THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

THE DESIGN OF A MULTIVARIATE MESOSCALE FIELD EXPERIMENT

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

BY

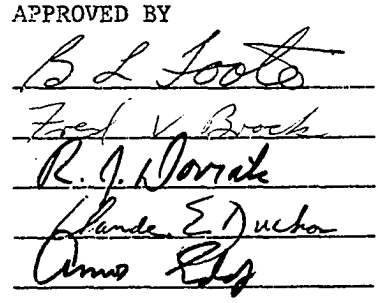
KENNETH C. CRAWFORD

Norman, Oklahoma

July 1977

THE DESIGN OF A MULTIVARIATE MESOSCALE FIELD EXPERIMENT

APPROVED BY



DISSERTATION COMMITTEE

ABSTRACT

THE DESIGN OF A MULTIVARIATE MESOSCALE FIELD EXPERIMENT

Radar reflectivities and rain gage data can be combined in many ways to estimate convective-storm surface rainfall. However, optimum (interpolation-produced) estimates used to evaluate modification experiments or to implement sampling techniques require that statistical properties of such estimates be known. Such properties derive from the storm structure, implied by the observed data, and are deduced by using space-time covariance and cross-covariance functions. A four dimensional Gaussian-damped function can reflect relevant characteristics (physical and statistical) of Southeastern Montana convective systems. Functional parameter values relate to system features of size, motion-speed and preferred storm track. An average Southeast Montana system compares and contrasts with the features of an Oklahoma counterpart.

The optimum interpolation methodology is enhanced to account for multivariate means and variances. Bivariate analyses that use radar/rain gage data sets are shown superior to the best univariate results. The analyses reflect patterns derived from radar rainfall estimates and scaled to rain gage magnitudes. The influence of a Z-R relationship on analysis accuracy is minimal and the model's signal recoverability qualities are shown. Consequences of filtering data set observations improperly are discussed.

The development of an experimental-design evaluation function is completed through modelling the parameter means and variances. Predictand-related sensors are shown essential to network design. Trade offs in multivariate sensor deployments (spatial and temporal) are explained. Deployment along and across a preferred storm track is related to covariance anisotropy, gage density, temporal sampling intervals, the availability of radar data, and the interrelationships among the multivariate predictor data sets. Moreover, optimal sensor orientation to sample a moving convective system is found best to observe the system's accumulated rainfall pattern.

ACKNOWLEDGEMENTS

This dissertation represents a culmination of efforts by many friends and associates. Surely the list of support would be long. However, significant financial support by the National Oceanic and Atmospheric Administration and the National Weather Service created an opportunity to accelerate the technique applications reported herein. I am also grateful to Dr. Bernie Silverman and his staff at the Bureau of Reclamation for enthusiastically providing a data base and computer resources.

Amos Eddy, through his creative genius, represents a fundamental force behind my work. Though philosophically we differ, I admire his many positive traits. Through the years he has patiently prodded my intellect with generous doses of his keen practical knowledge.

Dr. Pat Brady provided a foundation that insured my work would not take forever to complete. He never failed to answer and reanswer my many questions.

I have had the continual cooperation and interest of Dr. Fred Brock, Dr. Richard J. Doviak, Dr. Claude Duchon, and Dr. Bob Foote. Each, in his own way, kept me "honest".

The derivation in Appendix B is largely the work of Pat Avara. Ms. Helen J. Ardrey edited some of the words. Marg Eddy confidently and professionally created a beautiful manuscript during the last few hectic days.

Ultimately, support for this work rests with my wife Charlotte, my son Scott, and my daughter Kendra. Charlotte never doubted my ability. In her own way, she faithfully worked to see this work completed. I only hope the future provides the opportunity to repay their love and understanding.

TABLE OF CONTENTS

		Page
ACKNOWLEDGEMENTS		iv
LIST OF TABLES		vii
LIST OF ILLUSTRATIONS		ix
Chapter		
I.	THE INTRODUCTION	1
II.	THE DATA	5
III.	THE DEVELOPMENT OF METEOROLOGICAL EXPERIMENTAL DESIGN	9
	A. Objective Analysis Development.	9
	B. Early Sensor Placement	11
	C. Multivariate Experimental Design	13
IV.	COVARIANCE FUNCTION AND OBJECTIVE ANALYSIS ALGORITHMS: A REFINEMENT	27
	A. The Covariance Model	27
	B. The Analyzed Signal	35
V.	THE GENESIS OF A CONVECTIVE STORM CLIMATOLOGY	39
	A. Covariance Data Requirements	39
	B. Analysis Data Requirements	42
VI.	COVARIANCE CLIMATOLOGICAL STUDY - OBJECTIVE ANALYSIS RESULTS	48
	A. Univariate Analysis	48

Page

B. Covariance Study..... 65
C. Bivariate Analysis..... 82

VII. AN EXPERIMENTAL DESIGN EVALUATION FUNCTION: REVISITED..... 94

VIII. EXPERIMENTAL DESIGN RESULTS..... 102

A. Radar Observation Influence..... 105
B. Preferred Storm-Track Influence.. 114
C. Time-Domain Sampling Influence... 119
D. Sensors Deployed Balanced by Economy and Climatology..... 124

IX. THE CONCLUSIONS..... 126

X. ADDITIONAL RESEARCH: APPLICATIONS.... 131

A. Covariance Functions..... 131
B. Multivariate Analysis..... 132
C. Optimal Sampling..... 133

XI. THE BIBLIOGRAPHY..... 135

APPENDIX A..... 141

APPENDIX B..... 148

APPENDIX C..... 156

LIST OF TABLES

TABLE		Page
1.	Bureau of Reclamation HIPLEX radar/ rain gage data sets investigated herein.....	8
2.	Station surface precipitation analysis: radar influence controlled by varying the lag-zero cross-covariance value.....	46
3.	Covariance climatological-study results from investigating the Table 1 storm data sets.....	66
4.	Covariance climatological study results obtained from investigating the accumulated rainfall patterns of the Tables 1 and 3 storm data sets.....	67
5.	Type convective system (Table 1), motion-direction and speed, and optimal sampling orientation (positive with respect to north = 360^{o} and equal to 270^{o} - Tables 3 or 4 results).....	71
6.	The average Southeast Montana cross-covariance and expected gage network lag-zero covariance value for moving storm systems and accumulated rainfall patterns.....	79
7.	Filtered covariance parameter values obtained from the methods in Chapter V.B and the Table 3 convective systems.....	80

TABLE Page

8. Filtered total-storm-rainfall covariance parameter values obtained from the methods in Chapter V.B and the Table 4 convective systems..... 82

9. Characteristics of an "average" accumulated rainfall pattern from Southeast Montana..... 104

10. Contributions of each term to selected objective function values plotted in Fig. 29..... 110

11. Contributions of each term to selected objective function values plotted in Fig. 30..... 111

LIST OF ILLUSTRATIONS

FIGURE		Page
1.	The 1975 HIPLEX recording gage network of 58 stations near Miles City, Montana.....	7
2.	The 1976 HIPLEX recording gage network of 109 stations near Miles City, Montana.....	7
3.	Covariance model accuracy suggested by the sample covariance values...	30
4.	Covariance at lag-zero determined from data sets whose density was progressively decreased.....	41
5.	Signal-to-noise ratios of station (58 gages) bivariate surface rainfall analyses determined from radar data sets whose density was progressively decreased.....	45
6.	(A-H) Radar univariate analysis of 7 July 75 squall line moving from 350° at 25 km hr^{-1}.....	50
7.	(A-B) Radar univariate total-storm-rainfall analysis of 7 July squall line.....	52
8.	(A-E) Radar univariate analysis of 31 July 75 squall line moving from 190° at 74 km hr^{-1}.....	53
9.	(A-B) Radar univariate total-storm-rainfall analysis of 31 July 75 squall line.....	54

FIGURE Page

10. (A-F) Radar univariate analysis of 18 August 75 squall line moving from 250° at 59 km hr^{-1}..... 55

11. (A-B) Radar univariate total-storm-rainfall analysis of 18 August 75 squall line..... 56

12. (A-J) Constant Altitude PPI (CAPPI) analysis of the Fig. 10 squall line obtained from a four dimensional data subset, its Table 3 covariance function, and the data assimilation capability in the objective analysis model..... 57

13. (A-F) Radar univariate analysis of 17 July 75 embedded thunderstorm moving from 220° at 53 km hr^{-1}..... 60

14. (A-B) Radar univariate total-storm rainfall analysis of 17 July 75 embedded thunderstorm..... 61

15. (A-G) Radar univariate analysis of 18-19 August 75 embedded thunderstorm moving from 270° at 45 km hr^{-1}..... 62

16. (A-B) Radar univariate total-storm-rainfall analysis of 18-19 August 75 embedded thunderstorm..... 64

17. The relationship between convective system motion-direction and the sensor orientation for optimal sampling..... 72

18. The relationship between the sensor orientation for optimal sampling of accumulated rainfall patterns and moving convective systems.. 74

19. (A-F) Rain gage univariate analysis of 7 July 75 squall line moving from 330° at 22 km hr^{-1}..... 83

20. (A-F) Radar univariate analysis of 7 July 75 squall line moving from 330° at 22 km hr^{-1}..... 85

FIGURE		Page
21.	(A-F) Rain gage/radar bivariate analysis of 7 July 75 squall line moving 330° at 22 km hr^{-1}.....	87
22.	(A-B) Rain gage univariate total-storm-rainfall analysis of 7 July 75 squall line.....	90
23.	(A-B) Radar univariate total-storm-rainfall analysis of 7 July 75 squall line.....	90
24.	(A-B) Rain gage/radar bivariate total-storm rainfall analysis of 7 July 75 squall line.....	91
25.	(A-B) Rain gage univariate total-storm-rainfall analysis of 6 June 76 convective cell.....	92
26.	(A-B) Radar univariate total-storm-rainfall analysis of 6 June 76 convective cell.....	92
27.	(A-B) Rain gage/radar bivariate total-storm-rainfall analysis of 6 June 76 convective cell.....	93
28.	The method used to provide the experimental design algorithm a movable sensor's relative location (closed squares) in the presence of analysis grid (circles) and a permanently located (if any) sensor set (open squares).....	101
29.	(A-B) An evaluation of HIPLEX sampling requirements for observing accumulated rainfall patterns.....	108
30.	(A-B) An evaluation of HIPLEX sampling requirements for observing accumulated rainfall patterns.....	109

FIGURE Page

31. The influence of a preferred storm track on sensor deployment..... 116

32. The influence of a preferred storm track on sensor deployment..... 118

33. The influence of time-domain sampling on sensor deployment..... 122

34. A computational overview of the steps needed to produce an optimum interpolation analysis..... 156

35. A computational overview of the steps needed to evaluate optimal sampling requirements..... 157

THE DESIGN OF A MULTIVARIATE MESOSCALE FIELD EXPERIMENT

CHAPTER I

THE INTRODUCTION

Nearly 4000 years ago an Egyptian Pharaoh was forewarned of a weather-related catastrophe destined to affect the civilization of that day. The events happened as predicted. Time has come and gone. Clearly world civilization has changed and become complex during the intervening centuries. In spite of significant technological advances, weather-related catastrophes continue to plague twentieth century man. Ancient (e.g., John, AD 95) and contemporary writers (Ehrlich, 1971, and Schneider, 1976) now point to our climate-related crises that lie ahead. Indeed, increasing world population coupled with runaway inflation and an energy crises of immense proportions, brought to the forefront during the 1970s, added turmoil to other complex world problems. The message is clear to all who hear: we must use our known diminishing resources wisely as our world

seeks to move intact into the twenty-first century.

Resources available to meteorologists for use in gaining a fuller understanding of atmospheric interactions are limited and often difficult to obtain. The message is equally ours to share. Meteorological field programs designed to contribute solutions of climate-related problems must, of necessity, be sound theoretically as well as designed for efficiency.

Coordinated field programs are underway (the National Hail Research Experiment--NHRE and the High Plains Experiment--HIPLEX) with weather modification objectives. NHRE focuses on damaging hail and aims for its suppression; HIPLEX aims to enhance rainfall. Common to each program (indeed, to comprehensive field programs in general) is the need to deploy a variety of observational resources including radiosonde networks, surface stations, radars, aircraft, tethered balloons and other sensing systems, direct and indirect. Typically, these systems are deployed for periods of two to three months and provide the data base with which many scientists will test their hypotheses. Although the number and variety of instruments which can be devoted to these programs is substantial, the complexities of the atmosphere, its relevant scales of activity and scale interactions, dictate that observational compromises must be made. It follows that, because of the substantial manpower and instrument resources committed to such programs, the

instrument sets must be utilized in an efficient, if not optimum, fashion.

An equally compelling reason for optimum experimental design is provided by Barnes' well documented report (1974) that became the first coordinated life-cycle study of the immediate environmental structure in devastating storms. Barnes said "that our ensemble sounding might have yielded more information had attention been paid to their spacing relative to the storms."

This study is directed toward an optimum experimental design of the HIPLEX program; the study concentrates on areas of immediate concern to HIPLEX as it seeks to achieve the stated modification objectives: Can radar and rain gage observations be combined to produce analyses with qualities superior to those obtained in univariate analyses? What analysis methodology will provide consistent, high quality results essential to HIPLEX? How many gages are necessary to achieve a desired sampling and analysis quality? What is the best deployment of that gage quantity? Study results are based on the null hypothesis of convective storm features deduced from statistical structures implied by digital radar and surface gage data. The methodologies applied herein are universal and await application in other areas.

The primary positive contributions herein include:

1) more realistic modelling of storm statistical structure;

2) a convective storm climatology obtained using modelled space-time covariance functions whose parameters correlate well with system physical characteristics;

3) upgrading the analysis algorithm to account for realistic differences in multivariate parameter means and variances with the consequence that bivariate radar/rain gage analyses reflect radar-derived precipitation patterns scaled to gage magnitudes;

4) the development of an experimental design evaluation function which incorporates the effect of observational biases and the relative role to be played by each parameter in producing a final analysis;

and 5) the evaluation of trade offs involved in multivariate sensor deployments (spatial and temporal) using a combination of fixed and movable sensors.

The study reviews meteorological experimental design (Chapter III), discusses refinement of the covariance function and objective analysis algorithms (Chapter IV), evaluates analysis and sampling data requirements (Chapters V and VIII), reveals the analysis algorithm's positive characteristics (Chapter VI), and develops an experimental design evaluation function (Chapter VII and Appendix B). My results are illustrated by using a combination of radar/rain gage data sets (Chapter II) obtained from Southeast Montana during 1975-1976.

A computational overview is provided, in Appendix C, of the analysis and sampling methodologies applied herein.

CHAPTER II

THE DATA

The data sets used in this research consisted of precipitation observations obtained from quantitative radar reflectivity measurements and (only when supported by simultaneous storm radar data) from surface rainfall measurements. All observations were made in Southeast Montana in the Bureau of Reclamation's HIPLEX study area near Miles City, Custer County.

Data-acquisition procedures on the Miles City C-band (5 cm) radar produces a complete observation set every five minutes with spatial resolutions characterized by 1^{o} azimuthal and vertical separations and generally 0.5 km radial separations. Three dimensional (x-y-t) and four dimensional (x-y-z-t) radar data sets were obtained. Most data sets contained 1000 or more spatially separated radar bins. However, spatial and temporal data gaps frequently limited the data volume processed.

Reflectivity was converted into rainfall (in mm hr^{-1}) using Jones' (1956) relationship of $Z = 486R^{1.37}$, suggested by Battan (1973) as appropriate for "typical" convective storm systems. Reflectivities less than 20 dBz were converted into zero measured rainfall. The three-dimensional radar-rainfall data sets (typically taken from the 1^{o} elevation scan) were used to produce total-storm-rainfall data sets (x-y) by appropriately summing, over the storm-period of interest, the rainfall represented in each five-minute observation.

Fifteen-minute surface rainfall accumulations were obtained from a recording gage network of 58 gages (Fig. 1) in 1975 and 109 gages in 1976 (Fig. 2). Precipitation amounts were converted to mm hr^{-1} and total-storm-rainfall data sets calculated. Average gage separation-distance was approximately 3.6 km. Reflectivity data from storms passing over the gage network were interpolated linearly (for data processing convenience even though this is not required by the analysis algorithm) to be time-coincident with the corresponding gage data.

Three rain gage and 17 radar sets were obtained (Table 1) from the Bureau of Reclamation's CYBER 74-28 computer system. This quality user-oriented system possesses data archive and retrieval components capable of maximizing "thruput" and minimizing frustration (a rare system characteristic indeed!). A more complete description

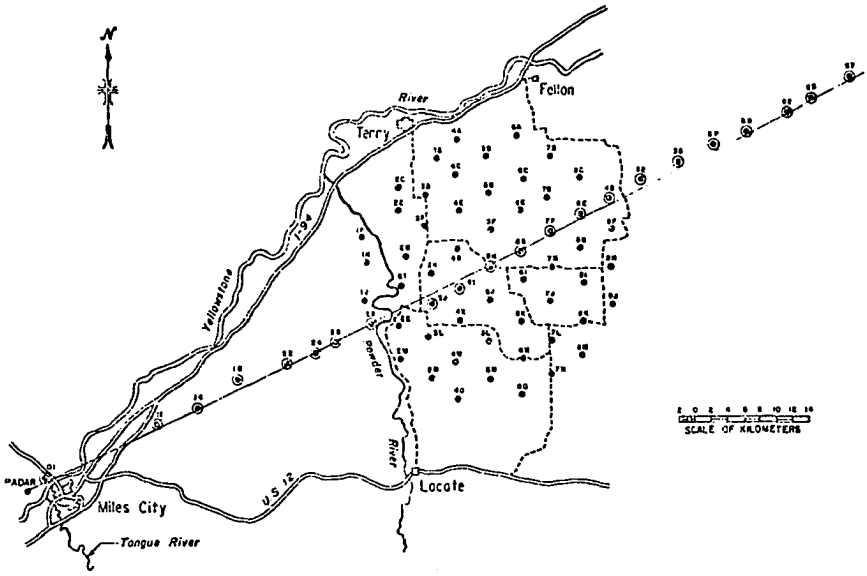


Fig. 1. The 1975 HIPLEX recording gage network of 58 stations near Miles City, Montana.

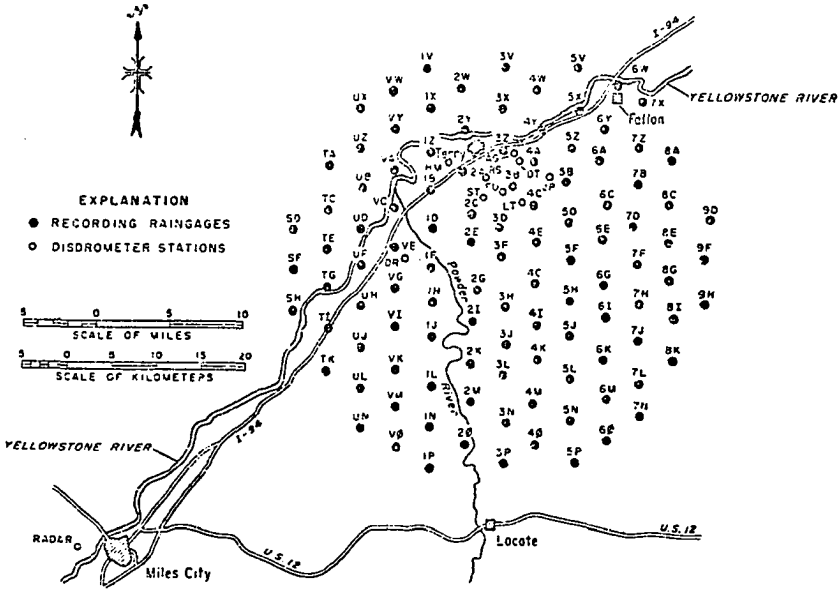


Fig. 2. The 1976 HIPLEX recording gage network of 109 stations near Miles City, Montana.

DATE	TIME	TYPE OF DATA
5 July 75	1930-2000	x-y-t and x-y-z-t radar data only
7 July 75	1307-1342	x-y-t radar data only
7 July 75	1900-2000	x-y-t radar data plus gage data
17 July 75	0007-0027	x-y-t radar data only
17 July 75	1545-1615	x-y-t radar data only
21 July 75	2252-2322	x-y-t radar data only
30 July 75	2003-2018	x-y-t and x-y-z-t radar data only
31 July 75	0658-0713	x-y-t and x-y-z-t radar data only
18 August 75	2213-2238	x-y-t and x-y-z-t radar data only
18 August 75	2347-0012	x-y-t radar data only
6 June 76	0530-0600	x-y-t radar data plus gage data
7 June 76	0430-0445	x-y-t radar data plus gage data
7 June 76	0515-0545	x-y-t radar data only

Table 1. Bureau of Reclamation HIPLEX radar/rain gage data sets investigated herein. All times are Greenwich Mean Time.

of Bureau of Reclamation HIPLEX data may be found in the HIPLEX Data Inventory (Interior Department, 1976).

Finally, similar radar and rain gage observations were obtained from the work of Pat Brady (1976) who investigated structural characteristics of Oklahoma thunderstorms.

CHAPTER III

THE DEVELOPMENT OF METEOROLOGICAL EXPERIMENTAL DESIGN

A. Objective Analysis Development

In the development stage of a field experiment, it is necessary to consider the methods used to assess the meteorological characteristics under study, such as size, shape, intensity, and duration, as well as to have the ability to determine instrument-induced noise-effect on observation quality. Also, it is important to measure objectively the role of a proposed instrument configuration in terms of instrument type and number, the sampling rate, and the sampling positions. The statistical objective analysis technique used herein fulfills these needs, and is an essential part of the experimental design concept.

Because an objective analysis is the weighted combination of observations taken at irregular intervals in space and time, it estimates a parameter field at some specified point (normally a lattice) in space and time.

The computer age and numerical weather prediction (NWP) prompted development of numerous "objective" analysis techniques. Panofsky (1949) made what was likely the first attempt at today's objective analysis by using least squares polynomial fitting to an observed data set to analyze objectively the data field at non-observation points. First guess fields and statistical weight functions were introduced by Bergthorsson and Döös (1955). The methodology, a successive correction technique later extended by Cressman (1959), was used to initialize NWP models for many years by applying a series of first-guess-field corrections based on the separation distance between grid points and observation locations. Barnes (1964) used the Cressman distance-weighting technique with his own data-density weighting factor to increase the detail in objectively analyzed fields. Asynoptic observations were later incorporated into the Barnes technique (1973) using mesoscale time-series observations.

Sasaki suggested a different analysis approach (1958) and used variational calculus to achieve dynamic consistency; Amos Eddy (1963) and L. S. Gandin (1963) formulated what is known today as "optimum interpolation" statistical objective analysis. The optimum interpolation technique effectively uses a data set's structure expressed in terms of the data auto- and cross-correlation functions. Eddy and Gandin's technique is termed "optimum" because it is the interpolation scheme that minimizes the interpolation

error variance compared to any other interpolation method.

The work of Endlich and Mancuso (1968) and Shapiro and Hastings (1973) used an approach similar to Panofsky while Stephens and Polan (1971) investigated the problem of spectrum degeneration induced by objective analysis.

Multivariate objective analysis has been pursued vigorously by many authors in the most recent literature including Eddy (1973), Thiebaux (1973, 1974), Schlatter (1975), Passi (1975), and Thiebaux (1977).

B. Early Sensor Placement Efforts

Determination of an aerological network's optimum spacing dates from the rawinsonde invention. Attempts and techniques used afterwards have been summarized concisely by Gandin et. al. (1967) and Gandin (1970). The first known quantitative attempt to solve the sensor deployment problem was made in 1936 in the Soviet Union by Drozdov (via Gandin et. al., 1967). Drozdov began by not allowing the linear interpolation standard error, mid-way between stations, to exceed a certain given amount. This interpolation error was related uniquely to a meteorological variable's structure function by Drozdov and Sepelevskij (1946). The important trade off between increasing station density and/or instrument quality was contributed by Póne (via Gandin et. al., 1967). After an elaborate theoretical study of network requirements, Bessemoulin (1960) determined that the objective analysis errors in differing network

densities were related to: (1) the atmospheric structure, (2) the instrumentation accuracy, and (3) the analysis technique.

By 1961 Gandin also related the sensor deployment problem to the objective analysis technique and recommended use of the optimum interpolation error instead of the linear interpolation error. Gandin (1963) felt that the advantage of optimum interpolation was not in its ability to minimize interpolation error variances, but in its ability to account for observational error effects. He stated that optimal interpolation made possible analyses whose accuracy exceeded that of the observations themselves. The reader can visualize the point by considering N observations from an analysis location, each possessing a known standard deviation σ. An analysis produced by a simple average of the observations has a standard deviation σ/N. A similar argument proceeds when the observations are relocated spatially. Optimum interpolation builds from this concept to produce analyses that are superior to all other interpolation schemes (fundamentally!) because variance between the analysis and the truth is minimized.

Additional information considered important in network design included (1) the use of auxiliary information such as parameter inter-correlations and data from multi-type sensors, (2) the actual station arrangement (and not only station density), (3) the observational frequency, and

(4) the likely meteorological scales probed by a proposed network. Gandin proposed sequential establishment of new stations at points where interpolation error variance was a maximum.

Other attempts at specifying optimal networks have been reported by Alaka (1970), Baer and Withee (1971), Steinitz et. al., (1971), Kasahara (1972), Alaka and Elvander (1972a), and Northrup et. al. (1972). The optimum measure usually includes a root-mean-square deviation of the resultant analyses from some arbitrary "true" analyses in the simulation experiments they performed. The Kasahara design algorithm, for example, referred to as Observing Systems Simulation Experiments (OSSE), is highly dependent on numerical model output (the true analyses) for "results often difficult to interpret."

C. Development of Multivariate Experimental Design

C.1 Fundamental Model Assumptions

The recent work in experimental design has been developed theoretically by Eddy (1973, 1974) and applied by Eddy (1976), Yerg (1973b), Kays (1974) and Brady (1975, 1976). The characteristic feature of their experimental design approach lies in its use of multivariate objective analysis. (A discussion of classical multivariate linear regression is given in Appendix A.) The objective analysis model represents an extension of classical multivariate linear

regression and is described by:

$$Y = X\beta + \varepsilon \tag{1}$$

where X is an n x m observation matrix, Y is an n x 1 predictand matrix, β is an m x 1 matrix of regression weights, ε is an n x 1 population error matrix, and $e = Y - \hat{Y}$ is the residual error matrix associated with sample regression. The noise or error matrix will reflect not only instrument errors and correlations, but geophysical noise representing scales (much more important!) not objectively analyzed, for one reason or another --- such as model inadequacy, multi-scale data, fine scale information. The objective analysis is given by:

$$\hat{Y} = X\hat{\beta} \tag{2}$$

where $\hat{\beta}$ estimates the population weights β and $\hat{Y}$ estimates the population values Y. Following the Appendix A discussion, $\hat{\beta}$ is found from:

$$\hat{\beta} = (X^t V^{-1} X)^{-1} X^t V^{-1} Y \tag{3}$$

where V is the error covariance matrix derived in Appendix A. Confidence in an analysis may be evaluated through:

$$\text{Var } \hat{\beta} = (X^t V^{-1} X)^{-1} \sigma^2 \tag{4}$$

or through:

$$\text{Var } \hat{Y} = V^{-1} X^t (\text{Var } \hat{\beta}) X V^{-1}. \tag{5}$$

Clearly, unless we know the nature of the residual variance σ^2, the confidence in an objective analysis cannot be evaluated!

The assumptions of this multivariate multiple linear regression objective analysis model are that:

(1) the underlying relationships between predictors and predictands are modelled (thus the need for predictand data, the grid point values, is avoided);

(2) the basic signal plus noise (the $[X^tV^{-1}X]$ and the $X^tV^{-1}Y$ of Eq. (3)) is modelled and referred to as the modelled covariance function. The function, expressed in terms of an observation set's spatial-temporal lag correlations, subsequently may be used to analyze that set or to place sensors optimally to sample phenomena assumed to have similar statistical characteristics;

(3) the covariance function is assumed homogeneous throughout a parameter field (points of equal distance-separation and the same relative direction have equal correlation). However, isotropy is not assumed because separation direction between two locations is a significant influence on the point-pair correlation coefficient (Thiebaux, 1975). The covariance fundamental to Eq. (3) is assumed by Eq. (1) to be a measure of the linear relationship that exists between two variates;

(4) the regression weights, which filter observations in space and time, are evaluated using the modelled

covariance function, thereby making the analysis methodology "portable";

(5) the model's residual errors are assumed distributed $\varepsilon \sim N(0,\sigma^2 V)$ where V is the error covariance matrix of Appendix A. Residual errors having the assumed distribution imply that a relative bias associated with any reporting instrument is assumed zero ($E[\mu_Y - \bar{\hat{Y}}] = E[\mu_1 - \bar{X}_1] = \ldots = E[\mu_n - \bar{X}_n] = 0$), and that the residual variance is assumed constant across the field ($\sigma_Y^2 = \sigma_1^2 = \ldots = \sigma_n^2 = \sigma^2$);

(6) the noise contributions are not necessarily mutually independent (see Appendix A);

(7) finally, the model in Eq. (1) is assumed adequate in the sense that ε does not contain coherent signal.

Extensive investigation of the interpolation results using multivariate analysis-of-variance techniques has shown $\Sigma e_i e_j = 0$. These results provide support of the analysis model's ability to produce unbiased results (Eddy and McDonald, 1977).

The residual variance in any objective analysis may be estimated from:

$$\hat{\sigma}^2 = e^t V^{-1} e = (Y - X\hat{\beta})^t V^{-1} (Y - X\hat{\beta}) \qquad (6)$$

If we let $A = X(X^t V^{-1} X)^{-1} X^t V^{-1}$, Eq. (6) may be written as:

$$\hat{\sigma}^2 = (Y^t - Y^t A^t) V^{-1} (Y - AY) \qquad (7)$$

which can be rewritten as:

$$\hat{\sigma}^2 = Y^t(I - A)\, V^{-1}\, Y \qquad (8)$$

since (I - A) is idempotent.

From Eq. (8) it is seen that the residual or unexplained variance is a function only of sampling positions, a concept developed in Section C.5 of this chapter.

C.2 The Covariance Function Formulation

The multivariate experimental design model implemented by Brady (1976) begins with a null hypothesis about the atmospheric signals to be sampled. The atmospheric covariance represented by the $X^tV^{-1}X$ and the $X^tV^{-1}Y$ in Eq. (3) is assumed to represent the population covariance for a given analysis or sampling problem. The covariance may be derived from an historical data set or from some numerical model output (Kreitzberg and Perkey, 1974). Clearly, the more the covariance function reflects the physical and statistical characteristics of the phenomenon being analyzed, the better the resultant analysis will be.

The k^{th}, ℓ^{th} element of the first term in Eq. (3) is given by:

$$(X^tV^{-1}X)_{k\ell} = \sum_{i}^{N} \sum_{j}^{N} X_{ik} X_{j\ell} V^{-1}_{ij} \qquad (9)$$

where $X_{ik}X_{j\ell}$ is the signal covariance between station k and station ℓ. V_{ij} is the noise anatomy. Using discrete time and space lags (typically a function of data density) raw correlation matrices (sometimes loosely referred to as covariance matrices) representative of the data structure are calculated. In the context of Eq. (1), correlations between stations the same distance apart contribute to the stability of any given correlation estimate. Then, the problem becomes one of incorporating the raw structure into the objective analysis/experimental design model. Brady (1976) used a functional relationship to express the climatological information content of these raw correlation matrices. A similar approach is used in this study.

