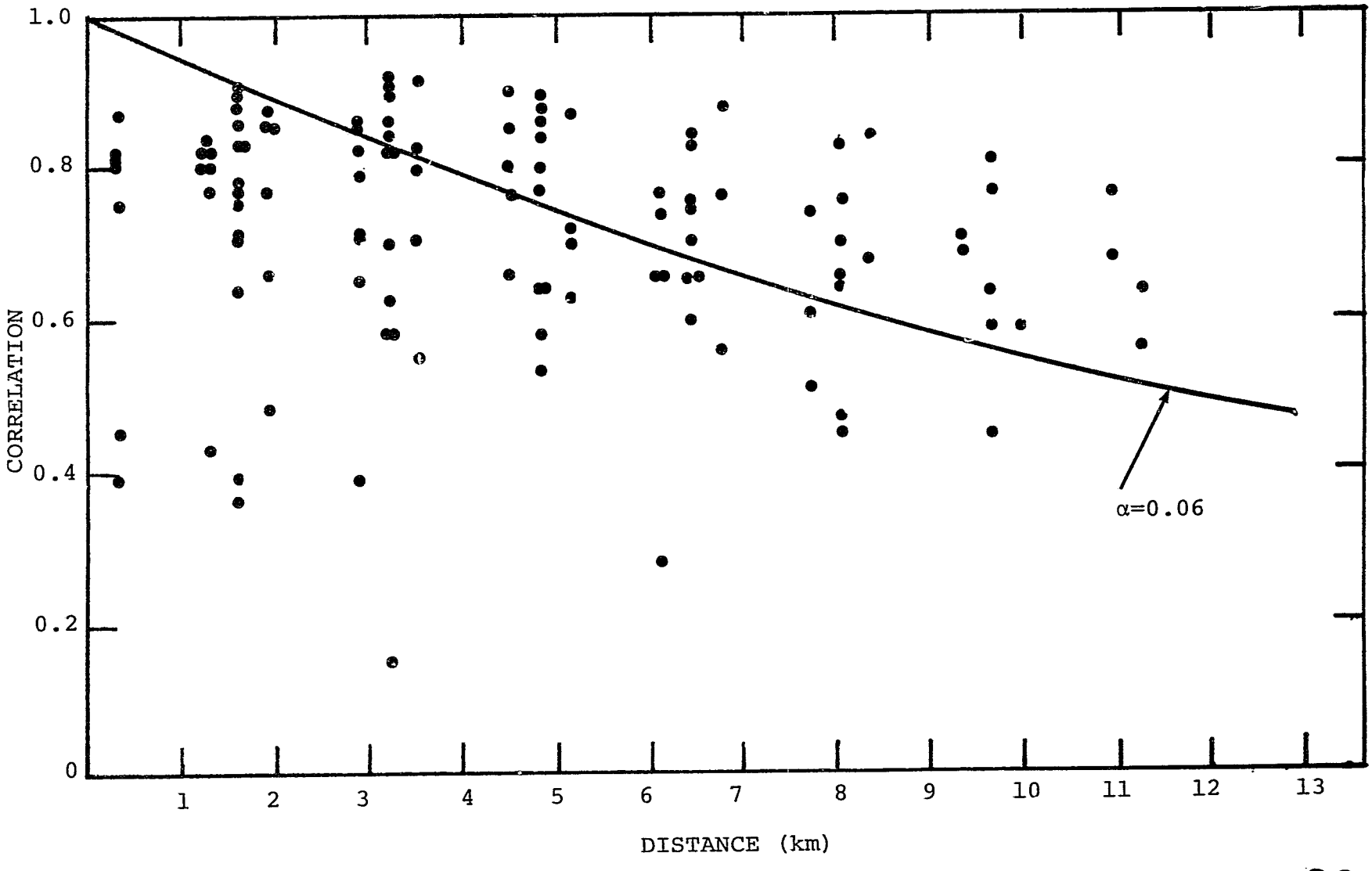


FIGURE 3.9 CORRELATION OF SOIL MOISTURE MEASUREMENTS WITH DISTANCE, LUVERNE, MINNESOTA (DRY PORTION OF BASIN)

8.5 Areal Measurements

At this time, there are no completely satisfactory remote sensing systems for making large-scale areal average soil moisture measurements. However, a variety of experiments have been conducted indicating that such measurements are highly feasible. A short summary of areal measurement techniques is presented here along with a discussion of the potential accuracy of the methods.

Aside from the gamma radiation measurement techniques discussed earlier, only three other methods appear feasible for soil moisture determination. These three methods are active microwave, passive microwave, and thermal infrared measurements. The relationship of all three measurements to soil moisture has been extensively researched.

Schmugge (19) presents an excellent summary of the current capabilities of passive microwave systems. The sensitivity of the soil's emissivity to its moisture content has been demonstrated with radiometers operating from gound-based, aircraft, and spacecraft platforms. This dependence of the soil's microwave emissivity on its moisture content is due to the large contrast between the dielectric properties of free water and those of dry soil. The dielectric constant of a soil-water mixture is predominantly influenced by the amount of water in the mixture. As a result, changes in soil emissivity are produced, varying from 0.95 for dry soils to 0.6 or less for wet soils.

Passive microwave radiometers measure the thermal radiation emitted by the soil. This radiation is generated within the

volume of the soil and is dependent on the moisture (i.e.,dielectric) and temperature profiles in the soil.

Passive microwave observations are effected both by vegetation cover and soil surface roughness. Schmugge states that the magnitude of the effect depends on the amount of vegetation and the wavelength of observation. A thick canopy would approximate a Lambertian black body, i.e., it would have an emissivity close to unity and show no angular or polarization effects. Basharinov and Shutko (20) and Kirdiashev et al. (21), have reported on observations made in the USSR over the 3- to 30-cm wavelength range for a variety of crops. Their results indicate that for small grains and grasses, the sensitivity to soil moisture is 80 to 90 percent of that expected for bare ground at wavelengths greater than 10 cm. Broad-leaf cultures, like mature corn or cotton, transmit only 20 to 30 percent of the radiation from the soil at wavelengths shorter than 10 cm and about 60 percent at the 30-cm wavelength. They observed 30 to 40 percent sensitivity for a forest at the 30-cm wavelength although they did not mention the type or height of trees. These results could be applied in determining α for the correlation area method.

In his summary, Schmugge states that surface roughness also tends to mask true soil moisture conditions, having its greatest effect on wet soil. Surface roughness can increase emissivity can be increased by 0.2 for very wet soils and only 0.02 for dry soils, resulting in decreased sensitivity of microwave emissivity to soil moisture variations. To effectively deal with the

problem of surface roughness effectively, it will be necessary to learn more about the range of roughness that occur in nature.

Passive microwave measurements from aircraft have the best potential for providing data sources for testing areal measurements in the correlation area method in the near future. Schmugg presents results from aircraft flights over the northern end of California's Imperial Valley. The brightness temperatures (T_B) of the individual fields were determined by averaging data acquired during the 3-second interval that the aircraft was over each field. These T_B values were compared with ground measurements of soil moisture typically made at four points in each field. The correlations between the T_B values and soil moisture in the top 2.5 cm of soil were generally greater than 0.8. The difficulty in making accurate ground measurements has hampered the determination of the accuracy of remote sensing techniques.

Some attempts have been made to correlate soil moisture with data from very coarse resolution satellite data. As reported by McFarland (22) and by Eagleman and Lin (23), the sensitivity of the 21 cm radiometer to soil moisture has been demonstrated from space during the SKYLAB mission. McFarland showed a definite relationship between the SKYLAB 21-cm brightness temperatures and the Antecedent Precipitation Index (API).

Eagleman and Lin compared the brightness temperature with estimates of the soil moisture over the radiometer footprint. The soil moisture estimates were based on a combination of actual ground measurements and calculations of the soil moisture using a

climatic water balance model. The correlation coefficient for 12 data points was 0.96, which is very good considering the difficulty of obtaining soil moisture information over a footprint of such a size and considering the fact that the brightness temperature was averaged over the wide range of cultural conditions that occurred over the area. Such results should, however, be viewed with some skepticism.

At this time, the potential of active microwave measurements is known, but the technique is less well developed than is the passive microwave technique. Considerable research published by Schmugge (12) among others indicates that active microwave sensors will offer better sensitivity under vegetative cover and better spatial resolution. A number of experiments from both aircraft and space vehicles will be required before such data can be accurately evaluated for use in models. Preliminary results indicate correlations of 0.8 or better between the reflected energy and soil moisture.

The final potential source of areal estimates of soil moisture is thermal infrared data. The technique for collecting such data is dependent on the measurement of the daily variation in soil surface temperature. The temperature variation is a function of the soil thermal mass, which is a function of the soil moisture content. Price (24) summarizes some current satellite capabilities and the means for using the data. The Heat Capacity Mapping Mission (HCMM) satellite and a similar one, TIROS-N, were launched in 1978. Both satellites acquire high-resolution thermal infrared data (10.5-12.5 km) at times of day

(early morning and early afternoon) more optimal for measuring daily maximum and minimum surface temperature. The HCMM has a ground resolution of 0.5 km, and a system noise figure of 0.4 degrees Kelvin.

Price describes how a numerical model may be used with satellite and other data to estimate evaporation and soil moisture. This procedure is, however, expensive both in data and computer time. The model computes the diurnal variation of surface temperature for sets of conditions representing variations in soil moisture and near-surface relative humidity. Agreement between theory and modeled results is reported to be good.

It seems possible that TIROS-N or HCMM data could be used in preliminary applications or tests of the correlation area method and the NWSRFS models. Little or nothing, however, is known about correct values for C_a from the method.

8.6 Standard Deviation of Random Field

Three possible approaches for estimating the standard deviation, σ, of a random field of a variable are the use of historical data (multiple-realization approach), real-time data (single-realization approach), and a conceptual model.

Of these three approaches, the historical approach and the conceptual model approach offer the most promise.

The standard deviations developed from the NWS operational program (64 flight lines and more than 1000 sampling points) are probably the best set of data that represent the North Cental

Plains region. Other data sets are generally for small areas such as 40-acre plots.

The values of 7.5 percent for the standard deviation in the winter months of January, February, and March and 4.9 percent for the rest of the year can be used as initial values for α (Figure 3.8) based on historical data.

The use of real-time data from satellites (areal mapping of soil moisture for the basin) could be used to compute a value for the standard deviation α. However, at the present time such areal mapped data are not available.

An interesting possibility is the use of a conceptual model approach that would not require large amounts of soil moisture data.

The key to this conceptual approach is to decompose the soil moisture field (x,y) into two components:

$$p(x,y)=f(x,y).r(x,y) \qquad (3.6)$$

where $f(x,y)$= field capacity at (x,y)

$r(x,y)$=fraction of field capacity filled at (x,y)

The field capacity $f(x,y)$ should be available from soil maps and is presumed constant (field capacity is probably greater than

that indicated by SCS during the winter months). Thus, the properties of the field $p(x,y)$ can be determined.

What is required is the covariance of $p(x,y)$ or $Cov_p(x_1,y_1;x_2,y_2)$ given by

$$Cov_p(x_1,y_1;x_2,y_2)=E[p(x_1,y_1)-Ep(x_1,y_1)] \quad (3.7)$$
$$[p(x_2,y_2)-Ep(x_2,y_2)]$$

The first question is whether the function Cov_p can be defined in terms of properties of $f(x,y)$ and $r(x,y)$. For such a definition, it will be necessary to make some assumptions about the relationship of $f(x,y)$ and $r(x,y)$ and about the form of each.

In the absence of any strong evidence to the contrary, the following simple assumptions will be made: $f(x,y)$ and $r(x,y)$ are independent Gaussian random fields with covariance funtion Cov_f and Cov_r, respectively, and homogeneous mean values $\bar{f}=Ef(x,y)$ and $\bar{r}=Er(x,y)$. Thus, the mean value of $p(x,y)$ is $\bar{p}=\bar{f}\cdot\bar{r}$.

The Gaussian property assumed for f and r along with their assumed independence allows the moment-generating function to be employed to derive:

$$\begin{aligned} Cov_p(x_1,y_1;x_2,y_2) &= Ef(x_1,y_1)r(x_1,y_1)f(x_2,y_2)r(x_2,y_2)-\mu^2 \\ &= Cov_r(x_1,y_1;x_2,y_2)\cdot Cov_f(x_1,y_1;x_2,y_2) \\ &\quad +\bar{f}^2 Cov_r(x_1,y_1;x_2,y_2) \\ &\quad +\bar{r}^2 Cov_f(x_1,y_1;x_2,y_2), \end{aligned} \quad (3.8)$$

which is the required relationship between the mean and covariances of $f(x,y)$ and $r(x,y)$ and the required covariance

function of p(x,y). The mean, $\bar{f}$, and covariance, Cov_f, are available from soil information, but what about $\bar{r}$ and Cov_r?

If it is necessary to have point soil moisture data to estimate $\bar{r}$ and Cov_r, then one might as well use that data to estimate Cov_p directly. However, there is evidence of a relationship between soil moisture and the API. It should be possible to locate, historical point soil moisture data at locations that also have a rain gage and thus investigate whether there exists the functional relationship

$$r(x,y)=g[API(x,y)] \qquad (3.9)$$

where g is a function to be determined and

API (x,y) is the antecedent precipitation index.

It is likely that such a relationship can be found to approximate $\bar{r}$ and Cov_r adequately from the API values at multiple sites. Recent work for NASA, for example, has demonstrated that API values can be estimated using passive microwave emissivity (16).

CHAPTER 4

SUMMARY

The material presented in this study is the first step in a program to incorporate remotely sensed data into rainfall/runoff and snowmelt models. A general statistical technique called the correlation area method is presented. It allows proper weighting of a variety of remotely sensed measurements and permits areal estimates, point observations, and line observations to be combined while accounting for measurement accuracy.

The correlation area method makes use of the statistical characteristics of each measurement type in order to properly weight it in the models. The statistics used are:

- the correlation of a point measurement with the true value;
- the correlation of a point in a line measurement and the true line average;
- the correlation of an areal average with the true areal average;
- the rate at which the point or line correlation decays with distance; and
- the measure of randomness in an areal distributed random field.

Guidelines are given for selecting the coefficients for common remotely sensed measurements used in hydrologic modeling. Measurement techniques and their statistical properties are discussed for soil moisture, snow covered area, snow water equivalent, water covered area, and wet area.

There has been no previous evaluation of the usefulness of soil moisture for improving operational hydrologic forcasting. The correlation area method provides a means for evaluating the use of soil moisture in hydrologic models. In addition, it offers a unique technique for evaluating the usefulness of remotely sensed hydrometeorological measurements for hydrologic modeling.

The data base for computing the statistical parameters required with the correlation area method for soil moisture and water equivalent of the snow cover is extremely limited. The error in one or more of the parameters could be very large. Intuitively, the accuracy of the parameters may not be too critical. The value of the correlation area method lies in the way in which the measurements are weighted in computing the average areal values. It may not be too critical if an area of 13 or 26 sq km (5 or 10 sq miles) is assigned to a point measurement for a 1300 sq km (500 sq miles) basin with five point measurements. In such a case, most of the area would be assigned to the remotely sensed measurements. Even with a small correlation value the areal measurement might greatly improve the estimate based on point measurements alone. Until actual field

tests are conducted on the correlation area method (as proposed for the third phase of this study), there is no objective method to determine the added advantage of combining the various measurements. The field testing of the updating procedures using the correlation area method could also provide more information on the natural variability of the soil moisture and water equivalent variables than probably could be obtained by direct research studies.

