76-24,359

BRADY, Patrick James, 1946-
MULTIVARIATE EXPERIMENTAL DESIGN IN METEOROLOGY.

The University of Oklahoma, Ph.D., 1976
Engineering, general

THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

MULTIVARIATE EXPERIMENTAL DESIGN IN METEOROLOGY

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

BY

PATRICK JAMES BRADY

Norman, Oklahoma

1976

MULTIVARIATE EXPERIMENTAL DESIGN

IN METEOROLOGY

APPROVED BY

Amos Eddy

Jay S. Fein

James A. Payne

DISSERTATION COMMITTEE

ABSTRACT

MULTIVARIATE EXPERIMENTAL DESIGN IN METEOROLOGY

An objective technique, which permits the analysis and investigation of multivariate as well as univariate data sets, is discussed. The technique necessitates the determination of the intrarelationships that exist within each single parameter data set, as well as the interrelationships that exist between these data sets. In general, these relationships are described in terms of anisotropic multidimensional correlation and cross-correlation functions which model the spatial (x,y,z) and temporal (t) structure of the phenomenon as reflected in the observations themselves.

The applicability of this technique is demonstrated using a multivariate data set comprised of surface and radar precipitation measurements. Derived spatial-temporal correlation and cross-correlation functions are presented; univariate and multivariate analyses of the observation set are compared; and the relative worth of different Z-R (radar reflectivity-rainfall) relationships is explored.

The combining of this analysis technique with a non-linear programming (NLP) algorithm for use as an experimental design tool is also discussed. A multidimensional model of the signal-plus-noise structure of a phenomenon of interest is sought. This model is then used by the objective analysis technique to determine the values of an objective function. The function in turn is minimized by an NLP algorithm and this minimization results in a station configuration for optimally sampling the phenomenon of interest. The potential of this design methodology is explored using the surface and radar precipitation measurements. Optimal spatial station configurations are presented and trade-offs between different sensor types are explored.

ACKNOWLEDGEMENTS

I would like to thank Dr. Amos Eddy for his creative supervision, enthusiasm, and constant interest in this research effort. The author's appreciation is also extended to Dr. Archie Kahan and Dr. Bernie Silverman of the Bureau of Reclamation, and Dr. Don Takeuchi of Meteorology Research, Inc. for providing the data used in this study. I am also grateful to the following members of my dissertation committee for their careful perusal of my work: Dr. Michael Devine, Dr. Jay Fein, Dr. James Payne, and Dr. Eugene Wilkins.

Mrs. Margaret Eddy's expert help in preparation of this manuscript was invaluable, and much thanks goes to her.

I would like to express my graditude to my wife, Michelle, for her patience and support during this period. It meant so much to the final accomplishment of this goal.

Lastly, acknowledgement is made to the National Science Foundation for funding this research.

TABLE OF CONTENTS

LIST OF TABLES

TABLE | Page

LIST OF ILLUSTRATIONS

FIGURE Page

FIGURE Page

FIGURE Page

FIGURE		Page

FIGURE Page

FIGURE		Page

FIGURE Page

FIGURE		Page

FIGURE | Page

FIGURE | Page

MULTIVARIATE EXPERIMENTAL DESIGN IN METEOROLOGY

CHAPTER I

INTRODUCTION

Since early in the 17th century, man has been attempting to measure various geophysical phenomena. During the last forty years, as technology has advanced, the availability of quality instrumentation, as well as the accuracy of measurements obtained, has largely improved. Development and implementation of sophisticated instrumentation packages, such as those found in satellites, on aircraft, and at radar sites, has greatly assisted the scientists' search for more knowledge. The quantification of geophysical phenomena is a vital step in the meteorologists' and/or hydrologists' continuing pursuit of a more complete understanding of his chosen field. The acquisition of pertinent observations, with respect to areas of investigation, is essential.

The field study or field experiment is the environment in which the sophisticated and expensive instrumentation is often used, in hopes of obtaining those vital, relevant, but often illusive observations. To hope for good results from such studies that expend hundreds of thousands of dollars and more, is no longer acceptable. The public's outcry over the waste of their tax monies for projects that are ill-conceived, ill-planned, and ill-managed will not lessen in the future. It would seem therefore that forthcoming studies in atmospheric and related sciences must be structured in such a way as to produce a benefit-cost relationship that is acceptable both to the scientist and to those members of society paying for the research. One area in which a minimum of expenditure, in terms of manpower and expense, could produce a maximum benefit, in terms of quality experimental results, is that portion of the field study concerned with the experiment-planning aspect of the research, sometimes referred to as experimental design.

In any field study, four steps are required.

1. The problem to be investigated must be defined in the most concrete and objective terms possible. This is crucial to the success of the study, for without definition a study has no purpose, and without purpose there is no need. (The investigator must formulate, in terms of the experimental design concept, his null hypothesis concerning the phenomenon to be investigated.)

2. A careful examination has to be made of the tools available for resolving the definable problem, and the best method of using those tools must be sought. This process is an essential (and often neglected) part of any study. The design of the study to be undertaken is equally as important as the problem definition and field deployment.

3. The study should be implemented to obtain those observations most likely to be of use in resolving the defined problem. An update of the study, using the design process in conjunction with incoming experimental results, is also usually warranted.

4. The resultant observations are analyzed, conclusions are drawn, and the initial hypothesis is either accepted, rejected, or refined.

Having defined a problem to be investigated, the researcher must plan the manner in which the investigation is to be carried out. It is in this design step of the field study that objectively-founded decisions should be the norm, but subjective decisions are often the rule. How does a scientist make the most efficient use of the instrumentation available to obtain a best set of observations for his study? Each instrument type has associated costs, noise characteristics, and data acquisition problems. Which ones are best for a particular experiment? How many are needed, as a minimum, to resolve the problem being investigated? What instrument deployment configuration will be most advantageous to the outcome of the study? Can interrelationships between various parameters be used to enhance the quality of the observation set obtained and/or to reduce the cost of the field operations? Is the study justifiable? That is, with the known instrument and given monetary restrictions, does the problem to be investigated have a good chance of being resolved?

During the past few years a number of numerical techniques have been developed which permit objective investigation of these types of questions. A mathematical clarification of these techniques as well as illustrative examples of their investigative potential are presented in the following chapters.

CHAPTER II

GENERAL DESIGN APPROACH

The experimental design approach to "optimizing" the return from a field observation study requires the following four processes. First, a statement is needed, as precise as possible, in some objective language, of the problem to be investigated, including all known parameters associated with the problem and some statement about the quality of estimation required for each parameter. Second, access is needed to the relevant information existent in the profession, as seen from the viewpoint of the user. Ideally this would include, but not be limited to, a mathematical model of the phenomenon of concern and/or mathematical models of the infrastructure of the phenomenon to be investigated. Third, a pre-experiment investigation of the phenomenon is required, using the tools of experimental design, in order to acquire sufficient information to answer those questions discussed in the introduction. Fourth, an active open-minded dialogue between the researcher, the experimental designer, and the system coordinator is necessary so as best to fulfill the needs of the user. Each of these steps is an essential link in the development of an acceptable experimental design. The research effort discussed herein, however, concerns itself with the "pre-experiment investigation" process of the design effort, and with those mathematical tools required in order to accomplish this task.

Two separate but fundamentally related approaches to the "pre-experiment investigation" process have been undertaken. The first approach was to develop techniques which would permit examination of previously collected observations in terms of current experiment objectives. This would allow earlier experimental results to be used in the design of forthcoming related investigations. Basically this approach necessitates the determination of the spatial-temporal signal-plus-noise structure (as reflected in the observation set) of the phenomenon of concern, and the comparison of univariate-multivariate analyses of that phenomenon. In Chapter IV these concepts are explained and their use as an investigative tool is explored in Chapter V.

The second approach to the experimental design problem was a logical extension of the techniques of Chapter IV. Those techniques were concerned with the analysis of already acquired data sets so as to investigate such things as signal/noise ratios, signal-plus-noise structures, and the interrelationships between various parameters of the phenomenon being studied. Having made those investigations, it would seem appropriate to use the information thus obtained, in the actual design of forthcoming experiments. By combining the analysis routine of Chapter IV with a nonlinear programming algorithm (NLP), the incorporation of results from past experiments into the design phase of related future investigations is possible. This design approach is discussed in Chapter VI and its potentialities, with respect to the problems of instrument number, instrument type, and instrument placement, are explored in Chapter VII.

CHAPTER III

DATA

The data used in this research consisted of precipitation observations which were obtained in the form of surface rainfall measurements and quantitative radar measurements of reflectivity. The rain gauges used were those in the Agriculture Research Service, Washita River Watershed Rain Gauge Network, which is located southwest of Norman, Oklahoma. The reflectivity information was obtained from the WSR-57 radar located at the National Severe Storms Laboratory in Norman. Figure 1 shows the respective radar-rain gauge positions.

Five minute rainfall accumulations, expressed in mm, were procured for the following four time periods:

1) 21 May, 1974, 0725-0850 local time,
2) 25 May, 1974, 0755-0905 local time,
3) 25 May, 1974, 0935-1010 local time,
4) 25 May, 1974, 1145-1310 local time.

The surface measurements were obtained using breakpoint (accumulated amounts at times of inflection) data from the weighing bucket gauges in the network for the appropriate times. The radar reflectivity was converted into rainfall

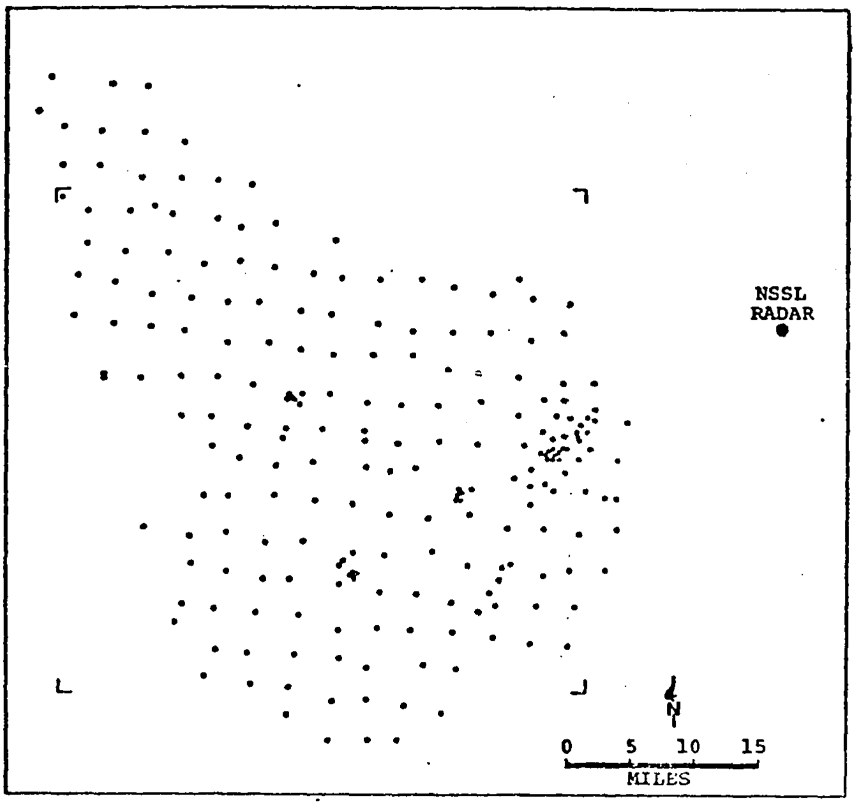

Fig. 1: The relative rainguage-radar positions.

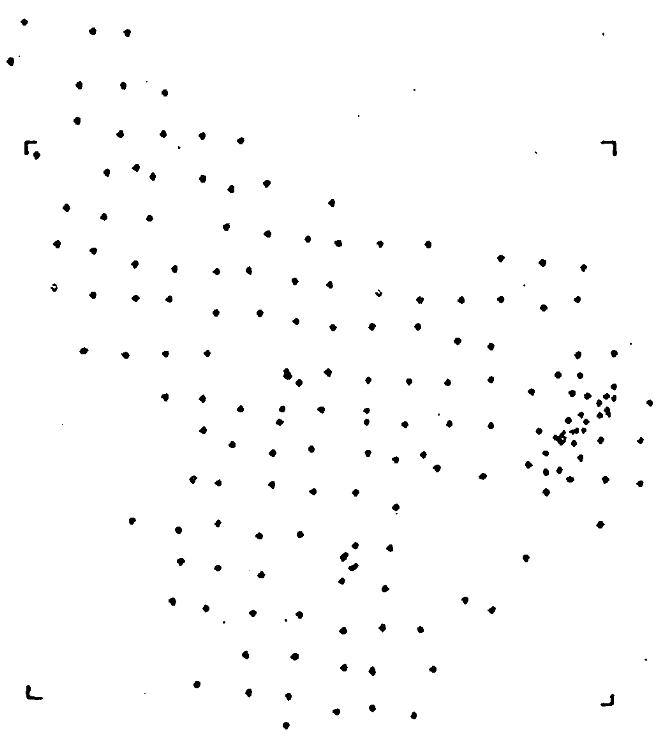

Fig. 2: The 181 sensor positions for the data on the 21st.

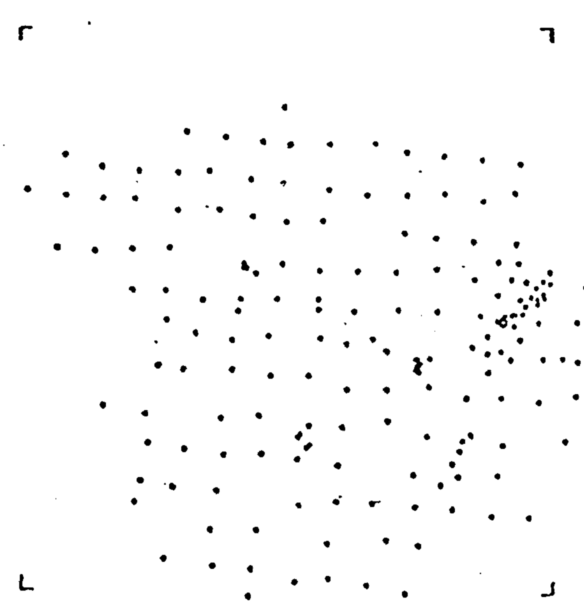

Fig. 3: The 187 sensor positions for the data on the 25th.

amounts using the Marshall-Palmer Z-R relationship, $R=.0365Z^{.625}$, as well as two empirically derived Z-R relationships; $R=.0131Z^{.687}$ for the data of the 21st and $R=.0148Z^{.696}$ for the data of the 25th. These latter relationships were determined by using an aircraft mounted Particle Measuring Systems Precipitation Spectrometer Probe to sample the precipitation patterns on the dates mentioned.

The network initially consisted of approximately 225 gauges. However, not all of these gauges provided rainfall measurements for the dates specified. Because of hardware and other problems, the network consisted of 181 gauges on the 21st (Figure 2) and 187 gauges on the 25th (Figure 3). The converted reflectivity amounts from the 181 and 187 radar bins, which were horizontally coincident with the 181 and 187 surface gauges, were included in the data sample.

The radar measurements were made at Plan Position Indicator (PPI) scan rates of 3 rpm at an elevation angle of 1°. The resultant data represented measurements at radial sectors of 2° and at gate lengths of about 1 km. An absence of reflectivity within a bin converted into 0 measured rainfall. A more complete description of the data can be found in Takeuchi, et al. (1974).

CHAPTER IV

OBJECTIVE ANALYSIS

In the planning and development stage of a field study, it is necessary to have some means of investigating the characteristics of the phenomenon of concern, such as size, shape, intensity, and duration, as well as having the ability to determine the effect of instrument induced noise on the quality of the measurements taken. It is also necessary to have some means of objectively measuring the worth of proposed instrument configurations, in terms of the instrumentation type, the instrument number, the sampling rate, and the sampling positions. An objective analysis technique fulfills these needs, and is an essential part of the experimental design concept.

A. Analysis Model

Objective analysis is the weighting of observations in space and time in order to estimate a parameter field at some specified point (normally a lattice) in space and time. Much effort has been expended during the past twenty-five years, in developing and investigating objective analysis techniques.

One of the earlier analysis schemes was Cressman's (1959), which is the basis for many of the numerical weather prediction (NWP) models of today. His method extended the work done by Bergthorssen and Doos (1955) and essentially applies a series of corrections to a first guess field during a series of scans of that field, where the corrections are based on the distance between the grid point and the observation.

Different approaches to the problem were proposed by Sasaki (1958), who suggested a variational analysis method which achieves dynamic consistency, and Gandin (1963) and Eddy (1963), who formulated the use of space-time distance correlations to analyze atmospheric phenomena.

The fitting of a polynomial to observed data using least-squares, for the purpose of objectively analyzing the data field at non-observation points, was first proposed by Panofsky (1949). The more current work using a similar analysis approach is reflected in the papers by Endlich and Mancuso (1968), and Shapiro and Hastings (1973).

Epstein and Pitcher (1972) adopted a Bayesian approach, in conjunction with a stochastic, dynamic prediction model, to the problem of analysis, while Stephens and Polan (1971) have investigated the problem of spectrum degeneration induced by objective analysis of the observed parameter field.

Of the analysis schemes available, the technique of Eddy (1973) was adopted here as the primary tool in the development of a systems package for use in field study design. This technique is an extension of classical multivariate multiple linear regression. The extension consists of modeling the basic signal-plus-noise structure that is representative of the information contained in the time series data set being analyzed, and then using that modeled "structure function" in the determination of the regression weights. Employing this approach, it is possible to filter a parameter field at positions where data is available, and also to analyze the parameter field at positions in space and time where no observations have been made.

Consider the linear regression model for the univariate case

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

where Y = a vector of predictands (Nx1)

X = a matrix of predictors (NxM)

β = the vector of regression weights (Mx1)

ε = the noise vector (Nx1).

In classical regression, if

$E[\varepsilon]=0$, $VAR[\varepsilon]=V\sigma^2$, and $\varepsilon \approx N(0,\sigma^2 V)$ where $V=E[\varepsilon\varepsilon^t]$

then the regression weights determined from the data are

$$\hat{\beta} = [X^t V^{-1} X]^{-1} [X^t V^{-1} Y], \tag{2}$$

and the objective analysis may be accomplished by

$$\hat{Y} = X\hat{\beta}. \tag{3}$$

Looking at the k^{th}, ℓ^{th} element of equation (2) we see that

$$[X^tV^{-1}X]_{k\ell} = \sum_{ij}^{NN} X_{ik}X_{j\ell}V_{ij}^{-1}, \text{ and} \tag{4}$$

$$[X^tV^{-1}Y]_k = \sum_{ij}^{NN} X_{ik}Y_jV_{ij}^{-1}. \tag{5}$$

The signal covariance $X_{ik}X_{j\ell}$ and $X_{ik}Y_j$, as well as the noise anatomy V_{ij}, can be computed from the previously modeled signal-plus-noise structure as reflected in the data set, within the limits of the model itself. The covariance function COV(XX) is usually sufficient for determining the regression weights of equation (2). As the noise structure reflected in the observations becomes more complex, the signal-plus-noise model could become less representative of the actual realization. In order to alleviate this problem, the noise matrix V can be modeled separately (Best, 1973).

Now consider the problem of using the linear regression model in analyzing $\hat{Y}$ of equation (3) when the predictor set is multivariate, as would be the case when different sensor types have been employed in measuring the phenomenon. For the two-parameter data set previously described, equation (3) can be rewritten as

$$\hat{Y} = X_1\hat{\beta}_1 + X_2\hat{\beta}_2, \tag{6}$$

where X_1 contains the rain gauge measurements and X_2 contains the reflectivity estimates.

Looking at the k^{th}, ℓ^{th} element of equation (2) in terms of this two-parameter data set we see that the signal covariance quantities $X_{ij}X_{j\ell}$ and $X_{ik}Y_j$ must now be expressed as

$$X_{1ik}X_{1j\ell} \text{ or } X_{1ik}X_{2j\ell} \text{ or } X_{2ik}X_{1j\ell} \text{ or } X_{2ik}X_{2j\ell} \text{ and}$$

$$X_{1ik}Y_j \text{ or } X_{2ik}Y_j \text{ ,}$$

depending on which parameter mix (gauge measurements and/or reflectivity estimates) can be best used to analyze $\hat{Y}$ in equation (6). In general for the n-parameter data set, n^2 covariance and cross-covariance functions are needed to determine any of the k^{th}, ℓ^{th} elements in equations (4) and (5). For the two-parameter precipitation data sets, two co-variance and two cross-covariance functions were needed:

$$\mathrm{COV}(X_1X_1),\ \mathrm{COV}(X_2X_2),\ \mathrm{COV}(X_1X_2),\ \mathrm{COV}(X_2X_1).$$

This analysis technique has a number of features which make it especially suitable for use as an experimental design tool.

1. The analyses produced are best with respect to minimizing the error sum of squares, and are maximum likelihood unbiased estimates if the errors are assumed to be independent and normally distributed with a mean of zero and variance σ^2. In many physical situations this assumption is quite sensible.

2. This technique allows investigation of the effect of instrument-induced noise on the analyses produced, which in turn can be used in the determination of those instruments best suited for a particular experiment. This also allows for investigation of trade offs that exist between the use of a large number of high noise instruments, which are usually less expensive per unit, and the use of a small number of low noise instruments, which usually cost more per unit.

3. Analysis of a parameter set, using the intrarelationships within that set, and the interrelationships between two comparable parameter sets can be accomplished. The extension of the analysis technique to be truly multivariate was a major area of this research.

4. Modeled estimates are possible, based on the structure function supplied, of the amount of variance in the original data sample that has been explained by a particular sampling placement in space and time, with respect to some specified grid lattice. This allows for the possibility of shifting sampling patterns in space and time so as to explain the greatest percentage of the original data variance. The "shifting" is accomplished by using the analysis technique in conjunction with a non-linear programming algorithm. It is this aspect of the analysis technique that permits the determination of the instrument locations, the sampling rate, and the instrument number best suited for a particular field study. The inclusion of the multivariate analysis technique, so as to be able to estimate the worth (if any) of interrelationships that exist between different parameter sets, was another major research area.

In addition to these features, a number of pragmatic reasons exist for using this particular analysis scheme.

During the past few years much effort has gone into the development of the mathematical, engineering, and programming concepts needed by this analysis technique. This effort had been directed toward the analysis of single parameter data sets (Eddy, 1975; Brady, 1975a). Nevertheless, the expertise acquired and the knowledge gained in the development of that analysis system package was beneficial in the development of a similar analysis system package for the multivariate case.

Also, much of the work accomplished in the field of experimental design as relates to the atmospheric sciences (Eddy, 1974; Kays, 1974; Yerg, 1973) makes specific use of the various analysis features discussed previously. As all of these features were not inherent in any other analysis model, the Eddy analysis scheme was vital to the experimental design process discussed herein.

B. Signal-Plus-Noise Model

It was previously stated that a modeled structure function, reflective of the information contained in the observation set of the phenomena being investigated, was needed in order to produce analyses from that observation set. The structure function idea has been used by different authors to express different concepts. For example, one definition of a structure function, in relation to atmospheric phenomena, was proposed by Gandin (1963). His interest was in measuring the structure of atmospheric phenomena in terms of the non-information in an observation set (noise) as a function of separation distance. Specifically, his structure function is defined to be the mean square of the difference in the values of the deviation of the meteorological elements from their respective means, for all observation-pair combinations in the parameter field. If C is the autocovariance function of an observation set, which is a measure of the information or degree of signal in the atmospheric phenomena, and homogeneity and isotrophy are assumed, then the structure function is $2(\sigma^2\text{-C.})$.

For the purpose of this discussion, the structure function of an observation set, which is subsequently used for analysis of that observation set, is expressed in terms of the spatial-temporal correlation of the data values of that set, determined for various lag distances in space and time. The parameter field is assumed to be stationary with respect to the structure function; that is points of equal distance-separation and the same relative direction have equal correlation. The structure function is not assumed to be isotropic that is direction of separation influences the point-pair correlation coefficient value.

Product-moment correlations have been used for more than half a century to study various geophysical phenomena. Until recently, quite a bit of the work done related to the science of hydrology. Indeed, since the first computations by Sir Ronald Fisher in 1922, correlation analysis of precipitation measurements has been used in the investigation of rainfall patterns. However, in all of the work accomplished using this technique, correlation analysis has been an end unto itself. That is, correlation coefficient fields for different separation distances have been determined, those fields have been studied to discern various characteristics of the geophysical phenomenon being investigated, and the results of the studies have been published with information about pattern shapes, sizes, etc. No doubt, for a hydrologist or others interested in such things, the information acquired from this type of analysis is worthwhile.

In relation to the Eddy analysis technique, however, the use of correlation analysis is somewhat different. It is not considered an end unto itself. Rather, it is but one step in the determination of an objective analysis of some atmospheric parameter. It is true, however, that the more reflective the structure function is of the physical and statistical characteristics of the phenomena being analyzed, the better the resultant analysis will be. Therefore, the objective is to use the appropriate knowledge available from the correlation analysis studies to determine the best model for the structure function, and then determine the parameters of the model appropriate to the entity being analyzed. With these ideas in mind, the basic concepts of correlation analysis are discussed, and the adaptations of these concepts for use in experimental design are presented.

In general, a mutual relationship exists between data values of an observation set. The correlation coefficient is a measure of the strength and direction of this relationship. More specifically, the correlation coefficient is a measure of the relative precision with which one variate can be estimated from corresponding values of another variate or variates. If two observations are "close" to one another in space and time, their data values would tend to be positively correlated. As the separation distance in space and time increases between observations, their correlation would be expected to decrease; and as the separation distance decreases, their correlation would be expected to increase. This is true for any regression model employed; linear, quadratic, logarithmic, etc. However, since almost all the work done in this area assumes a linear regression model, the correlation coefficient is usually a measure of the linear relationship that exists between two variates, normally displaced in space and/or time.

As previously mentioned, almost all of the early correlation pattern analysis was confined to the study of precipitation, as was that initial paper by Fisher-Mackenzie (1922). Besides discussing the correlation patterns of precipitation fields, Fisher-Mackenzie also expressed some interest in determining the functional relationship between the correlation coefficient and separation distance. Boyd (1939) expanded this work by introducing additional parameters into Fisher's original formulation. Unfortunately, neither Fisher's nor Boyd's formulas fit the data very well. In two papers on rainfall correlations in South Australia (Stenhouse and Cornish, 1958; Cornish, Hill, and Evans, 1961) the same formula was used without reference to Boyd's work and without much success.

Meanwhile, in the United States, rainfall correlations were being investigated by Horton (1923), Miller (1931), Bernard (1943), and Oltman and Tracy (1951). McDonald and Green (1960) were interested in the value of Spearman's rank-correlation coefficient in meteorological studies, while the work of Court (1958) and Buell (1972) accomplished correlation analysis of other meteorological parameters such as wind, pressure, and temperature. Even so, there exist problems with this line of study. For example, the determination of a functional form that adequately reflects the correlation pattern characteristics has been of great concern. The need for some such mathematical formula can be readily developed.

Consider the fact that the correlation fields generated from a single storm system (in the case of precipitation) for the number of spatial and temporal lags sufficient to discern the storm structure in terms of the correlation coefficient can be quite large; typically 2000-3000 discrete values. In determining an objective analysis of these storm systems the correlation coefficient for tens of thousands of data pair combinations is needed. Compound this with the fact that some type of scheme is necessary to interpolate the discrete lag correlation values to the continuous range of separation distance combinations possible between observation pairs, and the need for a functional model of the correlation structure becomes more apparent.

To be able to dispense with the data management problem of 2000-3000 data values is convenient; to be able to dispense with the interpolation problem, in terms of computational time required to perform an analysis, is paramount.

Also, the influence of small scale perturbations, as well as inadequate data samples, in determining correlation coefficients for some separation distances, can be a problem. If the value of the "raw" coefficient (which is not constrained by modeling) for certain lags does not fit the general pattern observed in the correlation field, then bias in the objective analysis will be the likely result. This bias is introduced by assigning abnormally high (or low) weights to predictors based on the correlation coefficient values found. A smoothing of these effects is needed in order to obtain consistent analyses. In fitting a functional model to the coefficient fields, and weighting this fit by the number of observation pairs used to determine the corresponding correlation values, the effect of unrepresentative perturbations can be reduced, if not eliminated.

And, finally, it is somewhat difficult to conceptualize the physical structure from the data set, by pondering over raw correlation coefficient matrices. On the other hand, the parameters of an appropriately chosen and fitted functional form can quite readily reflect the relevant physical and statistical information about the system being studied. This is a tremendous convenience, especially when investigating the station configuration problem.

A number of researchers have proposed formulae to reflect adequately the relationship between the correlation coefficient value and the observation-pair separation distance (Caffey, 1965; Hutchinson, 1969, 1970, 1971; Anderson, 1970; Hendrick and Comer, 1970; Steinitz, Huss, Manes, Sinai and Alerpson, 1971; Stol, 1972; Gringorten, 1973; Kraemer, 1974; Longley, 1974). Unfortunately, none of the suggested formulae take into account all of the characteristics required to represent the correlation fields.

It is essential that the function be four dimensional to delineate the spatially (x,y,z) and temporally (t) determinable correlation structure. Since the probability is great that the correlation patterns will be anisotropic, the mathematical model must express the correlation coefficient as a function, not only of distance, but also direction. And, since the values of the correlation coefficients tend to become negative as the separation distance increases, provision for this eventuality must be made.

A major hurdle in the use of the analysis technique employed, was the search for and determination of an adequate correlation model function. This concerned not only the actual model formulation, but also the problem of transforming four-dimensional observation values into anisotropic four-dimensional correlation coefficient fields, as well as the fitting of the model function to those fields. These problems were compounded by the fact that the observations are normally obtained from stations whose spatial placement is not uniform.

Two algorithms were developed to generate the needed correlation coefficient fields, the first for non-uniform station placements, as is the case with normal field observations, the second for symmetrically generated data arrays, which is the type of output expected from an objective analysis program or atmospheric numerical simulation model. The use of both algorithms allows for the initial estimate of the correlation structure of the observation field, as well as the determination of the correlation structure from the objectively analyzed grid values. After fitting the model function to these fields, an estimate of the ability of the analysis technique to retrieve the original data structure is possible. This provides some measure of the distortion introduced into an analyzed field by the analysis technique itself. The better the technique used, the less will be the amount of distortion detected.

The second problem, of fitting the model function to the generated correlation fields, was solved by employing an NLP algorithm. Let $r_i(u_x,u_y,u_z,u_t,)$ be the correlation coefficient for spatial lags u_x,u_y,u_z and temporal lag u_t, and let $n_i(u_x,u_y,u_z,u_t)$ be the number of observation pairs used in determining that value. Then the problem of fitting the model function $f_i(u_x,u_y,u_z,u_t)$ to each of the M lag-categories, reduces to the problem of minimizing

$$Q = \sum_{i}^{M} n_i(u_x,u_y,u_z,u_t)(r_i(u_x,u_y,u_z,u_t)-f_i(u_x,u_y,u_z,u_t))^2 \quad (7)$$

for all discrete lag values of u_x, u_y, u_z, and u_t. The NLP algorithm adjusts the parameters of f until the value of Q, the sum of the square of the errors, has been sufficiently reduced. The present form of the correlation model is given below.

Letting

$$g(u_x,u_y,u_z) = a_1(u_x^+/d) + a_2(u_y^+/d) + a_3(u_x^-/d) + a_4(u_y^-/d) + a_5(u_z^+/d) + a_6(u_z^-/d),$$

we obtain

$$f(u_x,u_y,u_z,u_t) = g(u_x,u_y,u_z) + (a_7 - g(u_x,u_y,u_z)) \exp\{-((u_x^+/a_8)^2 + (u_y^+/a_9)^2 + (u_x^-/a_{10})^2 + (u_y^-/a_{11})^2 + (u_z^+/a_{12})^2 + (u_z^-/a_{13})^2 + (u_t/a_{14})^2)\} \quad (8)$$

where $a_1, \ldots, a_6, a_8, \ldots, a_{13}$ permit the structure function to be anisotropic; $a_1, \ldots, a_6$ allow the function to become negative in the respective spatial directions; $a_8, \ldots, a_{13}$ determine the function's spatial shape and size; a_{14} determines the function's temporal shape and size; a_7 is the correlation value for a spatial-temporal lag of zero (See Chapter V, section B for a more detailed discussion of the a_7 parameter); and $d = |u_x|+|u_y|+|u_z|$. Since the particular data set discussed herein exhibits no vertical distribution, the u_z parameters are zero, and the appropriate form of the function is $f(u_x,u_y,u_t)$.

C. Nonlinear Programming (NLP) Algorithm

As has been stated, nonlinear programming plays an important role in the development of an experimental design systems package. Not only is it needed in the parameterization of the structure function, it is also vital for the investigation of the sensor configuration problem. A number of NLP algorithms exist, each appropriate to various nonlinear programming applications. The algorithm chosen for the particular applications previously described, was developed by Himmelblau (1972), and an updated version of this technique (Himmelblau, 1974) has been implemented.

The basic idea of the algorithm was described in a paper by Paviani and Himmelblau (1969), which extended and generalized the work of Nelder and Mead (1964). Likewise, the method of Nelder and Mead was an extension and generalization of the work accomplished by Spendley, Hext and Himsworth (1962). The algorithm is a direct search technique which improves the value of the objective function by using information provided by feasible points, as well as certain nonfeasible points (termed 'near feasible' points). The near-feasibility limits are gradually made more restrictive as the search proceeds toward the solution of the programming problem, until in the limit only feasible points remain in the solution space.

The algorithm utilizes a penalty function technique whereby the constrained problem is transformed into a sequence of unconstrained problems by employing a modified definition of the objective function. For the general case, the NLP problem is expressible as

$$\begin{aligned} &\text{minimize } f(\xi) && , \xi \in E^n \\ &\text{subject to } h_i(\xi) = 0 && , i=1,2, \ldots, m \\ &\qquad\qquad g_i(\xi) \geq 0 && , i=m+1,m+2, \ldots, \rho \end{aligned}$$

where the objective function $f(\xi)$ and the constraint functions $h(\xi)$ and $g(\xi)$ may be linear and/or nonlinear. Utilizing this notation, the definition of the modified objective function becomes

$$P(\xi,t)=\min[0,(t-f(\xi))]^2+\sum_{i=1}^{m} h_i^{\,2}(\xi)+\sum_{i=m+1}^{\rho} \min(0,g_i(\xi))^2 \qquad (9)$$

where $\rho(\xi,t)$ is the Moving Exterior Truncations (MET) penalty function, t is a truncation level, and the constraints representing variable bounds have been included as inequality constraints in the last term of equation (9). The basic procedure for utilizing the MET penalty function is to minimize equation (9) for a monotonic increasing sequence of truncation levels $\{t_k\}$ converging to f*, the value of the objective function at the constrained minimum, so that

$$t_{k-1} < t_k \leq f^*$$

where the subscript k denotes the k^{th} unconstrained minimization. A new technique for generating the sequence of truncation levels was incorporated into the algorithm in order to give a faster rate of convergence than previous methods utilizing the same penalty function.

In order to find the initial truncation level t_1, which is not known a priori, the parametric quadratic loss penalty function

$$P_i(\xi,r)=f(\xi)+\frac{1}{r}\sum_{i=1}^{m} h_i^2(\xi)+\frac{1}{r}\sum_{i=m+1}^{\rho} \min[0,g_i(\xi)]^2 \qquad (10)$$

is minimized on the first stage, where r is a weighting parameter. Because a point minimizing equation (10) also minimizes equation (9) for an appropriate value of t, the initial truncation level is found from

$$t_1 = f[\xi'(r_1)] - \frac{1}{2}r \qquad (11)$$

where $\xi'(r_1)$ denotes the point minimizing equation (10) for the parameter value r_1.

The method of Fletcher (1970) was used for the unconstrained minimizations of the penalty function. Fletcher's method is a variable metric method modified to eliminate unidimensional searches. His algorithm was modified to accept difference approximations for derivatives (Stewart, 1967) so that analytical derivatives need not be furnished in order to use the NLP algorithm.

The criteria which must be satisfied before final convergence to the problem solution is assumed are the following:

$$f[\xi'(t_k)] - t_k \leq \phi \quad (12)$$

$$|h_i[\xi'(t_k)]| \leq \phi,\ i=1,\ldots,m \quad (13)$$

$$g_i[\xi'(t_k)] \geq -\phi,\ i=m+1,\ \ldots,\ \rho \quad (14)$$

where ϕ is a small positive number and $\xi'(t_k)$ denotes the point minimizing the MET penalty function for the truncation level t_k.

Stocker (1969) made a comparison of five search methods, and the results appeared in a text by Beveridge and Schechter (1970). The five search codes tested were

1. A modification due to Paviani (1969) of the unconstrained search technique of Nelder and Mead, which combined all restrictions into a single tolerance.

2. A modification due to Barnes and Himmelblau (1966) of the basic method proposed by DiBella and Stevens (1965).

3. A modification of the Rosenbrock method (1960) to allow for constraints.

4. The method of Fiacco and McCormick (1964) using penalty functions.

5. The program developed by H.V. Smith and distributed through the IBM Share General Program Library.

The conclusion was that the method of Paviani was "the most reliable direct-search method" tested. This modification by Paviani formed the basis for the subsequent algorithms of Himmelblau (1972; 1974).

There were other reasons for the selection of this particular NLP code. For example, since the objective function of both problem applications mentioned is not expressible in a closed form, it was necessary to obtain an algorithm which allowed for this possibility. Also it was desirable to have the ability to include inequality as well as equality constraints in the problem statement. These reasons, as well as the fact that the algorithm was readily available as a computer code, influenced the choice of this NLP technique.

Finally, the previous work of Yerg (1973) and Eddy (1974) demonstrated the adaptability of this algorithm, and the results of that work clearly showed its potential.

It should be noted, that there exist specific NLP algorithms for solving least-squares problems similar to the correlation-model-fitting problem previously discussed. The implementation and subsequent testing of such an algorithm was not accomplished. It was decided that since the least-squares problem need be solved only once for each phenomenon analyzed, the "cost" of acquiring such a specialized technique, in terms of time required and effort expended, would exceed the advantages of having same. As Himmelblau's algorithm (1974) was already available, and as it could satisfactorily solve these types of problems, it was used in the fitting of the structure model to the correlation data.

CHAPTER V

ANALYSIS RESULTS

Objective analyses were obtained using the previously described Z-R relationships in three different data configurations. Analyses of the surface precipitation fields were accomplished by

1) using only the surface-gauge measurements,
2) using only the converted-reflectivity estimates, and then
3) using the multivariate predictor set of surface-gauge measurements and reflectivity estimates.

In order to compare objectively the results of these different analyses, the following criteria were used. Denoting the observed surface rainfall amount at the ith station as y_i, the analysis discrepancy for that station was determined by

$$e_i = y_i - \hat{y}_i$$

where $\hat{y}_i$ was the analyzed value at the ith station. The variance of these discrepancies for all stations, written

$$\sigma_e^2 = \sum_i^N (e_i - \bar{e})^2, \text{ where } \bar{e} = (\sum_i^N e_i)/N,$$

was used to compare the effectiveness of the three data configurations for each analysis time. For the May 21st data set N=181, and for the May 25th data sets N=187.