The need for such a functional relationship can be justified readily:

(1) analysis accuracy is improved when the analysis scheme uses a functional relationship instead of sample covariances (Thiebaux, 1977);

(2) the function dispenses with data management problems presented by several thousand discrete space/time correlation values sufficient to discern a given system's atmospheric structure;

(3) dispenses with the interpolation of discrete lag correlation values at all possible observation-pair/separation-distance combinations;

(4) incorporates small scale perturbations and

inadequate data samples realistically by requiring a function fit in proportion to the observation-pair count used to determine each correlation coefficient;

(5) and, the function, through its parameters, more easily reflects relevant physical and statistical information of a system under study than is gained by pondering over raw correlation coefficient matrices.

A modified negative exponential function that resolves the stated needs was developed and is discussed in Chapter IV. An engineering solution to the data-function fit uses non-linear programming (NLP) to achieve a least-squares functional determination through appropriate parameter value adjustments. These parameters become the NLP algorithm's independent decision variables whose final values minimize the objective function Q:

$$Q = \sum_{i}^{D} N_i \, (r_i - f_i)^2 \qquad (10)$$

where N_i pairs enter into each of the D correlation values r_i, and f_i is the negative exponential function.

C.3 The Use of Non-Linear Programming

The NLP algorithm is a key element in the objective analysis/experimental design package (Yerg, 1973a). It is used in the covariance function parameterization and plays a vital role in the sensor placement problem. The algorithm implemented was developed initially by Spendley et. al. (1962).

Nelder and Mead (1964) added a variation to this sequential simplex method which in turn was extended and generalized by Paviani and Himmelblau (1969). An updated version of the technique described in Himmelblau (1972) is the one implemented here. This flexible tolerance method:

(1) is a direct search method (it does not use analytical approaches such a gradients and second derivatives) that improves objective function values by using information from feasible points and certain "near" feasible points;

(2) uses a flexible polyhedron with W + 1 vertices (for W independent decision variables), evaluates the objective function at each vertex and rejects the highest objective function value;

(3) performs a continual check on the constraint set violation by points used in the objective function evaluation;

(4) and, in the limit, allows only feasible points to remain in the solution space as the flexible polyhedron contracts to within a given tolerance of the solution.

The Paviani-Himmelblau method is used because (1) it was "the most reliable direct-search method" tested by Stocker (1969), (2) the algorithm was available readily, (3) Yerg (1973) and Eddy (1974) showed its potential, and Brady (1976) implemented it.

One final note: the justification of this powerful technique as part of the experimental design package lies,

not simply in its ability to solve the correlation-model-fitting problem, but principally in its ability to place sensors optimally in a highly constrained environment. Incorporation of the NLP algorithm into a sensor placement algorithm is discussed in Section C.5.

C.4 The Objective Analysis Algorithm

Full utilization and implementation of the multivariate methodology necessitated expansion of the univariate model "represented" in Eq. (1). As a demonstration, bivariate analyses are produced using:

$$\hat{Y} = X_1\hat{\beta}_1 + X_2\hat{\beta}_2 \qquad (11)$$

where, for example, X_1 (rain gage observations) and X_2 (reflectivity precipitation estimates) are used to estimate $\hat{Y}$ (surface rainfall). In the bivariate case, two covariance functions and two cross-covariance functions are required:

$$COV(X_1X_1), \quad COV(X_2X_2), \quad COV(X_1X_2), \quad COV(X_2X_1)$$

All four functions are evaluated using Eq. (3) with actual use in Eq. (11) determined by the parameter mix (gage measurements and/or reflectivity estimates) producing the "best" analysis $\hat{Y}$ (best is determined through the multiple correlation coefficient). The extension of Eq. (11) to a multivariate case is clear.

Actual implementation of the algorithm proceeds

in several steps. For any given predictand location:

(1) variables enter into a possible predictor set if they are positively correlated with the predictand. This bit of engineering prevents the $X^tV^{-1}X$ matrix from becoming singular and difficult to invert when a disproportionate share of predictors have negative correlations;

(2) if the predictors found exceed program dimensions, selected predictors are eliminated using stepwise regression (Efroymson, 1960), the appropriate correlation function and its "radius of influence," defined as that spatial/temporal distance from a predictand to the negative exponential function's zero value;

(3) respective elements of the $X^tV^{-1}X$ matrix and the $X^tV^{-1}Y$ vector are computed and the regression coefficients determined. Thus the correlation function serves to determine not only predictor locations, but the actual model filtering properties;

(4) finally, the objective analysis using Eq. (2) or Eq. (11) is performed.

Extensive use of the analysis methodology reveals positive characteristics:

(1) data set structures compare favorably with objectively analyzed structures (Brady, 1976);

(2) signal-to-noise ratios are significantly larger in analyzed data sets than in raw data sets (Brady, 1976);

(3) the technique, compared to other analysis methods, performs best with minimum data persistence (Lacy, 1973). This point is supported strongly by Phillips (1976) and Thiebaux (1977);

(4) methodology apparently encounters no data boundary problems (Pasteris, 1975) compared to some analysis schemes (Brandes, 1975);

(5) residual errors essentially are Gaussian distributed (Eddy, 1967);

(6) analysis residuals exhibit no significant space autocorrelation implying the analysis abstracted the significant data-set signal and filtered most of the data-set noise (Eddy, 1967, and Eddy and McDonald, 1977). Additional positive characteristics are indicated by Alaka and Elvander (1972b) and Alaka (1974). The analysis technique and the NLP algorithm intermesh to produce a sensor placement package described in the next section.

C.5 The Sensor Placement Algorithm

Previously, Eq. (8) indicated that an analysis' unexplained variance was a function only of sampling positions. The result may be stated also in terms of the multiple correlation coefficient, R, which represents the proportion of variance of Y accounted for by $\hat{Y}$. By definition:

$$R = \frac{\text{Cov}(Y,\hat{Y})}{\{(\text{Var } Y)(\text{Var } \hat{Y})\}^{\frac{1}{2}}} \tag{12}$$

Now if $\sigma_Y = (Y^t V^{-1} Y)^{\frac{1}{2}}$ and $\sigma_{\hat{Y}} = (\hat{Y}^t V^{-1} \hat{Y})^{\frac{1}{2}}$, Eq. (12) may be written as:

$$R = \frac{Y^t V^{-1} \hat{Y}}{(Y^t V^{-1} Y)^{\frac{1}{2}} \, (\hat{Y}^t V^{-1} \hat{Y})^{\frac{1}{2}}} \tag{13}$$

If the results of Eq. (3) are used in Eq. (13), our result becomes:

$$R^2 = \frac{Y^t V^{-1} \hat{Y}}{Y^t V^{-1} Y} \tag{14}$$

The best objective analysis is obtained when the correlation between Y and $\hat{Y}$ is maximized or when $(1 - R^2)$ is minimized. Stated another way, the modelled residual variance, $(1 - R^2)\sigma^2$, unaccounted for by the regression and for any one grid point, depends upon the atmospheric structure, the error covariance matrix, and the objective analysis technique and may be minimized by an appropriate sensor deployment. The residual variance is modelled from the same components required to obtain $\hat{\beta}$ in Eq. (3). For an entire grid array and for any given set of sensor locations, the weighted mean residual variance is given by:

$$F = \frac{1}{G} \left[\sum_{i=1}^{G} \eta_i \, (1 - R_i^2)\sigma^2 \right] \tag{15}$$

where G is the number of grid points, η_i weights each grid point by its relative worth (as suggested by an a priori

climatology), and $\Sigma\eta_i = 1$. The Eq. (15) summation constitutes an objective function minimized using the NLP algorithm discussed previously through appropriate sensor relocation, subject to logistic and engineering constraints if necessary. The objective function represented by Eq. (15) is referred to as the experimental design model; model development is completed in Chapter VII and Appendix B.

The optimal design of a network requires that relative values be assigned to the data collected by competing networks. Eq. (15) indicates that the weighted sum of residual variances can be minimized by NLP to yield an optimum network. One realization from Eq. (15) becomes a measure of the relative value of that particular proposed network.

A particular network's relative value is obtained as follows: Given a set of perfect sensors and an adequate model, one station located at the space-time analysis location would predict exactly the predictand signal; there is no residual variance. Given an imperfect sensor set, the covariance function lag-zero correlation value (referred to as the A_7 parameter) now is less than 1.0. Thus, the one imperfect station located at the same space-time analysis location would predict a value proportional to A_7; the residual variance in this case would be proportional to $(1 - A_7)$. The same logic can extend to include all other proposed station locations, each evaluated for its effect on

the one predictand location; and then, the residual variance evaluated at all locations is summed as in Eq. (15). The NLP algorithm systematically shifts station locations, subject to any constraint set, until it discovers that placement set producing the minimum objective function. Sensors are located optimally in an expected value sense when climatology is used as the a priori weights in Eq. (15). Thus, actual data realizations are not needed to evaluate a proposed network's sampling qualities!

The extensive testing by Yerg (1973b) demonstrate merits of the approach in that:

(1) optimum sampling is most beneficial when system size is relatively small compared to the analysis grid;

(2) optimal sampling becomes more important as the signal-to-noise ratio decreases;

(3) optimal sensor deployment always reduces the objective analysis residual variance from that obtained using a good intuitive sensor deployment!

Additional characteristics are shown in Chapter VIII.

CHAPTER IV

COVARIANCE FUNCTION AND OBJECTIVE ANALYSIS ALGORITHMS: A REFINEMENT

A. The Covariance Model -- A Multivariate Gaussian-Damped Function

The fundamental concept of optimum objective analysis is it's use of data-set statistical structure to determine, not only the analysis model's filtering properties, but also the actual selection of predictor stations; thus, the necessary correlation analysis is not an end unto itself. The logical and straightforward approach toward an optimum analysis is through use of an analytical function assumed to model the statistical structure. Because analysis accuracy is highly dependent on the form of the correlation function (Lacy, 1973; Thiebaux, 1975 and 1977), care must be given to the analytical function's selection (Julian and Thiebaux, 1975; Brady, 1976; and Thiebaux 1976).

A covariance function model simple and compact, yet explicitly derivable from statistical principles, was desired

possessing the following characteristics:

(1) the function must be four dimensional (to delineate spatial and temporal determined correlation covariance) and capable of becoming negative at increasing lags;

(2) the analytical function must be capable of modelling possible asymmetry in a data set's structure. Thiebaux (1976 and 1977) indicated the deficiencies of isotropic weighting functions used in objective analysis. When anisotropy was included, she demonstrated a significant deviation reduction (74%) in observed correlation values around the best fitting isotropic model. Consequently, anisotropy will be permitted in the covariance model by allowing each Cartesian coordinate direction "influence radius" to assume different values in the respective positive and negative directions. Anisotropy in the time domain, however, is not modelled. While there is no difficulty in principle in accounting for the statistical structure of developing or dissipating storm systems, it remains to be shown that increased computer time and associated complexity would be compensated by a more accurate analysis;

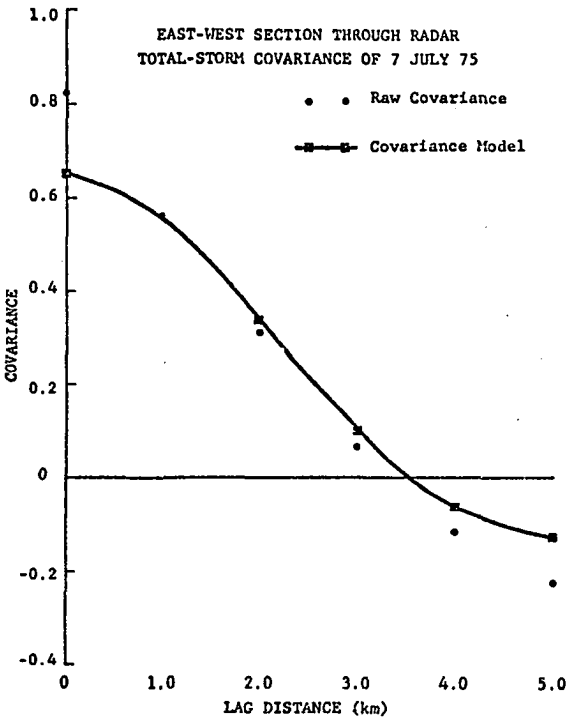
(3) in general, however, covariance anisotropy is not conveniently oriented along a coordinate axis. The study extends previous work and proposes that anisotropic covariance patterns, not oriented along a Cartesian coordinate axis, represent an additional significant analysis-

error-effect. Typically, such elliptical covariance data sets reflect system physical characteristics or result from data-set trends and predominant storm motion-tracks;

(4) the function must represent accurately the raw signal covariance (Fig. 3).

In the example, relatively large deviations at lag-zero between values of the covariance function and the raw covariance matrix does not indicate the model's inability to model parameter structure; rather, the model rejects the raw lag-zero covariance value because that value does not represent the data set's basic signal-to-noise ratio. The unrealistically high value resulted from observations that contained highly correlated information because they were closest to the radar, and hence, had smaller radial separations than did the more distant observations.

Function accuracy is demonstrated using three different forms of the covariance model. The accumulated rainfall covariance of 7 July 75 (2000 GMT) became the test case (see also Fig. 23). Anisotropic function-fit, constrained to be oriented along a coordinate axis, reduced the weighted least squares error $\sim$ 70% over the best fitting isotropic model. An additional significant deviation reduction of $\sim$ 12% (to $\sim$ 82%) was obtained when the covariance model permitted the elliptical field to be skewed with respect to the coordinate axis. It should be noted



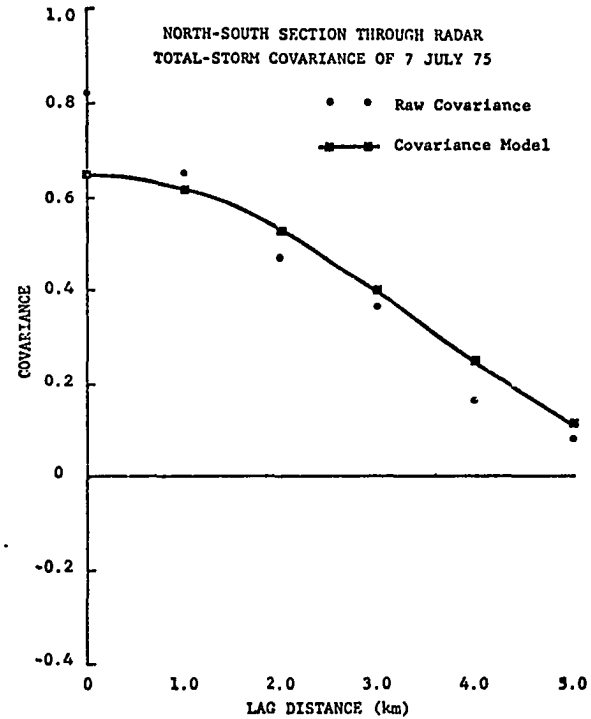


Fig. 3. Covariance model accuracy suggested by the sample covariance values. Function influence radii are 3.55 km east-west and 6.03 km north-south; and the function lag-zero covariance is 0.65.

that function parameter values are influenced strongly by raw covariance characteristics and the actual data volume used in a least squares fit. For example, the lag-zero correlation value is somewhat lower and the system influence radius somewhat longer the more the function is "influenced" by negative covariances which appear at large lag values.

A multivariate Gaussian-damped function possesses these characteristics, resolves the stated needs of Chapter III and is given by:

$$f(xyzt) = A_7 * \cos\left[\frac{2\pi\sqrt{K_1+K_2}}{4}\right] * \exp -\left[\frac{K_1}{2(1-\alpha^2)} + \frac{K_2}{2}\right] \quad (16)$$

where:

$$K_1 = \left[\frac{x^2}{\sigma_x^2} - \frac{2\alpha xy}{\sigma_x\sigma_y} + \frac{y^2}{\sigma_y^2}\right] \quad (17)$$

and:

$$K_2 = \left[\frac{z^2}{\sigma_z^2} + \frac{t^2}{\sigma_t^2}\right] \quad (18)$$

Eq. (16) represents the linear covariance between locations at separation distances of (x-y-z-t) and is generated from raw covariance matrices (Eq. (9)) assuming homogeneous and stationary parameter mean values. In this study, the positive X axis is east, positive Y is north, and positive Z is up while T is the time-domain axis. Function parameters in Eq. (16) are the NLP algorithm's independent

decision variables and are given by: σ_x, σ_y, σ_z, and σ_t which represent the analysis model's influence radii in the respective X-Y-Z and T directions. Study notation is given by: $\sigma_x^+ = A_1$, $\sigma_y^+ = A_2$, $\sigma_x^- = A_3$, $\sigma_y^- = A_4$, $\sigma_z^+ = A_5$, and $\sigma_z^- = A_6$ which permit anisotropy and determine the covariance function's size and shape; by A_7 which is the function's lag-zero correlation value; by $\alpha = A_8$ which reflects covariance function ellipticity not oriented along an X-Y plane coordinate axis (discussed shortly); and by $\sigma_t = A_9$ which is the time influence radius. Their final values are the result of a least-squares functional determination. The covariance function for two-dimensional (x-y) and three-dimensional data sets (x-y-t) may be expressed also by Eq. (16) when simplified through omission of appropriate components.

For the data sets processed, only covariance function ellipticity in the X-Y plane will be free to orientate as indicated by the data set characteristics ($\alpha_{xy} = \alpha$). The coefficient α becomes a measure of the product-moment correlation of the density contained in the X-Y covariance field. Anisotropy in the X-Z, Y-Z, or time domain planes will be constrained to be oriented along a respective coordinate axis (i.e., $\alpha_{xz} = \alpha_{yz} = \ldots = \alpha_{zt} = 0$). Consequently, some convective storm characteristics, such as storm tilt with height, are not modelled. The constraints imposed on the covariance modelling characteristics do not compromise the study effort since the major thrust is in the optimal sampling and analysis

of accumulated rainfall patterns. Thus, the four-variate normal function takes the form of Eq. (16).

Tatsuoka (1971) indicates the major axis of a bivariate normal curve is oriented at the following angle with respect to the positive X axis (east):

$$\theta = \begin{cases} \frac{1}{2} \text{ ARCTAN } \dfrac{2\alpha\sigma_x\sigma_y}{\sigma_x^2 - \sigma_y^2} & \text{when } \sigma_x \neq \sigma_y \\ \pm 45^{\circ} & \text{when } \sigma_x = \sigma_y \end{cases} \qquad (19)$$

The coefficient (α) is synonymous with the bivariate normal's correlation coefficient. The study refers to the orientation angle (θ) as the storm's preferred track, even though system physical characteristics, data set trends or actual storm tracks could produce similar covariance patterns. Both the orientation angle (θ) and the coefficient (α) assist in modelling covariance ellipticity in the X-Y plane; thus, Eq. (19) reflects orientation of the covariance function to the X axis.

The function has certain symmetries which always over-simplify the storm to be sampled. Clearly, every storm is different and whereas the function can mold itself to the main features of each storm, the rainshafts (which seem to occur nearly randomly in space and time) appear to the analysis as high frequency, high power noise. In addition,

there are three basic reasons why an accurately calibrated radar will not measure rainfall correctly and hence, produce cross-correlations between radar reflectivity patterns and surface rainfall patterns which often are weak: (1) the radar pulse volume is not full; (2) an incorrect Z-R relationship is not used; and (3) dry subcloud air may alter significantly the radar-revealed rainfall as it descends from the radar pulse volume to the surface. Moreover, storm scale asymmetries cause problems which are compounded when asymmetries in the radar data are transposed and modified by the time the rain reaches the ground.

The results of the non-linear program fit yield: (1) signal-to-noise ratios; (2) gage/radar calibration information; (3) storm size and lifetime; (4) orientation and motion direction; and (5) certain asymmetries in the storm structure.

Two covariance modelling deficiencies remain; first, storm motion-direction and speed has not been incorporated properly, thereby occasionally permitting corrupted horizontal spatial influence radii. A second known deficiency is the model's inability to permit maximum cross-covariance values at non-zero spatial and temporal lags. In other words, the model currently cannot express completely the atmosphere's physics inherent in intensifying or dissipating systems nor in situations with time-lagged responses between multivariate predictors (e.g., radar

indicated rainfall does not immediately reach ground level but is modified by wind flow patterns while enroute). This study is most concerned with total-storm-rainfall patterns; at this point these modelling deficiencies are not considered serious. However, these two deficiencies remain and are points for additional study.

A third point of future concern involves sampling and analysis of phenomena which contain interacting action scales (e.g., tornado cyclone circulation embedded and imposed upon a larger parent thunderstorm circulation). The current covariance/objective analysis model does not incorporate directly such realistic features. The analysis model (Eq. (1)) is no longer correct under such multiple scale circumstances and produces estimates biased by an alias matrix which contains secondary action-scale information (Draper and Smith, 1966).

Climatological characteristics of the covariance model are illustrated by the function parameter values in Chapter VI.

B. The Analyzed Signal -- A Perturbation About Its Mean

It is quite probable that multivariate predictor data sets do not exhibit the same population characteristics of equal means and variances. The analysis algorithm (Eq. (1)) must be generalized to model these characteristic differences by including a $\hat{\beta}_0$ term in the $\hat{\beta}$ matrix. For any

given predictand location, modelling of parameter means and variances is accomplished in several steps:

(1) predictor locations and corresponding weighting coefficients are determined as in Chapter III.C.4;

(2) each predictor data value is transformed to a standardized variable by using the appropriate parameter mean and pooled standard deviation $\{\tilde{x}_i = (x_i - \bar{x}_i)/\sigma_{x_i}\}$. The resulting data sets represent normalized signal plus noise perturbations;

(3) the selected predictor values are filtered to determine the final signal perturbation $\hat{Y}_p = \sum_i (x_i \hat{\beta}_i)$;

(4) finally, the estimated but standardized signal perturbation is scaled back to predictand magnitudes $(Y = \hat{Y}_p * \sigma_p + \bar{X}_p)$.

Thus, the modelling effort allows non-stationary parameter mean values to be incorporated in time series analysis. If realistic differences in parameter means and variances are considered, straightforward and impartial filtering of the multivariate data sets follow; and each predictor is allowed its proper influence in modulating the final analyzed field. The result is an analysis that possesses parameter characteristics in proportion to the strength of existing parameter interrelationships. As an example, bivariate radar/rain gage analyses will be shown to produce surface rainfall estimates from gage magnitudes modulated by observed radar patterns. The methodology (which

does go far beyond that proposed by Brandes, 1975) produces analyses that possess the desirable statistical properties of a maximum likelihood estimator; and it is determined from properly chosen, properly weighted predictors! First guess fields and other superfluous handwaving is not required.

The influence of the covariance model chosen on analysis accuracy was evaluated using the 7 July 75 radar data set. Rainfall estimates at each bin location were determined withholding the predictand bin observation from the analysis. Analysis error variances were determined using a covariance model that was: (1) constrained to be isotropic; (2) anisotropic but constrained to be orientated along a coordinate axis (i.e., $\alpha = 0$); and (3) anisotropic with no overriding restrictions (i.e., Eq. (16)). Analysis results of filtering the observations, using the latter covariance function, served as the standard. The actual anisotropy in the test data set is only slightly off a north-south orientation (illustrated later in Fig. 23).

Interpolation error variances increased only 0.3% when observations were filtered using a covariance model which was anisotropic along a coordinate axis. However, isotropic data analysis increased error variances by 4%. Indeed these results are significant when one considers the high data density (1040 bins spaced 0.5 km radially and 1^{o} azimuthally) and the fact that Lacy (1973), Yerg (1973b), and Phillips (1976) have found the optimum interpolation methodology to

perform best, compared to other methods, in sparse data areas. Also, the brief demonstration of covariance model influence on analysis accuracy confirms the results of Thiebaux (1977) who provided evidence of 7-16% interpolation error reductions for the multivariate scheme using anisotropic weighting over the same scheme using a Gaussian correlation function.

Objective analysis results are presented in Chapter VI.

CHAPTER V

THE GENESIS OF A CONVECTIVE STORM CLIMATOLOGY FROM SOUTHEAST MONTANA

A. Covariance-Generation Data Requirements

This study focuses primarily on the analysis and data sampling requirements needed to evaluate weather modification experiments. The methodology is universal; it only awaits implementation in other areas. Logically, the field experiment design begins with an early quantitative description of the phenomenon considered amenable to sampling. The definitive statement, expressed in terms of the statistical properties of convective-storm surface-rainfall estimates, will be derived from space-time covariance and cross-covariance functions. Also, the statement might take the form of preferred storm tracks, frequencies of occurrence, and seasonal and geographical variations. A climatological investigation of these properties using modelled covariance functions is started in this Chapter. All conclusions rest on the data sets investigated.

The first natural questions to arise center on the type of data sets necessary to approximate population statistical properties of surface rainfall estimates. Evidence presented later strongly suggests that storm covariances shown by gage observations are unreliable. Primarily, gage-determined covariances are unreliable because spatial (3.6 km average spacing) and temporal (15 minutes) gage sampling resolutions are inadequate to determine the covariance, even though the gages may be adequate for later objective analyses.

Thus, radar derived precipitation estimates become the prime source for revealing convective storm characteristics. Data requirements are evaluated using the 7 July 75 total-storm-rainfall data set. The covariance function determined from the original data set (1040 bins each covering approximately 0.325 km^2) will serve as the standard against which value judgements are made. Data set density was decreased progressively (all bins, then every other bin, every third, etc.) to approximately 10% of the original data (2.91 km^2/bin) and the covariance function was determined. Function parameters were compared to their original values. Differences, between the newly determined and original parameter values, of less than $\sim$ 10% were assumed subjectively to indicate no substantial difference between the two covariance functions. Most parameters became unstable and differed by more than 10% when less than 15% of the original

data remained for function determination and became difficult to determine when only 10% remained. A plot of the lag-zero covariance value (Fig. 4) suggests the inference. Therefore, it is concluded that the covariance function can be defined with acceptable accuracy using as little as $\sim$ 15% of available radar data (or range bins separated by $\sim$ 3.0 km) at the lowest elevation angle. Additional spatial data-requirement reductions become possible when three dimensional data sets (x-y-t) are examined.

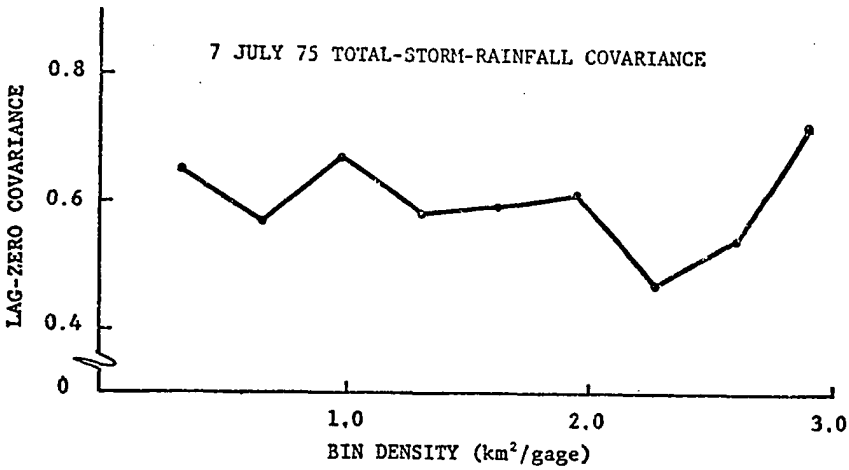


Fig. 4. Covariance at lag-zero determined from data sets whose density was progressively decreased. Parameter instability increased as bin-separation distance approached the storm influence radius.

It is interesting to notice that the average 3.6 km gage separation distance is just slightly longer than the 3.0 km bin separations required for adequate radar covariance determination. Widely spaced gage sensors, whose observing qualities can be suspect, were simply unable to detect the small spatial characteristics of Southeast Montana systems. Chaotic patterns in the raw covariance matrices do suggest a basic inability by the Miles City gage network to provide the convective storm structure.

Finally, the effect of using a different Z-R relationship was investigated. Using the Marshall-Palmer relationship of $Z = 200R^{1.6}$ and a modified Jones relationship of $Z = 3 * (486R^{1.37})$, the radar total-storm-rainfall was recalculated, the covariance functions were determined and compared with the original function. The modified Jones-determined function was identical to the original function; the Marshall-Palmer data set did produce very minor, almost random differences. Although only one data set has been processed in this manner, it does suggest that the Z-R relationship effect on covariance function determination is not substantial while only the α in $Z = \beta R^{\alpha}$ exerts an influence on analysis results.

B. Objective Analysis Data Requirements

Of equal importance are the data requirements necessary to produce analyses sufficiently accurate to permit

satisfactory discrimination between two storm analyses obtained when population rainfall amounts differed by a specified amount (Eddy, 1976). The radar data requirements were evaluated using the rain gage/radar 7 July 75 total-storm-rainfall and an approach similar to that in Section A. Gage data requirements are inferred later in Chapter VIII. At this place it is sufficient to understand that gage importance in analysis accuracy is minimal provided "enough" surface observations are available to adjust biased radar precipitation estimates adequately.

However, since previous results suggested gage-determined covariance functions as unreliable, it is important to evaluate the influence, beyond calibration, that gage observations exert on an analysis. Gage influence was assessed in the following manner. A grid-point bivariate surface rainfall analysis was performed by filtering gage observations with a gage-determined covariance function. Raw correlation matrices, determined using analysis grid-point values as input data, were used to calculate a "filtered" gage covariance function. (The lag-zero covariance value was left purposely unchanged since it indicated expected signal-to-noise ratios.) The filtered gage function was used in a second bivariate surface rainfall analysis and the same process repeated. At each step the filtered covariance function was compared with original function-parameter values.

Alternatively, it seems reasonable to assume that a gage network should reflect the same statistical structure as that shown by a corresponding radar data set. Thus a second series of bivariate surface rainfall analyses were initiated and used a gage covariance function that derived its shape from the radar data sets and its magnitude from the surface data.

Both procedures were repeated with the consequence that the two filtered gage covariance functions converged on what was essentially a common function and nearly identical in shape to the original radar structure. Thus, the storm covariances revealed by the Miles City network are again concluded as unreliable. Gage observations in this study are filtered accordingly by radar-derived storm structure.

Bivariate surface rainfall estimates at each gage location were determined (referred to as a station objective analysis), withholding the predictand gage observation from the analysis and using progressively decreased radar data-set densities. Analysis error and predictand variances were used to form a signal-to-noise ratio (SNR). The SNR determined using all original data became the standard against which value judgements were made.

The analysis model recovered progressively less signal, but the SNR reductions became less significant, as bin-separation distance increased (Fig. 5). Categorical

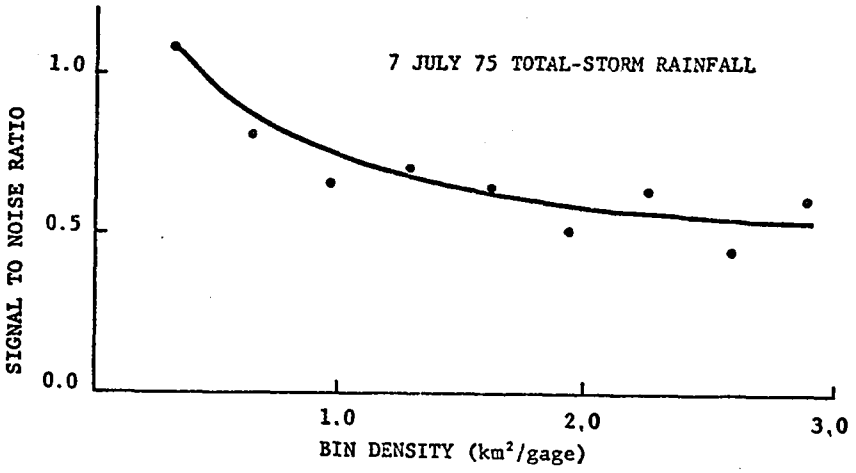


Fig. 5. Signal-to-noise ratios of station (58 gages) bivariate surface rainfall analyses determined from radar data sets whose density was progressively decreased. Solid curve represents a best fit of actual data values (circles).

interpretation (Fig. 5) of the optimum data-set size would require knowledge of increased signal recoverability costs compared with the benefits derived therefrom. However, it is plain that increasing data density in a sparse observational network (say from 3.0 km^2/bin to 2.0 km^2/bin) would provide minimal improvements in recovering the large scale signal. In fact, the realistic finer scale convective storm features can be recovered only when most of the available data is used. Thus, for a given sensor deployment, analysis accuracy appears more highly dependent on available data than does covariance-function determination.