ACKNOWLEDGEMENTS

The Authors greatefully acknowledge the leadership and support provide by Dr. Albert Rango (NASA), and the assistance given by Dr. Thomas Schmugge (NASA), and Dr. Thomas Jackson (ARS) who served with Dr. Rango as a review team for NASA.

Thanks are also extended to the staff of the Hydrologic Research Laboratory (NWS) for the assistance and cooperation provided by them; Dr. Eric Anderson's advice on the NWSRFS models was extremely valuable and the reviews by Mr. George Smith and Dr. Konstantine Georgakakos were very helpful.

REFERENCES

1. R. L. Bras and Ignacio Rodriguez-Iturbe, "Evaluation of Mean Square Error Involved in Approximating the Areal Average of a Rainfall Event by a Discrete Summation," Water Resources Research, 12(2), 1976, pp.181-183.

2. R. L. Bras and Ignacio Rodriguez-Iturbe, "Network Design for the Estimation of Areal Mean of Rainfall Events," Water Resources Research, 12(6), 1976, pp.1185-1195.

3. E. L. Peck, R. S. McQuivey, T. N. Keefer, E. R. Johnson and J. L. Erekson, "Review of Hydrologic Models for Evaluating Use of Remote Sensing Capabilities," NASA CR166674, Goddard Space Flight Center, Greenbelt, Maryland, March 1981, pp.99.

4. E. L. Peck, T. N. Keffer, and E. R. Johnson, "Strategies for Using Remotely Sensed Data in Hydrologic Models," NASA CR166729, Goddard Space Flight Center, Greenbelt, Maryland, July 1981, pp.84.

5. T. J. Jackson and R. H. McCuen, "Accuracy of Inpervious Area Values Estimated Using Remotely Sensed Data," Water Resources Bulletin, AWRA, Vol. 15, No. 2, April 1979, pp.436-446.

6. NASA, "Applications Systems Verfication and Transfer Project," NASA Technical Paper, December 1981, pp.1822-1826.

7. D. R. Wiesnet, D. F. McGinnis, and T. A. Pritchard, "Mapping of the 1973 Mississippi River Floods by the NOAA-2 Satellite," Water Resources Bulletin, Vol. 10, No. 5, October 1974, pp.1040-1049.

8. G. K. Moore and G. W. North, "Flood Inundation in the Southeastern United States from Aircraft and Satellite Imagery," Water Resources Bulletin, Vol. 10, No. 5, October 1974, pp.1082-1096.

9. E. L. Peck, "Snow Measurement Predicament," Water Resources Research, Vol. 8, No. 1, February 1972, pp.244-248.

10. T. R. Carroll and K. G. Vadnais, "Operational Airborne Measurement of Snow Water Equivalent Using Natural Terrestrial Gamma Radiation," Proceedings, Western Snow Conference, Laramie, Wyoming, April 1980, pp.97-106.

11. Proceedings WMO/IAHS Workshop on Remote Sensing of Snow and Soil Moisture by Nuclear Technique, VOSS, Norway, 23-27 April 1979.

12. T. J. Schmugge, "Microwave Approaches in Hydrology," Photogrammetric Engineering and Remote Sensing, Vol. 46, No. 4, April 1980, pp.495-507.

13. S. G. Reynolds, "The Gravimetric Method of Soil Moisture Determination; Part I: A Study of Equipment, and Methodological Problems; Part II: Typical Required Sample Sizes and Methods of Reducing Variabilities; Part III: An Examination of Factors Influencing Soil Moisture Variability," Journal of Hydrology, Vol. 11, North-Holland Publishing Co., Amsterdam, 1970.

14. N. Y. Leonova, "Estimation of the Error in the Determination of Water Storage at Observation Points," Trans. Trudy GGI No. 233, pp.118-126, Pub. Soviet Hydrology, Selected Papers, Vol. 16, No. 2, 1977, pp.98-107.

15. A. K. Filippova, "Results of Soil Moisture Surveys on Catchments," Trans. Trudy GGI, No. 92, pp.104-118, Pub. Soviet Hydrology, Selected Papers, Vol. 4, 1964, pp.351-363.

16. E. L. Peck, "Effect of Snow Cover on Upward Movement of Soil Moisture," Journal of the Irrigation and Drainage Division, ASCE, Vol. 100, No. IR4, December 1974, pp.405-412.

17. T. R. Carroll, "Airborne Soil Moisture Measurement Using Natural Terrestrial Gamma Radiation," Soil Science, Vol. 132, No. 5, Nov 1981.

18. E. L. Peck, L. W. Larson, R. K. Farnsworth, and T. L. Dietrich, "Comparison of Aerial Passive Gamma and Passive Microwave Techniques for Measurement of Soil Moisture," Proceedings, 10th International Symposium on Remote Sensing of Environment, October 1975, Research Institute of Michigan, Ann Arbor, pp. 1235-1243.

19. T. J. Schmugge, "Soil Moisture Sensing with Microwave Radiometers," Proceedings, Machine Processing of Remotely Sensed Data, June 1980, pp.346-354.

20. A. Ye Basharinov and H. M. Shutko, "Determination of the Moisture Content of the Earth's Cover by Superhigh Frequency (Microwave) Radiometric Methods," A Review. Radiotekhnika i Electronika, Vol. 23, pp.1778-1791. NASA Tech. Trans. TM-75411. Radio Engineering and Electronic Physics, 23 September 1978, pp.1-12, translation.

21. K. P. Kirdiashev, A. A. Chukhlantsev, and A. M. Shutko, "Microwave Radiation of the Earth's Surface in the Presence of Vegetation Cover." Radiotekhnika i Elektronika, Vol. 24, pp.256-264. NASA Tech. Trans. TM-75469.

22. M. J. McFarland, "The Correlation of Skylab L-Band Brightness Temperatures with Antecedent Precipitation," Preprints Conf. on Hydrometeorology, Fort Worth, Amer. Meteor. Soc., 60-65. 1976.

23. J. R. Eagleman and W. C. Lin, "Remote Sensing of Soil Moisture by a 21 cm Passive Radiometer," J. Geophys. Res., 81, 3660-3666, 1976.

24. J. C. Price, "The Potential of Remotely Sensed Thermal Infrared Data to Infer Surface Soil Moisture and Evaporation," Water Resources Research, Vol. 16, No. 4, August 1980, pp.787-795.

C-2

APPENDIX A

BASIN AVERAGES BY OBJECTIVE ANALYSIS

Consider a basin with area B and a random field $p(x,y)$ representing the variable of interest defined over the two spatial dimensions x and y. An estimate $\hat{P}$ is desired of the areal average of $p(x,y)$ over the basin area. The random field $p(x,y)$ could be soil moisture or snow water equivalent or any variable of interest. The true value of the areal average P is given by equation A.1

$$P = \frac{1}{B} \iint_B p(x,y)\ dxdy \qquad (A.1)$$

The estimate $\hat{P}$ will be formed as a weighted average of various measurements m_i in the form of equation A.2.

$$\hat{P} = \Sigma \lambda_i m_i \qquad (A.2)$$

Two obvious questions arise: What is the accuracy of $\hat{P}$? What should the weights be? The mean square error of estimate MSE can be formed as defined by equation A.3.

$$\begin{aligned} MSE &= E(\hat{P}-P)^2 \qquad (A.3) \\ &= E\ \hat{P}^2 - 2E\hat{P}P + EP^2 \\ &= \Sigma\Sigma \lambda_i \lambda_j Em_i m_j - 2\Sigma \lambda_i Em_i P + EP^2 \end{aligned}$$

where E is the expected value operator.

The weights (λ) can be defined to minimize MSE with respect to the λ's imposing, if desired, the constraint that the λ's sum to 1.0 (an unbiasedness constraint). The solution is a straightforward application of Lagrange multipliers. Introducing the Lagrange multiplier γ for the unbiasedness constraint produces n+1 linear equations in n+1 unknowns - the λ's for each of the n measurements plus the Lagrange multiplier γ:

$$2\Sigma_j \lambda_j Em_i m_j - 2EPm_i + \gamma = 0 \quad i=1,\ldots,n \tag{A.4}$$

$$\Sigma \lambda_i - 1 = 0 \tag{A.5}$$

There is certainly nothing new or conceptually difficult in applying equations A.3, A.4 and A.5 to the problem at hand (see for example Bras 1976a, 1976b).* Further, this technique would meet all the requirements of this study. The accuracy measure is given by the MSE of equation A.3. The accuracy of each measurement is implicit in the second term of equations A.4 The cross-correlations of measurements are included in the first terms of equations A.4.

The difficulty arises when the geometry of the measurements is considered. Each measurement can be represented as

$$m_i = \frac{1}{D_i} \iint_{D_i} p(x,y)\,dxdy + \varepsilon_i \tag{A.6}$$

where D_i is the "domain" of the measurement
ε_i is the measurement error.

*References 1. and 2. in the main text.

In all the applications to date of the objective analysis method described above, the measurements have all been point observations, i.e., the domain of the measurement is a single point (x_i, y_i). Thus, the required integrals, though cumbersome, have not been overbearing to compute. In our case, measurements will include satellite and airborne observations with sampling domains which are themselves areal averages and/or line averages. In this situation the evaluation of the terms $Em_i m_j$ and EPm_i are considerably more cumbersome.

APPENDIX B

BASIN AVERAGES BY THE CORRELATION AREA METHOD

An algorithm is desired to estimate weights (λ's) to be applied to various measurements (m_i's) to produce an estimate of the areal average in the form of equation B.1. (Equation A.2 is reproduced here as equation B.1).

$$\hat{P} = \Sigma \lambda_i m_i \qquad (B.1)$$

As has been shown in Appendix A, the λ's can be found in such a way as to minimize the mean square error (MSE) of the estimate $\hat{P}$, but only at the expense of rather cumbersome evaluation of several integrals. In this Appendix a heuristic approach is developed to estimate the λ's . This heuristic appraoch is called the "correlation area" technique. It does recognize measurement accuracy and the cross-correlation of the measurement. It does produce a measure of the accuracy of $\hat{P}$, but not the mean square error of $\hat{P}$. The computations required by the correlation area approach are considerably simpler than those needed for the objective analysis described in Appendix A, but they are still nontrivial.

The algorithm can be described in steps: The first step is to define the "sample area" S_i for each measurement. The sample area of a measurement is simply the portion of the basin with which that sample is more highly correlated than any other measurement. The sample area for each sampling geometry (point, flight lines, and areal average measurements) is derived from knowledge of the sampling covariance function. The sampling covariance functions are described in Appendix C and the geometries of the sample area of point, line, and areal samples are discussed in Appendices D, E, and F, respectively.

The next step is to define the "correlation area" A_i in equation B.2.

$$A_i = \iint_{S_i} cor_i(x,y)\,dxdy \qquad (B.2)$$

where $cor_i(x,y)$ is the correlation of measurement i with the variable $p(x,y)$

S_i is the sample area of measurement i

A_i is the correlation area of measurement i

thus A_i represents the area of the basin sampled by a perfect measurement with equivalent accuracy to measurement i and A_i can be considered independent of the other measurements since the sample area S_i includes only the portion of the basin which is more highly correlated with measurement m_i than it is with any other measurement.

The weight λ_i for a measurement m_i is found by

$$\lambda_i = \frac{A_i}{\Sigma A_i} \qquad (B.3)$$

thus the weights are forced to sum to 1.0.

Finally a measure of the overall accuracy of $\hat{P}$ is given by ω_m

$$\omega_m = \frac{\Sigma A_i}{B} \qquad (B.4)$$

By its definition $0 \leq \omega_m \leq 1$ with $\omega_m = 1$ representing perfect knowledge of the areal average over the basin. The value of ω_m will be called the "correlation" of $\hat{P}$ with P, using the term correlation in a rather loose sense.

The computational simplification of the correlation area approach is due to the ability to describe geometrically the sample area S_i assigned to each sample. The integrals required in equation B.2 to estimate the correlation area are still non-trivial, but they are considerably simpler than those required by the objective analysis of Appendix A.

The correlation area technique performs in an intuitively correct fashion. If some measurement technology is more accurate relative to another then its sample area S_i and correlation area A_i will increase relative to less accurate measurements, thus it will receive a higher weight. If the random field is highly correlated in space then the correlation function will presumably decay more slowly with distance giving a higher absolute values of the A_i's and a higher accuracy measure ω_m even though the weights themselves may not change. If two measurements are highly correlated with each other as a result of being located close together in space, then the sample area of each will be reduced over what it would be without the nearby sample; thus the correlation area does explicitly consider cross-correlation of measurements. Finally, the accuracy measure ω_m produced by the correlation area approach may not have the theoretical validity of the mean square error estimate produced by an objective analysis, but it is sensitive to the measurement accuracy and the correlation of the random field itself in an intuitively correct fashion and it possesses the convenient property of being bounded by zero and one.

APPENDIX C

SAMPLING CORRELATION FUNCTIONS

In order to implement the correlation area technique it is necessary to define the correlation of each measurement with the random field under study. This correlation function will be noted as $cor_i(x,y)$ for the correlation of i-th measurement with the value of $p(x,y)$, and it should be defined by

$$cor_i(x,y)=\frac{E(p(x,y)-\mu)(m_i-\mu)}{E(m_i-\mu)^2} \qquad (C.1)$$

where $\mu=Ep(x,y)$ is the mean value.

In fact the definition C.1 has introduced the first of many assumptions that will appear in this appendix, namely that the mean value of $p(x,y)$ is homogeneous. It is necessary to develop estimates of $cor_i(x,y)$ for each of three sampling geometries: point samples, line samples, and areal-average samples.

POINT SAMPLES

A sample m_i taken at a single point (x_i,y_i) will be regarded as*

$$m_i=p(x_i,y_i)+\varepsilon_i \qquad (C.2)$$

where ε_i=measurement error for measurement i

Further it will be assumed that

$$E\varepsilon_i=0$$

$$E\varepsilon_i^2=\sigma_i^2$$

i.e. ε_i is zero mean white noise.

---

*Note: The notation m_i and σ_i^2 will be used throughout regardless of sampling geometry. It is hoped that the exposition will make clear the distinction between point, line, and areal samples.

Introducing C.2 into C.1 gives

$$cor_i(x,y)=\frac{E(p(x,y)-\mu)((p(x_i,y_i)-\mu,+\varepsilon_i)}{E((p(x_i,y_i)-\mu)+\varepsilon_i)^2} \qquad (C.3)$$

using the assumption that ε_i is white noise, and introducing the notation $cov(x,y;x_i,y_i)$ for the covariance of the field itself gives

$$cor_i(x,y)=\frac{cov(x,y;x_i,y_i)}{\sigma^2+\sigma_i^2} \qquad (C.4)$$

where σ^2=variance of P

$$=E\ (p(x,y)-\mu)^2$$

$$cov(x,y;x_i,y_i)=E(P(x,y)-\mu)(P(x_i,y_i)-\mu)$$

Assuming that the random field P has homogeneous isotropic covariance gives

$$cov(x,y;x_i,y_i)=\sigma^2 cor(d) \qquad (C.5)$$

where $d=((x-x_i)^2+(y-y_i)^2)^{\frac{1}{2}}$ (C.6)

and cor(d) is the correlation function of the random field P.

