

THE EFFECT OF STORM ON RUNOFF

By NA Patrick

WRC NO 183 /5/93

THE EFFECT OF STORM PATTERNS ON RUNOFF

by

N.A. PATRICK

**Water Systems Research Group
University of the Witwatersrand
Johannesburg, 2000, South Africa**

**Water Systems Research Group
Report No.5
1992**

**Report to the Water Research Commission on the Project
"Effects on Urbanization on Catchment Water Balance"**

**Head of Department
and Project Leader : Professor D. Stephenson**

WRC REPORT NO. 183/5/93

ISBN 1 86845 033 3

ISBN SET NO. 1 86845 040 6

WATER RESEARCH COMMISSION

EFFECTS OF URBANIZATION ON CATCHMENT WATER BALANCE

LIST OF REPORTS

<u>Report No.</u>	<u>Title</u>	<u>By</u>
1	Analysis of effects of urbanisation on runoff	D. Stephenson
2	Description of catchments and methodology	J.J. Lambourne T. Coleman
3	Geohydrology of Catchments	W.A.J. Paling D. Stephenson
4	A hydrometeorological data management package Wits Data Management System WITDMS	J.J. Lambourne
5	The effect of storm patterns on runoff	N. Patrick
6	Runoff Modelling	A. Holden
7	Streamflow modelling	P. Kolovopoulos
8	Runoff management modelling	T. Coleman D. Stephenson
9	Catchment water balance	J.J. Lambourne F. Sutherland
10	Urban Runoff Quality & Modelling Methods	T. Coleman
11	Compendium of papers published on the research	
12	Executive Summary	D. Stephenson

ABSTRACT

Five methods of interpolating three-dimensional surfaces were examined in a literature survey, and as a result four of these were tested for their ability to satisfy a series of accuracy goals. The methods tested are known as Inverse Squared Distance, Multiquadratic interpolation, Polynomial surface fitting and Distance weighted least-squares interpolation. The method that performed the best, Inverse Squared Distance, was used for the subsequent research carried out.

A computer program was developed that enabled the plotting of contours of rainfall intensity over small time steps. This was applied to a 10.36Km² area peri-urban catchment using data from five rain-gauges monitoring the catchment. A modification of a well tested overland flow model, WITWAT, was used to generate runoff hydrographs for the catchment and hence the spatial variation of real storm events were maintained. The shapes of the resultant runoff hydrographs were compared with other methods of rainfall distribution.

SYNOPSIS

The estimation of runoff from a catchment is an important process that underlies the majority of hydrological planning and design studies. Estimation can be divided into three distinct aspects; rainfall falling to the ground, the movement of the water over and through the land surface and the movement of the water through channels. The second and third aspects form the focus of most runoff models, and this is being modelled with increasing degrees of detail and discretisation. The apparent gap in detail covering the first aspect, that of rainfall falling to the ground surface, forms the basis for this project.

Several computerised methods of describing detailed distributions of rainfall patterns from recorded rainfall data were studied. The most suitable method was selected using the criterion of how well it satisfied a series of general accuracy goals. To achieve a thorough evaluation, extensive testing was carried out. The results and conclusions of this testing process is reported upon.

A method of producing contour maps of rainfall intensity for time steps over the duration of the storm event was developed. Rainfall data from a peri-urban catchment of about 10km² area were used and a large number of storm events studied in this way. It was discovered that a single storm event was composed of several cells, whose behaviour is extremely erratic. These cells are of small aerial extent (about 4km diameter) and short duration (mostly less than 20 minutes). The study of cells has formed the subject of several papers by other authors, and pertinent points from these papers have been included in this study.

The visual display of changing storm patterns with time provides a means of analysing features of storm events which is not possible with statistically based methods, such as cross correlation techniques. The development and decay of storm cells could be examined, and characteristics determined. Widely varying shapes of contour maps were noticed in the study: Each storm intensity contour map represented 5 minutes of storm event, and changes from one 5 minute interval to the

subsequent showed convoluted, spotty, selective and planar distributions of rainfall over a small catchment and over small time spans. This presented a far more complex interpretation of storms than depicted by simplistic design storms which have constant intensity over the storm duration or Thiessen distributions which allocate rainfall to weighted areas contributed to by the nearest gauge. Both these methods are commonly used to describe rainfall input for rainfall/runoff process models.

An adapted version of a well tested runoff model was used to determine the effects that different shapes of rainfall patterns used as an input would have on runoff hydrographs. This made it possible to distribute real rainfall events spatially over the catchment, as well as temporally. Hence real storms were broken down to five minute time intervals, and the rainfall intensity from each time interval distributed over the catchment in accordance with the actual shape of the storm for that time interval. It was also possible to use conventional rainfall distribution methods as an input to the runoff model.

In this way several types of rainfall input were used for the purposes of comparison; single gauges to represent the whole catchment, Thiessen weighted averages of the existing gauges and the spatial distribution method selected (Inverse Squared Distances). The resultant hydrographs exhibited widely varying shapes for the same event. This indicated that current popular methods of rainfall representation for runoff models do not agree with each other (when studying the same event), or with a more accurate representation of real rainfall events.

The main conclusion, and suggestion for further study is that a new generation of runoff models should be developed which include a spatial distribution of rainfall that accurately describes real events. The benefits derived from the current trend to describe the physical parameters of catchments (slope, ground cover, soil permeability properties etc.) in great detail over smaller and smaller areas or sub-areas within a study catchment is largely negated when the rainfall input is distributed as a lumped occurrence over the catchment.

CONTENTS

Page

ABSTRACT	i
SYNOPSIS	ii
CONTENTS	iv
LIST OF FIGURES AND TABLES	vi

1 GENERAL INTRODUCTION

1

2 STORM ORIGINS

1

2.1 CAUSES OF PRECIPITATION	4
2.1.1 Convection	4
2.1.2 Orographic	5
2.1.3 Frontal	5
2.2 CLOUD TYPES	7
2.3 SCALES OF STORM TYPES	9

3 PREVIOUS WORK ON STORM CHARACTERISTICS

11

3.1 WEATHER PREDICTION	11
3.2 MICROSCALE STUDIES	12
3.2.1 England	13
3.2.2 Tunisia	14
3.2.3 Israel	15
3.3 OTHER FACTORS	16
3.4 EFFECTS ON RUNOFF	17
3.5 STUDY PROPOSAL	18

4 GRIDDING PROCESS

20

4.1 INTRODUCTION	20
4.2 LITERATURE SURVEY	24
4.3 POLYNOMIAL SURFACES	28
4.4 MULTIQUADRATIC SURFACES	33
4.5 INVERSE DISTANCE WEIGHTED SURFACE	38

<u>CONTENTS</u>	Page
4.6 DISTANCE WEIGHTED QUADRATIC SURFACES	41
4.7 KRIGGING	44
4.8 THIESSEN WEIGHTS	47
4.9 TESTING PROCESS	50
4.10 DISCUSSION ON GRIDDING METHODS	55
<u>5 CONTOURING PROCESS</u>	71
<u>6 STUDY CATCHMENT</u>	73
6.1 STORM PATTERN CHARACTERISTICS	75
6.1.1 Findings from contour maps	77
6.2 TOPOGRAPHICAL EFFECTS	87
6.3 CELL SIZES AND DURATIONS	88
6.4 EFFECTS OF HYETOGRAPHS ON RUNOFF	91
6.5 DISCUSSION	98
<u>7 DISCUSSION</u>	100
<u>REFERENCES</u>	104

LIST OF FIGURES AND TABLES

LIST OF FIGURES

Figure	Page
2.1 Cold front	6
2.2 Warm front	6
2.3 Occluded front	7
2.4 Cloud types	8
4.1 Raw data	21
4.2 Surface one	21
4.3 Surface two	22
4.4 Surface three	22
4.5 Block diagram of Krumbein-type coefficients	30
4.6a Data set 1	57
4.6b Contours for data set 1	57
4.7a Data set 2	58
4.7b Contours for data set 2	58
4.8a Data set 3	59
4.8b Contours for data set 3	59
4.9a Inverse Squared Distance (ISD) applied to test1a Sum of Squared Residuals	61
4.9b Polynomial (POL) applied to test1a Sum of Squared Residuals	61
4.9c Multiquadratic (MQUAD) applied to test1a Sum of Squared Residuals	61
4.9d Distance Weighted (DIST) applied to test1a Sum of Squared Residuals	61
4.10a Inverse Squared Distance (ISD) applied to test2a Sum of Squared Residuals	62
4.10b Polynomial (POL) applied to test2a Sum of Squared Residuals	62
4.10c Multiquadratic (MQUAD) applied to test2a Sum of Squared Residuals	62
4.10d Distance Weighted (DIST) applied to test2a Sum of Squared Residuals	62

LIST OF FIGURES

Figure	Page
4.11a Inverse Squared Distance (ISD) applied to test3a Sum of Squared Residuals	63
4.11b Polynomial (POL) applied to test3a Sum of Squared Residuals	63
4.11c Multiquadratic (MQUAD) applied to test3a Sum of Squared Residuals	63
4.11d Distance Weighted (DIST) applied to test3a Sum of Squared Residuals	63
4.12a Comparison of methods - test1a Percentage of (AAME/MEAN RAIN)	64
4.12b Comparison of methods - test2a Percentage of (AAME/MEAN RAIN)	64
4.12c Comparison of methods - test3a Percentage of (AAME/MEAN RAIN)	64
4.13 Comparison of methods AAME (100 data points)	64
4.14a Plot of test1a using all 100 gauges Method used is: ISD	67
4.14b Plot of test1a using all 100 gauges Method used is: POL	67
4.14c Plot of test1a using all 100 gauges Method used is: MQUAD	68
4.14d Plot of test1a using all 100 gauges Method used is: DIST	68
4.15a Actual contours for set: test1a	69
4.15b Polynomial method applied to set: test1a All one hundred data points used in interpolation	69
4.16 Comparison of methods Time taken for gridding process in minutes	70

LIST OF FIGURES

Figure	Page
6.1 Contour map of study catchment: Montgomery Park	74
6.2 Thiessen Weights and areas for Montgomery Park	76
6.3 Contour maps for a segment of a storm	78
6.4 Absence of central gauge makes analysis difficult in this case	81
6.5 Spotty rainfall for the storm on 27/11/1987	81
6.6 Absence of cellular structure for the storm on 21/01/1988	82
6.7 Erratic behaviour of a portion of the storm on 24/01/1988	83
6.8 Movement of a cell over the study catchment	84
6.9 Partial coverage of catchment by storm event/cell	85
6.10 A central rain-gauge would clarify these two contour maps	86
6.11 Elevation effects on rainfall	87
6.12 Heytographs for the five gauges for the same storm event	89
6.13 Intensity-Duration-Frequency curves for Montgomery Park	93
6.14 Design storm Hydrographs	94
6.15 Hydrographs for the storm event on 15/10/1987	95
6.16 Hydrographs for the storm event on 6/11/1987	96
6.17 Hydrographs for the storm event on 10/11/1987	97

LIST OF TABLES

Table	Page
6.1 Land use percentages	73

1 GENERAL INTRODUCTION

This report is concerned with the effect that storm shape and movement has on runoff from a catchment. Runoff from a catchment is becoming more important in urban drainage design as pipe networks and retention structures have to be designed with the least cost for the least risk possible, and in rural catchment studies where volumes required for storage structures, erosion quantities and siltation of storage structures need to be evaluated.

Required storage volumes, erosion and siltation quantities, water quality variations and peak flow rates all depend on the runoff that occurs from a catchment, and more particularly after a storm event. To be able to estimate the volume of topsoil eroded from a field one has to be able to estimate the amount and energy of water available to move soil particles - i.e. runoff. Similarly, the load carried in a channel, siltation rates, maximum flow depths and maximum flow rates depend on the estimation of runoff. The understanding and modelling of the process that relates incident precipitation to resultant runoff is therefore vitally important to Hydrologists, Engineers and Planners.

The maximum rate of runoff possible from a storm event is particularly important, especially in flood analysis studies. This is very dependant on storm pattern - i.e. the position of the storm event over the catchment and how this changes with time, the distribution of rainfall intensity throughout the aerial coverage of the storm event and how the characteristics of the storm event change during the lifetime of the storm event. Very little attention is paid to these aspects and this project attempts to remedy this situation.

Our understanding of the physical processes that will produce rainfall from any set of atmospheric conditions are limited. Thus from the hydrologists point of view this aspect of the problem becomes empirical, or stochastic, as opposed to deterministic. There is an added complication in that collection of data is dependent on waiting for events to happen. By comparison, other branches of engineering can

often produce more samples of materials that are to be tested if there is a gap in the available data. Hence there is a large probability component in hydrological modelling. To reduce possible errors of estimation and design (of runoff peaks, volumes etc.), the physical (deterministic) processes that produce runoff from rainfall are being analysed in increasing degrees of accuracy. A concurrent study of the stochastic aspects of the process (patterns of rainfall) to the same degree of accuracy and analysis is called for.

Classical techniques for modelling the rainfall-runoff process involve arithmetic averaging of raingauge data at some stage. Whether it is an aerial averaging where the rainfall is averaged out over the area it fell in, as in the case of thiessen polygons for example, or if it is a temporal and occurrence averaging, where the number of events with a certain depth are grouped together and probabilities of occurrence are postulated, as in the case of design storms, a means of aggregating and simplifying the rainfall data is employed. With the growth in popularity of micro-computers and computer based packages it is possible to study events in smaller time steps and smaller spatial areas, hence increasing the degree of discretisation and accuracy.