In addition to data set densities, analysis results are controlled strongly by the cross-covariance between gage and radar observations. Cross-covariance influence was evaluated using the SNR from station objective analyses. The results (Table 2) clearly show the significant radar influence on surface precipitation analysis. Nevertheless, too much influence by the radar ($\rho(0) = 0.84$) or too little influence ($\rho(0) = 0.42$) gives sub-optimal results. One must use a proper sensor mix for optimum analysis results.

PREDICTOR VARIABLES	COVARIANCE SOURCE	SIGNAL-TO-NOISE RATIO
Gages only	Gage data	0.29
Gages only	Pattern from radar Magnitude from gage	0.15
Radar plus gages	Cross-covariance $\rho(0) = 0.42$	0.56
Radar plus gages	Cross-covariance $\rho(0) = 0.63$	1.06
Radar plus gages	Cross-covariance $\rho(0) = 0.84$	0.96

Table 2. Station surface precipitation analysis: radar influence controlled by varying the lag-zero cross-covariance value.

Also, the effect on analysis accuracy of using a different Z-R relationship was investigated. Using the Marshall-Palmer relationship and a modified Jones relationship of $Z = 3 * (486^{1.37})$, the radar total-storm-rainfall was recalculated, a bivariate station analysis performed and the SNR calculated. The modified Jones-determined SNR was identical with the original SNR; the Marshall-Palmer derived SNR was improved slightly. The implication seems that the β of $Z = \beta R^{\alpha}$ is unimportant for analysis results and that only the α exerts an influence. In fact through the SNR, the analysis model provides an objective, realistic approach toward evaluation of an optimum Z-R relationship. Although only one data set has been processed in this manner, it seems to suggest that the Z-R relationship effect on analysis accuracy, when using the analysis methodology developed in this study, is minimal.

CHAPTER VI

COVARIANCE CLIMATOLOGICAL STUDY AND OBJECTIVE ANALYSIS RESULTS

A. Univariate Objective Analysis Results

The climatological characteristics of surface rainfall estimates from Southeast Montana are dominated by the statistical structure from the many radar data sets processed. Radar univariate analysis results and covariance-function plots, typical of Southeast Montana, are illustrated graphically (Section A); climatological characteristics are discussed in Section B. Bivariate radar/rain gage analyses follow (Section C) and illustrate the analysis algorithm's power and capability.

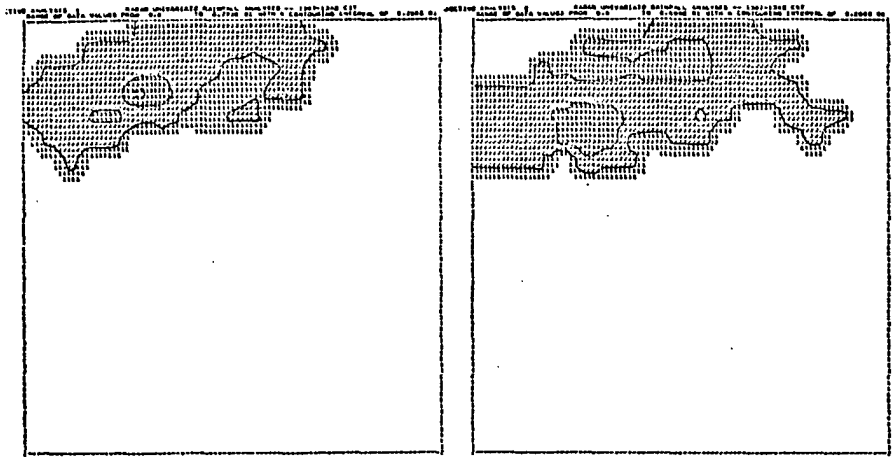
Univariate analyses of three organized lines and two embedded convective cells follow. Also presented is a typical set of rainfall patterns analyzed from a four-dimensional data subset; these results illustrate the model's data assimilation capability. All figures presented have

true north toward the page top. The storms' motion-directions and speeds were determined by tracking movements of the salient storm features. Unless otherwise noted in figure captions, analysis grids are 22.5 km x 22.5 km and covariance plot mesh sizes are 1.0 km x 1.0 km.

On 5 July 75 a well organized convective line moved southward toward Miles City and produced maximum steady rainfall rates near 10 mm hr^{-1} (Fig. 6). Little intensity or feature change was noted as the system progressed steadily southward. By 1338 GMT (not shown), cells moving from 350^{o} and 010^{o} at 25 km hr^{-1} merged into a distinct core and accounted for the total rainfall swath observed (Fig. 7A). Maximum analyzed rainfall was less than 2.0 mm.

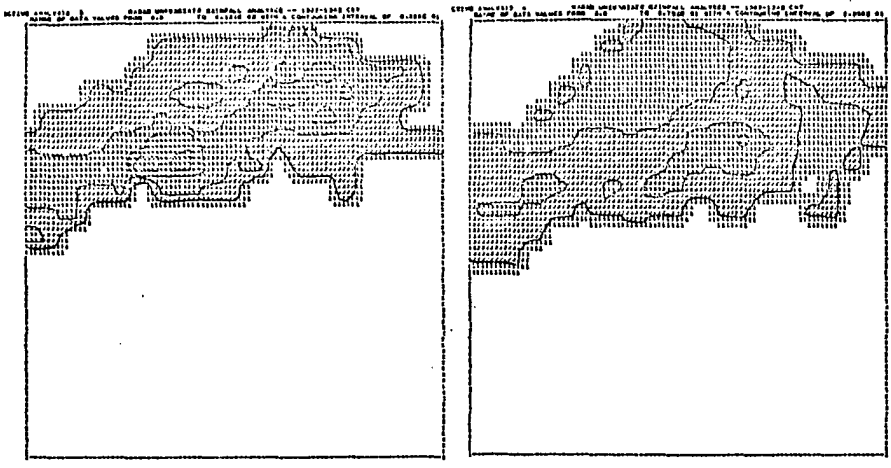
On 31 July 75 a northwest-southeast squall line moved rapidly northward. Maximum rainfall rate observed was 12.7 mm hr^{-1} (Fig. 8B). Rapid system movement resulted in minimal surface rainfall accumulations (Fig. 9A). Analyzed amounts were barely 1.0 mm. Notice the highly skewed and anisotropic squall line features reflected in the optimal filtering function (Fig. 8E).

The third squall line presented (18 August 75) intensified rapidly (Fig. 10D) to rates exceeding 110 mm hr^{-1} as it moved quickly east northeastward. Rainfall rates diminished to near 75 mm hr^{-1} by 2238 GMT (Fig. 10E). Total-storm-rainfall exceeded 10 mm (Fig. 11) and was oriented in northwest-southeast bands that were coincident with successive



6A: 1307 GMT

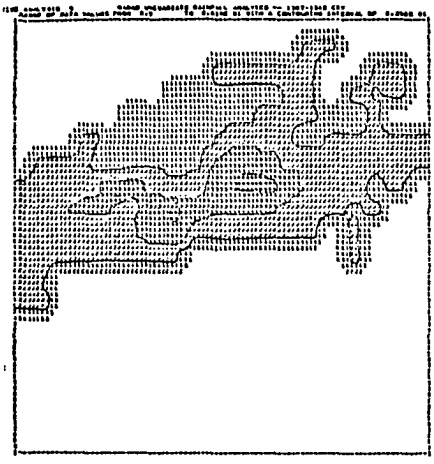
6B: 1312 GMT



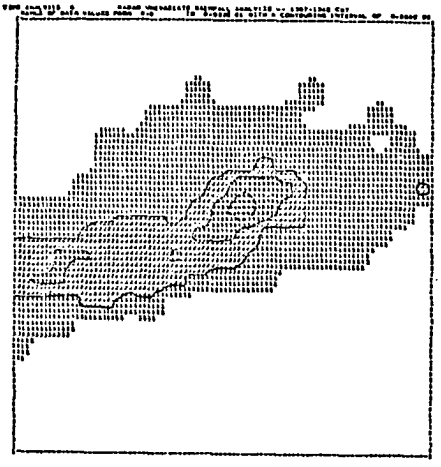
6C: 1317 GMT

6D: 1322 GMT

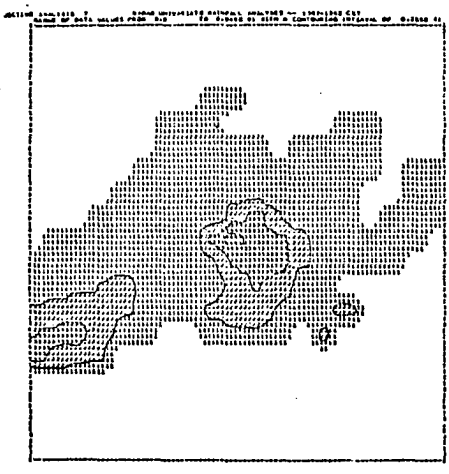
Fig. 6 (A-H). Radar univariate analysis of 7 July 75 squall line moving from 350° at 25 km hr^{-1}. Isohyet contouring interval is 2.5 mm hr^{-1}. Modelled covariance function values (H) range from 0.0 to 0.40.



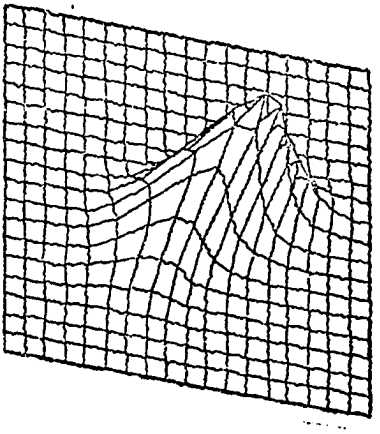
6E: 1327 GMT



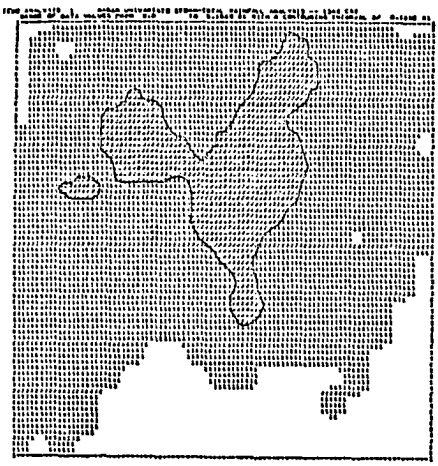
6F: 1332 GMT



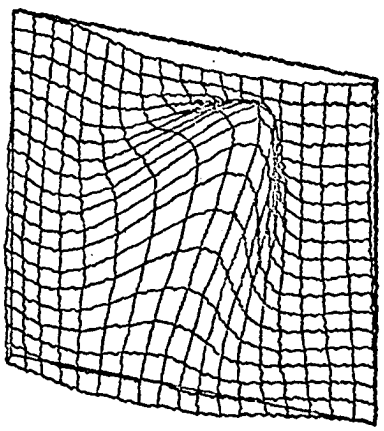
6G: 1337 GMT



6H: System Covariance



7A: 1342 GMT

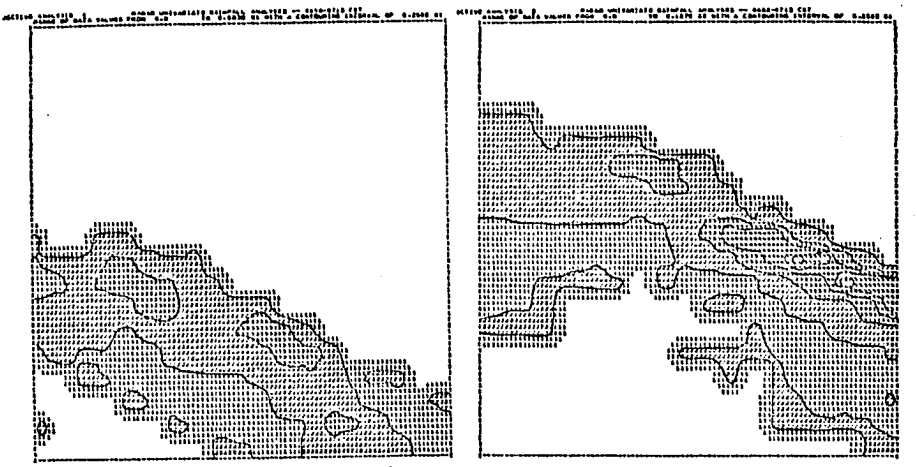


7B: System Covariance

Fig. 7 (A-B). Radar univariate total-storm-rainfall analysis of 7 July 75 squall line. Isohyet contouring interval is 1.0 mm. Modelled covariance function values (B) range from -0.094 to 0.460.

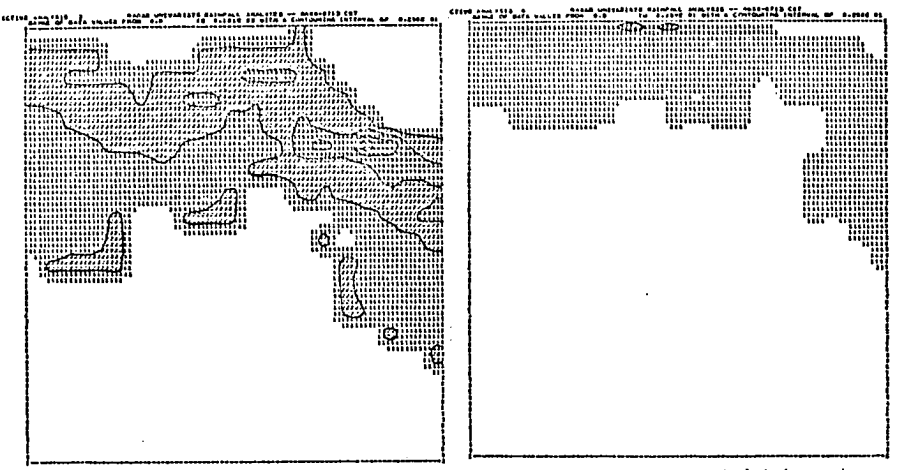
line positions. An explanation for the precipitation banding is offered in Section B.

The model's four-dimensional data-assimilation capability is demonstrated by interpolation analyses on horizontal grid arrays placed a desired distance above ground level (for the 18 August 75 squall line of Figs. 10-11 now illustrated in Fig. 12). Interpolation estimates derive from predictor information (in a four-dimensional data subset) filtered in proportion to the predictor's spatial-temporal separation distance from the analysis location. The plotted results become a true constant-altitude planned-position-indicator (CAPPI) analysis of the radar information. These



8A: 0658 GMT

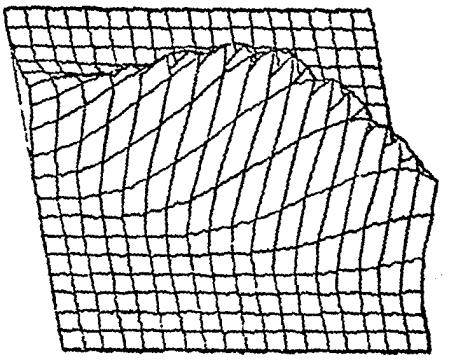
8B: 0703 GMT



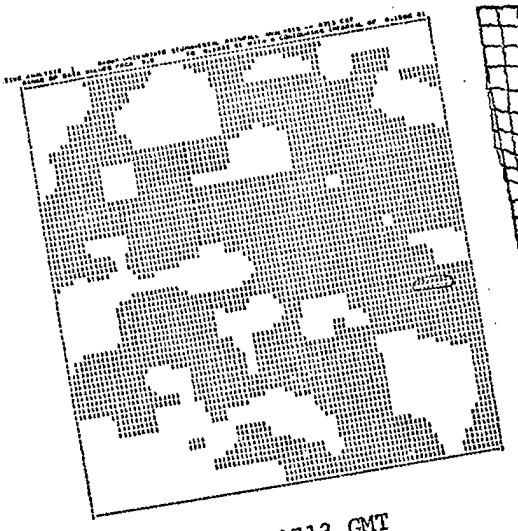
8C: 0708 GMT

8D: 0713 GMT

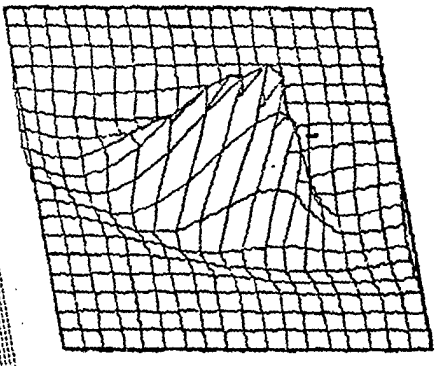
Fig. 8 (A-E). Radar univariate analysis of 31 July 75 squall line moving from 190° at 74 km hr^{-1}. Isohyet contouring interval is 2.5 mm hr^{-1}. Modelled covariance function values (E) range from 0.0 to 0.40. Covariance mesh is 0.75 km square.



8E: System Covariance

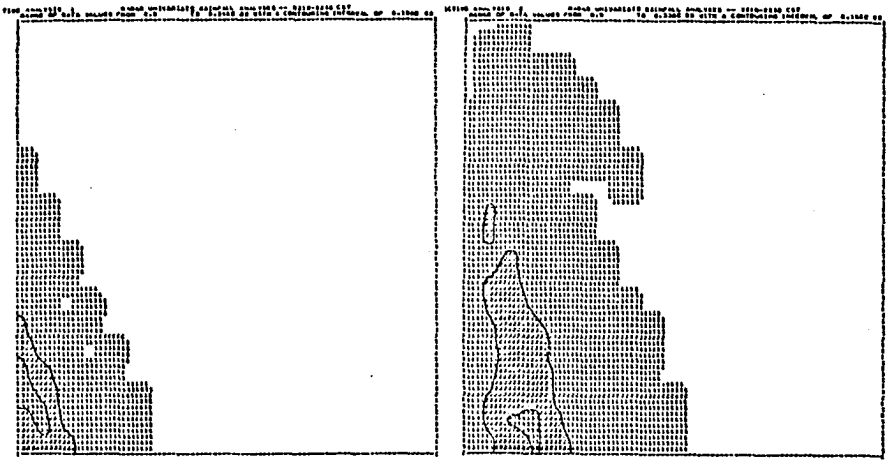


9A: 0713 GMT



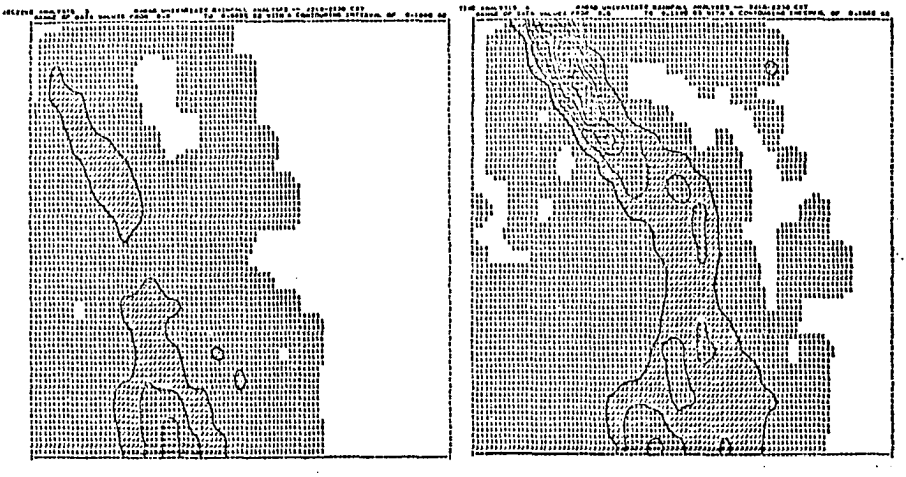
9B: System Covariance

Fig. 9 (A-B). Radar univariate total-storm-rainfall analysis of 31 July 75 squall line. Isohyet contouring interval is 1.0 mm. Modelled covariance function values (B) range from -0.065 to 0.440.



10A: 2218 GMT

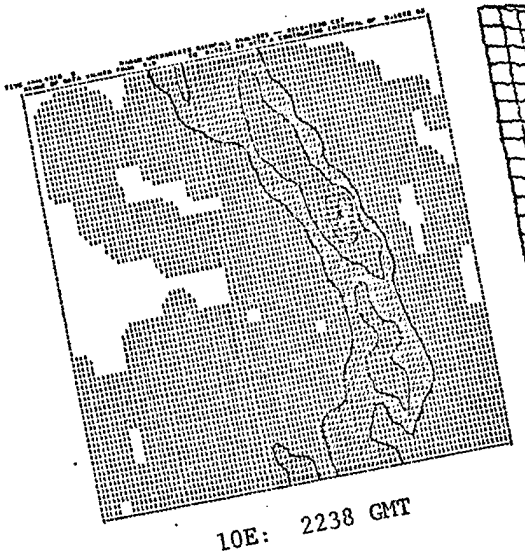
10B: 2223GMT



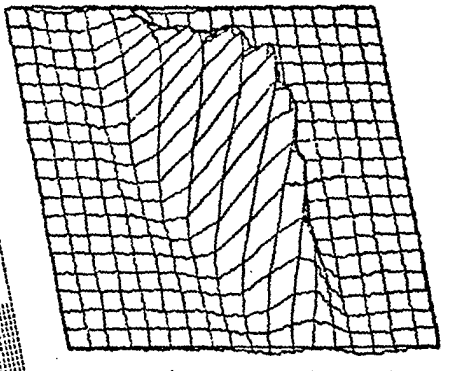
10C: 2228 GMT

10D: 2233 GMT

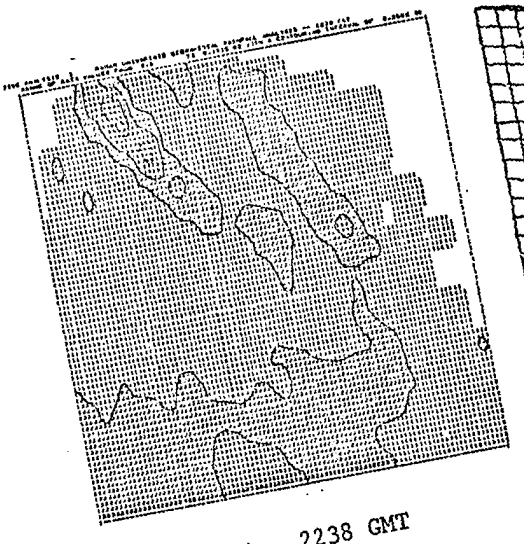
Fig. 10 (A-F). Radar univariate analysis of 18 August 75 squall line moving from 250° at 59 km hr^{-1}. Isohyet contouring interval is 15.0 mm hr^{-1}. Modelled covariance function values (F) range from -0.031 to 0.650.



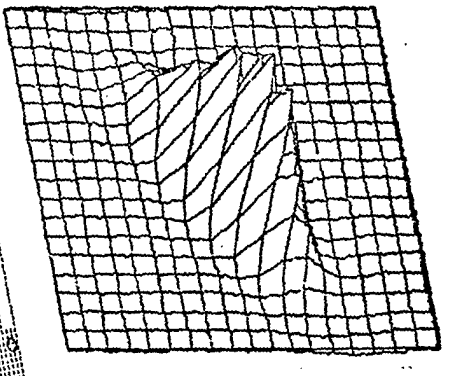
10E: 2238 GMT



10F: System Covariance

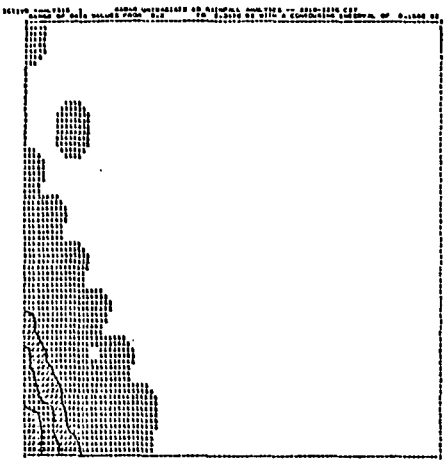


11A: 2238 GMT

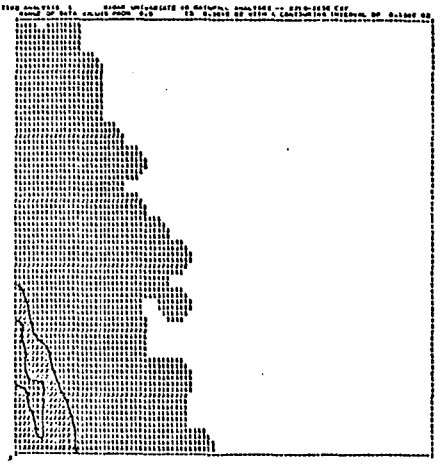


11B: System Covariance

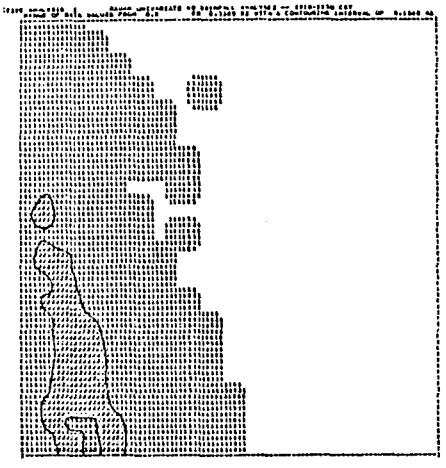
Fig. 11 (A-B). Radar univariate total-storm-rainfall analysis of 18 August 75 squall line. Isohyet contouring interval is 2.5 mm. Modelled covariance function values (B) range from -0.026 to 0.610.



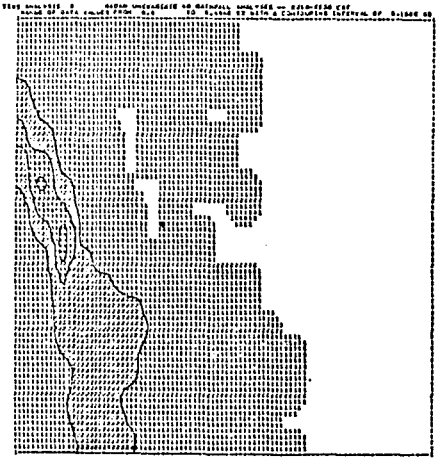
12A: 2218 GMT, 1.0 km



12F: 2218 GMT, 2.25 km

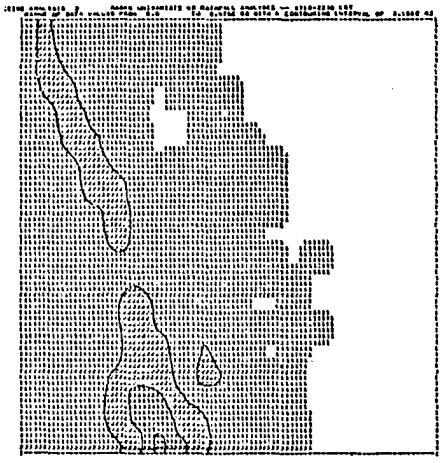


12B: 2223 GMT, 1.0 km

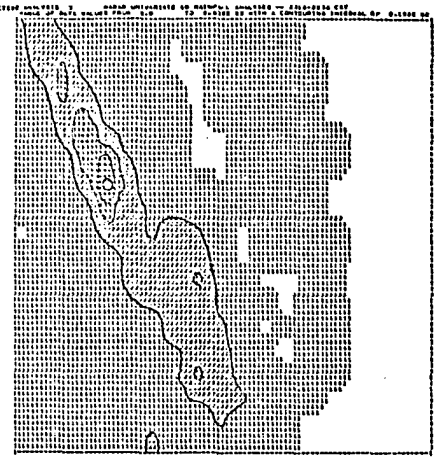


12G: 2223 GMT, 2.25 km

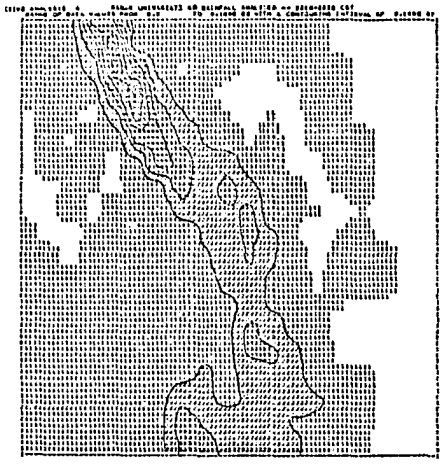
Fig. 12 (A-J). Constant Altitude PPI (CAPPI) analysis of the Fig. 10 squall line obtained from a four dimensional data subset, its Table 3 covariance function, and the data assimilation capability in the objective analysis model. Isohyet contouring interval is 15.0 mm hr^{-1}.



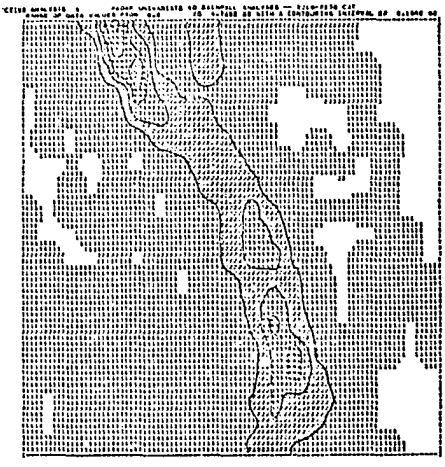
12C: 2228 GMT, 1.0 km



12H: 2228 GMT, 2.25 km



12D: 2233 GMT, 1.0 km

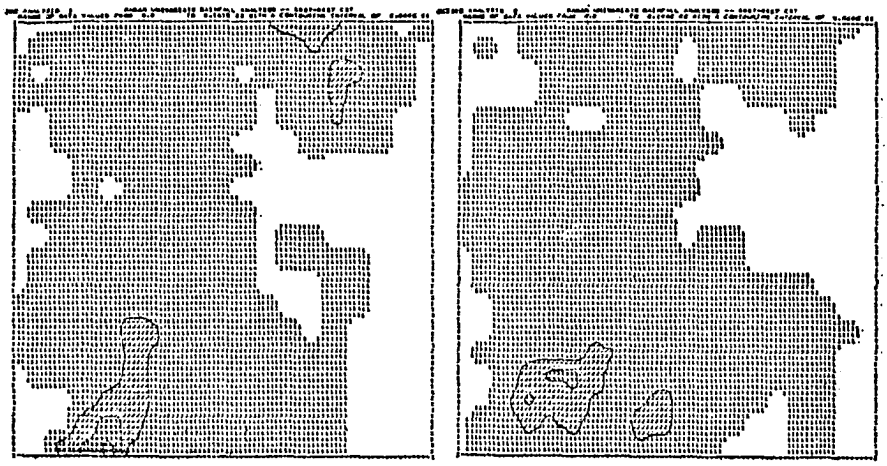


12I: 2233 GMT, 2.25 km

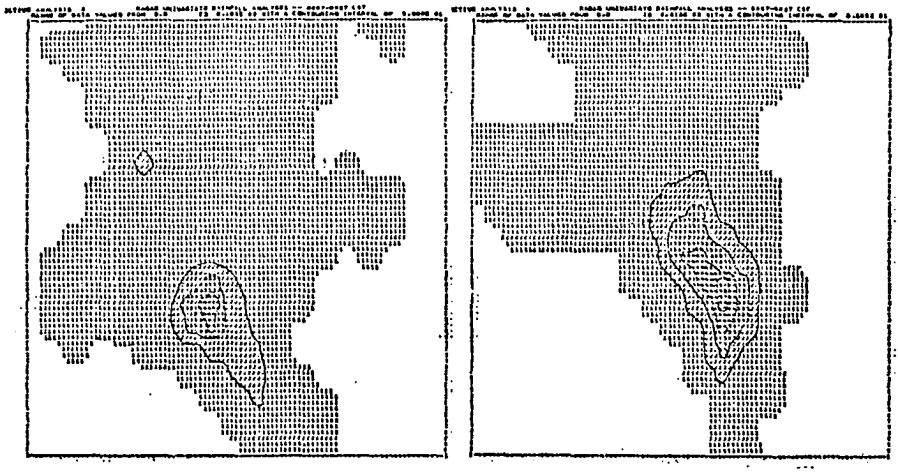
12E: 2238 GMT, 1.0 km 12J: 2238 GMT, 2.25 km

results may be termed "true" because some investigators allow only the closest observation to become the final value at a CAPPI grid location. Other investigators simply average the closest N observations for a grid point estimate. Features commonly associated with advancing convective systems are noted. Storm intensities are stronger 2.25 km into the storm (Figs. 12F-J) than at cloud base (Figs. 12A-E). Lower level features are retarded somewhat over the corresponding entities aloft.