In determining $\hat{y}_i$, two procedures were used. Procedure 1 was to withhold from the predictor set the observed value at the ith station for the time at which the analysis was accomplished. For example, the May 21st data set had 18 time series observations for each of the 181 surface gauges, and for the ith station, the vector of observations could be written $Y_i = [y_i(0725), y_i(0730), y_i(0735), \ldots, y_i(0845), y_i(0850)]$. When the analysis at the ith station was then required for some specific time, say 0735, the observation at that station for the time specified, in this case $y_i(0735)$, was withheld from the set of possible predictors. This procedure permitted a comparison of the worth of the different analysis methods per se; i.e., a comparison of the univariate vs. the multivariate analysis approach.

Procedure 2 was a slight modification of the first. Instead of withholding just the surface observation at the <u>time</u> of the analysis, the <u>entire</u> time series observation set was withheld for the station being analyzed. As an example, when accomplishing the analysis at the ith station for the May 21st data set, the entire vector of 18 observations (Y_i) was withheld from the predictor set. By using this second procedure, a comparison could be made of the worth of the different data configurations for analyzing surface precipitation.

In all the results presented, it is assumed that the best analysis was the one which produced the smaller value of σ_e^2. In other words, a smaller variance of the computed error field implies that the analysis technique used and/or the data configuration employed made it possible to better recover the observed precipitation measurement values. The rest of this chapter presents the results obtained for the four storm systems analyzed.

A. 21 May, 1974, 0725-0850

The deduced correlation and cross-correlation model parameters for this particular precipitation data set are shown in Table 1. Parameters A8, A9, A10, and A11 are in

	POSITIVE X		POSITIVE Y		NEGATIVE X		NEGATIVE Y		TIME	
	A1	A8	A2	A9	A3	A10	A4	A11	A14	A7
CORR{X_1X_1}	.05	9.3	-.14	10.6	.09	10.4	-.17	8.6	34.0	.35
CORR(X_2X_2)	-.05	10.1	-.03	8.5	.14	8.4	-.02	4.7	18.1	.58
CORR(X_1X_2)	.02	11.3	-.03	10.5	.06	7.9	-.13	5.8	24.5	.39
CORR(X_2X_1)	-.01	7.1	-.13	9.7	.09	14.2	-.03	8.7	23.3	.35
CORR[X_2X_2]	-.05	12.2	-.07	9.0	.09	11.6	.04	4.2	19.2	.53
CORR[X_1X_2]	.05	12.1	-.05	10.5	.03	9.5	-.10	5.2	26.5	.38
CORR[X_2X_1]	-.05	-6.8	-.11	8.5	.06	16.1	-.01	7.5	25.0	.34

TABLE 1.--The signal-plus-noise model parameters for 21 May, 0725-0850. The parenthesized model designation, CORR(X_iX_j), denotes the parameters obtained from the Marshall-Palmer Z-R conversion, while the bracketed model designation, CORR[X_iX_j], denotes the parameters obtained from the empirically derived Z-R conversion.

nautical miles, while parameter A14 is in minutes. Figures 4 and 5 visualize the anisotropic CORR(X_2X_2) model for a temporal lag of 0 minutes and a temporal lag of 10 minutes, respectively. As the spatial and/or temporal distance in-

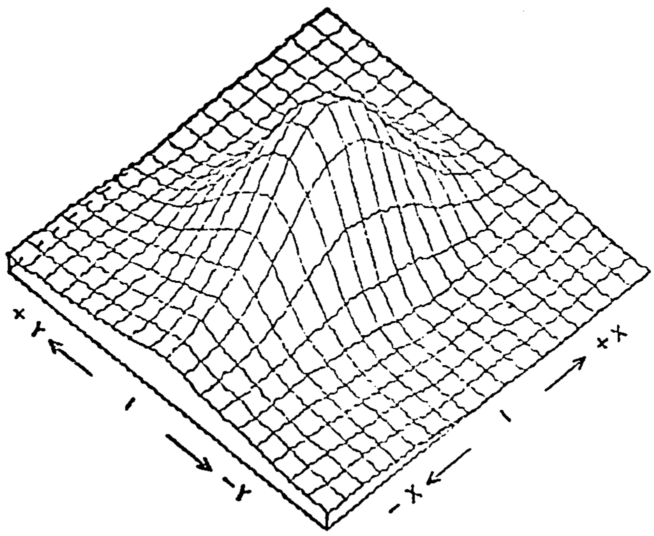

Fig. 4: The May 21st CORR(X_2X_2) model for u_t = 0 min. Model values range from 0.0-0.578.

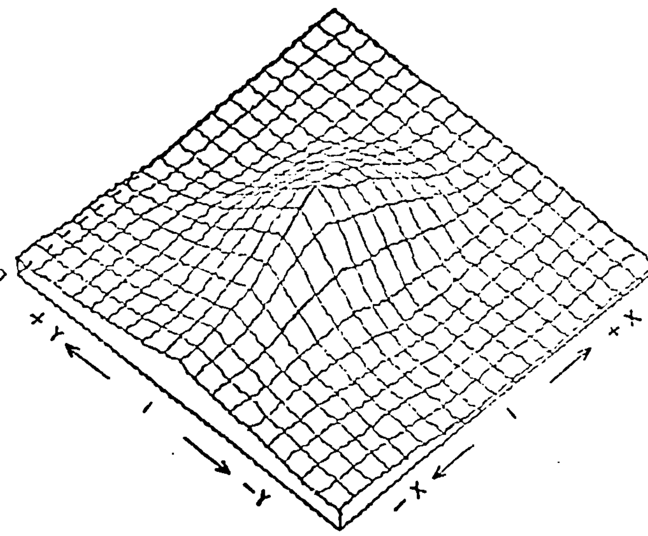

Fig. 5: The May 21st CORR(X_2X_2) model for u_t = 10 min. Model values range from 0.0-0.426.

creases, the correlation between observation pairs decreases, as illustrated in these two figures. The spacing between grid lines corresponds to a horizontal distance of 2.0 nautical miles.

Grid analyses for the entire May 21st storm are presented in Figures 6 thru 23 in order to show the type of precipitation pattern of that date. These analyses were produced using the Marshall-Palmer derived multivariate data configuration as were all the grid analyses presented in this chapter. In general the rainfall rate for this data set was between 0.2 and 1.3 inches/hour, with the higher rates

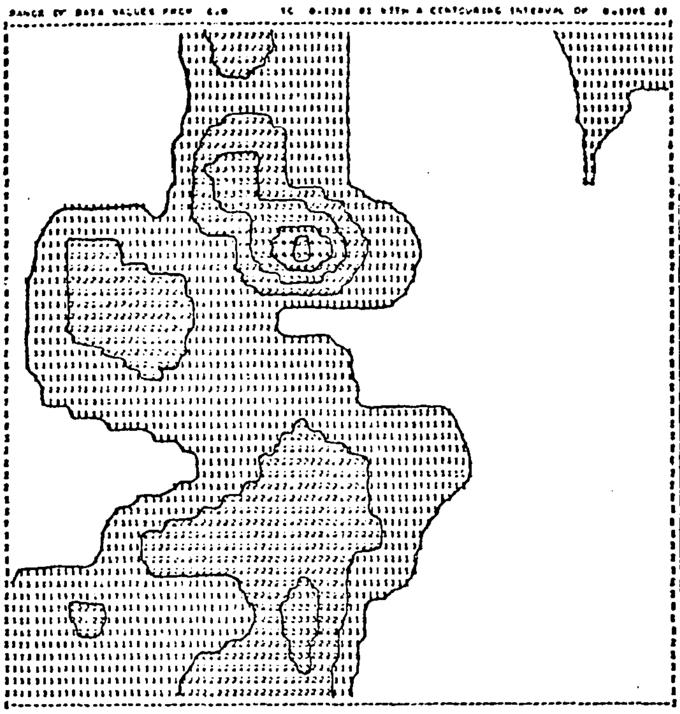

Fig. 6 : 0725 - 21 May.

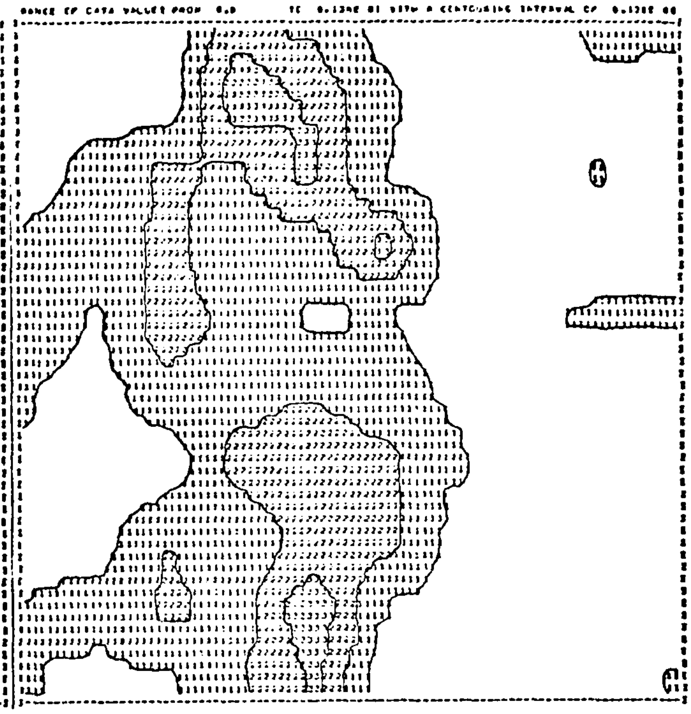

Fig. 7 : 0730 - 21 May.

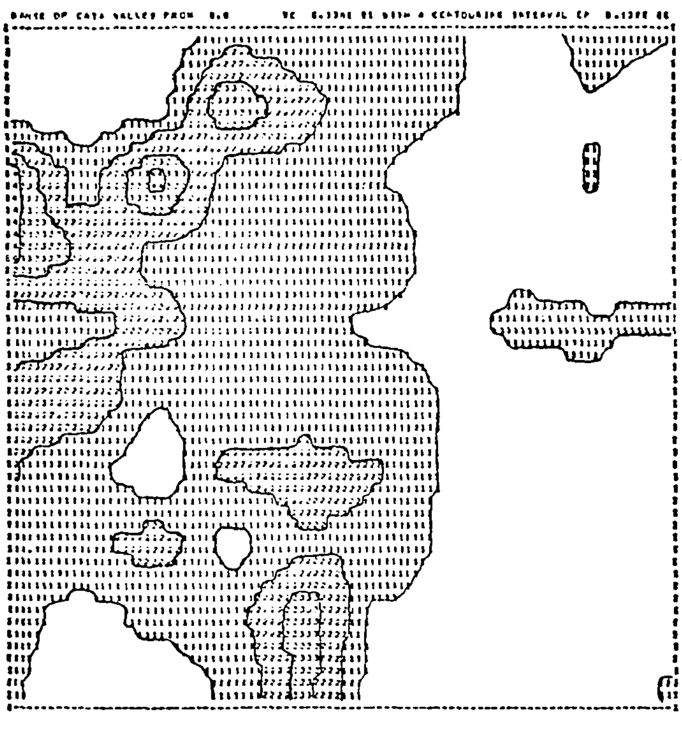

Fig. 8 : 0735 - 21 May.

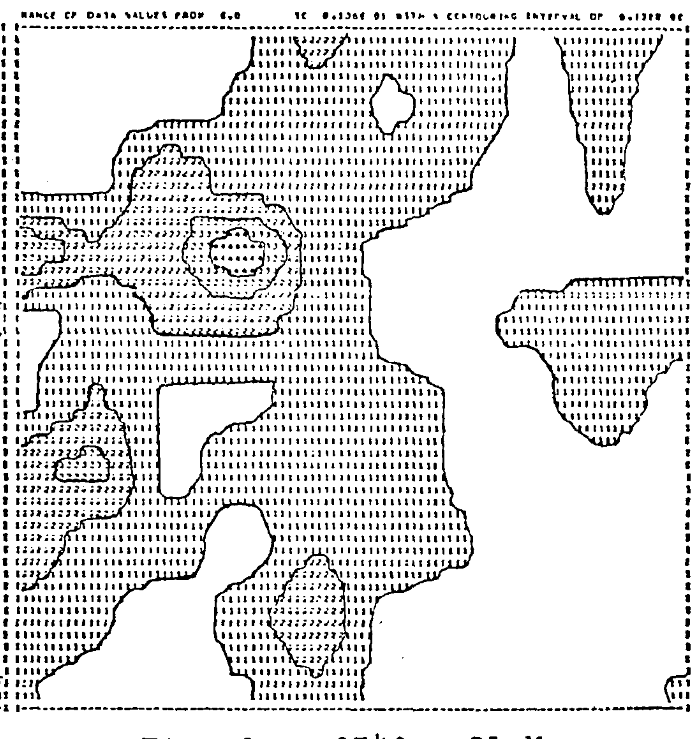

Fig. 9 : 0740 - 21 May.

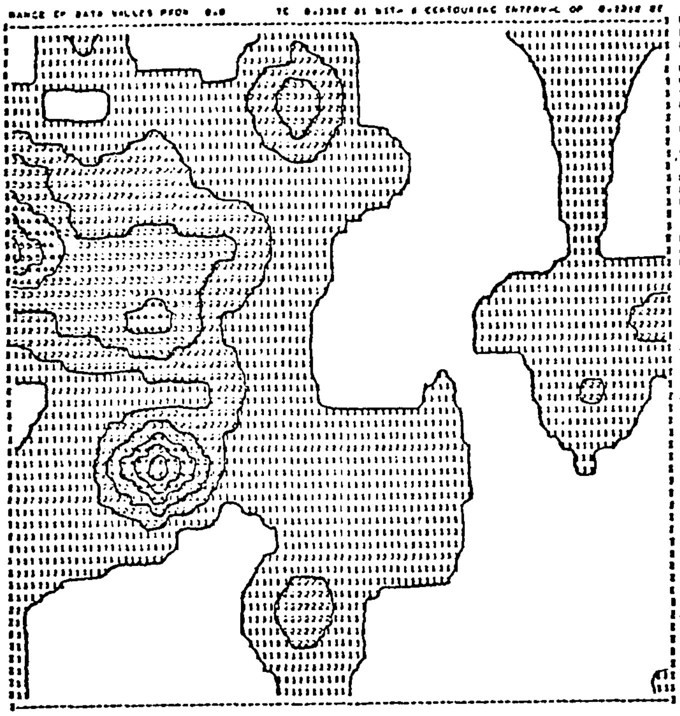

Fig. 10: 0745 - 21 May.

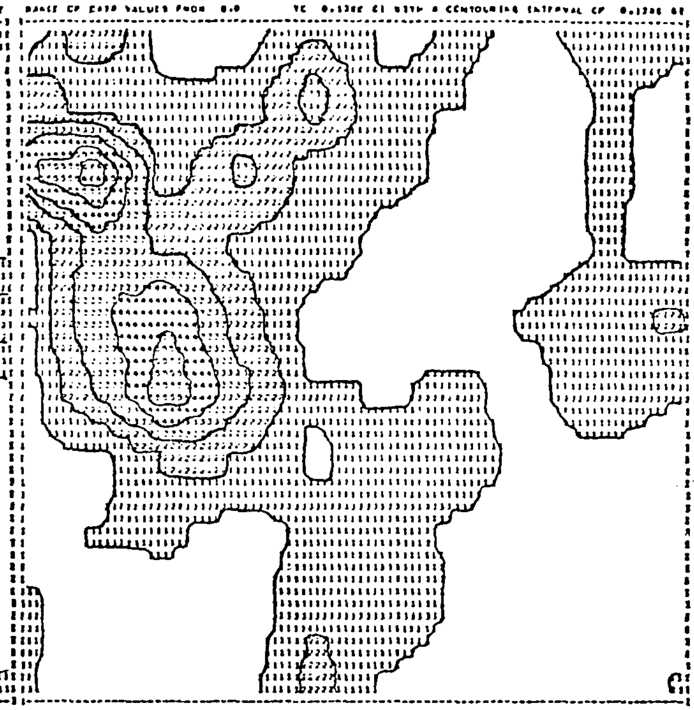

Fig. 11: 0750 - 21 May.

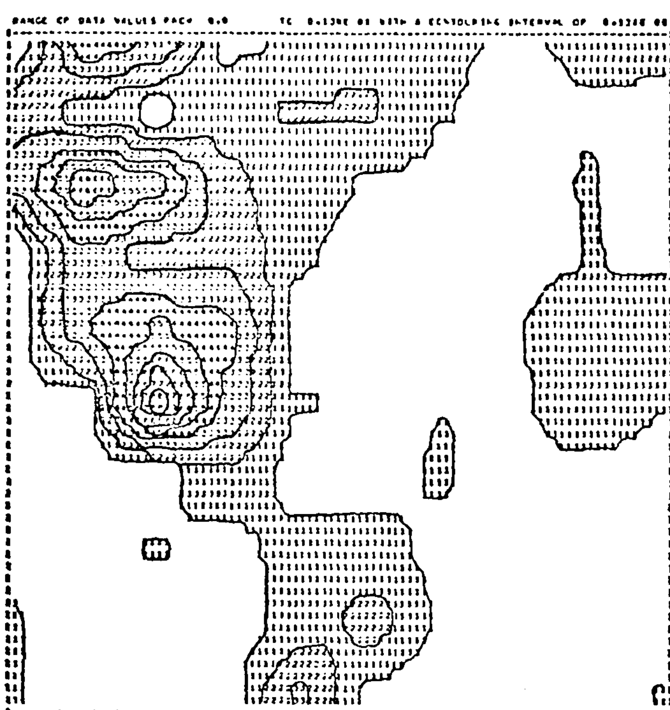

Fig. 12: 0755 - 21 May.

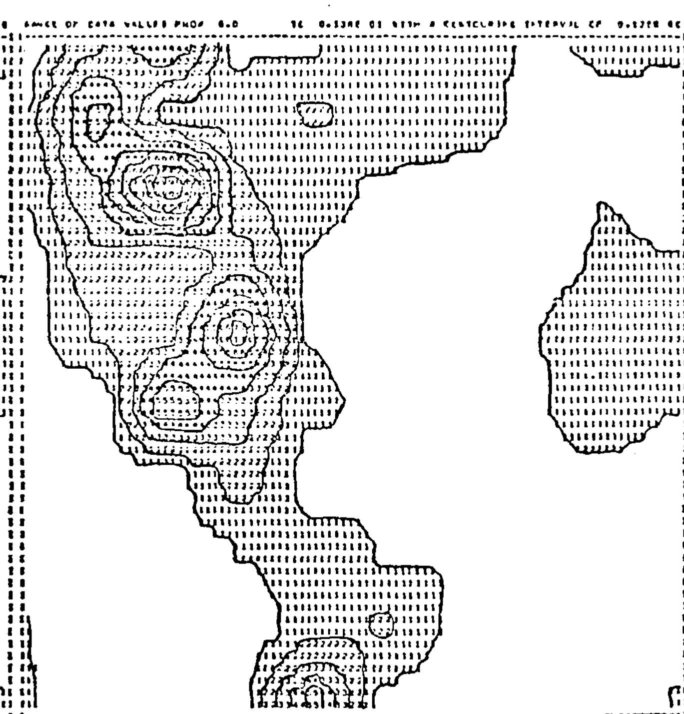

Fig. 13: 0800 - 21 May.

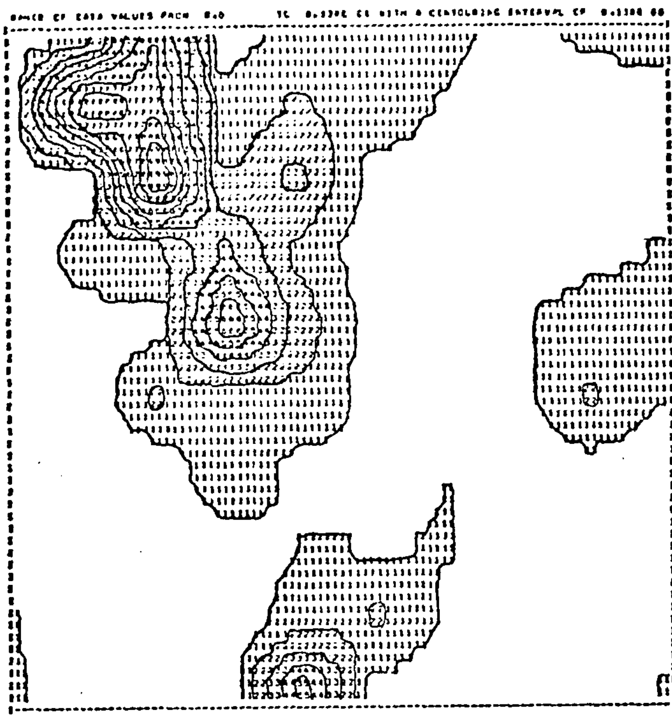

Fig. 14: 0805 - 21 May.

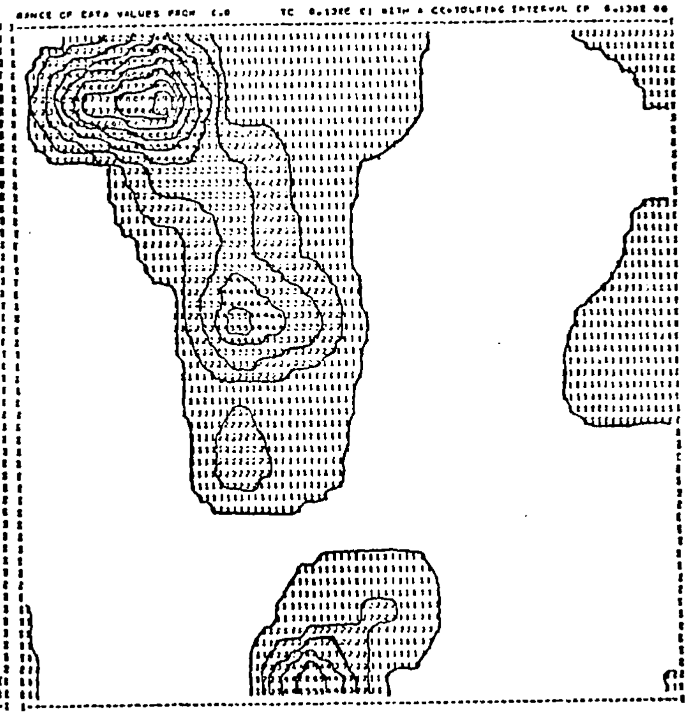

Fig. 15: 0810 - 21 May.

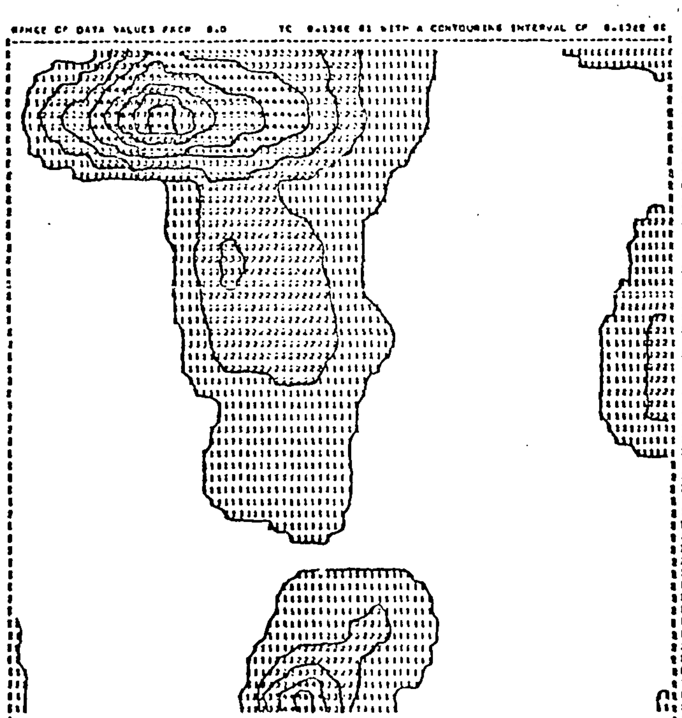

Fig. 16: 0815 - 21 May.

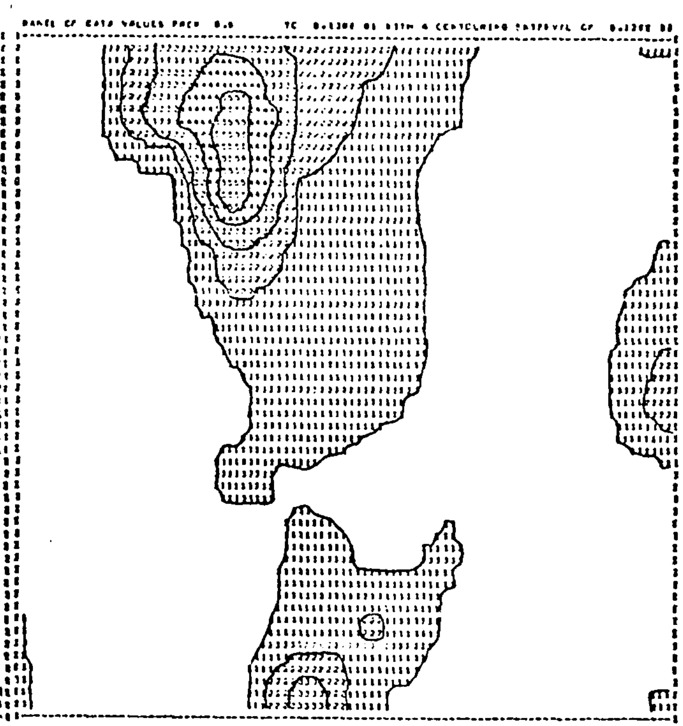

Fig. 17: 0820 - 21 May.

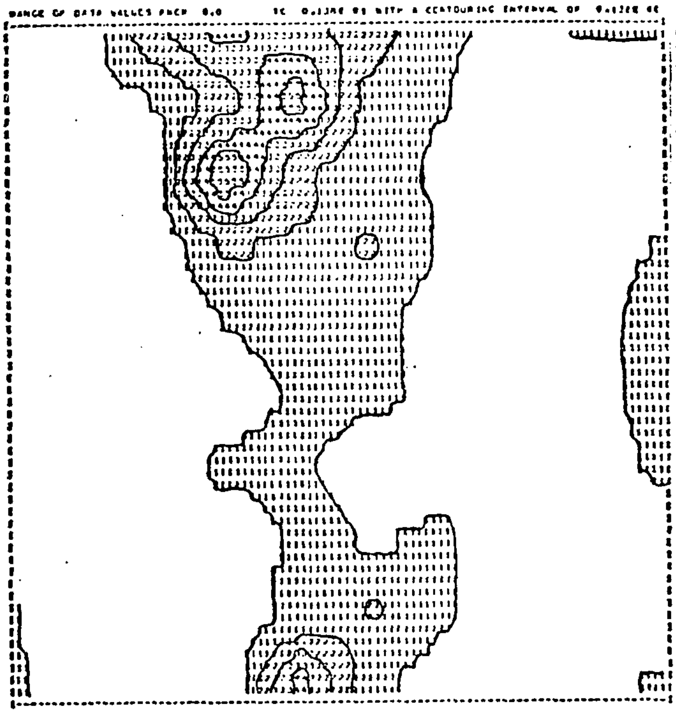

Fig. 18: 0825 - 21 May.

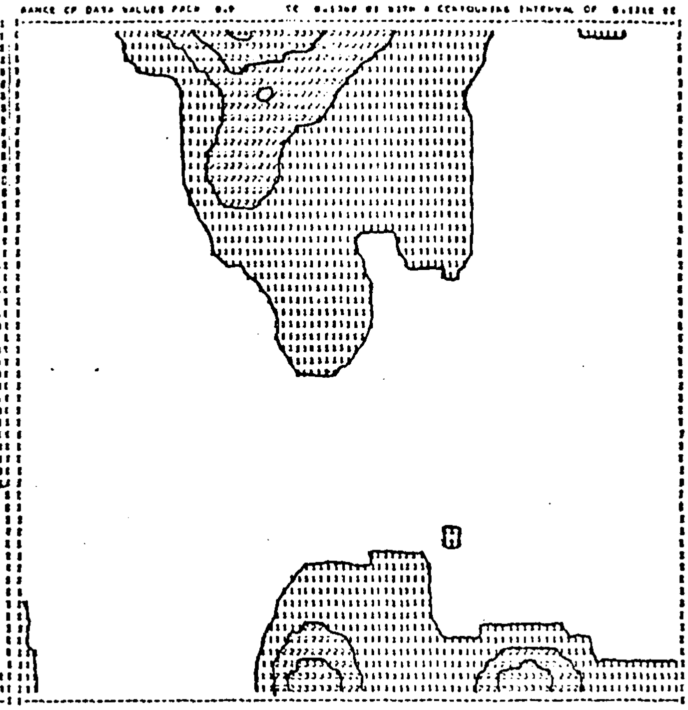

Fig. 19: 0830 - 21 May.

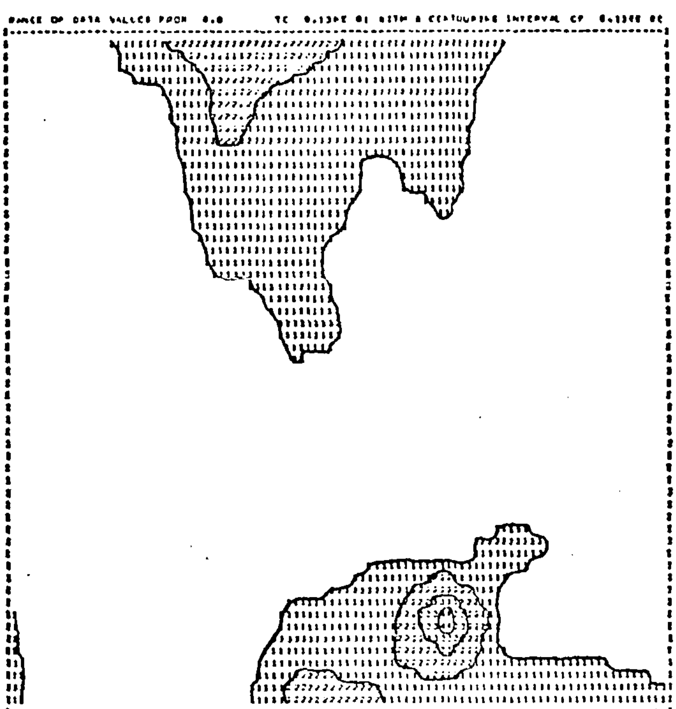

Fig. 20: 0835 - 21 May.

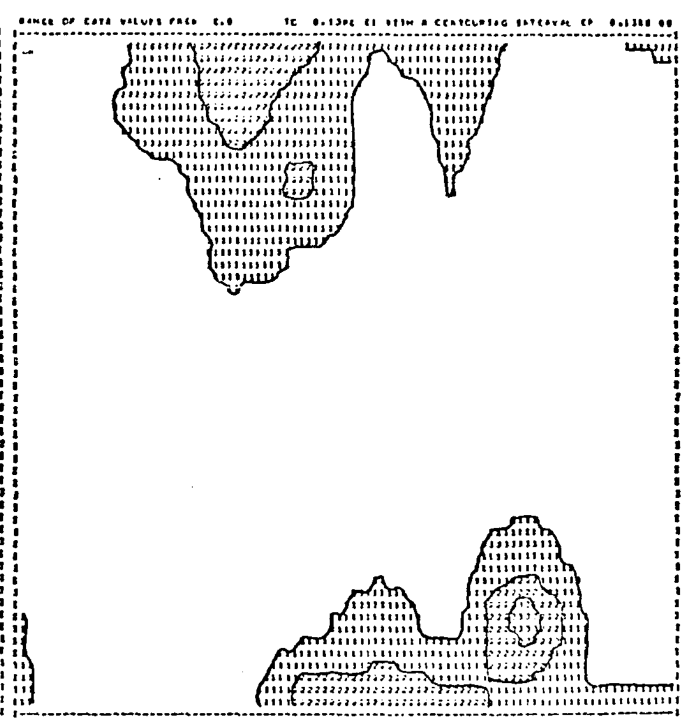

Fig. 21: 0840 - 21 May.

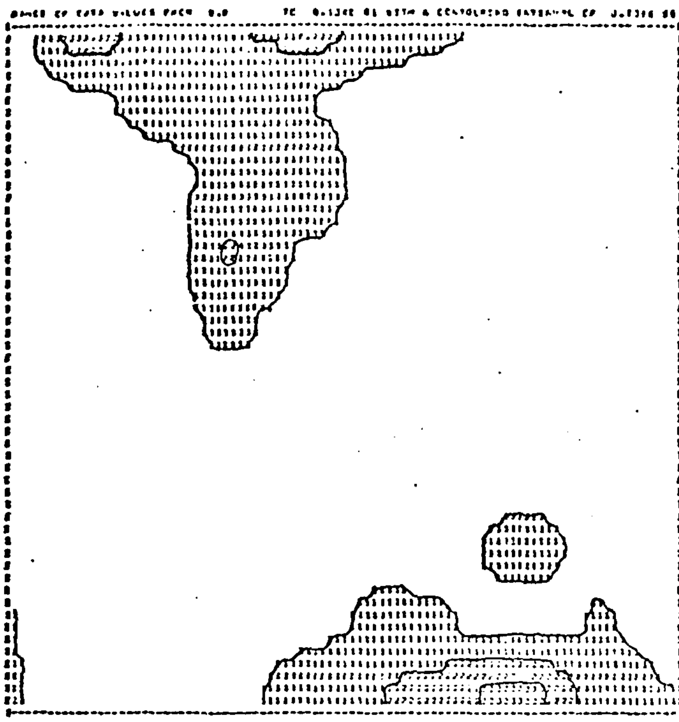

Fig. 22: 0845 - 21 May.

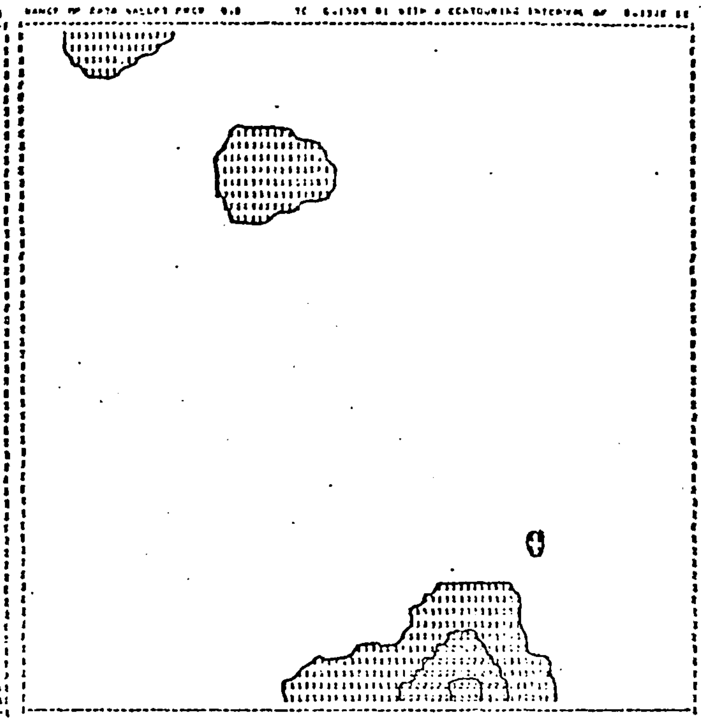

Fig. 23: 0850 - 21 May.

occuring but a few times at a few stations. The area covered by each grid analysis presented in this chapter, corresponds to the area enclosed by the corner brackets shown in Figure 1, page 9; approximately 41 x 41 square miles. Isohyetal lines correspond to rainfall increments of ≃ .14 mm. It should be noted that the precipitation pattern covers a rather large portion of the sensor network.

Table 2 shows the analyses results for the May 21st data set obtained when using procedure 1; that is withholding only the pertinent observation from the prediction set. It is seen that for the 18 analyzed times, the multivariate data configuration using the Marshall-Palmer Z-R relationship was best 11 times, the multivariate data configuration using the air-

ANALYSIS TIME	OBSERVATION VARIANCE	MARSHALL-PALMER Z-R RELATIONSHIP RADAR	MARSHALL-PALMER Z-R RELATIONSHIP GAUGES &RADAR	GAUGES	AIRCRAFT-DETERMINED Z-R RELATIONSHIP RADAR	AIRCRAFT-DETERMINED Z-R RELATIONSHIP GAUGES &RADAR
725	0.0681	0.0448	0.0419*	0.0478	0.0527	0.0458
		(0.0509)	(0.0519)	(0.0527)	(0.0528)	(0.0519)
730	0.0441	0.0288	0.0219*	0.0281	0.0315	0.0226
		(0.0327)	(0.0336)	(0.0340)	(0.0341)	(0.0337)
735	0.0426	0.0312	0.0256	0.0284	0.0351	0.0255*
		(0.0316)	(0.0324)	(0.0328)	(0.0329)	(0.0325)
740	0.1416	0.1177	0.1181	0.1070*	0.1229	0.1173
		(0.1050)	(0.1078)	(0.1092)	(0.1095)	(0.1080)
745	0.0565	0.0522	0.0441	0.0357*	0.0484	0.0396
		(0.0419)	(0.0430)	(0.0436)	(0.0437)	(0.0431)
750	0.0731	0.0624	0.0443	0.0460	0.0550	0.0406*
		(0.0542)	(0.0556)	(0.0563)	(0.0565)	(0.0557)
755	0.0946	0.0715	0.0571*	0.0675	0.0655	0.0595
		(0.0701)	(0.0720)	(0.0729)	(0.0731)	(0.0721)
800	0.3131	0.2509	0.2323*	0.2435	0.2580	0.2329
		(0.2322)	(0.2383)	(0.2414)	(0.2421)	(0.2388)
805	0.1715	0.1182	0.0626*	0.0908	0.1190	0.0732
		(0.1272)	(0.1306)	(0.1322)	(0.1326)	(0.1308)
810	0.1567	0.1109	0.0855*	0.0932	0.1166	0.0907
		(0.1162)	(0.1192)	(0.1208)	(0.1211)	(0.1195)
815	0.1182	0.0845	0.0572*	0.0680	0.0870	0.0615
		(0.0877)	(0.0900)	(0.0911)	(0.0914)	(0.0901)
820	0.1174	0.0710	0.0607*	0.0741	0.0820	0.0667
		(0.0871)	(0.0894)	(0.0905)	(0.0908)	(0.0896)
825	0.1021	0.0734	0.0536*	0.0748	0.0793	0.0558
		(0.0757)	(0.0777)	(0.0787)	(0.0790)	(0.0779)
830	0.0740	0.0551	0.0419*	0.0600	0.0596	0.0428
		(0.0549)	(0.0563)	(0.0570)	(0.0572)	(0.0564)
835	0.0576	0.0454	0.0436	0.0494	0.0479	0.0434*
		(0.0427)	(0.0439)	(0.0444)	(0.0446)	(0.0439)
840	0.0227	0.0143	0.0120	0.0163	0.0169	0.0115*
		(0.0169)	(0.0173)	(0.0175)	(0.0176)	(0.0173)
845	0.0315	0.0189	0.0163	0.0201	0.0223	0.0160*
		(0.0233)	(0.0239)	(0.0243)	(0.0243)	(0.0240)
850	0.0402	0.0260	0.0257*	0.0340	0.0312	0.0263
		(0.0303)	(0.0307)	(0.0311)	(0.0313)	(0.0307)

TABLE 2. A LISTING OF ANALYSIS ERROR VARIANCES COMPUTED USING THE DATA SETS INDICATED BY THE VARIOUS COLUMN HEADINGS. THE OBSERVATION VARIANCES ARE INCLUDED FOR COMPARISON PURPOSES. THE LOWEST VARIANCE FOR EACH ANALYSIS TIME IS MARKED (*). THIS INDICATES WHICH ANALYSIS DATA CONFIGURATION PRODUCED THE BEST RESULTS. THE VALUES IN THE PARENTHESES ARE THE MODELED VARIANCE ESTIMATES. THIS DATA IS FROM THE 21ST OF MAY.

craft determined Z-R relationship was best 5 times, and the surface-gauge data set produced the minimum error variance for the times 0740 and 0745. Graphics of these results are shown in Figures 24 and 25 for the Marshall-Palmer Z-R relationship and in Figures 26 and 27 for the aircraft determined Z-R relationship. Figures 24 and 26 compare the analyses results with the original data variance. Since the variance in the original data set is due to the measurement of the signal plus the noise in the phenomenon, and since the variance of the error field reflects only the <u>noise</u> in the observation set, the error variance should always be less than the variance of the data set itself, assuming that some 'signal' was indeed observed. This is one way to assess the worth of an analysis technique (See section E of this chapter for a more detailed discussion of analysis filtering properties). The observation variance is included in Figures 24 and 26 (and similar figures throughout this chapter) in order to make the above assessment more apparent. Figures 25 and 27 express the analysis error variance as a percentage of the variance of the respective observation set for each time period analyzed.