Hydrological models which process infiltration, evapotranspiration, subsurface flow, overland flow and channel routing are available, and attempt to model the rainfall-runoff process to varying degrees of accuracy (e.g. Green and Stephenson, 1986). The complexity of some models and calculation requirements make these models ideal for computerisation.

Rainfall-runoff process models use rainfall as a cause in a cause-effect relationship. Hence the nature of the rainfall input will have a significant effect on the nature of the output of these process models. Commonly, "design storms" are used as an input and are characterized by constant rainfall intensities corrected by an aerial reduction factor and applied to the whole catchment. Other types of inputs include weighted averages of hyetographs from raingauges in the catchment and characteristic shapes or "templates" used to generate rainfall data (e.g. Chicago, bi-modal).

It is one intention of this report to show that there are several factors associated with storms that will have an effect on the shape of the runoff hydrograph for a catchment which are not accounted for with present models. Such factors include; the physical size of the catchment in relation to the real size of the storm event, the variation of ground levels within the catchment, and the duration, movement and type of storm event that occurs in the area.

For the purposes of this study, then, the rainfall-runoff process can be grouped into two concepts; the first being the pattern of rainfall as it strikes the ground at various levels of intensity, the second being the mechanisms that explain how the water moves from one area to another via the land surface and eventually through channels. The first concept is the basis for the research in this project, while the second is used in a "black-box" approach to show the effects of the first.

A computerised method that models detailed distributions of rainfall intensities was used to determine the effects of storm shape and movement on runoff. Contours of rainfall intensities over five-minute time intervals were plotted for several storm events, and the resultant runoff determined by an existing and well tested runoff model. Several other methods for distributing rainfall over a catchment were also used and these results compared.

2 STORM ORIGINS

The nature of the precipitation that reaches the ground must depend on the nature and the origin of the precipitation. Hence a brief introduction to storm types is given here. There are three basic causes of precipitation from the upper atmosphere; orographic, convective and frontal systems (which are again divided into warm fronts, cold fronts and occluded fronts). The size and nature of storm origins will influence the intensity and area affected by a storm event, and this effect is significant for the modelling of catchments.

This introductory chapter is largely based on "The Earth Sciences" by Strahler (1963) unless otherwise stated.

2.1 CAUSES OF PRECIPITATION

2.1.1 Convection

Convection storms are caused by warm, moist air rising and becoming unstable. These storms produce some of the most severe storm events with torrential downpours, high wind speeds and are often accompanied by lightning.

A trigger is necessary to cause ascent of an air mass, and this is usually ascribed to a warm surface that the air mass passes over, thus heating the air and causing it to rise. There are several other causes though, such as latent heat of condensation, which will cause heating of an air mass and also produce convectational movement.

Once a trigger exists to initiate the vertical movement of air, the next criterion occurs if the air becomes unstable. This implies that a parcel of air becomes warmer than the surrounding air (for reasons intrinsic to the air mass, and relative to the surrounding air) and rises of its own accord. The scale and rate of rising and amount of moisture present in the air will determine the severity of the storm event.

2.1.2 Orographic

In this type of event, a physical obstruction (such as a mountainside or escarpment) forces an air mass to rise, when prevailing winds cause movement of air horizontally. The resultant precipitation is generally of two degrees of severity; a light persistent rain where the air mass is essentially stable or heavy convection type showers where the air mass has become unstable.

This effect may be visible on any continent where there is a mountain range near the coastline (e.g. California area on the North American continent and Natal area on Southern Africa) with the windward side of the mountain range characterized by high rainfall areas and the leeward side characterized by low rainfall and hot, dry conditions.

In equatorial and tropical regions orographic rainfall is mostly of the severe type and produces violent monsoon events, and is the source of many record storm events.

2.1.3 Frontal

A front is the term applied to the boundary or interface where two dissimilar air masses meet. The main differences between air masses is in temperature, density and moisture content. These differences appear trivial, but in actual fact air masses do not mix well at all. It is better to imagine two air masses meeting as in water and oil meeting, rather than two similar miscible air masses of differing temperatures.

There are three types of frontal systems; cold fronts, warm fronts and occluded fronts, each with different characteristics with potential for producing different storm events.

.Cold fronts A cold air mass invades an area occupied a warm air mass, and because of the difference in temperatures, forces the warm air to rise.

Horizontal movement of the air and ground friction produces a blunting of the cold wedge near the ground surface, which can result in an abrupt lifting of the warm air and the production of unstable air (see figure 2.1).

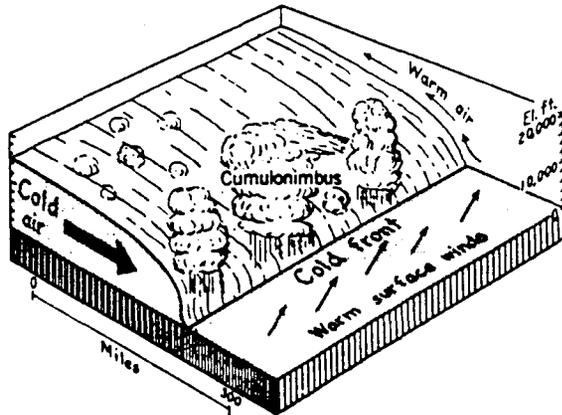


FIGURE 2.1 COLD FRONT
(After Strahler, 1963)

.Warm fronts

A warm air mass invades an area occupied by a cold air mass. A more gentle gradient is produced as the warm air moves up the face of the cold air. These fronts tend to move more slowly and cover a wider expanse than the cold fronts (see figure 2.2).

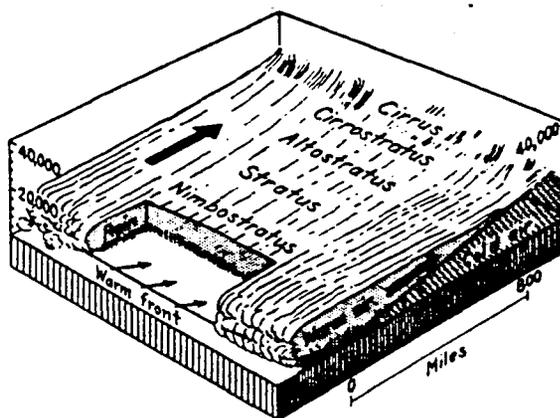


FIGURE 2.2 WARM FRONT
(After Strahler, 1963)

.Occluded fronts This occurs when a cold front has caught up with and pushed into a warm front. Relatively speaking there are three separate masses of air; a cold mass, a less cold mass and a warm mass. The warm mass is trapped between the other two masses and forced up (see figure 2.3).

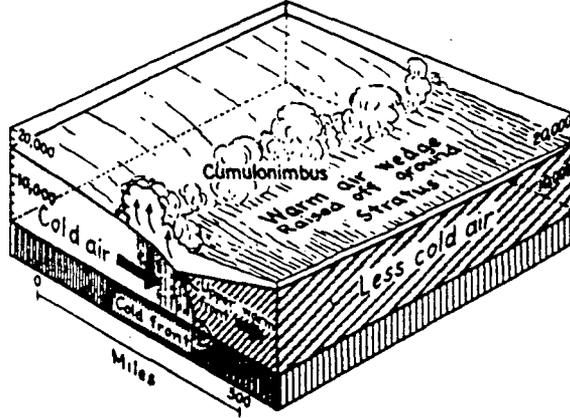


FIGURE 2.3 OCCLUDED FRONT
(After Strahler, 1963)

The severity of the storm event is determined by the degree of instability in the air masses. Other factors such as moisture content, prevailing atmospheric conditions and topographical effects will determine the size, duration and movement of the storm event. These factors establish the nature of the resultant precipitation.

2.2 CLOUD TYPES

Visible evidence of water in the atmosphere is seen by the shape or type of clouds, and the resultant nature of precipitation. To become visible, clouds need water droplets or ice crystals in the order of 0.02mm to 0.06mm diameter. These form by condensation around tiny hygroscopic nuclei (such as chemical compounds suspended in the atmosphere). Growth occurs (when there is water available) by more condensation and collision of particles. This growth becomes rapid as the saturation point of the air mass is approached.

In the upper layers, water can exist in a supercooled state as a liquid far below zero degrees Celsius. Clouds higher than 6.5km are however generally composed entirely of ice.

There are two generic terms for cloud shapes:

- .Stratiform These are blanket-like layers of cloud and occur when the air moves in laminar type layers. There is some mixing, but the extent of vertical movement is small.
- .Cumuloform These are flat-based globular shaped clouds. They often extend to very high levels and are formed by rising air currents. The clouds are comprised of cells that rise in a series of bubble-like masses rather than a continuous column. As they move vertically upward, air is entrained laterally from the surrounding air.

Clouds are divided into four categories, defined by the height at which the cloud is situated and by the process of their formation. This is shown in figure 2.4.

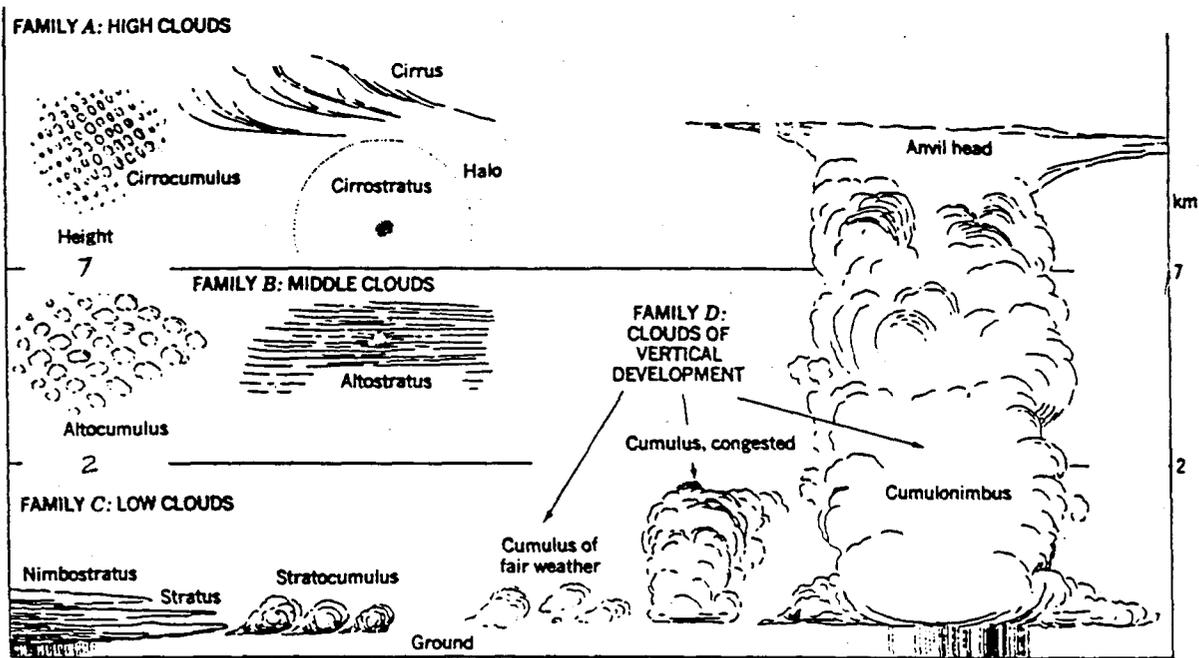


FIGURE 2.4 CLOUD TYPES
(After Strahler, 1963)

- .High clouds These are usually higher than 7km and are composed entirely of ice crystals.

- .Middle clouds These are found at heights between 2km and 7km. Altcumulus usually indicate fair weather while altostratus frequently produce rain.

- .Low clouds These occur from ground round level to 2km above the surface. Stratocumulus are best seen in the clearing period following a storm, while nimbostratus are evident as a dense, grey blanket and may produce heavy rain. They may extend up into the middle layer.

- .Vertical clouds Clouds formed as a result of vertical movement of air are the thunderstorm type. They may grow from cumulus (with cauliflower shaped tops) to cumulonimbus or thunderheads, which produce violent rains and storms. These clouds occasionally extend to heights of 18km in the tropics.

2.3 SCALES OF STORM EVENTS

The above discussion is important to the hydrologist. The physical size of the catchment, the geographical location and nature of relief changes within the catchment to be modelled will influence the type of storm event which can be expected in that area. The aerial extent of the storm event must also be related to the size of the catchment. If a catchment is to be realistically modelled, then the characteristics of possible storm events must be modelled as accurately as necessary, in addition to the description of the physical parameters of the catchment.

The discussion above is related to intensity of rainfall that can be expected from cloud cover. The discussion below is related to the aerial extent of the cloud cover. It is convenient to adopt the definitions of size or scale used by meteorologists. Eagleson (1970) describes these as:

.Microscale This describes areas with diameters less than about 5km or 6km and convective cells predominate. A stationary observer would see short intense bursts of rainfall. The cells move with an apparently random motion, with speeds of 30 to 65km/h. The propagation is not normally in the same direction as the mean wind in the cloud layer.

.Mesoscale This describes areas of about 5km to 50km. The conditions that enable convective cells to occur remain for longer periods than the individual cells, hence an agglomeration of cells would define a mesoscale event. A thunderstorm is a mesoscale event.

.Synoptic scale This describes areas of up to several hundred kilometers in extent, and is a description of frontal systems. This is the scale of most satellite photographs of weather conditions, and shows many mesoscale events linked together.

Thus for a more accurate study on a catchment, the size of the catchment should be compared with the scales listed above, and this information used in conjunction with an estimation of expected rainfall intensity to determine the nature of the storm events to be applied to the area under study.

The focus of this project is to determine what effect storm shape variations can have on the resultant runoff from a catchment.