Late on 17 July 75 and on the south side of a large rain mass, a small cell became organized, intensified, and moved quickly northeastward (Fig. 13). Rainfall rates reached 21.6 mm hr^{-1} by 0022 GMT (Fig. 13D). A well defined precipitation maximum trailed the storm-core path. Rapid cell movement prevented accumulations from exceeding 3.0 mm (Fig. 14).

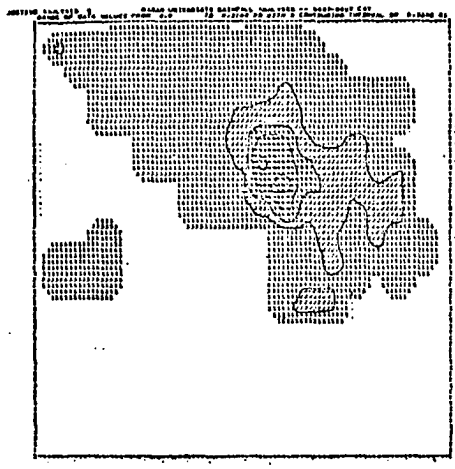


13A: 0007 GMT 13B: 0012 GMT

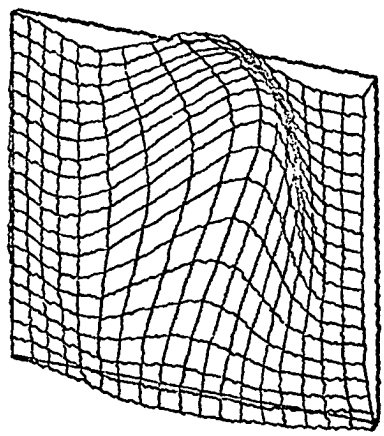


13C: 0017 GMT 13D: 0022 GMT

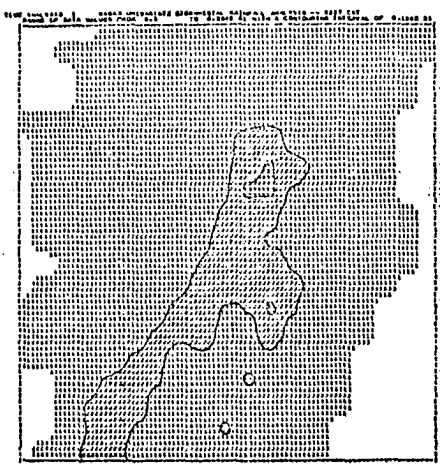
Fig. 13 (A-F). Radar univariate analysis of 17 July 75 embedded thunderstorm moving from 220° at 53 km hr^{-1}. Isohyet contouring interval is 5.0 mm hr^{-1}. Modelled covariance function values (F) range from -0.080 to 0.350.



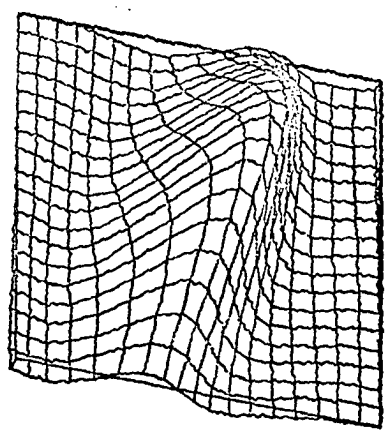
13E: 0027 GMT



13F: System Covariance

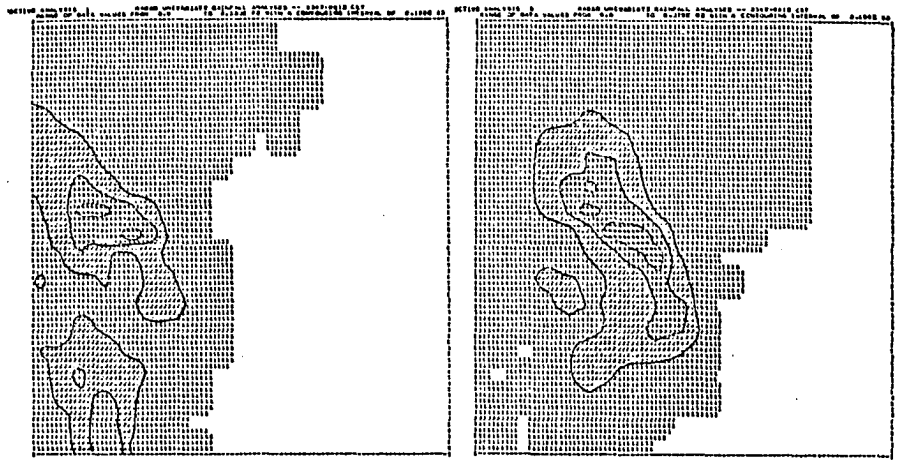


14A: 0027 GMT

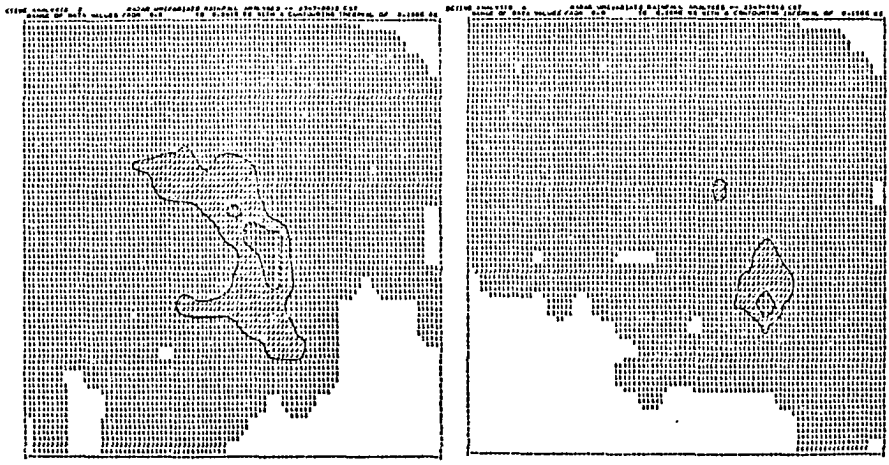


14B: System Covariance

Fig. 14 (A-B). Radar univariate total-storm-rainfall analysis of 17 July 75 embedded thunderstorm. Isohyet contouring interval is 1.0 mm. Modelled covariance function values (B) range from -0.044 to 0.340.

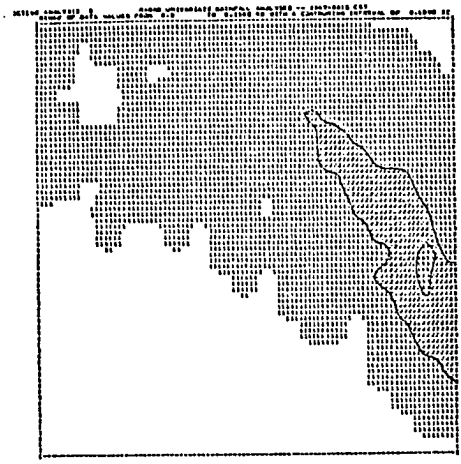


15A: 2347 GMT 15B: 2352 GMT

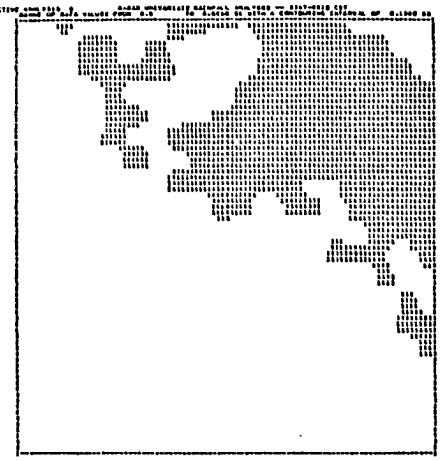


15C: 2357 GMT 15D: 0002 GMT

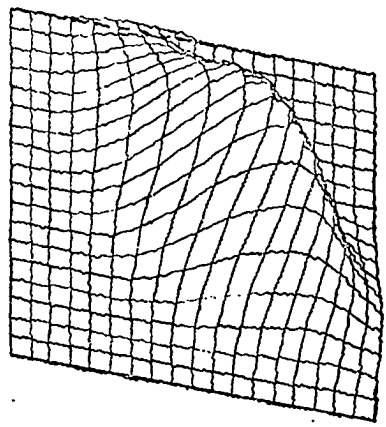
Fig. 15 (A-G). Radar univariate analysis of 18-19 August 75 embedded thunderstorm moving from 270° at 45 km hr^{-1}. Isohyet contouring interval is 10.0 mm hr^{-1}. Modelled covariance function values (G) range from -0.003 to 0.600.



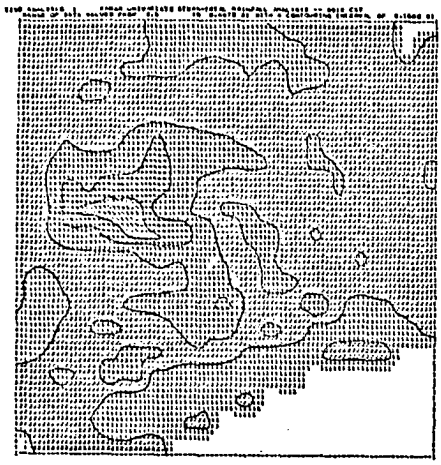
15E: 0007 GMT



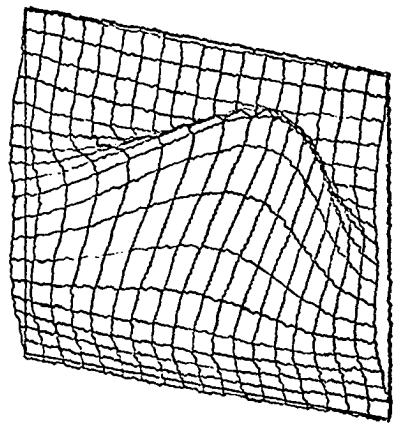
15F: 0012 GMT



15G: System Covariance



16A: 0012 GMT



16B: System Covariance

Fig. 16 (A-B). Radar univariate total-storm-rainfall analysis of 18-19 August 75 embedded thunderstorm. Isohyet contouring interval is 1.0 mm. Modelled covariance function values (B) range from -0.130 to 0.570.

An embedded convective system (possibly a short line) moved quickly toward Miles City on the afternoon of 18 August 75. At 2352 GMT, after reaching maximum intensity of 38 mm hr^{-1} (Fig. 15B), the system weakened slowly to 25 mm hr^{-1} by 0007 GMT (Fig. 15E) Surface rainfall accumulations reflected the weakening trend of the eastward moving system (Fig. 16A). Maximum observed precipitation was almost 5.0 mm.

This study documents (Section B) the covariance model's capability to reflect system characteristics. The reader is urged to study the Table 3 and 4 summaries and contrast the covariance function plots: for (1) small (Fig. 9B) and large (Fig. 15G) systems; (2) symmetrical (Fig. 20F) and anisotropic (Fig. 23B) systems; and (3) anisotropic systems

oriented off the coordinate axes (Figs. 8E, 14B, and 15G) and parallel to a major axis (Fig. 10F).

B. Covariance Climatological-Study Results

The statistical properties of surface rainfall estimates from Southeast Montana derive from storm structure implied by observed data and then deduced by using space-time covariance and cross-covariance functions. These properties were investigated using a variety of convective storm systems (Table 1). Included in the study were dissipating and intensifying squall lines and isolated convective cells moving from a variety of motion-directions and speeds. Study results are summarized in Table 3 (three and four dimensional covariances), Table 4 (total-storm-rainfall covariances) and Table 5 (motion-direction and speeds and calculated preferred storm track (Eq. (19)). A sensor array orientated along the preferred storm track of a convective system will permit the optimal sampling and analysis of similar phenomena; this point is clearly shown in Chapter VIII. Please note that two dimensional data sets always possess symmetric covariance properties of the form $A_1 = A_3$ and $A_2 = A_4$. System physical characteristics are termed "coherent" over those spatial or temporal lag distances that exhibit positive covariance. Drufuca and Zawadzki (1975) provide similar definitions of pattern coherency in their study of the climatological features in a large rain gage data set.

MILES CITY MONTANA RAINFALL X-Y-Z-T COVARIANCE STATISTICS

DATE	EAST	NORTH WEST	SOUTH	UP	DOWN	CORR	ELIPSE	TIME	AREA	ROTATE	SNR
5 JULY 75 1930-2000 GMT	6.7	4.9	3.5	5.5	0.0	0.0	0.65	-0.30	41.3	83.2	-28.5	1.86
7 JULY 75 1307-1342 GMT	12.0	4.2	7.5	8.0	0.0	0.0	0.40	0.97	11.0	187.2	33.6	0.67
7 JULY 75 1900-2000 GMT	4.1	4.0	4.3	5.2	0.0	0.0	0.59	-0.53	18.6	61.1	-54.0	1.44
17 JULY 75 0007-0027 GMT	5.6	10.1	3.8	5.6	0.0	0.0	0.35	0.07	10.0	115.5	86.8	0.54
17 JULY 75 1545-1615 GMT	5.2	5.8	3.6	3.9	0.0	0.0	0.59	-0.92	23.3	67.5	-48.5	1.44
21 JULY 75 2252-2322 GMT	14.7	3.6	4.7	3.4	0.0	0.0	0.45	0.83	19.8	107.2	11.6	0.82
30 JULY 75 2003-2018 GMT	5.1	5.1	4.1	4.9	0.0	0.0	0.41	0.58	23.6	73.3	45.2	0.69
31 JULY 75 0658-0713 GMT	12.4	5.1	10.0	4.6	0.0	0.0	0.40	-0.94	6.6	170.9	-21.4	0.67
18 AUGUST 75 2213-2238 GMT	2.8	5.8	2.8	5.8	0.0	0.0	0.65	-0.79	3.3	50.8	-67.1	1.86
19 AUGUST 75 2347-0012 GMT	9.6	9.9	5.9	7.1	0.0	0.0	0.60	-0.91	6.8	207.0	-45.9	1.50
5 JULY 75 1930-2000 GMT	6.6	5.2	3.6	5.9	7.7	7.7	0.56	-0.40	40.9	89.7	-38.2	1.27
30 JULY 75 2003-2018 GMT	5.0	5.1	4.2	4.9	6.0	6.1	0.44	0.67	20.6	73.2	46.3	0.79
31 JULY 75 0658-0713 GMT	8.9	4.5	9.5	4.8	5.2	5.2	0.46	-0.94	4.5	134.9	-26.0	0.85
18 AUGUST 75 2213-2238 GMT	3.4	8.7	3.5	8.8	3.0	3.1	0.63	-0.88	3.3	95.9	-70.0	1.70
6 JUNE 76 0530-0600 GMT	6.9	9.7	5.9	7.6	0.0	0.0	0.69	-0.39	35.1	175.3	-65.6	2.23
7 JUNE 76 0430-0445 GMT	4.8	6.5	5.3	7.2	0.0	0.0	0.55	-0.42	36.4	108.7	-63.3	1.22
7 JUNE 76 0515-0545 GMT	12.5	8.3	12.8	8.5	0.0	0.0	0.71	-0.33	10.2	334.2	-18.8	2.45
PARAMETER AVERAGES	7.4	6.3	5.6	6.0	5.5	5.5	0.54	-0.27	18.6	125.6	-19.1	1.29
STANDARD DEVIATIONS	3.5	2.1	2.7	1.0	1.7	1.6	0.11	0.64	12.8	69.4	45.8	0.57

Table 3. Covariance climatological-study results from investigating the Table 1 storm data sets. Parameters correspond to the variables A_1 through A_9 in Chapter IV.A. Units are km and minutes. Covariance area, preferred storm track (Eq. (19)), relative data-set signal, and parameter averages and standard deviations are included.

MILES CITY MONTANA STORM-TOTAL RAINFALL COVARIANCE STATISTICS

DATE	EAST	NORTH	WEST	SOUTH	UP	DOWN	CORR	ELIPSE	TIME	AREA	ROTATE	SNR
5 JULY 75 2000 GMT	4.8	5.2	4.8	5.2	0.0	0.0	0.79	-0.33	0.0	78.5	-52.6	3.76
7 JULY 75 1342 GMT	4.1	6.0	4.1	6.0	0.0	0.0	0.46	0.30	0.0	76.6	71.0	0.85
7 JULY 75 2000 GMT	3.5	6.0	3.5	6.0	0.0	0.0	0.65	-0.37	0.0	67.2	-73.1	1.86
17 JULY 75 0027 GMT	4.1	8.1	4.1	8.1	0.0	0.0	0.34	0.59	0.0	105.4	70.3	0.52
17 JULY 75 1615 GMT	3.6	4.4	3.6	4.4	0.0	0.0	0.51	-0.47	0.0	49.9	-55.9	1.04
21 JULY 75 2322 GMT	12.1	3.6	12.1	3.6	0.0	0.0	0.72	0.23	0.0	136.6	4.2	2.57
30 JULY 75 2018 GMT	4.4	5.0	4.4	5.0	0.0	0.0	0.65	0.65	0.0	69.0	50.9	1.86
31 JULY 75 0713 GMT	3.9	2.2	3.9	2.2	0.0	0.0	0.44	-0.54	0.0	27.4	-20.9	0.79
18 AUGUST 75 2238 GMT	2.5	3.8	2.5	3.8	0.0	0.0	0.61	-0.80	0.0	29.9	-58.7	1.56
19 AUGUST 75 0012 GMT	7.2	3.8	7.2	3.8	0.0	0.0	0.57	0.08	0.0	84.6	3.4	1.33
6 JUNE 76 0600 GMT	5.2	7.1	5.2	7.1	0.0	0.0	0.78	-0.32	0.0	115.8	-67.5	3.55
7 JUNE 76 0445 GMT	4.7	6.8	4.7	6.8	0.0	0.0	0.53	-0.18	0.0	101.6	-77.1	1.13
7 JUNE 76 0545 GMT	6.4	4.6	6.4	4.6	0.0	0.0	0.64	-0.20	0.0	94.2	-15.5	1.78
PARAMETER AVERAGES	5.1	5.1	5.1	5.1	0.0	0.0	0.59	-0.10	0.0	79.8	-17.0	1.74
STANDARD DEVIATIONS	2.3	1.6	2.3	1.6	0.0	0.0	0.13	0.43	0.0	30.8	51.7	0.97

Table 4. Covariance climatological-study results obtained from investigating the accumulated rainfall patterns of the Tables 1 and 3 storm data sets. Parameters correspond to the variables A_1 through A_9 in Chapter IV.A. Units are km and minutes. Covariance area, preferred storm track (Eq. (19)), relative data set signal, and parameter averages and standard deviations are included.

The climatological characteristics (Tables 3 and 4) suggest a very close relationship exists between parameter values of the covariance function and convective-storm system dimensions. Physical features such as system size and asymmetry, motion-direction and speed, and calculated preferred storm track reflect similar features in a system's statistical structure. The implication seems that optimal sampling and analysis is strongly dependent upon a system's physical characteristics.

The average cloud base area within the 20, 25, and 30 dBz reflectivity contours was calculated for the three dimensional storm data sets and time intervals in Table 3. The corresponding covariance function area was also calculated as the function's ellipsoidal area within the spatial influence radii (i.e., the area containing positive covariance at time lag zero). A simple correlation between the covariance function area and the cloud base area within the 20, 25, and 30 dBz reflectivity contours produced correlation coefficients of 0.45, 0.49, and 0.52, respectively. (Selected storm data sets were eliminated subjectively from this and following computations for several reasons--unable to determine motion speed, system life-time too short). The results presented suggest that weak precipitation rates are not part of the salient storm statistical structure. Moreover, the results indicate covariance function responses to cardinal system features.

The simple correlation between the storm motion-speed and the covariance time influence radius (also referred to as the decorrelation time) was -0.76. Without regard to storm size, the strong implication is that fast moving storms decorrelate quickly. As an example, the 31 July 75 squall line (a large system illustrated in Fig. 8) moved at 74 km hr^{-1} and decorrelated in 6.6 minutes. A statistically smaller storm on 18 August 75 (Fig. 10) moved at 59 km hr^{-1} and decorrelated in 3.3 minutes. At the other extreme, the 5 July 75 convective cell (not shown) drifted southeastward at 11 km hr^{-1} and was coherent statistically for over 40 minutes. These results indicate that small storms which move slowly are equivalent statistically in time to faster moving large storms. The simple correlation between motion speed, normalized by the storm's spatially coherent dimensions, and decorrelation time verifies the point (correlation coefficient of 0.79) and is a step toward universal applicability of these results.

Motion-speed, even for a well defined intense system, is a significant influence on the total-storm-rainfall footprint. The 31 July 75 system moved northward at 74 km hr^{-1} (Fig. 8), and produced maximum rainfall rates near 15 mm hr^{-1}. However, the rainfall left in its wake appeared as a collection of random noise (Fig. 9). The covariance model reflected the lack of spatial coherency and indicated influence radii (Table 4) significantly below the climatological

average. A simple correlation between total-storm-rainfall covariance area and storm motion-speed (correlation of -0.54) points to the same conclusion.

Proper interpretation of the parameter averages requires recognition of the small sample size which contains individual cells and organized lines, weak systems and stronger ones. Nonetheless, squall-line covariance statistics do dominate parameter averages, principally because organized lines are about 40% more spatially coherent than are individual cells (average covariance area of 160 km^2 vs 116 km^2). The high variability in the east-west influence radius (Table 3) and the relative high uncertainty in implied covariance orientation (standard deviation of 0.64) is attributed to non-homogeneous system averaging. Four of five highly elongated convection lines moved either northward or southward (Table 5); the more symmetrical convective cells moved from a basic westerly direction.

The primary area of investigation involves the design of a multivariate mesoscale field experiment, which is expressed in terms of optimal sensor density and orientation requirements. Now the point is addressed partially; later a more comprehensive discussion follows (Chapter VIII). The motion direction of a convective system and its calculated optimal sampling orientation (Eq. (19)) is presented in Fig. 17. At first glance the plot looks unimpressive and suggests only a weak positive relationship. Two (of five)

DATA	MOTION DIRECTION AND SPEED (km hr^{-1})	TYPE SYSTEM	SAMPLING--ORIENTATION (Table 3)	(Table 4)
5 Jul 75	310/11	Cell	299°	323°
7 Jul 75	350/25	Line	236	199
7 Jul 75	330/22	Line	324	343
17 Jul 75	220/53	Cell	183	200
17 Jul 75	240/45	Cell	319	326
21 Jul 75	240/22	Cell	258	266
30 Jul 75	?	Line*	225*	219*
31 Jul 75	190/74	Line	291	291
18 Aug 75	250/59	Line	337	329
19 Aug 75	270/45	Cell(?)	316	267
6 Jun 76	?	Cell*	336*	338*
7 Jun 76	180/21	Dissipated*	333*	347*
7 Jun 76	190/51	Line	289	286
MODAL VALUE	240/45		289-291°	286-291°

Table 5. Type convective system (Table 1), motion-direction and speed, and optimal sampling orientation (positive with respect to north = 360° and equal to 270° - Tables 3 or 4 results). Superscript table values (*) not used in modal value determination.

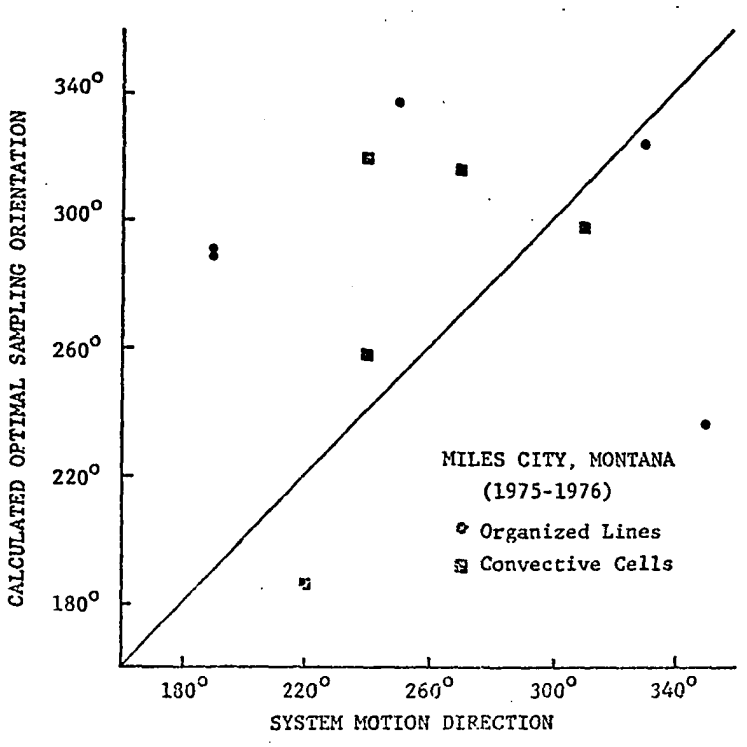


Fig. 17. The relationship between convective system motion-direction and the sensor orientation for optimal sampling. Organized line (circles) and individual cell (squares) results taken from Table 5. Solid line is for user reference.

cells depart from the relationship; conceivably, one cell could have been classified appropriately as an embedded line (Fig. 15). The other cell (not shown but from 17 July 75, 1545-1615 GMT) was of marginal quality primarily because of temporal data gaps.

The relationship between a convective line's

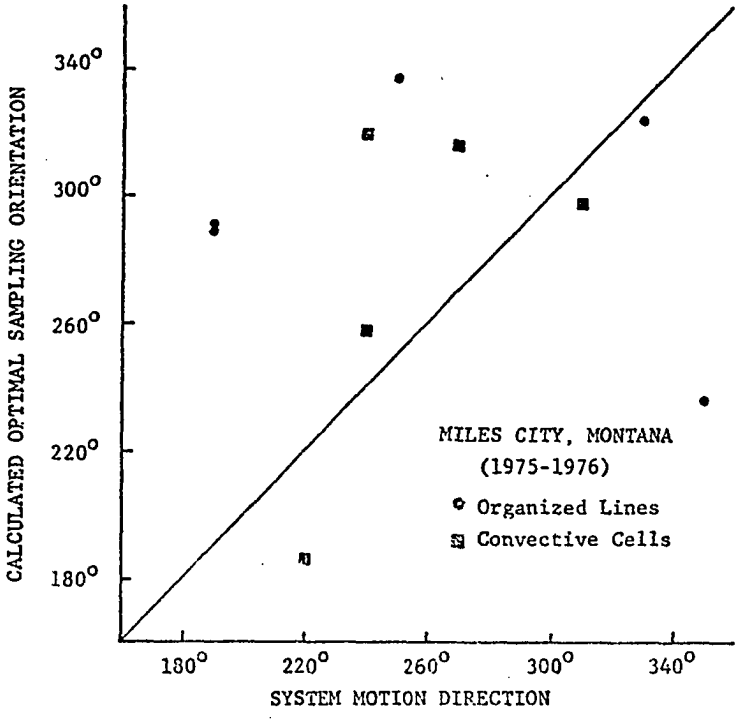


Fig. 17. The relationship between convective system motion-direction and the sensor orientation for optimal sampling. Organized line (circles) and individual cell (squares) results taken from Table 5. Solid line is for user reference.

cells depart from the relationship; conceivably, one cell could have been classified appropriately as an embedded line (Fig. 15). The other cell (not shown but from 17 July 75, 1545-1615 GMT) was of marginal quality primarily because of temporal data gaps.

The relationship between a convective line's

motion-direction and the sensor orientation that permits optimal sampling is certainly not obvious. Only the 7 July 75 squall line indicated a close relationship between covariance orientation and system motion-direction (see Fig. 20); primarily that relationship resulted from a well defined strong cell curving southeast through the analysis grid at the same basic velocity as the parent squall line. A closer subjective inspection (Figs. 6, 8 and 10) supports strongly the calculated orientation as that of the organized line. In most cases the line orientation appears orthogonal to its movement path. Yet, implications are that convection line sampling does benefit from the near non-orthogonality of covariance ellipticity and system motion-direction. The results are encouraging considering the limited sample and the inelegant methodology used to determine a system's velocity.

But what is the relationship of these parameters to that orientation necessary to observe optimally an accumulated rainfall pattern? Figure 18 provides convincing evidence that the sensor orientation, which produces optimal sampling in a moving convective system, is also best to observe the system's accumulated rainfall pattern--an important concept supported by nearly identical modal values of the preferred storm track in accumulated rainfall patterns and in the orientation of moving convective systems (Table 5). Eleven of thirteen Table 5 moving-system orientations differed by less

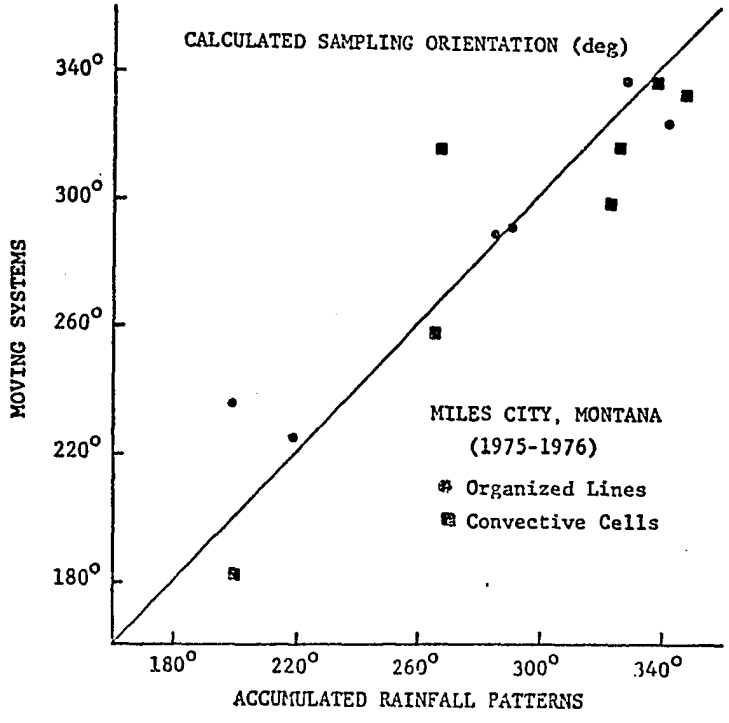


Fig. 18. The relationship between the sensor orientation for optimal sampling of accumulated rainfall patterns and moving convective systems. Organized lines (circles) and individual cell (squares) results taken from Table 5. Solid line is for user reference.

than 25° from those sensor orientations that would permit optimal sampling of the accumulated rainfall patterns. The two cases that differed by more than 25° were organized as lines (Figs. 6-7 and 15-16) whose rainfall swath was dominated by the motion direction of an embedded cell. The 18 August 75 squall line results conform but are misleading

(Figs. 10-11). A closer study indicates an intense squall line (rainfall rates exceeded 110 mm hr^{-1}) that progressed moderately northeastward. System intensity and northeast movement-speed resulted in an accumulated rainfall pattern dominated by relative precipitation maxima and oriented along successive line positions (Figs. 10-11). Radar cloud base sampling at five-minute intervals in this case was not sufficient to detect the true rainfall swath. One quick solution to the problem created by insufficient temporal sampling would be to incorporate (optimally) information above cloud base in estimating the accumulated rainfall pattern. But sampling implications, for the 18 August case only, are judged unrealistic and are dismissed.