From these results, it might be conjectured that the multivariate approach does indeed produce better analyses of the surface precipitation field than does either of the other analysis approaches presented. It can be readily seen from Figures 25 and 27 that the multivariate approach (M) produces as small or smaller an error variance, for 16 out of

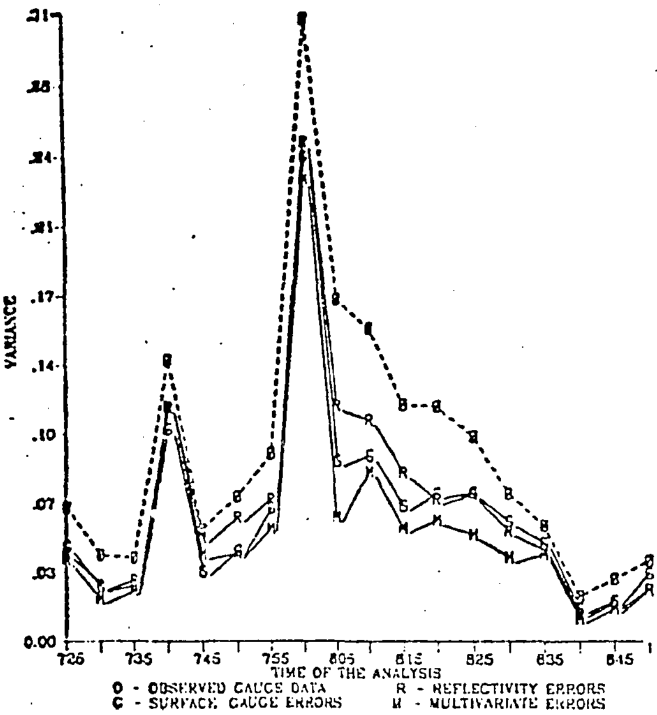

Fig. 24: A comparison of the computed error variances and the respective data variance for the May 21st observation set using the Marshall-Palmer Z-R conversion and procedure 1.

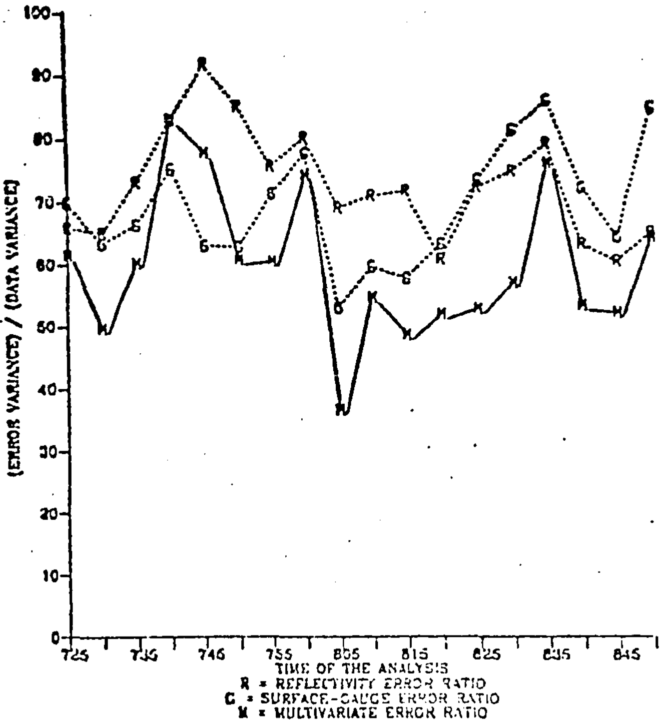

Fig. 25: The error variances of Figure 24 expressed as a percentage of the respective data variance.

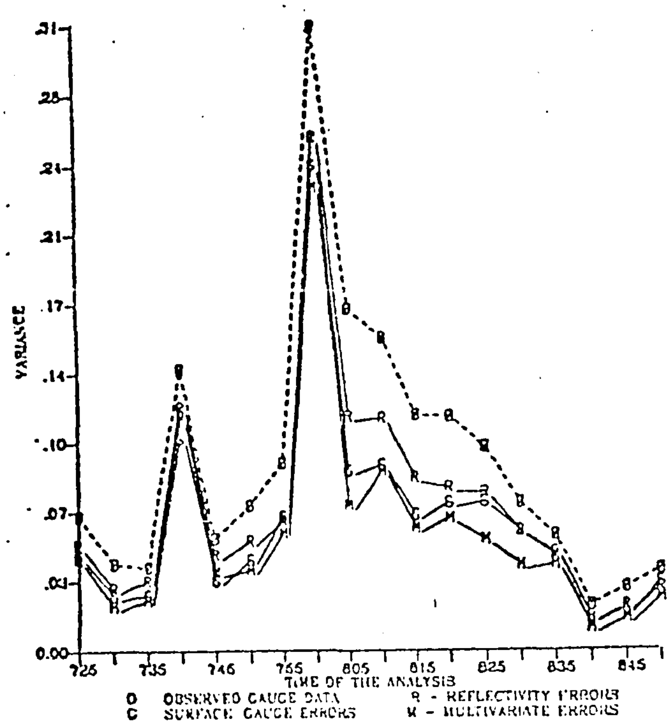

Fig. 26: A comparison of the computed error variances and the respective data variance for the May 21st observation set using the aircraft-determined Z-R conversion and procedure 1.

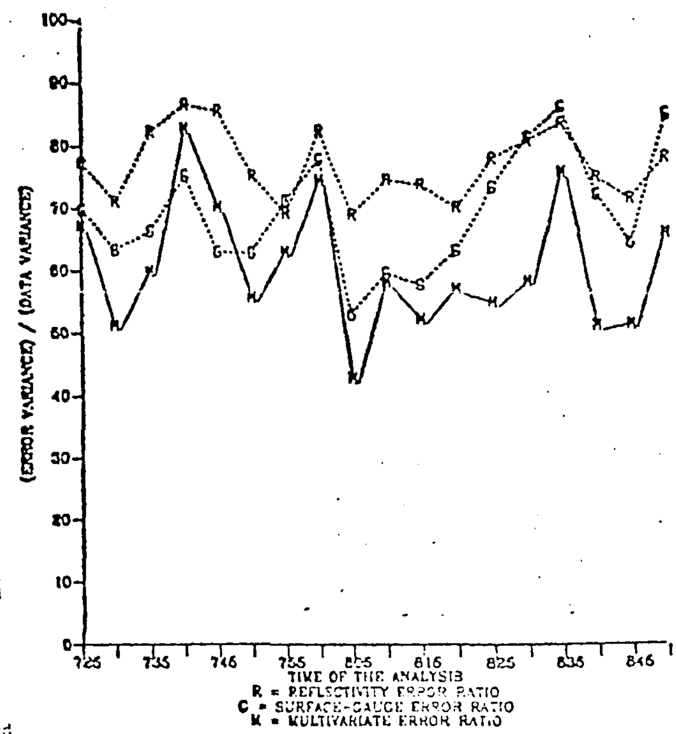

Fig. 27: The error variances of Figure 26 expressed as a percentage of the respective data variance.

the 18 analyzed cases, no matter which Z-R relationship is employed. However, it would be tenuous at best, to make any definite conclusions based on this one storm realization.

Table 3 shows the analyses results for the May 21st data set obtained when using procedure 2; that is withholding the entire time series observation vector from the predictor set. It is seen that for the 18 analyzed times: when using the Marshall-Palmer Z-R conversion, the multivariate data configuration was best 4 times while the radar data alone produced the better result 9 times; when using the empirically derived Z-R conversion, the multivariate data configuration was best only 1 time while the radar data alone produced the better result 3 times; and the surface-gauge data set produced the minimum error variance for the analysis time 0820. Graphics of these results are shown in Figures 28 and 29 for the Marshall-Palmer Z-R relationship and in Figures 30 and 31 for the aircraft determined Z-R relationship. Figures 28 and 30 compare the analyses results with the original data variance while Figures 29 and 31 express the analysis error variance as a percentage of the variance of the respective observation set for each time period analyzed.

From these results, it might be conjectured that the converted radar precipitation estimates constitute a better predictor set for the observed surface precipitation field, than does either of the other data configurations utilized. As can be seen from Figures 29 and 31 the radar data set

ANALYSIS TIME	OBSERVATION VARIANCE	MARSHALL-PALMER Z-R RELATIONSHIP RADAR	MARSHALL-PALMER Z-R RELATIONSHIP GAUGES &RADAR	GAUGES	AIRCRAFT-DETERMINED Z-R RELATIONSHIP RADAR	AIRCRAFT-DETERMINED Z-R RELATIONSHIP GAUGES &RADAR
725	0.0681	0.0448*	0.0483	0.0525	0.0527	0.0528
		(0.0509)	(0.0496)	(0.0494)	(0.0528)	(0.0522)
730	0.0441	0.0288*	0.0312	0.0375	0.0315	0.0359
		(0.0327)	(0.0323)	(0.0319)	(0.0341)	(0.0339)
735	0.0426	0.0312*	0.0343	0.0405	0.0351	0.0346
		(0.0316)	(0.0312)	(0.0307)	(0.0329)	(0.0327)
740	0.1416	0.1177	0.1142*	0.1222	0.1229	0.1161
		(0.1050)	(0.1037)	(0.1022)	(C.1095)	(0.1088)
745	0.0565	0.0522	0.0537	0.0528	0.0484*	0.0512
		(0.0419)	(0.0414)	(0.0408)	(0.0437)	(0.0434)
750	0.0731	0.0624	0.0651	0.0694	0.0550*	0.0636
		(0.0542)	(0.0535)	(0.0528)	(0.0565)	(0.0561)
755	0.0946	0.0715	0.0709	0.0838	0.0655*	0.0759
		(0.0701)	(0.0692)	(0.0683)	(0.0731)	(0.0726)
800	0.3131	0.2509*	0.2596	0.2807	0.2580	0.2640
		(0.2322)	(0.2292)	(0.2260)	(0.2421)	(0.2405)
805	0.1715	0.1182	0.1088*	0.1372	0.1180	0.1269
		(0.1272)	(0.1255)	(0.1238)	(0.1326)	(0.1318)
810	0.1567	0.1109	0.1057*	0.1107	0.1166	0.1156
		(0.1162)	(0.1147)	(0.1131)	(0.1211)	(0.1203)
815	0.1182	0.0845	0.0846	0.0840	0.0870	0.0834*
		(0.0877)	(0.0865)	(0.0853)	(0.0914)	(0.0908)
820	0.1174	0.0710	0.0703	0.0699*	0.0820	0.0867
		(0.0871)	(0.0859)	(0.0848)	(0.0908)	(0.0902)
825	0.1021	0.0734	0.0707*	0.0820	0.0793	0.0793
		(0.0757)	(0.0748)	(0.0737)	(0.0790)	(0.0785)
830	0.0740	0.0551*	0.0615	0.0651	0.0596	0.0649
		(C.0549)	(0.0542)	(0.0534)	(0.0572)	(0.0568)
835	0.0576	0.0454*	0.0515	0.0538	0.0479	0.0546
		(0.0427)	(0.0422)	(0.0416)	(C.0446)	(0.0443)
840	0.0227	0.0143*	0.0176	0.0207	0.0169	0.0174
		(0.0169)	(0.0166)	(0.0164)	(0.0176)	(0.0175)
845	0.0315	C.0189*	0.0200	0.0261	0.0223	0.0233
		(0.0233)	(0.0230)	(0.0227)	(0.0243)	(0.0242)
850	0.0402	0.0260*	0.0280	0.0348	0.0312	0.0304
		(0.0303)	(0.0293)	(0.0292)	(0.0313)	(0.0308)

TABLE 3. A LISTING OF ANALYSIS ERROR VARIANCES COMPUTED USING THE DATA SETS INDICATED BY THE VARICUS COLUMN HEADINGS. THE OBSERVATION VARIANCES ARE INCLUDED FOR COMPARISCN PURPOSES. THE LOWEST VARIANCE FOR EACH ANALYSIS TIME IS MARKED (*). THIS INDICATES WHICH ANALYSIS DATA CONFIGURATION PRODUCED THE BEST RESULTS. THE VALUES IN THE PARENTHESES ARE THE MODELED VARIANCE ESTIMATES. THIS DATA IS FROM THE 21ST OF MAY.

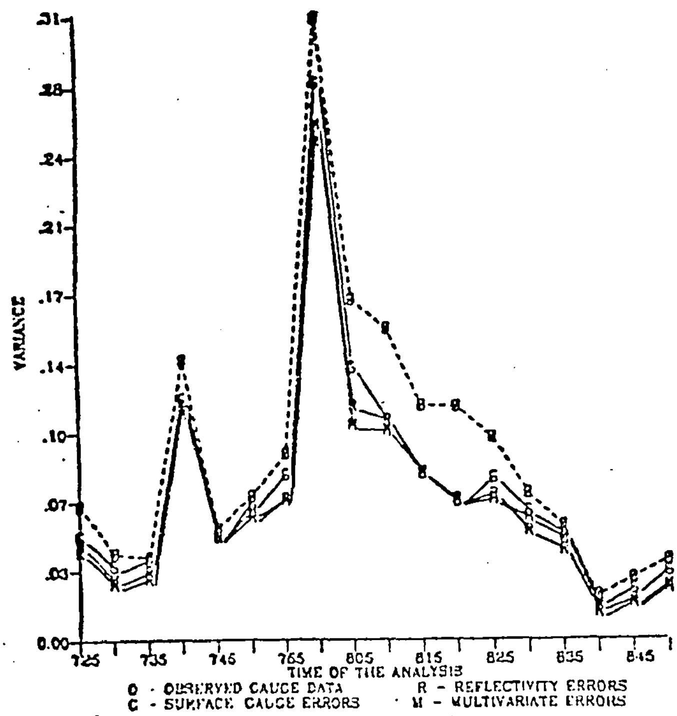

Fig. 28: A comparison of the computed error variances and the respective data variance for the May 21st observation set using the Marshall-Palmer Z-R conversion and procedure 2.

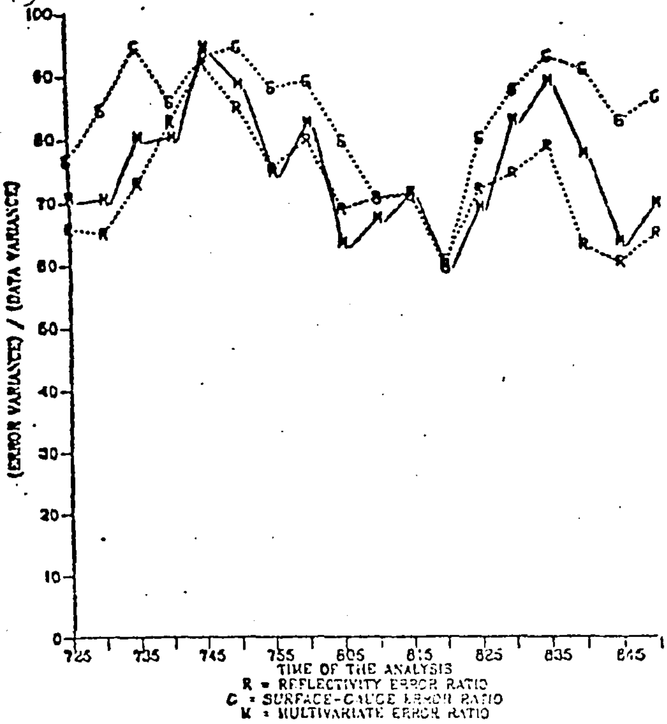

Fig. 29: The error variances of Figure 28 expressed as a percentage of the respective data variance.

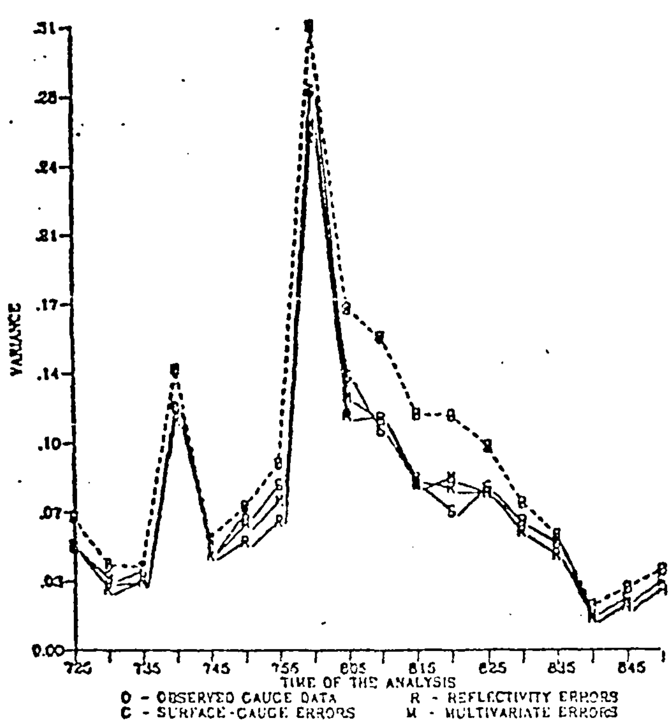

Fig. 30: A comparison of the computed error variances and the respective data variance for the May 21st observation set using the aircraft-determined Z-R conversion and procedure 2.

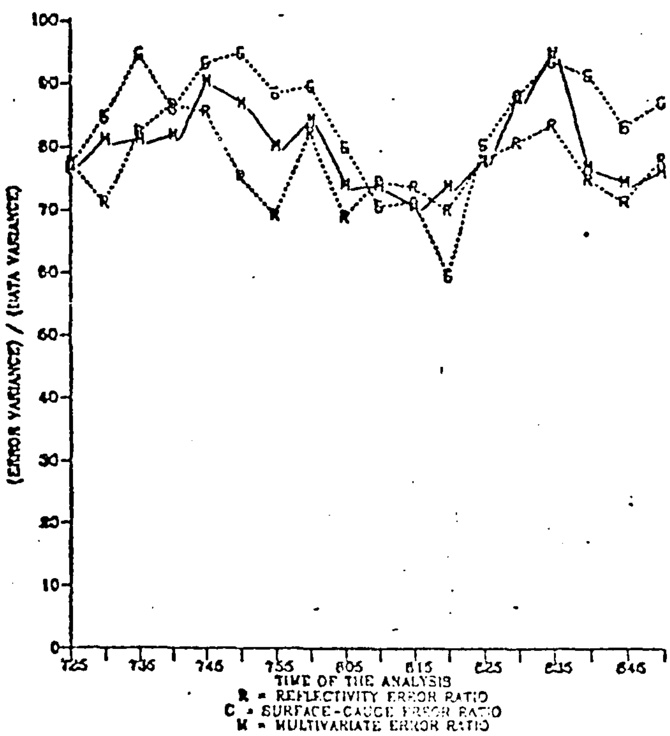

Fig. 31: The error variances of Figure 30 expressed as a percentage of the respective data variance.

produces as small or smaller an error variance 14 and 12 times respectively, for the 18 analyzed time periods. Once again, however, it would be hazardous to make any conclusions based on this single storm realization.

In addition to the comparisons already mentioned, it was also possible to investigate the relative worth of the different Z-R relationships used, the results of which are shown in Figures 32 and 33. Both graphics compare the error variance using the multivariate data analysis configuration. The first figure is the result of employing procedure 1, and the second figure is the result of employing procedure 2. These comparisons would tend to imply that neither Z-R conversion was predominantly better than the other when used in a multivariate analysis environment.

And finally, employing procedure 1 for this particular storm system, the effect of offsetting the radar derived data set by 1 time unit (5 minutes) was explored. The results, using the aircraft determined Z-R relationship, are given in Table 4. For the 17 analyzed time periods, the "normal" analysis produced the smallest error variance 11 times, the offset radar data set produced a better result 4 times, and the surface-gauge data set was best for the times 0740 and 0745. From these comparisons it was tentatively concluded that poorer analyses were produced by offsetting the radar data set a minus 5 minutes. This was consistent with the results found in Takeuchi, the implication being that it

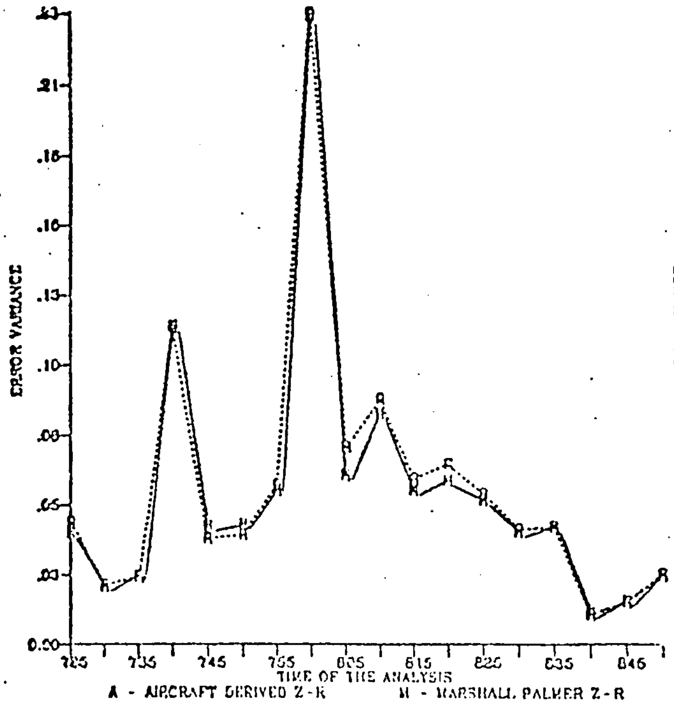

Fig. 32: A comparison of the effect of different Z-R relationships on the analysis of surface precipitation for May 21st when procedure 1 was used.

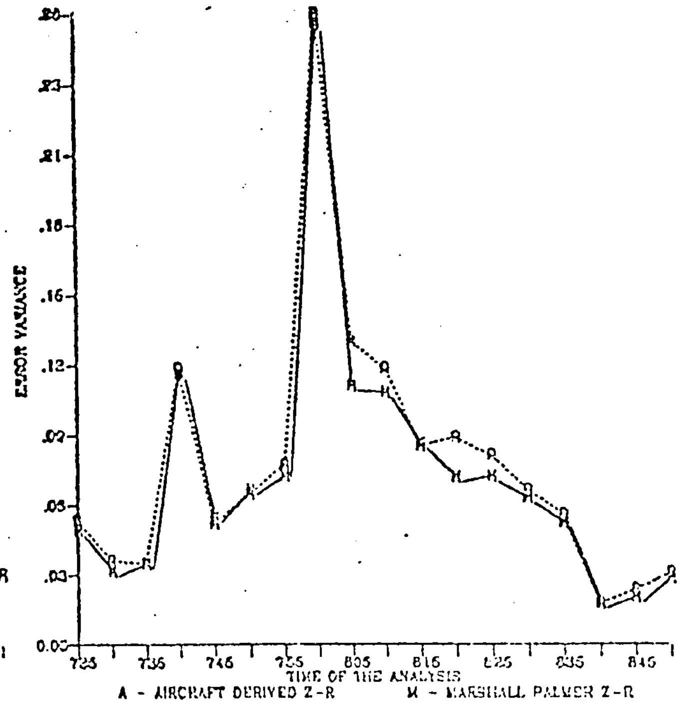

Fig. 33: A comparison of the effect of different Z-R relationships on the analysis of surface precipitation for May 21st when procedure 2 was used.

took considerably less than 5 minutes for the radar detected precipitation to reach the surface. (Remember the radar elevation angle was only 1°.) Based on these findings, the observed temporal relationship between all data sets was maintained for all other analyses presented in this chapter.

B. 25 May, 1974, 0755-0905

The deduced correlation and cross-correlation model parameters for this particular data set are shown in Table 5. Parameters A8, A9, A10, and A11 are in nautical miles, while parameter A14 is in minutes.

Grid analyses for this storm system are shown in Figures 34 thru 48. In general the rainfall rate for this

ANALYSIS TIME	OBSERVATION VARIANCE	A.D. Z-R RELATIONSHIP REGULAR DATA SET RADAR	GAUGES &RADAR	GAUGES	A.C. Z-R RELATIONSHIP OFFSET DATA SET RADAR	GAUGES &RADAR
725	0.0681	0.0527	0.0458	0.0478	0.0000*	0.0000*
		(0.0528)	(0.0519)	(0.0527)	(0.0000)	(0.0000)
730	0.0441	0.0315	0.0226*	0.0281	0.0321	0.0286
		(0.0341)	(0.0337)	(0.0340)	(0.0344)	(0.0335)
735	0.0426	0.0351	0.0255	0.0284	0.0311	0.0232*
		(0.0329)	(0.0325)	(0.0328)	(0.0330)	(0.0323)
740	0.1416	0.1229	0.1173	0.1070*	0.1328	0.1195
		(0.1095)	(0.1080)	(0.1092)	(0.1099)	(0.1076)
745	0.0565	0.0484	0.0396	0.0357*	0.0491	0.0371
		(0.0437)	(0.0431)	(0.0436)	(0.0438)	(0.0429)
750	0.0731	0.0550	0.0406*	0.0460	0.0646	0.0423
		(0.0565)	(0.0557)	(0.0563)	(0.0567)	(0.0555)
755	0.0946	0.0655	0.0595*	0.0675	0.0783	0.0611
		(0.0731)	(0.0721)	(0.0729)	(0.0734)	(0.0718)
800	0.3131	0.2580	0.2329*	0.2435	0.2868	0.2419
		(0.2421)	(0.2388)	(0.2414)	(0.2429)	(0.2379)
805	0.1715	0.1180	0.0732*	0.0908	0.1444	0.0818
		(0.1326)	(0.1308)	(0.1322)	(0.1331)	(0.1303)
810	0.1567	0.1166	0.0907*	0.0932	0.1217	0.0915
		(0.1211)	(0.1195)	(0.1208)	(0.1215)	(0.1190)
815	0.1182	0.0870	0.0615*	0.0680	0.0976	0.0715
		(0.0914)	(0.0901)	(0.0911)	(0.0917)	(0.0898)
820	0.1174	0.0820	0.0667*	0.0741	0.0923	0.0671
		(0.0908)	(0.0896)	(0.0905)	(0.0911)	(0.0892)
825	0.1021	0.0793	0.0558	0.0748	0.0790	0.0543*
		(0.0790)	(0.0779)	(0.0787)	(0.0792)	(0.0776)
830	0.0740	0.0596	0.0428*	0.0600	0.0616	0.0432
		(0.0572)	(0.0564)	(0.0570)	(0.0574)	(0.0562)
835	0.0576	0.0479	0.0434	0.0494	0.0511	0.0431*
		(0.0446)	(0.0439)	(0.0444)	(0.0447)	(0.0438)
840	0.0227	0.0169	0.0115	0.0162	0.0163	0.0111*
		(0.0176)	(0.0173)	(0.0175)	(0.0176)	(0.0173)
845	0.0315	0.0223	0.0160*	0.0201	0.0228	0.0166
		(0.0243)	(0.0240)	(0.0243)	(0.0244)	(0.0239)
850	0.0402	0.0312	0.0263*	0.0340	0.0290	0.0272
		(0.0313)	(0.0307)	(0.0311)	(0.0314)	(0.0305)

TABLE 4. A LISTING OF ANALYSIS ERROR VARIANCES COMPUTED USING THE DATA SETS INDICATED BY THE VARIOUS COLUMN HEADINGS. THE OBSERVATION VARIANCES ARE INCLUDED FOR COMPARISON PURPOSES. THE LOWEST VARIANCE FOR EACH ANALYSIS TIME IS MARKED (*). THIS INDICATES WHICH ANALYSIS DATA CONFIGURATION PRODUCED THE BEST RESULTS. THE VALUES IN THE PARENTHESES ARE THE MODELED VARIANCE ESTIMATES. THIS DATA IS FROM THE 21ST OF MAY.

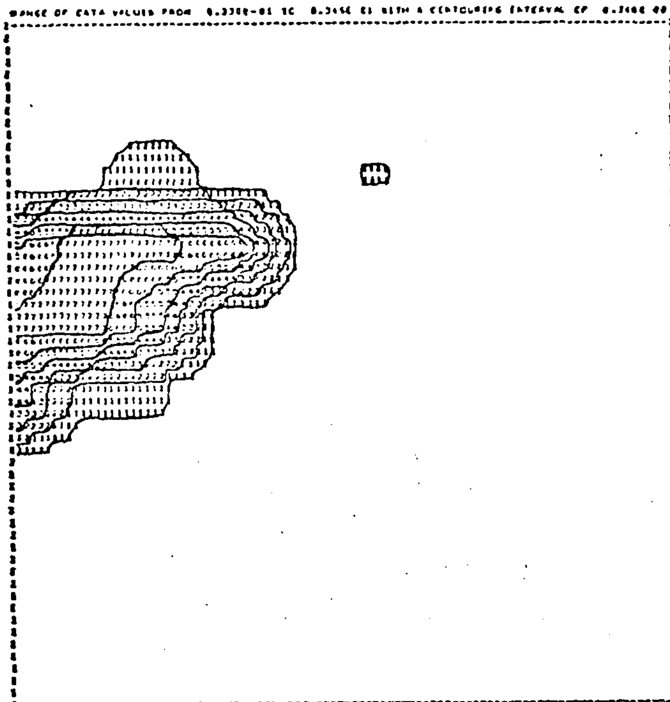

Fig. 34: 0755 - 25 May.

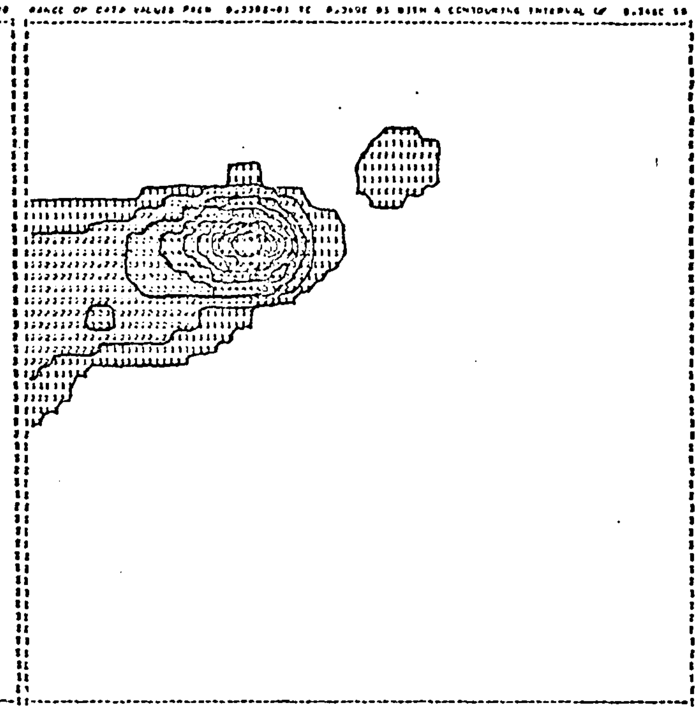

Fig. 35: 0800 - 25 May.

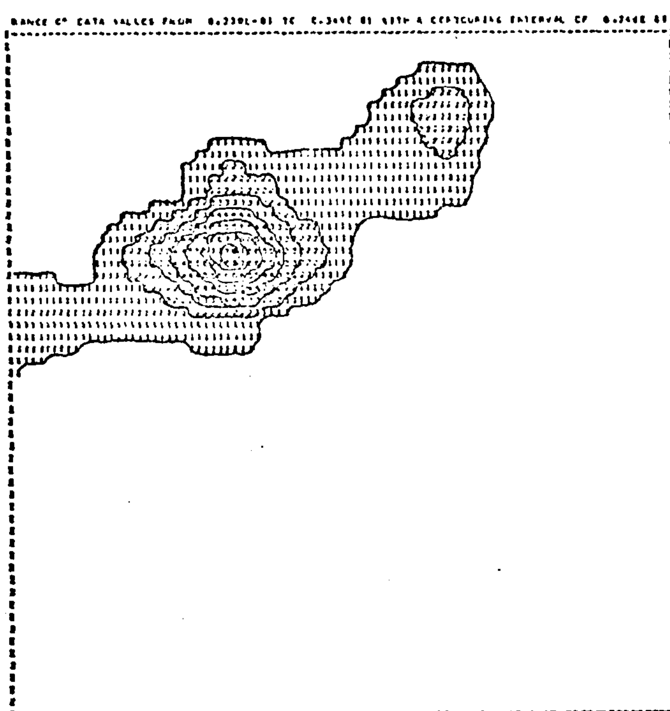

Fig. 36: 0805 - 25 May.

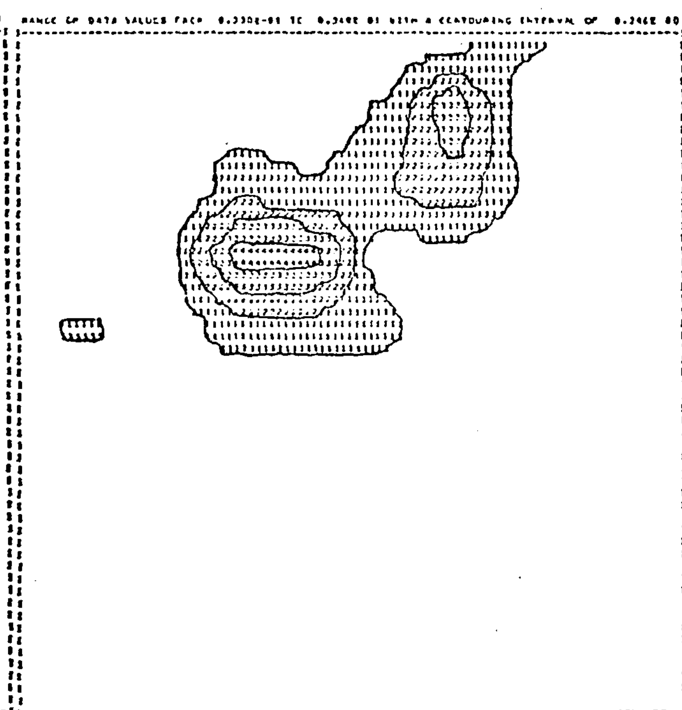

Fig. 37: 0810 - 25 May.

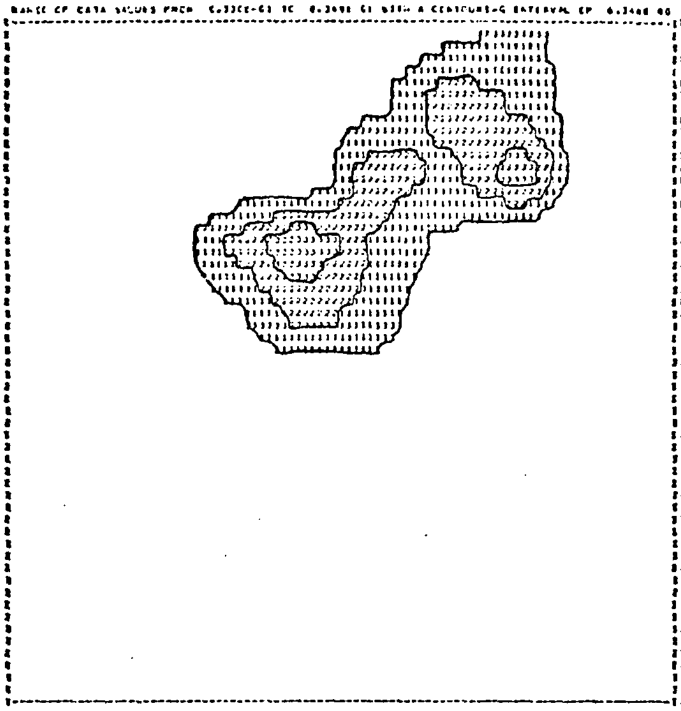

Fig. 38: 0815 - 25 May.

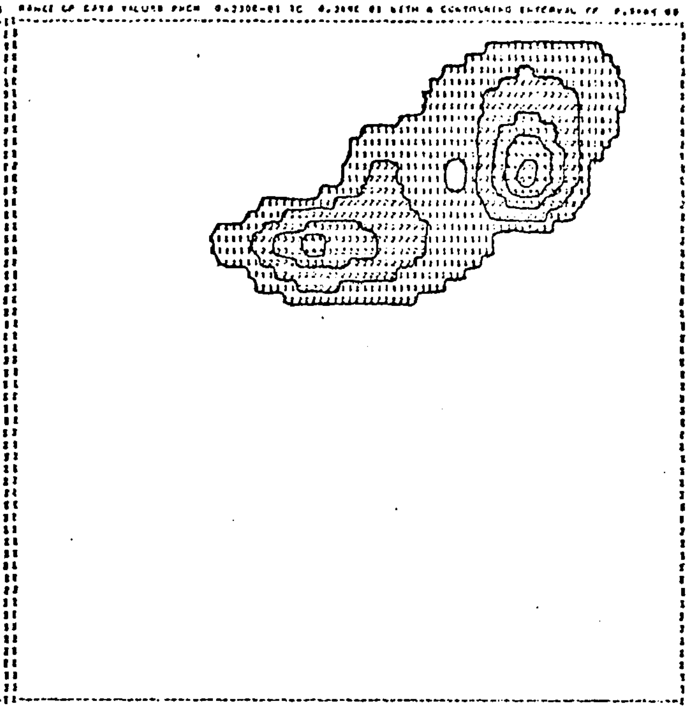

Fig. 39: 0820 - 25 May.