Further, assuming that the form of cor(d) is a simple exponential with parameter α, and combining C.6, C.5, C.4 gives

$$cor_i(x,y)=\frac{\sigma^2}{\sigma^2+\sigma_i^2}\exp(-\alpha d) \qquad (C.7)$$

ORIGINAL PAGE IS OF POOR QUALITY

Introducing the notation C_p for the correlation of a point measurement with the true value at that point gives

$$cor_i(x,y)=C_p\exp(-\alpha d) \tag{C.8}$$

for the correlation of a point sample at (x_i,y_i) with the value at (x,y) with d defined by C.6.

Thus, there are two parameters required to describe the correlation function for point samples:

C_p - The correlation of a point sample with the true value at that point.

α - The correlation decay parameter of the random field under study.

LINE SAMPLES

If the "correct" form of a line sample measurement is inserted into equation C.1 then the measurement m_i becomes

$$m_i=\frac{1}{D_i}\iint_{D_i} p(x,y)\,dxdy+\varepsilon_i \tag{C.9}$$

with D_i the flight line, and the integral is then an average over the flight line. If equation C.9 is inserted into equation C.1 the result is taking expectations over exactly the type of integrals that the correlation area approach was designed to avoid. To avoid this sad state of affairs, the sampling covariance function for line samples will be assumed to take the form

$$cor_i(x,y)=C_\ell\exp(-\alpha d_\ell) \tag{C.10}$$

where d_ℓ is the distance from (x,y) to the flight line.

α is the same decay parameter as before, i.e. the decay parameter applicable to point values of the random field P.

C_ℓ is the correlation of m_i with a point value on the flight line itself.

At least some justification for the assumptions inherent in equation C.10 can be provided. Consider the point (x^*,y^*) located on the flight line at the point closest to (x,y). The best linear estimate of $P(x^*,y^*)$ based on the measurement m_i is $\hat{P}(x^*,y^*)$ given by

$$\hat{P}(x^*,y^*)=C_\ell(m_i-\mu)+\mu \qquad (C.11)$$

and the best estimate of P(x,y) given $P(x^*,y^*)$ is

$$\hat{P}(x,y)=\exp(-\alpha d_\ell)(P(x^*,y^*)-\mu)+\mu \qquad (C.12)$$

Replacing $P(x^*,y^*)$ in equation C.12 with $\hat{P}(x^*y^*)$ from equation C.11 gives

$$\hat{P}(x,y)=C_\ell \exp(-\alpha d_\ell)(m_i-\mu)+\mu, \qquad (C.13)$$

which shows that the form of the sampling correlation function for line samples assumed in C.10 is correct.

AREAL SAMPLES

Consider Figure C.1 which shows an areal sample covering domain D_i, the basin with area B, and the portion of the sample inside the basin $B \wedge D_i$. Presume that the areal measurement m_i has a correlation of C_a with the true areal average over D_i. But the correlation of m_i with the areal average over $B \wedge D_i$ should be less than C_a due to the "corruption" of the areal sample over the part of D_i which is not in the basin. For simplicity the correlation of measurement m_i with the basin portion $B \wedge D_i$ will be assumed to be

$$C_{ai} = C_a \frac{B \wedge D_i}{D_i} \qquad (C.14)$$

The sampling correlation functions for each sampling geometry are summarized below

POINT SAMPLES

$$cor_i(x,y) = C_p \exp(-\alpha((x-x_i)^2 + (y-y_i)^2)^{\frac{1}{2}})$$

LINE SAMPLES

$$cor_i(x,y) = C_\ell \exp(-\alpha d_\ell)$$

where d_ℓ=distance from (x,y) to the sample line

AREAL SAMPLES

$$cor_i(x,y) = C_a \frac{(B \wedge D_i)}{D_i} \qquad (x,y) \text{ in } (B \wedge D_i)$$

$$= 0 \text{ otherwise}$$

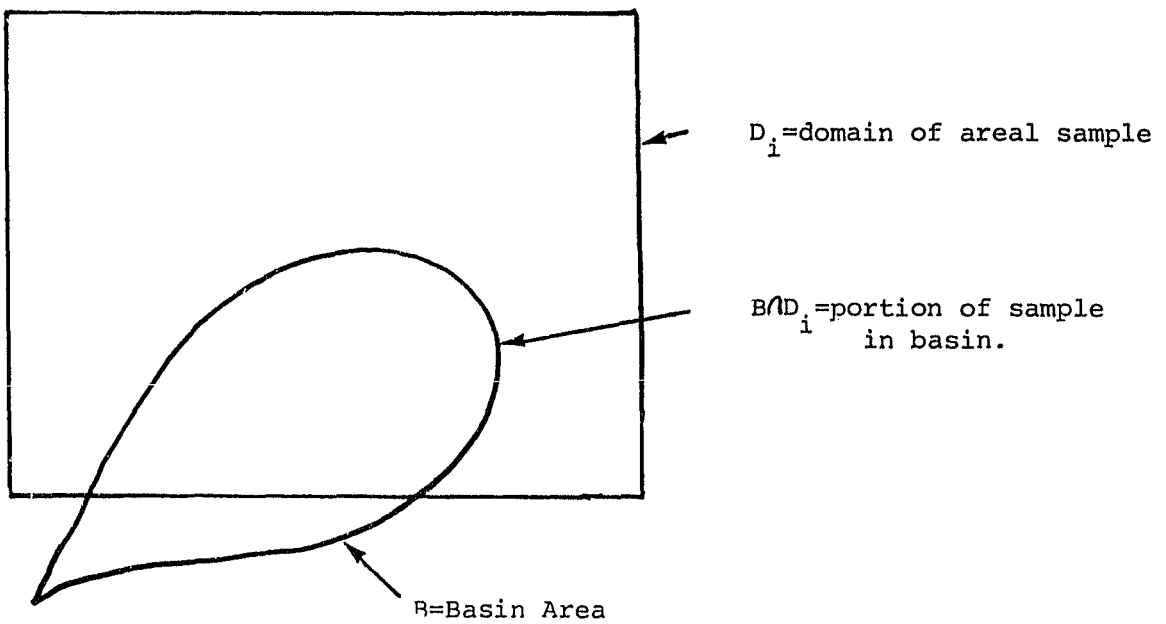


FIGURE C.1. INTERSECTION OF BASIN AND AREAL SAMPLE DOMAIN

APPENDIX D

CORRELATION AREA OF POINT SAMPLES

INTRODUCTION

The first task in finding the correlation area of point samples is to find the sample area - that portion of the basin with which the particular sample is more highly correlated than any other measurement. This will be developed by considering in turn the basin boundaries, the effect of other point samples, the effect of line samples, and finally the effect of areal samples.

DIVISION INTO SUBREGIONS

Before proceeding with this, a coordinate transformation will be introduced since the natural coordinate system for considering point samples is a polar system centered on the sample point (x_i, y_i).

For ease of reference, equation B.2 for the correlation area (A_i) of the point sample i is reproduced here:

$$A_i = \iint_{S_i} cor_i(x,y)\,dxdy$$

introducing the form of the point correlation function gives

$$A_i = \iint_{S_i} C_p \exp(-\alpha((x-x_i)^2+(y-y_i)^2)^{\frac{1}{2}})\,dxdy \qquad (D.1)$$

Transforming to polar coordinates in (r,θ) centered on (x_i,y_i) gives

$$A_i=\int_0^{2\pi}\int_0^{r(\theta)} C_p\exp(-\alpha r)\, r\, dr\, d\theta \qquad (D.2)$$

with $r(\theta)$ describing the shape of the sample area. In fact, the sample area will not be described by a single function $r(\theta)$ valid over $0\leq\theta\leq 2\pi$; rather it will be described by a number of functions $r^j(\theta)$ describing the shape of portions of the sample area valid over $\theta_1^j<\theta<\theta_2^j$.

Thus the correlation area will be defined by

$$A_i=\sum_{j=1}^{J} A_i^j \qquad (D.3)$$

$$A_i^j=\int_{\theta_1^j}^{\theta_2^j}\int_0^{r^j(\theta)} C_p\, r\exp(-\alpha r)\, dr\, d\theta \qquad (D.4)$$

Unfortunately, there is no guarantee that the sample area will be a convex region. When it is not, $r(\theta)$ will not be a single-valued function over some ranges of θ. Figure D.1 shows such a situation for a portion of the sample area bounded by three straight lines forming a nonconvex region. The boundaries are formed by lines from (x_1,y_1) to (x_2,y_2) from (x_2,y_2) to (x_3,y_3) and from (x_3,y_3) to (x_4,y_4). It is fairly easy to see that the correct correlation area for the region considered is

$$A_i=A_i^1-A_i^2+A_i^3 \qquad (D.5)$$

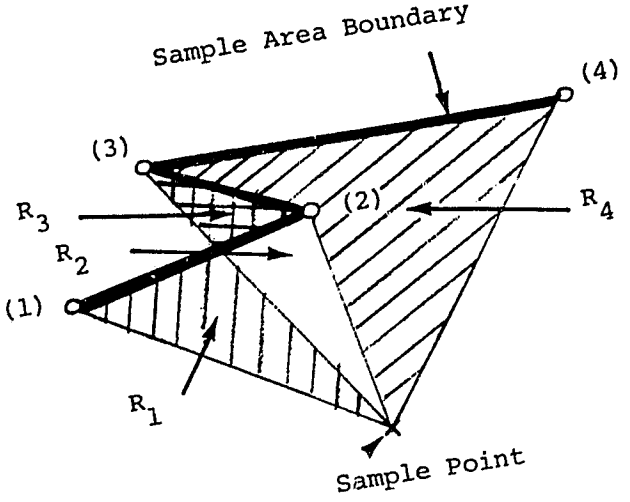


FIGURE D.1. NON--CONVEX SAMPLE AREA

The value of A_i^1 includes regions labelled R_1 and R_2. The value of A_i^2 includes regions labelled R_2 and R_3. The value of A_i^3 includes regions R_2, R_3, and R_4. Thus the computation of equation D.5 will include regions

$$R_1+R_2-R_2-R_3+R_2+R_3+R_4$$

giving regions $R_1+R_2+R_4$ which is correct.

The point is that care must be taken in defining the sign of the correlation area of each subregion.

It is certainly possible for a point sample taken outside the basin boundary to be more highly correlated with some nearby portion of the basin than any other measurement. Thus it is possible for a point measurement to be outside the sample area S_i for that measurement. Again, the same form of the integrals for each subregion of the sample area can be used, but great care in defining the sign for each subregion is needed.

Having described the necessity for care in the sign convention for the correlation area of each subregion j, the superscript j will be dropped to simplify notation, and the exposition will concentrate on defining $r(\theta)$ for various boundary shapes. Thus we have

$$Ai = \int_{\theta_1}^{\theta_2}\int_{o}^{r(\theta)} C_p r \exp(-\alpha r)\,dr\,d\theta$$

$$= \frac{Cp}{\alpha^2}\left[(\theta_2-\theta_1) - \int_{\theta_1}^{\theta_2}(1+\alpha r(\theta))\exp(-\alpha r(\theta))\,d\theta\right] \qquad (D.6)$$

It is understood that equation D.6 may apply in a piecemeal fashion to various subregions of the sample area.

In addition to finding a value for the correlation area A, it will be necessary to find the total area of the sample area. This area is S_i defined by

$$S_i = \int_{\theta_1}^{\theta_2} \int_{o}^{r(\theta)} r\,dr\,d\theta \qquad (D.7)$$

$$= \int_{\theta_1}^{\theta_2} \frac{r^2(\theta)}{2} d\theta$$

Depending on the form of $r(\theta)$, the integrals of equations D.6 and D.7 may need to be evaluated numerically.

The sample area of point measurement i will be described by an ordered set of points (x_1,y_1), $(x_2,y_2),\ldots(x_n,y_n)$ defined in counter clockwise fashion with each pair of points connected by a line. The lines between the boundary points will be either straight lines, parabolic arcs, or circular arcs. Thus, the definition of the correlation area proceeds in two steps: defining the sample area boundary, and evaluating A_i and S_i via equations D.6 and D.7 for each subregion of the sample area.

It is fairly simple to keep the sign of each subregions value of A_i and S_i correct. Consider the sample area boundary between point (x_1,y_1) and (x_{i+1},y_{i+1}). Due to the manner in which the sample area boundary is defined, the measurement point will always be "inside" any portion of the boundary which is either a circular or parabolic arc, thus the correlation area and sample area of any subregion bounded by a circular or parabolic arc will always be positive. For straight line boundaries, the sign should be positive whenever (x_{i+1},y_{i+1}) is to the left of point (x_i,y_i) relative to the measurement point (x_p,y_p). This can be tested by computing the azimuth of (x_i,y_i) relative to point (x_p,y_p) noted as ϕ_i $(0\leq\phi_i\leq360)$ and the azimuth of (x_{i+1},y_{i+1}) in a similar fashion noted ϕ_{i+1}. If the value of (ϕ_i+180) modulo 360 is less than ϕ_{i+1} then the sign for the subregion is positive.

The sample area will be defined in an iterative fashion. Beginning with the entire basin boundary as the first estimate of the sample area, the sample area will be reduced by considering the effect of each measurement in turn. As each measurement is considered, the algorithm of Appendix H will be used to define the new sample area as the region included inside both the old sample area and the portion of the x - y plane more highly correlated with the sample point than it is with the particular measurement being considered. This process is illustrated in Figure D.2.