The results were anticipated for observing individual cells (Figs. 13-16). Frequently, organized lines sweeping across the countryside deposit precipitation maxima parallel to system motion direction (reflecting embedded cell movement); so, appropriately, such features are detected by the covariance model (see Figs. 6-7, 15-16, 20, and 23). Chapter VIII results provide significant information relating a sensor set's deployment to a phenomenon's preferred storm-track.

Clearly, the results justify procurement of an adequate climatology deduced from space-time covariances and so necessary for <u>any</u> optimal experimental design. In particular, the results point to the need for occurrence

frequencies of (1) system motion-directions and speeds, and to a lesser extent, (2) squall line/convective cell relative frequency and (3) squall line orientation. Additional climatological requirements are mentioned in Chapter VIII.

Radar information has been used to determine characteristic features of convective systems on the High Plains. For the present, the results suggest that the average moving storm in Southeast Montana (Table 3) is almost spatially symmetric (x-y-z), time coherent for almost 20 minutes, and has salient features covering 125 km^2. These systems would have been sampled and analyzed more optimally if sensors had been oriented west northwest-east southeast (289-109°; see Chapter VIII for more conclusive results), and contained, in their raw data sets, a third more signal than noise (SNR = $A_7/(1 - A_7)$).

Similarly, the average accumulated rainfall pattern from Southeast Montana (Table 4) is symmetric, and has salient features covering 80 km^2. These rainfall footprints would have been sampled and analyzed more optimally if stations had been oriented west northwest-east southeast (287-107°; also see Chapter VIII) and contained, in their raw data sets, almost twice as much signal as noise.

The fact that moving convective systems "apparently" are more spatially coherent than their accumulated rainfall patterns (125 km^2 vs 80 km^2) was a surprising result. For most data sets (Figs. 8-9 provide a notable acception), the

implications are misleading and result primarily from relative SNR differences between the two type data sets. The spatial influence radii in any system are restricted somewhat by engineering aspects in achieving the least squares functional fit to a raw covariance matrix set. The NLP algorithm has more freedom to adjust values of the spatial influence radii when a data set's signal is low than when the data set's signal is high. For example, in data sets containing marginal signal qualities, the engineered spatial influence radii asympotically may approach the correct value and still inflict only minor least square errors on the function's determination. For systems of similar size but where signal qualities are increased, minor adjustments by the NLP algorithm in its search for the spatial influence radii will produce increased errors between the raw covariance matrices and the NLP-determined function values. Data sets where signal qualities are high produce an added benefit of high confidence in the calculated spatial influence radii. Thus, there is higher confidence in the characteristic dimensions of accumulated rainfall patterns than in the physical dimensions of a moving system--simply because accumulated rainfall patterns generally contain more signal than is observed in its moving-system observation set. It appears then, that the covariance model over estimates spatial coherency in moving convective systems.

In contrast, Central Oklahoma convective systems

present a striking difference to those selected from Southeast Montana. Brady's (1976, 1977) early morning storm on 21 May 74 was noticeably anisotropic, time coherent for near two-thirds an hour, and had salient features covering almost 3000 km^2 (25 times as large as the average moving Montana system). The Oklahoma system contained equal amounts of signal and noise and was sampled and analyzed optimally if the sensors were oriented southwest-northeast (245-065^o). Similarly, the corresponding accumulated rainfall pattern was anisotropic, had salient features covering almost 2100 km^2 (also 25 times as large as a Montana rainfall field), and was sampled and analyzed optimally for sensors oriented west southwest-east northeast (250-070^o). Also, the pattern contained four times as much signal as noise.

The covariance model has the inherent ability to discriminate objectively between parameter characteristics which are controlled, perhaps, by geographical location, season or modification experiments. This particular ability of the covariance model is a quality that seems well worth evaluating in future experimental work.

The majority of 1975-early 1976 convective systems judged suitable for this study did not pass over the recording gage network. Only three rain gage/radar data sets were secured for study (Table 1). Consequently, any development of a cross-covariance climatology and expected gage network signal-to-noise ratio was not practical. Only average

parameter values are presented (Table 6); the reader may wish to compare these limited Southeast Montana results with Brady's (1976) similarly limited Oklahoma gage and cross-covariance results. All three storm data sets enter into the average accumulated rainfall cross-covariances; the moving storm results are solely from the 7 July 75 squall line case. Interpretation of the results follow in Section C.

	EAST	NORTH	WEST	SOUTH	CORR	ELLIPSE	TIME	GAGE A_7
MOVING SYSTEM	4.4	8.4	5.0	6.0	0.40	-0.80	31.0	0.50
ACCUMULATED RAINFALL	6.6	7.7	7.5	9.1	0.55	-0.25	0.0	0.50

Table 6. The average Southeast Montana cross-covariance and expected gage network lag-zero covariance value for moving storm systems and accumulated rainfall patterns. Influence radii are in km and minutes.

Additional comment on cross-covariance importance is noted in Chapter VIII.

A final point worth discussing concerns the analysis model's ability to recover a data set's signal. User confidence in model filtering properties is evaluated objectively by system ability to filter data set noise (geophysical and otherwise) and still recover the basic observed signal. Model performance is evaluated by comparing parameter values for the covariance function obtained from the original data and from grid-point estimates of the analyzed data set

(using the methodology of Chapter V.B). Parameter values of the filtered covariance function and for the moving systems (Table 3) and for the accumulated rainfall patterns (Table 4) are summarized in Tables 7 and 8.

FILTERED MILES CITY MONTANA RAINFALL X-Y-Z-T COVARIANCE STATISTICS

DATE	EAST	NORTH	WEST	SOUTH	UP	DOWN	CONN	ELIPSE	TIME	AREA	ROTATE	SNR
5 JULY 75 1930-2000 GMT	7.3	6.0	5.1	6.1	0.0	0.0	0.75	-0.37	57.6	117.6	-32.2	3.00
7 JULY 75 1307-1342 GMT	16.3	4.2	10.8	9.3	0.0	0.0	0.71	0.64	21.1	286.9	23.8	2.45
7 JULY 75 1900-2000 GMT	7.5	6.3	8.1	10.1	0.0	0.0	0.69	-0.47	21.1	199.8	-57.7	2.23
17 JULY 75 0007-0027 GMT	7.8	14.3	4.9	6.9	0.0	0.0	0.58	-0.53	10.8	211.1	-70.1	1.38
17 JULY 75 1545-1615 GMT	11.0	12.2	8.7	11.4	0.0	0.0	0.59	-0.28	18.4	366.9	-55.0	1.44
21 JULY 75 2252-2322 GMT	16.8	5.7	9.6	6.2	0.0	0.0	0.61	0.49	25.0	247.2	11.4	1.56
30 JULY 75 2003-2018 GMT	11.2	22.3	7.5	14.6	0.0	0.0	0.66	0.02	29.8	541.1	89.3	1.94
31 JULY 75 0658-0713 GMT	17.4	11.5	20.0	15.4	0.0	0.0	0.79	-0.97	4.4	790.1	-37.4	3.76
18 AUGUST 75 2213-2238 GMT	14.1	27.7	11.1	22.4	0.0	0.0	0.74	-0.98	6.8	992.8	-63.3	2.85
19 AUGUST 75 2347-0012 GMT	9.9	13.1	5.8	9.8	0.0	0.0	0.66	-0.59	10.7	283.1	-57.4	1.94
6 JUNE 76 0530-0600 GMT	10.1	18.1	8.6	10.3	0.0	0.0	0.74	-0.57	42.3	417.2	-68.5	2.85
7 JUNE 76 0430-0445 GMT	9.5	14.3	8.9	9.5	0.0	0.0	0.61	-0.41	25.8	343.4	-68.1	1.56
7 JUNE 76 0515-0545 GMT	19.2	10.9	18.0	10.0	0.0	0.0	0.79	-0.32	13.0	610.6	-14.3	3.76
PARAMETER AVERAGES	12.2	12.8	9.8	10.9	0.0	0.0	0.69	-0.34	22.1	416.0	-30.7	2.36
STANDARD DEVIATIONS	3.9	6.6	4.4	4.3	0.0	0.0	0.07	0.46	14.3	244.1	45.5	0.80

Table 7. Filtered covariance parameter values obtained from the methods in Chapter V.B and the Table 3 convective systems. Variables listed correspond to Table 3.

A comparison of original and filtered parameter values of the covariance model provides strong evidence of model ability to abstract observed signal from significantly noisy data sets. The filtered signal of the average moving-storm (Table 7) closely resembles characteristics of the original signal. Salient features are slightly larger than original data set features because (1) filtered small-scale fluctuations corrupted the predominant system signals and (2) because grid-point data resolution was about half the original data resolution. Optimal sampling station orientation was unchanged. Albeit, the analysis did recover almost twice as much signal (compared with analyzed noise) as was apparent in the original data.

Conversely, almost 50% of the noise contributions to the original data were discarded by the analysis. Model performance was exceedingly high for most data sets (e.g., the SNR increased by 561% and 366% after analysis of the squall line data sets that occurred on 31 July 75 and 7 July 75). Poor data set quality likely restricted model performance on 17 July 75 (1545 GMT).

Similar results were found (Tables 4 and 8) in the analyzed accumulated rainfall patterns. Clearly, these filtering results convincingly point to the high ability of optimum interpolation methods in retaining essential data-set features while tossing aside nearly 50% of the undesired data set properties.

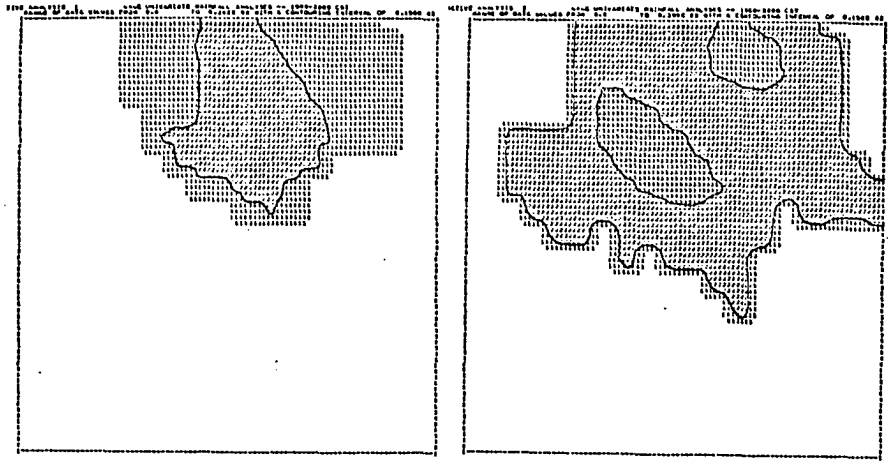
FILTERED MILES CITY MONTANA STORM-TOTAL RAINFALL COVARIANCE STATISTICS

DATE	EAST	NORTH WEST	SOUTH	UP	DOWN	CORR	ELIPSE	TIME	AREA	ROTATE	SNR
5 JULY 75 2000 GMT	6.2	6.2	6.2	6.2	0.0	0.0	0.83	-0.30	0.0	120.6	-45.2	4.88
7 JULY 75 1342 GMT	8.7	7.6	8.7	7.6	0.0	0.0	0.75	0.13	0.0	207.9	23.1	3.00
7 JULY 75 2000 GMT	4.0	8.6	4.0	8.6	0.0	0.0	0.82	-0.51	0.0	107.4	-74.5	4.56
17 JULY 75 0027 GMT	4.9	9.3	4.9	9.3	0.0	0.0	0.73	0.53	0.0	142.6	71.2	2.70
17 JULY 75 1615 GMT	7.6	7.7	7.6	7.7	0.0	0.0	0.53	-0.38	0.0	183.9	-45.6	1.13
21 JULY 75 2322 GMT	14.6	5.0	14.6	5.0	0.0	0.0	0.80	0.23	0.0	229.3	5.1	4.00
30 JULY 75 2018 GMT	8.3	20.1	8.3	20.1	0.0	0.0	0.69	0.33	0.0	522.4	81.1	1.50
31 JULY 75 0713 GMT	5.1	2.8	5.1	2.8	0.0	0.0	0.73	-0.69	0.0	43.9	-23.4	2.70
18 AUGUST 75 2238 GMT	6.3	7.8	6.3	7.8	0.0	0.0	0.59	-0.29	0.0	155.6	-62.9	1.44
19 AUGUST 75 0012 GMT	13.5	5.6	13.5	5.6	0.0	0.0	0.77	-0.08	0.0	237.5	-2.2	3.35
6 JUNE 76 0600 GMT	8.1	13.7	8.1	13.7	0.0	0.0	0.71	-0.31	0.0	349.8	-75.0	2.45
7 JUNE 76 0445 GMT	8.1	8.7	8.1	8.7	0.0	0.0	0.56	-0.16	0.0	219.3	-56.9	1.94
7 JUNE 76 0545 GMT	8.3	7.2	8.3	7.2	0.0	0.0	0.77	-0.03	0.0	187.1	-6.0	3.35
PARAMETER AVERAGES	8.0	8.5	8.0	8.5	0.0	0.0	0.71	-0.12	0.0	208.2	-16.2	2.85
STANDARD DEVIATIONS	3.0	4.2	3.0	4.2	0.0	0.0	0.09	0.34	0.0	115.4	49.5	1.13

Table 8. Filtered total-storm-rainfall covariance parameter values obtained from the methods in Chapter V.B and the Table 4 convective systems. Variables listed correspond to Table 4.

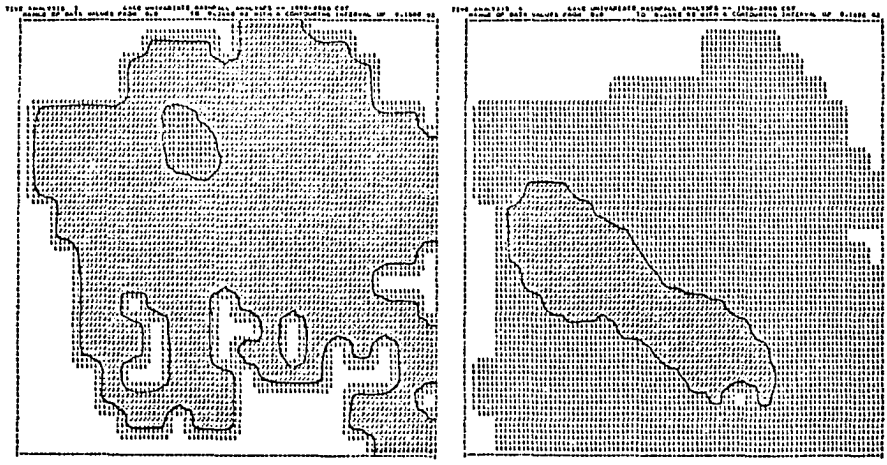
C. Bivariate Objective Analysis Results

On 7 July 75 a squall line moved south and crossed the rain gage network (between 1900-2000 GMT). After 1930 GMT, the convective line's core split into two distinct cells with the strongest portion curving southeast before moving beyond the network. Gage and radar univariate analyses at 15-minute intervals are shown in Figs. 19-20 and the corresponding bivariate analyses in Fig. 21.



19A: 1900 GMT

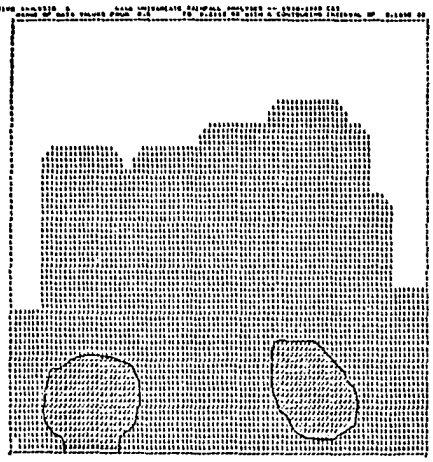
19B: 1915 GMT



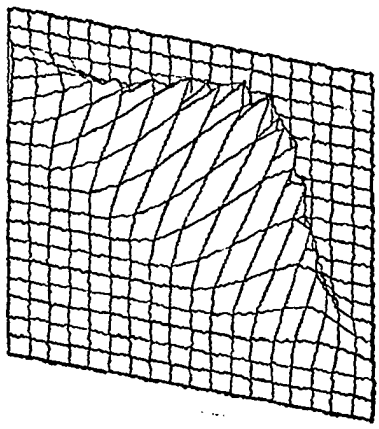
19C: 1930 GMT

19D: 1945 GMT

Fig. 19 (A-F). Rain gage univariate analysis of 7 July 75 squall line moving from 330° at 22 km hr^{-1}. Isohyet contouring interval is 15.0 mm hr^{-1}. Modelled covariance function values (F) range from -0.001 to 0.310. Covariance mesh is 0.5 km square.



19E: 2000 GMT

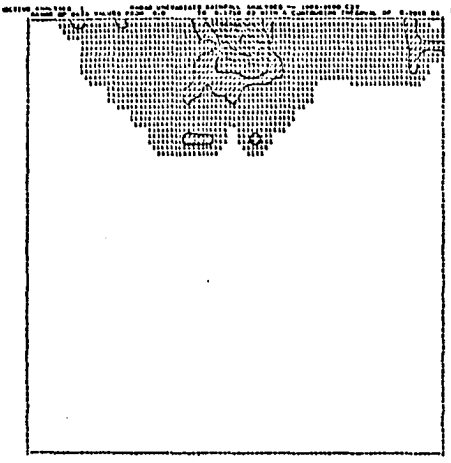


19F: System Covariance

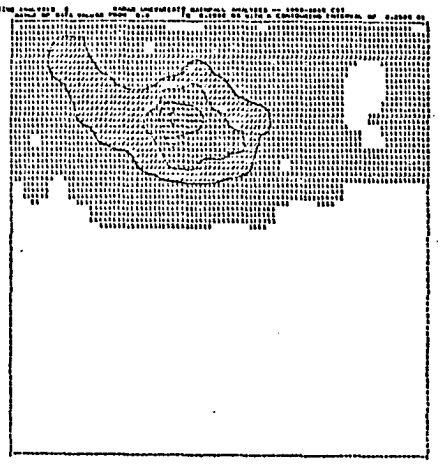
Figures 19-21 are presented basically for reader information; demonstration of bivariate objective analysis superiority is achieved with the accumulated rainfall analyses presented shortly.

All the same, Fig. 19F (the actual calculated gage covariance) and Fig. 20F (the "true" storm structure) illustrate graphically the misleading structures obtained from an inadequately sampled system. Improper signal filtering of the system (not shown) resulted in displaced maxima broadly spread (true gradients relaxed) over the network and inconsistent with the true analyses.

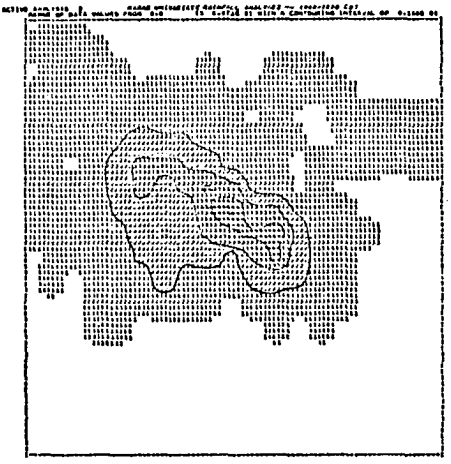
System physics are reflected by the gage/radar cross-covariance function (Fig. 21F and Table 6). Accuracy in surface rainfall estimates from radar observations is possible in this case only when radar predictor information



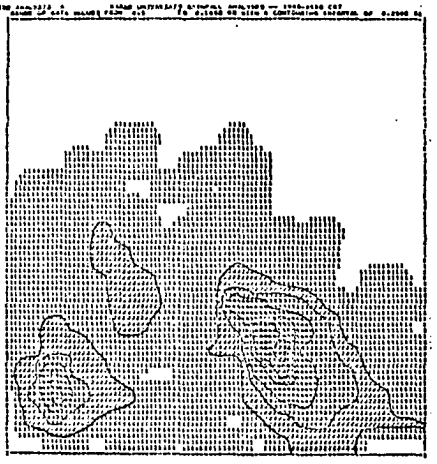
20A: 1900 GMT



20B: 1915 GMT

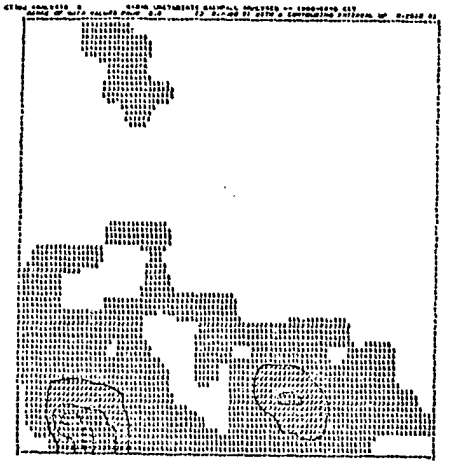


20C: 1930 GMT

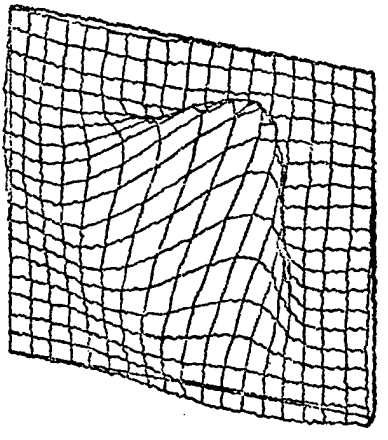


20D: 1945 GMT

Fig. 20 (A-F). Radar univariate analysis of 7 July 75 squall line moving from 330° at 22 km hr^{-1}. Isohyet contouring interval is 2.5 mm hr^{-1}. Modelled covariance function values (F) range from -0.089 to 0.590.



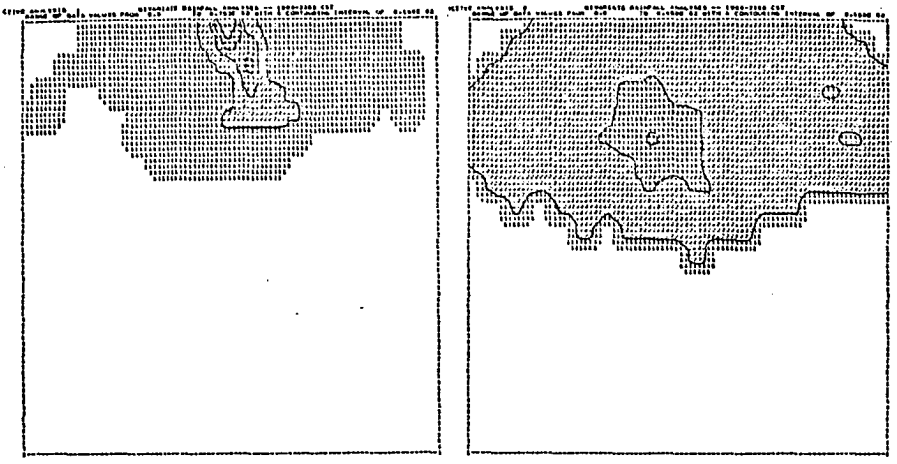
20E: 2000 GMT



20F: System Covariance

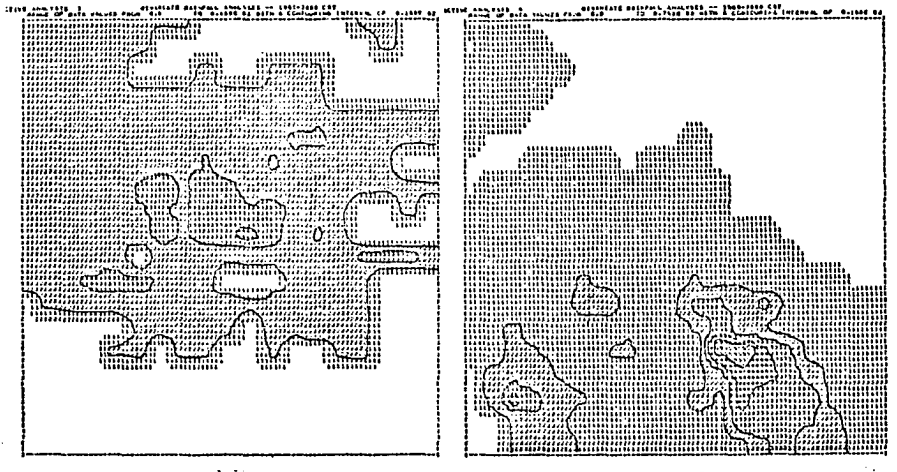
is obtained from locations north northwest of the gage (direction from which storm moved). Predictors chosen from the south would have relaxed rainfall gradients along the storm's leading edge (where typically gradients are strongest) and artificially increased gradients on the storm's backside. Moving east or west for radar predictors would have produced similar undesirable consequences.

From evidence gained during the gage/radar case studies, there appears a strong possibility that radar precipitation estimates were significantly too low. Inclusion of observations as low as 20 dBz could conceivably have accounted for some bias in the rainfall estimates. Covariance results and author experience suggest more explanations are needed. However, the beauty of my analysis model is that none are needed! Data sets that possess different



21A: 1900 GMT

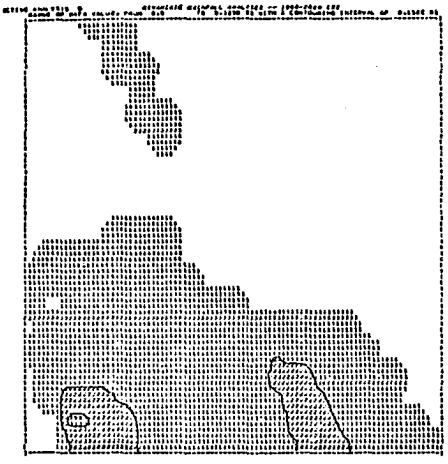
21B: 1915 GMT



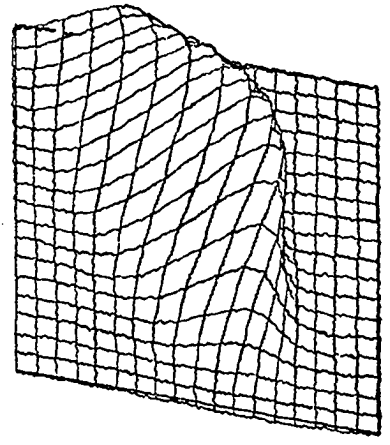
21C: 1930 GMT

21D: 1945 GMT

Fig. 21 (A-F). Rain gage/radar bivariate analysis of 7 July 75 squall line moving from 330° at 22 km hr^{-1}. Isohyet contouring interval is 15.0 mm hr^{-1}. Modelled cross-covariance function values (F) range from -0.017 to 0.400. Cross-covariance mesh is 1.0 km square.



21E: 2000 GMT



21F: System Cross-Covariance

parameter means or variances are considered properly by the methods introduced in Chapter IV.B. But, the reader should understand that accurately calibrated radar information properly converted to rainfall is important to analysis quality. Zawadzki (1975) suggests an approach whereby gage information can be correlated with radar data to enhance the radar's value as a hydrological tool.

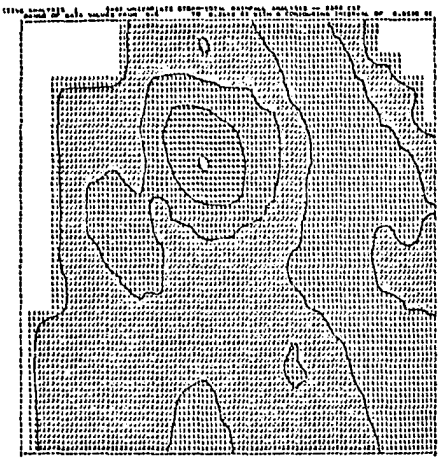
Gage and radar univariate accumulated rainfall analyses (Figs. 22-23) reflect the same general patterns and indicate the movement path of the line's significant cell. (Contrast Figs. 22B and 23B and notice the misleading structure revealed by the gage network.) The gage analysis maximum (20.4 mm) coincides with observations but lacks enough detail given by the univariate analysis of closely spaced radar observations. Maximum radar analyzed rainfall

was 4.67 mm (an amount 1/4 as large as the surface measured precipitation).

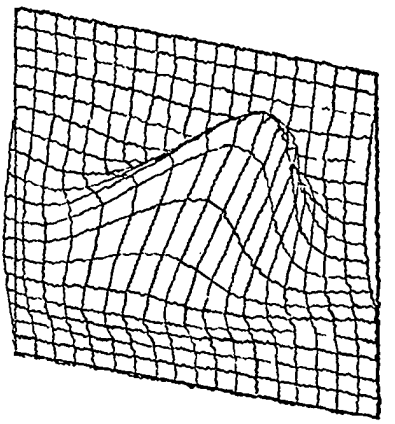
The resulting bivariate analysis (Fig. 24A) maintains essential features of both data sets and those in which the meteorologist has most confidence: the radar indicated rainfall pattern adjusted by the mean observed surface rainfall. Maximum analyzed rainfall of 20.6 mm was colocated with a similar surface maximum (Figs. 22A and 24A). In addition, the bivariate analysis indicated a relative rainfall maximum that apparently was missed by the coarse gage network. The accompanying cross-covariance function (Fig. 24B) suggested optimum radar predictor location directions along storm-motion-path and not the sub-optimal isotropic or across movement-path direction.

Bivariate analysis capability is demonstrated again using a rain gage/radar data set from a slowly moving convective cell on 6 June 76. Gage accumulated rainfall analysis (Fig. 25A) indicates a diffuse rainfall core basically oriented northwest-southeast. Maximum amounts of 19.4 mm occurred in the east central portion of the analysis grid. Radar observed patterns (Fig. 26A), basically similar, suggested maximum surface amounts should have occurred northwest of the primary gage indicated maxima.

The bivariate analysis again maintained essential data set features (Fig. 27A) and produced 29.3 mm at the location of maximum precipitation in the radar analyzed pattern.