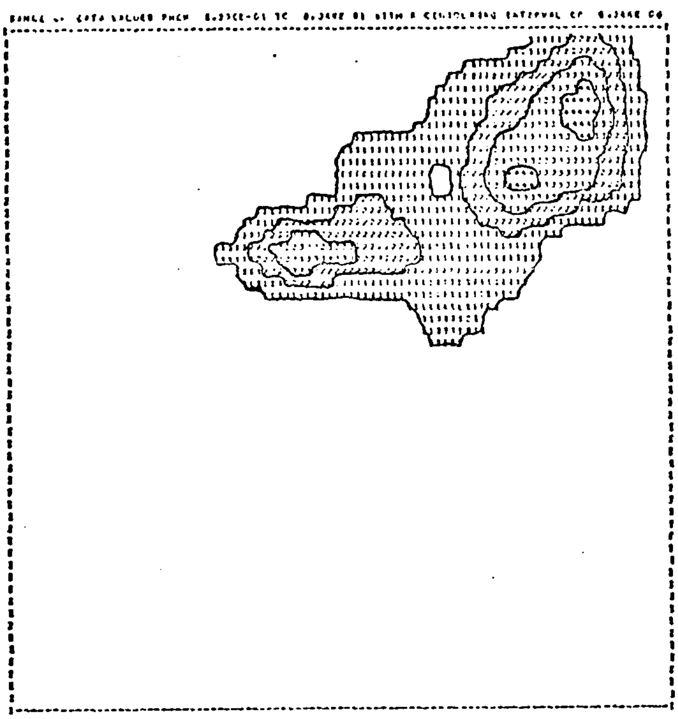

Fig. 40: 0825 - 25 May.

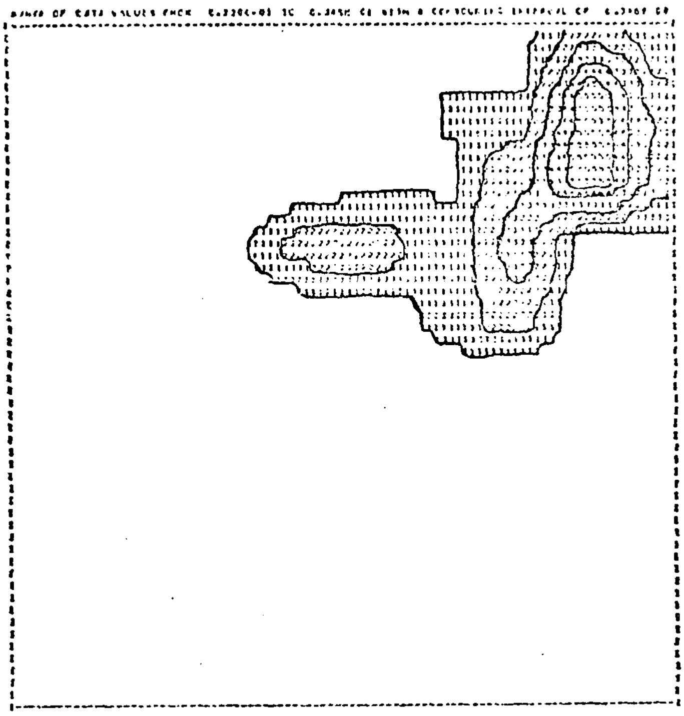

Fig. 41: 0830 - 25 May.

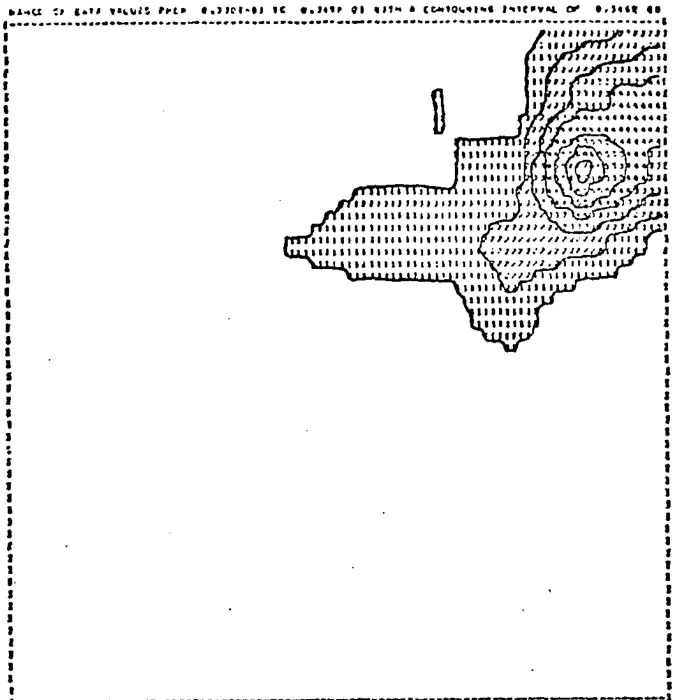

Fig. 42: 0835 - 25 May.

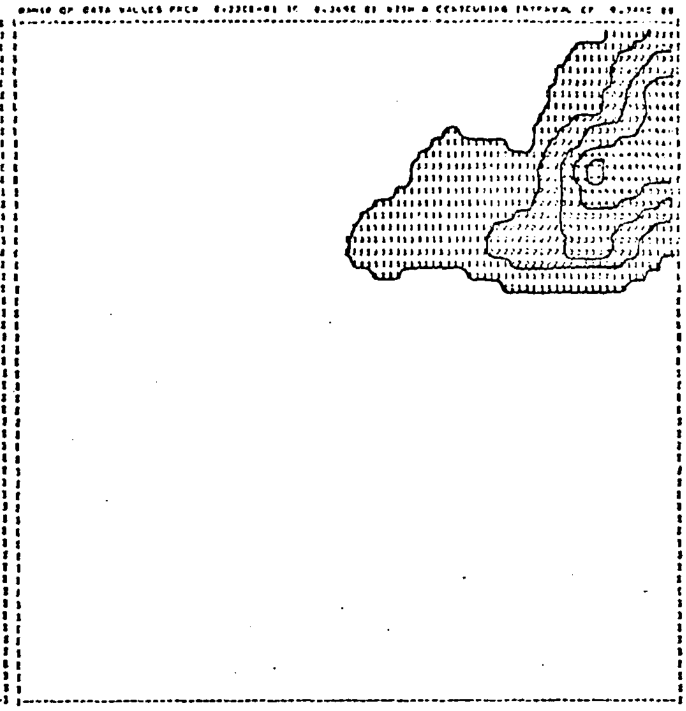

Fig. 43: 0840 - 25 May.

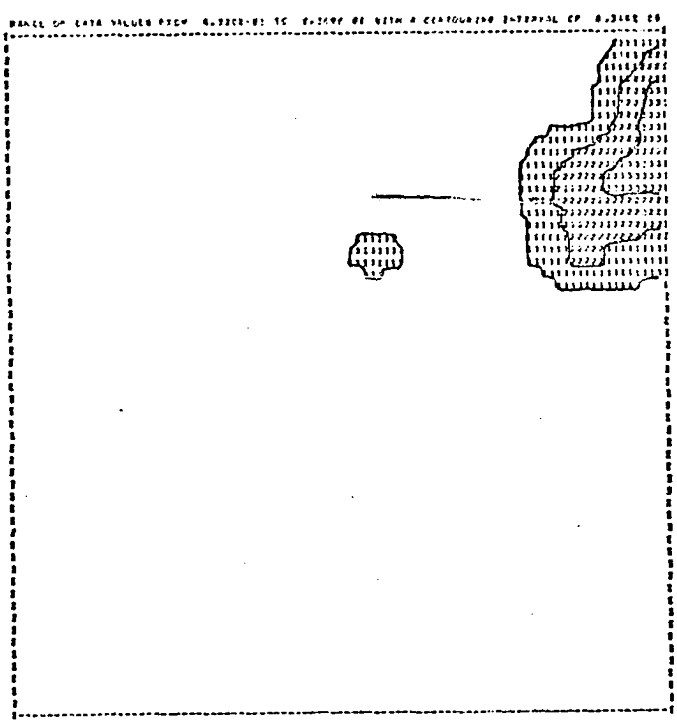

Fig. 44: 0845 - 25 May.

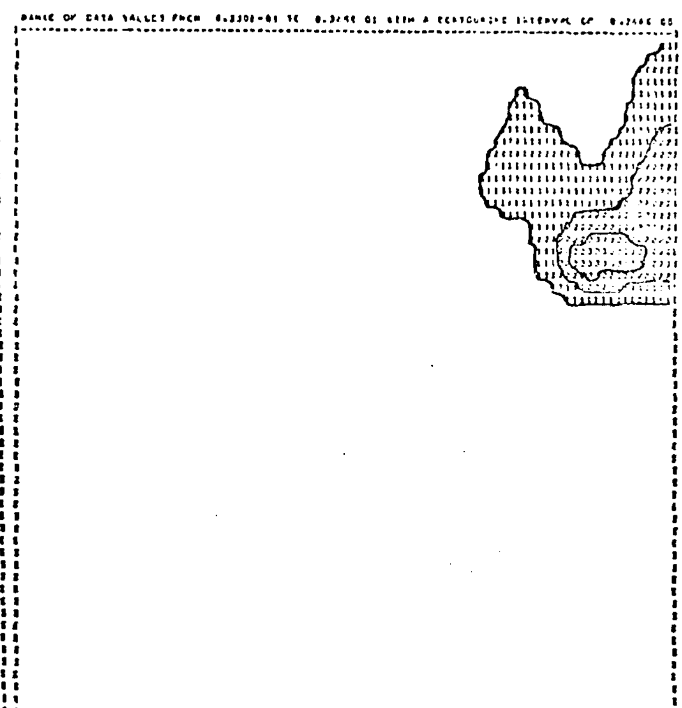

Fig. 45: 0850 - 25 May.

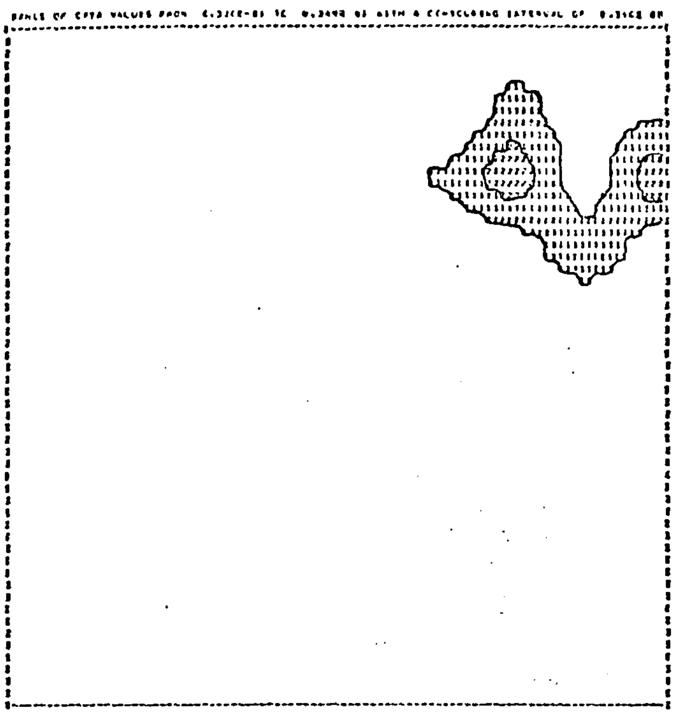

Fig. 46: 0855 - 25 May.

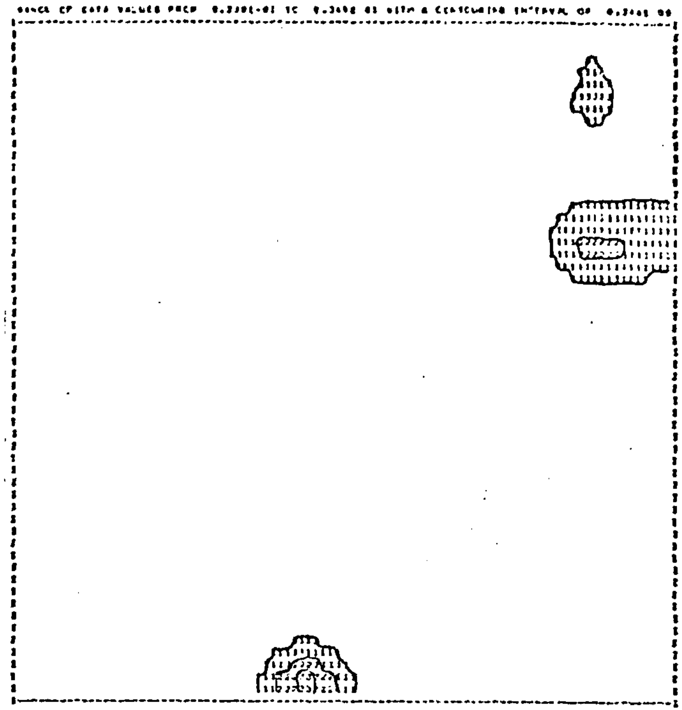

Fig. 47 : 0900 - 25 May.

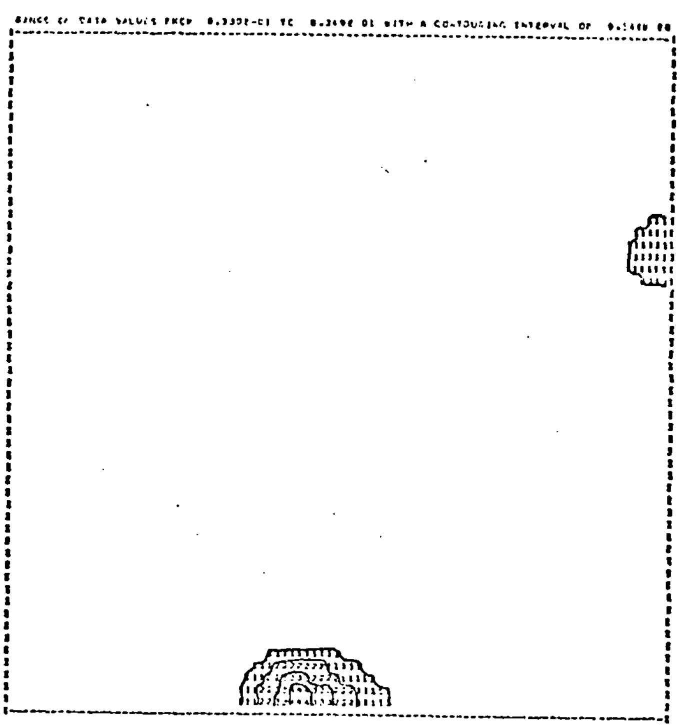

Fig. 48: 0905 - 25 May.

	POSITIVE X		POSITIVE Y		NEGATIVE X		NEGATIVE Y		TIME	
	A1	A8	A2	A9	A3	A10	A4	A11	A14	A7
CORR$\{X_1X_1\}$	-.11	9.0	.09	9.0	-.19	8.9	-.04	9.1	25.1	.41
CORR(X_2X_2)	-.08	6.9	.01	12.1	-.28	7.6	-.09	6.1	23.0	.41
CORR(X_1X_2)	-.04	5.2	-.00	9.1	-.16	8.9	-.01	8.7	25.0	.36
CORR(X_2X_1)	.01	9.0	.07	9.0	-.19	8.8	-.17	8.9	24.7	.14
CORR$[X_2X_2]$	-.08	5.7	.09	7.2	-.24	7.0	-.05	4.8	10.7	.53
CORR$[X_1X_2]$	-.04	5.2	-.01	8.7	-.20	8.8	-.03	8.5	24.8	.35
CORR$[X_2X_1]$	-.05	9.0	.09	9.0	-.20	8.8	-.15	8.9	24.7	.14

TABLE 5.--The signal-plus-noise model parameters for 25 May, 0755-0905. The parenthesized model designation, CORR(X_iX_j), denotes the parameters obtained from the Marshall-Palmer Z-R conversion, while the bracketed model designation, CORR$[X_iX_j]$, denotes the parameters obtained from the empirically derived Z-R conversion.

data set varied from 0.3 inches/hour to as much as 3.5 inches/hour, with these higher rates occurring only a very few times at but a couple of stations. Isohyetal lines correspond to rainfall increments of ≈ .35 mm. Notice that this precipitation pattern, in contrast to the one of 21 May, covers a much smaller percentage of the sensor network at any given time. This is true for all three storm conditions that occurred on this date.

Table 6 shows the analyses results for the first storm occurrence of May 25th obtained when using procedure 1. For the 15 analyzed times, the Marshall-Palmer Z-R relationship produced the best analysis 2 times using the radar data alone and 4 times when the multivariate approach was used, while the aircraft determined Z-R relationship produced the best analysis the remaining 9 times when used in

ANALYSIS TIME	OBSERVATION VARIANCE	MARSHALL-PALMER Z-R RELATIONSHIP RADAR	MARSHALL-PALMER Z-R RELATIONSHIP GAUGES &RADAR	GAUGES	AIRCRAFT-DETERMINED Z-R RELATIONSHIP RADAR	AIRCRAFT-DETERMINED Z-R RELATIONSHIP GAUGES &RADAR
755	1.0420	0.6444 (0.8048)	0.6133* (0.7426)	0.6884 (0.7438)	0.7036 (0.8188)	0.6644 (0.7429)
800	0.7600	0.4184 (0.5834)	0.2409 (0.5393)	0.2409 (0.5394)	0.4353 (0.5776)	0.2408* (0.5393)
805	0.4556	0.2675 (0.3497)	0.1985 (0.3232)	0.1986 (0.3233)	0.2598 (0.3462)	0.1985* (0.3233)
810	0.3308	0.1920 (0.2540)	0.1427 (0.2347)	0.1428 (0.2348)	0.2134 (0.2514)	0.1427* (0.2348)
815	0.1997	0.0566* (0.1533)	0.0616 (0.1417)	0.0616 (0.1417)	0.0763 (0.1518)	0.0616 (0.1417)
820	0.2607	0.1339 (0.2002)	0.1227 (0.1850)	0.1228 (0.1850)	0.1265 (0.1992)	0.1227* (0.1850)
825	0.2659	0.1228 (0.2041)	0.1042* (0.1887)	0.1043 (0.1887)	0.1372 (0.2021)	0.1043 (0.1887)
830	0.3386	0.1324* (0.2599)	0.1478 (0.2402)	0.1478 (0.2403)	0.1756 (0.2573)	0.1478 (0.2403)
835	0.2269	0.0766 (0.1742)	0.0703 (0.1610)	0.0703 (0.1610)	0.0977 (0.1724)	0.0703* (0.1610)
840	0.2379	0.1171 (0.1826)	0.1132 (0.1688)	0.1132 (0.1688)	0.1277 (0.1808)	0.1132* (0.1688)
845	0.0806	0.0161 (0.0619)	0.0114 (0.0572)	0.0114* (0.0572)	0.0266 (0.0612)	0.0114* (0.0572)
850	0.1776	0.1026 (0.1365)	0.0737 (0.1262)	0.0737 (0.1262)	0.0884 (0.1351)	0.0737* (0.1262)
855	0.0983	0.0696 (0.0755)	0.0610 (0.0698)	0.0610 (0.0698)	0.0775 (0.0747)	0.0609* (0.0698)
900	0.0676	0.0506 (0.0519)	0.0452* (0.0480)	0.0453 (0.0480)	0.0600 (0.0514)	0.0453 (0.0480)
905	0.0943	0.0692 (0.0726)	0.0583* (0.0673)	0.0745 (0.0674)	0.0723 (0.0732)	0.0604 (0.0673)

TABLE 6. A LISTING OF ANALYSIS ERROR VARIANCES COMPUTED USING THE DATA SETS INDICATED BY THE VARIOUS COLUMN HEADINGS. THE OBSERVATION VARIANCES ARE INCLUDED FOR COMPARISON PURPOSES. THE LOWEST VARIANCE FOR EACH ANALYSIS TIME IS MARKED (*). THIS INDICATES WHICH ANALYSIS DATA CONFIGURATION PRODUCED THE BEST RESULTS. THE VALUES IN THE PARENTHESES ARE THE MODELED VARIANCE ESTIMATES. THIS DATA IS FROM THE 2[illegible]TH OF MAY.

a multivariate environment. The surface-gauge univariate analysis produced as good a result as the aircraft determined multivariate analysis for the time 0845. Figures 49 thru 52 compare these results in the same manner as did Figures 24 thru 27 for the May 21st data set.

It should be noticed from Figures 50 and 52, that although the multivariate analysis approach gave the consistently better results, it is difficult to discern this fact from these graphs. A closer look at Table 6 will show that the multivariate analysis approach generally produced the better results only in the 4th or less significant digit, and for all practical purposes, the surface-gauge univariate analysis did just as well. The reason for this can be found in Table 5. It is seen that the A7 cross-correlation values, $CORR(X_1X_2) = .36$ and $CORR[X_1X_2] = .35$, were smaller than the A7 correlation value .41, which was found for the surface-gauge data set alone. The A7 parameter is an indication of the amount of signal, or the amount of usable information, within the respective data set, and the quantity, A7/(1-A7), is an estimate of the signal/noise ratio. In an observation network when the measurements taken reflect perfectly the signal present (a situation that is rather difficult to achieve) the value of the A7 parameter is exactly 1, and when measurements of white noise are made, the value of the A7 parameter is exactly 0. For the four storms analyzed, the values ranged from a low of .09 to a high of .87.

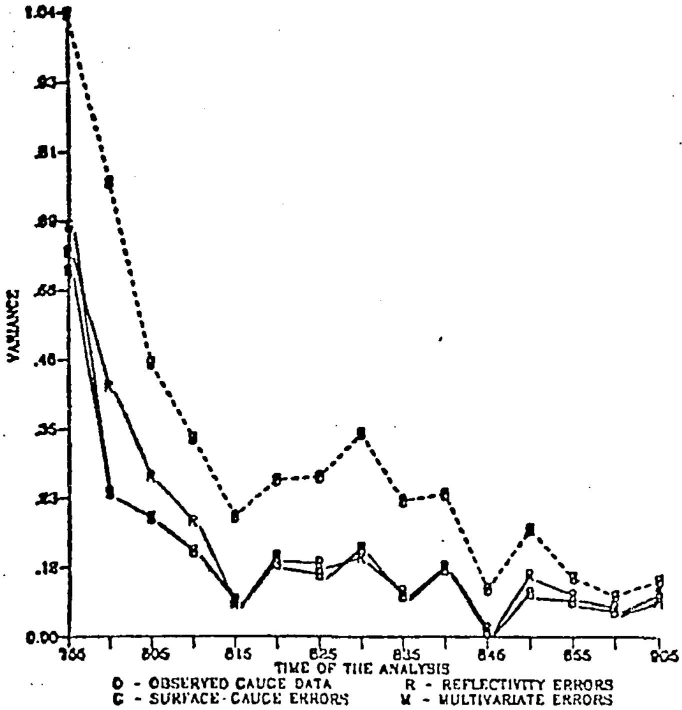

Fig. 49: A comparison of the computed error variances and the respective data variance for the May 25th, 0755-0905, observation set using the Marshall-Palmer Z-R conversion and procedure 1.

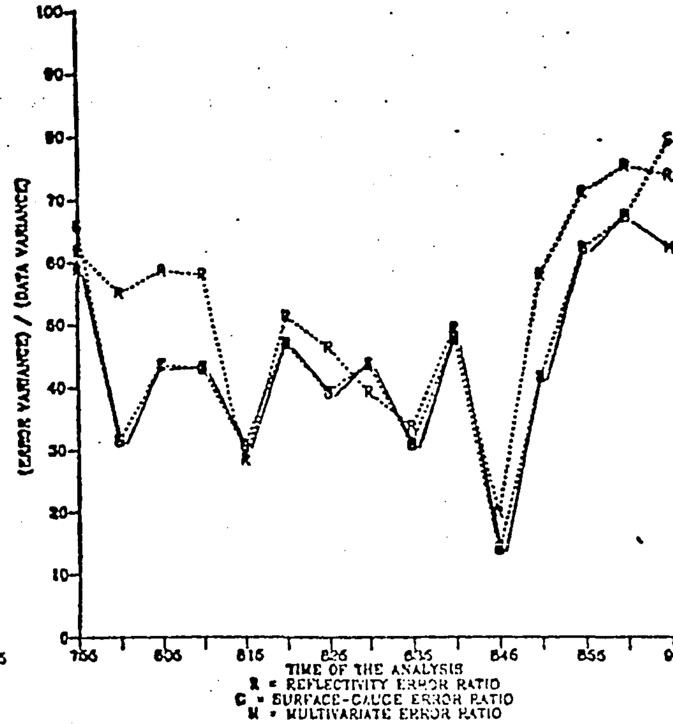

Fig. 50: The error variances of Figure 49 expressed as a percentage of the respective data variance.

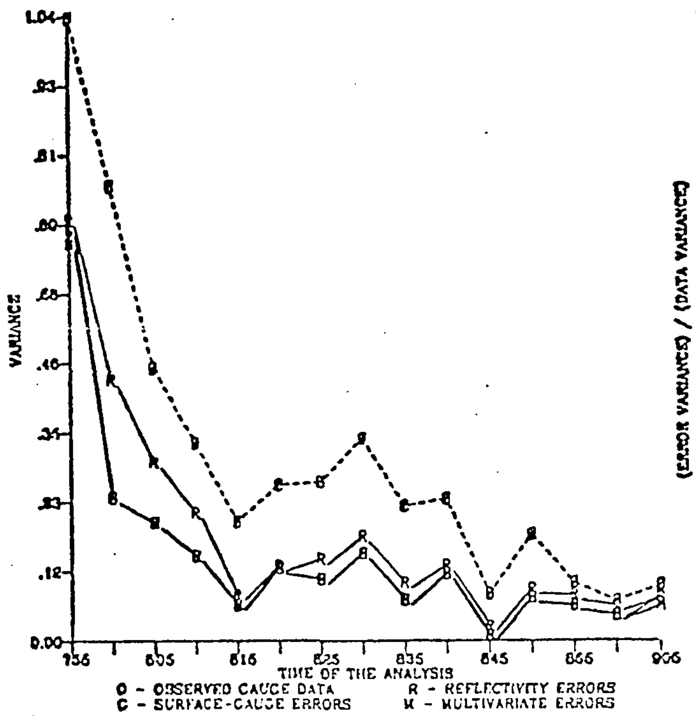

Fig. 51: A comparison of the computed error variances and the respective data variance for the May 25th, 0755-0905, observation set using the aircraft-determined Z-R conversion and procedure 1.

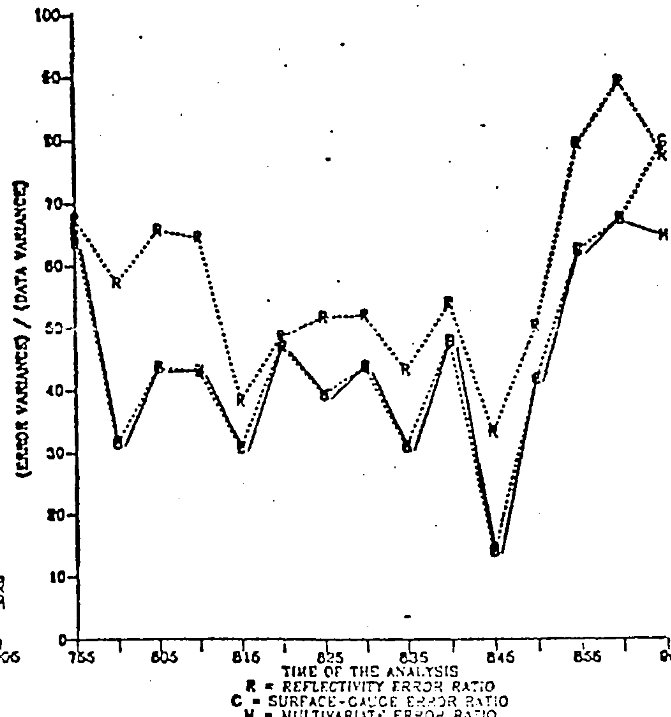

Fig. 52: The error variances of Figure 51 expressed as a percentage of the respective data variance.

The relatively smaller cross-correlation values for this particular storm imply that if an analysis of surface precipitation is desired, it is generally more beneficial to use a surface-gauge predictor in the regression equation as compared to a reflectivity derived predictor. However, as the spatial-temporal separation between a predictand and possible predictors becomes greater within the surface-gauge data set itself, it is possible for a radar bin position to have a relatively higher correlation value and to therefore be included in the predictor set. The reason that the multivariate analysis was almost always just a bit better than the surface-gauge univariate approach, was because occasionally a radar bin position was indeed closer to the predictand than the next available surface-gauge position. This radar bin position was then incorporated into the predictor set, a better regression equation resulted (as compared to the regression equation determined for the surface-gauge univariate approach), and the multivariate analysis produced slightly better results. The separation distance at which a predictor is chosen from one parameter set as opposed to another, is a function of the relative A7 values of the respective correlation and cross-correlation models.

For the May 21st data set, as seen in Table 1, the A7 cross-correlation values, .39 and .38 respectively, were larger than the A7 surface-gauge correlation value of

.35. There was therefore some advantage in including the reflectivity data in the predictor set, as opposed to using the surface-gauge data by itself. Based on this information, the multivariate analyses should have produced consistenly better results than did the univariate approach, which was indeed the case.

The results using procedure 2 are shown in Table 7. Employing the Marshall-Palmer Z-R conversion, the radar univariate configuration and the multivariate approach produced the best results 6 times each. For the times 0820 and 0850, the univariate configuration using the empirically derived Z-R conversion of the radar data produced the best results, while the surface-gauge data was best at 0855. Figures 53 thru 56 are a comparison of these results.

Notice from Figures 54 and 56 that either the multivariate approach or the radar reflectivity univariate configuration produced the consistently better analyses. Also note, that by withholding the entire time series surface observation vector from the set of possible predictors, the close correspondence between the multivariate analysis and the surface-gauge univariate results is no longer evident. This is because the A7 cross-correlation values were just slightly worse than the A7 surface-gauge correlation value, reducing the possible predictors that could be included from the surface-gauge data compensated for the relative weakness of the cross-correlation term. For a given pre-

ANALYSIS TIME	OBSERVATION VARIANCE	MARSHALL-PALMER Z-R RELATIONSHIP RADAR	MARSHALL-PALMER Z-R RELATIONSHIP GAUGES &RADAR	GAUGES	AIRCRAFT-DETERMINED Z-R RELATIONSHIP RADAR	AIRCRAFT-DETERMINED Z-R RELATIONSHIP GAUGES &RADAR
755	1.0420	0.6444	0.5437*	0.6470	0.7036	0.6291
		(0.8048)	(0.6688)	(0.6924)	(0.8188)	(0.6710)
800	0.7600	0.4184	0.3861*	0.4354	0.4353	0.4063
		(0.5834)	(0.4850)	(0.5014)	(0.5776)	(0.4875)
805	0.4556	0.2675	0.2241*	0.2768	0.2998	0.2463
		(0.3497)	(0.2907)	(0.3006)	(0.3462)	(0.2922)
810	0.3308	0.1920	0.1905*	0.2097	0.2134	0.1960
		(0.2540)	(0.2111)	(0.2183)	(0.2514)	(0.2122)
815	0.1997	0.0566*	0.0991	0.0946	0.0763	0.0946
		(0.1533)	(0.1275)	(0.1318)	(0.1518)	(0.1281)
820	0.2607	0.1339	0.1504	0.1908	0.1265*	0.1629
		(0.2002)	(0.1664)	(0.1720)	(0.1982)	(0.1672)
825	0.2659	0.1228	0.1133*	0.1474	0.1372	0.1378
		(0.2041)	(0.1697)	(0.1755)	(0.2021)	(0.1706)
830	0.3386	0.1324*	0.1628	0.1727	0.1756	0.1704
		(0.2599)	(0.2161)	(0.2234)	(0.2573)	(0.2172)
835	0.2269	0.0766*	0.0815	0.0963	0.0977	0.0909
		(0.1742)	(0.1448)	(0.1497)	(0.1724)	(0.1455)
840	0.2379	0.1171	0.0891*	0.1142	0.1277	0.0914
		(0.1826)	(0.1518)	(0.1570)	(0.1808)	(0.1526)
845	0.0806	0.0161*	0.0189	0.0171	0.0266	0.0182
		(0.0619)	(0.0514)	(0.0532)	(0.0612)	(0.0517)
850	0.1778	0.1026	0.1006	0.1029	0.0884*	0.1037
		(0.1365)	(0.1135)	(0.1173)	(0.1351)	(0.1141)
855	0.0983	0.0696	0.0703	0.0675*	0.0775	0.0799
		(0.0755)	(0.0628)	(0.0649)	(0.0747)	(0.0631)
900	0.0676	0.0506*	0.0550	0.0520	0.0600	0.0573
		(0.0519)	(0.0431)	(0.0446)	(0.0514)	(0.0434)
905	0.0943	0.0692*	0.0704	0.0694	0.0723	0.0711
		(0.0726)	(0.0605)	(0.0627)	(0.0732)	(0.0608)

TABLE 7. A LISTING OF ANALYSIS ERROR VARIANCES COMPUTED USING THE DATA SETS INDICATED BY THE VARIOUS COLUMN HEADINGS. THE OBSERVATION VARIANCES ARE INCLUDED FOR COMPARISON PURPOSES. THE LOWEST VARIANCE FOR EACH ANALYSIS TIME IS MARKED (*). THIS INDICATES WHICH ANALYSIS DATA CONFIGURATION PRODUCED THE BEST RESULTS. THE VALUES IN THE PARENTHESES ARE THE MODELED VARIANCE ESTIMATES. THIS DATA IS FROM THE 25TH OF MAY.

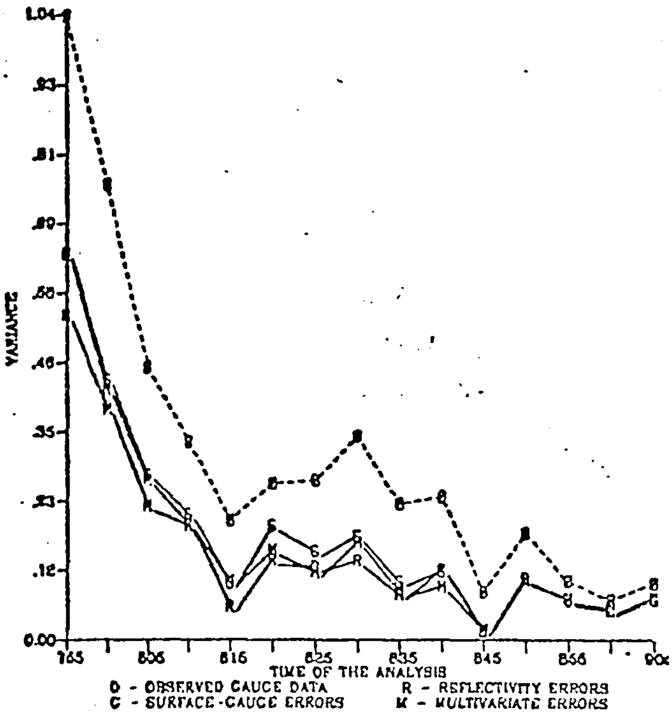

Fig. 53: A comparison of the computed error variances and the respective data variance for the May 25th, 0755-0905, observation set using the Marshall-Palmer Z-R conversion and procedure 2.

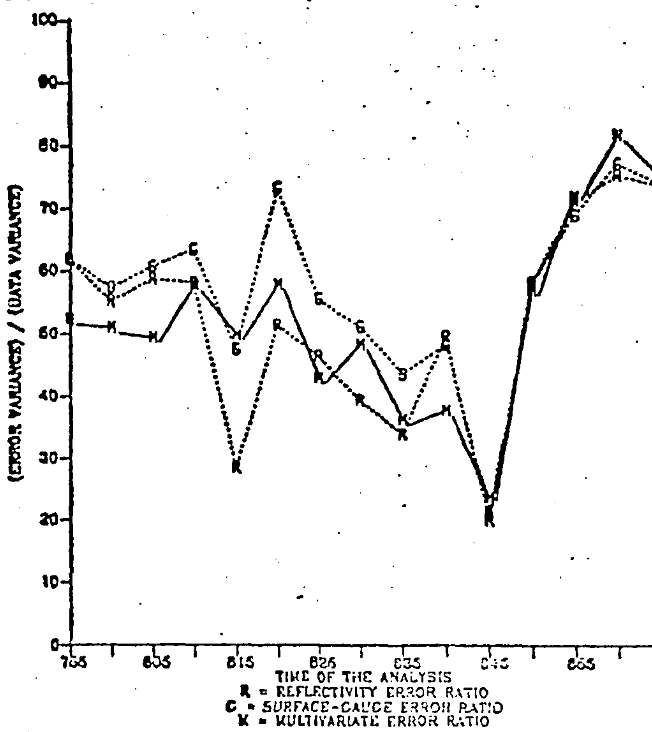

Fig. 54: The error variances fo Figure 53 expressed as a percentage of the respective data variance.

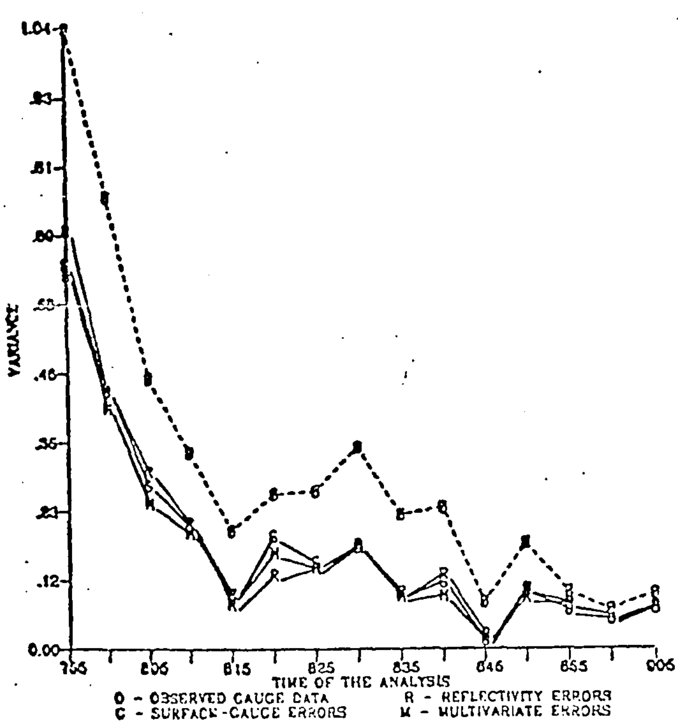

Fig. 55: A comparison of the computed error variances and the respective data variance for the May 25th, 0755-0905, observation set using the aircraft-determined Z-R conversion and procedure 2.

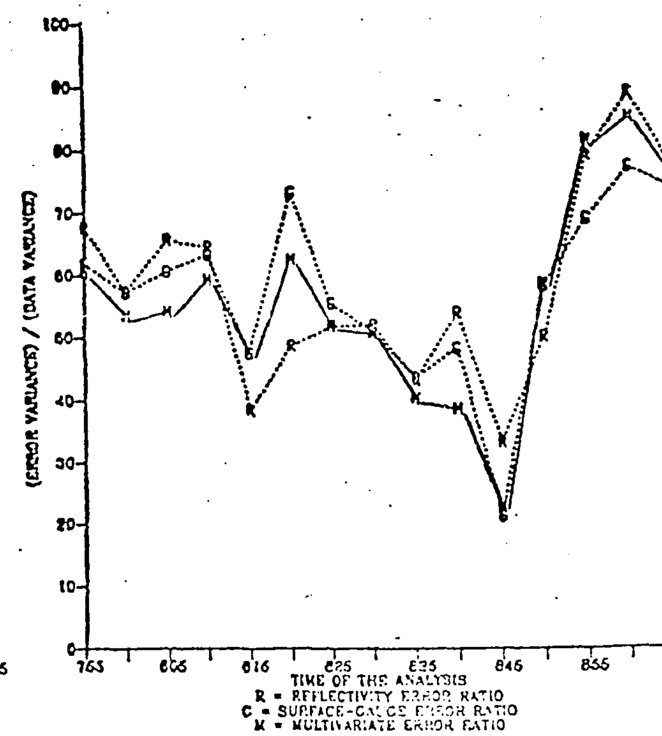

Fig. 56: The error variances of Figure 55 expressed as a percentage of the respective data variance.