3 PREVIOUS WORK ON STORM CHARACTERISTICS

Several diverse aspects of spatial and temporal phenomena of storms have been researched. What follows is a summary of the most pertinent to this project.

3.1 WEATHER PREDICTION

(After Gauntlett and Leslie, 1975)

Models to predict weather patterns for areas have been developed. These models are mostly numerical in nature and fall under the general heading of Numerical Weather Prediction (NWP) models. This forms an alternative to statistically and empirically based inputs for catchment simulation models. In this case, NWP models exclude Planetary Boundary Layer models (which deal with very low-level effects such as sea-breezes, fog, urban pollution etc.).

The lower limit of time resolutions available with these models is down to about one hour, and the lower limit of spatial resolutions is down to areas of about two hundred square kilometers. Hence these models deal with macroscale (synoptic) and mesoscale events. Microscale, or single cell events are not modelled. A drawback of macroscale NWP models is that they tend not to be able to incorporate the effects of topography with sufficient accuracy.

At the time Gauntlett and Leslie wrote their paper, there were three main drawbacks to NWP models:

- . It is generally both difficult and expensive to measure atmospheric conditions at sub-synoptic scales.
- . Mesoscale systems are markedly non-linear, and difficult to analyze. In addition, major theoretical advances in meteorology concern planetary and synoptic scale phenomena and not mesoscale systems.

- . The demands on computing power are high. Increasing resolution in the vertical and horizontal direction by a factor of two requires approximately one order of magnitude increase in the number of calculations.

Recent advances in computer technology will have made the computation times much shorter, but in 1975 a comprehensive limited area (of about one thousand square kilometers) as an input for an NWP model, when run on a third generation CDC 7600 computer would take in the order of seven hours per simulated day in computing time.

In a more recent study, Corradini (1985) studied a hilly area in the Upper Tiber River in Italy. The model incorporated the essential physics of orographic rainfall processes on a synoptic to mesoscale level. Case studies were limited to events occurring ahead of surface warm fronts, which unfortunately excludes convective type events. The results showed that the parameterized numerical model used was acceptable for simulating orographic effects down to a mesoscale level, estimating mean aerial rainfall and also for designing raingauge networks.

On a synoptic scale, NWP models are very accurate when predicting weather patterns, and enjoy considerable accuracy on a mesoscale. But in light of the above factors, it unlikely that NWP models will be applied to microscale events in the near future.

3.2 MICROSCALE STUDIES

Several authors have studied microscale phenomena, as they are of chief concern to the hydrologist, especially for small catchments (in the order of 10km² and less). Microscale phenomena are usually treated as single convection cell events and several cells are grouped together to make up the average thunderstorm event. In the absence of techniques such as NWP models, statistically based methods and interpolation methods are applied firstly to real data, and then often used to generate synthetic data. Two examples of this are full correlation analysis and surface fitting.

In general, cells are all small scale (less than 10km in diameter), have varying shapes (spherical to elliptical to irregular), and have varying but short lifetimes (in the order of thirty minutes). Berndtsson and Niemczynowics (1986) suggest that it is important to distinguish between humid climates and semi-arid climates. In the following paragraphs, studies carried out in England, Tunisia and southern Israel will be examined, thus covering a wide range of climatic zones. It is possible that a correlation may exist that can be transported from catchment to catchment: The overriding climatic conditions may determine the prevalent nature of the microscale events.

Climatic conditions are well known for areas throughout the globe, whereas the cellular structure of storms is not known in such detail. Definition of the cell characteristics constituting a storm event would describe the storm event accurately and this may lead to a definition of storm types for study areas in general. Several authors have studied the characteristics of microscale (cellular) events. Some differing climatic zones are examined below:

3.2.1 ENGLAND

(After Shaw, 1982)

Shaw used the method of full correlation analysis to study cell parameters in two areas; Cardington in Bedfordshire and Winchcombe in the Cotswold, England. It was found that cells passing over a dense network of gauges produced large variations of intensities within small aerial distances and small time differences. Significant changes were observed sometimes over distances of less than 1km and times of less than 2 minutes.

Cells were found to be typically elliptical and had nominal dimensions of 3km x 1.5km on the major and minor axes. Lifetimes were typically from 4 to 6 minutes. The computed velocity of drift of the surface rainfall pattern was found to agree well with the mean 700mbar wind velocity for that day. The estimated direction of movement of the storm pattern as a whole was found to agree strongly with the upper air wind directions. This does not agree with work done by Marshall (1975) over the same area.

The randomness of these events is borne out by the standard deviations Shaw lists next to the parameters measured. For example; for lengths of major axes, standard deviations can be as high as about half the quoted value.

Shaw does state that the correlation technique may not be applicable if the storm center does not pass over the area studied, and this may account for some of the deviations. It appears that the nature of the storm event itself affects the correlation (i.e. if it exhibits a steady or convectional storm pattern) as well.

3.2.2 TUNISIA

(After Berndtsson and Niemczynowicz, 1986)

Berndtsson and Niemczynowicz studied a small catchment (just under 20km²) in northern Tunisia over a 2 year period. The cross-correlation technique was used to determine spatial relationships between gauges for the storms studied. Storms with the highest intensities over a five minute time interval were selected for the study. In general, depths of 0mm to 27mm were experienced, with an average intensity of 7.2mm/hr. The 10 study storms used represented 10% to 30% of the mean annual precipitation between 1982 and 1983.

High intensities were found between 5 and 15 minute time spans. Total durations were in the range of 20 to 90 minutes, but for the purposes of their study the duration was limited to 1 hour as subsequent intensities were deemed inconsequentially low by Berndtsson and Niemczynowicz. The typical cell size for this 1 hour duration was found to be about 6km² to 7km². Correlations on a monthly basis were also examined, but no cellular patterns were determined. Instead a NE-SW parallel type pattern emerged.

Monthly, daily and hourly values were cross-correlated and it was found that although monthly and daily values were adequately correlated up to distances of 2km to 3km, hourly values were not well correlated at these distances. At distances of less than 1km however, correlations of 0.8 to 0.9 were found for hourly events. Unfortunately, the study does not say whether shorter time intervals were cross-correlated or not, and what the results might have been.

They found that high intensity rainstorms were over represented when real data was compared with theoretical intensity-duration-frequency curves. Berndtsson and Niemczynowicz suggest that this may have been an unusual year, or that the theoretical curves have a bias. Either way there is a large variation, both spatially and temporally, between the theoretical curves and the observed data. This means that use of the theoretical curves may lead to inaccurate assessments for short duration events.

3.2.3 ISRAEL

(After Sharon, 1972)

Sharon studied a very arid region in southern Israel, Southern Arava. The area has mean annual precipitation figures of 30mm to 35mm, which are fairly constant throughout the region over long periods. The study area occurs within a small region which has a uniform long-term average rainfall and weather condition history.

Sharon was investigating the occurrence of 'spotty rain', a term that describes large spatial variations in rainfall where small percentages of area receive high intensities of rainfall. Three years of data from three stations, all within 15km of each other were used. Events when all three gauges had less than 3mm total rainfall were ignored. The following was found:

- . The highest value for a storm event was used as a reference value, and compared to the other gauge values. The results of this set of comparisons are as follows:

- On 70% of the days, the other 2 stations recorded less than 1/2 of the highest value.
 - On 60% of the days, the other 2 stations recorded less than 1/3 of the highest value
 - On 50% of the days, the other 2 stations recorded less than 1/7 of the highest value
- . On 4 of the 21 days of rain that were studied, all three stations received the same order of magnitude rainfall. On 11 of the 21 days, the ratio between highest and lowest depth exceeded 20:1.
- . This describes the nature of spotty rain, and it was found that 60% of the total rainfall comes from spotty rainfall events.

Two general types of rainfall were distinguished in the study; one being a spatially uniform event lasting a few hours, the other being of a localized or spotty nature associated with high intensities and small durations.

3.3 OTHER FACTORS

It has been shown that rainfall varies considerably, both spatially and temporally. The most common methods of determining how this rainfall varies is by recording stations or gauges. Unfortunately this is in itself an inaccurate procedure:

In a study of 12 gauges over a 9m x 15m grid it was found that a 95% confidence limit experienced errors of up to 13% for large storms. High ground level wind velocities are frequently associated with thunderstorm events and this can produce large errors in rain gauge readings. It has been suggested that a 20% increase in measured aerial rainfall be allowed because of this (Hall and Barclay, 1975).

In many cases, gauges are located close to areas of settlement. These tend to be in valleys or on shallow slopes, which are on elevations lower than the average elevation for the basin as a whole. This leads to calculated averages being lower than actual averages (Eagleson,

1970). Berndtsson and Niemczynowicz (1986) found that for monthly and daily values, correlation coefficients of about 0.6 were found for the influence of altitude on rainfall depth.

A linear relationship between altitude and total rainfall depth has been suggested and is of the form:

$$D_h = D_o + V.h$$

where h is the height above a datum in meters, D_o is the rainfall depth at that datum in millimeters, D_h is the new expected depth in millimeters at the height h and V is a factor that varies between 0.2 and 1.0. V has to be calibrated for the particular area. This has been proposed for long-term (e.g. yearly) values (Chebotarev, 1960).

The various time distributions of design storms to be routed over a study catchment has also been studied in detail (e.g. Lambourne and Stephenson, 1986). Generic types as the Chicago, Triangular, Moment and Bimodal hyetograph shapes have been investigated. It was found that the bimodal variation proposed in the report performed better than the others tested for use in highveld areas. A simple statistically based model describing cellular structures of storms was also utilized in the study. While this is a considerable improvement on constant intensity events, it does not address a key aspect of this report: The investigation of spatial and temporal variations with an accuracy comparable to the level of discretization employed in most overland runoff models when describing the ground surface properties.

3.4 EFFECTS ON RUNOFF

In a study by Sieker (1978) it was found that there was considerable effect when design-storms and measured storms were routed over the same catchment.

Design storms are characterized by;

- a specified duration which is related to a real catchment by (generally) the estimate of time it takes for a drop of water to travel the length of the catchment,
- a frequency function of the total rainfall and
- a constant intensity over the duration of the time estimate.

This event is superimposed over the whole catchment, taking aerial reduction factors into account where necessary. By contrast, real events have varying intensities and durations and large random components to their movement.

Sieker found that while studying a small catchment (0,54km²), 50% errors in runoff estimation were found when comparing naturally based storms and design storms for ten year recurrence intervals. The naturally based storms gave the higher runoff value.

3.5 STUDY PROPOSAL

Studies have shown that there are sharp drops in rainfall intensities at small distances from local peaks (see above). It has also been shown that using data based on so-called design storms (of constant intensities) and data based on real storms can give significantly different results for runoff hydrographs. Therefore current techniques for modelling storms do not describe the events adequately.

The author wished to study the pattern of rainfall and the resultant runoff. It is ultimately the pattern of rainfall impinging on the ground that determines runoff. This is not the same as the pattern of rainfall as it leaves the base of the clouds, as high wind speeds and local effects can divert the distribution of rain on it's path to the ground. This could not satisfactorily be studied using correlation techniques, numerical weather studies, individual hyetographs or radar tracking. This defined the direction of study for this project: The effect spacial and temporal distribution of real storms has on runoff.

A representation of how the rainfall intensity varied with time and space on the ground surface of a catchment was needed. In order to achieve this a means of fitting a mathematical surface to the measured rainfall over a study catchment was developed.

This mathematical model was used to produce contour maps of rainfall intensity for several storm events and used as an input to a computer model to produce runoff hydrographs for a study catchment.

Contour maps of storm events give a qualitative evaluation of storm shapes and movements, provided that the storm events can be divided into small time intervals with several maps produced for each storm event. General conclusions about cell size, duration and movement can be determined from examining these contour maps.

The comparison of runoff hydrographs, where the rainfall input for the computer runoff model is determined by a rainfall distribution model and other more conventional distributions of rainfall gives a quantitative evaluation of the effects of storm shape and movement, and current techniques of distributing rainfall over a catchment.

The following chapters deal with proposed methods for three-dimensional rainfall interpolation and the evaluation of the proposed methods to satisfy a series of goals outlined in the next chapter.

4 GRIDDING PROCESS

4.1 INTRODUCTION

Researchers have shown that the variation of rainfall intensity over short distances and time durations, and therefore the characteristics of rainfall resulting from storm events varies greatly (Shaw, 1982, Berndtsson and Niemczynowicz, 1986 and Sharon, 1972). This has been examined by studying several parameters (see chapter three). A computer method is developed for studying the effects of spatial variations of rainfall on runoff. This relies on measured rainfall data from rainfall recording stations.

Methods of data collection are becoming more and more sophisticated. Modern data loggers have the facility to measure rainfall depths accurately at very short time intervals (down to fractions of a second with the better loggers). Synchronization of a network of gauges is also vastly improved over mechanical methods as modern data loggers have accurate built-in clocks.

An important factor when modelling storm events on a computer is the nature of the raw data to be interpreted. This can be classified by several categories, but we will only look at the geographical classifications of data points here. Ripley (1981) lists some patterns such as; uniform random, centric systematic, stratified random and nonaligned systems. These patterns have a significant impact on the accuracy of the output of the various models.

The most fundamental distinction, though, is whether the data is regular or not. In this research project, regular implies that the data points are set out on a regular grid in the X-Y plane, and there is one z-value associated with each (x,y) point. Any other configuration is not regular and will be called irregular.

It is convenient to have the data in this regular format, especially for computer contouring programs. It is, however, seldom possible to obtain data in this form. This introduces the need for methods to convert the raw, irregular data into a convenient regular grid pattern. This is then in a suitable format for plotting contours.