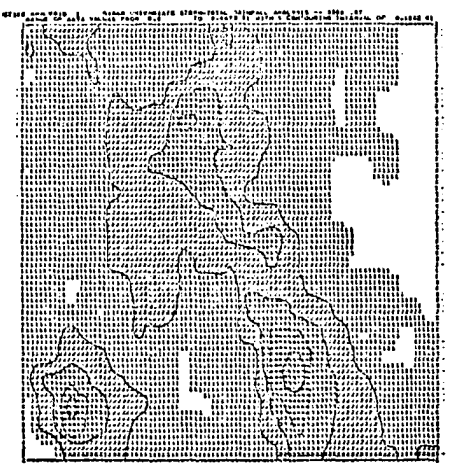


22A: 2000 GMT

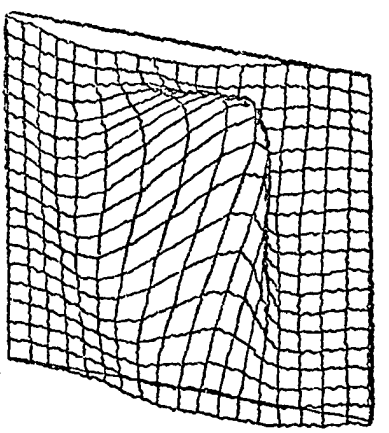


22B: System Covariance

Fig. 22 (A-B). Rain gage univariate total-storm rainfall analysis of 7 July 75 squall line. Isohyet contouring interval is 5.0 mm. Modelled covariance function values (B) range from -0.104 to 0.570. Covariance mesh is 0.5 km square.

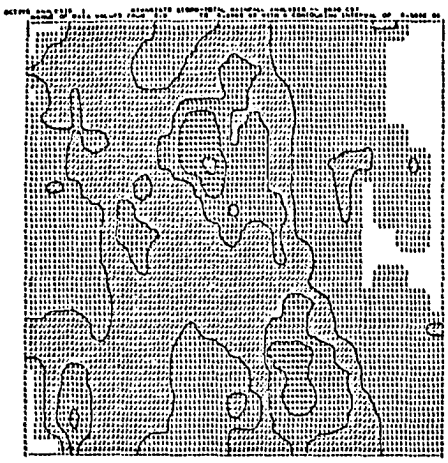


23A: 2000 GMT

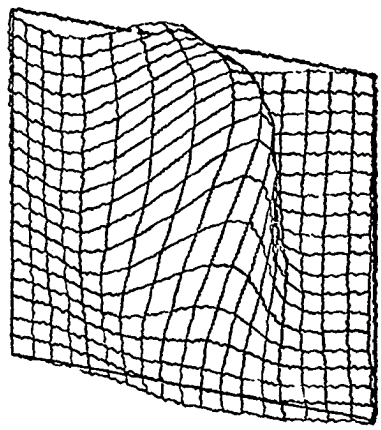


23B: System Covariance

Fig. 23 (A-B). Radar univariate total-storm-rainfall analysis of 7 July 75 squall line. Isohyet contouring interval is 1.0 mm. Modelled covariance function values (B) range from -0.125 to 0.650.



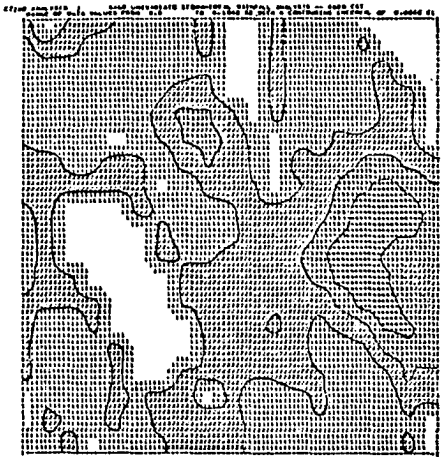
24A: 2000 GMT



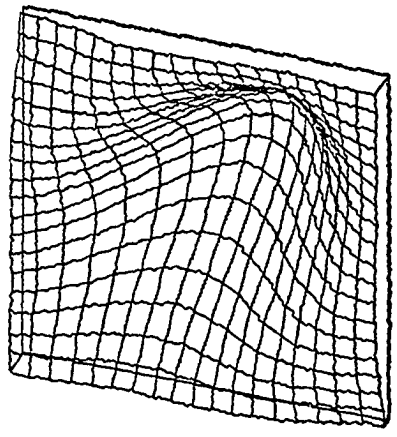
24B: System Cross-Covariance

Fig. 24 (A-B). Rain gage/radar bivariate total-storm rainfall analysis of 7 July 75 squall line. Isohyet contouring interval is 5.0 mm. Modelled cross-covariance function values (B) range from -0.078 to 0.450. Cross-covariance mesh is 1.0 km square.

The rainfall swath trailed southeast to a secondary maximum near 15 mm and colocated with the gage analyzed maximum. Differences between the two univariate analyses and the bivariate results appear significant; in fact, the results agree with observed data when the model's filtering properties are considered. The bivariate maximum of 29.3 mm is consistent with gage observations in the area. The coarse network gage analysis did not reflect the larger rainfall volume because gage-location/grid-point separation distance was large enough to dilute the linearly interpolated observations. Further, the analysis methods have been shown to be fully capable of predicting values larger (or smaller) than observations suggest

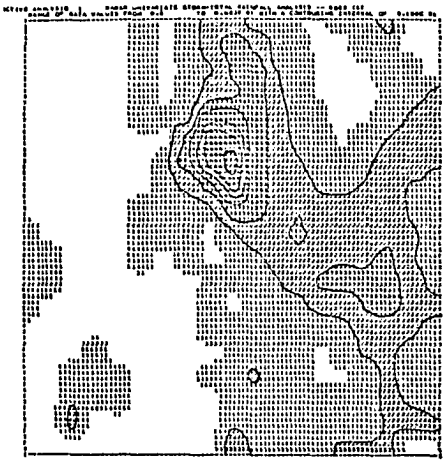


25A: 0600 GMT

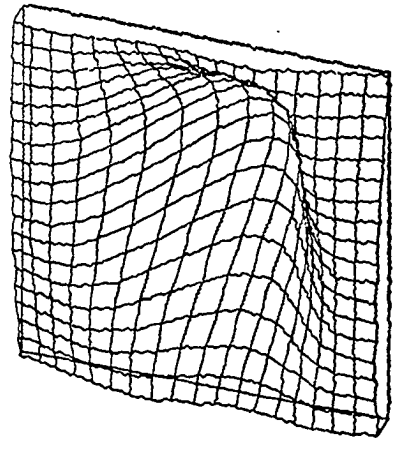


25B: System Covariance

Fig. 25 (A-B). Rain gage univariate total-storm-rainfall analysis of 6 June 76 convective cell. Isohyet contouring interval is 5.0 mm. Modelled covariance function values (B) range from -0.136 to 0.710. Covariance mesh is 0.5 km square. Analysis grid is 36.0 km square.

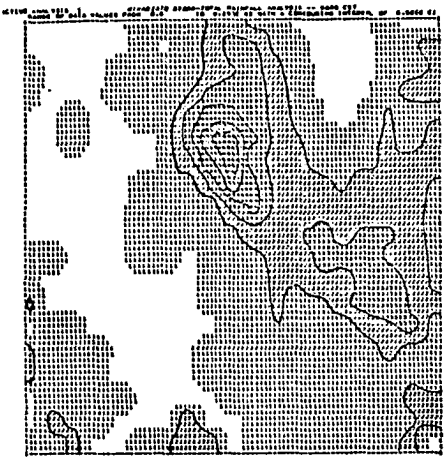


26A: 0600 GMT

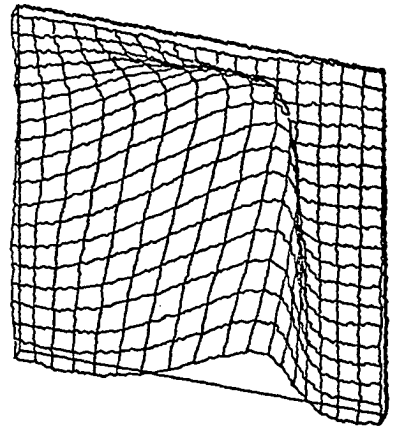


26B: System Covariance

Fig. 26 (A-B). Radar univariate total-storm-rainfall analysis of 6 June 76 convective cell. Isohyet contouring interval is 1.0 mm. Modelled covariance function values (B) range from -0.158 to 0.780. Analysis grid is 36.0 km square.



27A: 0600 GMT



27B: System Cross-Covariance

Fig. 27 (A-B). Rain gage/radar bivariate total-storm-rainfall analysis of 6 June 76 convective cell. Isohyet contouring interval is 5.0 mm. Modelled cross-covariance function values (B) range from -0.065 to 0.680. Cross-covariance mesh is 1.0 km square. Analysis grid is 36.0 km square.

as possible (Eddy, 1967 page 36). Thus, the bivariate analysis benefits from details revealed by the very dense radar observation set.

CHAPTER VII

AN EXPERIMENTAL DESIGN EVALUATION FUNCTION: REVISITED

Rain gage data and radar reflectivity information have been combined previously to produce optimal surface rainfall estimates. Now, deployment of a statistically adequate rain gage set is investigated. Previous discussions have shown that the task critically depends on the chosen analysis procedure, storm physics, and sensor engineering aspects. In addition, decisions by a project director significantly interact with features critical to sensor deployment and influence a field experiment's design.

The experimental design model (Chapter III, Eq. (15)) evaluates the relative value of a proposed instrument configuration. Sampling quality is measured through modelled estimates of explained signal variance. Brady (1976) demonstrated the practical validity of this approach. The model assumes the value of each sensor is determined only by parameter statistical structure. Primarily, the model examines

the effect of sensor density reduction on the ability of a sampling system to detect signal variation. The model does not account for observational biases nor the relative role played by each parameter in producing the final analysis.

Brady (1976) extended the model's ability to detect signal variation by accounting for the predictand's mean value, which scales the analyzed field. He assumed unbiased observations where the mean values and variances of both parameters are equal. Now, development of the experimental design model is completed when parameter means and variances are modelled. Chapter VIII results report on the trade offs involved in the final use of a sensor set. At this point, the only assumptions are those fundamental to the objective analysis algorithm.

Derivation of the experimental design model is given in Appendix B. The design evaluation function becomes:

$$F = \frac{1}{G} \sum_{i}^{I=G} \eta_i \left[(1-R_i^2) + \frac{(1- \sum_{j=1}^{J} \hat{\beta}_{G_j})^2}{M_G} + \frac{(\sigma_R/\sigma_G \sum_{k=1}^{K} \hat{\beta}_{R_k})^2}{M_R} \right.$$

$$\left. - \frac{2(\sigma_R/\sigma_G)\bar{\rho}}{(M_G M_R)^{\frac{1}{2}}} \left(\sum_{k=1}^{K} \hat{\beta}_{R_k} - \sum_{k=1}^{K} \hat{\beta}_{R_k} \sum_{j=1}^{J} \hat{\beta}_{G_j} \right) \right] \qquad (20)$$

which represents the weighted-mean, residual-variance fraction over the space-time sample field (G points). Each term has the following interpretation:

$$\underline{\text{TERM 1}}: \quad \frac{1}{G} \sum_{i}^{I=G} \eta_i (1 - R_i^2)$$

The first term represents the unexplained variance fraction contributed by perturbations in the phenomenon under analysis. Also, it measures the sensor reduction effect on the ability of a sampling system to detect signal variation.

$$\underline{\text{TERM 2}}: \quad \frac{1}{G} \sum_{i}^{I=G} \eta_i \left[\frac{(1 - \sum_{j=1}^{J} \hat{\beta}_{G_j})^2}{M_G} \right] \quad \text{where J is a subset of } M_G.$$

The second term represents the unexplained variance fraction contributed by uncertainty in the mean value of gage observations. Clearly, a multivariate designed experiment is impossible without minimal predictand-related sensors (M_G); otherwise, the term is infinite and the final analysis cannot be scaled. This implies, for the particular case of radar/rain gage sensor deployments, that gage observations are essential to relate radar information to surface rainfall estimates. Our interpretation of M_G is the effective number of gages within the statistically coherent area of a convective storm. Enough sensors must be within this area for an analysis of surface rainfall adequate to meet the requirements of a principal investigator. If no gages are chosen for the

perturbation analysis (i.e., $\sum_{j=1}^{J} \hat{\beta}_{G_j} = 0$), uncertainty in the analyzed parameter mean value still exists and is represented by $1/M_G$.

$$\underline{\text{TERM 3}}: \quad \frac{1}{G} \sum_{i}^{I=G} \eta_i \left[\frac{(\sigma_R/\sigma_G \sum_{k=1}^{K} \hat{\beta}_{R_k})^2}{M_R} \right]$$

This term's contribution to the residual variance fraction has an interpretation similar to Term 2; that σ_R^2/M_R represents uncertainty in the mean value of precipitation estimated from reflectivity information. Numerous radar-derived rainfall observations give high confidence to its mean value field; the residual variance contribution is relatively small. The term is zero if no reflectivity predictors are chosen. Also notice that a relative perturbation power ratio (σ_R^2/σ_G^2), determined by the radar and gage data, is required if radar observations are to be used. Relatively high variances among the gage observations increases the radar's importance to system sampling ability.

$$\underline{\text{TERM 4}}: \quad \frac{1}{G} \sum_{i}^{I=G} \eta_i \left[\frac{-2(\sigma_R/\sigma_G)\bar{\rho}}{(M_R M_G)^{\frac{1}{2}}} \left(\sum_{k=1}^{K} \hat{\beta}_{R_k} - \sum_{k=1}^{K} \hat{\beta}_{R_k} \sum_{j=1}^{J} \hat{\beta}_{G_j} \right) \right]$$

The final term represents residual variance contributions from uncertainty in the observation's mean value. The only new variable is $\bar{\rho}$--the expected correlation between modelling errors in estimating the mean value of rain gage and radar derived rainfall.

The important control exercised by the objective function on the placement algorithm exists because it considers: (1) the uncertainty in the parameter mean values (required for proper scaling of a bivariate analysis); and (2) the uncertainty in the perturbation analysis (measured by the residual variance).

The following general result is inferred. For a given gage covariance function, a low lag-zero cross-correlation value results in an objective function dominated by the first two terms. A high lag-zero cross-correlation value lowers the first term's contribution (radar bin density is much greater than gage network density), raises the second term's contribution (because gages are not chosen for the perturbation analysis) and brings the third and fourth terms into play. These features are shown in Chapter VIII.

At this point, recall that the sample residual variance is a biased estimate of the population unexplained variance. The sample residual variance $\{\hat{\sigma}^2 = \frac{1}{n} E (e^t e)\}$ easily is shown equal to $\hat{\sigma}^2 = \frac{1}{n-m} \sigma^2$ by extending the Eq. (8) results with the methods of Appendix B. The $1/(n-m)$ term represents the effective degrees of freedom in an analysis. The term's effect insures that sample residual variances are increased to more accurately represent the population value. Because sensors are deployed in a climatological sense, occasionally, the modelled residual variances are larger than the observation error variances.

Since storms may occur throughout the study area, the sensor deployment must consider the climate of the region; otherwise, experiment requirements may not be satisfied. Because an a priori climatology of preferred storm track or genesis information was not available, a weighted residual variance fraction was not obtained (i.e., η_i of Eq. (20) was kept uniform across the analysis grid).

Also, storm size variety implies that over sampling of one storm is under sampling in another. Thus, the gage density and orientation requirements must be adequate to define a sufficient storm subset. These requirements consider "average" accumulated rainfall patterns (Chapter VI) of prime interest to HIPLEX. Clearly, the storm climatology is too sparse at the present to produce stable results (Chapter VIII).

The actual sensor deployment of this study proceeds differently from Brady's methods (1976). He placed sensors by using latitude/longitude coordinate positions as the NLP decision variables. Efficiency in determining the optimal deployment of a sensor set relates closely to the number of decision variables involved. When each station is free to find its optimal location, placement of large sensor sets demands excessive computer resources.

Because study objectives are different than were Brady's, the current method uses only two NLP decision variables to place sensors. Initial deployment proceeds as follows: the first sensor is placed at the center of the

analysis grid. Then, the desired additional sensors are added in the symmetrical fashion suggested (Fig. 28). Sensor separation-distances (east-west {Δx} and north-south {Δy}) relate directly to the influence radii of the predictand's covariance function. Finally, the sensors are rotated to lie along the preferred storm-track (Eq. (19)). The algorithm's two decision variables become the sensor-separation distances parallel and normal to a preferred storm track. Now, sensor set size does not impact (theoretically) on placement efficiency (and hence computer resources).

Deployment of multivariate sensor sets, fixed and movable, are possible. Extensive tests of the placement methods are presented in Chapter VIII.

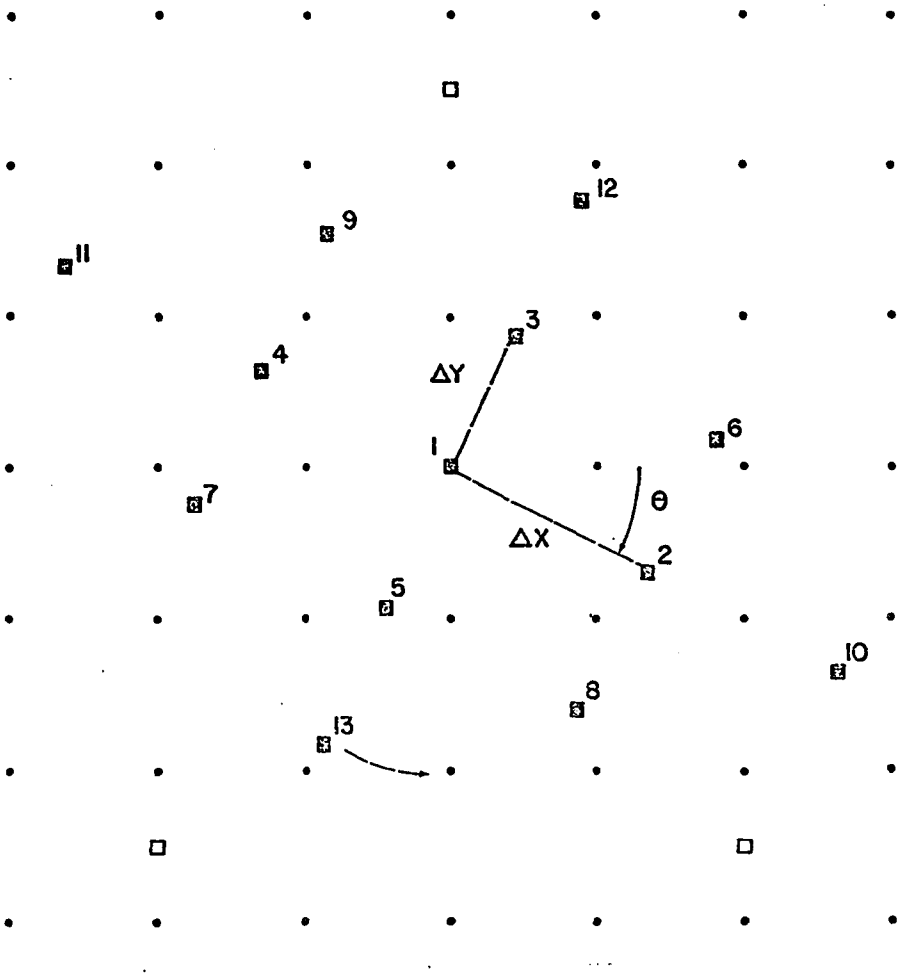


Fig. 28. The method used to provide the experimental design algorithm a movable sensor's relative location (closed squares) in the presence of an analysis grid (circles) and a permanently located (if any) sensor set (open squares).

CHAPTER VIII

EXPERIMENTAL DESIGN RESULTS

The objectives of this chapter seek to evaluate the HIPLEX data requirements in observing accumulated rainfall patterns. Because HIPLEX is concerned with rainfall enhancement, evaluation of modification experiments involves optimal sampling and analysis. Previous discussions suggested that these methods are interdependent; both techniques are vital for HIPLEX to achieve its objectives.

Satisfactory discrimination between storm analyses, obtained from systems whose population rainfall amounts differ, rests on the accurate analysis of an adequate data set. Logically, obtaining a statistically adequate data set precedes its optimal analysis; in reality, this rarely occurs. Thus, selected data sets of interest, and representative of the study area climatology, are used to derive objective criteria that a principal investigator could use to determine and to deploy his "best" sensor set (quality, quantity, type

and location). Consequently, the experimental design results relate to the covariance study and objective analysis results presented earlier. Now, design of the HIPLEX sampling program is evaluated by using an "average" accumulated rainfall pattern (Table 4) and the experimental design model (Chapter VII).

The illustrative gage pattern possesses symmetric structural characteristics (i.e., $A_1 = A_2 = A_3 = A_4$) and is only slightly anisotropic ($A_8 = -0.10$). Since "filtered" covariance functions accurately represent the population structure of a system (Chapter V.B), features of the accumulated rainfall pattern were enlarged to evaluate sampling requirements of systems statistically coherent over 150 km^2 (i.e., influence radii of 6.9 km). Thus, the illustrative storm "sampled" in this chapter possesses the Table 9 features (assembled subjectively from Tables 4 and 6). Other attempts at specifying gage network requirements to observe spatial variations in precipitation also are based on parameter characteristics (Hendrick and Comer, 1970).

A natural question to arise concerns analysis grid characteristics, since design results depend on sensor locations relative to the grid. Its size must be sufficient to contain the storm. In reality, the grid must be adequate to "catch" enough storms to satisfy experiment requirements. Storm characteristics are considered homogeneous over the Miles City study area; thus, sensor density and deployment

	EAST	NORTH	WEST	SOUTH	CORR	ELLIPSE	TIME
GAGE COV.	6.9	6.9	6.9	6.9	0.50	-0.10	0.0
RADAR COV.	6.9	6.9	6.9	6.9	VRBL	-0.10	0.0
GAGE/RADAR CROSS-COV.	7.0	8.5	7.0	8.5	VRBL	-0.25	0.0
RADAR/GAGE CROSS-COV.	7.0	8.5	7.0	8.5	VRBL	-0.25	0.0

Table 9: Characteristics of an "average" accumulated rainfall pattern from Southeast Montana. Values of the radar covariance and cross-covariance at lag-zero are controlled subjectively to assess climatology's impact on sampling requirements.

over a minimum size grid represents deployment requirements that extend to a larger grid. The only limitations are economics and the relevant climatology.

An illustrative "round storm" is inscribed inside a square grid containing $\sim$ 190 km^2 (13.8 km x 13.8 km). As the mesh size of a grid approaches the average station spacing, interpolation analysis accuracy becomes more stable. Analyses on grids, that have a mesh size smaller than the average station spacing, produce little new information and serve only to reduce computational efficiency. The results presented use a 7 x 7 x 2.3 km square grid.

A. Radar Observation Influence on Optimal Gage Deployment

A fundamental question faced by HIPLEX investigators concerns gage density requirements and the subsequent deployment sufficient to meet and evaluate stated HIPLEX objectives. A decision on the basic optimal gage density is expressed as the gage quantity (M_G) required within the storm. In part, the decision is governed by parameter SNR characteristics; to a larger degree, the decision is influenced by the value, at lag-zero, of the cross-correlation between radar and gage data. The results represent the objective criteria a project director should use to obtain an efficient match of available resources and experiment requirements.

Sensor deployment in a highly constrained environment is complex at best. The design function invariably contains several local optima in a solution space dependent on the sensor quantity, grid characteristics, and multivariate parameter structure. However, our NLP algorithm guarantees that any sensor deployment will find at least a local optimum (Himmelblau, 1972). Indeed, the local optimum problem was encountered during the sensor placement investigation. The model's occasional sensitivity to different sensor starting vectors became a source of irritation. With experience, internally consistent results were obtained; they are considered valid and accurate.

Now, gage density requirements are evaluated for the

Table 9 system. Radar SNR values and the strength of gage/radar interrelationships are controlled subjectively, along with movable gage and fixed radar bin quantity, to illustrate the consequent impact on sampling requirements. Typical results obtained are presented when 391 and 63 fixed radar bins are spaced uniformly throughout the analysis grid (Figs. 29-30). For reader convenience, univariate gage deployment results are repeated in each figure. When the radar covariance equalled 0.80, the bivariate design requirements matched the gage univariate results; these results are not shown in Fig. 29A and 30A. Values of each objective function term are provided (Tables 10-11) when approximately 5 and 10 gages are within the statistically coherent accumulated rainfall pattern; the results derive from Figs. 29-30.

Gage observations from Miles City were found to contain four times more variance than observed in corresponding radar data sets (Eddy and McDonald (1977)); thus, the ratio σ_R/σ_G was set to 0.25. The correlation coefficient $\bar{\rho}$ was set to 0.4, a value that reflects only marginal correlation in modelling errors of the two mean values. The following points are noted for designers of multivariate mesoscale field experiments:

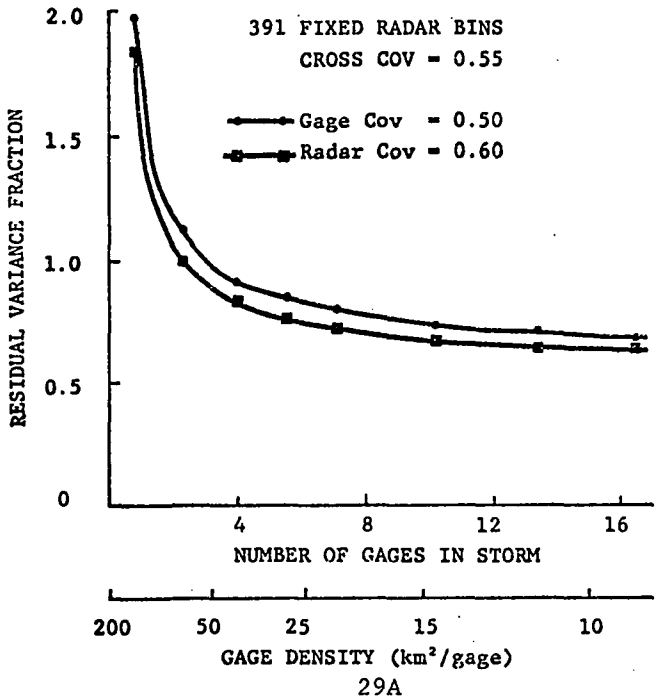
In general, elliptically shaped storms require widely separated gages oriented along the storms' preferred track and spaced more closely along the minor axis of the covariance ellipse. As sensors are added to a network, the placement

algorithm permits more closely spaced sensors. The algorithm actually "forces" sensors to remain in the analysis grid. However, a point of diminishing returns is soon reached. Now, sensors added no longer provide independent sampling information. For a system possessing Table 9 characteristics, a maximum of 10 gage sensors is ample to explain and sample the significant signal variance. Additional sensors, placed initially outside the analysis grid, no longer are relocated inside, simply because the gages are not needed to explain the signal variance.

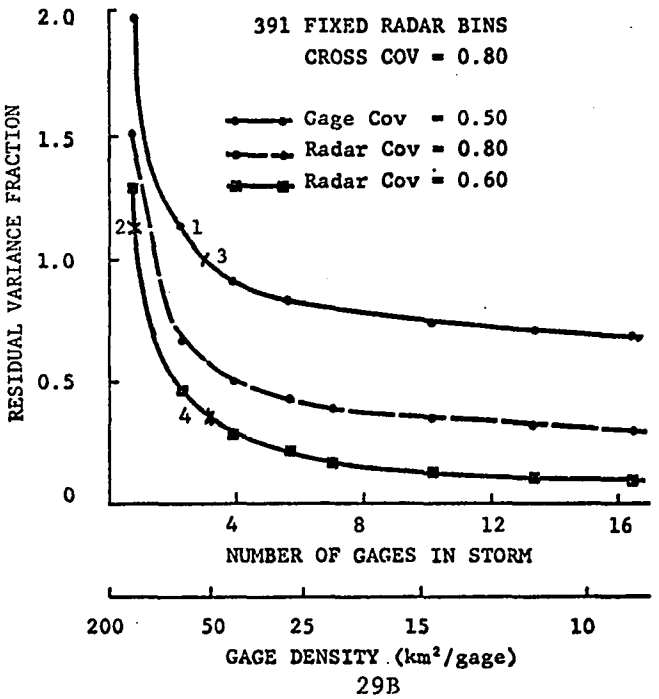
Under the conditions which all study results are derived, sampling quality is not diminished by minor location adjustments to an optimally deployed sensor set. Should trees, hills, and obstinate land owners cause problems, a project director could, in good faith, relocate selected sensors by several hundreds of meters and not compromise the design effort.

Reader attention is now called to Figs. 29-30. A first point to notice is that no matter how many radar bins are available for analysis, unless a certain minimum gage quantity is present to stabilize the mean value estimates, these radar data are simply of no use in accurately estimating surface rainfall.

A second point is made when one considers the cross-covariance between the two different data sets to be significantly better than the gage autocovariance. In such

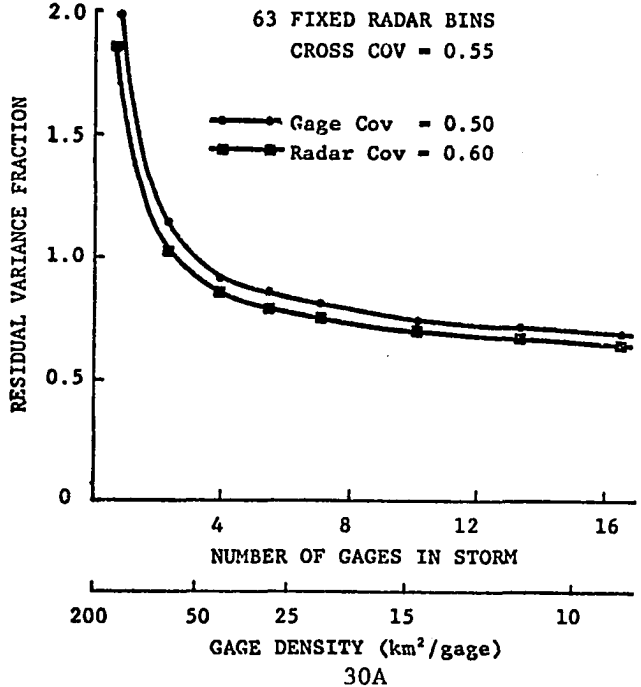


29A

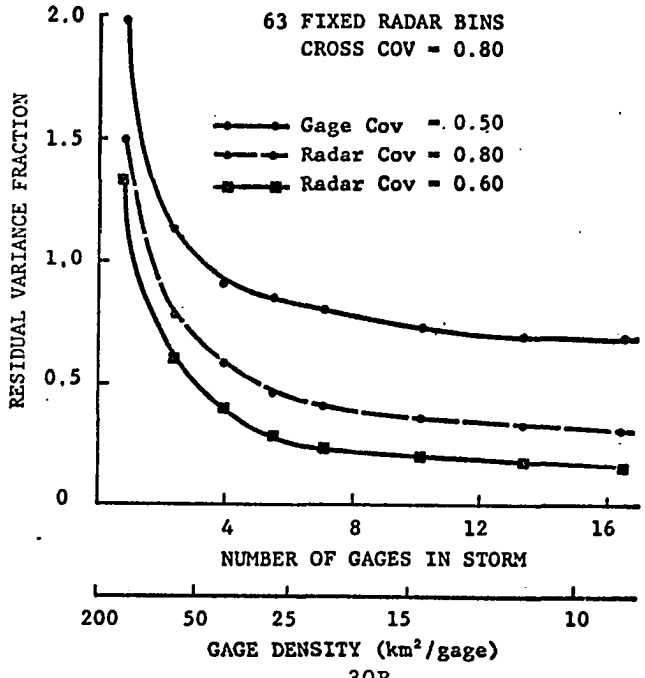


29B

Fig. 29 (A-B). An evaluation of HIPLEX sampling requirements for observing accumulated rainfall patterns. Sensor set quantity and quality varied to assess the impact on sampling requirements (391 radar bins spaced uniformly throughout the analysis grid).