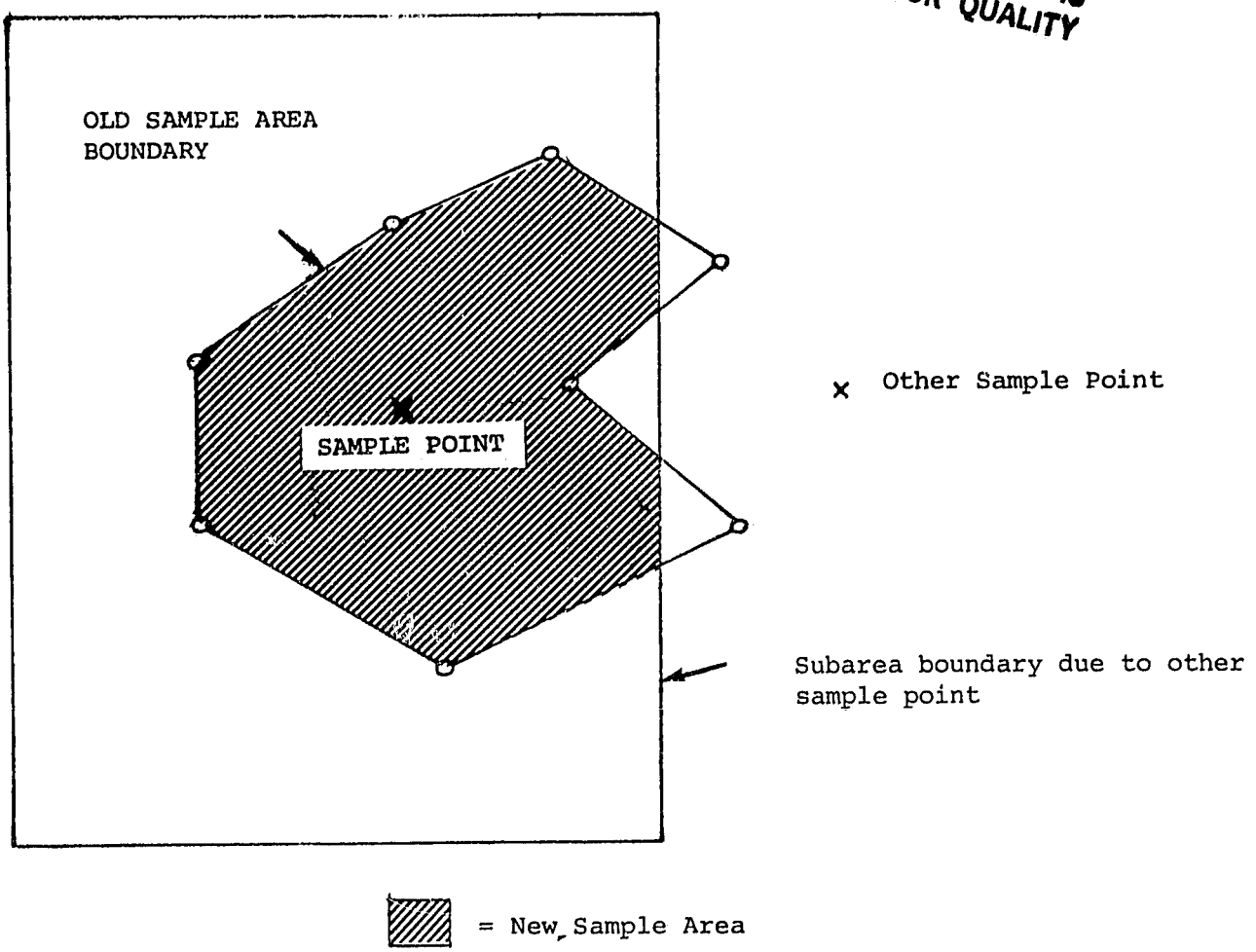


FIGURE D.2. EXAMPLE OF ITERATIVE DEFINITION OF SAMPLE AREA

Sections below consider the effects of other point samples, line samples, and areal samples on the sample area. Then the computation of A_i and S_i for linear, parabolic, and circular boundaries of the sample area are described.

POINT MEASUREMENTS

Presuming that all point samples have the same sampling correlation function, then the locus of points equally correlated with two samples is the same as the locus of points equidistant to the two samples, i.e., the perpendicular bisector of the line joining the two. As shown in Figure D.3 it is fairly simple to define the sample area for this case as a closed boundary. Points 2 and 3 lie on the perpendicular bisector of the line joining the two sample points. The points are selected to be far outside any possible basin boundary (say $\pm$50,000 miles). Then points (1) and (4) are taken on lines perpendicular to (2) and (3) again a large distance away.

Note that the final result of this process, when only point samples are considered, is the definition of sample areas equivalent to the Thiessen polygons frequently employed in estimating mean areal precipitation from raingage data.

FLIGHT LINE SAMPLES

As developed in Appendix C, the sample correlation function for line samples is

$$cor_i(x,y)=C_\ell \exp(-\alpha d_\ell) \quad \text{(line sample } m_i\text{)} \qquad \text{(D.8)}$$

where d_ℓ=distance from (x,y) to flight line.

and the sample correlation function for point samples is

$$cor_i(x,y)=C_p \exp(-\alpha d_p) \quad \text{(point sample } m_i\text{)} \qquad \text{(D.9)}$$

where d_p=distance from (x,y) to (x_i,y_i)

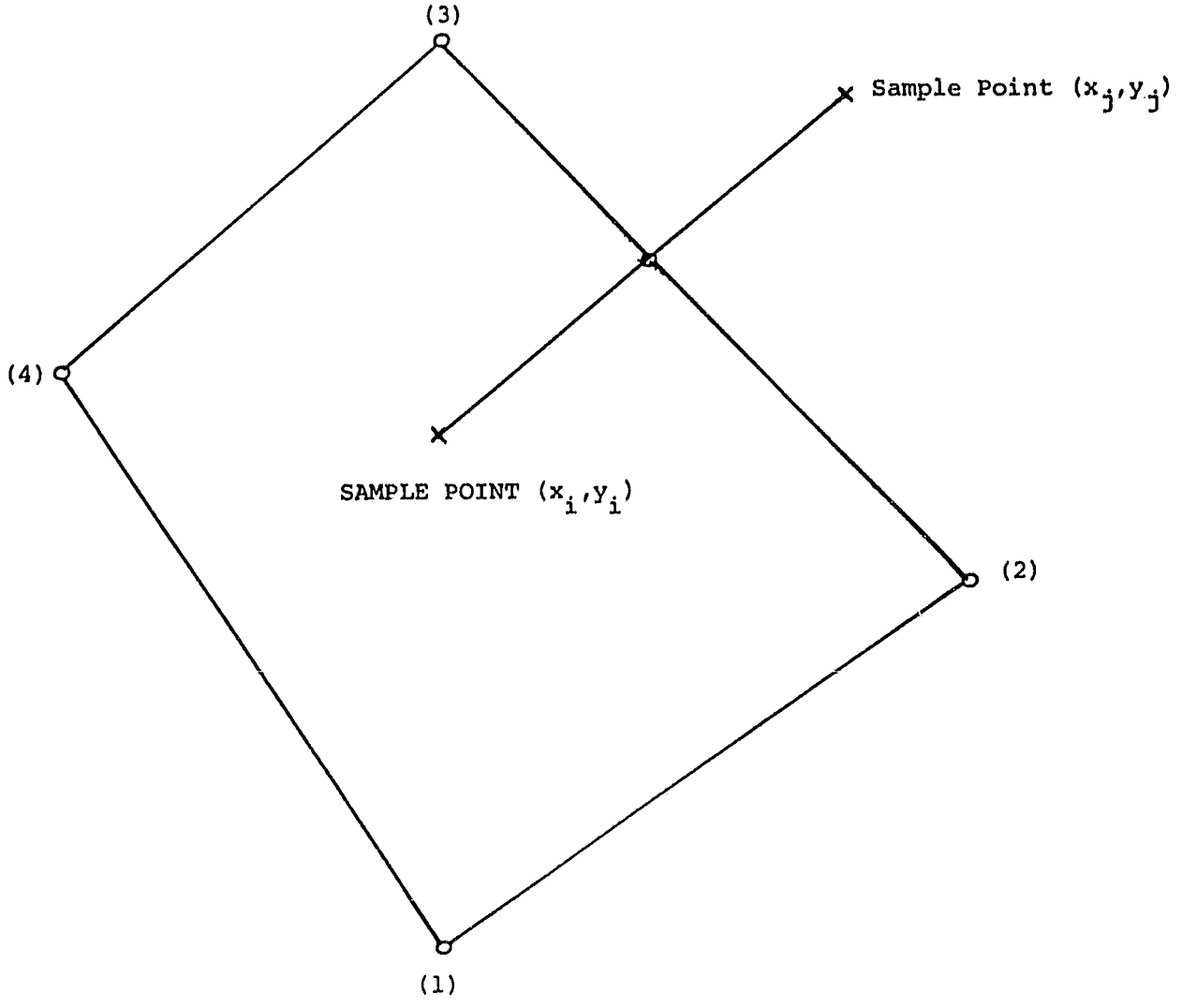


FIGURE D.3. INTERACTION OF TWO POINT SAMPLES

Because the value of α is the same for both the line and point correlation functions, the locus of points of equal correlation will be a parabola. If $C_p=C_\ell$ then the flight line will be the directrix, but more generally the directrix of the parabola will be parallel to the flight line. Referring to Figure D.4, the distance to the directrix is 2d where d is defined as the distance to the point of equal correlation with both the flight line and the point sample. Thus

$$C_p \exp(-\alpha d)=C_\ell \exp(-\alpha(D-d)) \tag{D.10}$$

which gives

$$d=\frac{1}{2}\left[\frac{1}{\alpha}\ell n\left(\frac{C_p}{C_\ell}\right)+D\right] \tag{D.11}$$

with the flight line defined by the line

$$ay=bx+c \tag{D.12}$$

the value of D is given by

$$D=\left(\frac{(ay_p-bx_p-c)^2}{a^2+b^2}\right)^{\frac{1}{2}} \tag{D.13}$$

The directrix is a line given by

$$ay=bx+c' \tag{D.14}$$

which defines d as

$$2d=\left(\frac{(ay_p-bx_p-c')^2}{a^2+b^2}\right)^{\frac{1}{2}} \tag{D.15}$$

Combining equations D.15, D.13, and D.11 the value of c' can be determined and inserted in equation D.14 to give the equation of the directrix as

$$ay=bx+c-\frac{\sqrt{a^2+b^2}}{\alpha}\ell n\left(\frac{C_p}{C_\ell}\right) \tag{D.16}$$

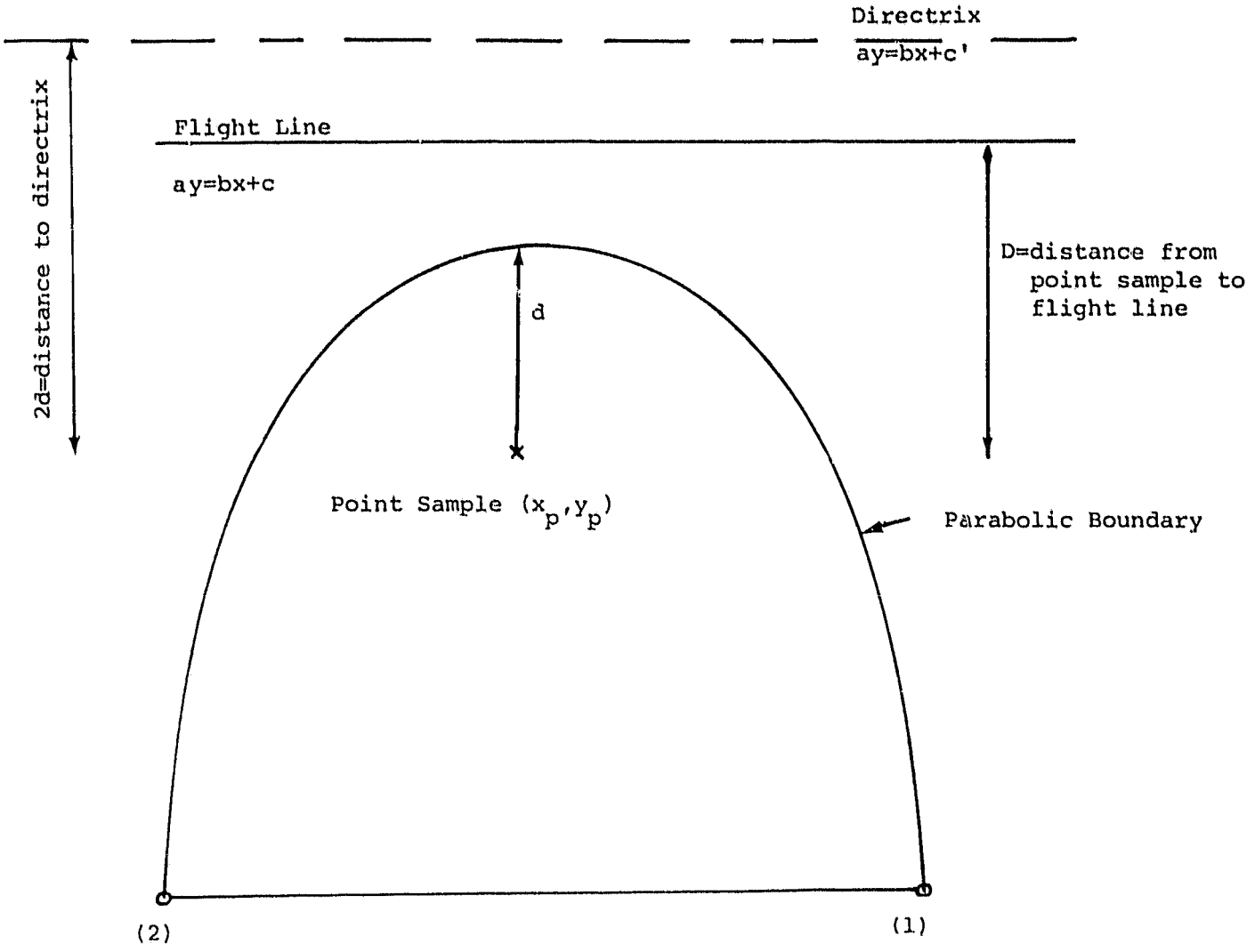


FIGURE D.4. INTERACTIVE OF POINT SAMPLE AND FLIGHT LINE

The complete equation of the parabolic boundary of the sample area is

$$(a^2+b^2)x^2+(a^2+b^2)y^2+(b-2x_p(a^2+b^2))x$$
$$-(a+2y_p(a^2+b^2))y$$
$$+\left[(x_p^2+y_p^2)(a^2+b^2)+c-\frac{\sqrt{a^2+b^2}}{\alpha}\ln\left(\frac{C_p}{C_\ell}\right)\right]=0 \qquad (D.17)$$

In a manner similar to that employed for the intersection of two point samples, the sample area is made into a closed boundary by selection of points (1) and (2) connected by straight lines and located "very far" from the basin.

AREAL SAMPLES

As developed in Appendix C, the sample correlation function for areal samples is

$$cor_i(x,y)=C_a\left(\frac{B\cap D_i}{D_i}\right) \quad (x,y) \text{ in } (B\cap D_i) \qquad (D.18)$$
$$=0 \quad \text{otherwise}$$

for areal sample m_i with sampling domain D_i and B the basin area. The entire basin area B will therefore be subdivided into basins each of which belongs to the domain of a particular areal sample m_i. All points within this subbasin are considered to be correlated with the areal sample at the value given by equation D.18. Thus, the sample area of point samples must not include any parts of the subbasin that are less correlated with the point sample being considered. Thus,

$$C_a\left(\frac{B\cap D_i}{D_i}\right)=C_p\exp(-\alpha R)$$

defines the boundary of the sample area of the point sample where R is the distance from the point sample. Thus the sample area boundary is a circle of radius R where R is given by

$$R=\frac{1}{\alpha}\left\{\ell nC_p-\ell n\left[C_a\left(\frac{B\Lambda D_i}{D_i}\right)\right]\right\} \tag{D.19}$$

and the equation of this circle is given by

$$(x-x_p)^2+(y-y_p)^2=R^2 \tag{D.20}$$

for the point sample at (x_p,y_p).