The process of converting irregular data to regular data is termed gridding and this is where the focus of the programming component of this project lies. As an introduction to the difficulties associated with interpreting data, examine figures 4.1 to 4.4 below:

As can be seen, there are many possible surfaces that can be fitted to the simple example given here. It would be more accurate if a data point was available for the center of the study area. When dealing with random events such as storms, this advance knowledge is not possible and gauges are fixed in position and number.

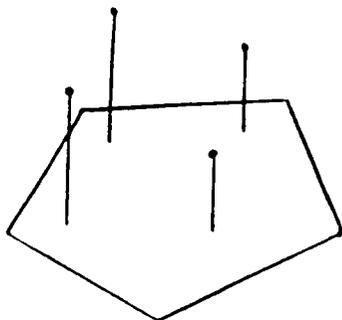


FIGURE 4.1 RAW DATA

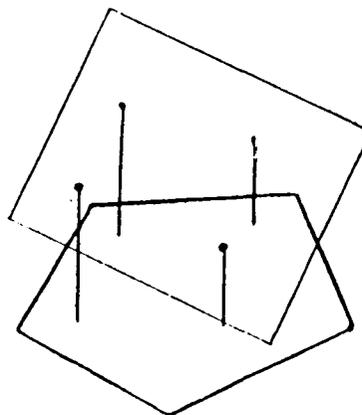


FIGURE 4.2 SURFACE ONE

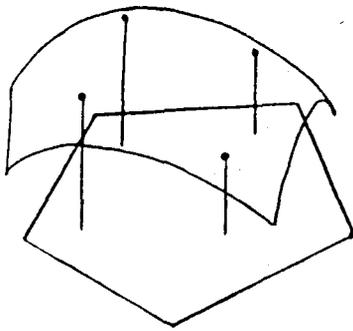


FIGURE 4.3 SURFACE TWO

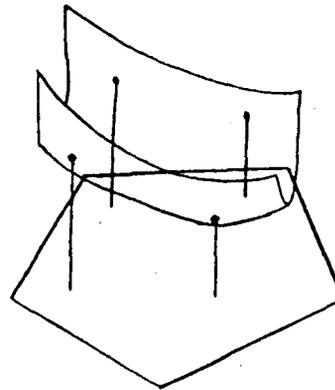


FIGURE 4.4 SURFACE THREE

In the case of rainfall measurement, there can be anything from say three to three hundred rain-gauges in the study catchment. This is mostly a function of catchment size and money available to set up recording networks. The resultant accuracy of interpolation is strongly influenced by this, and this factor is dealt with in the discussion at the end of this section.

Once the raw data has been gridded, it remains to take the grid and interpolate contours. Thus the production of contour maps of a three-dimensional surface traditionally is a two step process; gridding and contouring. For most micro-computers, memory space limitations requires that these two processes be separate for large systems of data.

There are commercial software packages that can do this, but these are not suitable to the further requirements of importing and exporting data to and from the gridding-contouring processes. These packages also appear to give time of computation less emphasis than is appropriate here.

In order to evaluate the gridding methods to be compared, a set of goals was defined. These goals are:

- 1) The interpolated storm must be the same size as the real storm.
- 2) The interpolated depth-values must equal the known depth-values where they coincide spatially.
- 3) The method must not be too sensitive to missing data.
- 4) The average depth-value over the whole study area for the interpolation storm must be the same as for the real storm.
- 5) The method must be fast.

The methods examined for the gridding process are discussed in more detail in their own chapters, but a list follows below:

- i) Polynomial surfaces
- ii) Multiquadratic fitting
- iii) Inverse squared distances
- iv) Distance weighted quadratic surfaces
- v) Krigging
- vi) Thiessen polygon method (though not a gridding process)

The first two and fourth and fifth methods all attempt to find a "trend" of one form or another, based on the available data, for the interpolation. What this means, is that the methods postulate a general form or trend that the rainfall event adheres to. It is the trend that creates the shape of the interpolation surface in the absence of data, and influences the shape in the presence of data.

This is not so noticeable when there are a large number of data points, but in sparse areas, this trend can be identified. With the smaller catchments (where there are say ten or less gauges) there are large areas which are not close to any rain-gauge. In these areas, it is often this very trend that is a cause for large inaccuracies. This is dealt with more fully later in this chapter.

These are by no means the only methods available, but on the basis of a literature survey, these appear to be the most popular. There are many variations of each of these methods and also other methods that are very similar although derived from different roots.

For some studies, the mean precipitation over a catchment is needed, and the methods also have to be compared with each other by some means for relative merits and accuracy. For these reasons, and because of its high level of performance, the Thiessen polygon weighting method was also examined.

The method of Krigging was not included in the comparative testing of methods. There were several reasons for this, the chief one being whether the method was in fact valid when applied to storms. This is dealt with in more detail in a later chapter. As the testing of various methods was not intended to be the main thrust of this project, and in view of the fact that four other methods were tested, this did not detract from the value of the overall project .

It is also intended that the method finally used for the gridding process can be used to provide a more detailed input for spacial overland flow models - a more rationally spread out grid of rainfall points can be input as opposed to fewer isolated node points.

4.2 LITERATURE SURVEY

This portion of the literature survey focussed on which method for the gridding processes was to be used. Three publications were examined initially; Heymann and Markham (1982), Maaren et al (1984) and Ripley (1981). These publications indicated that possibly five gridding methods and the Thiessen polygon method would be appropriate to the subject and therefore suitable for study. The publications also gave indications of how to write the necessary programs that were used in the study.

Of the three publications, the book by Ripley gives the most mathematically rigorous description of the theory behind some of the gridding methods. The other two give indications of the accuracy of the gridding methods, although not in great detail. The Council for Scientific and Industrial Research (CSIR) report by Heymann and Markham (1982) gives the most detailed description of the process of Krigging.

While the five methods were being researched, it was noted that many publications referred to the same "root" sources. This indicated to the author that the origins of these methods were few, while adaptations were numerous.

The five methods described in the literature survey are:

4.2.1 Polynomial surfaces

The fitting of polynomial surfaces by the method of least squares is well documented, and can be found in most numerical methods textbooks and some statistics textbooks. The most frequently discussed variations of the method concern what order of polynomial surface is to be fitted to the data. In the absence of a more detailed study on gridding methods, this method is most commonly used where a method is required to interpret three-dimensional data.

4.2.2 Inverse squared distance weighted method

The inverse squared distance weighted method is, mathematically, the most straightforward. The method is easy to program and fast to apply as it does not require the solving of any matrices, which the others do. There is some discussion as to what function of inverse distance to use (e.g. fourth power, exponentials or combinations). The method is otherwise very simple. The method is not restricted to the gridding application and is used in other disciplines such as surveying.

4.2.3 Multiquadratic method

The multiquadratic method relies on the fitting of a series of right circular quadratic cones to the study catchment (hence the term multiquadratic surface). The cones then have differing impact to the point being interpolated depending on their relative distances from the point .

4.2.4 Distance weighted quadratic surfaces

This method is a combination of the first two and fits a "roving" polynomial to the point of interest, the known data points being weighted by distance for their significance in setting up the polynomial surface. This is repeated for each point to be interpolated.

4.2.5 Krigging

The Krigging method, is a statistical method that makes use of correlations inherent in the data. It derives its name from Dr. Krige, who adapted work by Matheron and Delhomme and applied it to the mining industry. There still exists some doubt as to the validity of the method for application to storms (Heymann and Markham, 1982). This method was not tested for this reason.

4.2.6 Thiessen polygon weights

While not actually a gridding method, most literature covering this topic rate the method very highly for estimation of mean precipitation over a catchment (Heymann and Markham, 1982). Since mean precipitation is also one of the desired outputs of the gridding-contouring process, the method was also used, as a comparison for the other gridding methods.

The methods can be classified in broad categories, frequently referred to in the literature. Examples of these categories are; numerical integration (Thiessen method), smoothing techniques (polygon method and distance weighted quadratic surfaces), interpolation techniques (inverse squared distance and multiquadratic) and correlation

techniques (Krigging) (after Heymann and Markham, 1982). These classifications, while being instructive, are not germane to the issues in this project. Consequently the categories are used very loosely throughout the project and do not specify any particular method.

Once the methods to be studied were determined, it was necessary to find ways of evaluating and comparing them in order to make rational, educated decisions about their relative merits and shortcomings. An analytical approach was feasible, but the mathematics required would be excessive and difficult to extrapolate into a useful comparison of the methods. Hydrology is an applied science, and it was deemed more useful to produce some comparisons that have more immediate meaning to most hydrologists (such as sum of squared residuals, correlation coefficients, efficiency coefficients etc.).

In this respect, the methods of comparisons were adaptations of standard statistical methods that are mostly used for two-dimensional studies. These may be applied to the three-dimensional case without loss of accuracy, as in both cases they will reflect a relationship between a measured and calculated parameter. There was very little literature available on the testing of three-dimensional interpolation models.

The publications examined which discussed the gridding processes used testing techniques that suited each particular study. The author feels that for a research project such as this, more general methods were needed. It was necessary to determine a "test statistic" that would make it possible to evaluate the gridding processes. However, a fact that was to emerge from the literature survey on testing techniques was that there are numerous statistically based tests that fulfill the same function; that of determining the accuracy of fit of an interpolation method.

Without prior experience it is difficult to determine if any one test statistic is superior to another. A single "best test" was in fact impossible to find, as some test statistics behave better in some situations than in others. It became evident that any test statistic would be insensitive to reflecting some particular type of error under some condition.

A simple example of this is that the arithmetic mean of residuals might give a fair indication of how good the fit is for an average precipitation figure, but is inadequate for showing how good the fit is for discreet points, where the absolute arithmetic mean or sum of squared residuals would be more appropriate. The decision was made to use several tests statistics and extract the most meaningful at the completion of the study.

The following sections of chapter four reflect the detailed findings of the literature survey on the gridding methods and test statistic parameters.

4.3 POLYNOMIAL SURFACES

This method uses the process of least squares to fit a three-dimensional polynomial surface to the available data. The size or order of polynomial surface is a function of the number of data points. In this chapter, order means the highest power that any term in the equation is raised to. The order of equations used varies greatly, depending on the application. Sixth order polynomials are, however, commonly used.

The general form of the polynomial equation used by the test computer program for this method is:

$$z = a_0 + a_1x + a_2y + a_3x^2 + a_4xy + a_5y^2 + \dots$$

The polynomial surface equation is evaluated for the unknown coefficients a_i by the method of least squares (Chatfield 1983, pp. 179). This method minimizes the square of the distance between the observed data points and the calculated data points. To set up the set of simultaneous equations used to solve for the unknown coefficients, it is necessary to differentiate the polynomial surface equation with respect to each unknown coefficient.

Each differential equation is summed over the total number of data points, the corresponding values from the known (x,y,z) data set being substituted into the simultaneous equations. The sum associated with an unknown coefficient is placed in the column corresponding to that coefficient.

In this way, as many simultaneous equations are generated as there are unknowns, and this defines the size of the square matrix that will have to be solved for the unknown coefficients. It also defines the minimum number of data points necessary to get a unique solution to the coefficients. An unfortunate consequence of using least squares to determine the unknown coefficients, is that it often produces ill-conditioned matrices for solution. The Hewlett Packard used for the solution and processing of data worked to sixteen decimal places. This facility reduced the errors associated with ill-conditioned matrices considerably

Conventional matrix notation can be used to describe the system of simultaneous equations to be solved as:

$$Z=Q.A$$

Where A is the vector of unknown coefficients a_i , Z contains the rainfall terms and Q the coordinate terms generated from the least squares equations.

The system of simultaneous equations is solved by the method of LU-Decomposition. Each coefficient in the polynomial surface equation makes a different contribution to the surface proportional to the magnitude of the coefficient and the magnitude of the product of the x and y terms.

A primary consideration when using this method is which terms to include in the polynomial surface equation. Krumbein (1966) has shown that the orders of the respective x-y terms can be generated from a block diagram as shown in figure 4.5. If the equation is generated using this block diagram, then the size of matrix is related to the order of equation by:

$$\text{matrix size} = (\text{order}+1)(\text{order}+2)/2$$

Implicit in this approach, is that there exists an equation which accurately describes the known data. The method relies heavily on this assumption, as one equation is fitted to the whole data set. In many cases this may not be valid, and some terms included in the polynomial surface equation may not be as important as others that have been left out.

An alternative approach, then, is to evaluate terms in descending order of influence. In this context influence is proportional to the relative magnitudes of the coefficients. Miesch and Connor (1968) have shown that a stepwise regression procedure generally results in a more efficient evaluation of coefficients, as far as their influence goes. However, when they compared this to a Krumbein type equation of similar number of terms, the two approaches produced markedly differing patterns of residuals.