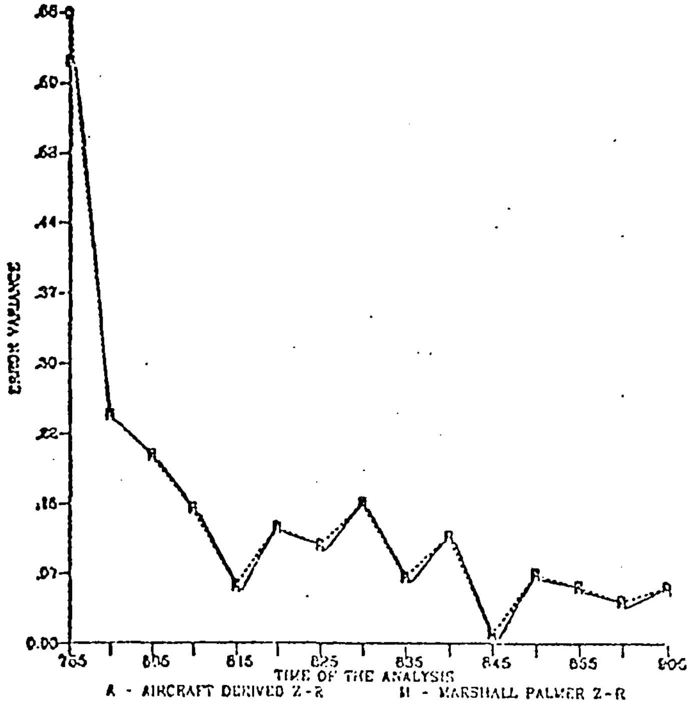

Fig. 57: A comparison of the effect of the different Z-R relationships on the analysis of surface precipitation for May 25th, 0755-0905, when procedure 1 was used.

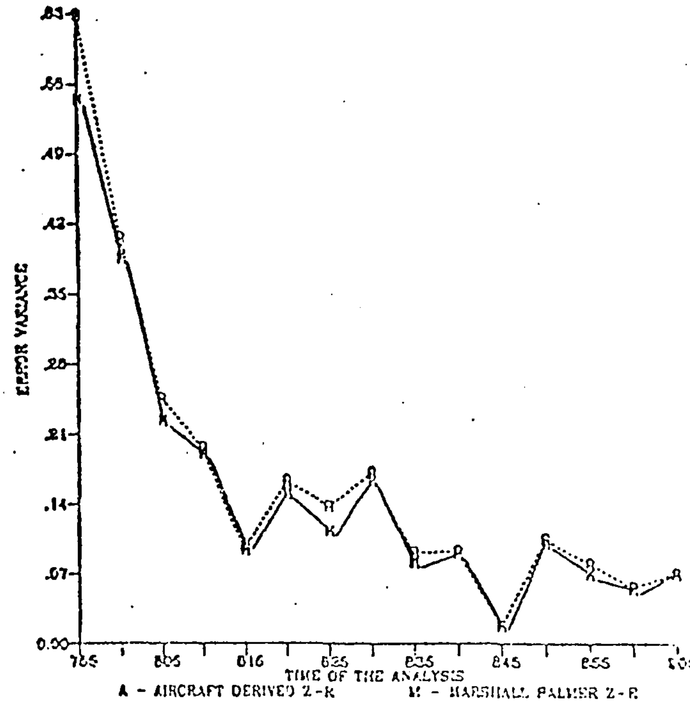

Fig. 58: A comparison of the effect of the different Z-R relationships on the analysis of surface precipitation for May 25th. 0755-0905, when procedure 2 was used.

dictand then, the worth of a reflectivity estimate increased, and a greater number of these were used in determining the regression equations, resulting in different regression weights and different analyzed fields.

Figures 57 and 58 compare the relative worth of the different Z-R relationships, as was done for the May 21st data set. The first figure corresponds to procedure 1 results; Figure 58 corresponds to procedure 2 results. As before, no predominant difference exists between these two Z-R conversions when used in a multivariate analysis configuration.

C. 25 May, 1974, 0935-1010

The deduced correlation and cross-correlation model parameters for this particular data set are shown in Table 8. Parameters A8, A9, A10, and A11 are in nautical miles, while parameter A14 is in minutes.

	POSITIVE X		POSITIVE Y		NEGATIVE X		NEGATIVE Y		TIME	
	A1	A8	A2	A9	A3	A10	A4	A11	A14	A7
CORR{X_1X_1}	.01	8.9	.09	9.0	-.01	9.0	.03	8.9	25.0	.36
CORR(X_2X_2)	-.04	9.0	-.14	8.9	-.00	9.0	-.56	8.9	25.0	.38
CORR(X_1X_2)	.01	9.0	-.03	8.9	.05	9.0	-.04	9.0	25.0	.14
CORR(X_2X_1)	-.02	9.0	.15	9.0	.06	9.0	-.11	9.0	24.9	.12
CORR[X_2X_2]	-.03	9.0	-.17	9.0	-.02	9.0	-.09	9.0	25.0	.37
CORR[X_1X_2]	.03	9.0	-.00	9.0	.02	9.0	-.05	9.0	25.0	.15
CORR[X_2X_1]	.04	9.0	.06	9.0	.08	9.0	-.04	9.0	24.9	.09

TABLE 8.--The signal-plus-noise model parameters for 25 May, 0935-1010. The parenthesized model designation, CORR(X_iX_j), denotes the parameters obtained from the Marshall-Palmer Z-R conversion, while the bracketed model designation, CORR[X_iX_j], denotes the parameters obtained from the empirically derived Z-R conversion.

Grid analyses for this storm system are shown in Figures 59 thru 66. In general the rainfall rate for this data set was between 0.1 and 1.0 inches/hour. Isohyetal lines correspond to rainfall increments of ≈ .06 mm.

The data set consisted of only 8 observation times, with the majority of the precipitation occuring at 0935 and 0940. The temporal size of the data set was so short, and the number of sensors involved in the measurements were so few, that accurate signal-plus-noise models of the precipitation field were difficult to obtain. If the temporal length

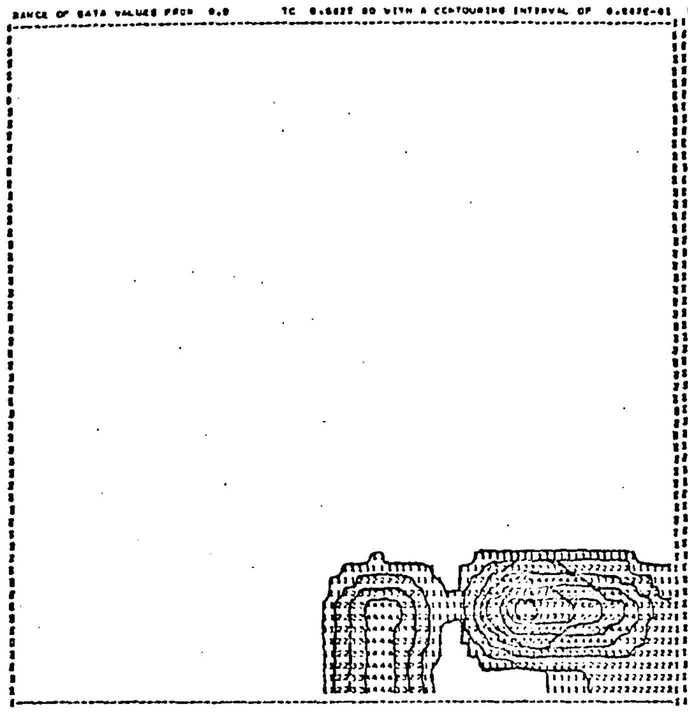

Fig. 59: 0935 - 25 May.

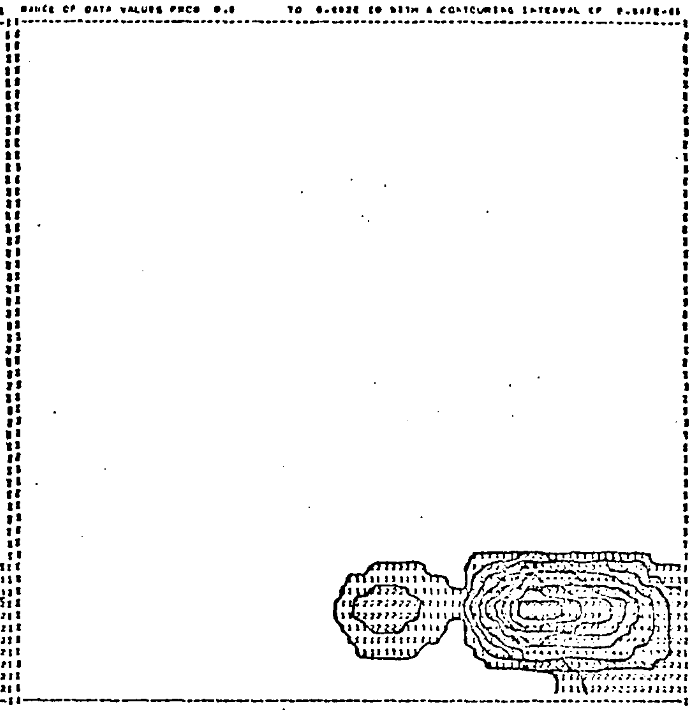

Fig. 60: 0940 - 25 May.

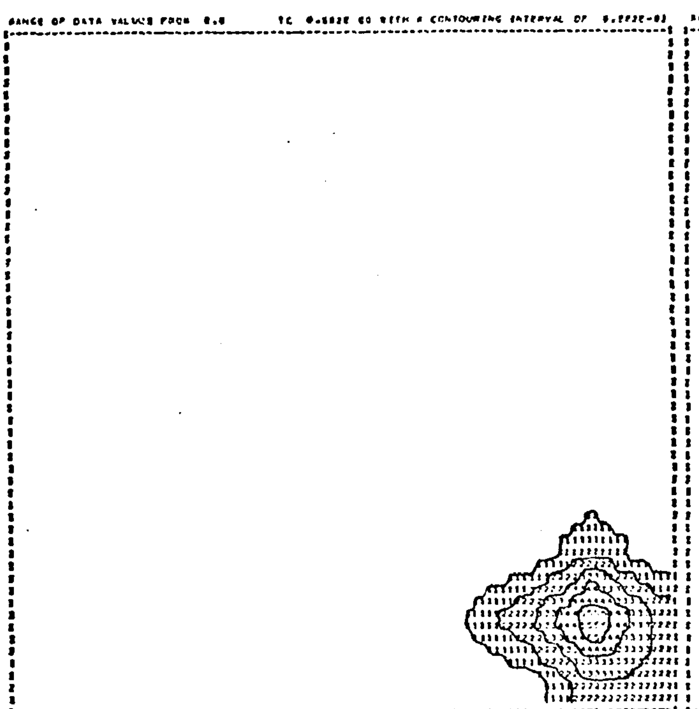

Fig. 61: 0945 - 25 May.

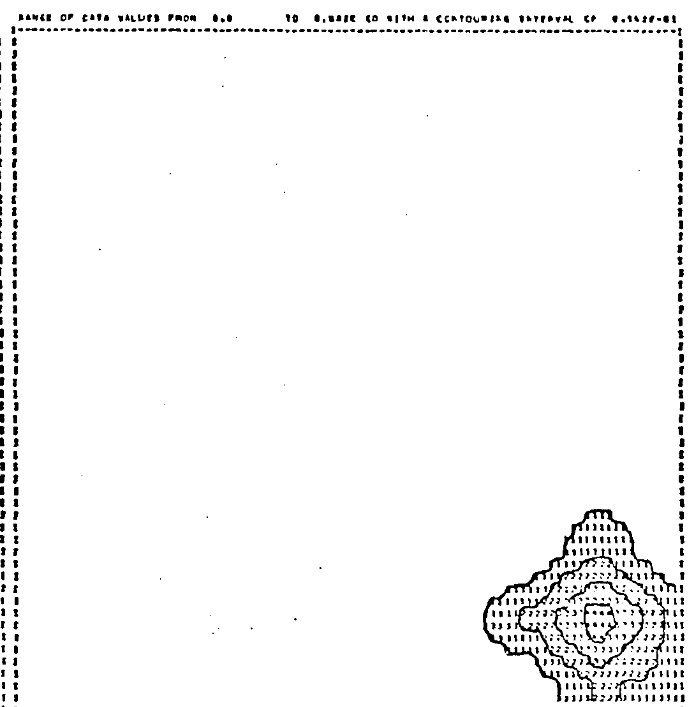

Fig. 62: 0950 - 25 May.

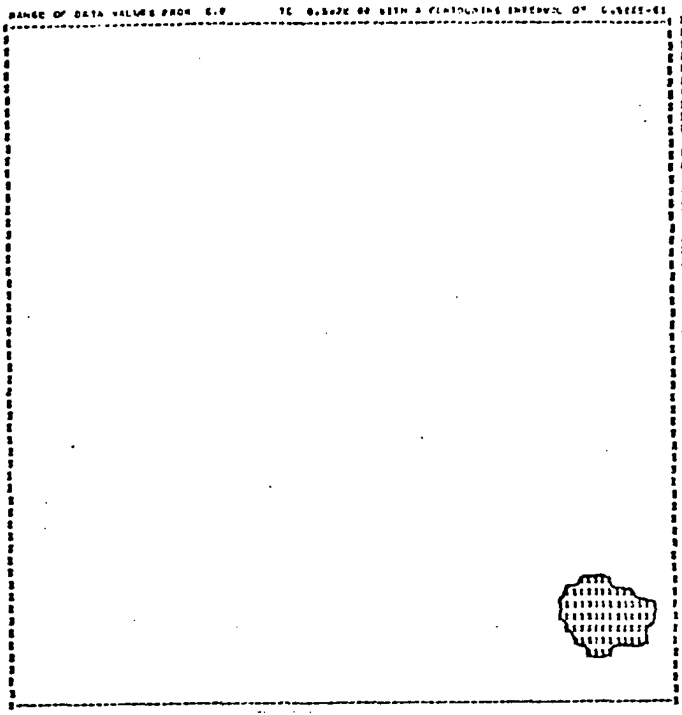

Fig. 63: 0955 - 25 May.

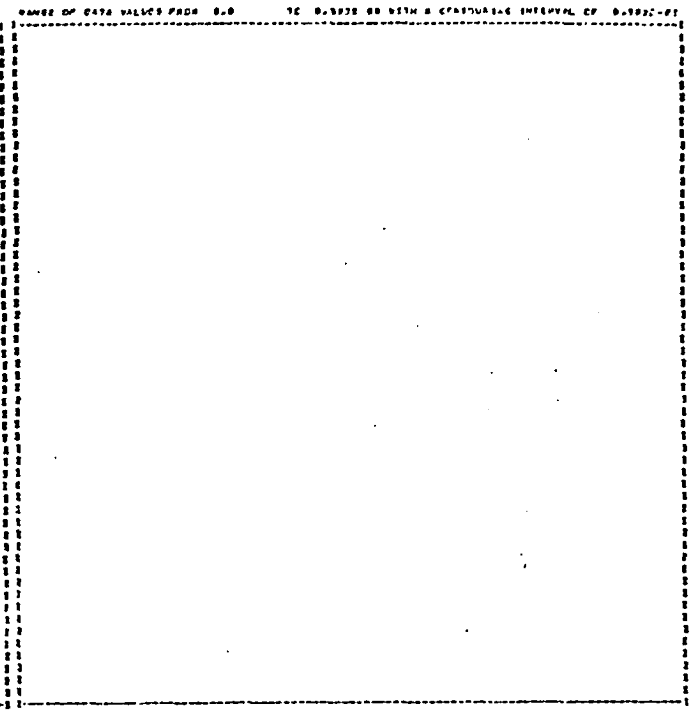

Fig. 64: 1000 - 25 May.

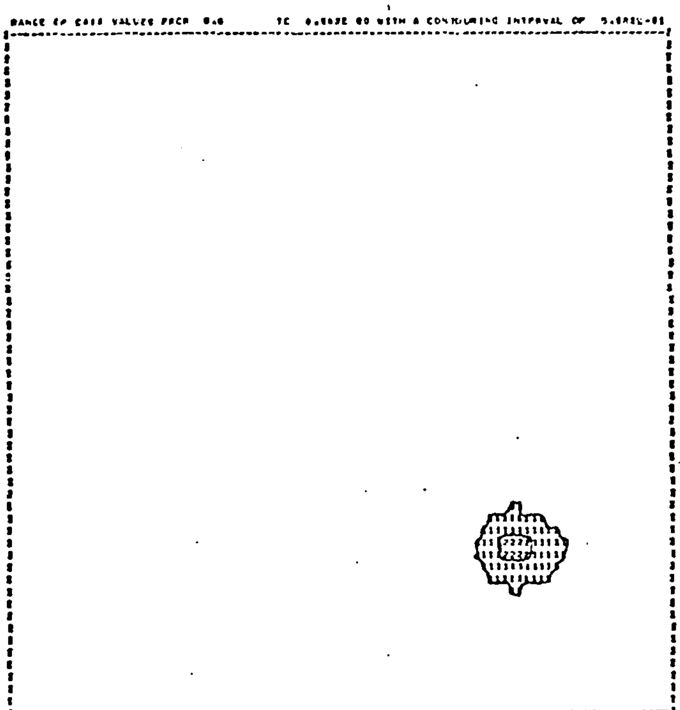

Fig. 65: 1005 - 25 May.

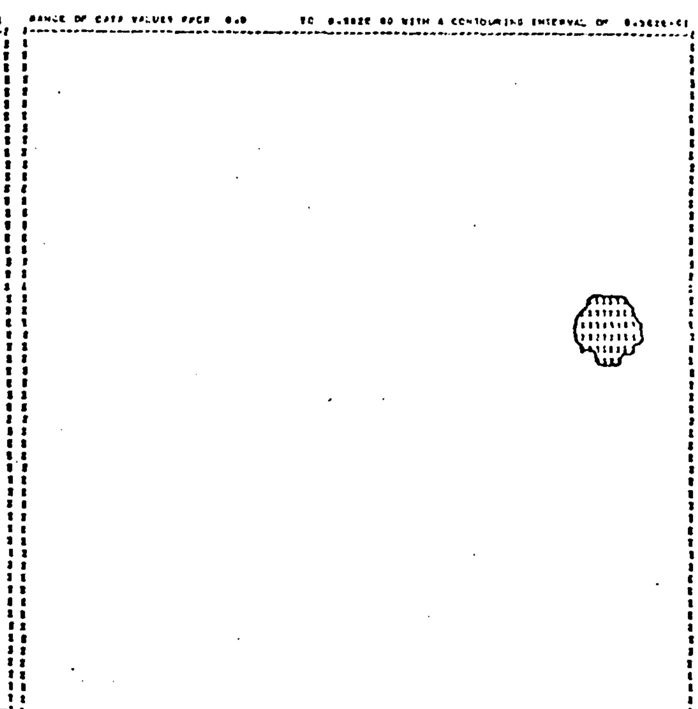

Fig. 66: 1010 - 25 May.

of the observation set is short, but the number of stations involved in the measurement of the phenomenon is relatively large, or if the number of stations is small, but the temporal length of the observation set is relatively long, then accurate model estimates are possible. The size of the sensor network and/or the length of the observation set necessary for determining these accurate model parameters is a function of the phenomenon being investigated, the characteristics of the storm situation, the network spacing, the sampling rate, etc. For this particular case the data set was border line. Of the four storms studied, this one had the shortest recorded life span, was confined to the smallest area, had the smallest rainfall rates, and subsequently produced the lowest correlation and cross-correlation values.

Table 9 shows the results obtained when using procedure 1. It should be noticed that the multivariate analyses for both Z-R relationships were identical to the results obtained by using the surface-gauge univariate configuration, for all but the last time analyzed. By comparing the A7 correlation and cross-correlation values of Table 9 respective of the discussion in section B of this chapter, the reason for this development is obvious. The strength of the cross-correlation structure was so weak compared to the correlation within the surface-gauge data set itself, that the reflectivity data was never included in the regression equation even in a multivariate environment.

ANALYSIS TIME	OBSERVATION VARIANCE	MARSHALL-PALMER Z-R RELATIONSHIP RADAR	MARSHALL-PALMER Z-R RELATIONSHIP GAUGES &RADAR	GAUGES	AIRCRAFT-DETERMINED Z-R RELATIONSHIP RADAR	AIRCRAFT-DETERMINED Z-R RELATIONSHIP GAUGES &RADAR
935	0.0298	0.0263 (0.0284)	0.0204* (0.0227)	0.0204* (0.0227)	0.0251 (0.0282)	0.0204* (0.0227)
940	0.0291	0.0237 (0.0278)	0.0144* (0.0221)	0.0144* (0.0221)	0.0225 (0.0276)	0.0144* (0.0221)
945	0.0064	0.0040 (0.0061)	0.0020* (0.0048)	0.0020* (0.0048)	0.0035 (0.0060)	0.0020* (0.0048)
950	0.0051	0.0030 (0.0049)	0.0023* (0.0039)	0.0023* (0.0039)	0.0035 (0.0048)	0.0023* (0.0039)
955	0.0006	0.0005 (0.0005)	0.0004* (0.0004)	0.0004* (0.0004)	0.0005 (0.0005)	0.0004* (0.0004)
1000	0.0004	0.0004 (0.0004)	0.0004* (0.0003)	0.0004* (0.0003)	0.0004 (0.0004)	0.0004* (0.0003)
1005	0.0035	0.0034 (0.0033)	0.0031* (0.0027)	0.0031* (0.0027)	0.0035 (0.0033)	0.0031* (0.0027)
1010	0.0028	0.0028* (0.0027)	0.0028 (0.0022)	0.0028 (0.0022)	0.0028 (0.0027)	0.0028 (0.0022)

TABLE 9. A LISTING OF ANALYSIS ERROR VARIANCES COMPUTED USING THE DATA SETS INDICATED BY THE VARIOUS COLUMN HEADINGS. THE OBSERVATION VARIANCES ARE INCLUDED FOR COMPARISON PURPOSES. THE LOWEST VARIANCE FOR EACH ANALYSIS TIME IS MARKED (*). THIS INDICATES WHICH ANALYSIS DATA CONFIGURATION PRODUCED THE BEST RESULTS. THE VALUES IN THE PARENTHESES ARE THE MODELED VARIANCE ESTIMATES. THIS DATA IS FROM THE 25TH OF MAY.

The analyses results, when procedure 2 was used, are shown in Table 10. Once again the multivariate analyses for both Z-R relationships were identical to the results obtained using the surface-gauge univariate configuration. Notice that for all the analyses, the radar data alone generally produced as good or better results, no matter which Z-R conversion was used.

D. 25 May, 1974, 1145-1310

The deduced correlation and cross-correlation model parameters for this particular data set are shown in Table 11. Parameters A8, A9, A10, and A11 are in nautical miles, while parameter A14 is in minutes.

	POSITIVE X		POSITIVE Y		NEGATIVE X		NEGATIVE Y		TIME	
	A1	A8	A2	A9	A3	A10	A4	A11	A14	A7
CORR(X_1X_1)	-.01	6.8	-.03	10.6	-.16	4.8	-.06	8.4	18.6	.51
CORR(X_2X_2)	-.07	6.0	-.04	7.0	-.14	3.0	-.02	4.6	10.4	.87
CORR(X_1X_2)	-.02	5.2	.00	9.6	-.17	5.2	-.04	5.9	15.8	.58
CORR(X_2X_1)	.04	7.6	-.01	8.6	-.13	3.7	-.04	6.6	14.9	.55
CORR[X_2X_2]	-.04	5.4	.02	7.5	-.21	4.3	-.03	4.5	11.0	.87
CORR[X_1X_2]	-.04	5.4	-.03	10.3	-.18	4.8	-.03	5.8	16.7	.57
CORR[X_2X_1]	.05	7.8	-.01	9.5	-.12	3.6	-.05	6.3	13.8	.55

TABLE 11.--The signal-plus-noise model parameters for 25 May, 1145-1310. The parenthesized model designation, CORR(X_iX_j), denotes the parameters obtained from the Marshall-Palmer Z-R conversion, while the bracketed model designation, CORR[X_iX_j], denotes the parameters obtained from the empirically derived Z-R conversion.

Grid analyses for this storm are shown in Figures 67 thru 84. In general the rainfall rate for this data set

ANALYSIS TIME	OBSERVATION VARIANCE	MARSHALL-PALMER Z-R RELATIONSHIP RADAR	MARSHALL-PALMER Z-R RELATIONSHIP GAUGES &RADAR	GAUGES	AIRCRAFT-DETERMINED Z-R RELATIONSHIP RADAR	AIRCRAFT-DETERMINED Z-R RELATIONSHIP GAUGES &RADAR
935	0.0298	0.0263	0.0282	0.0282	0.0251*	0.028[illegible]
		(0.0284)	(0.0212)	(0.0212)	(0.0282)	(0.0212)
940	0.0291	0.0237	0.0230	0.0230	0.0225*	0.0230
		(0.0278)	(0.0206)	(0.0206)	(0.0276)	(0.0206)
945	0.0064	0.0040	0.0045	0.0045	0.0035*	0.0045
		(0.0061)	(0.0045)	(0.0045)	(0.0060)	(0.0045)
950	0.0051	0.0030*	0.0035	0.0035	0.0035	0.0035
		(0.0049)	(0.0036)	(0.0036)	(0.0048)	(0.0036)
955	0.0006	0.0005	0.0005	0.0005	0.0005*	0.0005
		(0.0005)	(0.0004)	(0.0004)	(0.0005)	(0.0004)
1000	0.0004	0.0004*	0.0004	0.0004	0.0004	0.0004
		(0.0004)	(0.0003)	(0.0003)	(0.0004)	(0.0003)
1005	0.0035	0.0034*	0.0037	0.0037	0.0035	0.0037
		(0.0033)	(0.0025)	(0.0025)	(0.0033)	(0.0025)
1010	0.0028	0.0028*	0.0031	0.0031	0.0028	0.0031
		(0.0027)	(0.0020)	(0.0020)	(0.0027)	(0.0020)

TABLE 10. A LISTING OF ANALYSIS ERROR VARIANCES COMPUTED USING THE DATA SETS INDICATED BY THE VARIOUS COLUMN HEADINGS. THE OBSERVATION VARIANCES ARE INCLUDED FOR COMPARISON PURPOSES. THE LOWEST VARIANCE FOR EACH ANALYSIS TIME IS MARKED (*). THIS INDICATES WHICH ANALYSIS DATA CONFIGURATION PRODUCED THE BEST RESULTS. THE VALUES IN THE PARENTHESES ARE THE MODELED VARIANCE ESTIMATES. THIS DATA IS FROM THE 25TH OF MAY.

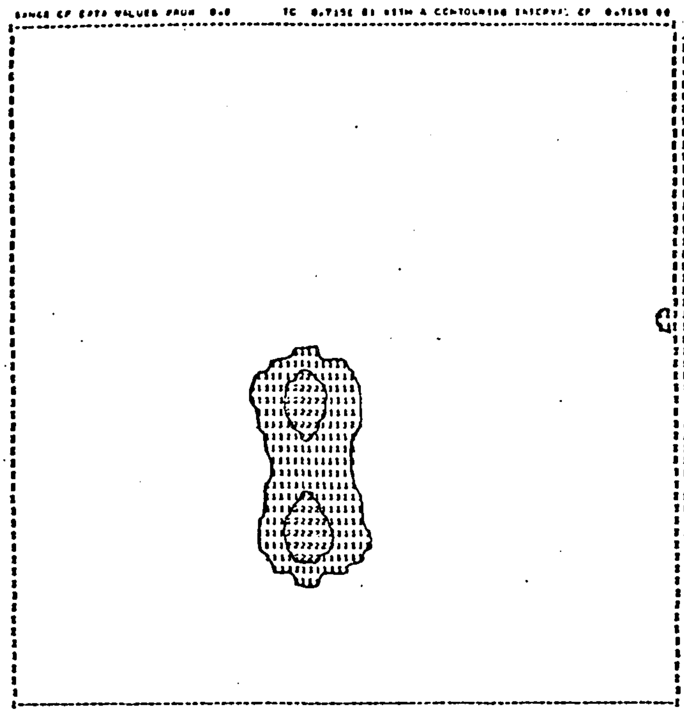

Fig. 67: 1145 - 25 May.

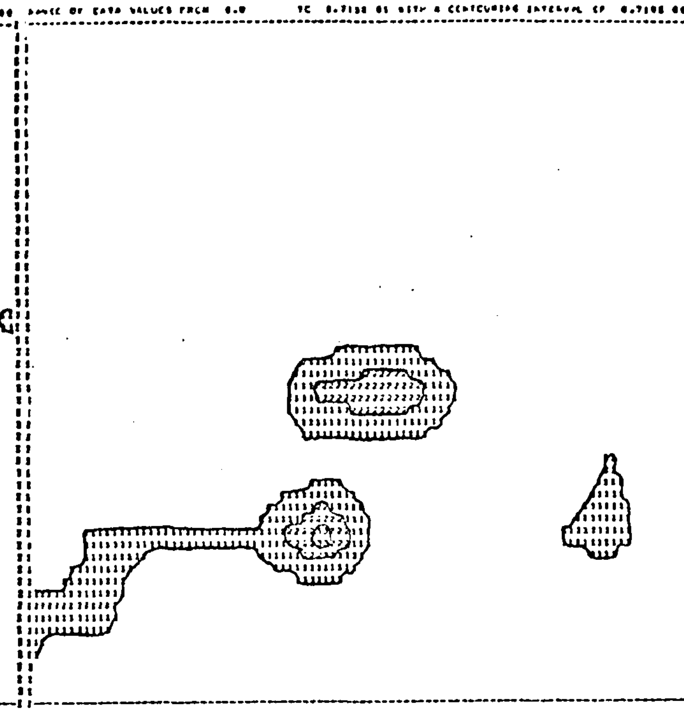

Fig. 68: 1150 - 25 May.

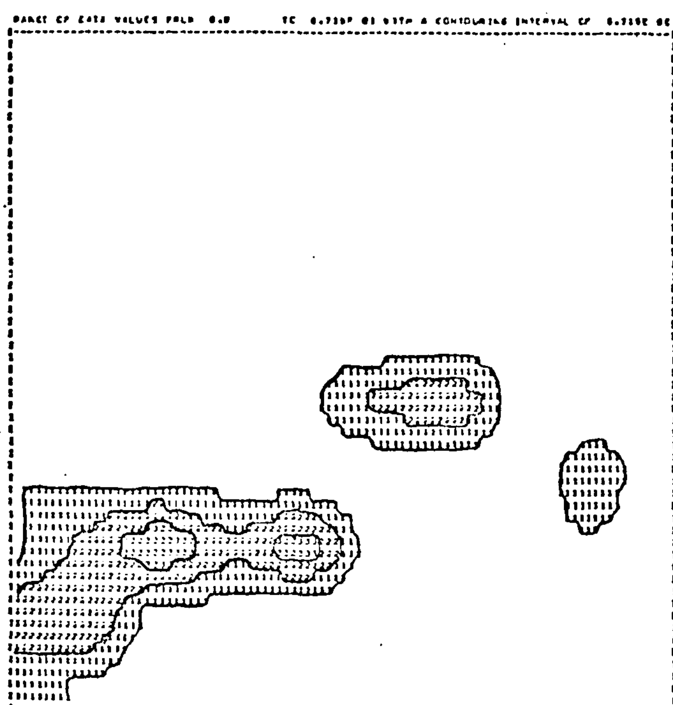

Fig. 69: 1155 - 25 May.

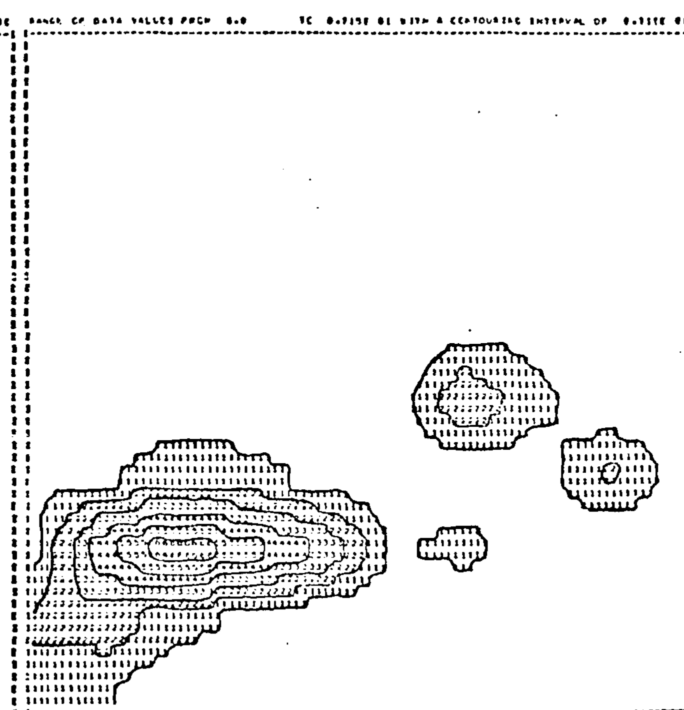

Fig. 70: 1200 - 25 May.

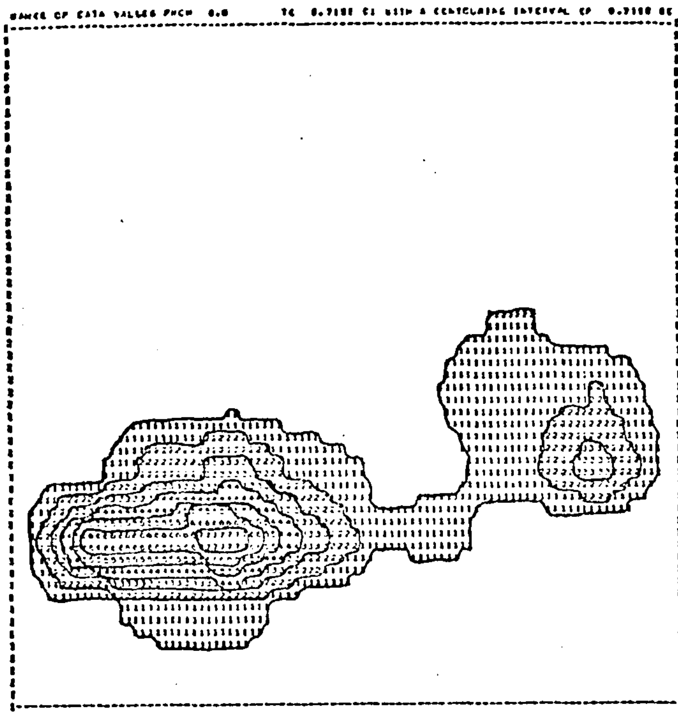

Fig. 71: 1205 - 25 May.

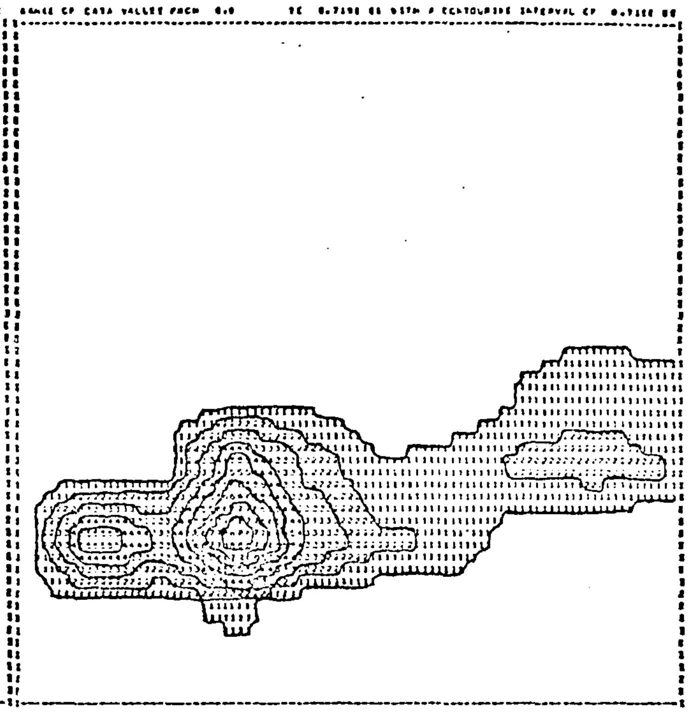

Fig. 72: 1210 - 25 May.

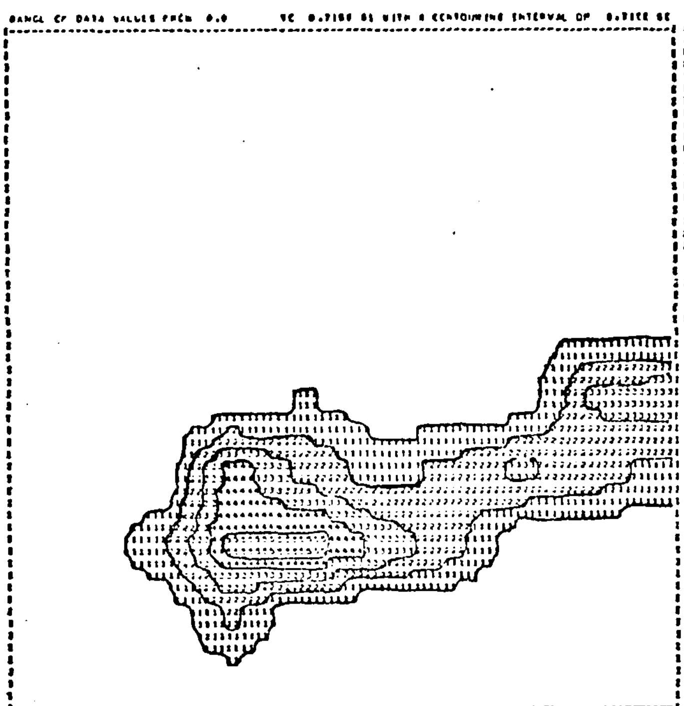

Fig. 73: 1215 - 25 May.

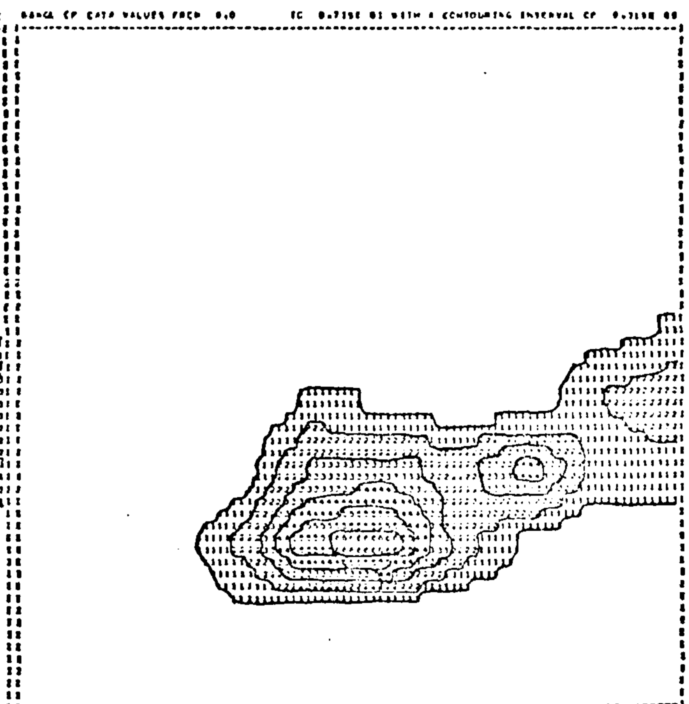

Fig. 74: 1220 - 25 May.