This completes the consideration of the interaction of all possible types of samples with a point samples. Starting with the basin boundary described by connected straight lines the sample area(s) for each point sample is described in iterative fashion:

(1) Subdivide the basin into subbasins, one for each domain of an areal sample.

(2) Reduce each subbasin by imposing the straight lines resulting from each point sample.

(3) Further reduce each subbasin by imposing the parabolic sample area resulting from each line sample.

(4) Finally reduce each subbasin further by imposing the circular sample area resulting from the areal samples.

The exposition now turns to evaluating the correlation area and sample area integrals for each type of line segment connecting the sample area points defined by the above process.

LINEAR BOUNDARIES

Figure D.5 shows a point sample at (x_p,y_p) and a straight line boundary of the sample area between points (x_i,Y_i) and (x_{i+1},y_{i+1}). Recall that the entire sample area boundary is described by the lines between the vertex points given by the ordered pairs $(x_1,y_1),(x_2,y_2)\ldots(x_n,y_n)$. A counter clockwise convention has been adopted to order the set of sample area boundary points. (See Appendix I).

The lengths of the sides of the triangle formed by (x_p,y_p), (x_i,y_i) and (x_{i+1},y_{i+1}) are

$$r_{i,p}=\left[(x_i-x_p)^2+(y_i-y_p)^2\right]^{\frac{1}{2}} \qquad \text{(D.20a)}$$

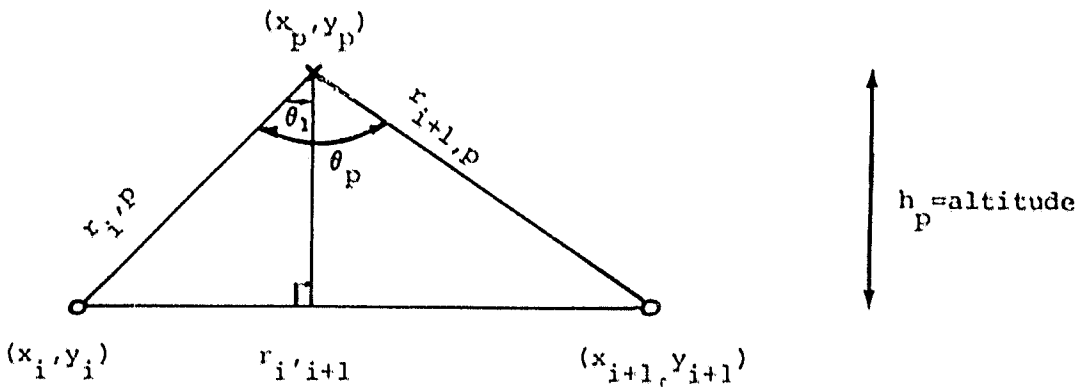


FIGURE D.5. POINT SAMPLE WITH STRAIGHT LINE BOUNDARY

$$r_{i+1,p}=\left[(x_{i+1}-x_p)^2+(y_{i+1}-y_p)^2\right]^{\frac{1}{2}} \qquad \text{(D.20b)}$$

$$r_{i,i+1}=\left[(x_i-x_{i+1})^2+(y_i-y_{i+1})^2\right]^{\frac{1}{2}} \qquad \text{(D.20c)}$$

The angle θ_p is given by

$$\theta_p=\cos^{-1}\left(\frac{r_{i,p}^2+r_{i+1,p}^2-r_{i,i+1}^2}{2r_{i,p}\,r_{i+1,p}}\right) \qquad \text{(D.21)}$$

where θ is taken as positive by convention.

Then h_p=height of altitude from (x_p,y_p) to opposite side

$$=\frac{r_{i,p}\,r_{i+1,p}\,\sin\theta_p}{r_{i,i+1}} \qquad \text{(D.22)}$$

Now define θ_1 as the angle between the altitude h_p and the point (x_i,y_i) given by

$$\theta_1=\pm\cos^{-1}\left(\frac{h_p}{r_{ip}}\right) \qquad \text{(D.23)}$$

$\theta_1>0$ for $r_{i+1,p}^2 > r_{i,p}^2+r_{i,i+1}^2$

$\theta_1<0$ otherwise

Then $\theta_2=\theta_1+\theta_p$ (D.24)

and finally

$$r(\theta)=\frac{h_p}{\cos(\theta)} \qquad \theta_1\leq\theta\leq\theta_2 \qquad \text{(D.25)}$$

The definition for $r(\theta)$ given by equation D.25 over the limits given by D.23 and D.24 is then used in equation D.6 giving

$$A_i=\frac{C_p}{\alpha^2}\left[\theta_p-\int_{\theta_1}^{\theta_2}\left(\frac{1+\alpha h_p}{\cos\theta}\right)\exp\left(\frac{-\alpha h_p}{\cos\theta}\right)d\theta\right] \tag{D.26}$$

which must be evaluated numerically.

Of course S_i for this geometry is simply

$$S_i=\frac{h_p\ r_{i,i+1}}{2} \tag{D.27}$$

As each point measurement is considered, then equations D.26 and D.27 give the correlation area and the sample area respectively for each subregion formed by a straight line boundary of the sample area. If the appropriate sign is chosen for each subregions A_i and S_i value, the sum of these will then be the total correlation area and the total sample area for the point sample over those portions of the sample area bounded by straight lines.

PARABOLIC BOUNDARIES

Consider Figure D.6 showing the sample point (x_p,y_p) and the directrix. Geometrically we have

$$\frac{2d-r}{r}=\cos\theta \tag{D.28}$$

Solving for r gives

$$r=\frac{2d}{\cos\theta+1} \tag{D.29}$$

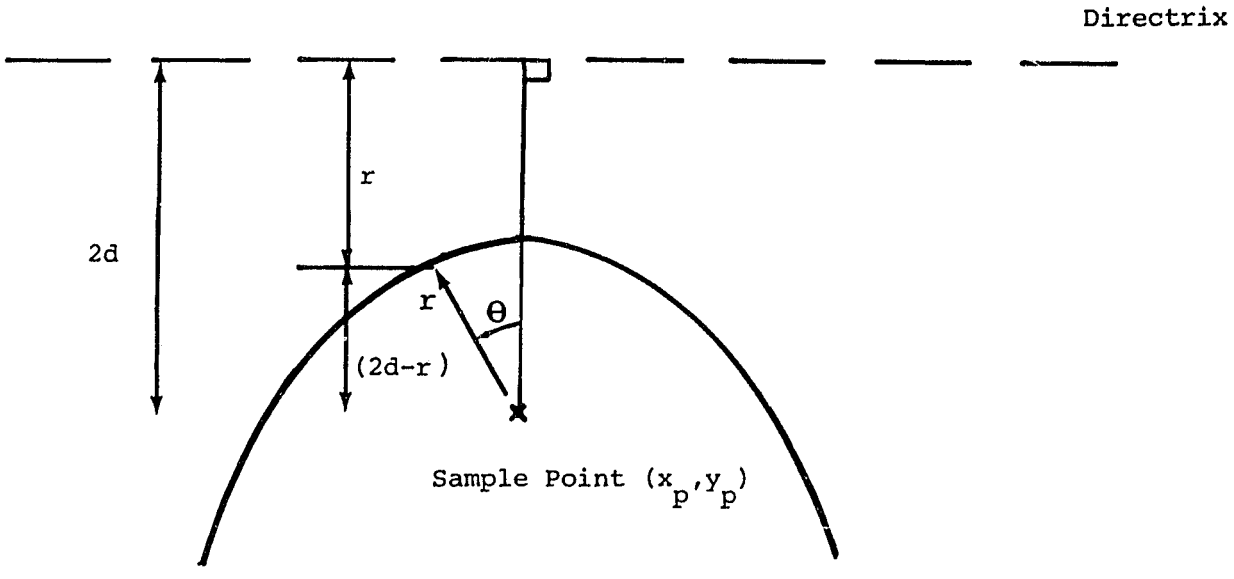


FIGURE D.6. TRANSFORMATION OF PARABOLA TO POLAR COORDINATES

Employing equations D.11 and D.13 gives the final result

$$r(\theta)=\frac{1}{\cos\theta+1}\left[\frac{1}{\alpha}\ell n\left(\frac{C_p}{C_\ell}\right)+\left(\frac{(ay_p-bx_p-c)^2}{a^2+b^2}\right)^{\frac{1}{2}}\right] \qquad -\pi<\theta<\pi \qquad (D.30)$$

Equation D.30 is used in equations D.6 and D.7 to define the correlation area and sample area of subregions with parabolic boundary. The form of D.30 forces the integrals of D.6 and D.7 to be evaluated numerically.

CIRCULAR BOUNDARY

The circular boundary gives the simplest form for $r(\theta)$. It is simply

$$r(\theta)=R \qquad (D.31)$$

where R is defined by equation D.19.
when inserted into equation D.6 this gives

$$A_i=\frac{C_p(\theta_2-\theta_1)}{\alpha^2}\left[1-(1+\alpha R)\exp(-\alpha R)\right] \qquad (D.32)$$

for the correlation area of the subregion of the sample area bounded by the circular arc between points (x_1,y_1) (x_2,y_2).
The sample area itself is simply

$$S_i=\frac{R^2}{2}(\theta_2-\theta_1) \qquad (D.33)$$

APPENDIX E

CORRELATION AREA OF LINE SAMPLES

INTRODUCTION

Defining the correlation area and sample area for line samples proceeds in much the same way as it does for point samples. It is necessary first to define the shape of the sample area, then to evaluate the correlation area and sample area over conveniently shaped subregions of the sample area.

Again, it is convenient to perform a coordinate transformation. In this case the usual (x,y) coordinate plane will be rotated and translated to align the x axis along the flight line itself with x=o at the start of the flight line. This transformed coordinate frame is depicted in Figure E.1.

Recall that the form of the sampling correlation function from Appendix C for line samples is

$$cor_i(x,y)=C_\ell \exp(-\alpha d_\ell) \tag{E.1}$$

where d_ℓ= distance from (x,y) to flight line

In the transformed coordinate system this will become

$$cor_i(y)=C_\ell \exp(-\alpha|y|) \tag{E.2}$$

where the absoluate value in the exponent can be dropped by the expedient of considering separately the area above and below the flight line and defining y positive for each case.

Then the correlation area of the j-th subregion is simply

$$A_i^j = \int_o^{L^j_{max}} L^j(y) C_\ell \exp(-\alpha y)\, dy \tag{E.3}$$

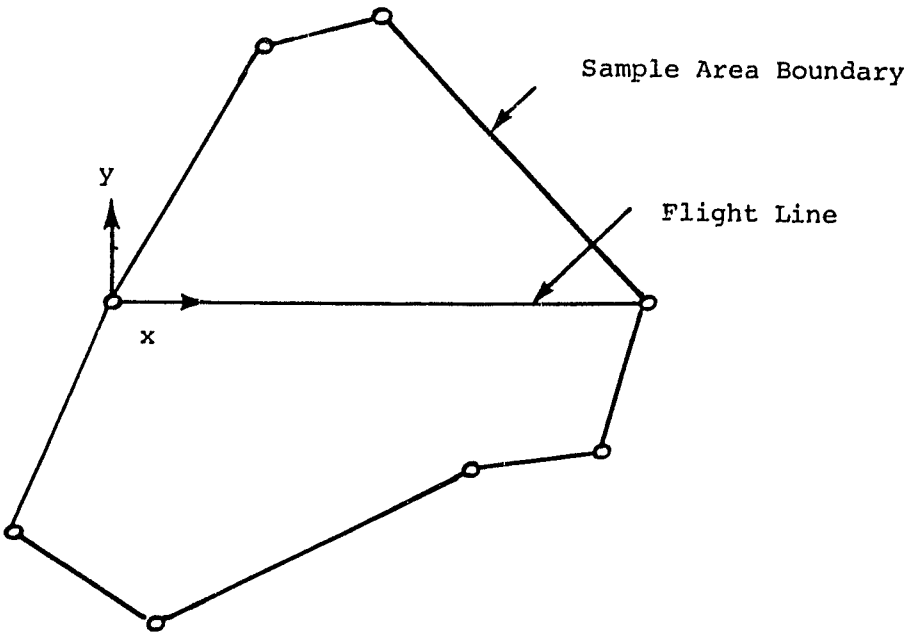


FIGURE E.1 COORDINATE SYSTEM FOR LINE SAMPLES

where L(y) gives the length of the j-th subregion in the x direction as a function of y.

L^j_{max} is the maximum distance in the y direction.

Further the sample area of the j-th subregion is

$$S^j_i = \int_0^{L^j_{max}} L^j(y)\,dy \qquad (E.4)$$

All that is required to evaluate E.3 and E.4 is the description of the shape of each subregion as defined by L(y). This is now done by considering in order the effect of areal samples, other flight lines, and point samples. Finally, the treatment of flight lines which terminate within the basin boundary rather than passing across the entire basin is discussed.

INTERACTION WITH AREAL SAMPLES

By analogy to the consideration of point samples with an areal sample, the effect of an areal sample is to terminate the sample area of the line sample at a distance R from the flight line. R is given by

$$R=\frac{1}{\alpha}\left\{\ell n C_\ell - \ell n\left[C_a\left(\frac{B \wedge D_i}{D_i}\right)\right]\right\} \qquad (E.5)$$

Two straight lines are constructed parallel to the flight line at distance +R and -R. These lines are used to reduce the size of the sample area in exactly the same fashion that the perpendicular bisector between two point samples was used to limit the sample area of the point samples. Thus if the original basin boundary is described by straight line sections, the reduced sample area after considering the effect of areal samples is also described by straight line sections.

INTERACTION WITH FLIGHT LINES

Figure E.2 shows two flight lines. Presuming that both flight lines have identical sampling correlation functions, the locus of points equally correlated with each flight line is defined by the line that bisects the angle between them. Again this is a straight line boundary which acts to further reduce the sample area for the flight line.

INTERACTION WITH POINT SAMPLES

As developed in Appendix C, the sample area boundary between a point sample and a line sample is a parabola. Whereas the sample area for the point sample is the region inside the parabola, the sample area for the line sample is simply the area outside the parabola. This case is covered by the algorithm of Appendix H.

FLIGHT LINES TERMINATING INSIDE BASIN

Figure E.3 shows a flight line terminating inside the basin boundary. For this case, the basin will be subdivided into three regions. Region B is inside the region bounded by lines perpendicular to the flight line passing through the terminal points of the flight line. Thus Region B can be treated as above. The correlation area of the flight line with Region A will be found by treating the terminus of the flight line as a point sample. However, this point sample will have a sampling correlation function of

$$Cor_i(x,y)=C_\ell \exp(-\alpha d) \qquad (E.6)$$

i.e. the sample correlation function applicable to the line sample, not point samples.