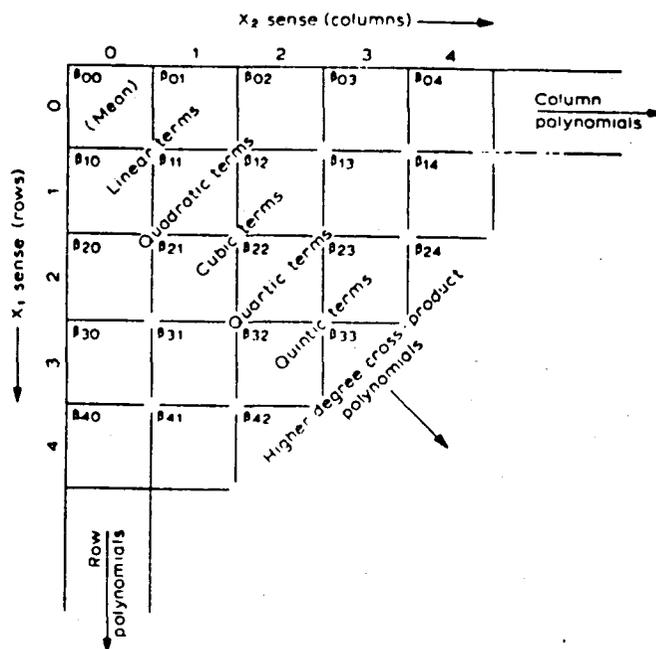


FIGURE 4.5 BLOCK DIAGRAM OF KRUMBEIN-TYPE COEFFICIENTS
(After Cliff, 1975)

The stepwise regression approach tends also to be more dependant on the nature of the data. The main drawback of this is that the terms which will have the most significant influence must be reevaluated each time the method is used, which is time-consuming if the process is repeated often. If there are no other criteria, then the choice of method would be which gives the best results, and that requires some experience in the particular nature of the data being analysed. For this study, calculation time was important, and the block diagram approach to generating equations was tested.

In either approach, it is still assumed that the initial polynomial surface equation describes the known data accurately. It is the experience of the author that this method performs very well if the data is of a known family of curves, and this family is used as the initial surface equation.

Each coefficient can be identified as producing different spacial phenomena, or three-dimensional form that contributes to the surface shape as a whole. This will be explained using a slightly different notation for the polynomial surface equation, one that suits Krumbein's block diagram better. In this notation, the general form of the equation is:

$$z = a_{00} + a_{10}x + a_{01}y + a_{20}x^2 + a_{11}xy + a_{02}y^2 + \dots$$

The subscript serves a dual purpose, firstly it shows where the coefficient fits in on Krumbein's block diagram as a (row,column) coordinate system, and, secondly, gives the exponent value for the x and y terms relating to that coefficient in the form $a_{x \text{ exp}, y \text{ exp}}$. The significance of the first six coefficients are then:

- | | |
|--------------------|---|
| a_{00} ; | Moves the surface up and down the Z-axis as a whole. |
| a_{01}, a_{10} ; | Represent two flat sheets that slope left to right and slope top to bottom, the high side depending on the sign of the coefficient. |

- a_{20}, a_{02} ; Represent two horizontal u-shaped channels (parabolic cross-section), one running from left to right, the other from right to left. The sign determines whether the parabola is inverted or not.
- a_{11} ; Represents a saddle-shaped term.

Subsequent terms of higher orders produce more complicated components to the surface. The example above shows a second order or quadratic three-dimensional surface equation. As can be seen, the magnitude of the coefficients and the (x,y) values determines the relative importance of the spatial phenomena at that (x,y) coordinate position. For such an (x,y) coordinate set, the interpolated Z-value is built up out of the collection of differing three-dimensional forms to represent the surface at that point.

Intuitively, the higher the order of equation used, and the more data points used, the more accurate the interpolated surface will be. The former is unfortunately often untrue, while the latter does have some validity, but requires clarification.

In many practical cases there is only a small number of rain-gauges, say less than ten, in the catchment. This limits the order of equation to cubic or less, and limits the potential accuracy of the fit: If the actual surface is convoluted by nature, it is unlikely that a quadratic or cubic equation will be a good fit. In such cases there would be an improvement if a higher order equation were to be used since the higher order surface could interpolate the convolutions better, but only up to a point. (This is only possible if there are enough rain-gauges to allow unique solutions to the simultaneous equations.)

With higher order equations, there is a tendency for oscillations or ripples to appear between known data points. This is even more pronounced at the edges of the interpolated area. The method then becomes unsuitable for interpolation with higher order equations.

With an infinite number of data points a surface could be found to fit the data exactly, but this is clearly impossible under the circumstances. A popular compromise is to use sixth order polynomials for the interpolation.

The least-squares method minimizes the square of the distance between known and calculated points and tries to produce a best-fit. This is done for the whole area covered, so the value that is minimized is better understood as an average difference, and this average difference is minimized. The result of this is that the average error of interpolation is often very small, but the error at discrete points is sometimes very large.

An increase in the number of data points will certainly reduce the mean error of interpolation, but may have little effect on interpolation at discrete points, as the type and order of polynomial surface equation is not changed.

An advantage of this method is that it is simple to integrate the resulting surface equation to determine a volume of rainfall for the study area. The boundaries of the area for integration do have to be definable mathematical curves, as they determine the limits in the double integral necessary to find the volume.

4.4 MULTIQUADRATIC SURFACES

Multiquadratic analysis is a term coined by Hardy (1971), and is used to describe methods where surface equations of one or two sheets are utilized to represent a known set of surface data. The method was first applied to topographical problems, but now enjoys a much wider field of application. Typical surface equations that could be used are hyperbolic paraboloids, hyperboloids, ellipsoids, spheres, paraboloids and cones of various description.

The basic concept is to fit a series of equations to the known data points, with one equation fitted to one data point. The shape of each particular equation depends on the nature of the surrounding data - hence the irregularity of the actual surface determines how flat or steep the fitted equations become.

The effects of these equations is then summed over the whole area of interest and produces the interpolation surface. It has been common practice to use equations of quadratic order, hence the name multiquadratic, but the process can equally be applied to cubic, quartic and higher order equations.

Hardy (1971) suggests that different types of equations could be suited to different types of surfaces, but also states that his research has not yet enabled him to form definite conclusions as to what type of quadratic surface best represents a type of topography. A detailed mathematical analysis of the theory of multiquadratic surfaces can be found in Lee et al (1974).

The multiquadratic method was first applied to areal rainfall estimation by Shaw and Lynn (1972). This work was used as a basis by Adamson (1978) for the analysis of a five day storm in the Crocodile River catchment and the regional evaluation of annual mean rainfall in an area just north of Nelspruit, both areas being in southern Africa. Work done by Hardy (1971) suggests that a special case of the right circular cone quadratic be used for the interpolation. Other quadrics tested tended to displace maximums and minimums.

Two relevant quadric surfaces are:

Circular hyperboloids where:

$$z = c_j[(x_j - x)^2 + (y_j - y)^2 + a]^{0.5} \dots\dots\dots(1)$$

and Circular paraboloids where:

$$z = c_j[(x_j - x)^2 + (y_j - y)^2] \dots\dots\dots(2)$$

As suggested by Hardy, Shaw and Lynn (1972) used the right circular cone which is a special case of (1) with $a = 0$. Hence equation (1) becomes:

$$z = c_j[(x_j - x)^2 + (y_j - y)^2]^{0.5} \dots\dots\dots(3)$$

What remains then is to apply equation (3) to each point where interpolation is desired, in such a way that all the known data points have an influence in the value of z . This process is repeated at every point where an interpolated value is required. The unknowns are the constants c_j , which influence the shape of the fitted cone, and which are dependant on the known data.

To determine the values of the constants c_j , equation (3) can be rewritten in a form that is more convenient for computerization where:

$$z_i = \sum_{j=1}^n c_j[(x_j - x_i)^2 + (y_j - y_i)^2]^{0.5} \dots\dots\dots(4)$$

In the above formula, n is the number of known data points, i the subscript that pertains to the known coordinate set currently being worked with, and j the subscript that pertains to all the known coordinate sets.

An algorithm for setting up this method for use can be written as:

- . Select the coordinates of the i^{th} known data set; this gives the value (x_i, y_i) .
- . For each known data point, including the i^{th} set, calculate the value of $[(x_j - x_i)^2 + (y_j - y_i)^2]^{0.5}$, with the known data points giving the values for (x_i, y_i, z_i) .

The value of c_j is then the only unknown and a system of equations can be set up which can be conveniently written in matrix notation as:

$$Z=A.C$$

where: A is the result of each $[(x_j - x_i)^2 + (y_j - y_i)^2]^{0.5}$ calculation, with i giving the row position and j the column position value.

C is the column vector of the unknowns c_j .

Z is the column vector of the known z_i values.

The solution to this system of equations is determined by solving:

$$C=A^{-1}.Z$$

For the coefficients c_j . This system of equations was solved by LU-Decomposition. Once this has been achieved, the surface can be interpolated for any coordinate set (x,y) by summing equation (3) over all the known data points. Thus equation (3) becomes:

$$z = \sum_{j=1}^n c_j [(x_j - x)^2 + (y_j - y)^2]^{0.5} \dots\dots\dots(5)$$

In this case, the point to be interpolated at gives the value of the coordinate set (x,y), and the known data points give the values of the coordinates and constants (x_j, y_j, c_j) . Thus the interpolated value z is built up from the influences of all the component cones in relation to their distance from the interpolation point.

A conceptual picture of how this method works can be described as follows:

Imagine a flat plate in the shape of the study area at an elevation of $z = 0$. Also imagine a dot on this plate at the (x,y) coordinates of each known data point (i.e. all the known data points).

At each known data point a cone has been fitted, the vertex of these cones is at $z = 0$ (by definition since $a = 0$). The steepness or shallowness of the walls depends on the magnitude of c_j . Depending on the sign of c_j , these cones may be above or below the flat plate. Hence a series of "funnels" can

be placed at each known data point, with the vertex on the flat plate. Each of the cones extends over the whole study area.

What results is a flat plate with as many cones as there are known data points attached to the plate at the (x,y) coordinates of the known data points. The cones may be above or below the flat plate, but the walls will extend over the whole study area.

To interpolate a z value at an arbitrary (x,y) point, one must first draw an imaginary line perpendicular to the flat plate at that point, and extend it both sides of the flat plate. Consistent with the method, this perpendicular line must intersect all the fitted cones. The interpolated z -value is then a weighted sum of the z -values of the intersection points with each of the fitted cones.

The weights are determined by the distance from the apex of the cone to the (x,y) coordinates of the interpolation point; with a decreasing influence for an increasing distance.

It follows from the above that the size of the matrix to be solved for the c_j coefficients is equal to the number of rain-gauges. This, in turn, has a serious effect on the accuracy of the method. For small numbers of rain-gauges an equally small number of cones can be set up for the interpolation. Thus interpolation at a needed (x,y) coordinate set results from the sum of a small number of cones, and the influence of these dependant on their distance from the interpolation point.

Another drawback of the method is this; because of the summing of the many cones for an interpolation point, it is seldom possible for the interpolated points to equal the known data points exactly. The other cones will always have an influence on this point, albeit small.

Adamson (1978) includes a routine for determining the volume under the interpolation surface. The routine is limited to areas of a rectangular shape though. This is not an insurmountable problem, as the study catchment can be broken up into a series of rectangular shapes. The main problem with this routine is the time taken for computation.

It was found that the time taken to compute the volume under the whole study area was excessive, and that the Thiessen polygon method produced as reliable results but in a faster time. The author suggests that an alternative method be used for calculating the volume under the surface.

4.5 INVERSE DISTANCE WEIGHTED SURFACE

This method produces an interpolated value by taking a weighted sum of all the known z-values. The nature of the weighting function is one of decreasing influence with increasing distance, and is normally some function of the inverse of the distance between the interpolation point and the known data points. Hence the name inverse distance weighting.

The method is the simplest mathematically and also the simplest to program, and does not require the solving of any matrices. There are no surface equations that are fitted to the data, and therefore, this method does not attempt to define a trend for the data.

If one considers a point (x,y) at which the interpolated z-value is desired and:

- . Let d_1, d_2, \dots, d_n be the distances between the point (x,y) and the n known data points.
- . Let $f(d)$ be a function of d (such as $f(d) = d^{-2}$).

Then the general form of this interpolation method can be written as:

$$z = \frac{\sum_{i=1}^n z_i \cdot f(d_i)}{\sum_{i=1}^n f(d_i)}$$

Where z is the interpolated value for the coordinates (x,y) , z_i are all the known z values, $f(d_i)$ are the distances between the interpolation points (x,y) and the known data points acted on by the function $f(d)$ and $f(d_i)$ is the sum of all these distances acted on by the same function.

It can be seen that each known z -value is multiplied by a weight and then incorporated into the summation. After all the known z -values have been summed in this manner, the total sum is divided by the separate summation of all the weights. This is the standard inverse distance technique.

In general the distance weighting function can be written as;

$$f(d) = d^{-a}$$

where a is some positive constant, commonly between one and four.

The desired result is for the value of $f(d)$ to be very large when d is small and to decay to zero when d is large. The rate of decay leaves considerable scope for experimentation with this method. To this end, there have been many different types of distance function suggested.

Ripley (1981) states that common choices for $f(d)$ are d^{-a} , e^{-ad} , and e^{-ad^2} , where e is the base of natural logarithms and a is some positive constant. In this project the first form of weighting function was used, with $a = 2$. This approach has had success in the mining industry (Heymann and Markham 1982), and because of the similarities in the problems of these two fields, was tried here. This gives rise to the specific inverse squared distance or ISD method.

There are more esoteric functions that have been suggested such as $f(d)=1/(d^2 + c)^2$, where c is some small constant (McLain 1971). These are suggested because of the practical limitations of computer precision, and the desired prevention of overflow errors that may result from dividing by very small numbers.

Because the ISD method relies solely on the decaying influence of z-value with distance, the method is very dependant on the physical size of the study catchment. Real distances are calculated and performed on by the inverse squared distance function.

Consider a real catchment. If the coordinates of the gauges are in terms of millimeters, meters, or kilometers, there is going to be a significant difference in the interpolation result. Although the relative displacement of the gauges does not change, the order of magnitude of the distances does. Since this value is to be squared, inverted, and used as a weight in the summation of a z-value estimate, this becomes a corner-stone in the accuracy of the method.