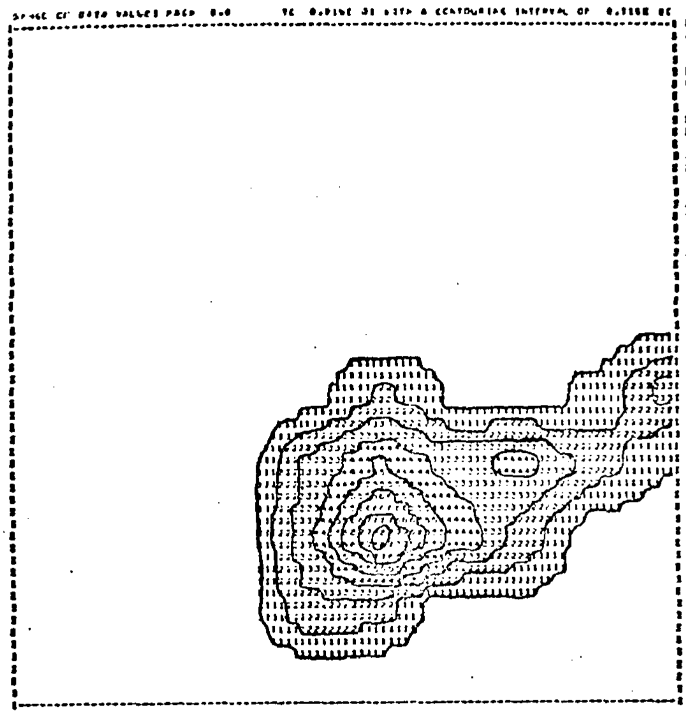

Fig. 75: 1225 - 25 May.

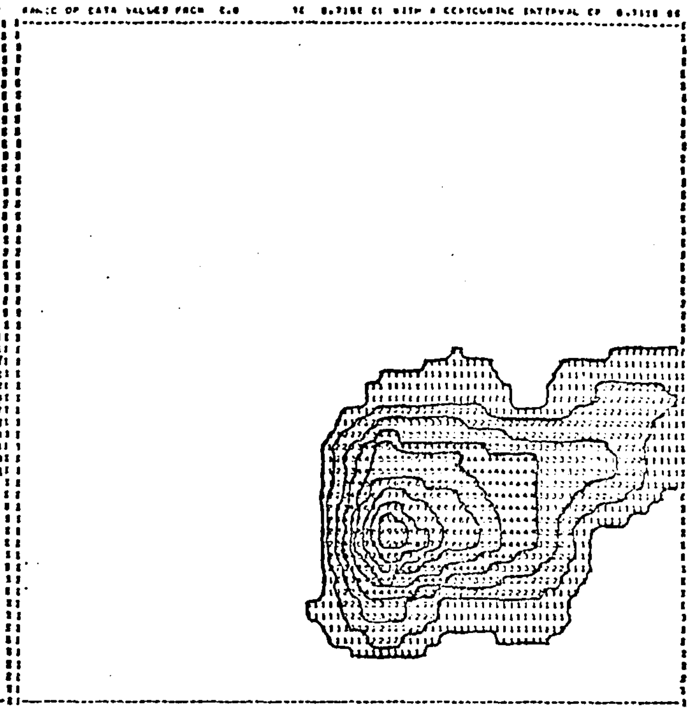

Fig. 76: 1230 - 25 May.

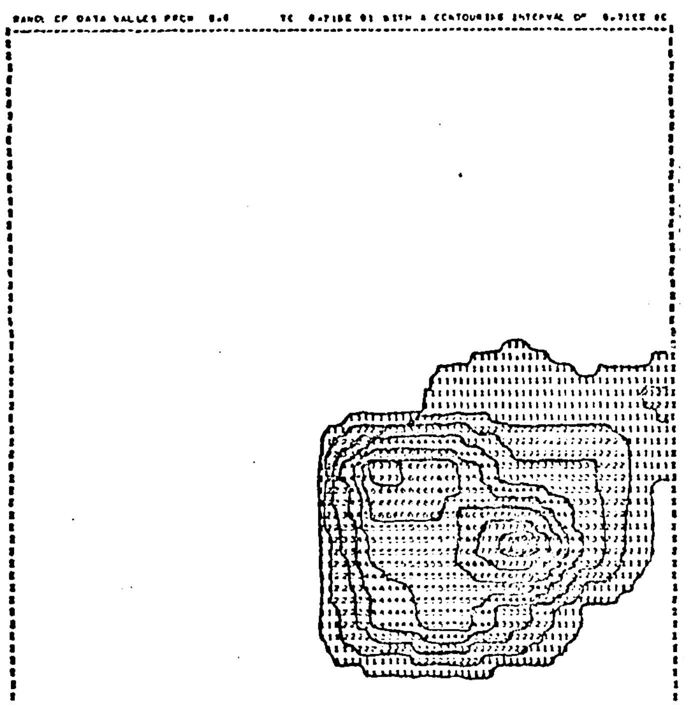

Fig. 77: 1235 - 25 May.

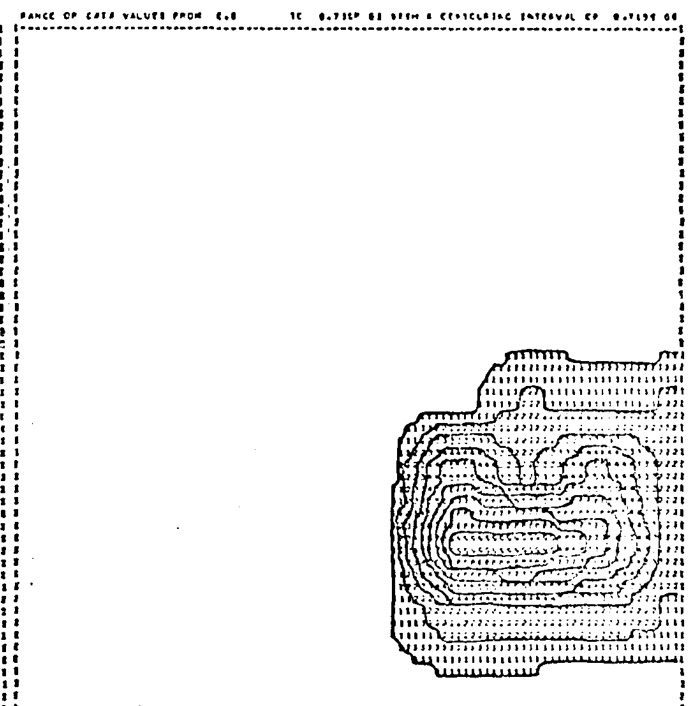

Fig. 78: 1240 - 25 May.

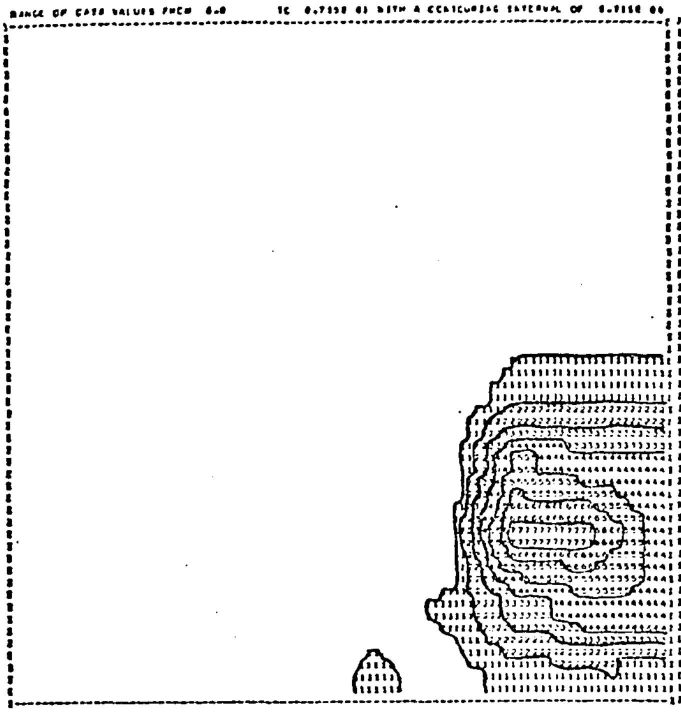

Fig. 79: 1245 - 25 May.

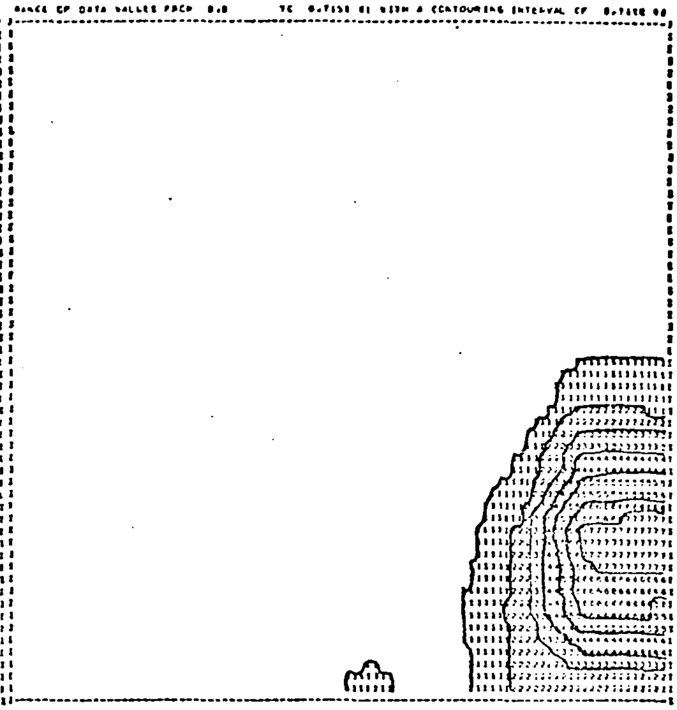

Fig. 80: 1250 - 25 May.

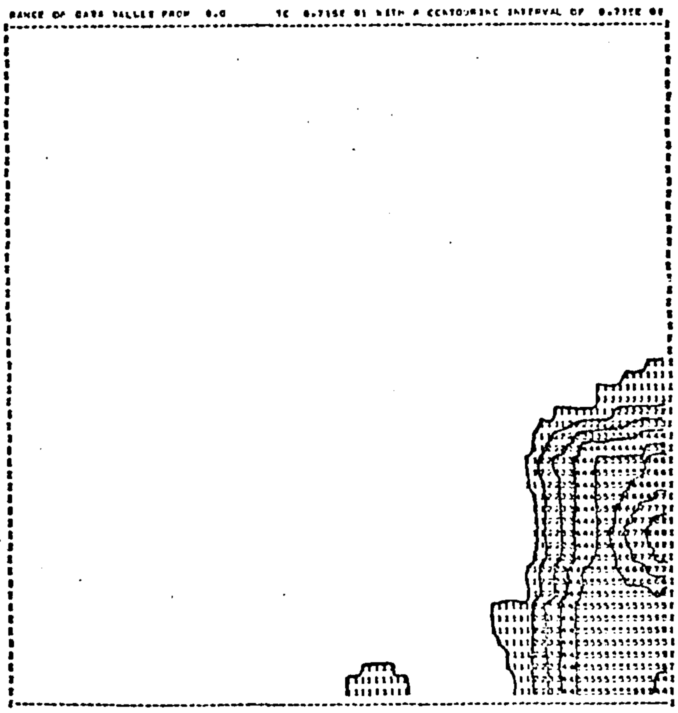

Fig. 81: 1255 - 25 May.

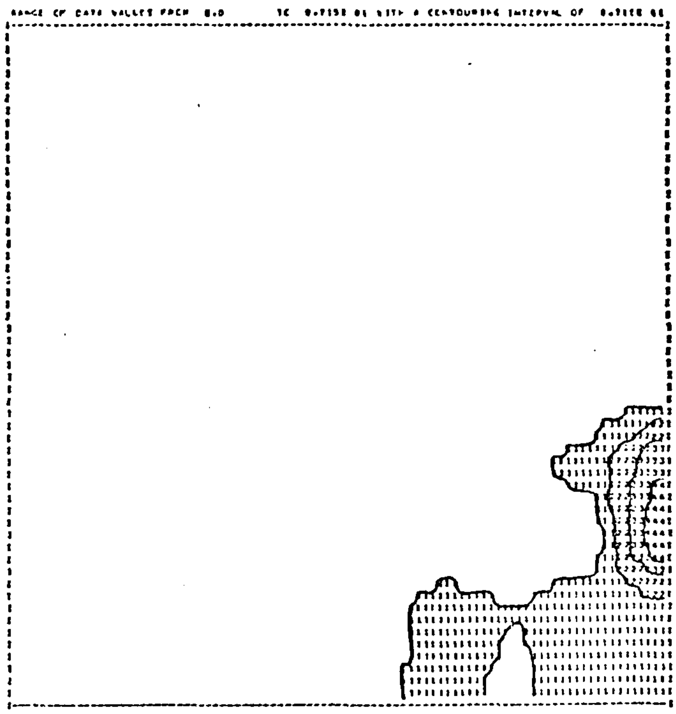

Fig. 82: 1300 - 25 May.

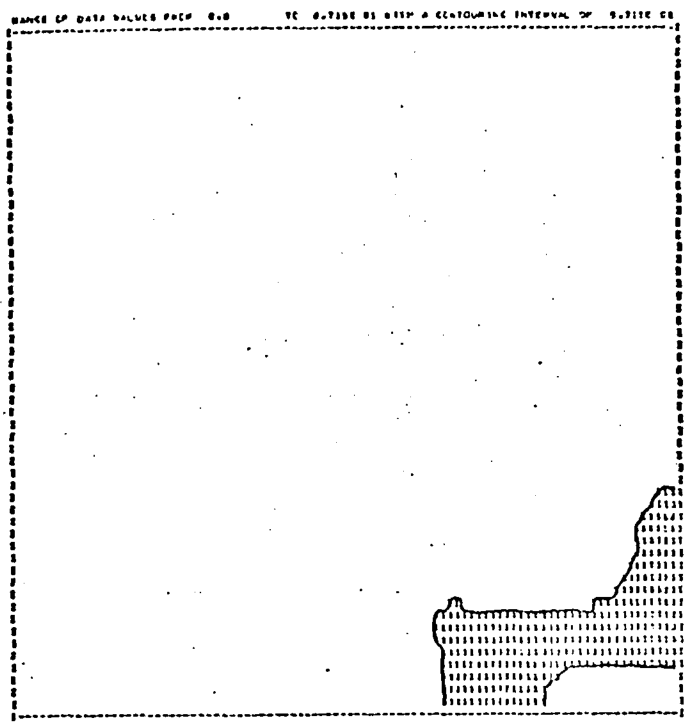

Fig. 83: 1305 - 25 May.

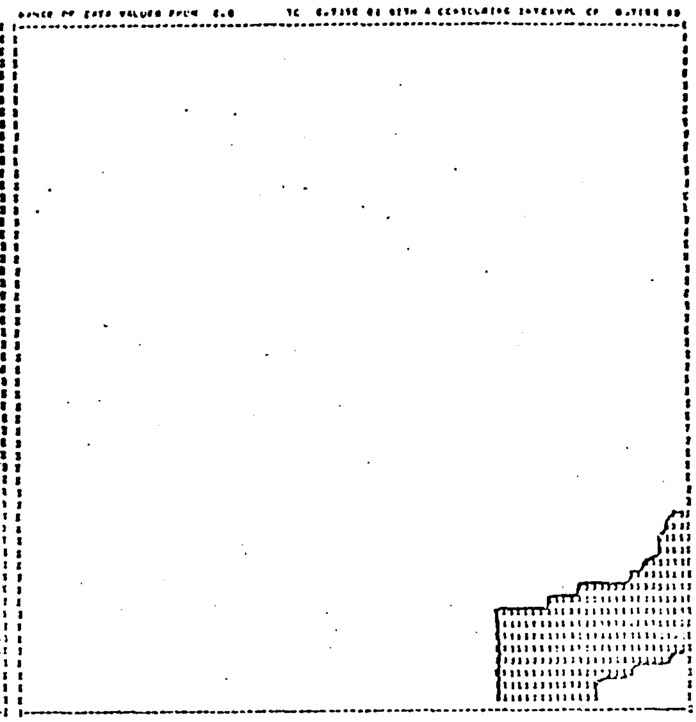

Fig. 84: 1310 - 25 May.

varied from 0.5 inches/hour to as much as 7.0 inches/hour. The higher rates invariably were isolated anomalies. For a specific time the rainfall rate measured at the surface varied by as much as 5.0 inches/hour from the rainfall rate estimated at the horizontally collocated radar bin. Nevertheless, this storm exhibited the strongest correlation and cross-correlation of the four storms examined, a fact reflected in the magnitudes of the A7 parameters for the respective signal-plus-noise models. Isohyetal lines in Figures 67 thru 84 correspond to rainfall increments of ≈ .72 mm.

Table 12 shows the results obtained when using procedure 1. It is seen that for the 18 analyzed times, the

ANALYSIS TIME	OBSERVATION VARIANCE	MARSHALL-PALMER Z-R RELATIONSHIP RADAR	MARSHALL-PALMER Z-R RELATIONSHIP GAUGES &RADAR	GAUGES	AIRCRAFT-DETERMINED Z-R RELATIONSHIP RADAR	AIRCRAFT-DETERMINED Z-R RELATIONSHIP GAUGES &RADAR
1145	0.5952	0.6034 (0.3344)	0.5205 (0.3443)	0.5421 (0.3297)	0.6468 (0.3358)	0.5090* (0.3478)
1150	0.3783	0.3553 (0.2107)	0.2104 (0.2168)	0.2497 (0.2066)	0.3332 (0.2114)	0.1843* (0.2205)
1155	0.7263	0.5813 (0.4045)	0.3368 (0.4164)	0.3509 (0.3967)	0.5271 (0.4059)	0.2935* (0.4233)
1200	1.0320	0.6220 (0.5746)	0.4693* (0.5914)	0.5920 (0.5634)	0.6048 (0.5765)	0.4829 (0.6013)
1205	1.0710	0.6849 (0.5968)	0.5846 (0.6142)	0.4624 (0.5851)	0.4377 (0.5987)	0.3982* (0.6245)
1210	1.4490	0.8169 (0.8068)	0.4382 (0.8304)	0.4970 (0.7911)	0.5111 (0.8095)	0.4166* (0.8443)
1215	2.4860	1.3933 (1.3848)	1.4670 (1.4253)	1.1336* (1.3578)	1.3680 (1.3893)	1.1903 (1.4491)
1220	3.6660	1.3468 (2.0421)	1.0717 (2.1018)	1.1607 (2.0023)	1.6083 (2.0487)	1.0398* (2.1369)
1225	3.4080	1.3688 (1.8983)	1.1118 (1.9538)	0.9491* (1.8613)	1.5784 (1.9044)	1.1728 (1.9864)
1230	2.2910	1.5884 (1.2759)	1.2997 (1.3132)	1.0243* (1.2510)	1.5457 (1.2801)	1.2569 (1.3352)
1235	3.5570	3.3646 (1.9813)	2.5768 (2.0392)	1.8965* (1.9426)	2.9928 (1.9877)	2.1738 (2.0733)
1240	2.7660	1.3941 (1.5404)	1.3115 (1.5855)	1.3985 (1.5104)	1.3235 (1.5454)	1.2118* (1.6119)
1245	2.7020	0.9713 (1.5049)	0.9049* (1.5489)	1.0897 (1.4756)	1.1325 (1.5098)	1.0130 (1.5748)
1250	4.1400	2.0165 (2.3060)	1.3308* (2.3735)	1.3568 (2.2611)	1.7685 (2.3135)	1.3895 (2.4131)
1255	2.6080	1.4195 (1.4524)	1.1532* (1.4949)	1.1545 (1.4241)	1.6560 (1.4572)	1.1824 (1.5199)
1300	0.8416	0.4606 (0.4687)	0.3687 (0.4824)	0.3732 (0.4596)	0.5343 (0.4703)	0.3492* (0.4905)
1305	0.4973	0.2546 (0.2770)	0.1613* (0.2851)	0.1791 (0.2716)	0.2685 (0.2779)	0.1868 (0.2898)
1310	0.3970	0.3049 (0.2249)	0.3254 (0.2296)	0.1977 (0.2235)	0.3139 (0.2411)	0.1944* (0.2347)

TABLE 12. A LISTING OF ANALYSIS ERROR VARIANCES COMPUTED USING THE DATA SETS INDICATED BY THE VARIOUS COLUMN HEADINGS. THE OBSERVATION VARIANCES ARE INCLUDED FOR COMPARISON PURPOSES. THE LOWEST VARIANCE FOR EACH ANALYSIS TIME IS MARKED (*). THIS INDICATES WHICH ANALYSIS DATA CONFIGURATION PRODUCED THE BEST RESULTS. THE VALUES IN THE PARENTHESES ARE THE MODELED VARIANCE ESTIMATES. THIS DATA IS FROM THE 25TH OF MAY.

multivariate data configuration using the Marshall-Palmer Z-R relationship was best 5 times, the multivariate data configuration using the aircraft determined Z-R relationship was best 9 times, and the surface-gauge data set produced the minimum error variance for the times 1215, 1225, 1230, and 1235. A comparison of these results are shown in Figures 85 thru 88.

Referring to Figure 86 (the Marshall-Palmer Z-R relationship results), the multivariate approach produced the smallest error variance, 12 of the 18 times, while Figure 88 shows that this analysis approach was best 11 times when using the aircraft derived Z-R conversion. Once again, indications are that the multivariate analysis configuration produces consistently better results than either of the univariate approaches considered.

The results using procedure 2 are shown in Table 13. Employing the Marshall-Palmer Z-R conversion, the radar univariate configuration was best for 1245 while the multivariate approach was best for the times 1200 and 1300. The aircraft determined Z-R conversion produced the best results 5 times and 4 times for the univariate and multivariate configurations respectively, while the surface-gauge univariate approach was best the remaining 6 times. Figures 89-92 are a comparison of these results.

For the Marshall-Palmer Z-R relationship the radar univariate approach or the multivariate approach produced

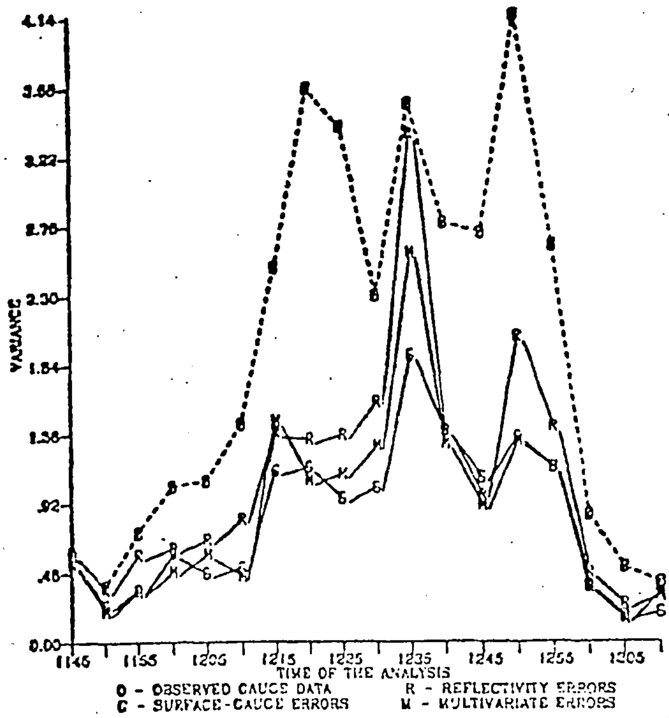

Fig. 85: A comparison of the computed error variances and the respective data variance for the May 25th, 1145-1310, observation set using the Marshall-Palmer Z-R conversion and procedure 1.

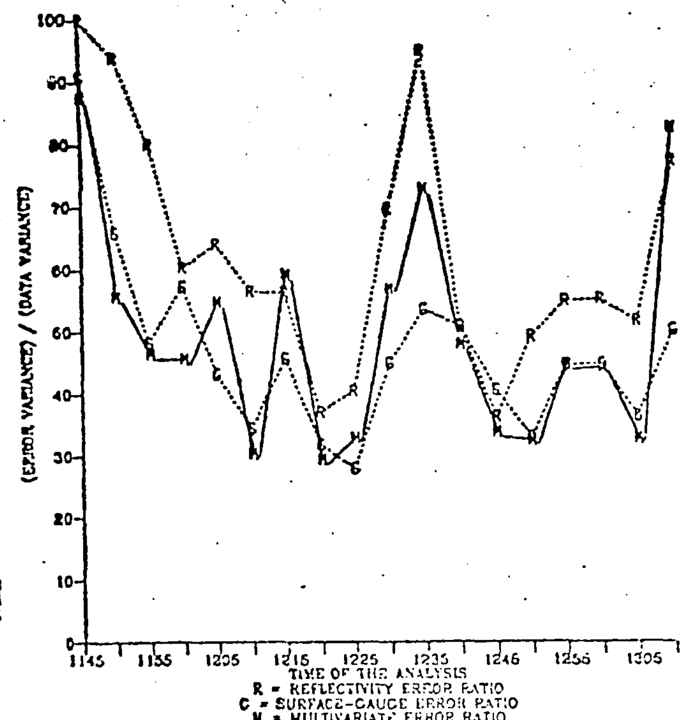

Fig. 86: The error variances of Figure 85 expressed as a percentage of the respective data variance.

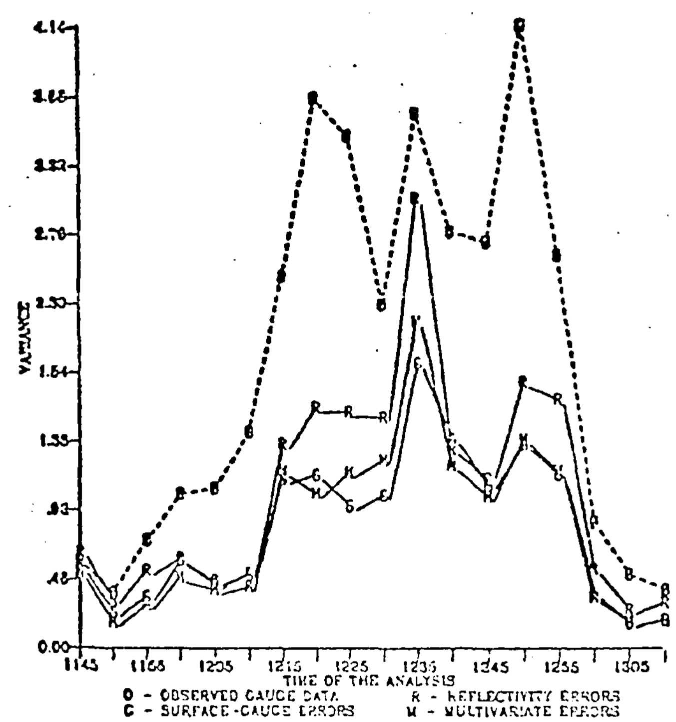

Fig. 87: A comparison of the computed error variances and the respective data variance for the May 25th, 1145-1310, observation set using the aircraft-determined Z-R conversion and procedure 1.

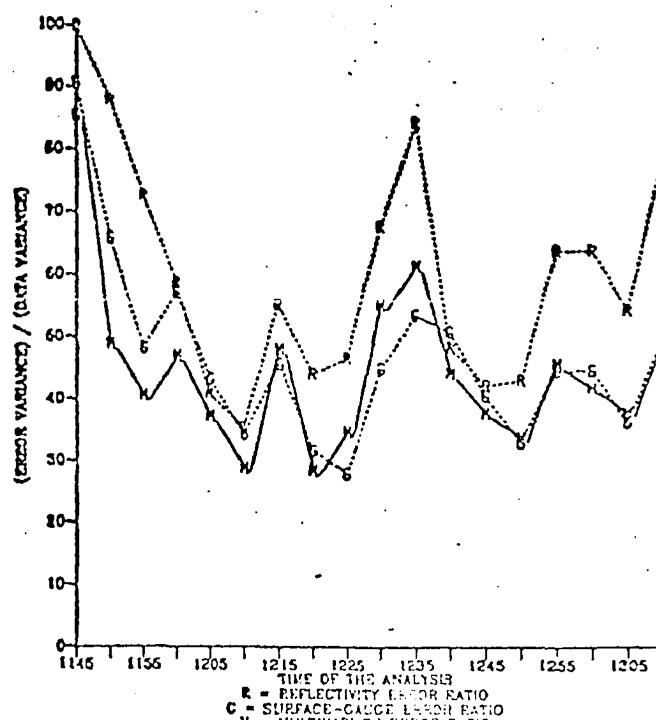

Fig. 88: The error variances of Figure 87 expressed as a percentage of the respective data variance.

ANALYSIS TIME	OBSERVATION VARIANCE	MARSHALL-PALMER Z-R RELATIONSHIP RADAR	MARSHALL-PALMER Z-R RELATIONSHIP GAUGES &RADAR	GAUGES	AIRCRAFT-DETERMINED Z-R RELATIONSHIP RADAR	AIRCRAFT-DETERMINED Z-R RELATIONSHIP GAUGES &RADAR
1145	0.5952	0.6034 (0.3344)	0.6012 (0.3437)	0.6123 (0.3356)	0.6468 (0.3398)	0.5753* (0.3507)
1150	0.3783	0.3553 (0.2107)	0.3410 (0.2162)	0.4041 (0.2116)	0.3332* (0.2114)	0.3662 (0.2216)
1155	0.7263	0.5813 (0.4045)	0.5861 (0.4152)	0.5872 (0.4063)	0.5271 (0.4059)	0.5124* (0.4255)
1200	1.0320	0.6220 (0.5746)	0.5992* (0.5898)	0.7604 (0.5771)	0.6048 (0.5765)	0.6073 (0.6044)
1205	1.0710	0.6849 (0.5968)	0.6533 (0.6125)	0.6162 (0.5994)	0.4377* (0.5987)	0.4626 (0.6277)
1210	1.4490	0.8169 (0.8068)	0.5815 (0.8282)	0.7924 (0.8104)	0.5111* (0.8095)	0.5873 (0.8486)
1215	2.4860	1.3933 (1.3848)	1.6806 (1.4214)	1.5755 (1.3909)	1.3680* (1.3893)	1.4921 (1.4565)
1220	3.6660	1.3468 (2.0421)	1.4245 (2.0961)	1.4332 (2.0510)	1.6083 (2.0487)	1.3013* (2.1478)
1225	3.4080	1.3688 (1.8983)	1.2342 (1.9485)	1.1507* (1.9066)	1.5784 (1.9044)	1.2371 (1.9966)
1230	2.2910	1.5884 (1.2759)	1.5191 (1.3096)	1.3056* (1.2815)	1.5457 (1.2801)	1.5166 (1.3420)
1235	3.5570	3.3646 (1.9813)	2.7194 (2.0337)	2.0784* (1.9899)	2.9928 (1.9877)	2.4746 (2.0838)
1240	2.7660	1.3941 (1.5404)	1.3944 (1.5811)	1.4336 (1.5471)	1.3235* (1.5454)	1.3335 (1.6202)
1245	2.7020	0.9713* (1.5049)	1.1200 (1.5447)	1.0168 (1.5115)	1.1325 (1.5098)	1.3520 (1.5828)
1250	4.1400	2.0165 (2.3060)	1.6232 (2.3670)	1.5193* (2.3161)	1.7685 (2.3135)	1.5910 (2.4254)
1255	2.6080	1.4195 (1.4524)	1.3347 (1.4908)	1.3537 (1.4588)	1.6560 (1.4572)	1.3322* (1.5276)
1300	0.8416	0.4606 (0.4687)	0.5280 (0.4811)	0.4469* (0.4708)	0.5343 (0.4703)	0.5483 (0.4930)
1305	0.4973	0.2546 (0.2770)	0.2436* (0.2843)	0.2933 (0.2782)	0.2685 (0.2779)	0.2856 (0.2913)
1310	0.3970	0.3049 (0.2249)	0.3222 (0.2286)	0.2377* (0.2259)	0.3139 (0.2411)	0.2711 (0.2355)

TABLE 13. A LISTING OF ANALYSIS ERROR VARIANCES COMPUTED USING THE DATA SETS INDICATED BY THE VARIOUS COLUMN HEADINGS. THE OBSERVATION VARIANCES ARE INCLUDED FOR COMPARISON PURPOSES. THE LOWEST VARIANCE FOR EACH ANALYSIS TIME IS MARKED (*). THIS INDICATES WHICH ANALYSIS DATA CONFIGURATION PRODUCED THE BEST RESULTS. THE VALUES IN THE PARENTHESES ARE THE MODELED VARIANCE ESTIMATES. THIS DATA IS FROM THE 25TH OF MAY.

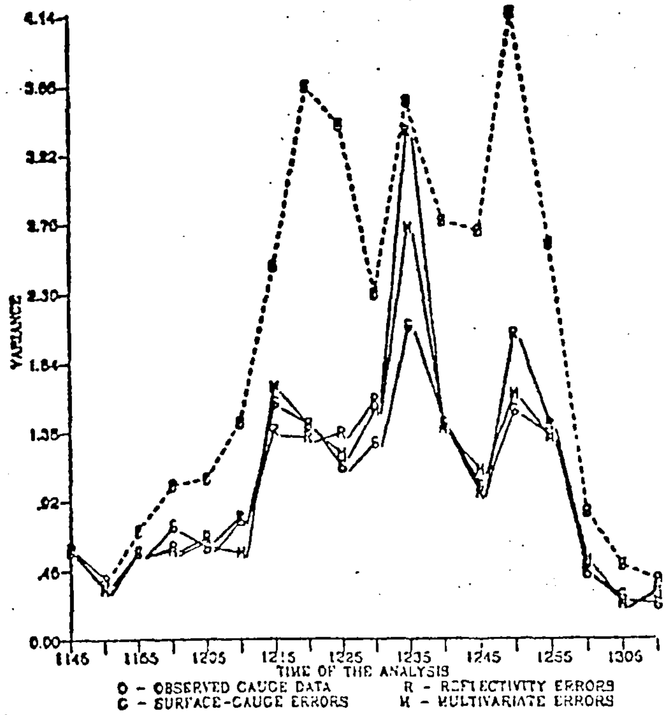

Fig. 89: A comparison of the computed error variances and the respective data variance for the May 25th, 1145-1310, observation set using the Marshall-Palmer Z-R conversion and procedure 2.

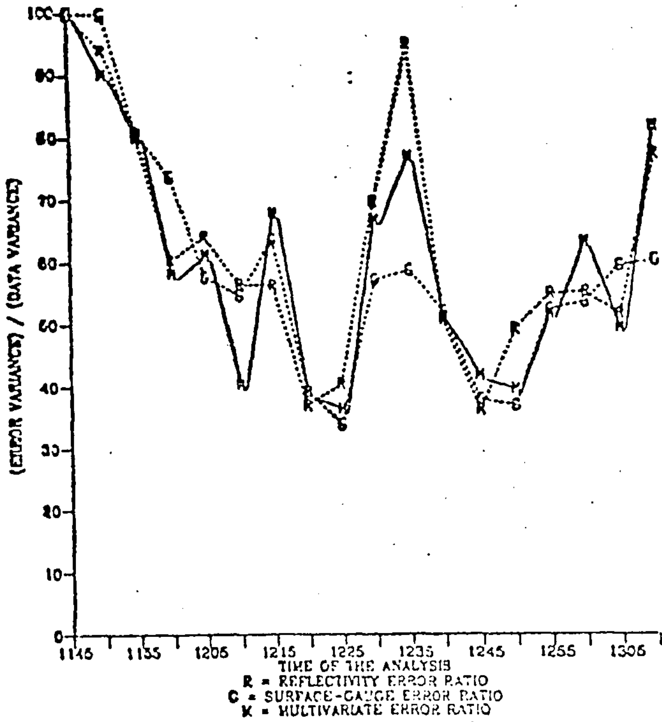

Fig. 90: The error variances of Figure 89 expressed as a percentage of the respective data variance.

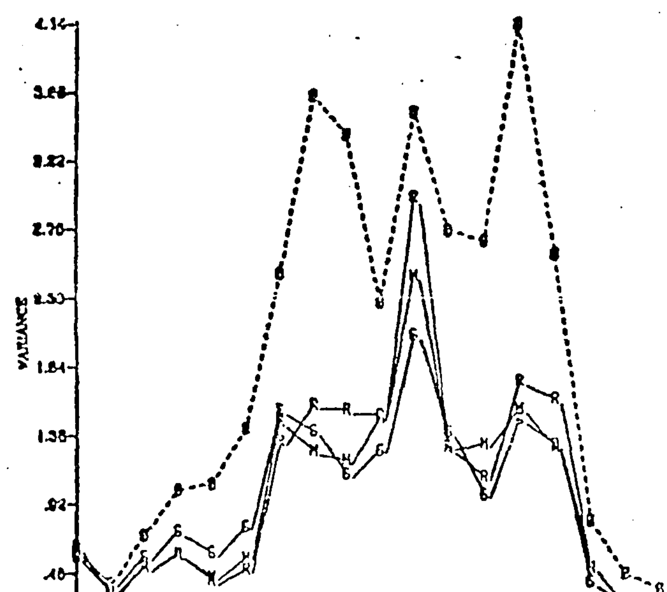

Fig. 91: A comparison of the computed error variances and the respective data variance for the May 25th, 1145-1310, observation set using the aircraft-determined Z-R conversion and procedure 2.

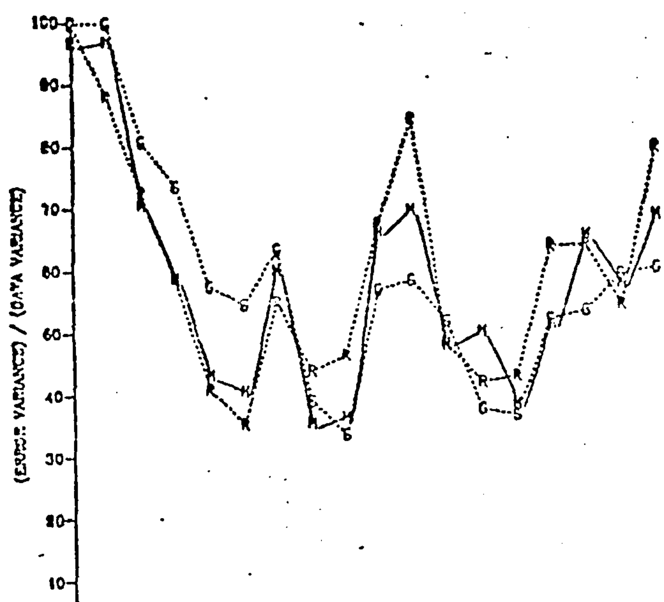

Fig. 92: The error variances of Figure 91 expressed as a percentage of the respective data variance.

the best results 12 of the 18 times, while for the aircraft derived Z-R conversion of the radar data, the univariate or multivariate approach was best 11 times. Notice that for the 11th analysis (1235) the radar data produced a rather large error variance when used in either the univariate or multivariate configuration. It was at this time that the differences between the surface-gauge rainfall rates and the reflectivity rainfall rates were the highest, a little over 5.0 inches/hour for some stations. This analysis technique (or any other) simply could not adequately handle this particular situation. When the problem stations (there were two in number) were deleted from the data sample, and the analyses using all three configurations were recomputed, the results were more consistent. However, since no data editing was accomplished on any other data set, the original results were kept and these appear in Table 13.