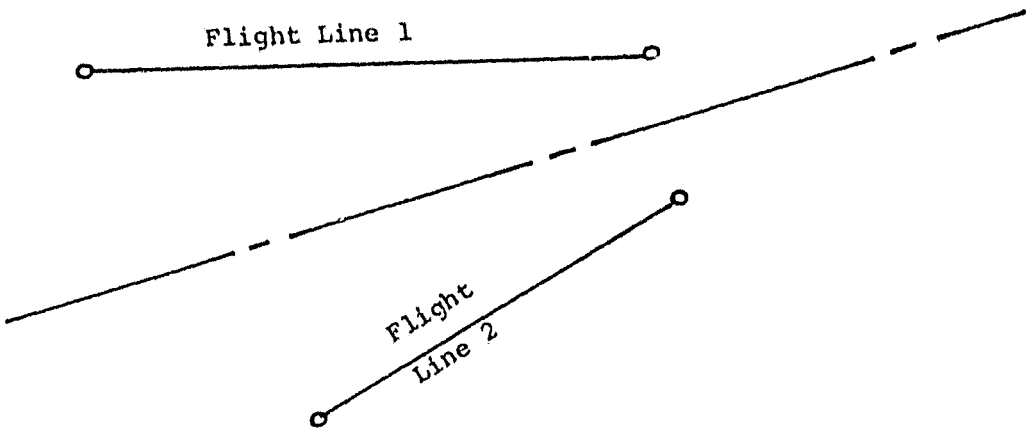


FIGURE E.2 INTERACTION OF TWO FLIGHT LINES

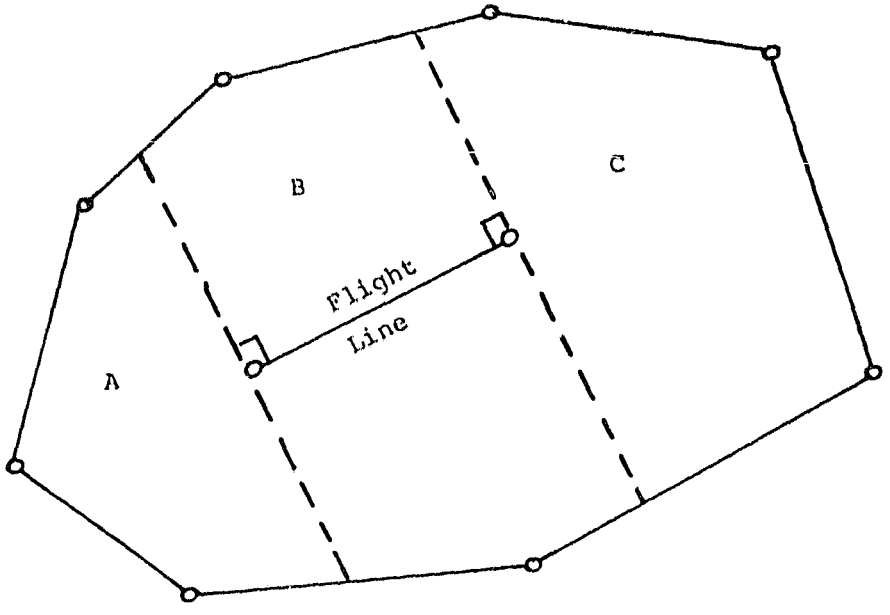


FIGURE E.3 FLIGHT LINE TERMINATING INSIDE BASIN

CORRELATION AREAS OF SUBREGIONS WITH LINEAR BOUNDARIES

Figure E.4 shows the flight line and two points (x_i, y_i) and (x_{i+1}, y_{i+1}) defining a straight line portion of the subregion. Note that (x_i, y_i) and (x_{i+1}, y_{i+1}) are in the transformed coordinate system with the x axis oriented along the flight line. The definition of L(y) is then straightforward.

$$L_{max} = (x_{i+1} - x_i) \tag{E.7}$$

$$y_{min} = \text{Min}(y_i, y_{i+1}) \tag{E.8}$$

$$y_{max} = \text{Max}(y_i, y_{i+1}) \tag{E.9}$$

$$L(y) = L_{max} \qquad 0 \leq y \leq y_{min} \tag{E.10}$$

$$L(y) = \frac{L_{max}(y_{max} - y)}{y_{max} - y_{min}} \qquad y_{min} \leq y \leq y_{max} \tag{E.11}$$

Both equations E.10 and E.11 give simple forms for the correlation area A_i when inserted into equation E.3. Of course the value of S_i is simply defined as

$$S_i = L_{max}\left(\frac{y_i + y_{i+1}}{2}\right) \tag{E.12}$$

Note that the value of L_{max} as defined by equation E.7 may result in $L_{max} < 0$. This implies adoption of a sign convention for the correlation areas of subregions. Again the sample area may not be convex as illustrated by Figure E.5 which considers three straight line segments.

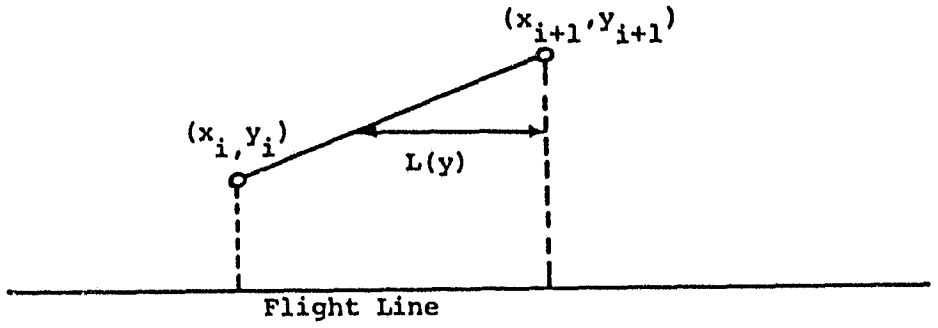


FIGURE E.4 CORRELATION AREA OF SUBREGION WITH LINEAR BOUNDARY

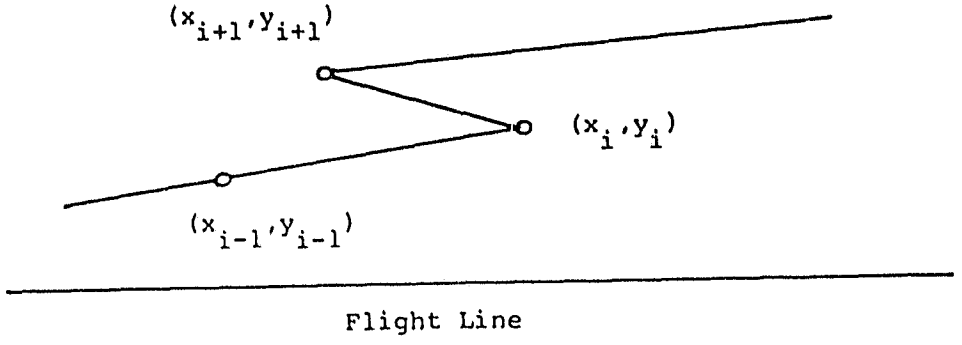


FIGURE E.5 NONCONVEX SAMPLE AREA BOUNDARY OF LINE SAMPLE

PARABOLIC BOUNDARIES

Figure E.6 shows a parabolic boundary section between points (x_i, y_i) and (x_{i+1}, y_{i+1}) produced by a point sample at (x_p, y_p). The full parabola is extended in dashed lines. The point where the parabola is closest to the flight line is (x^*, y^*) and it will either be outside the region from (x_i, y_i) to (x_{i+1}, y_{i+1}) as shown in Figure E.6a or inside as shown in Figure E.6b. In either case, the value of the correlation area will be described over four regions - two regions for (x_i, y_i) with (x^*, y^*) and two for (x_{i+1}, y_{i+1}) with (x^*, y^*). The four regions are noted I,II,III and IV on Figure E.6b. In case b the total correlation area and total sample area is simply the sum of the four regions. In case a it is simply region I plus region II minus region III minus region IV. The values of L(y) for each region are given below:

Region I-

$$L_{max} = |x^* - x_i|$$

$$L(y) = L_{max} \qquad o \leq y \leq y^*$$

Region II-

$$L_{max} = |x^* - x_i|$$

$$L(y) = L_{max} - (y - y^*)^2 \qquad y^* \leq y \leq y_i$$

Region III-

$$L_{max} = |x^* - x_{i+1}|$$

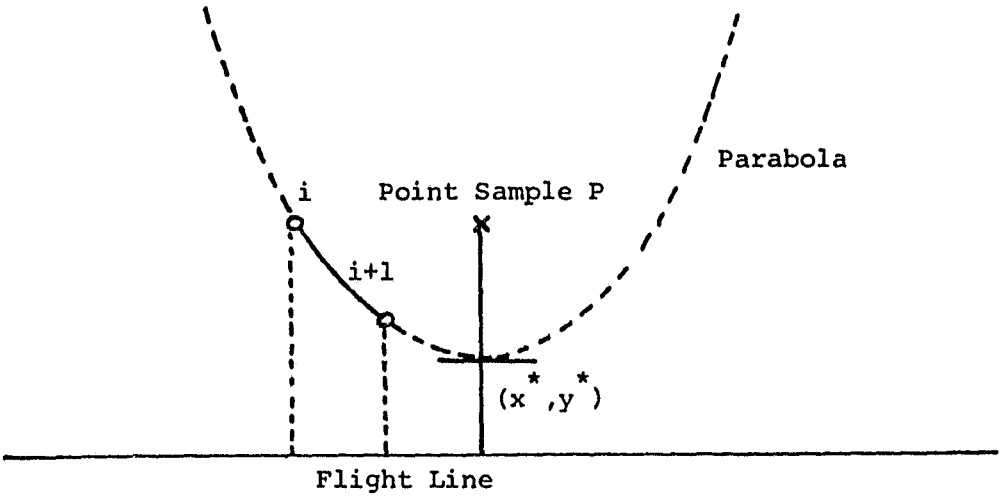
$$L(y) = L_{max} \qquad o \leq y \leq y^*$$

Region IV-

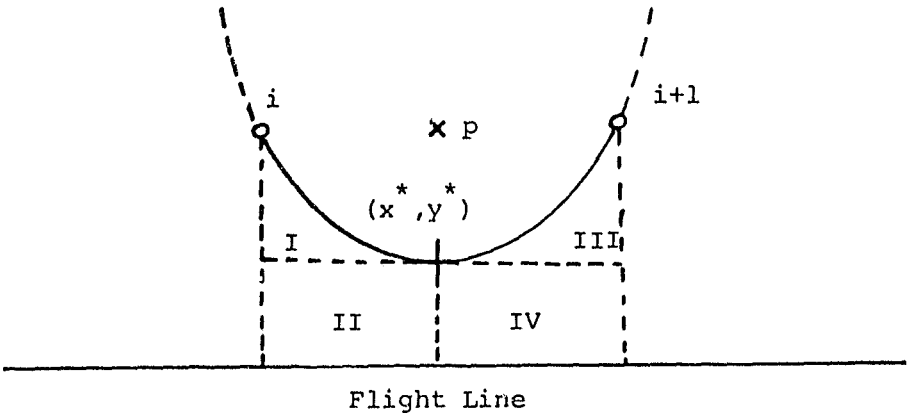
$$L_{max} = |x^* - x_{i+1}|$$

$$L(y) = L_{max} - (y - y^*)^2 \qquad y^* \leq y \leq y_{i+1}$$

The sign convention is simple to determine. Merely use the sign of $(x_{i+1} - x^*)$ as the sign adopted for regions III and IV.



(a) Parabola Is Closest to Flight Line Outside the Region from (x_i, y_i) to (x_{i+1}, y_{i+1})



(b) Parabola Is Closest to Flight Line Inside the Region from (x_i, y_i) to (x_{i+1}, y_{i+1})

FIGURE E.6. PARABOLIC BOUNDARY OF LINE SAMPLE AREA

Of course the form for L(y) for each region is inserted into equations E.3 and E.4 to find the correlation area and sample area of each region.

APPENDIX F

CORRELATION AREA OF AREAL SAMPLES

INTRODUCTION

Having computed the sample areas of all point and line samples inside the boundary of an areal sample it is quite simple to compute the correlation area of the areal sample. Referring to Figure F.1, the sample area of the areal sample is simply

$$S_i=(B\cap D_i)-\Sigma\text{sample areas of point and line samples} \qquad (F.1)$$

According to the sample correlation function developed in Appendix C, the areal sample is correlated with all points in the sample area at a value $cor_i(x,y)$ given by

$$Cor_i(x,y)=C_a\left(\frac{B\cap D_i}{D_i}\right) \qquad (x,y) \text{ in sample area} \qquad (F.2)$$

Thus the correlation area is A_i given by

$$A_i=C_a\left(\frac{B\cap D_i}{D_i}\right)S_i \qquad (F.3)$$

with S_i given by equation F.1.

This will finish the processing of areal average samples if only one type of areal sampling technology is in use. If more than one areal sampling technology is available it will be necessary to combine all the areal samples into a single set of areal samples. This is described in the following section.

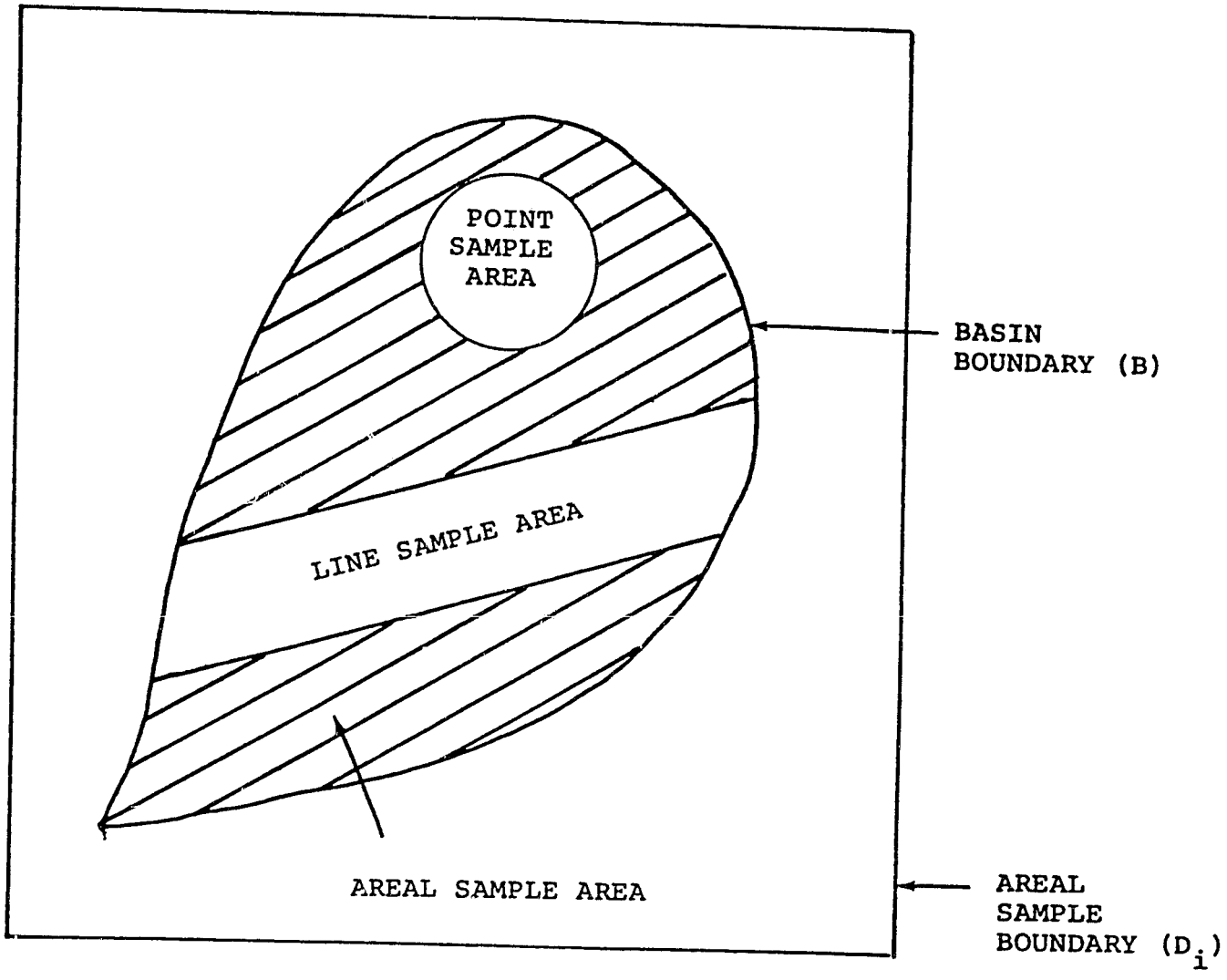


FIGURE F.1 EXAMPLE OF AREAL SAMPLE AREA

COMBINING AREAL SAMPLES

Figure F.2 shows two overlapping areal samples m_1 and m_2, sample 1 with domain D_1 and sample 2 with domain D_2. The area of overlap is $D_1 \cap D_2$. An estimate $\hat{A}$ of the true areal average A over $D_1 \cap D_2$ is desired as a combination of m_1 and m_2. Further, the sample correlation of $\hat{A}$ with A is needed. It will be assumed that m_1 and m_2 are independent, unbiased estimates of A with uncorrelated white noise estimation errors thus