This becomes an added problem if the method is to be applied to many catchments, all of differing sizes. Intuitively one would assume that it is desirable for gauges that are close together to have similarly large effects on interpolation points, as would be the case in a small catchment. However, in small catchments maintained by Water Systems Research Group at the University of the Witwatersrand it was found that there could be large differences in recorded rainfall between rain-gauges less than one kilometre apart for the same storm event.

This indicates that these small catchments have similar interpolation problems to larger catchments, and should be treated in a similar way. If this is not done, then the effects of these differing amounts of rainfall in small catchments are not reflected during interpolation and the method becomes insensitive and therefore inaccurate.

A suggestion is to scale all catchments in such a way that the resultant catchment which the computer program works with is of a set size. This means that any catchment is mapped onto a representative catchment of fixed size best suited to the accuracy of the method. The actual size of this catchment can be determined empirically as the one that yields the best results. This will be discussed later in more detail.

It is possible for a desired interpolation point to coincide with a known data point (e.g. if a regular grid is being set up for the whole study area). In this case, a distance of zero is calculated and a computational error results when trying to divide by this distance squared. It is easy to trap for this condition by setting the ISD function = 1 if $d = 0$. In this project, this was extended to include the case of $f(d) = 1$ if $d \leq 1$ as well.

If this is not done, then on the occasions that d is less than one, but not equal to zero, a weight much larger than one can result, and this may seriously effect the accuracy of interpolation at that point. The broader significance of including this condition is that the ISD function can never be greater than one. This means that no interpolated z -value can ever be greater than the maximum known z -value. Hence this method cannot produce z -values higher than any of the known data points, a characteristic that is not shared by any of the other methods tested. This not true for the reverse case of interpolated values less than the minimum known data point. The condition does occur, since the influence of any rain-gauge decreases with distance.

This could be extended to the extreme where the study area is surrounded by a gradually flattening surface with an increase in distance from the study area, until at an infinite distance, the interpolation surface is completely flat with $z = 0$.

The disadvantage of this method is that it does not lend itself to integration of a convenient function to calculate the volume under the interpolation surface. A method such a Thiessen polygons would have to be employed to calculate a mean rainfall estimate for the study area. An alternative to this is to fit a series of splines to the now regular grid over the study area and integrate this over the relevant area. Simpson's approximation can also be applied, with faster results.

4.6 DISTANCE WEIGHTED QUADRATIC SURFACES

This method can be considered as a combination of the first and third methods discussed i.e. polynomial surfaces and inverse distance weighted method. The method also sets up a polynomial surface equation, but gives known data points that are closer to the desired interpolation point more influence than those that are further away.

The main difference between this method and the polynomial surface method is that it sets up a new polynomial surface for every desired interpolation point. The method does also use the least-squares approach to determine the unknown coefficients for the polynomial surface equations.

As before, there is flexibility in the order of the polynomial surface equation and which terms to include or exclude. If there are many desired points of interpolation (in this project the test programs used 100), then the methods can be time consuming. This is aggravated with this method when large order polynomial surfaces are used each time. To alleviate this, it was decided to limit the polynomial surface equation to a quadratic.

The process then becomes:

- . Choose the desired interpolation point.
- . Work out a series of weights inversely proportional to the distance between the interpolation point and all the known data points.
- . Use the least-squares method to set up and solve a quadratic surface equation, where the component terms are premultiplied by the relevant weight.
- . Back-substitute the coordinates of the interpolation point into this now defined quadratic surface and calculate the interpolation z-value directly.

This whole procedure has to be repeated for each new interpolation point.

The following mathematical description of the method is vastly simplified if preceding work is utilized. The discussion in chapter 4.3 on polynomial surfaces for which terms to be used in generating the polynomial surface equation applies equally to this method. Hence in this project, a Krumbein-type quadratic surface equation was used. The discussion in chapter 4.5 on the inverse distance weighted method for which function of distance to use similarly applies here. Hence in this project, an inverse distance function of $f(d) = d^{-2}$ was used.

The method is described by McLain (1974) as follows:

The Krumbein-type quadratic surface equation can be written as:

$$z = a_{00} + a_{10}x + a_{01}y + a_{20}x^2 + a_{11}xy + a_{02}y^2 \quad \dots(1)$$

From this, the least squares equation can be postulated as:

$$Q = \sum_{i=1}^n (z_i - z_{ki})^2 \quad \dots(2)$$

Where z_{ki} are the known z-values and z_i are the interpolated z-values as defined by equation (1) (and adding an i subscript to each z, x, and y term). The standard procedure is to differentiate equation (2) with respect to each unknown coefficient and use the known data to fill the resulting square matrix (in this case a six by six matrix) built up from the differentiation and then to solve for the unknowns a_{im} . With this method one first has to postulate a new type of least-squares equation where:

$$Q = \sum_{i=1}^n (z_{ki} - z_i)^2 \cdot f(d) \cdot ((x_{ki} - x_i)^2 + (y_{ki} - y_i)^2) \quad \dots(3)$$

Where $f(d)$ is the inverse distance function, (x_i, y_i, z_i) are the coordinates and z-values of the interpolation point and (x_{ki}, y_{ki}, z_{ki}) are the coordinates and z-values of all the known data points.

As before, z_i is defined by equation (1) with an i subscript added to the z, x, and y terms. Equation (3) is differentiated with respect to each coefficient a_{im} in turn, and the resulting six equations are

each set equal to zero, and solved for the z_i terms. The six equations are summed over all the values of i , and this setup forms the basis of the least-squares matrix to be solved. Written in conventional matrix notation this becomes:

$$Z = B.A$$

Where Z contains all the terms with a z_i in, B contains all the terms with $x_i y_i$ terms and A contains the unknowns a_{im} . This has solution:

$$A = B^{-1}.Z$$

This was solved by LU-Decomposition. The method also suffers from the problem of ill-conditioned matrices, as did the polynomial surface method. However, since the polynomials have been limited to quadratics, the magnitudes of the elements of the solution matrix never reach the extremes that a solution matrix for a sixth order polynomial equation would. The ill-conditioning is introduced by the weighting factors that premultiply the component terms. Many of these weights are going to be very small numbers, and this produces the differences in magnitudes of the solution matrix, and therefore the ill-conditioning.

This method produces as many polynomial surfaces as there are desired interpolation points. It would be possible to integrate each of these over their respective areas of influence and obtain a total volume figure for the whole catchment. The determination of the areas of influence may be difficult to determine and for many interpolation points the process would be time consuming. It is suggested that an alternative method for calculating total volumes be employed.

4.7 KRIGGING

This method differs from the others discussed because of its strong statistical basis. The name krigging comes from Dr. Krige, who adapted work done by Delhomme and Matheron on spatial correlation. The method was first applied to the field of mineral resource estimation, and krigging can be considered to cover both best linear unbiased estimators of a point and the best linear weighted moving average of a block. The former considers estimates at a point while the latter considers estimates within a predefined block or area.

Point krigging, as it is sometimes referred to, was first introduced to the hydrological world by Gambolati and Volpi (1979) in a study of groundwater contour mapping in Venice (Heymann and Markham, 1982).

The method is a stochastic one. A stochastic process is a physical process whose structure involves a random mechanism (Chatfield, 1983 pp. 349). It uses the correlation existing between any two points in the area of a storm (Heymann and Markham, 1982). The general problem can be stated as:

Consider a storm event S having true unknown characteristics $Z(s)$, and a series of m samples of known magnitude $Z(s_i)$, with $i = 1, 2, \dots, m$ (i.e. rain-gauge data).

It is desirable to find a series of weights a_i such that

$$S^* = \sum_{i=1}^n a_i Z(s_i)$$

is the best estimator of $Z(S)$. This is done by computing the variance of the error of the method (after David 1977). The pursuit of this estimation of variance of $Z(S)$ by Z^* then becomes the problem of finding what is termed a variogram. The determination of the nature of this variogram forms the basis of the accuracy and applicability of the method.

It is convenient to conceptualize the variogram as a statistical description of the influence any known data point has over the desired estimation point. There are many different types of variograms such as

the spherical model, the De Wijsian model, the linear model and many exponential models, to name a few (David, 1977). These different variogram models would have a critical effect on the continuing accuracy of this method when applied to storm events.

This method was not used as a test method in this project. There were several reasons for this:

- . There was no direction given in the literature as to which variogram would be appropriate for the storms to be studied. It is possible that with extensive testing, a most appropriate variogram could be found, but it was felt that this exceeded the bounds of this project.
- . Ideally, large quantities of data would be used to empirically determine the most appropriate variogram specific to this study. However, this kind of information would be available as a consequence of this very project, and could not, therefore, be used in it's production.
- . The technique is based on the assumption that the rainfall events measured are stationary. This implies that the correlation pattern may show a direction, but that this pattern must be the same at all points in the storm. It is the condition of stationarity that may not be satisfied in convection storms (Heymann and Markham). As mentioned above, this could be verified or not as separate result of this study, but it was felt that this exceeded the bounds for this report.

In view of the fact that four other methods were going to be tested, it was also decided that leaving out this method would not have a detrimental effect on the final outcome of the project.

For a detailed discussion of this method, refer to the book by David (1977).

4.8 THIESSEN WEIGHTS

This section presents a brief discussion on the computation of the Thiessen weight or polygon method. The method is very simple in concept, and assigns a weight to each rain-gauge associated with the study area for a particular event, the sum of all the weights being one. Values obtained from each rain-gauge are multiplied by this weight and then summed to give an average value for the whole study area.

The value of the weight is determined by dividing the study area into a series of polygons. The dimensions of the polygons are defined by the condition that the line defining the border of two adjacent polygons must be equidistant to the two nearest rain-gauges. This divides the study area into a set of polygons, whose areas can be determined. These individual areas are each divided by the total area to give the weight of that rain-gauge. The study area boundary forms the outside edge for the external polygons.

Heymann and Markham (1982) found that overall, this method provided the most accurate estimates of rainfall volumes for their tests. The method also appeared to be robust, as tests were done with gauges left out, and the method performed well under these conditions.

The programming of the method is based on two papers by Diskin (1969 and 1970) the second being a revision of the concept introduced in the first paper. In the first paper he suggests a Monte Carlo procedure for estimation of the weights: Random points are generated within the study area, and a counter is incremented for whichever gauge is nearest to the random point. This is repeated for many such random points until a reasonable spread of points, and therefore, a reasonable accuracy has been obtained.

The second paper does away with the necessity of random points, and suggests the superposition of a regular grid of points over the study area. This is in fact the best outcome of the Monte Carlo procedure, and eliminates large amounts of computing time.

As before, each point within the study area is considered, and a counter incremented for the nearest gauge. Each point is considered once, and it becomes convenient to work in a series of horizontal rows. Accuracy is determined by the number of points used for the grid. This number is dependent on the number of gauges in the study area, but obviously has a minimum value.

More sophisticated approaches for defining polygons exist, such as the edge defining algorithm proposed by Croley and Hartmann, and the modifications to the Monte Carlo method suggested by Shih and Hamrick. For this project, the method as described by Diskin's second paper was more than adequate. Any Thiessen weights calculated for this project were done using the program written on this basis.

In a later chapter, the testing process is described and it is mentioned that the test data sets were composed of one hundred data points in a regular grid pattern. This provided a very fast method of working out a mean depth-value for the study area, if the Thiessen method is employed: Because the data is spaced in a regular grid pattern, the area of each polygon reduces to a square whose dimensions are the same as those of the grid point spacing.

This is further simplified if one examines the general form of the Thiessen method:

$$Z_{\text{mean}} = \frac{\sum_{i=1}^n (w_i \cdot z_i)}{\sum_{i=1}^n w_i}$$

Where w_i is the weight (calculated as the fraction of the associated area of the polygon over the total area) and z_i the depth-value associated with the rain-gauge i .

If the area associated with an internal data point is called dA then the weight associated with this depth-value becomes $dA/(\text{total area})$. The total area value can also be written in terms of dA and is of the form:

total area = (no. of y rows - 1).(no. of x rows - 1).dA

The weight for the internal depth-value then becomes:

weight value = $1/((\text{no. of y rows} - 1).(\text{no. of x rows} - 1))$

Which is a simple number to calculate, is constant for all internal depth-values and is independent of the actual areas involved. It can be shown that edge point weights are half this value and corner point weights one quarter of this value by applying the Thiessen polygon method. The calculation of a mean depth-value then becomes the process of summing all internal depth values and multiplying this by the weight value, summing all the edge depth-values and multiplying this by half the weight value, summing the four corner depth-values and multiplying by a quarter of the depth-value and summing these three results.

This is faster than any other method suggested so far but only has application if the depth-values are in a regular grid formation.

In many studies there are several rain-gauges within or near the study area that can be used to contribute towards rainfall data. It is often necessary to average the data collected from such rain-gauges in some way before the recorded data can be used as an input to a rainfall-runoff process model. Thiessen weights provide a method of doing this:

The records from each rain-gauge are divided into synchronised time steps. The weighted average for each time step is calculated by multiplying the rainfall intensity (or depth) for each time step by the appropriate Thiessen weight and summing the values from a time step together.

This averaged value for a time step is applied to the study area as a whole, and is sometimes further modified by an aerial reduction factor for larger catchment areas.

This approach is tested in chapter 6.

4.9 TESTING PROCESS

With four alternative methods that could perform the same task, that of gridding, it became necessary to devise a process to evaluate each method and determine the most suitable. The remainder of this project relies heavily on this section of programming, and it was decided to test these methods as thoroughly as possible for this reason.

Two alternatives were available; a rigorous mathematical analysis of the performance of the methods or a series of tests giving results in some meaningful variables (e.g. correlation coefficient). It was decided that the latter would be more useful in this case, as it would be more immediately obvious how the methods compared if this type of testing was undertaken. It was also desirable to get a "feel" for the different methods and how they behaved.