Figures 93-94 compare the relative worth of the different Z-R relationships, as was done for the May 21st data set. The first figure corresponds to procedure 1 results; Figure 94 corresponds to procedure 2 results. As before, no predominant difference exists between these two Z-R conversions when used in a multivariate analysis configuration.

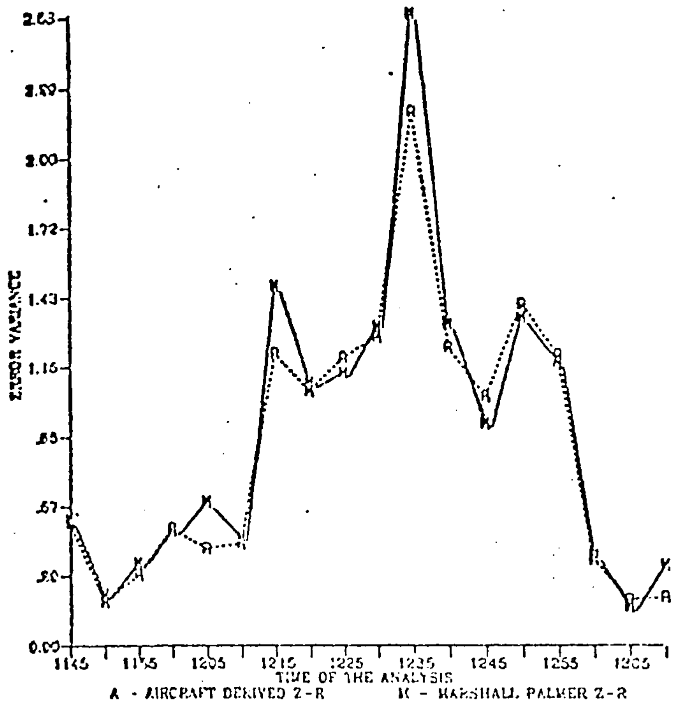

Fig. 93: A comparison of the effect of different Z-R relationships on the analysis of surface precipitation for May 25th, 1145-1310, when procedure 1 was used.

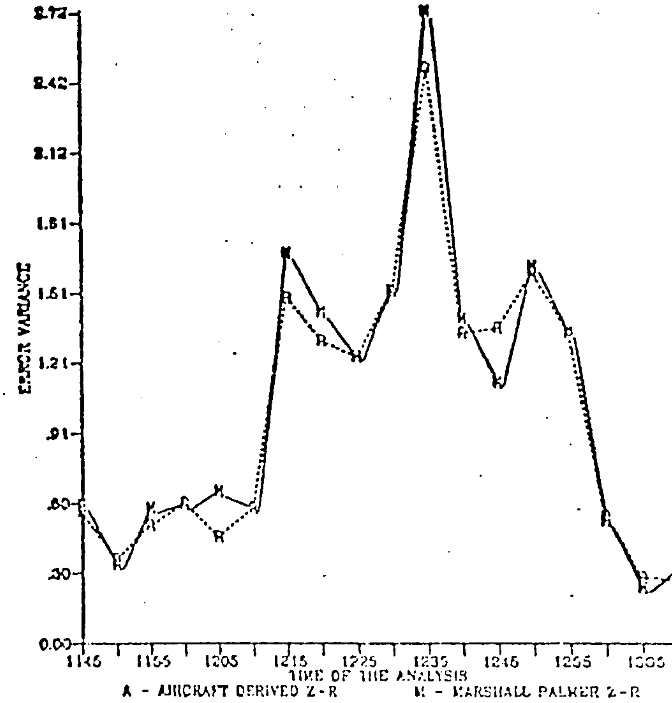

Fig. 94: A comparison of the effect of different Z-R relationships on the analysis of surface precipitation for May 25th, 1145-1310, when procedure 2 was used.

E. Analysis' Filtering Properties

It is of some importance to be able to make a quantitative statement concerning the filtering properties of the technique used in the analysis of a phenomenon of interest. Without a quantitative measure of the ability of the analysis scheme to filter the signal from the signal-plus-noise observation set, it would be difficult to attach much credibility to any results obtained. The two assessment techniques discussed in this section can be used to help evaluate the reliability of any analysis model employed.

The first assessment technique discussed is a necessary condition for an analysis scheme to be considered acceptable. In the results presented in sections A thru D of this chapter, it will be noticed that in every analysis, the error variance is always less than the respective data variance for the time period analyzed. This is an expected result. Obviously, if the error variance determined from an analysis was larger than the variance of the corresponding data set itself, either the technique or the data sample should be suspect. The fact that this was not the case is encouraging. This simple comparison, however, is rather negative in its interpretation. Favorable results imply that the analysis scheme exhibits no gross discontinuities. Favorable results do not imply that the analysis scheme is worthwhile.

The second assessment technique discussed is both a necessary and sufficient condition for an analysis scheme to be considered acceptable. This technique compares the signal-plus-noise structure obtained from the observation set, with the signal-plus-noise structure obtained from the analysis of that set. Two things will become apparent from this comparison. First, and most important, it is possible to determine the amount of distortion introduced into the analyzed field by the analysis technique itself; the poorer the technique used, the greater will be the distortion detected. A good technique should be essentially distortion free. The second thing to be noticed from this comparison, is the relative value of the structure model A7 parameter for the two structure functions obtained. Since any analysis technique should do some filtering of the noise from the signal, the A7 parameter value for the analyzed field would be expected to be greater than the A7 value for the corresponding observation set. The larger the analyzed A7 parameter value is, respective of a distortion free structure function, the better the analysis technique employed.

Figures 95 thru 102 show the time ordered signal-plus-noise structures obtained from the four storm realizations. The dimensions and orientation of these figures are the same as those of Figures 4 and 5. It will be noticed that in every case the signal-plus-noise structure obtained from the analyzed field, Figures 96, 98, 100, and 102, is very homologous to the original signal-plus-noise structure obtained from the observations themselves, Figures 95, 97, 99, and 101. It will also be noticed that the analyzed correlation coefficient A7 parameter is always larger, as was expected.

The results of these two assessment techniques, indicate that the analysis scheme employed does indeed filter the signal from the signal-plus-noise observation set in a distortion free manner. The fact that the observed signal-plus-noise structure was consistently recovered from the analyzed field, plus the fact that a substantial amount of noise was consistently filtered from the observation set, would imply that the results discussed in the earlier sections of this chapter have a high degree of credibility.

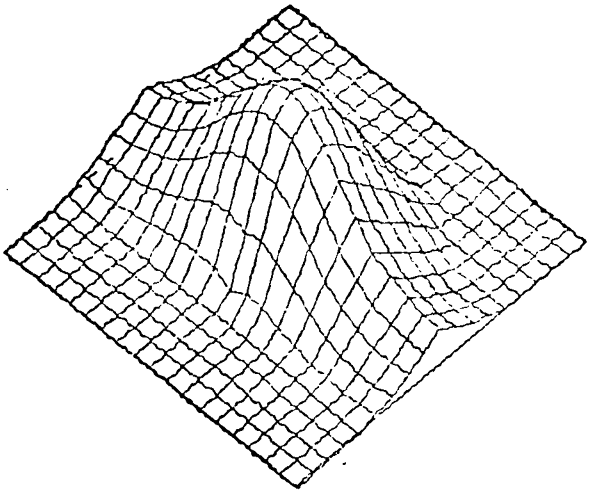

Fig. 95: The structure model of the observed surface precipitation for May 21st; U_t=0 min. and A7=.3545.

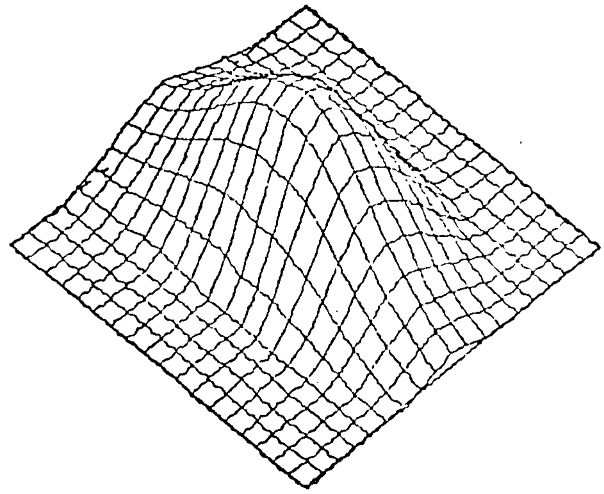

Fig. 96: The structure model of the analyzed surface precipitation for May 21st; U_t=0 min. and A7=.6011.

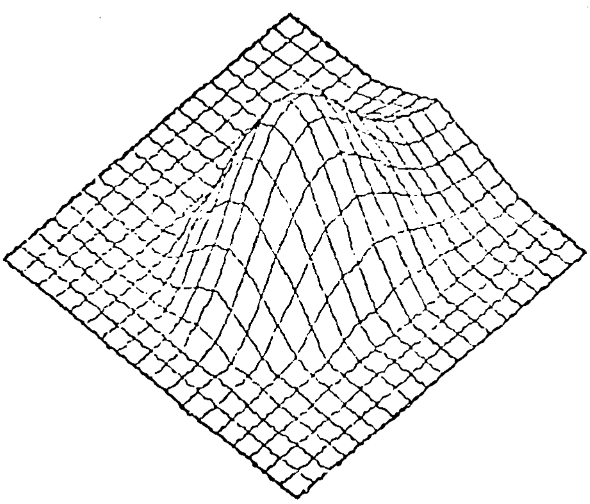

Fig. 97: The structure model of the observed surface precipitation for May 25th, 0755-0905; U_t=0 min. and A7=.4969.

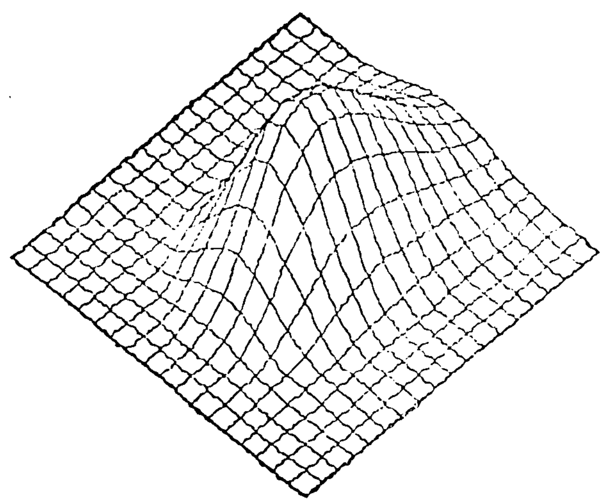

Fig. 98: The structure model of the analyzed surface precipitation for May 25th, 0755-0905; U_t=0 min. and A7=.7481.

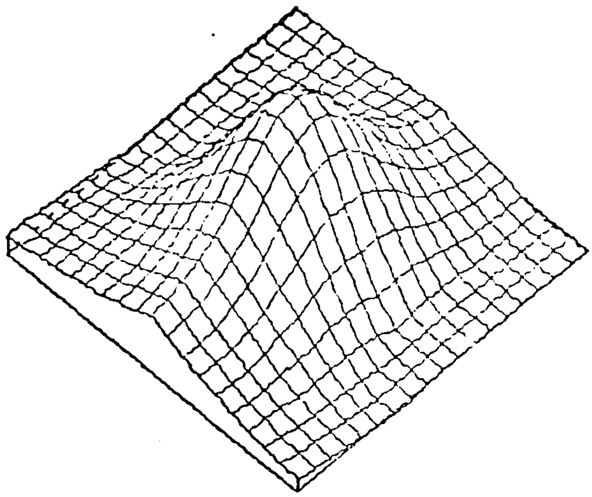

Fig. 99: The sturcture model of the observed surface precipitation for May 25th, 0935-1010; U_t=0 min. and A7=.3855.

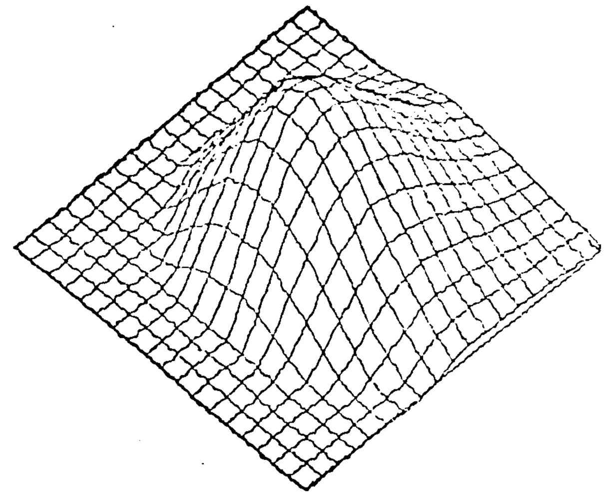

Fig. 100: The structure model of the analyzed surface precipitation for May 25th, 0935-1010; U_t=0 min. and A7=.5661.

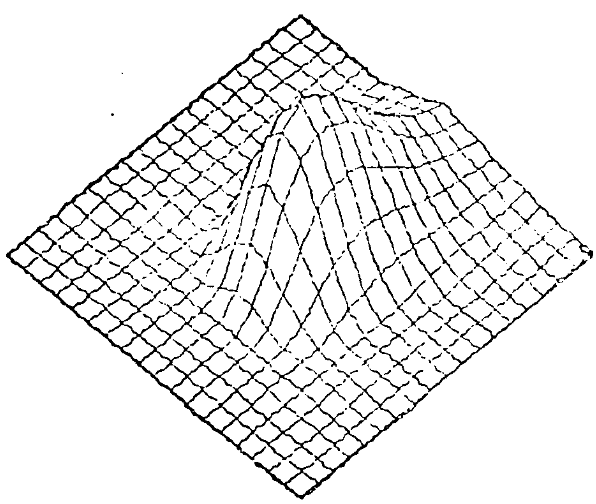

Fig. 101: The structure model of the observed surface precipitation for May 25th, 1145-1310; U_t=0 min. and A7=.5264.

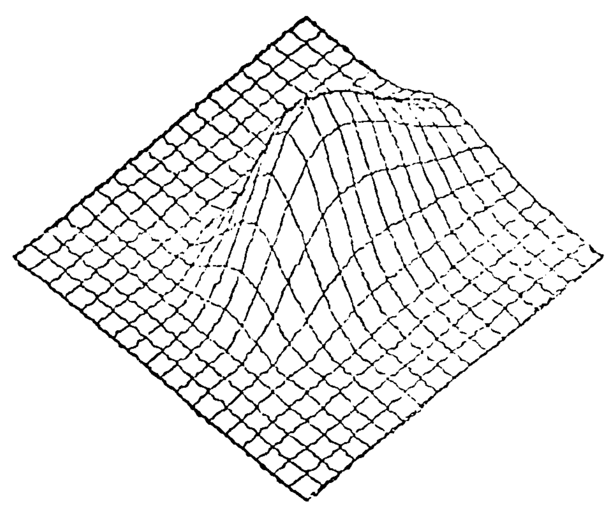

Fig. 102: The structure model of the analyzed surface precipitation for May 25th, 1145-1310; U_t=0 min. and A7=.8253.

CHAPTER VI

EXPERIMENTAL DESIGN MODEL

The objective analysis technique previously discussed can not only determine estimates of the parameter field, but also modeled estimates of the amount of variance in the original data sample that has been explained by a particular sampling pattern in space and time with respect to some specified grid lattice. Since this determination is not dependent on having observations of the parameter field, but is only dependent on the definition of the structure model supplied, it is possible to shift sampling patterns in space and time with respect to the analysis grid, so as to explain the maximum percentage of the original data variance (Eddy, 1974). The incorporation of this analysis technique with an NLP algorithm constitutes the experimental design model alluded to earlier, the particulars of which are discussed in this chapter.

The multiple correlation coefficient, R, can be defined as the proportion of variance of the observation Y accounted for by the objective analysis $\hat{Y}$. If

$$Y^t V^{-1} \hat{Y} = \hat{\sigma}_y \hat{\sigma}_{\hat{y}} R, \tag{15}$$

then $$R = Y^t V^{-1} \hat{Y} \,/\, \{(Y^t V^{-1} Y)^{1/2} \, (\hat{Y}^t V^{-1} \hat{Y})^{1/2}\}, \tag{16}$$

where $$\hat{\sigma}_y = (Y^t V^{-1} Y)^{1/2} \text{ and } \hat{\sigma}_{\hat{y}} = (\hat{Y}^t V^{-1} \hat{Y})^{1/2}. \tag{17}$$

Using the definition of $\hat{\beta}$ given in Chapter IV, we see that

$$\hat{Y} = X\hat{\beta} = X\{(X^t V^{-1} X)^{-1} (X^t V^{-1} Y)\}. \tag{18}$$

Multiplying both sides of this equation by $Y^t V^{-1}$, we obtain

$$Y^t V^{-1} \hat{Y} = Y^t V^{-1} X\{(X^t V^{-1} X)^{-1} (X^t V^{-1} Y)\}. \tag{19}$$

Now notice that $\hat{Y}^t V^{-1} \hat{Y}$ can be written as

$$\begin{aligned} \hat{Y}^t V^{-1} \hat{Y} &= \{(Y^t V^{-1} X)(X^t V^{-1} X)^{-1} X^t\} V^{-1} \{X(X^t V^{-1} X)^{-1} (X^t V^{-1} Y)\} \\ &= Y^t V^{-1} X\{(X^t V^{-1} X)^{-1} (X^t V^{-1} Y)\}, \end{aligned} \tag{20}$$

which is the same as equation (19). Therefore

$$Y^t V^{-1} \hat{Y} = \hat{Y}^t V^{-1} \hat{Y}, \text{ and} \tag{21}$$

$$R^2 = Y^t V^{-1} \hat{Y} \,/\, Y^t V^{-1} Y. \tag{22}$$

The fraction of variance unaccounted for by the regression is then given by $1-R^2$.

Denoting the multiple correlation coefficient between the observation Y and the objective analysis $\hat{Y}$ at the ith grid point by R_i^2, the modeled fractional unexplained variance or "residual" variance for the entire grid array is $[\sum_i^G (1-R_i^2)] \,/\, G$, where G is the number of grid points. This summation constitutes an objective function which can be minimized, subject to logistic and engineering constraints if necessary, by employing the NLP algorithm discussed in Chapter IV. Specifically,

$$f(\xi) = [\sum_{i}^{G} \alpha_i (1-R_i^2)] \;/\; G, \tag{23}$$

where α_i is a weighting factor assigned to each grid location in the analysis space, and

$$\xi = (x_1, y_1, z_1, t_1, \;\ldots\ldots,\; x_s, y_s, z_s, t_s) \tag{24}$$

is a vector of s sensor positions in 4-space. Given some starting vector of sensor positions ξ_1, the NLP algorithm repositions these sensors in space and time so as to minimize the value of the objective function $f(\xi)$. The α_i are used to reflect the pertinent climatology for the area relevant to the forthcoming experiment. The final vector of sensor positions corresponding to the minimized value of $f(\xi)$ is an optimal sensor configuration for sampling the phenomenon specified by the structure function used in the objective analysis. A modeled estimate of the amount of variance un-accounted for by the regression, with respect to a particular sampling scheme, is given by

$$\{[\sum_{i}^{G} \alpha_i (1-R_i^2)] \;/\; G\}\hat{\sigma}_y^2. \tag{25}$$

Figures 103 thru 110 compare the computed error

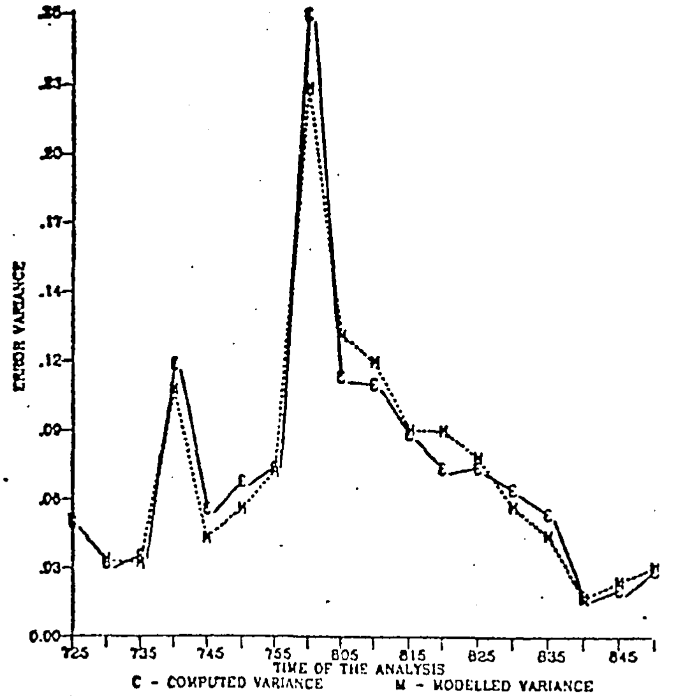

Fig. 103: A comparison of the computed error variance with the modeled error variance for May 21st using the Marshall-Palmer Z-R conversion.

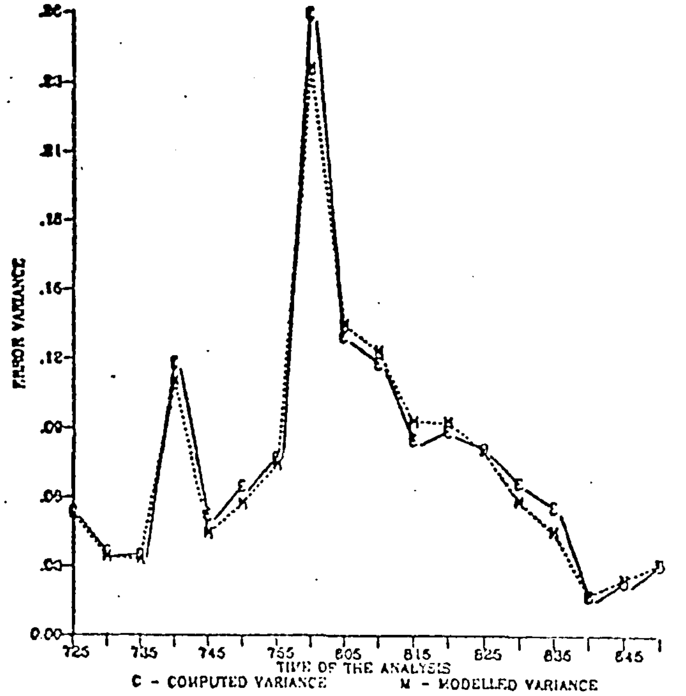

Fig. 104: A comparison of the computed error variance with the modeled error variance for May 21st using the aircraft-derived Z-R conversion.

Fig. 105: A comparison of the computed error variance with the modeled error variance for May 25th, 0755-0905, using the Marshall-Palmer Z-R conversion.

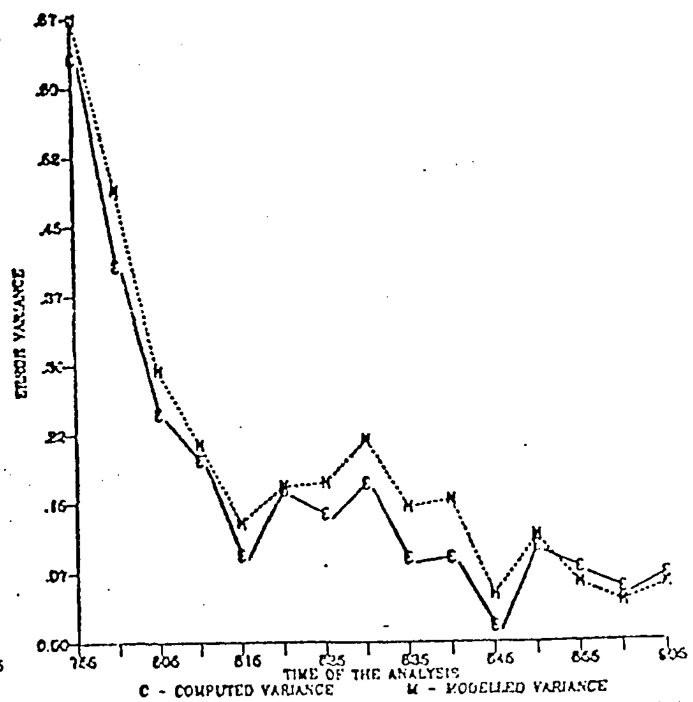

Fig. 106: A comparison of the computed error variance with the modeled error variance for May 25th, 0755-0905, using the aircraft-derived Z-R conversion.

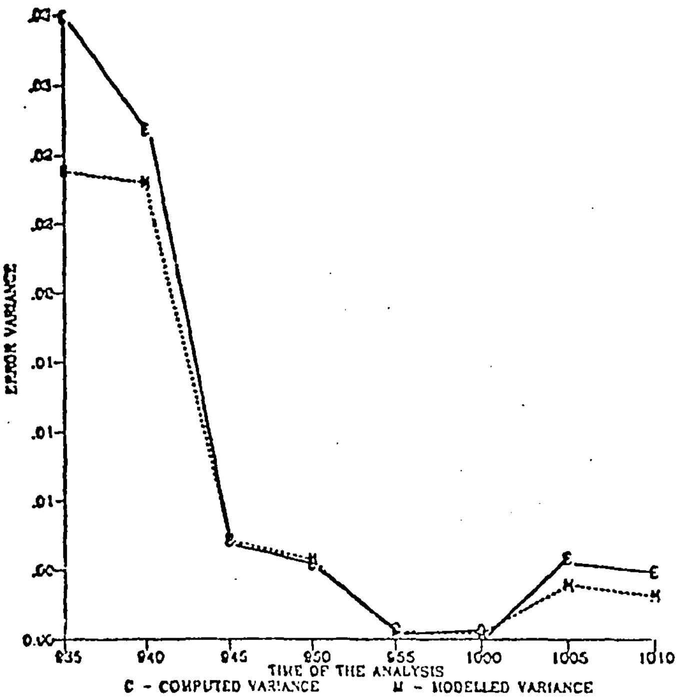

Fig. 107: A comparison of the computed error variance with the modeled error variance for May 25th, 0935-1010, using the Marshall-Palmer Z-R conversion.

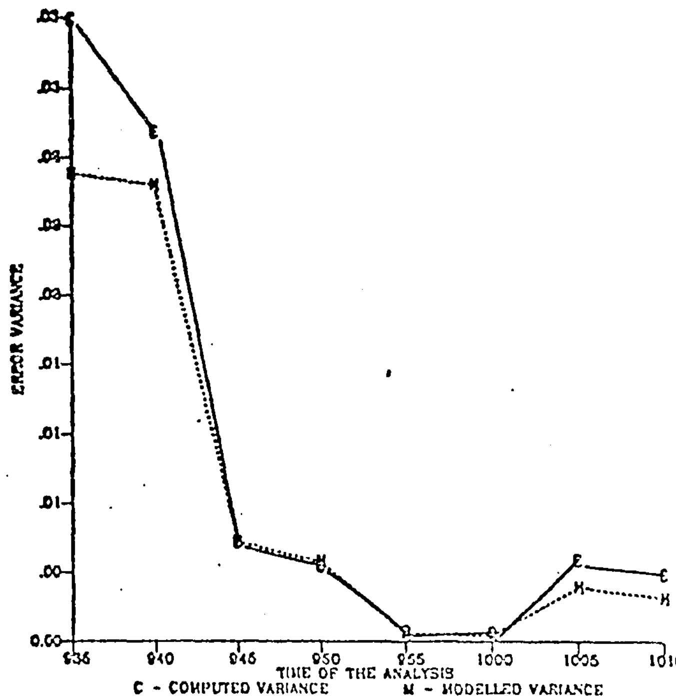

Fig. 108: A comparison of the computed error variance with the modeled error variance for May 25th, 0935-1010, using the aircraft-derived Z-R conversion.

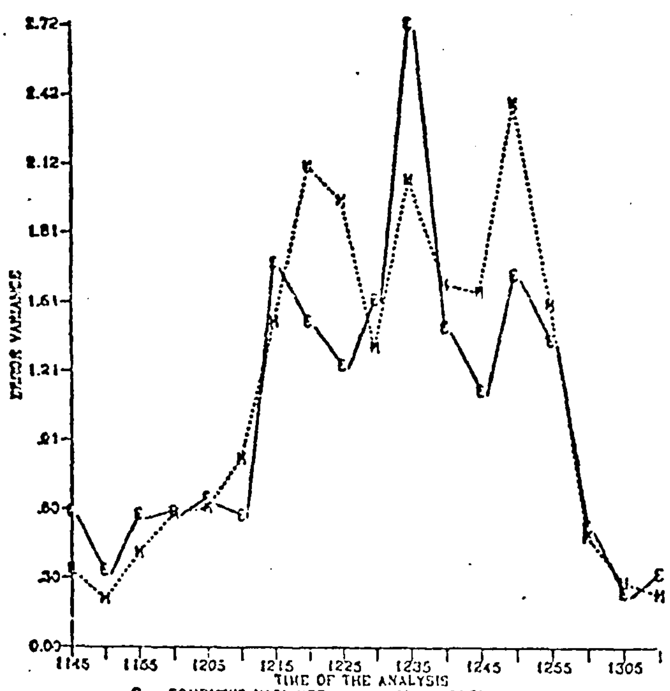

Fig. 109: A comparison of the computed error variance with the modeled error variance for May 25th, 1145-1310, using the Marshall-Palmer Z-R conversion.

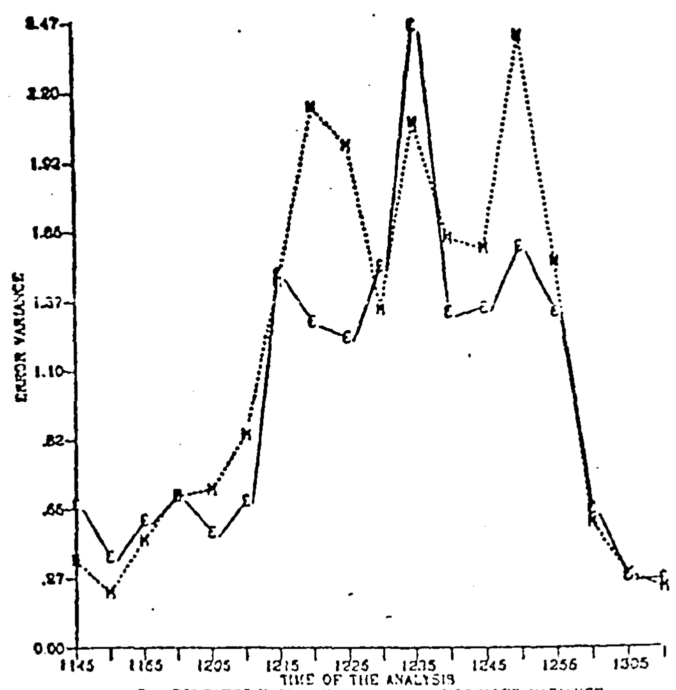

Fig. 110: A comparison of the computed error variance with the modeled error variance for May 25th, 1145-1310, using the aircraft-derived Z-R conversion.

variance, σ_e^2, with the modeled error variance (obtained using formula (25)), for the four precipitation data sets, in order to illustrate the adequacy of the residual variance model used. The error variances were obtained from the multivariate data configuration using procedure 2. Figures 103, 105, 107, and 109 are the time ordered comparisons for the Marshall-Palmer Z-R conversion, while Figures 104, 106, 108, and 110 show the results when empirically derived Z-R conversions were used.

The May 21st results, Figures 103 and 104, and the May 25th results for the time 0935-1010, Figures 107 and 108, show that the modeled error variance values correspond very well with the values of the computed error variance. The coefficients of correlation for these variance comparisons are .982, .993, .992, and .992, respectively.

The results from the May 25th 0755-0905 realization, Figure 105, seem to exhibit a positive bias in the modeled variance values. This indicates that the actual analysis results were somewhat better than what could be expected (in a statistical sense) from the regression model and data configuration employed. The empirically derived Z-R conversion for this time period, Figure 106, appears to have reduced this bias somewhat. The correspondence of relative variance fluctuations is still very good in either case, the correlation coefficients being .989 and .988, respectively.

For the May 25th 1145-1310 realization, Figures 109 and 110, the results correspond surprisingly well, considering the large amount of variation in this data set, as well as the peculiarities of the actual measurements, which were previously discussed in Chapter V, section D. Notice that for the 7th, 10th, and 11th analyses, the regression model predicted better results than those obtained, while for the 8th, 9th, and 12th thru 15th analyses, the results obtained were better than what would be expected. The correlation coefficient values for these variance quantities are .857 and .881, respectively, which is still quite high.

When investigating the station configuration problem with respect to a specified sensor mix, the objective function as given in equation (23) is used. However, when the purpose of the investigation is to determine the relative worth of different sensor types, a modified form of this equation is needed. As an illustration of this point, consider the problem of using a surface-gauge radar-reflectivity configuration for estimating surface precipitation. Two things must be properly "measured" in order to obtain an acceptable analysis of the precipitation field; the <u>variation</u> in the signal, which allows the analysis model to reflect the detail of the phenomenon correctly, and the mean of the signal, which allows for a proper scaling of the analyzed field. As the overall number of sensors decreases, the ability to detect the variation in the signal decreases. As the number of surface sensors decrease, the ability to recover the mean of the signal (as observed at the surface) decreases. Equation (23) models the effect of any overall sensor reduction. The development of a model to reflect the effect of reducing the number of surface sensors, and the incorporation of this second model with the first, is presented below.

Let $\bar{y}$ be the sample mean of the signal being analyzed, then $\sigma^2_{\bar{y}}$ is the variance of the mean, and

$$\sigma^2_{\bar{y}} = \sigma^2_y/N^*, \qquad (26)$$

where σ^2_y is the population variance and N^* is the number of effective degrees of freedom. For the general case

$$N^* = N\{1 + \frac{2}{N} \sum_{\tau}^{N-1} (N-\tau)\ Y(\tau)\}\ , \qquad (27)$$

where N is the number of surface sensors, Y is a vector of N surface observations, and $Y(\tau) = Y(i, i+\tau)$ for all i. If the observations are statistically independent (white noise) then $Y(\tau) = 0$ for all $\tau \neq 0$, and $N^* = N$. For the purpose of this discussion, white noise is assumed, and the proper incorporation of the mean variance model with the signal variance model results in the following formula,

$$\{[\sum_{i}^{G} \alpha_i(1-R_i^2)] \;/\; G\}\hat{\sigma}_y^2 + \hat{\sigma}_y^2 \;/\; N. \qquad (28)$$

It is a well known property, that if x* minimizes f(x), then x* minimizes cf(x) for any constant c > 0. Note that in formula (28) σ_y^2 is a constant > 0 for any analysis period. Therefore, the objective function can be rewritten as

$$f(\xi) = [\sum_{i}^{G} \alpha_i(1-R_i^2)] + 1/N, \qquad (29)$$

and it is this formulation that is used in the comparison of different sensor mixes.

By using this design technique, it is possible to discover optimal station configurations in space and time, in order to detect and analyze a phenomenon whose spatial-temporal structure can be described by the signal-plus-noise model supplied to the design algorithm. The results of this approach can only be as good as the accuracy of the structure models used, and they can only be as meaningful physically as the quality and quantity of the climatology information given. The resultant station configuration will be tuned to the specified "modeled" phenomenon as well as the climatology of the region of interest.

CHAPTER VII

DESIGN RESULTS

In order to illustrate the potentials of the design algorithm, the results of a pre-experiment investigation are presented. The experimental design for this simulated field study included: the spatial determination of optimal sensor positions, the examination of various sampling rates, and the investigation of different sensor mixes. Specifically, the surface-gauge signal-plus-noise model for the May 25th 1145-1310 realization, was chosen to be the type of precipitation pattern with which the field study was concerned; and for the investigation of the sensor mix problem, the respective correlation and cross-correlation models from the Marshall-Palmer Z-R relationship were employed. The purpose of the pre-experiment investigation was to present to the principal investigator some objective criteria which could be used in the determination of a best sensor configuration for sampling the specified phenomenon with respect to the assumed climatology of the region. The area covered by the field study was approximately 31x31 square miles.

The first problem to be investigated was the spatial determination of optimal sensor positions for i sensors, where

$i = 1, 3, \ldots, 11, 13, 17$, and 21. Figures 111 thru 114 show the optimal positions found using 5, 9, 13, and 21 sensors respectively. It was assumed that the climatology of the region indicated a preferred northeasterly storm track for the phenomenon of concern, that the probability of a storm moving on a particular southwest to northeast track through the specified area could be characterized by a normal distribution about the southwest-northeast diagonal, and that approximately 68% of the storm tracks would occur between the dashed lines shown in the aforementioned figures. Figure 115 shows the original positions of 13 sensors which were optimally relocated to the positions shown in Figure 116. Notice that the final positions for these 13 sensors do not coincide with the positions for the 13 sensors shown in Figure 113. This is because different storm track probability distributions were supplied to the design algorithm, resulting in the two dissimilar placements. The distribution for the latter placement had a much smaller standard deviation resulting in the much "tighter" configuration of Figure 116. A comparison of the two deployments illustrates the influence of the climatology information on the sensor positions determined. The same starting vector was used in each case.

The final objective function values are plotted in Figure 117 for the different sensor quantities investigated. Notice that as the number of surface sensors increase, the objective function values decrease, implying that a better

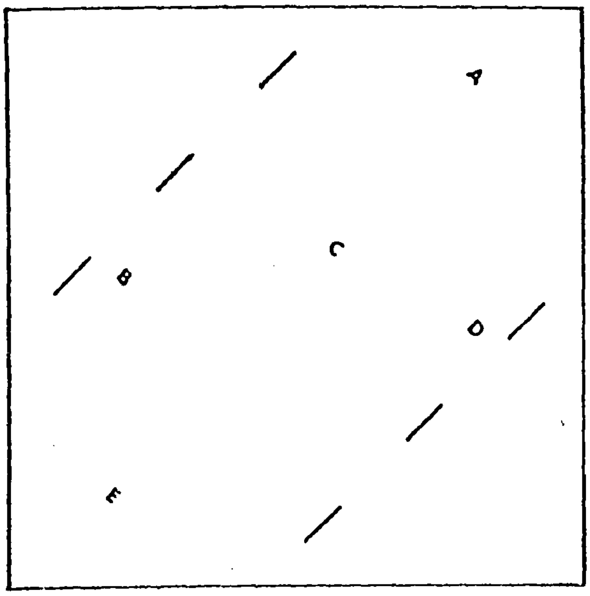

Fig. 111: The final design configuration for 5 sensors.