$$m_1 = A + \varepsilon_1 \tag{F.4}$$

$$m_2 = A + \varepsilon_2 \tag{F.5}$$

$$E\varepsilon_1 = E\varepsilon_2 = 0 \tag{F.6}$$

$$E\varepsilon_1^2 = \sigma_1^2$$

$$E\varepsilon_2^2 = \sigma_2^2$$

$$E\varepsilon_1\varepsilon_2 = 0$$

Further a homogeneous mean uncorrelated with $\varepsilon_1, \varepsilon_2$ will be assumed, thus

$$EA = Em_1 = Em_2 = \mu \tag{F.7}$$

$$E\varepsilon_1 A = 0$$

$$E\varepsilon_2 A = 0$$

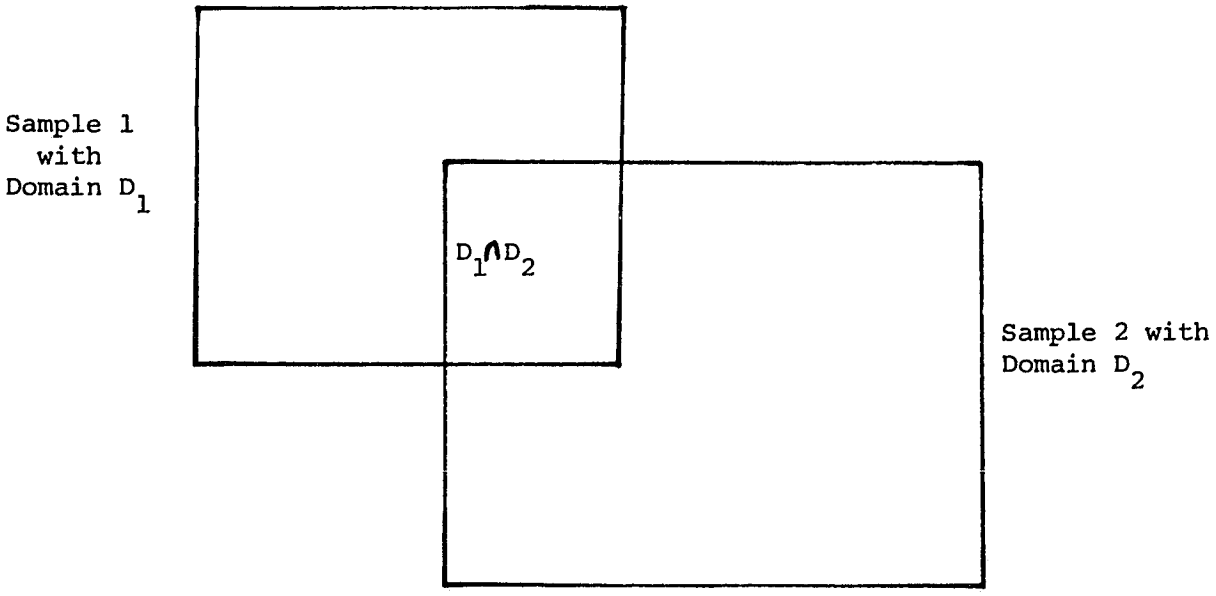


FIGURE F.2 TWO OVERLAPPING AREAL SAMPLES

ORIGINAL PAGE IS OF POOR QUALITY

In a manner consistent with the development of Appendix C, it is possible to define the correlation of each measurement with the true value A as follows

$$C_1 = C_{a1}\left(\frac{D_1 \cap D_2}{D_1}\right) \tag{F.8}$$

where C_{a1} = correlation of measurement m_1 with areal average over D_1

$$C_2 = C_{a2}\left(\frac{D_1 \cap D_2}{D_2}\right) \tag{F.9}$$

where C_{a2} = correlation of measurement m_2 with areal average over D_2

By these definitions, and employing the assumptions introduced in F.6 and F.7

$$C_1 = \frac{E(m_1-\mu)(A-\mu)}{E(m_2-\mu)^2} = \frac{\sigma^2}{\sigma^2+\sigma_1^2} \tag{F.10}$$

$$C_2 = \frac{E(m_2-\mu)(A-\mu)}{E(m_2-\mu)^2} = \frac{\sigma^2}{\sigma^2+\sigma_2^2} \tag{F.11}$$

where $\sigma^2 = E(A-\mu)^2$

Thus we have alternative definitions of σ_1^2 and σ_2^2 as

$$\sigma_1^2=\sigma^2\frac{(1-C_1)}{C_1} \tag{F.12}$$

$$\sigma_2^2=\sigma^2\frac{(1-C_2)}{C_2} \tag{F.13}$$

Assuming m_1 and m_2 to be independent, the value of A, the best estimator of $\hat{A}$ best estimator of A in the mean square sense, will be

$$A=m_1\left(\frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}\right)+m_2\left(\frac{\sigma_1^2}{\sigma_1^2+\sigma_2^2}\right) \tag{F.14}$$

when equations F.12 and F.13 are employed in F.14 the form

$$\hat{A}=m_1\beta+m_2(1-\beta) \tag{F.15}$$

where $\beta=\dfrac{C_1(1-C_2)}{C_1(1-C_2)+C_2(1-C_1)}$

is derived. This solves the first part of the problem - how to combine m_1 and m_2 to estimate A.

Next an estimate of the correlation of $\hat{A}$ to A is needed. Calling this C_A, we have

$$C_A=\frac{E(\hat{A}-\mu)(A-\mu)}{E(\hat{A}-\mu)^2} \tag{F.16}$$

Employing equations F.4, F.5, and F.15 gives

$$C_A = \frac{E(A-\mu+\varepsilon_1\beta+\varepsilon_2(1-\beta))(A-\mu)}{E((A-\mu)+\varepsilon_1\beta+\varepsilon_2(1-\beta))^2} \qquad (F.17)$$

Due to the independence assumptions on m_1 and m_2, equation F.17 reduces to

$$C_A = \frac{\sigma^2}{\sigma^2+\beta^2\sigma_1^2+(1-\beta)^2\sigma_2^2} \qquad (F.18)$$

Employing F.12 and F.13 removes σ^2 and produces

$$C_A = \frac{1}{1+\beta\ \frac{(1-C_1)}{C_1} + (1-\beta)^2\frac{(1-C_2)}{C_2}} \qquad (F.19)$$

which completes the analysis.

Having found $\hat{A}$ over the domain $(D_1 \cap D_2)$ and its corresponding correlation estimate C_A, $\hat{A}$ can be processed just like any other areal sample. Clearly the process of combining areal samples in this fashion is completely general regardless of the degree of overlap of two samples m_1 and m_2. Also it can be extended to include more than two measurement technologies by simply combining two at a time until finished.

APPENDIX G

TESTING A POINT FOR INCLUSION IN THE BASIN

INTRODUCTION

It is necessary at some points in the overall algorithm that points be tested for whether or not they are within the basin boundary. This appendix describes the algorithm that performs this test. The basin boundary is a closed, not necessarily convex, polygon described by the ordered pairs of vertex points $(x_1,y_1),(x_2,y_2),\ldots(x_n,y_n)$. The point to be tested will be noted as (x_t,y_t). The algorithm relies on the topologic property that any half-line from (x_t,y_t) which crosses the basin boundary at least once, will cross the boundary an even number of times if (x_t,y_t) is outside the basin and will cross the boundary an odd number of times if (x_t,y_t) is inside the basin.

To illustrate this property, consider Figure G.1. Test line 1 crosses the boundary 2 times indicating that point 1 is outside the boundary. Test line 2 crosses 3 times indicating that point 2 is inside the boundary.

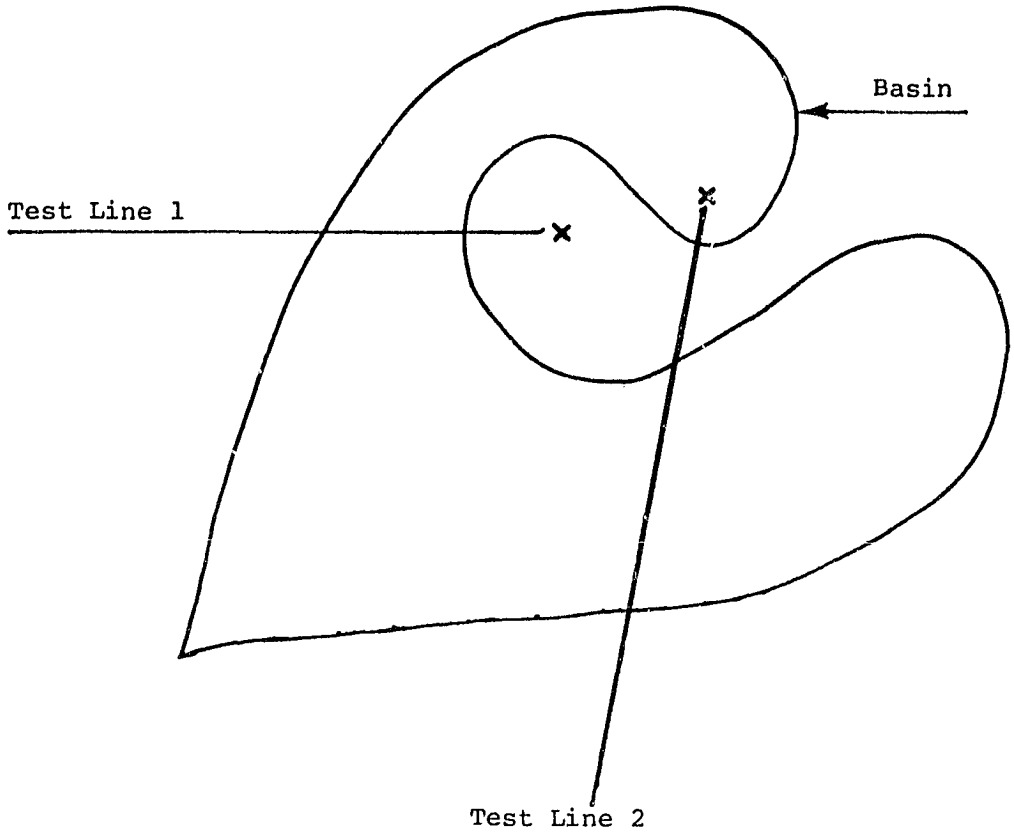


FIGURE G.1 ILLUSTRATION OF CROSSING PROPERTY OF CLOSED POLYGONS

BASIN BOUNDARIES

The basin boundary between the points $(x_i y_i)$ and (x_{i+1}, y_{i+1}) is given by

$$(x_i - x_{i+1})y = (y_i - y_{i+1})x + (x_i y_{i+1} - y_i x_{i+1}) \qquad (G.1)$$

It is understood that

$$x_1 = x_{n+1}$$

$$y_1 = y_{n+1}$$

The point (x,y) must also satisfy the limits

$$0 \leq \frac{(x_i - x)}{(x_i - x_{i+1})} \leq 1 \qquad x_i \neq x_{i+1} \qquad (G.2)$$

and $$0 \leq \frac{(y_i - y)}{(y_i - y_{i+1})} \leq 1 \qquad y_i \neq y_{i+1} \qquad (G.3)$$

where both G.2 and G.3 will be satisfied in all cases except $(x_i = x_{i+1})$ or $(y_i = y_{i+1})$ when only one condition can be tested.

For simplication of notation, equation G.1 will be noted as

$$a_i y = b_i x + c_i \qquad (G.4)$$

where $a_i = x_i - x_{i+1}$

$b_i = y_i - y_{i+1}$

$c_i = x_i y_{i+1} - y_i x_{i+1}$

TEST LINES.

It is necessary to define a test line which will cross the basin boundary at least once. In the unlikely event that the selected test line is colinear with a basin boundary or intersects a basin boundary at a vertex it will be necessary to select a new test line. As a simple expedient to assure that the test line does intersect the basin, define the point (x_c,y_c) on the basin boundary as

$$T=1 \tag{G.5}$$

$$x_c=\frac{x_T+x_{T+1}}{2}, y_c=\frac{y_T+y_{T+1}}{2} \tag{G.6}$$

The test line is then the line from (x_t,y_t) to (x_c,y_c) given by

$$(x_t-x_c)y=(y_t-y_c)x+(x_t\ y_c-y_t x_c) \tag{G.7}$$

Limited by $$\frac{x_t-x}{x_t-x_c} \geq o \quad x_t \neq x_c \tag{G.8}$$

and $$\frac{y_t-y}{y_t-y_c} \geq o \quad y_t \neq y_c \tag{G.9}$$

Again, both conditions G.8 and G.9 will be satisfied except for cases $(x_t=x_c)$ or $(y_t=y_c)$ when only one condition applies. If $(x_t=x_c)$ and $(y_t=y_c)$ then(x_t,y_t) is on the basin boundary itself.

Again, to simplify notation, the test line will be noted as

$$a_t y=b_t x+c_t \tag{G.10}$$

where $a_t=x_t-x_c$

$b_t=y_t-y_c$

$c_t=x_t y_c-y_t x_c$

If a new test line is required, simply redefine T=T+1 and continue at equation G.6.