Adopting a testing by comparison process requires two things; firstly there must be a raw data set to be used as an input, and secondly the true values for a full description of the whole system to be predicted must be known (in this context, a full description of the whole system means knowing the depth-values over the whole study area for the same points that the gridding process is going to interpolate at). The process then becomes; use the raw data to predict a description of the system and compare this with the known description of the system by some means. This must be repeated with different data sets to achieve a reasonable degree of confidence in the results.

Relating this to a real catchment introduces the fundamental problem of this type of testing: It is possible to collect a reliable series of records for several storm events, but almost impossible to have the corresponding full description of the whole system. It becomes necessary then to either work from an existing contour map and extract the relevant raw data and known grid points, or to generate artificial data.

In other studies of this nature, the method adopted was to work with known storm events and to leave out various combinations of gauges, and to determine the accuracy of fit to the now unknown gauge. It is the opinion of the author that this is more a test of the robustness of the method than the inherent accuracy or lack thereof. While this approach was used, a more realistic means of testing was also defined and used in the early stages of this study.

The option of working from contour maps of storm events was examined but the process of extracting grid points and reasonable values to simulate gauges was deemed impractical as many such data sets would be needed and this method would be time consuming. An alternative approach based on generated data was used.

Several reasons made this approach more attractive as a testing process:

- . The four methods were computer based and it would be convenient to generate data directly as a computer-readable file.
- . To make the testing as general as possible, it would be necessary to test different shaped catchments (and different gauge locations) as well as different storm events for the same catchments. It would more time efficient if these could be generated by computer and processed automatically.
- . There is a wide fluctuation in the number of gauges per catchment in real studies. In order not to prejudice the testing process it would be necessary to work with many different numbers of gauges per catchment. This flexibility was only allowed with generated data (in a reasonable time-span).
- . The methods of comparison are statistically based techniques and require large amounts of computation. This could easily be done at the same time as the gridding process, if the full description of the system was known.

In the end, a compromise between time taken for generating data and thorough testing for accuracy determined the final process. Three raw data sets were generated - each comprised of a regular ten by ten grid

of one hundred depth-values. The x and y difference between grid points for each data set was changed, thus creating three different shape areas. The three sets were named test1a, test2a and test3a, and are referred to by these names in the testing process. It was convenient to work with up to one hundred gridded points as this could be converted into a percentage figure to represent a required density of rain-gauges for a required accuracy of fit.

Test1a was derived from a generated contour map simulating a storm cell on the edge of a catchment, where grid points were extracted from the map by interpolating a depth-value according to the contour heights nearby. Test2a was generated from a pre-defined three-dimensional quadratic surface given by:

$$z = 15 + 7x - 3y + 9x^2 + 2y^2$$

The various z-values for the corresponding grid points were calculated by substituting in the coordinates of the grid points and finding the z-value. This equation is easy to integrate over an area for a total volume of precipitation figure, and this could be useful later on. On the basis of the literature survey the common use of the polynomial method suggested that storm patterns may follow a polynomial trend. For this reason a polynomial type storm pattern was included to try and simulate a real storm pattern.

Test3a was generated from a more complex function given by:

$$z = 6.SIN(x) - COS(3y) + LOG(x) - 1.002y$$

The idea was to create an undulating surface with a disguised trend to determine how well the methods could interpret hills and troughs in the data. This function could also be easily integrated for a volumetric figure.

To satisfy the conditions set out above, the following process of selecting data was used:

The full data set (i.e. test1a, test2a or test3a) was read in by the gridding process currently being tested.

- . A varying but controlled number of (x,y,z) coordinate sets were selected at random from the full data set, thus simulating the known rain-gauge depth amounts from a storm event. For this portion of the testing between three and one hundred points were selected at fixed intervals.
- . The gridding process was applied to this new data set and the values for the whole area generated.
- . The generated values could then be compared with the original test data set read in (i.e. test1a, test2a or test3a) and the accuracy of the fit determined.
- . As mentioned above, a varying but controlled number of coordinate sets were selected randomly from the test data set. This was controlled by the subset $n \in \{3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40, 50, 60, 80, 100\}$, where n is the number of points selected at any time.
- . Over and above this, when a particular number of coordinate sets was being worked with, the process of randomly selecting simulated gauges and gridding was repeated twenty times, each time with a new set of randomly selected points. Each time the generated grid points were compared with the known test set.

The varying number of points selected (from three to one hundred) simulated the varying number of gauges found in real situations. The randomly selected coordinate sets for each number of points simulated the different catchment shapes by the varying (x,y) coordinates of the coordinate set. The fact that each of the twenty cases were interpolating to the same simulated storm event does not detract from the value of the tests. When ninety points were being used for the interpolation, it was considered that this was equivalent to the robustness tests mentioned above. In this way maximum use was made of the three data sets.

One would expect that as the number of coordinate sets approached one hundred the accuracy of fit would also increase - this was true in general. The final case (when the full set of one hundred data points

are used to grid the study area) would give an indication of how well the method is able to interpolate the known data exactly. It was hoped that all the methods would be able to interpolate exactly at all known data points, but this was not the case, as will be shown later.

Having set up the framework for testing the methods it became necessary to decide what variables would be used for the comparisons. To this end the paper by Green and Stephenson (1986) was used, as it gives a detailed summary of criterion for comparison of single event models.

It became evident that there was no single test statistic that gave a consistently reliable indication of how well the method was performing. Some test statistics behaved well under certain conditions but badly in others. In view of the extensive amount of testing that was to be carried out it was deemed prudent to use several test statistics at every test and later determine the most meaningful, rather than leave something out and have to re-test the whole lot again at considerable waste of time.

On the basis of recommendations made in the article mentioned above, and the author's own experience, the following test statistics were used:

- . Arithmetic mean of errors.
- . Absolute arithmetic mean of errors.
- . Sum of squared residuals.
- . Root mean square error.
- . Coefficient of variation.
- . Mean of known depth-values.
- . Mean of interpolated depth-values.
- . Standard deviation of errors.
- . Efficiency.
- . Correlation coefficient.

Every time the data was gridded, the interpolated values were compared with the known values and the above ten test statistics calculated. This was done twenty times for each number of coordinate sets, and so there would be twenty values for a test statistic for a set number of

simulated gauges and test data set. The maximum, minimum and mean of the twenty values were also calculated where appropriate. It is arguable whether the errors followed a normal distribution, and so the standard deviation of errors was not used to compare methods directly, but more to give a "feel" to the variations in accuracy.

Several other aspects of the methods were also tested: The sensitivity of the methods to real distances between gauges was examined, the shape of the gridded surface was compared to the actual shape, the percentage of the absolute arithmetic mean of errors over the mean of the known depth-values was calculated and the time taken to complete the gridding process are some examples. This last value became important in selecting which method would be used in the remainder of the project.

The output of the testing process was printed and formed the basis of the comparative graphs and discussion in the concluding chapter for the gridding methods.

4.10 DISCUSSION ON GRIDDING METHODS

This chapter is based on the findings of the author while completing this project. The method of up setting test data sets is described in the previous chapter, and this chapter includes the discussion of the results and implications of the testing process.

The gridding method that was used for the remainder of the project was selected based on how closely it satisfied the goals outlined in chapter 4.1. It is convenient at this stage to repeat these goals:

- 1) The interpolated storm must be the same size and shape as the real storm.
- 2) The interpolated depth-values must equal the known depth-values where they coincide spatially.
- 3) The method must not be too sensitive to missing data.
- 4) The average depth-value over the whole study area for the interpolation storm must be the same as for the real storm.

- 5) The method must be fast.

Figures 4.6 to 4.8 show the one hundred depth-values for each test data set. They also show the contour shapes that were drawn from the test data sets, using a contouring program developed by the author. The numerical values of the three test data sets were used for a quantitative comparison, as has been discussed in the previous chapter. The three contour maps were used for a qualitative comparison of the four methods tested.

The four methods tested were inverse squared distances (ISD), multiquadratic method (MQUAD), polynomial surface fitting (POL) and distance weighted least-squares quadratic surface fitting (DIST). For convenience, the methods will be referred to by their abbreviations shown in parentheses.

To give an indication of how well the methods satisfy goals two to four, which demand various degrees of numerical accuracy, the sum of squared residuals for all the tests has been reproduced in figures 4.9) to 4.11). Each figure corresponds to one test data set, the four diagrams (a, b, c & d) correspond to each method's behavior while gridding that data set. The horizontal axis, labelled number of data points, represents the number of rain-gauges in a catchment

Several general conclusions can be determined from these diagrams:

- . The most obvious indication is how badly all methods behave with small numbers of rain-gauges.
- . The distance between the maximum and minimum graphs indicate a kind of deviation, or consistency in the method - the wider the band the less confidence there is in the output of the method.
- . The degree of oscillations in the early portions of many of the graphs is more an indication that twenty different tests were not adequate per number of data points, than any conclusive statement about accuracy. What is reasonable to say though, is that for less than ten gauges one can expect widely fluctuating degrees of accuracy.

9.6	9.6	9.7	10	11	11	10	9.3	9.5	9.3
9.2	9.1	9.1	9.4	9.9	10	9.6	8.5	7.8	7.8
8.8	8.7	8.5	8.5	8.8	8.9	8.4	7.7	7.1	7
8.4	8.4	8	7.7	7.6	7.5	7.2	6.8	6.6	6.6
8	8.1	7.5	6.8	6.4	6.1	6	6	6.1	6.3
7.3	7.3	6.8	5.9	5.2	4.7	4.7	5.1	5.5	5.9
6.3	6.2	5.8	4.9	4	3.3	3.5	4.2	4.9	5.5
5.1	5	4.8	4	2.9	2.1	2.4	3.4	4.3	5.4
4.1	4	4	3.4	2.3	1.4	1.9	2.9	4	4.8
3.5	3.6	3.6	3.1	2.2	1.6	2	2.8	3.8	4.6

FIGURE 4.6a DATA SET 1: - TEST1A
 PLOT OF DEPTH VALUES DIRECTLY FROM THE DATA SET
 (EDGE OF CATCHMENT DEFINED BY DOTTED LINE)

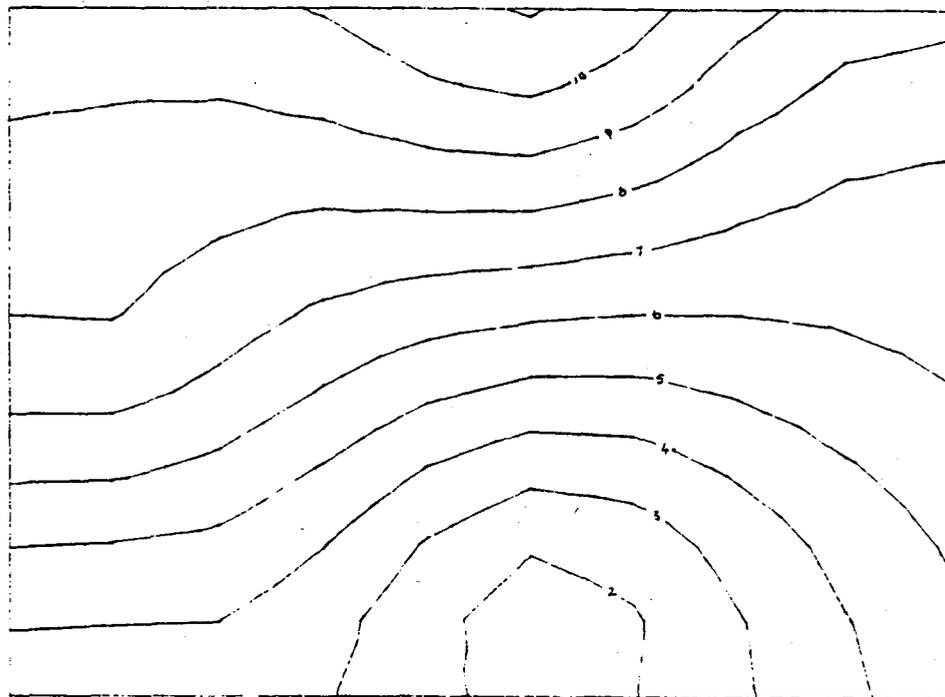


FIGURE 4.6b CONTOURS FOR DATA SET: - TEST1A

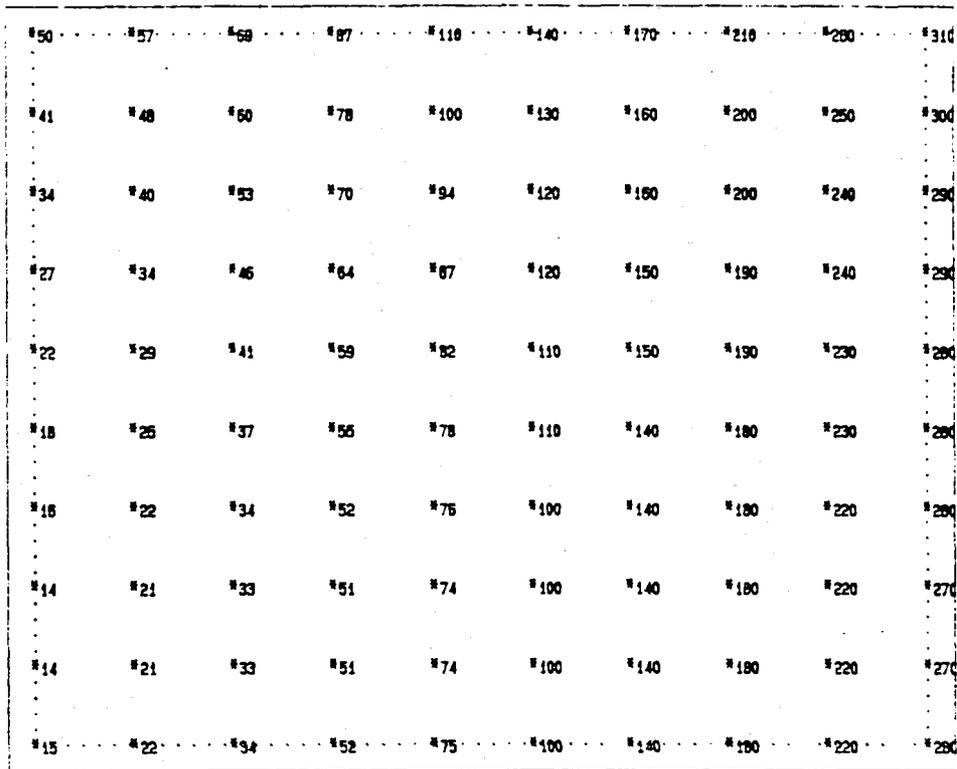


FIGURE 4.7a DATA SET 2: - TEST2A
 PLOT OF DEPTH VALUES DIRECTLY FROM THE DATA SET
 (EDGE OF CATCHMENT DEFINED BY DOTTED LINE)