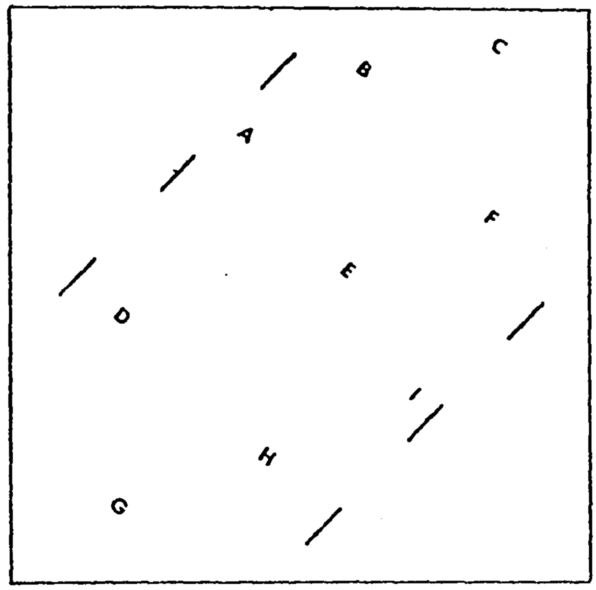

Fig. 112: The final design configuration for 9 sensors

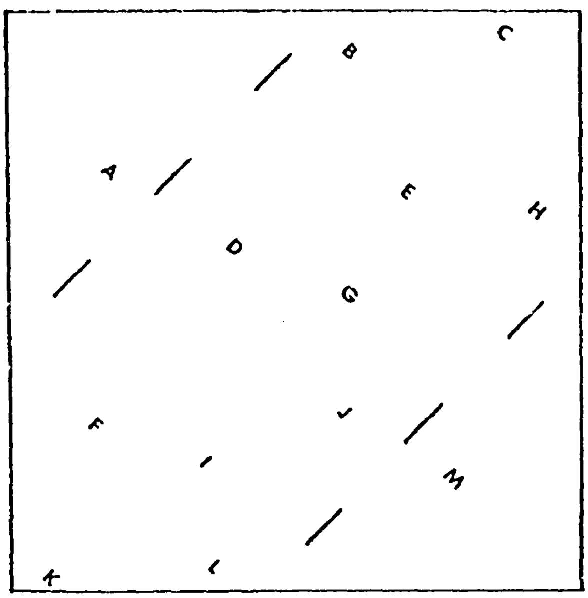

Fig. 113: The final design configuration for 13 sensors.

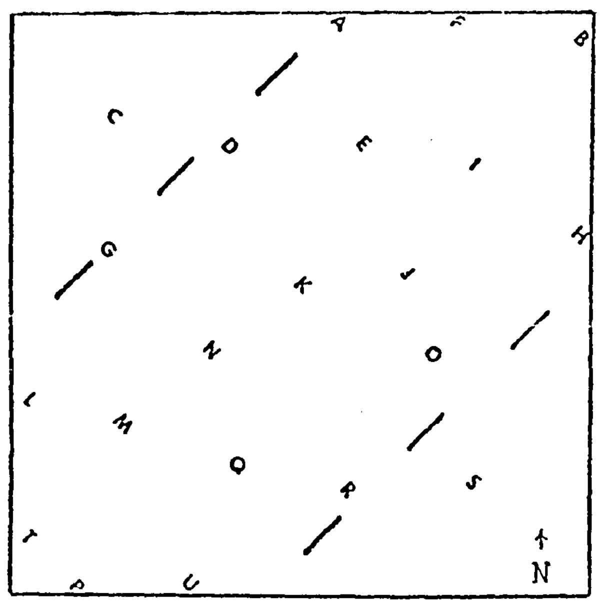

Fig. 114: The final design configuration for 21 sensors.

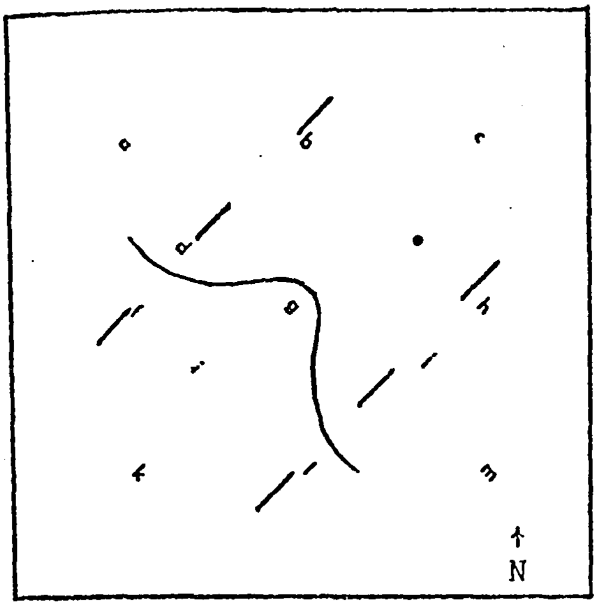

Fig. 115: The position starting vector used for the placement of 13 sensors.

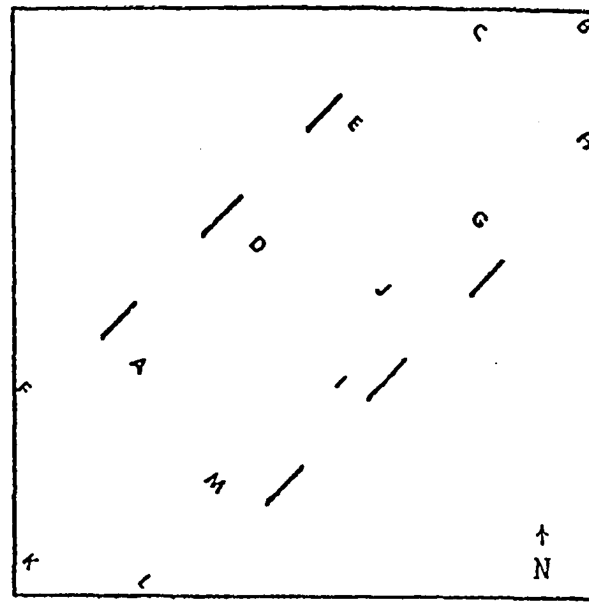

Fig. 116: The final design configuration for the 13 sensors of Figure 115.

sampling of the phenomenon is possible, which was the expected result.

A problem peculiar to the use of this design model, concerns the sensitivity of the NLP algorithm to different sensor starting vectors. In the function fitting of Chapter IV, the use of different starting vectors had very little influence on the final objective function values or on the parameters of the signal-plus-noise models obtained. This was because the function being minimized usually had a single local optimum which was also the global optimum, and regardless of the starting vector used, the algorithm would eventually converge to the optimal solution. In the sensor placement problem, however, the solution space invariably has a number of local optimums, depending on the number of sensors being

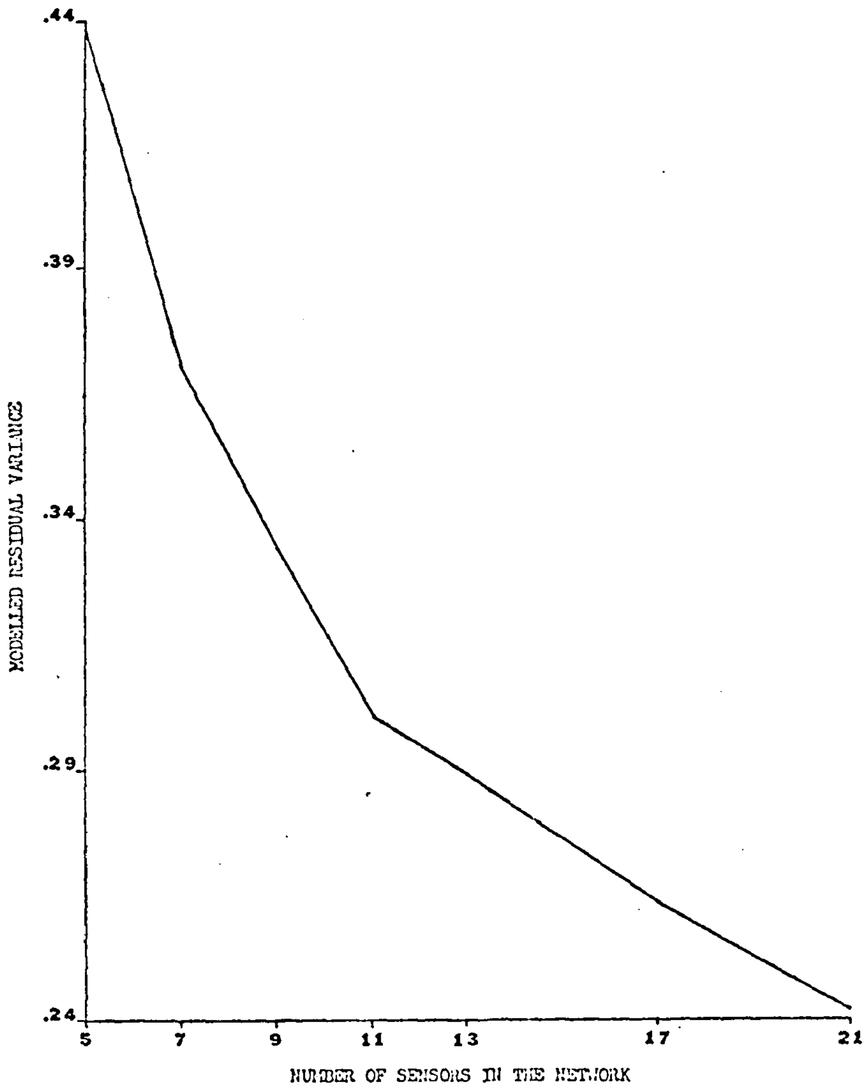

Fig. 117: The modeled residual variance as a function of the number of sensors in the sampling network.

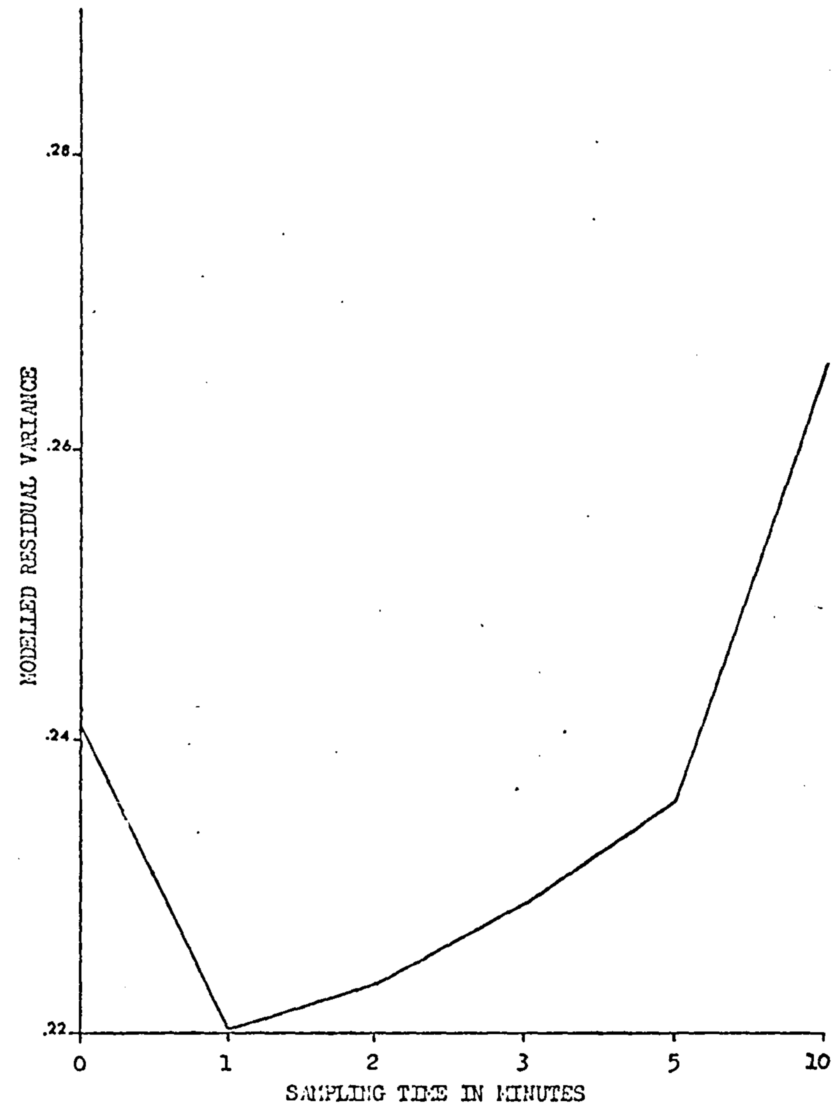

Fig. 118: The modeled residual variance as a function of the sampling rate for a network of 21 sensors.

investigated, the particular grid network used, the peculiarities of the signal-plus-noise model employed, etc. Therefore, the best statement that can be made for any station deployment is that a local optimum has been achieved; the NLP algorithm guarantees this (Himmelblau, 1972). Realizing this fact, a number of "good" sensor starting vectors can be used in different iterations of the design algorithm, and a comparison of the resulting objective function values would indicate the better final deployment. The results given in this chapter are the best deployments achieved for the several different starting vectors used.

Having examined the spatial arrangement for a variety of network sizes, various sampling rates were then investigated. The results of one such investigation, where 21 sensors were deployed, are shown in Figure 118. The better objective function values were achieved at the faster sampling rates, while the poorer objective function values occurred at the slower rates. The objective function value for the sampling time of 0 minutes is the result obtained from a spatial analysis of the phenomenon, as opposed to the spatial-temporal analyses results otherwise shown.

In Figure 118, it will be noticed that the objective function was determined for accumulation times of 1, 2, and 3 minutes as well as for accumulation times of 5 and 10 minutes. The May 25th signal-plus-noise model used in this simulated investigation was actually determined for an accumulation time of 5 minutes, which is usually the shortest accumulation time achievable in a precipitation measurement field experiment, especially when radar is involved. The investigation of the faster sampling rates was an exercise to demonstrate that the temporal dimensionality of the network could also be considered when using this design model. Obviously, in an actual application, the engineering limitations on the sensors used (minimum sampling rate, etc.) would be incorporated into the experimental design process.

An investigation of different sensor mixes was also made, using the correlation and cross-correlation functions previously mentioned. The cross-correlation model A7 parameters were modified, in order to present the sensor mix problem as a function of the relative magnitude of the signal/noise ratio. The relevant A7 parameters were changed from .577 and .551 to .707 and .681 for a first example and to .857 and .831 for the second. Notice that since the computed values of the A7 parameter for the surface-precipitation correlation model and the radar deduced correlation model were .506 and .865, respectively, the augmented cross-correlation values still remained within this range; the implication being that such cross-correlation values could possibly, if not probably, be achieved.

The 21 surface sensors (type 1 sensor) of Figure 114 were replaced one by one, from the center outward, with the corresponding radar estimates (type 2 sensor), and the value of the objective function was recomputed. As the sensors were replaced, since the cross-correlation A7 parameter values were relatively high, it was expected that a better objective function value would result, as is shown in Figure 119. For example 1, the objective function value decreases until 14% of the surface gauges have been replaced with type 2 sensors, while for example 2, the objective function minimum occurs when 43% of the sampling network is composed of type 2 sensors. In each example, as the relative number of surface sensors continues to decrease, the ability to estimate the mean of the surface precipitation field also decreases, as reflected in the upward swing of the respective graphs.

From additional work accomplished in the course of this research, similar to the type of investigation presented in Figure 119, the effect of the magnitude of the cross-correlation A7 parameters on the sensor mix problem became apparent. It was found that as the relative value of the cross-correlation A7 parameters increased, the minimum of the sensor mix curve was shifted to the right with an appropriately lower (better) residual variance value. A decrease in the relative value of these parameters resulted in an increase in the residual variance minimum of the sensor mix curve, with a corresponding shift to the left. These shifts, along with the respective changes in the minimum variance values, reflected the relative worth of the two sensor types being examined.

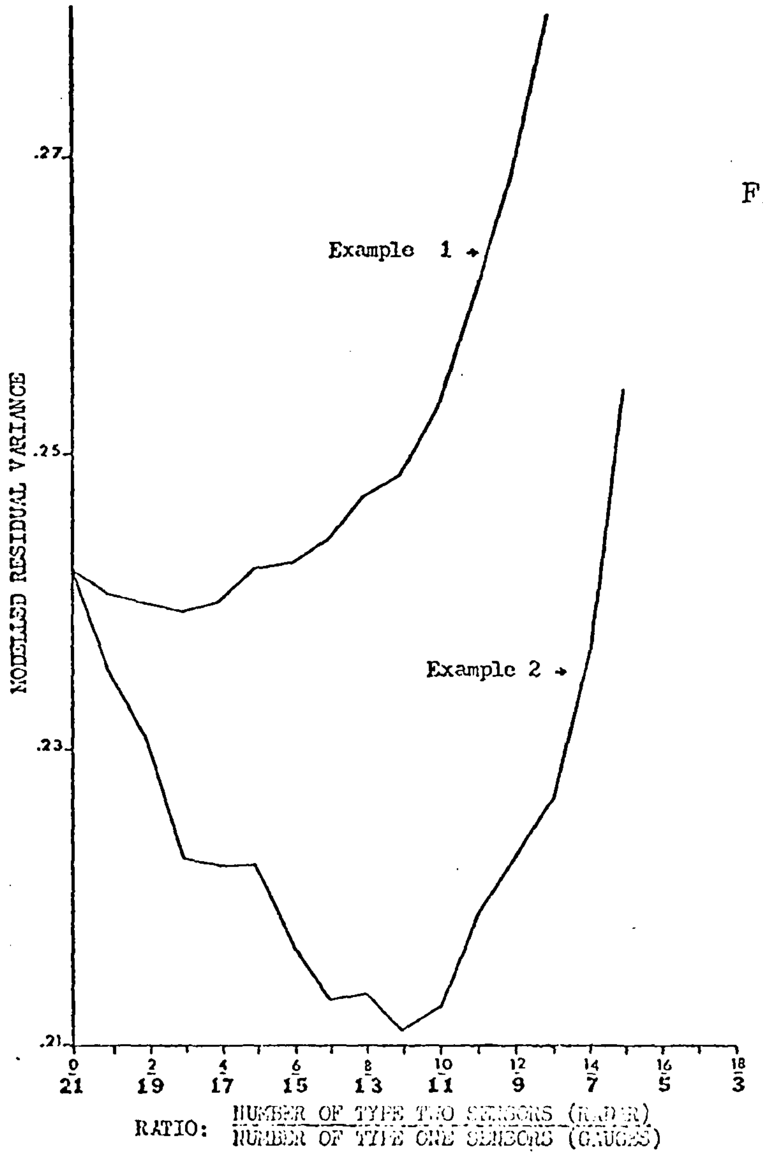

Fig. 119: The modeled residual variance as a function of the ratio of type 2 sensors (reflectivity estimates) to type 1 sensors (surface-gauge measurements) for a sampling network of 21 sensors. Example 1 and example 2 are results for two different cross-correlation models.

A right-shift would have included more type two sensors in the final sensor mix, while a left-shift would have included more type one sensors. It should also be noted, that the closer the A7 parameter values for all four signal-plus-noise models (CORR(X_1,X_1), CORR(X_2,X_2), CORR(X_1,X_2), CORR(X_2,X_1)) were, the flatter was the sensor mix curve determined, reflective of the fact that neither sensor type, under such circumstances, was predominantly better than the other.

Obviously the more sensors used and the faster the sampling rate employed, the better will be the results obtained from the field study. However, the financial, engineering, and logistic constraints limit the number of stations and the sampling rates available, making alternative configurations necessary. The decisions which determine these configurations can be made more easily, more correctly, and more objectively if the kind of information discussed in this chapter was available.

CHAPTER VIII

SUMMARY AND CONCLUSIONS

The purpose of this research was the development of a design package which could be used as an aid in the planning of forthcoming field experiments. The resultant package consisted of two major components; an analysis technique which allowed for the investigation of forthcoming experiments in terms of previous experimental results, and a design technique which permitted the incorporation of these investigative results into the determination of a best sampling configuration in order to resolve the proposed null hypothesis of the intended field study.

In Chapter IV, the multiple regression model was developed for the analysis of single as well as multiparameter data sets. It was shown that the technique presupposed no structure on the data set to be analyzed, but instead determined a signal-plus-noise model of the observed phenomenon, and then imprinted this information on the resultant regression weights. The structure model employed was four dimensional, as well as anisotropic, in order to reflect the spatial-temporal dimensionality encountered in actual experimental environments. An NLP algorithm was implemented in order to determine the pertinent structure function parameters.

Using this analysis approach, the inter- and intra-correlation structures were explored for a two parameter data set. The observations consisted of surface-gauge precipitation measurements as well as reflectivity estimates of rainfall rates. A determination of the signal-plus-noise configuration for the four storm realizations was accomplished, and analyses of these fields were subsequently produced. It is essential to understand that the purpose of this research was not to investigate precipitation patterns nor Z-R relationships nor the relative merits of univariate vs. multivariate data configurations; the purpose was to demonstrate the potentials of the developed analysis package. The results presented in Chapter V show the type of investigations that could be made if this analysis technique were employed. These results indicate that:

1. The multivariate data configuration produced the consistently better analyses when compared to the univariate configurations explored, the conclusion being that the analysis approach used added an additional and worthwhile dimensionality to the investigation of the geophysical data.

2. The Marshall-Palmer Z-R relationship, as opposed to the empirically derived Z-R relationship, did not normally produce noticeable differences in the results obtained when used in a multivariate environment. Considering the cost of obtaining the aircraft derived conversions, the resultant benefits could be questioned. However, whether or not the differences obtained were "significant" is a decision which must be based on the experiment for which the measurements were made.

3. The radar data alone produced surprisingly good analyses when compared to the surface-gauge results, if the purpose of the analysis was to estimate parameter values (procedure 2) as opposed to filtering the data field (procedure 1). Couple this with the fact that many more radar bins of information were potentially available than were actually used in obtaining these comparisons, and the relative worth of the reflectivity rainfall estimates becomes even more attractive. A thorough quantitative investigation of the quality of radar precipitation measurements would seem to be a worthwhile undertaking. Using this analysis package in the manner demonstrated, could aid in this task.

4. The analysis technique not only filtered the noise from the signal-plus-noise observation set, but also produced filtered analyzed fields that were distortion free. This said much for the credibility of the technique itself.

It should be noted that the above interpretations of the results of Chapter V were based on the assumption that "the best analysis was the one which produced the smaller error variance."

In Chapter VI, a design model was developed which made use of much of the information discerned from the previous chapters. The model incorporated the analysis scheme with an NLP algorithm in order to determine and investigate potential network configurations, including sensor number, sensor placement, sampling rate, etc. The designs developed were based on the comparison of modeled residual variance values, which tuned the sensor configurations to the specified modeled phenomenon and the climatology supplied.

As before, the results of a design effort were presented, not for the sake of the design itself, but only as a demonstration of the potentials of using this approach. Based on the type of information supplied, tuned to a particular experimental problem, the decisions faced by the principal investigator would be much easier to make. The design approach can be used for deployment of stationary sensor systems, mobile sensor systems (radiosondes, instrumented aircraft, etc.), or combinations thereof. Configurations for a single phenomenon were explored, but configurations based on a variety of intermeshed and related "system" specifications can be produced.

The development of the aforementioned design concepts into a workable systems configuration was the primary goal of this research. It is hoped that forthcoming investigations would make use of these design tools. Future investigations might include:

1) a continuation of the investigation of the precipitation measurement problem, in order to quantify the worth of radar reflectivity in estimating surface precipitation. Numerous cases for a variety of storm types could be examined, using the null hypothesis that the ability of the radar to measure surface precipitation accurately is a function of storm type, as well as storm size, shape, orientation, and intensity.

2) extending the design configuration to be able to investigate multivariate data sets of 3 and more parameters.

3) an actual design of a field experiment, in terms of the types of sensors to be used, the sensor numbers and the sensor configurations.

4) further investigation of the effect of non-white noise on the design process.

5) extending the structure function models, so that they more accurately reflect any non-stationarity in the phenomenon being investigated.

In conclusion, the techniques described herein have already been used, by Dr. Eddy's research group, in a variety of investigations for a variety of reasons. The results obtained have clearly demonstrated the worth of this design package. For experiments involving geophysical phenomena, the potential returns of employing the algorithms discussed would appear to be substantial. They have merely to be used.

REFERENCES

Objective Analysis

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

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

Brady, P.J., 1975a: "Implementation of an Objective Analysis Package on the WSMR Univac 1108," Technical Report for Task Order 75-213, White Sands Missile Range Atmospheric Science Group, B. Miers.

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.

Druyan, L.M., 1972: "Objective Analysis of Sea-Level Winds and Pressures Derived from Simulated Observations of a Satellite Radar-Radiometer and Actual Conventional Data," J. Appl. Meteor., Vol. 11, pp 413-428.

Druyan, L., 1973: "Objective Sea-Level Pressure Analysis for Data-Sparse Areas," J. Appl. Meteor., Vol. 12, pp 255-263.

Duquet, R.T., Danielsen, E.F., and N.R. Phares, 1966: "Objective Cross Section Analysis," J. Appl. Meteor., Vol. 5, pp 233-245.

Eddy, G.A., 1963: "Kinematic Divergence and Large Scale Energy Conversion," Arctic Meteorological Research Group, Department of Meteorology, McGill University, Montreal, Publication #60, Scientific Report #8.

Eddy, G.A., 1967: "The Statistical Objective Analysis of Scalar Data Fields," J. Appl. Meteor., Vol. 6, pp 597-609.

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

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

Eddy, G.A., 1975: "Rain Gauge Network Analysis," Contract No. 14-060D-7633, Atmospheric Water Resources Management Division U.S. Bureau of Reclamation.

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

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.

Epstein, E.S., 1969: "Stochastic Dynamic Prediction," Tellus, Vol. 21, pp 739-759.

Epstein, E.S., and E.J. Pitcher, 1972: "Stochastic Analysis of Meteorological Fields," J. Atmos. Sci., Vol. 29, pp 244-257.

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

Inman, R.L., 1970: "Papers on Operational Objective Analysis Schemes at the National Severe Storms Forecast Center," NOAA Technical Memorandum ERLTM-NSSL 51.

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

Klein, W.H., 1963: "Specification of Precipitation from the 700 Millibar Correlation," Mon. Wea. Rev., Vol. 91, pp 527-536.

Kruger, H.B., 1968: "General and Special Approaches to the Problem of Objective Analysis of Meteorological Variables," Quart. J. Roy. Meteor. Soc., Vol. 95, pp 21-39.

Leary, C. and Rory Thompson, 1973: "Shortcomings of an Objective Analysis Scheme," J. Appl. Meteor., Vol. 12, pp 589-594.

Lewis, J., and T. Grayson, 1972: "The Adjustment of Surface Wind and Pressure by Sasaki's Variational Matching Technique," J. Appl. Meteor., Vol. 11, pp 585-597.

Maine, R., and D.J. Gauntlett, 1968: "Modification to an Operational Numerical Weather Analysis System and Application to Rainfall," J. Appl. Meteor., Vol. 7, pp 18-28.

Morrison, D.P., 1967: Multivariate Statistical Methods, McGraw-Hill, Inc., New York.

Olsen, A.R., 1975: "Bayesian and Classical Statistical Methods Applied to Randomized Weather Modification Experiments," J. Appl. Meteor., Vol. 14, pp 970-973.

Paegle, J.N., 1974: "Prediction of Precipitation Probability Based on 500-mb Flow Types," J. Appl. Meteor., Vol. 13, pp 213-220.

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

Pasteris, P.A., 1975: "Objective Analysis of Northeast Oklahoma Rainfall," M.S. Thesis, Department of Meteorology, University of Oklahoma.

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

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.

Snee, R.D., 1973: "Some Aspects of Nonorthogonal Data Analysis," J. Quality Technology, Vol. 5, pp 67-79.

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

Stout, G.E., and E.A. Mueller, 1968: "Survey of Relationship between Rainfall Rate and Radar Reflectivity in the Measurement of Precipitation," J. Appl. Meteor., Vol. 7, pp 465-473.

Takeuchi, D., R. Peace and S. Howard, 1974: "Evaluation of Airborne Raindrop Size-Distributions Measurements for Improving Radar Rainfall Estimates," technical report MRI 75 FR-1294, Atmospheric Water Resources Management Division, U.S. Bureau of Reclamation.

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

Wiggert, V., and S. Ostlund, 1975: "Computerized Rain Assessment and Tracking of South Florida Weather Radar Echoes," Bull. Amer. Meteor. Soc., Vol. 56, pp 17-26.

Wilson, J.W., 1970: "Integration of Radar and Raingage Data for Improved Rainfall Measurement," J. Appl. Meteor., Vol. 9, pp 489-497.

Woodley, W.L., and A. Herndon, 1970: "A Raingage Evaluation of the Miami Reflectivity Rainfall Rate Relation," J. Appl. Meteor., Vol. 9, pp 242-257.

Woodley, W.L., Olsen, A.R., Herndon, A., and V. Wiggert, 1975: "Comparison of Gage and Radar Methods of Convective Rain Measurement," J. Appl. Meteor., Vol. 14, pp 909-928.

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

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

Zawadzki, I.I., 1973: "Statistical Properties of Precipitation Patterns," J. Appl. Meteor., Vol. 12, pp 459-472.

Correlation Analysis

Anderson, J.R., 1970: "Rainfall Correlations in the Pastoral Zone of Eastern Australia," Australian Meteor. Magazine, Vol. 18, pp 94-101.

Bernard, M., 1943: "Discussion of 'Precipitation Runoff Relationships as a Basis for Water-Supply Forecasting'," by George D. Clyde and R.A. Work, Trans. Amer. Geoph. Union, Part 111, pp 49-55.

Boyd, D.A., 1939: "Correlations between Monthly Rainfall at Eleven Stations in the British Isles," Memoirs Roy. Met. Soc., Vol. 4, pp 143-160.

Brady, P.J., 1975b: "Matching Rain Gauge Placement to Precipitation Patterns," Proceedings of the National Symposium on Precipitation Analysis for Hydrologic Modeling, Davis, California, June 26-28, pp 111-122.

Buell, C.E., 1972: "Correlation Functions for Wind and Geopotential on Isobaric Surfaces," J. Appl. Meteor., Vol. 1, pp 51-59.

Caffey, J.E., 1965: "Inter-station Correlations in Annual Precipitation and in Annual Effective Precipitation," Colorado State University Hydrology Papers, No. 6.

Cornish, E.A., Hill, G.W., and M.J. Evans, 1961: "Interstation Correlations of Rainfall in Southern Australia," Commonwealth Scientific and Industrial Research Organization, Div. Math. Stat., Tech. Paper 10.

Court, A., 1958: "Wind Correlations," Final Report, contract 19(604)-2060, Cooperative Research Foundation, San Francisco.

Court, A., 1975: "Rainfall Correlations, 1922-1961," Proceedings of the National Symposium on Precipitation Analysis for Hydrologic Modeling, Davis, California, June 26-28, pp 163-170.

Fisher, R.A., and W.A. Mackenzie, 1922: "The Correlation of Weekly Rainfall," Quart. J. Roy. Meteor. Soc., Vol. 48, pp 234-245.

Gandin, L.S., Mashkovich, S.A., Alaka, M.A., 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.

Gringorten, I.I., 1973: "Stochastic Modelling of the Areal Extent of Weather Conditions," U.S. Air Force Cambridge Research Labs., Environmental Research Paper No. 459.

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

Horton, R.E., 1923: "Rainfall Interpolation," Mon. Wea. Rev., Vol. 51, pp 291-304.

Hutchinson, P., 1969: "Estimation of Rainfall in Sparsely Gauged Areas," Bull. Int. Assn. Sci. Hydrol., Vol. 14, pp 101-119.

Hutchinson, P., 1970: "A Contribution to the Problem of Spacing Raingages in Rugged Terrain," J. Hydrol., Vol. 12, pp 1-14.

Hutchinson, P., 1971: "Accuracy of Estimates of Daily Areal Mean Rainfall," Weather, Vol. 26, pp 366-371.

Jenkins, G.M., and D.G. Watts, 1968: Spectral Analysis and its Applications, Holden-Day, San Francisco.

Kraemer, H.C., 1974: "The Non-null Distribution of the Spearman Rank Correlation Coefficient," J. Amer. Stat. Assn., Vol. 69, pp 114-117.

Longley, R.W., 1974: "Spatial Variation of Precipitation over the Canadian Prairies," Mon. Wea. Rev., Vol. 102, pp 307-312.

McDonald, J.E., and C.R. Green, 1960: "A Comparison of Rank-Difference and Product-Moment Correlation of Precipitation Data," J. Geophys. Res., Vol. 65, pp 333-336.

Miller, E.R., 1931: "Interpolation of Rainfall Data by the Method of Correlation," Mon. Wea. Rev., Vol. 59, pp 35-36.

Mitropol'skii, A.K., 1966: Correlation Equations for Statistical Computations, Consultants Bureau, New York.

Oltman, R.E., and H.J. Tracey, 1951: "Trends in Climate and in Precipitation-Runoff Relation in Missouri River Basin," USGS Circ. 98.

Rodriguez-Iturbe, I., 1974: "The Design of Rainfall Networks in Time and Space," Water Resources Research, Vol. 10, pp 713-728.

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

Stenhouse, N.S., and E.A. Cornish, 1958: "Inter-station Correlation of Monthly Rainfall in South Australia," Commonwealth Scientific and Industrial Research Organization, Div. Math. Stat., Tech. Paper 5.

Stidd, C.K., 1954: "The Use of Correlation Fields in Relating Precipitation to Circulation," J. Meteor., Vol. 11, pp 202-21[illegible]

Stol, P.T., 1972: "The Relative Efficiency of the Density of Rain-Gage Networks," J. Hydrol., Vol. 15, pp 193-208.

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: "Anisotropic Correlation Functions for Objective Analysis," Proceedings of the Fourth Conference on Probability and Statistics in the Atmospheric Sciences, Tallahassee, Florida, Noverber 18-21, pp 92-95.

Zawadzki, I.I., 1973: "Measurement and Application of Rainfall Autocorrelation Functions," Stormy Weather Group, Scientific Report MW-76, McGill University.

NLP

Barnes, G.K., and D.M. Himmelblau, 1966: "Optimization of a Kraft Pulping Mill," paper at Canadian Institute of Chemistry Conference, October, 17.

Beveridge, G.S.G., and R.S. Schechter, 1970: Optimization: Theory and Practice, McGraw-Hill, Inc., New York.

DiBella, C.W., and W.F. Stevens, 1965: "Process Optimization by Non-linear Programming," Ind. Eng. Chem. (Process Design Develop.), Vol. 4, pp 16-20.

Fiacco, A.V., and G.P. McCormick, 1964: "The Sequential Unconstrained Minimization Technique for Non-linear Programming: A Primal Dual Method," Management Sci., Vol. 10, pp 360-366.

Fletcher, R., and C.M. Reeves, 1964: "Function Minimization by Conjugate Gradients," Computer J., Vol. 7, pp 149-154.

Fletcher, R., 1965: "Function Minimization without Evaluating Derivatives - a Review," Computer J., Vol. 8, pp 33-41.

Fletcher, R., 1970: "A New Approach to Variable Metric Algorithms," Computer J., Vol. 13, pp 317-322.

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

Himmelblau, D.M., 1974: Personal communication.

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

Paviani, D.A., 1969: Nonlinear programming, Ph.D. dissertation, University of Texas, Austin, Texas.

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

Powell, M.J.D., 1964: "An Efficient Method for finding the Minimum of a Function of Several Variables without Calculating Derivatives," Computer J., Vol. 7, pp 155-162.

Rosenbrock, H.H., 1960: "An Automatic Method for finding the Greatest or Least Value of a Function," Computer J., Vol. 3, pp 175-184.

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

Stewart III, G.W., 1967: "A Modification of Davidon's Minimization Method to Accept Difference Approximations of Derivatives," J.A.C.M., Vol. 14, pp 72-83.

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

Zangwill, W.I., 1969: Nonlinear Programming - A Unified Approach, Prentice-Hall, Inc., New Jersey.

APPENDIX

In the development of the analysis technique for implementation as a computer algorithm, a number of unforeseen practical considerations were made apparent. These problems, along with their respective solutions, are the subject of the following discussion.

1. Axis Rotation

When the implemented analysis technique was initially used for analyzing modeled as well as observed precipitation patterns, it was found that the shape and form of the analyzed patterns did not correspond very well with the shape and form of the original precipitation fields. For example, if a modeled circular system was analyzed, the resultant pattern would often turn out to be eliptical. This implied that the analysis algorithm was somehow distorting the original parameter field. The solution to this problem was twofold. First, it was found that by performing an appropriate axis rotation, it was possible to improve the analyzed patterns considerably. The development of the rotation algorithm used was based on the work of Tatsuoka (1971), where the basic idea was to determine a rotated axis which would maximize the variance of the original station placement pattern. This rotated axis was used in all subsequent analysis procedures for that particular station placement pattern. The results of this rotation can be seen by comparing the structure functions of the May 21st data set in Figure 4 for the rotated CORR(X_2X_2) model and in Figure 95 for the non-rotated CORR(X_1X_1) model. The shape and size are similar (they won't be exact because one correlation is for the radar-reflectivity data and the second is for the surface-gauge data) but the respective orientation is different. The analyses produced are always presented with respect to the original spatial axis.

The second part of the solution to the problem was to make the correlation function model anisotropic, as was discussed in Chapter IV, section B.

2. Signal-plus-noise model

In generating the raw correlation matrices which are used to determine the appropriate parameters of the structure model, it is necessary to exercise some care in the selection of the number of spatial and temporal lags determined, as well as the separation distance to which each of these lags corresponds. It is essential to make these selections so that a sufficient number of station pairs is included in each discrete lag category to discern the signal-plus-noise structure of the observation set adequately, keeping in mind that too many lags, especially in time, will tend to smooth the structure function so as to destroy any actual pattern present.

For example, if the average station spacing in the observation network is 5.0 nautical miles between sensors, then an appropriate lag "distance" might be approximately 5.0 nautical miles. A larger lag distance, say 50.0 nautical miles, would group too much information into each discrete lag bin, while a smaller lag distance, say 0.5 nautical miles, would not group nearly enough information to be able to discern the structure present. The appropriate lag distances as well as the number of lags to be determined can be rather easily deduced by comparing the output of the correlation generation programs with the structure function parameters found by the N.L.P. algorithm. If the correlation pattern coincides with the structure function parameters, then the appropriate lags and distances were used. If they do not coincide, then different sets of lag values and/or lag distances should be investigated.

3. Analysis program

If a disproportionate number of predictors used in the determination of the regression weights have negative correlation coefficients, the $X^tV^{-1}X$ matrix tends to become singular and difficult to invert, and the elements of the $\hat{\beta}$ vector tend to become unstable. In order to alleviate this situation, two subroutines are used in the analysis program for the determination of the structure function values, the first reduces the "radius" of influence so as to include only those potential data values whose correlation coefficients are positive, the second permits the computation of the respective elements of the $X^tV^{-1}X$ matrix and the $X^tV^{-1}Y$ vector, regardless of distance separation. The "radius" of influence is defined to be that distance from a predictand out to the inflection point of the negative exponential function for each respective spatial and temporal direction.