INTERSECTIONS

The intersection between the test line (G.10) and the i-th basin boundary (G.9), if it exists, is the point (x_i^*, x_i^*) given by

$$x_i^* = \frac{c_i a_t - c_t a_i}{b_t a_i - b_i a_t} \tag{G.11}$$

$$y_i^* = \frac{a_t b_t c_i - a_t b_i c_t}{b_t a_i - a_t b_i} \tag{G.12}$$

whenever $b_t a_i = a_t b_i$ the intersection point (x_i^*, y_i^*) is undefined. There are two cases where this will occur.

Case 1: Colinear. It is possible that lines G.4 and G.10 are in fact the same line. If this is the case then all three conditions below apply.

$$b_t a_i = a_t b_i \tag{G.13}$$

$$c_t a_i = a_t c_i \tag{G.14}$$

$$c_t b_i = b_t c_i \tag{G.15}$$

If the lines G.4 and G.10 are colinear, a new test line is required.

Case 2: Parallel. If lines G.4 and G.10 are parallel but not colinear, then either G.14 or G.15 or both will not be true. In this case there are no crossings of the i-th boundary by the test line, i.e., the intersection point (x_i^*, y_i^*) is not a crossing.

CROSSINGS

Having found the intersection point (x_i^*, y_i^*) via equations G.11 and G.12 it is necessary to see if this intersection point is a crossing point. It is a crossing if it, (1) lies on the correct portion of the test line and also, (2) lies on the basin boundary. The test proceeds in two steps.

Step 1 is to test (x_i^*, y_i^*) for being in the correct portion of the test line. This is done by testing the conditions G.8 and G.9 at the point $(x=x_i^*, y=y_i^*)$. Each of these tests has four possible outcomes:

(1) Not Satisfied. The condition is simply not true.

(2) Impossible. The condition cannot be tested. For example, if $x_t=x_c$ condition G.8 cannot be tested.

(3) Satisfied=. The condition is true as an equality. For example, $x_t - x_i^*=0$.

(4) Satisfied$\neq$. The condition is true as a strict inequality.

Since each test has four outcomes, there are 16 possibilities as summarized in Table G.1. As noted, three of these cases cannot occur. Seven of these cases, noted in Table G.1 as "No" indicate that (x_i^*,y_i^*) is not a crossing point. Three of the cases, noted in Table G.1, as "Boundary", indicate that (x_i^*,y_i^*) is on the boundary of the basin if it passes the crossing test of Step 2. The final three cases, noted as "Step 2" in Table indicate that (x_i^*,y_i^*) is a crossing point if it passes the crossing test of Step 2.

Step 2 is to test (x_i^*, y_i^*) for being on the basin boundary. This is done by testing the conditions G.2 and G.3 at the point $(x=x_i^*, y=y_i^*)$. Again each test has four outcomes with 16 possibilities for both tests as indicated in Table G.2. Three

TABLE G.1 TEST FOR POINT ON TEST LINE

Condition G.8 \ Condition G.9	Not Satisfied	Impossible	Satisfied=	Satisfied¬
Not Satisfied	No	No	No	No
Impossible	No	(1)	Boundary	Step 2
Satisfied=	No	Boundary	Boundary	(1)
Satisfied¬	No	Step 2	(1)	Step 2

(1) This condition cannot occur.

TABLE G.2 TEST FOR CROSSING POINT

Condition G.2 \ Condition G.3	Not Satisfied	Impossible	Satisfied=	Satisfied≠
Not Satisfied	No	No	No	No
Impossible	No	(1)	Vertex	Crossing
Satisfied=	No	Vertex	Vertex	(1)
Satisfied≠	No	Crossing	(1)	Crossing

(1) This condition cannot occur.

cases are impossible occurences. Seven cases are not crossings. If any of the other 6 cases occur for a point indicated as a boundary point in Step 1 then (x_i^*, y_i^*) is the same as (x_t, y_t) and the test point lies on the basin boundary itself. If (x_t, y_t) is not on the basin boundary and the point (x_i^*, y_i^*) is a vertex of the basin boundary then a new test line is required. Finally, if (x_t, y_t) is not on the boundary (noted as "Step 2" in Table G.1) and the point (x_i^*, y_i^*) is noted as a "crossing" in Table G.2, then a simple crossing of the test line and the basin boundary has been found.

INCLUSION TEST

Form the set of all intersection points $((x_i^*, y_i^*), i=1...n)$ which are also crossings. If this set has an even number of members, then (x_t, y_t) is outside the basin. If this set has an odd number of members, then (x_t, y_t) is inside the basin.

ADDENDUM

Although the algorithm described above is complete, it is possible to improve its computational performance by a modification in the choice of a test line. Simply drop a vertical line from (x_t, y_t) defined by

$$x = x_t \tag{G.16}$$

$$y \leq y_t$$

If this test line fails to intersect the boundary, then (x_t, y_t) is surely outside the boundary. Further, it is not necessary to locate the intersection of the test line given by G.16, with each boundary line as only those straight boundary lines for which

$$0 \leq \frac{(x_i - x_t)}{(x_i - x_{i+1})} \leq 1 \tag{G.17a}$$

and

$$y_t \geq \text{Min}(y_i, y_{i+1}) \tag{G.17b}$$

need be considered. These conditions should quickly eliminate most boundary lines from consideration. Considering only boundary lines which satisfy conditions G.17, it is easy to find the intersection (x_t, y^*) of the boundary and test lines by simply solving equation G.1

for y^* at $x=x_t$. Then the intersection will be a crossing if y^* satisfies

$$0 \leq \frac{y_i - y^*}{y_i - y_{i+1}} \leq 1 \qquad \text{(G.18)}$$

If condition G.18 is satisfied as an equality, then the intersection point is at a vertex and a new test line is required.

APPENDIX H

SUBDIVIDING THE BASIN

It will be necessary at numerous points in the correlation area algorithm to subdivide the basin by defining subbasins which are included within both the basin boundary and within the boundary of some sub-area. Examples are finding the portion of the basin covered by an areal sample and finding the portion of the basin on one side of the line separating the sample areas of two point samples. This appendix describes the algorithm which is employed to perform this subdivisions.

The basin boundary is described by lines connecting an ordered set of points. This set will be noted F1 and is composed of points ($\underline{f}_1^1$, $\underline{f}_1^2$, $\underline{f}_1^3 \ldots \underline{f}_1^n$) defined in a counter clockwise fashion where $\underline{f}_1^i = (x^i, y^i)$ are the vertex points of the boundary. The lines connecting the vertex points may be straight lines, parabolic arcs, or circular arcs. The region bounded by F_1 is not necessarily a convex region, thus there may be multiple sub-basins included within both the basin boundary and the sub-area boundary.

The sub-area boundary is also described by lines connecting an ordered set of points defined counter clockwise. This set will be noted as F_2 using notation similar to that adopted above. Cases where the sub-area is unbounded can be included by the simple expedient of adding points at infinity; for example, the sub-area description below the X axis is described by the points $[(-\infty,-\infty), (+\infty,-\infty), (+\infty,0), (-\infty,0)]$ connected by straight lines.

The first step in the algorithm is to augment the sets F_1 and F_2 to include all the intersections of the basin and sub-area boundaries maintaining the same counter clockwise convention. The

augmented sets will be noted G_1 for the basin boundary + intersection points and G_2 for the sub-area boundary + intersection points. For example, assume that the basin boundary line between points $\underline{f}_1^3$ and $\underline{f}_1^4$ intersects the sub-area boundary lines between points $\underline{f}_2^9$ and $\underline{f}_2^{10}$ and that the intersection occurs at point (x^*,y^*). Then the set G_1 is $(\ldots\underline{f}_1^3,(x^*,y^*),\underline{f}_1^4,\ldots)$ and the set G_2 is $(\ldots,\underline{f}_2^9,(x^*,y^*),\underline{f}_2^{10},\ldots)$. Thus, all intersection points will be included in both G_1 and G_2, and F_1 is a subset of G_1, and F_2 is a subset of G_2.

In Appendix B a complete algorithm for locating the intersections of a straight line within the boundaries of a basin which is also composed of straight lines has been developed. It is a fairly straightforward extension to find the intersections of straight lines, parabolic arcs, and/or circular arcs, although the algebra does get more complex in the case of parabolic or circular arcs.

The purpose of the algorithm is to produce one or more sets of points H^i which are the boundaries of sub-basins within both the basin boundary and the sub-area boundary. It will be necessary to create a set of markers for each point in F_1 which indicate whether each point has been "used".

The algorithm proceeds in several steps as described below:

Step 1. Mark all points in F_1 as unused.

Step 2. Scan the points in F_1 one at a time searching for a point in F_1 which is inside the sub-area boundary (See Appendix G). As each point is considered it is marked "used". Only points which are initially marked "unused" need be checked. This step has three possible results:

(1) On the first pass through this step there are no points in F_1 which are inside the sub-area. In this case, test any point in F_2 for being inside the basin. If any point in F_2 is inside the basin then <u>all</u> of F_2 is inside the basin and thus the sub-basin boundary set H is identical to the sub-area boundary set F_2. If any point in F_2 is outside the basin, then <u>all</u> of F_2 is outside the basin and there is no subarea at all.

(2) On subsequent passes through Step 2 if all unused points in F_1 are outside the sub-area then the algorithm is complete.

(3) A point in F_1 has been found that is inside the sub-area. Call this point the start point noted by $\underline{f}_1^s$, and proceed to Step 3.

<u>Step 3</u>. This step will produce a new sub-area boundary set H^i. The first point in this set will be $\underline{f}_1^s$ as defined in Step 2. The set H^i will be formed by moving points in order from either G_1 or G_2 into the set H. The first point is $\underline{f}_1^s$. The next point to be moved to H is the next point in order in G_1. This process continues to move in order through G_1 until one of two possible event occurs.

(1) The algorithm returns to the point $\underline{f}_1^s$. This terminates the description of sub-area boundary set H^i. The algorithm must then return to Step 2 to check for multiple sub-areas.

(2) An intersection point is reached. The intersection point is added to the set H^i as usual, but the algorithm switches from the set G_1 to the set G_2 to continue the sub-area definition. Recall that each intersection point belongs to <u>both</u> G_1 and G_2.

The algorithm proceeds in order through either G_1 or G_2, switching from one to the other at each intersection point, until it returns to the starting point $\underline{f}_1^s$. When $\underline{f}_1^s$ is reached, Step 3 terminates by returning to Step 2.

Any point in G_1 which is not an intersection point will also be a member of F_1. Each such point should be marked "used" as it is added to the sub-area boundary set H^i.

A final complication arises when it is easier to define a sub-area as all points <u>not</u> in the sub-area boundary. This case is easily handled through two changes in the above algorithm:

(1) In Step 2, the points in F_1 should be tested for being outside the defined sub-area boundary rather than inside.

(2) In Step 3, points should be taken from the set G_2 in a clockwise (backwards) fashion rather than a counterclockwise fashion. G_1 should still be scanned counterclockwise (forward).

The definition of the order of points the sub-area boundary set F_2 must still be in the correct (counterclockwise) order.

APPENDIX J
DESCRIPTION OF SOILS ALONG LUVERNE, MINNESOTA RESEARCH FLIGHT LINE A

BISCAY SERIES

The Biscay Series consists of poorly drained soils formed in loamy over gravelly sediments under tall grass vegetation on glacial outwash plains, valley trains and river terraces. The surface soil is black and very dark gray loam 20 inches thick. The subsoil is olive gray and gray mottled loam and gravelly loam 16 inches thick. The substratum is dark gray and olive gray mottled gravelly coarse sand. Slops range from 0 to 2 percent. Most areas are used for cropland.

COMFREY SERIES

The Comfrey Series consists of poorly and very poorly drained soils formed in alluvium on bottomlands. The surface soil is black clay loam 26 inches thick. The substratum is dark gray mottled clay loam. Slopes range from 0 to 2 percent.

FAIRHAVEN SERIES

The Fairhaven Series consists of well drained soils formed in glacial outwash sediments on uplands. The surface layer is black and very dark grayish brown silt loam 14 inches thick. The subsoil is brown silt loam in upper 10 inches and brown and yellowish brown gravelly sandy loam in lower 3 inches. The substratum is grayish brown, brown, yellowish-brown and pale brown coarse sand and gravel. Slopes range from 0 to 12 percent. Most areas are used for cropland.

GRACEVILLE SERIES

The Graceville Series consists of deep, well and moderately well drained soils formed in 40 to 60 inches of silty alluvium over sand and gravel on stream terraces and outwash plains. The surface layer is dark grayish brown silty clay loam 20 inches thick. The subsoil is dark grayish brown, brown and light yellowish brown silty clay loam 33 inches thick. The substratum is dark brown gravelly sand. Slopes are less than 2 percent. Most areas are used for cropland.

APPENDIX J
DESCRIPTION OF SOILS
ALONG LUVERNE, MINNESOTA
RESEARCH FLIGHT LINE A (Continued)

MOODY SERIES

The Moody Series consists of deep, well-drained soils formed in silty loess on uplands. The surface layer is dark grayish-brown silty clay loam 5 inches thick. The subsoil is dark grayish-brown and brown, friable silty clay loam in upper 14 inches and light yellowish-brown, friable silty loam in lower 8 inches. The underlying material is light yellowish-brown, calcareous silt loam. Slopes range from 0 to 17 percent. Most areas are cultivated.

SAC SERIES

The SAC Series consists of deep, well drained soils formed uner prairie vegetation in loess over glacial till on uplands. The surface layer is black and very dark grayish brown silty clay loam 11 inches thick. The subsoil is very dark grayish brown, brown and black silty clay loam in upper 17 inches and brown and dark yellowish brown clay loam in lower 16 inches. The substratum is yellowish brown clay loam. Slopes range from 2 to 14 percent. Most areas are used for cropland.

TRENT SERIES

The Trent Series, nonflooded, consists of deep moderately well drained soils formed in loess on uplands. The surface layer is dark gray silty clay loam 16 inches thick. The subsoil is dark grayish brown, grayish brown and pale brown silty clay loam in upper 23 inches and pal brown silt loam in lower 7 inches. The substratum is light brownish gray and light gray silt loam. Slopes range from 0 to 2 percent. Cropland is the main use.

WHITEWOOD SERIES

The Whitewood Series consists of deep, somewhat poorly drained soils formed in local silty alluvium on flats and in swales and depressed upland drainageways. These soils have black, silty clay loam surface layers 16 inches thick: black, dark-grey and olive-gray, friable, silty clay loam subsoils 27 inches thick: and olive-gray silty clay loam underlying material. Slopes are less than 2 percent. Areas are used as cropland and pasture.

APPENDIX I

AREA OF POLYGONS

If a polygon is defined by its vertex points $(x_1,y_1),(x_2,y_2)$ $\ldots(x_n,y_n)$ with points defined in a counter clockwise order, the area of the polygon can be easily determined. The area is A given by

$$A = \frac{1}{2}\left[\Sigma x_i y_{i+1} - \Sigma x_{i+1} y_i\right] \qquad \text{(I.1)}$$

with $x_{n+1} \equiv x_1$

$y_{n+1} \equiv y_1$

Equation I.1 applies to convex and non-convex polygons.

This provides a simple test for the "correct" counter clockwise definition. If the area computed by equation I.1 is negative, then the points have been defined clockwise and could therefore be reordered.