30A



30B

Fig. 30 (A-B). An evaluation of HIPLEX sampling requirements for observing accumulated rainfall patterns. Sensor set quantity and quality varied to assess the impact on sampling requirements (63 radar bins spaced uniformly throughout the analysis grid).

GAGES IN STORM (M_G)	GAGE COV	RADAR COV	CROSS COV	PERTURBATIONS		MEAN		TOTAL
				TERM 1	TERM 2	TERM 3	TERM 4	OF
5.51	0.50	0	0	0.815	0.042	0	0	0.857
5.51	0.50	0.80	0.55	0.667	0.182	0.00	-0.003	0.846
5.51	0.50	0.60	0.55	0.594	0.182	0.00	-0.004	0.772
5.51	0.50	0.80	0.80	0.257	0.182	0.00	-0.005	0.434
5.51	0.50	0.60	0.80	0.042	0.182	0.00	-0.006	0.218
10.24	0.50	0	0	0.726	0.009	0	0	0.735
10.24	0.50	0.80	0.55	0.667	0.094	0.00	-0.002	0.759
10.24	0.50	0.60	0.55	0.594	0.098	0.00	-0.003	0.689
10.24	0.50	0.80	0.80	0.257	0.098	0.00	-0.003	0.352
10.24	0.50	0.60	0.80	0.042	0.098	0.00	-0.004	0.136

Table 10. Contributions of each term to selected objective function values plotted in Fig. 29.

GAGES IN STORM (M_G)	GAGE COV	RADAR COV	CROSS COV	PERTURBATIONS		MEAN		TOTAL
				TERM 1	TERM 2	TERM 3	TERM 4	OF
5.51	0.50	0	0	0.815	0.042	0	0	0.857
5.51	0.50	0.80	0.55	0.680	0.157	0.000	-0.006	0.831
5.51	0.50	0.60	0.55	0.624	0.180	0.001	-0.009	0.796
5.51	0.50	0.80	0.80	0.244	0.229	0.002	-0.016	0.459
5.51	0.50	0.60	0.80	0.090	0.225	0.002	-0.018	0.299
10.24	0.50	0	0	0.726	0.009	0	0	0.735
10.24	0.50	0.80	0.55	0.668	0.071	0.000	-0.004	0.735
10.24	0.50	0.60	0.55	0.620	0.092	0.001	-0.007	0.706
10.24	0.50	0.80	0.80	0.243	0.135	0.002	-0.012	0.368
10.24	0.50	0.60	0.80	0.080	0.150	0.003	-0.016	0.217

Table 11. Contributions of each term to selected objective function values plotted in Fig. 30.

cases, a simple replacement of gages as predictors by radar reflectivities as predictors does not improve significantly the sampling accuracy. Increased confidence in the radar field's mean value (terms 3 and 4 of Eq. (20)) is far outweighed by decreased confidence in the gage field's mean value (term 2 of Eq. (20)). This second point results from an abundance of radar observations (M_R), from a ratio σ_R/σ_G that is less than unity, and to a lesser extent, from a correlation ($\bar{\rho}$) between $\bar{X}_G$ and $\bar{X}_R$ that also is less than unity.

When the cross-covariance function is equal to or higher than the gage covariance function, an analysis reflects radar data characteristics, primarily because the grid points are surrounded by abundant radar data. Recall that: $\beta_o = \bar{Y} - \Sigma\hat{\beta}_G\bar{X}_G - \Sigma\hat{\beta}_R\bar{X}_R$. If the cross-covariance and gage covariance functions are nearly equal, a replacement of gages by radar data as the predictors may increase sampling pattern accuracy. However, the pattern accuracy increases only at the expense of accuracy (over many analyses) in the mean value of the field (i.e., $\Sigma\hat{\beta}_G = 0$).

Notice a third point. If the cross-covariance is high, the gages serve only to calibrate the radar; the placement algorithm responds by failing to relocate any gage sensor. In this case, an adequate placement solution is a uniform deployment across the analysis grid of at least the minimum gage quantity (point 1 noted previously). However, some storms have low cross-covariance values (dry sub-cloud air, incorrect

Z-R relation, radar pulse volume not full, gusty surface winds). Now, an optimal deployment pays dividends because the gage observations, in such cases, both define the perturbations and stabilize the mean value.

A fourth point occurs as the gage's covariance value at lag zero decreases (plotted results not shown). This implies that gage accuracy has decreased; consequently, deployment of more gages per unit area is needed for calibration purposes. Unless the gage quantity is increased, even good radar data will not improve the analysis quality over one obtained from more accurate gages.

The fifth point is rather subtle. It arises from differences in the bivariate placement results of Fig. 29B (391 bins) and Fig. 30B (63 bins). Network requirements evaluated from highly intracorrelated data (cov = 0.80) is accurate only if predictor redundancy is minimized. Stepwise regression procedures, which eliminate predictor redundancy, were compromised somewhat to maintain a computationally efficient algorithm (details in Section C). The effect of predictor redundancy becomes clear when sampling requirement results are not altered--even though evaluated from only 16% of an available, but highly intracorrelated radar data set (Tables 10-11 might assist the reader at this point). Predictor subsets, obtained from sparse radar data, have more intra-observation independence than an equivalent number of closely spaced radar predictors.

Screening procedures, permitted to perform efficiently, produce similar results, regardless of initial predictor quantity or parameter intrarelationships. Consequently, each analysis point in the bivariate curves of Fig. 29B used highly intracorrelated radar data, whereas in Fig. 30B, the radar data contained more useful information (more degrees of freedom). Statistically significant pattern features are not lost, even though less than 20% of the available radar data may be used. Bergman and Bonner (1976) have encountered high redundancy in their predictor data sets and have used optimum interpolation methods to assess the impact of predictor redundancy on analysis quality. Alaka (1974) shows that, the ability to consider data quality, is a significant advantage in using optimum interpolation schemes.

Tables 10-11 support these results. For example, 5-10 gages, combined with highly intracorrelated but widely spaced radar observations, observe the same basic signal variation (term 1 of Eq. (20)) as observed in a dense sampling environment. Decreased parameter intracorrelations (limited by logic of course) assist the design algorithm in providing favorable decision criteria to a project director.

The second point noted should be considered with this fifth point to reach a desired compromise.

B. Preferred Storm-Track Influence On Sensor Deployment

Does a preferred storm track influence the sampling

quality of a sensor array? Do sensors oriented along such a preferred track produce the best sampling results? Brady's results (1976) imply that the answers are yes--covariance anisotropy is an important consideration if experiment goals are to be achieved.

The current placement methods evaluate a preferred storm track's influence on sampling quality. Pitfalls of subjective network design are illustrated with two different anisotropic covariance functions: the first is characterized by symmetry ($A_1 = A_2 = A_3 = A_4$) and the second by asymmetry ($A_1 = A_3$ and $A_2 = A_4$ but $A_1 \neq A_2$). A sparse (5 gages spaced 38.1 km^2/gage) and a dense (13 gages spaced 14.7 km^2/gage network is deployed optimally to observe systems possessing these characteristics. Relative gage locations are inferred from Fig. 28. The two optimally deployed gage networks, each derived from two different system covariances, were reoriented progressively around the original preferred storm path. At each new sensor orientation, the value of the objective function was determined.

Sampling anisotropic systems with symmetrical properties (Table 9) is always optimal (Fig. 31) when sensors, regardless of density, are oriented parallel to the storm structure ellipse. Analysis and sampling errors decrease nearly 4% when parameter climatology is considered and the sensors are not misoriented by 45° (the optimal orientation is 325-135°). However, increased sensor density tends to offset sampling errors that result from misorienting a sensor set

(compare plotted result amplitudes). Also, the errors indicate little response to increased covariance function anisotropy (A_8 varied from 0.0 to -0.50) if the system possess symmetric structural characteristics ($A_1 = A_2 = A_3 = A_4 = 6.9$ km).

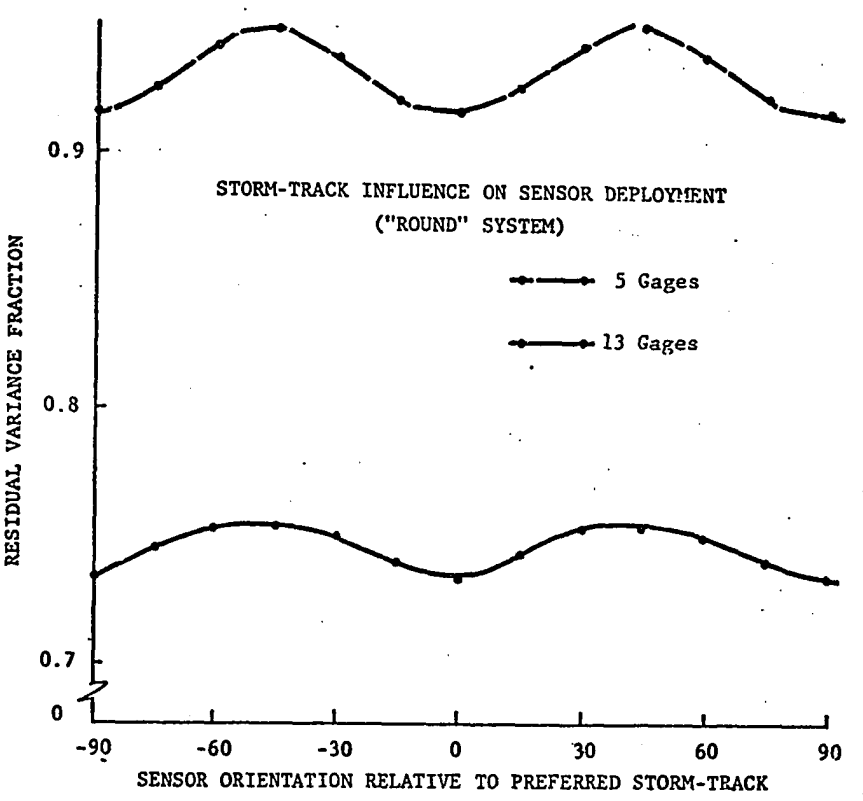


Fig. 31. The influence of a preferred storm track on sensor deployment. Sensors rotated clockwise when oriented positive relative to the preferred track.

A similar evaluation was performed for an anisotropic system that possessed properties of the 7 July 75 accumulated rainfall pattern (Fig. 23 and Table 4). On this date, $A_1 \neq A_2$ and the storm's preferred track was parallel to a 343-163° line (Table 5).

The same sparse and dense gage networks were deployed optimally, systematically reoriented, and objective function values recalculated. The results (Fig. 32) reveal a pronounced sensor-density influence on a network's ability to detect signal variation. The sparse network still performs best when oriented near the preferred storm track. Consideration of parameter climatology improves sampling and analysis quality by almost 6% over the least optimal sensor orientation.

Now, increased sensor density does not have the same positive effect noticed earlier. Additional sensors do not offset the sampling errors that result from misorientation of a sensor set (note the plotted result amplitudes). The thirteen sensors, as deployed originally, represent a best sampling and analysis configuration. The design algorithm searches for the best deployment configuration, but is restricted by severe (though not unrealistic) constraints which include: (1) placement of the first sensor at the grid center; (2) the order which sensors are placed {Fig. 28}; and (3) a relative sensor separation-distance imposed along and across the preferred storm track. These constraints seem to represent a logical strategy for use in sensor deployments based on limited knowledge of the parameter a priori climatology or of the parameter characteristics.

detection of signal variation (term 1 of Eq. (20)). If $A_1 \neq A_2$, increased ellipticity in a system's features (evaluated varying A_8 from -0.07 to -0.67) exerts a stronger influence on the sampling and analysis quality of accumulated rainfall patterns. For parameters with highly anisotropic structural characteristics, sampling and analysis quality is highest only if sensors are oriented along the preferred storm track.

Our conclusion is: sampling and analysis quality is influenced significantly by systems that possess exotic structural characteristics (e.g., Figs. 8E and 14B). Clearly, the deployment problem is complicated and depends significantly on parameter characteristics. The results provide indisputable evidence that coordinated field programs which propose pretty sampling configurations likely will fail to achieve their full potential. The implications are worth considering.

C. Time-Domain Sampling Influence On Sensor Deployment

Sensors deployed to sample accumulated rainfall patterns also provide the information used to investigate details of selected case-study systems. If the morphology of a convective storm were of prime importance, sampling requirements could incorporate time-domain sampling. The concept illustrated in the following example uses the Table 9 covariance function parameter values. For illustrative purposes, the cross-covariance and radar covariance values were set higher than observed (both equal 0.80). Each time influence radius

was set equal to 20.0 minutes.

Three sensor set combinations are presented to emphasize trade offs involved in sampling a moving system: (1) a coarse univariate gage deployment (7 gages spaced 27.2 km^2/gage); (2) a dense univariate gage deployment (13 gages spaced 14.7 km^2/gage); and (3) a bivariate deployment using 7 gages and 63 fixed radar bins spaced 3.0 km^2/bin. Sampling errors for each deployment are generated by using radar observation frequencies and gage accumulation times of 5.0, 7.5, 10.0, 12.5, and 15.0 minutes. The predictor location subset is constrained to be displaced no more than one time unit from the predictand location.

The results assume each parameter's SNR is constant for the temporally spaced "data sets" used. The concept of improved sampling quality through increased sampling frequency is alluded to by Eddy (1964). Improving the SNR of a data set through increased temporal (or spatial) sampling is limited, of course, by each sensor's engineering qualities. However, proper SNR consideration in evaluating the spatial or temporal sampling requirements would improve the results. Eddy (1976) viewed the temporal sampling problem from a similar perspective. He noted "that the uncertainty in observed rainfall amount (and in the objective analysis) increases with increasing rain amount....(and) performing two five-minute accumulation time analyses and adding them together will give a result with greater uncertainty than that associated with

performing one ten-minute accumulation time analysis. Thus, ... one needs more gages if he is going to sum two five-minute analyses." For rainfall patterns on our scale of interest, the SNR has been found to increase with decreasing rainfall volume (limited by sensor engineering qualities). However, the increased SNR with decreased rainfall is not sufficient to produce higher confidence in the sum of four fifteen-minute analyses than in an analysis of one hour rainfall data. Thus, the results represent a sampling quality that is achieved when weak to moderate rainfall producing systems are sampled (intense storms were not part of the Table 4 results).

As a general result, increased temporal sampling reduces spatial data requirements; the placement algorithm produces larger separations between sensors.

More specific results are noted in Fig. 33. An identical sampling quality is achieved by the three following sensor sets: (1) a coarse gage network that accumulates rainfall at five-minute intervals (point 1); (2) a dense gage network that accumulates at eight-minute intervals (point 2); or a bivariate sensor set that observes at 9.5 minute intervals.

If observations at five-minute intervals are needed, gages plus good radar data (point 5) reduce sampling error variances by $\sim$ 34% from those that occur in a coarse gage network (point 1). If only gages are available, sampling accuracy is improved $\sim$ 11% when gage network density is doubled (point 4). Thus, the illustrative results indicate

the trade offs involved when time-domain sampling is incorporated into the deployment of sensors.

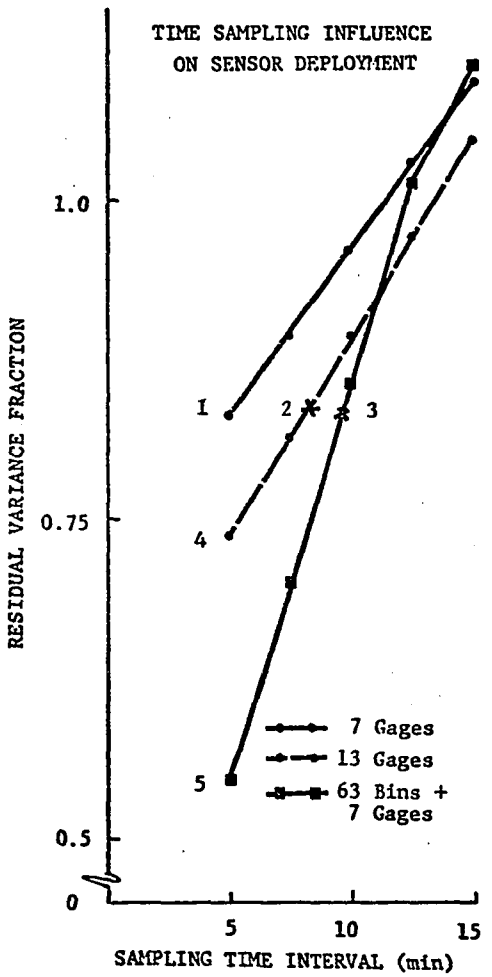


Fig. 33. The influence of time-domain sampling on sensor deployment.

One point of concern is noted. The bivariate placement results "appear" inferior to the univariate results when temporal sampling is infrequent. Clearly, the feature is not realistic. The discrepancy apparently originates in the screening regression algorithm; its purpose is to eliminate the insignificant and redundant predictors. Likely, the predictor array dimensions combined with the relative data paucity in the example and the F-test criteria (used to accept or reject possible predictors) to produce the sub-optimal results.

Because the predictand/radar relationship was so strong (cross-covariance of 0.80), core storage limitations prevented the screening procedure from even considering gage information. F-test criteria, set unrealistically low, reduced efficiency in the screening procedure. Some predictor redundancy likely remained (radar covariance of 0.80). As the sampling time interval decreased, the predictor subset produced represented analysis time information only. Some predictors chosen were separated widely, even though within the analysis location's spatial influence radii. In such instances, predictor intercorrelations (the off diagonal elements of the $\{X^tX\}$ covariance matrix) likely are negative. Because few predictors were available, the net result was an unstable $\{X^tX\}$ matrix that, inturn, corrupts the calculated regression coefficients. Hopefully, these points underscore the pitfalls of blatantly analyzing meteorological data without provisions for quality control.

D. The Sensors Deployed - An Economic/Climatological Balance

So, what is the optimum gage density necessary to meet HIPLEX sampling requirements? That question's answer is approached as follows. Consider Fig. 29B and the climatology it represents. Suppose a reasonably accurate gage analysis is obtained from sensors spaced 60 km^2/gage (point 1). An equivalent analysis is obtained from good radar data and a gage density of 190 km^2/gage (point 2). Analysis accuracy in the gage network is improved only 10% when the gage density is increased to 50 km^2/gage (point 3). At this gage density (point 3), quality radar data, combined with gage information, produces an analysis containing 65% less sampling error variance (point 4) than obtained in the best sampling from gage data alone. If these marginal increases in accuracy are deemed worth the extra resources required to obtain it, then one simply proceeds along the Fig. 29B curve to a sensor density where the project director calls a halt.

The sampling requirements illustrated do not pretend to meet the climatological data requirements of HIPLEX in observing accumulated rainfall patterns. Though a "representative" selection of HIPLEX storms has been investigated throughout the study, the sampling requirements apply only to systems possessing the Table 9 characteristics. Multi-parameter data set qualities, preferred storm tracks, system anisotropy, and spatial and temporal sampling resolution all interact to a

significant degree to influence final determination of a "best" sensor set deployment. However, the evidence is clear--sampling requirements are controlled indisputably by parameter climatology. Clearly, a climatological study focusing on system characteristics of prime interest is justified as an essential and early phase in HIPLEX (and all proposed coordinated field programs as well). Thus, final sensor requirements represent a balance between the demands of economy and climatology.

CHAPTER IX

THE CONCLUSIONS

An optimal sampling and analysis technique has been applied to a fundamental problem faced by HIPLEX investigators in observing and analyzing accumulated rainfall patterns from convective storms on the High Plains. The techniques employed are founded on the theory of optimum interpolation and efficiently used nonlinear programming and concepts of multiple correlation coefficients and stepwise regression analysis. Implementing such techniques requires knowledge of parameter statistical characteristics (climatology). Analyses produced have the desirable statistical properties of a maximum likelihood estimator and are determined from properly chosen, properly weighted predictors.

Here, climatological characteristics of convective storms are evaluated by using numerous data sets from Southeast Montana. A four-dimensional Gaussian-damped function is developed to reflect the relevant physical and statistical characteristics of High Plains convective systems. This

function is enhanced to model parameter covariances that exhibit anisotropy whose major axis isn't necessarily oriented along a Cartesian axis. Structural characteristics of accumulated rainfall patterns are definable with acceptable accuracy using less than 20% of available radar data (from the lowest elevation angle). Storm structure deduced using rain gage observations is accurately represented when the covariance function derives the shape from radar data sets and the signal-to-noise ratio from surface observations. The Z-R relationship effect on covariance function determination is not substantial.

The simple correlation between the covariance function area of moving convective systems and the cloud base area within the 20, 25 and 30 dBz reflectivity contours was 0.45, 0.49 and 0.52, respectively. The correlations indicate covariance function responses to cardinal system features and suggest that weak precipitation rates are not part of any salient storm statistical structure. Without regard to storm size, a simple correlation (-0.76) between storm-motion speed and covariance decorrelation time strongly suggests that any fast moving system decorrelates quickly. When motion speed is normalized by storm statistical dimensions and correlated with system decorrelation time, the coefficient of 0.79 reveals that slower moving small systems are statistically equal in time to faster moving large systems. Finally, the simple correlation (-0.54) between the covariance area of accumulated rainfall patterns

and storm-motion speed provides evidence that slow moving systems generally produce more coherent accumulated rainfall patterns than do the faster moving systems.

In Southeast Montana, the average radar accumulated rainfall pattern was spatially symmetric, was coherent over 80 km^2, would have been sampled and analyzed most optimally if sensors were oriented west northwest-east southeast, and contained almost twice as much signal as noise in their raw data sets. These selected High Plains systems were found to be much smaller than Brady's Oklahoma counterparts.

The optimum interpolation method is generalized to model characteristic differences in parameter means and variances by including a $\hat{\beta}_o$ term in the $\hat{\beta}$ matrix. Considering realistic differences in parameter means and variance permits straightforward and impartial filtering of multivariate data sets; and each predictor is allowed its proper influence in modulating the final analyzed field. Here, a limited gage network, more sparse than that needed to produce a satisfactory analysis on its own, is combined with radar observations to produce satisfactory surface rainfall analyses. Consequently, the bivariate radar/rain gage analyses produced reflect radar-derived precipitation patterns scaled to gage magnitudes.

The interpolation methods demonstrate a minimal Z-R relationship influence on analysis accuracy; in fact, the methodology reveals an objective, very realistic approach toward evaluating an optimum Z-R relationship. The model's signal

recoverability properties point convincingly to its high ability to retain essential data set features while tossing aside nearly 50% of the undesired data set properties.

An experimental design evaluation function is developed which incorporates the effect of observation biases and the relative role played by each parameter in producing a final sensor deployment. Predictand-related sensors are shown essential to network design. Deployment of sensors along and across the preferred storm track is related to covariance anisotropy, gage density, temporal sampling intervals, the availability of radar data, and the interrelationships among the multivariate predictor data sets.

A maximum of 10 gage sensors within the coherent convective storm area is sufficient to fully explain and sample the significant signal variance of a representative High Plains convective system.

As surface gage density is increased, the marginal increase in sampling accuracy is shown for the same representative system; also, results are indicated when radar data is available. A reasonably accurate surface rainfall analysis is possible using gage sensors spaced 60 km^2/gage. An equivalent analysis is obtained from good radar data and a gage density of 190 km^2/gage. Analysis accuracy in the gage network is improved only 10% when the gage density is increased to 50 km^2/gage. Quality radar data, combined with gage information spaced 50 km^2 apart, produces 65% less sampling error variance than

observed in the best sampling from gage information alone.

Misorientation of a sensor set exerts minimal influence on sampling quality when phenomena possess symmetric structural characteristics. However, as system anisotropy increases, sampling quality is influenced substantially by sensor set density and orientation. Sampling quality deteriorates by 12% when parameter climatology, only moderately anisotropic, is not considered. Moreover, the optimal sensor orientation to sample a moving convective system is found best to observe the system's accumulated rainfall pattern.

If radar information is available, sampling quality is maintained even though gage observation frequency is reduced fifty percent.

The sampling and analysis techniques herein have been enhanced significantly over the last several years; they can digest complex data sets and provide realistic solutions to atmospheric sampling and analysis problems. The techniques are not perfect. A number of colleagues, however, share the view that the techniques and money have a lot in common--"they aren't everything, but they're way ahead of what's in second place." Application of the techniques to the problems of HIPLEX have, I trust, demonstrated the techniques' merit. Their application in other areas is invited.

CHAPTER X

ADDITION RESEARCH: APPLICATIONS

Analysis techniques herein present highly specialized methods which now can be applied to all sorts of complex problems; but care must be taken as to the limitations noted in Chapter IV. Some applications follow.

A. Use of Modelled Covariance Functions

The covariance model could be upgraded to account for system motion directions and speeds which would be incorporated as additional decision variables to describe a data set's covariance. More decision variables permit cross-covariance function models to realize maximum values at non-zero spatial or temporal lags. Failure to consider such relative pattern lags might wreak havoc in bivariate analyses involving vertical motion patterns on the synoptic scale and their corresponding temperature data sets. Finally, the ability to model and hence, analyze multivariate phenomena with interacting

action scales is important in any phase of meteorological analysis; the covariance function and analysis model should be upgraded to permit this possibility.

Parameter characteristics shown by modelled covariance functions may be used as a tool to evaluate modification experiments. A climatology of nature's evolution inside the convective storm and manifested by covariance and/or cross-covariance function could be used to detect any abnormal changes in the physics of seeded clouds.

Operationally, the covariance modelling technique is useful to find optimally the information in real-time radar precipitation data, in turn used to update numerical precipitation guidance.

B. Use of the Multivariate Analysis Methodology

A next step for analysis might be trivariate analysis involving (among parameters) surface divergence patterns analyzed in conjunction with radar and surface gage data sets (or satellite information combined with radar and surface data). Already, the National Meteorological Center (NMC) of the National Weather Service has performed trivariate analysis of temperature and wind field components.

For certain purposes, dynamically constrained analyses offer an improved analysis quality. Analyses consistent with atmospheric physics are possible and might use a covariance function model that would contain the constraining information (i.e., an analytical function derived from physical equations

describing the atmosphere). Actual constraint on a final analysis responds to the predictor set chosen for analysis Dynamic constraints might also follow the statistical methods proposed by Eddy (1973).

Analyses, using sequential updates of the covariance filtering properties, seem possible on many atmospheric scales. Surface rainfall analysis initially might use climatological filtering properties which are updated to reflect a given storm's characteristics as the storm evolved over the analysis area. On a larger scale, synoptic analysis could use a similar technique. Currently, the NMC statistically types each day's synoptic pattern with 500 mb correlations. As the atmosphere evolved (or was forecast to evolve), analysis quality could improve by updating the model's filtering properties with new information.

The impact on analysis accuracy from sensors of differing qualities (e.g., satellite vs rawinsonde information) or from data sets with nonstationarity characteristics could be evaluated by incorporating the error covariance matrix into the analysis routines (see Appendix A).

C. Use of the Optimal Sampling Methodology

The current techniques are set to evaluate data requirements on many levels. What data volume (for instance) from the indirect probing sensors currently available will NMC require for use in its numerical prediction models? Clearly, maximum improved weather forecast accuracy involves optimal

trade offs between standard high quality data sets and between a proper mix of temporally and spatially selected satellite data. Would a simple relocation of NMC's coarse or fine mesh grids reduce interpolation errors (and hence improve forecast accuracy)?

The sampling method can address the cost-benefit in changing networks now existing. How many sensors can be withdrawn from a gage network without significant cost to defining any storm? What would be the value associated with adding N new sensors to observe mesoscale systems across our United States? Where would be the best place for remote automated meteorological observing systems in the vast ocean areas of the world?

Thus, our list goes on and on.

XI. THE BIBLIOGRAPHY

Alaka, M.A., 1970: "Theoretical and Practical Considerations for Network Design," Meteor. Monogr., Vol. 11, No. 33, pp 20-27.

Alaka, M.A., and R.C. Elvander, 1972a: "Matching Observational Accuracy and Sampling Resolution in Meteorological Data Acquisition Experiments," J. Appl. Meteor., Vol. 11, pp 567-577.

Alaka, M.A., and R.C. Elvander, 1972b: "Optimum Interpolation from Observations of Mixed Quality," Mon. Wea. Rev., Vol. 100, pp 612-624.

Alaka, M.A., 1974: "The Advantages and Requirements of Optimum Interpolation in the Objective Analysis of Meteorological Fields," Proceedings of Fifth Conference on Weather Forecasting and Analysis, March 1974, pp 148-151.

Baer, L., and G.W. Withee, 1971: "A Methodology for Defining Operational Synoptic Temporal Oceanic Sampling Systems, I. Stationary Conditions, II. Nonstationary Conditions," J. Appl. Meteor., Vol. 10, pp 1053-1065.

Barnes, S.L., 1964: "A Technique for Maximizing Details in Numerical Weather Map Analysis," J. Appl. Meteor., Vol. 3, pp 396-409.

Barnes, S.L., 1973: "Mesoscale Objective Map Analysis Using Weighted Time-Series Observations," NOAA Technical Memorandum ERLTM-NSSL 62.

Barnes, S.L., editor, 1974: "Papers on Oklahoma Thunderstorms, April 29-30, 1970," NOAA Technical Memorandum ERLTM-NSSL 69.

Battan, L.J., 1973: Radar Observations of the Atmosphere, The University of Chicago Press, Chicago.

Bergman, K.H., and W.D. Bonner, 1976: "Analysis Error as a Function of Observation Density for Satellite Temperature Soundings with Spatially Correlated Errors," Mon. Wea. Rev., Vol. 104, pp 1308-1316.

Bergthorsson, P., and B. Dòòs, 1955: "Numerical Weather Map Analysis," Tellus, Vol. 7, pp 329-340.

Bessemoulin, J., 1960: "Contribution to the Study of Meteorological Networks," WMO Technical Note No. 30, Annex 9, pp 76-82.

Best, D.L., 1973: An Evaluation of Noise Characteristics, The University of Oklahoma, Department of Meteorology Report.

Brady, P.J., 1975: "Matching Rain Gauge Placement to Precipitation Patterns," Proceedings of the National Symposium on Precipitation Analysis for Hydrologic Modelling, June 1975, pp 111-122.

Brady, P.J., 1976: Multivariate Experimental Design in Meteorology, The University of Oklahoma, Department of Meteorology Report.

Brady, P.J., 1977: "Optimal Sampling and Analysis Using Two Variables and Modeled Cross-Covariance Functions," J. Appl. Meteor., (in press).

Brandes, E.A., 1975: "Optimizing Rainfall Estimates with the Aid of Radar," J. Appl. Meteor., Vol. 14, pp 1339-1345.

Cressman, G.P., 1959: "An Operational Objective Analysis System," Mon. Wea. Rev., Vol. 87, pp 367-374.

Draper, N.R., and H. Smith, 1966: Applied Regression Analysis, John Wiley and Sons, Inc., New York.

Drozdov, O.A., and A.A. Sepelevskij, 1946: "The Theory of the Interpolation of Meteorological Elements in a Stochastic Field and their Application to Questions of Weather Maps and Network Rationalization," Trudy NIU GUGMS, Series I, No. 13.

Drufuca, G., and I.I. Zawadzki, 1975: "Statistics of Raingage Data," J. Appl. Meteor., Vol. 14, pp 1419-1429.

Eddy, A., 1963: "Kinematic Divergence and Large Scale Energy Conservation, "Arctic Meteorological Research Group, Department of Meteorology, McGill University, Montreal, Publication No. 60, Scientific Report No. 8.

Eddy, A., 1964: "The Objective Analysis of Horizontal Wind Divergence Fields," Q. J. R. M. S., Vol. 90, pp 424-440.

Eddy, A., 1967: "Two-Dimensional Statistical Objective Analysis of Isotropic Scalar Data Fields," Report No. 5, Atmospheric Science Group, The University of Texas; (condensed version published as "The Statistical Objective Analysis of Scalar Data Fields," J. Appl. Meteor., Vol. 6, pp 597-609).

Eddy, A., 1970: "The Statistical Evaluation of Observational Data," Meteor. Monogr., Vol. 11, No 33, pp 110-120.

Eddy, A., 1973: "The Objective Analysis of Atmospheric Structure," J. Meteor. Soc. Japan, Vol. 51, pp 450-457.

Eddy, A., 1974: "An Approach to the Design of Meteorological Field Experiments," Mon. Wea. Rev., Vol. 102, pp 702-707.

Eddy, A., 1976: "Optimal Raingage Densities and Accumulation Times: A Decision-Making Procedure," J. Appl. Meteor., Vol. 15, pp 962-971.

Eddy, A., and P. McDonald, 1977: "An Examination of Objective Analysis Accuracy Using Manova," Proceedings of Fifth Conference on Probability and Statistics in Atmospheric Sciences, November 1977, (in press).

Efroymson, M.A., 1960: Mathematical Methods for Digital Computers, Ralston, A., and H.S. Wilf, John Wiley and Sons, New York.

Ehrlich, P.R., 1971: The Population Bomb, Ballantine Books, New York.

Endlich, R.M., and R.L. Mancuso, 1968: "Objective Analysis of Environmental Conditions Associated with Severe Thunderstorms and Tornadoes," Mon. Wea. Rev., Vol. 96, pp 342-350.

Gandin, L.S., 1963: Objective Analysis of Meteorological Fields, Leningrad (Translated by Israel Program for Scientific Translations, Jerusalem, 1965).

Gandin, L.S., S.A. Mashkovich, M.A. Alaka, and F. Lewis, 1967: "Design of Optimum Networks for Aerological Observing Stations," World Meteorological Organization, Planning Report No. 21.

Gandin, L.S., 1970: "The Planning of Meteorological Station Networks," World Meteorological Organization, Technical Note No. 111.

Gilman, D.L., F.J. Fuglister and J.M. Mitchell, Jr., 1963: "On the Power Spectrum of 'Red Noise'," J.A.S., Vol. 20, pp 182-184.

Graybill, F.A., 1961: An Introduction to Linear Statistical Models, McGraw-Hill Book Company, Inc., New York.

Hendrick, R.L., and G.H. Comer, 1970: "Space Variations of Precipitation and Implications for Network Design," J. Hydro., Vol. 10, pp 151-163.

Himmelblau, D.M., 1972: Applied Nonlinear Programming, McGraw-Hill Book Company, Inc., New York.

Interior Department, 1976: "HIPLEX Data Inventory," Bureau of Reclamation, Denver.

John, 95: "The Revelation," The Bible, (New Testament).

Jones, D.M.A., 1956: "Rainfall Drop-Size Distribution and Radar Reflectivity," Res. Rept. No. 6, Illinois State Water Survey.

Julian, P.R., and H.J. Thiebaux, 1975: "On Some Properties of Correlation Functions used in Optimum Interpolation Schemes," Mon. Wea. Rev., Vol. 103, pp 605-616.

Kasahara, A., 1972: "Simulation Experiments for Meteorological Observing Systems for Garp," Bull. Amer. Meteor. Soc., Vol. 53, pp 252-264.

Kays, M.D., 1974: Optimal Sampling of a Stratospheric Sudden Warming, The University of Oklahoma, Department of Meteorology Report.

Kreitzberg, C., and D. Perkey, 1974: "Numerical Weather Prediction Including Mesoscale Convective Interaction," Colloquium on Subsynoptic Extratropical Weather Systems; Observation, Analysis, Modeling and Prediction.

Lacy, C., 1973: "Objective Analysis Using Modeled Space-Time Covariances: An Evaluation," ECOM-5514, Atmospheric Sciences Laboratory, U.S. Army Electronics Command, White Sands Missile Range, New Mexico.

Nelder, J.A., and R. Mead, 1964: "A Simplex Method for Function Minimization," Computer J., Vol. 7, pp 308-313.

Northrup, G.M., E.L. Davis and E.R. Sweeton, 1972: Environmental Models/Systems Effectiveness Study (EM/SE), Center for Environment and Man, No. 4052-436.

Panofsky, H.A., 1949: "Objective Weather-Map Analysis," J. Meteor., Vol. 6, pp 386-392.

Passi, R.M., 1975: "Statistical Estimation of Meteorological Parameters with Correlated Observations," Mon. Wea. Rev., Vol. 103, pp 521-527.

Pasteris, P.A., 1975: Objective Analysis of Northeast Oklahoma Rainfall, The University of Oklahoma, Department of Meteorology Report.

Paviani, D.A., and D.M. Himmelblau, 1969: "Constrained Non-Linear Optimization by Heuristic Programming," Operations Research, Vol. 17, pp 872-882.

Phillips, N.A., 1976: "The Impact of Synoptic Observing and Analysis Systems on Flow Pattern Forecasts," Bull. Amer. Meteor. Soc., Vol. 57, pp 1225-1240.

Sasaki, Y., 1958: "An Objective Analysis Based on the Variational Method," J. Meteor. Soc. Japan, Vol. 36, pp 77-88.

Schlatter, T.W., 1975: "Some Experiments with a Multivariate Statistical Objective Analysis Scheme," Mon. Wea. Rev., Vol. 103, pp 246-257.

Schneider, S.H., 1976: The Genesis Strategy, Plenum Press, New York.

Shapiro, M.A., and J.T. Hastings, 1973: "Objective Cross-Section Analyses by Hermite Polynomial Interpolation on Isentropic Surfaces," J. Appl. Meteor., Vol. 12, pp 753-762.

Spendley, W., G.R. Hext, and F.R. Himsworth, 1962: "Sequential Application of Simplex Designs in Optimization and Evolutionary Operations," Technometrics, Vol. 4, pp 441-461.

Steinitz, G., A. Huss, A. Manes, R. Sinai, and Z. Alerpson, 1971: "Optimum Station Network in the Tropics," J. Appl. Meteor., Vol. 10, pp 364-369.

Stephens, J.J., and A.L. Polan, 1971: "Spectral Modification by Objective Analysis," Mon. Wea. Rev., Vol. 99, pp 374-378.

Stocker, D.C., 1969: "A Comparative Study of Nonlinear Programming Codes," M.S. Thesis, The University of Texas, Austin, Texas.

Tatsuoka, M.M., 1971: Multivariate Analysis, John Wiley and Sons, Inc., New York.

Thiebaux, H.J., 1973: "Maximally Stable Estimation of Meteorological Parameters at Grid Points," J.A.S., Vol. 30, pp 1710-1714.

Thiebaux, H.J., 1974: "Estimation of Covariances of Meteorological Parameters Using Local-Time Averages," J. Appl. Meteor., Vol. 13, pp 592-600.

Thiebaux, H.J., 1975: "Experiments with Correlation Representations for Objective Analysis," Mon. Wea. Rev., Vol. 103, pp 617-627.

Thiebaux, H.J., 1976: "Anisotropic Correlation Functions for Objective Analysis," Mon. Wea. Rev., Vol. 104, pp 994-1002.

Thiebaux, H.J., 1977: "Extending Estimation Accuracy with Anisotropic Interpolation," Mon. Wea. Rev., (in press).

Yerg, M.C., 1973a: A Systems Approach to Optimal Experimental Design in Meteorology, The University of Oklahoma, Department of Meteorology Report.

Yerg, M.C., 1973b: An Optimal Sampling and Analysis Methodology, The University of Oklahoma, Department of Meteorology Report.

Zawadzki, I.I., 1975: "On Radar-Raingage Comparison," J. Appl. Meteor., Vol. 14, pp 1430-1436.

APPENDIX A

The multivariate multiple linear regression model of Chapter III evolved from classical multivariate analysis models developed theoretically in texts such as Graybill (1961) and Tatsuoka (1971) and is represented by:

$$Y = X\beta + \varepsilon \qquad \text{(A1)}$$

where X is an n x m matrix containing the independent stochastic variables, Y is an n x 1 matrix of dependent stochastic variable, β is an m x 1 matrix of regression coeffients and ε is an n x 1 population error term. Assume X and Y have a bivariate distribution. The basic theory of classical multiple linear regression is founded on the following assumptions:

(1) the fundamental assumption is that X and Y are linearly related;

(2) ε is an independent stochastic disturbance term whose distribution is not specified;

(3) the expected mean value of ε is zero (i.e., $E\{\varepsilon\} = 0$);

(4) homoscedasticity or constant residual variance of ε is assumed (i.e., $\mathrm{Var}\{\varepsilon\} = \mathrm{Var}\{Y\} = \sigma^2$);

(5) statistical independence is assumed among the disturbance terms (i.e., $E\{e_i e_j\} = 0$). Because Y is a linear function of ε, Y_i and Y_j also are independent statistically;

(6) the number of observations (that is, the sample size n) must be larger than the number of regression coefficients (m plus 1, where m independent predictor variables, with non-zero mean values, are chosen for the regression analysis).

The regression function of Y on X is expressed by:

$$\hat{Y} = X\hat{\beta} \qquad \text{(A2)}$$

where $\hat{\beta}$ estimates the population weights β and $\hat{Y}$ estimates the population values Y. Residual errors associated with sample regression become $e = Y - \hat{Y}$. Frequently residual errors are assumed Gaussian distributed with $\varepsilon \sim N(0,\sigma^2 I)$. The works of Eddy (1967, p. 39) and Eddy and McDonald (1977) support the Gaussian assumption.

If the error sum of squares $e^t e$ (residual variance) are minimized, and if e is $\sim N(0,\sigma^2 I)$, then:

$$\hat{\beta} = (X^t X)^{-1} X^t Y \qquad \text{(A3)}$$

which is an unbiased maximum-likelihood estimate $E(\hat{\beta}) = \beta$. The likelihood function, when maximized, produces a minimum value in the quantity $e^t e$ and justifies use of the least squares

principle. The $\hat{\beta}$ chosen has minimum variance among all the unbiased $\hat{\beta}$'s that could have been chosen (i.e., its an efficient estimator). As a maximum likelihood estimator, the chosen $\hat{\beta}$ also has desirable properties of consistency (estimator value approaches population value as the sample size increases) and sufficiency (an estimator that utilizes all information a sample contains about the parameter to be estimated).

Meteorological systems in general are not stationary. For example, seasonal and geographical differences produce non-stationary characteristics in meteorological data. In the current study, the data sets represent systems from one geographical area (Southeast Montana) and one season (summer) and are judged not influenced by stationarity considerations.

However, if the residual variance appears non-stationary or if the residual errors are correlated in space and time (as would occur when systematic small scale perturbations are imposed on the basic signal), then the ordinary least squares estimates in Eq. (A3) do not apply. The effect of autocorrelated residuals or of statistically dependent observations (multicollinearity) is the same: the regression coefficient variance (Eq. (4)) is larger than it would otherwise be. Analyses which fail to consider this problem will give false impressions of unrealistically high analysis significance (Eq. (5))! Thus, if the residuals, e, are autocorrelated or if the residual variances are non-stationary, the data must be

transformed to other variables which then satisfy the usual assumptions of $\varepsilon \sim N(0,\sigma^2 I)$. The transformation is accomplished using the error covariance matrix $V = E(\varepsilon\varepsilon^t)$. Eq. (A3) becomes:

$$\hat{\beta} = (X^t V^{-1} X)^{-1} X^t V^{-1} Y \tag{A4}$$

When $V \neq I$, the residual variance no longer has its minimum value because $e^t e$ is minimized to give a $\hat{\beta}$ free from the biasing influence of correlated residuals. Gilman et. al. (1963) and Eddy (1970) also illustrate the point wherein:

$$\hat{\sigma}^2 = \frac{\sigma_\varepsilon^2}{N}\left[1 + \frac{2}{N}\sum_{\tau=1}^{N-1}(N - \tau)\rho(\tau)\right] \tag{A5}$$

where the estimated or sample residual variance $\hat{\sigma}^2$ is a function of the population error variance σ_ε^2, the population autocorrelation $\rho(\tau)$, and the residual sample size N. For correlated residuals, the residual variance is larger by a factor:

$$\frac{2}{N}\sum_{\tau=1}^{N-1}(N - \tau)\rho(\tau) \tag{A6}$$

where $\rho(\tau)$ decreases monotonically with τ.

Draper and Smith (1966) refer to this situation as weighted least squares. They indicate that failure to consider the effects contained in the covariance matrix V will produce objective analyses whose variance with the signal (and not the observations) is not a minimum. An important concept for

needing V in the model is that its use minimizes regression coefficient variance (Eq. (4)). Our aim is minimum variance between the analysis and the truth (Eq. (5)) and not minimum variance in the residuals. The manner in which V is estimated is described by Best (1973), who assumed residuals that possessed stationary variances and were autocorrelated as a first order Markov process. Origin of the error covariance matrix for a two dimensional case is now described:

A bivariate normal probability distribution of two stochastic processes $y_1(t)$ and $y_2(t)$, each possessing a zero mean value, may be written as:

$$f(y_1 y_2) = AMP * EXP \left\{ \frac{-1}{2(1-\rho^2)} \left[\left(\frac{y_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{y_1 y_2}{\sigma_1 \sigma_2}\right) + \left(\frac{y_2}{\sigma_2}\right)^2 \right] \right\} \quad \text{(A7)}$$

where $AMP^{-1} = 2\pi\, \sigma_1 \sigma_2 \sqrt{(1 - \rho^2)}$ and Var $y_1 = \sigma_1^2$, Var $y_2 = \sigma_2^2$ and ρ is the correlation between y_1 and y_2. If R is defined as:

$$R = \begin{pmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{pmatrix} \quad \text{(A8)}$$

where:

$$r_{11} = 1/\sigma_1^2(1-\rho^2) \qquad r_{22} = 1/\sigma_2^2(1-\rho^2)$$

$$r_{12} = r_{21} = -\rho/\sigma_1\sigma_2(1-\rho^2) \tag{A9}$$

then Eq. (A7 may be written in quadratic notation as:

$$f(y_1y_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \text{ EXP } \left[- \tfrac{1}{2}\{(y_1y_2)R(y_1y_2)^t\}\right] \tag{A10}$$

Graybill (1961) proves $R^{-1} = V$, defined to be the variance-covariance matrix given by:

$$R^{-1} = V = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \tag{A11}$$

From Eq. (A11) note that Var $y_1 = \sigma_1^2$, Var $y_2 = \sigma_2^2$ and $\text{Cov}(y_1y_2) = \text{Cov}(y_2y_1) = \rho\sigma_1\sigma_2$. Extension to the multivariate case is straightforward.

What interpretations may be given to V? In classical regression, a trivial point is made when y_1 and y_2 are completely uncorrelated. In this case V becomes diagonal and the two processes are independent.

The interpretation is somewhat different when y_1 and y_2 are stochastic containing signal plus random autocorrelated

noise contributions. Best (1973) shows the noise structure of observations (Eq. (A11)) to be identical with the analysis model's residual structure; individual values are somewhat larger than observed in Eq. (A11). Thus, Eq. (A11) is relabelled an error covariance matrix and is given by:

$$V = \begin{pmatrix} \varepsilon_1\varepsilon_1 & \varepsilon_1\varepsilon_2 \\ \varepsilon_2\varepsilon_1 & \varepsilon_2\varepsilon_2 \end{pmatrix} \qquad \text{(A12)}$$

where ε is the population error term (Eq. (A1)). In addition $\varepsilon_1\varepsilon_1$ represents the model's residual variance at one spatial or temporal location <u>set</u> (i.e., the realization $\hat{Y}_1 = X\hat{\beta}$) and $\varepsilon_2\varepsilon_2$ the residual variance at a second realization ($\hat{Y}_2 = X\hat{\beta}$).

APPENDIX B

The completed analysis model is given by:

$$Y = X_{AG}\ \beta_{AG} + X_{AR}\ \beta_{AR} + \varepsilon \qquad \text{(B1)}$$

where X_{AG} represents rain gage observations, X_{AR} the radar reflectivity precipitation estimates, β_{AG} and β_{AR} the population regression coefficients, and ε the population error matrix. Rainfall estimates at desired locations are given by Y. The model is termed "completed" because the objective function used in sensor placement is developed to "account" for parameters with non-zero mean values and differing variances.

Previously, Chapter III.C.1 notation defined:

$\beta^t = \{\beta_1\ \beta_2\ \ldots\ \beta_j\ \ldots\ \beta_M\}$ with dimensions 1 x M

$X = \{X_1\ X_2\ \ldots\ X_j\ \ldots\ X_M\}$ with dimensions N x M

Let μ_Y equal the population Y mean, $\mu = \{\mu_1\ \mu_2\ \ldots\ \mu_j\ \ldots\ \mu_M\}$ represent the population X_j mean and L^t be a 1 x N utility row matrix = {1 1 ... 1}. X and β^t are partitioned as $\{X_G, X_R\}$

and $\{\beta_G^t, \beta_R^t\}$. Now let:

$$X_A = \{1\ X_1\ X_2\ \ldots\ X_M\} \quad \text{and} \quad \beta_A^t = \{\beta_o\ \beta_1\ \beta_2\ \ldots\ \beta_M\}$$

where β_o carries the parameter mean values. Eq. (B1) can be rewritten:

$$Y = \mu_Y L - L\mu\beta + X\beta + \varepsilon \tag{B2}$$

with $\mu_Y L - L\mu\beta$ accounting for regression parameter-mean values whose sample estimate is given by $\bar{Y} - \bar{X}\beta$. If $Z = X - L\mu$, the minimization process yields:

$$\hat{\beta} = \{(X-L\mu)^t(X-L\mu)\}^{-1}\ \{X-L\mu\}^t\ \{Y-\ _YL\}$$

$$\hat{\beta} = (Z^tZ)^{-1}\ Z^t(Y-\mu_YL) \tag{B3}$$

where $\hat{\beta}$ estimates β. Substituting for Y from Eq. (B2):

$$\hat{\beta} = \beta + (Z^tZ)^{-1}\ Z^t\varepsilon \tag{B4}$$

The analysis equation obtained from Eq. (B2) and Eq. (B4) is now:

$$\hat{Y} = (X-L\mu)\hat{\beta} + \mu_YL \tag{B5}$$

Suppose $\bar{X}$ (sample mean) is used in place of μ (population mean),

and $\bar{Y}$ in place of μ_Y. Eq. (B5) becomes:

$$\tilde{Y} = (X-L\bar{X})\hat{\beta} + \bar{Y}L \tag{B6}$$

Note that $X - L\bar{X}$ may be written as $(X-L\mu) - L(\bar{X}-\mu)$ and, when used in Eq. (B6), yields:

$$\tilde{Y} = Z\hat{\beta} - L(\bar{X}-\mu)\hat{\beta} + (\bar{Y}-\mu_Y)L + \mu_Y L \tag{B7}$$

Sample regression residual error is defined as:

$$e = Y - \tilde{Y} \tag{B8}$$

Use of Eqs. (B2) and (B6) permits Eq. (B8) to be rewritten as:

$$e = Z(\beta-\hat{\beta}) + L\{(\bar{X}-\mu)\hat{\beta} - (\bar{Y}-\mu_Y)\} + \varepsilon$$

$$e = \{L(\bar{X}-\mu)-Z\}\{\hat{\beta}-\beta\} + L\{(\bar{X}-\mu)\beta - (\bar{Y}-\mu_Y)\} + \varepsilon$$

which in turn is rewritten using Eq. (B4):

$$e = \left[\{L(\bar{X}-\mu)-Z\}\{Z^tZ\}^{-1}Z^t+I\right]\varepsilon + L\{(\bar{X}-\mu)\beta - (\bar{Y}-\mu_Y)\} \tag{B9}$$

Note that $\bar{X} = (L^tX)/N$ and $L\mu = (LL^t/N)L\mu$. Thus, $L(\bar{X}-\mu)$ becomes $(LL^t/N)\ (X-L\mu) = (LL^t/N)Z$. Substituting Eq. (B9) it gives:

$$e = \left[\{(LL^t/N)-I\}Z\{Z^tZ\}^{-1}Z^t + I\right]\varepsilon + L\{(\bar{X}-\mu)\beta - (\bar{Y}-\mu_Y)\} \quad \text{(B10)}$$

Now assume $\varepsilon \sim N(0,\sigma_\varepsilon^2)$. For one particular analysis situation, $\bar{X}$ and $\bar{Y}$ are modelled and thus, are constants.

The residual sum of squares from Eq. (B10)) can be expanded and simplified to yield:

$$E(e^te) = E\Big[\varepsilon^t\{I+Z(Z^tZ)^{-1}Z^t(LL^t/N-I)\}\varepsilon$$
$$+ 2\varepsilon^t\{I+Z(Z^tZ)^{-1}Z^t(L^tL/N-I)\}L\{(\bar{X}-\mu)\beta - (\bar{Y}-\mu_Y)\}$$
$$+ \{\beta^t(\bar{X}^t-\mu^t) - (\bar{Y}^t-\mu_Y^t)L^t\}L\{(\bar{X}-\mu)\beta - (\bar{Y}-\mu_Y)\}\Big]$$

$$E(e^te) = E\Big[\varepsilon^t\{I+Z(Z^tZ)^{-1}Z^t(LL^t/N-I)\}\varepsilon + N\{(\bar{X}-\mu)\beta - (\bar{Y}-\mu_Y)\}^2\Big] \quad \text{(B11)}$$

since $I + Z\,(Z^tZ)^{-1}Z^t\,(LL^t/N - I)$ is idempotent. Eq. (B11) further expands to:

$$E(e^te) = E\Big[\varepsilon^tI\varepsilon + \varepsilon^t(\bar{X}-\mu)\{Z^tZ\}^{-1}(\bar{X}-\mu)^t\varepsilon - \varepsilon^tZ(Z^tZ)^{-1}Z^t\varepsilon + N\{(\bar{X}-\mu)\beta - (Y-\mu_Y)\}^2\Big] \quad \text{(B12)}$$

Graybill (1961, page 87) proves that the expected value of a quadratic form e^tAe is equal to σ_e^2 trace (A). If A also is an

idempotent matrix of rank k, the expected value reduces to $k\sigma_e^2$. Because ε is assumed $\sim N(0,\sigma_\varepsilon^2)$, Eq. (12) simplifies to:

$$E(e^te) = \{N + N(\bar{X}-\mu)\{Z^tZ\}^{-1}(\bar{X}-\mu)^t - M\}\sigma_\varepsilon^2 + N\{(\bar{X}-\mu)\beta + (Y-\mu_Y)\}^2 \quad \text{(B13)}$$

since $Z\ (Z^tZ)^{-1}\ Z^t$ is idempotent of rank M--the actual number of predictors (M_G gages and M_R radar bins) used to determine residual variance at a predictand location. Now $\sigma_e^2 = E(e^te)/N$, so:

$$\sigma_e^2 = \{1 - \frac{M}{N} + (\bar{X}-\mu)\{Z^tZ\}^{-1}(\bar{X}-\mu)^t\}\sigma_\varepsilon^2 + \{(\bar{X}-\mu)\beta - (\bar{Y}-\mu_Y)\}^2 \quad \text{(B14)}$$

Eq. (B14) states that, in the limit as $N\to\infty$, σ_e^2 is a biased estimate of σ_ε^2 with the bias arising from modelling errors in $\bar{X}$ and $\bar{Y}$ (unless $\bar{X} = \mu$ and $\bar{Y} = \mu_Y$).

Consider the effect of random modelling errors in $\bar{X}$ and $\bar{Y}$. Since our analyses are surface rainfall estimates, model $\bar{X}_G = \bar{Y}$ because $\mu_G = \mu_Y$. Thus:

$$\mu = \{\mu_{G_1}\ \mu_{G_2}\ \cdots\ ,\ \mu_{R_1}\ \mu_{R_2}\ \cdots\}$$

which has dimensions N x $(M_G + M_R)$. Let the grid point rain estimate be derived from $X = \{X_{G_1}\ X_{G_2}\ \cdots\ X_{G_{M_G}}\ ,\ X_{R_1}\ X_{R_2}\ \cdots\ X_{R_{M_R}}\}$

and define:

$$C = E\{(\bar{X}-\mu)^t(\bar{X}-\mu)\} \tag{B15}$$

Assume a random modelling error over many analyses such that:

$$\bar{X}_G \sim N(\mu_G, \sigma^2_{\bar{G}}) \sim N(\mu_G, \sigma^2_G/M_G)$$

and

$$\bar{X}_R \sim N(\mu_R, \sigma^2_R/M_R).$$

The expected residual variance over many analyses (Eq. (B13)) is:

$$E(\sigma^2_e) = \{1 - \frac{M}{N} + (\bar{X}-\mu)\{Z^tZ\}^{-1}(\bar{X}-\mu)^t\}\sigma^2_\varepsilon + \beta^tC\beta - 2(\bar{X}-\mu)\beta(\bar{Y}-\mu_Y) + \sigma^2_G/M_G \tag{B16}$$

Now,

$$\sigma^2_{\bar{G}_i} = \sigma^2_{\bar{G}_j} \qquad \rho_{\bar{G}_{ij}} = 1 \qquad \sigma^2_{\bar{R}_i} = \sigma^2_{\bar{R}_j} \qquad \rho_{\bar{R}_{ij}} = 1$$

$$\text{and} \quad \rho_{\bar{R}_i\bar{G}_i} = \bar{\rho} \text{ for all ij.}$$

Briefly stated, our assumptions do not permit relative biases among the gages or among the radar observations nor do they allow non-homoscedastic properties in the observation sets. The correlation between variations of $\bar{X}_G$ and $\bar{X}_R$ about their

population mean values is denoted by $\bar{\rho}$.

Note that $\bar{X}-\mu = L^t(X-\mu)/N = L^tZ/N$. So from Eq. (B16):

$$\lim_{N\to\infty} E(\sigma_e^2) = \{1 - \frac{M}{N} + L^tZ(Z^tZ)^{-1}Z^tL/N^2\}\sigma_\varepsilon^2 \qquad \text{(B17)}$$

$$+ \beta^tC\beta - 2(\bar{X}-\mu)\beta(\bar{Y}-\mu_Y) + \sigma_G^2/M_G$$

Since $\sigma_\varepsilon^2 = (1 - R^2)\sigma_G^2$, divide Eq. (B17) by σ_G^2 to get (after some non-obvious simplifications):

$$E(\sigma_e^2/\sigma_G^2) = 1-R^2 + (1-\Sigma\hat{\beta}_G)^2/M_G + (\sigma_R/\sigma_G)^2(\Sigma\hat{\beta}_R)^2/M_R \qquad \text{(B18)}$$

$$- 2(\sigma_R/\sigma_G)\bar{\rho}\{(\Sigma\hat{\beta}_R)-(\Sigma\hat{\beta}_R)(\Sigma\hat{\beta}_G)\}/(\sqrt{M_GM_R}).$$

Equation (B18) becomes a contribution from each analysis location to the objective function that is minimized by an optimal sensor-site selection. The $\hat{\beta}$'s determined use sensor-site location and the modelled covariance function. Additionally, σ_R, σ_G, $\bar{X}_R$, $\bar{X}_G$ and $\bar{\rho}$ must now be modelled.

The ratio σ_e^2/σ_G^2 represents the average fraction of unexplained variance at each analysis location (grid point or station). For an entire grid array and for any sensor location set, the weighted mean residual variance in space and time is given by:

$$F = \frac{1}{G}\left[\sum_{i=1}^{G} \eta_i \, E(\sigma_e^2/\sigma_G^2)\right] \tag{B19}$$

where G is the number of grid points, η_i weights each grid point by its relative worth and $\Sigma\eta_i = 1$. The summation constitutes the objective function minimized for the results presented in Chapter VIII. Sensor placement implications from each objective function term (Eq. (B19)) is discussed in Chapter VII.

APPENDIX C

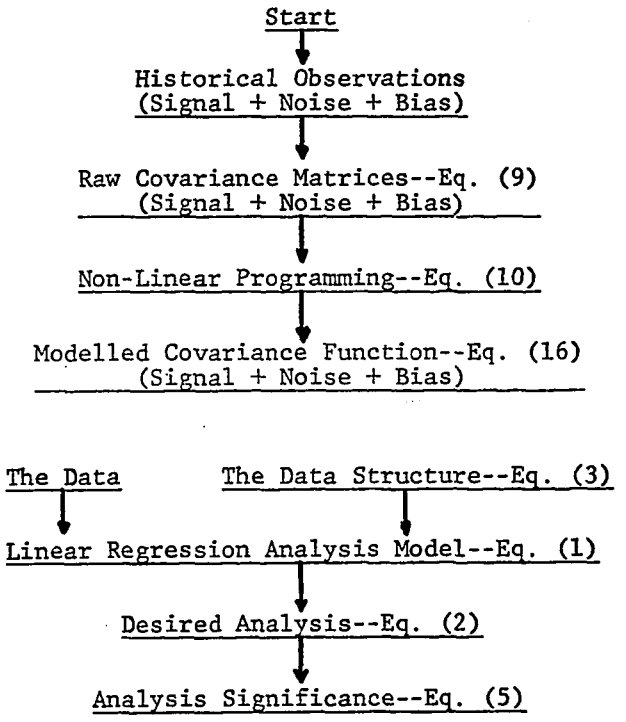


Fig. 34. A computational overview of the steps needed to produce an optimum interpolation analysis. Equations refer to those in the text.

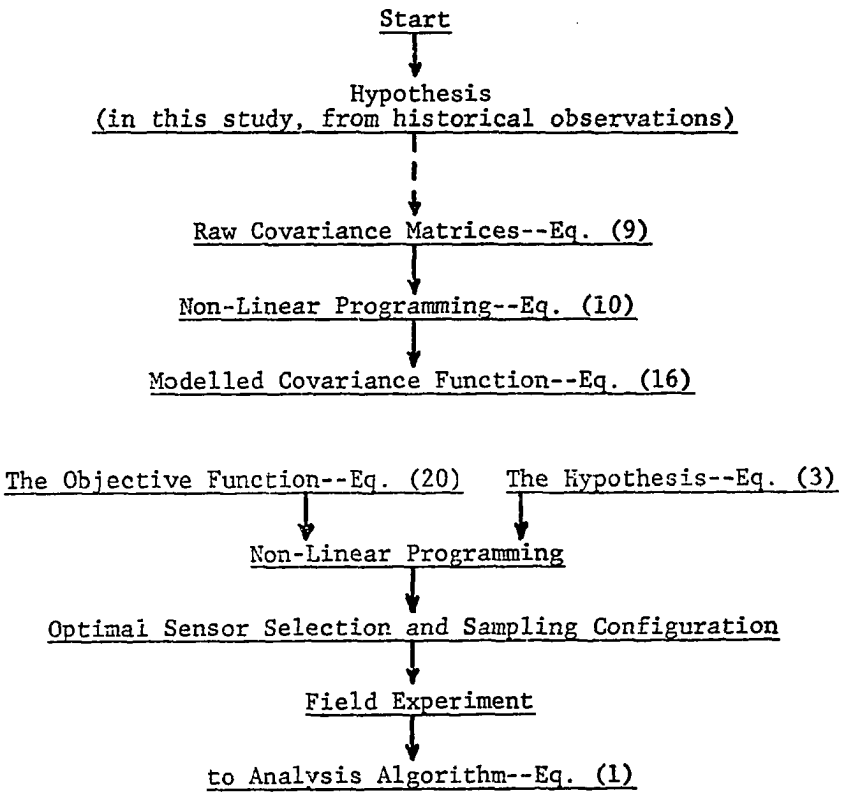


Fig. 35. A computational overview of the steps needed to evaluate optimal sampling requirements. Equations refer to those in the text.