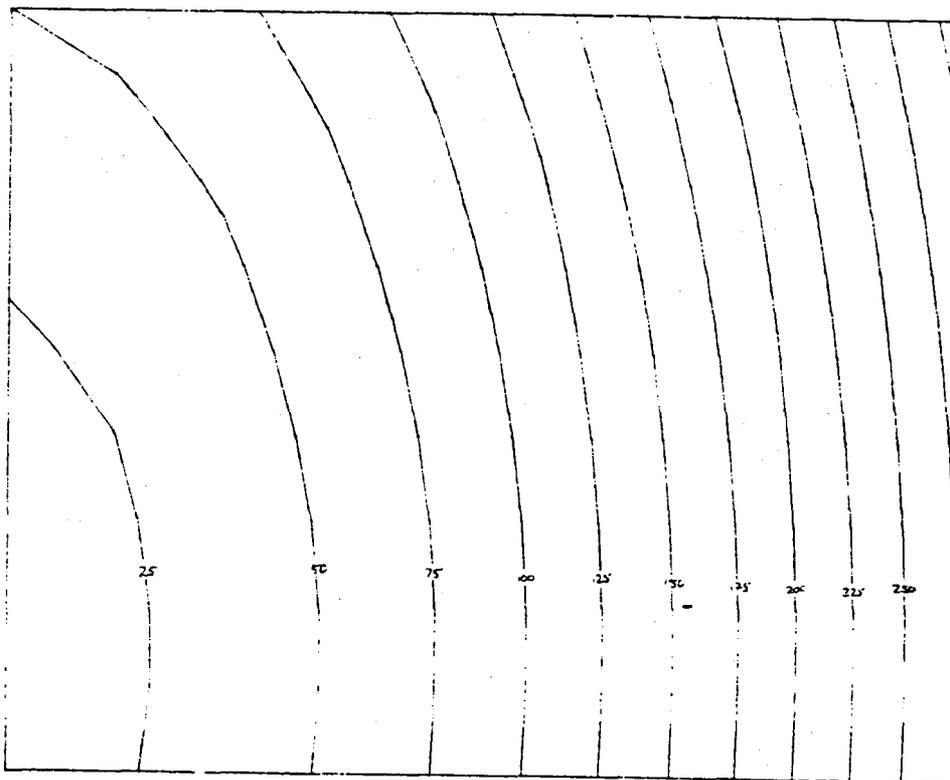


FIGURE 4.7b CONTOURS FOR DATA SET: - TEST2A

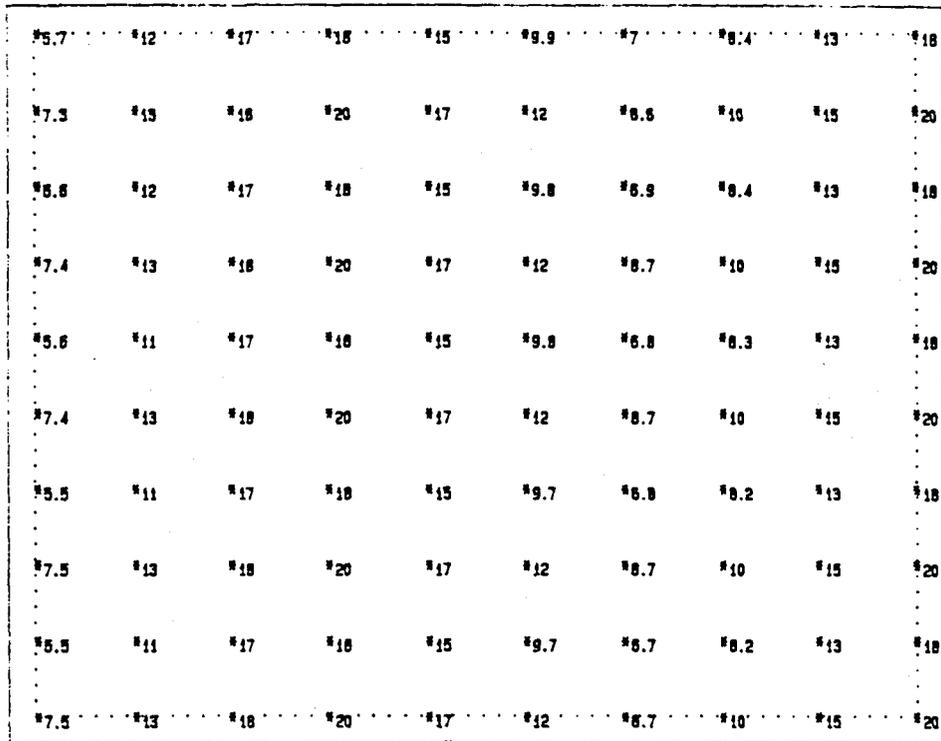


FIGURE 4.8a DATA SET 3: - TEST3A
 PLOT OF DEPTH VALUES DIRECTLY FROM THE DATA SET
 (EDGE OF CATCHMENT DEFINED BY DOTTED LINE)

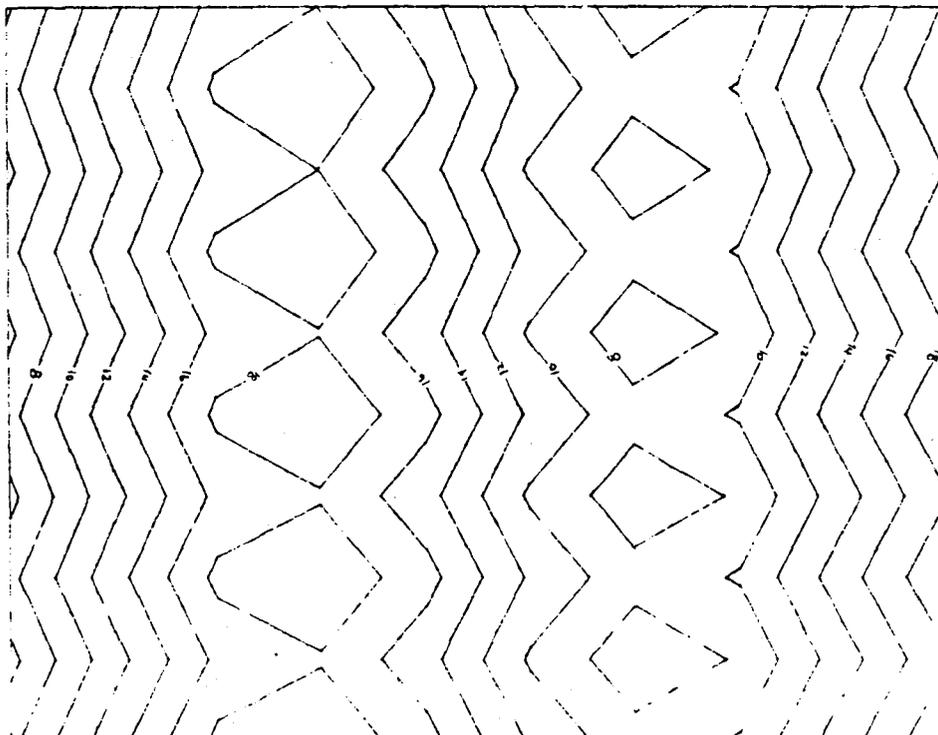


FIGURE 4.8b CONTOURS FOR DATA SET: - TEST3A

- . Many of the tests show that the methods behave badly even when all one hundred depth-values were used. It appears that this is a characteristic of the method, and not the data set.

In an attempt to evaluate goal number four, that of average estimates being the same for interpolated storms and real storms, two statistics are presented. The ratio of the absolute arithmetic mean of errors over the mean of depth values for the real storm are presented as a percentage for all the methods and test cases (figure 4.12a to figure 4.12c).

The choice of the ratio of absolute arithmetic mean of errors (AAME) over the mean of all depth values was prompted by a need to conceptualize the relative magnitude of the error of interpolation. The AAME value on it's own gives an indication of the magnitude of the average fit for the whole area, but this can be misleading in the case of two extremes when the actual depth-values are large in magnitude (121.8 for test2a) or small in magnitude (6.2 for test1a). The AAME value was then compared with the mean depth-value to give an indication of the scale of the error. At it's best, this value should be as small as possible as soon as possible, thus showing the magnitude of error is small in comparison with the values being measured.

In addition, figure 4.13 shows the arithmetic mean of error values for all three test sets for one hundred data points. This is intended to show how the methods perform when all the data is available for interpolation, and shows what is likely to be the best performance possible for the methods. This figure was instructive in showing which methods were undesirable for the latter part of this project, working on the rationale that if the method does not behave well with all one hundred data points present, it wouldn't behave as well as others will with less than one hundred data points available for interpolation.

The ISD method behaves consistently throughout. Although the percentage error starts at about 45% each time, it is never higher than this. The other methods sometimes fluctuate wildly in the areas of low numbers of

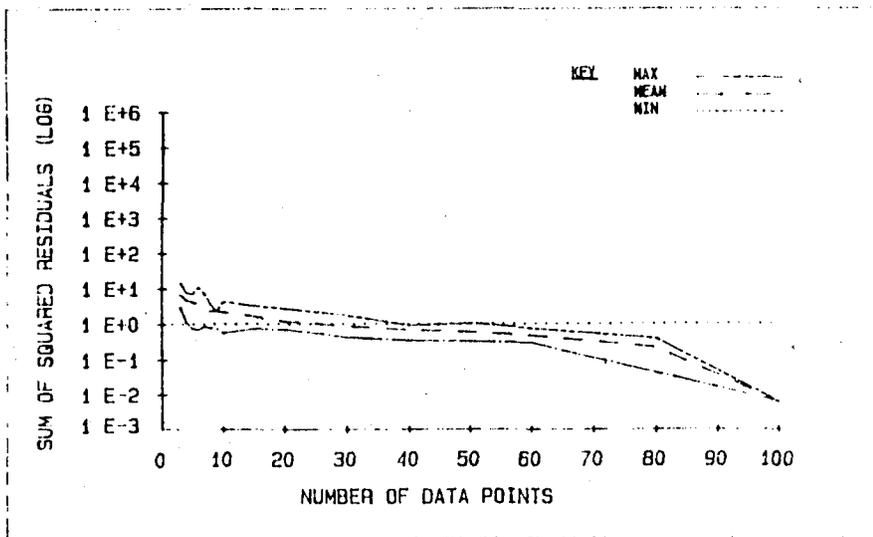


FIGURE 4.9a INVERSE SQUARED DISTANCE (ISD) APPLIED TO TEST1A
SUM OF SQUARED RESIDUALS

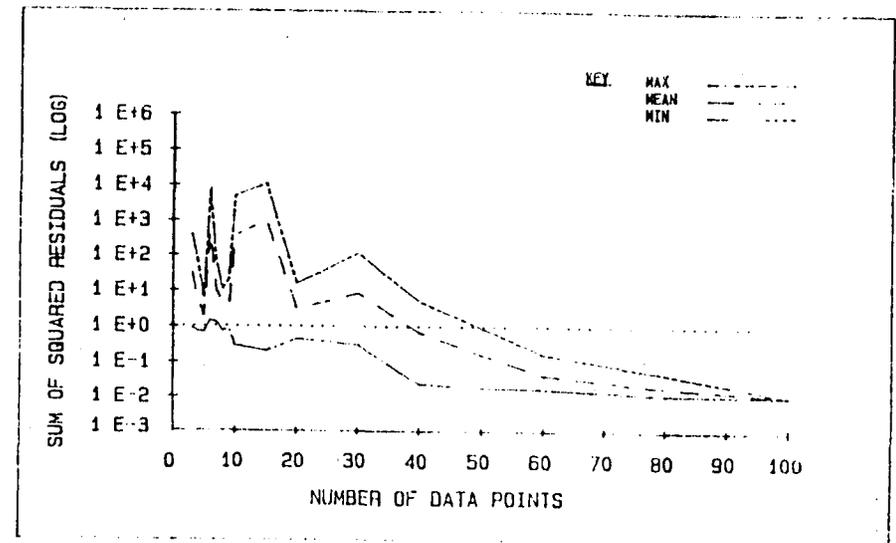


FIGURE 4.9b POLYNOMIAL (POL) APPLIED TO TEST1A
SUM OF SQUARED RESIDUALS

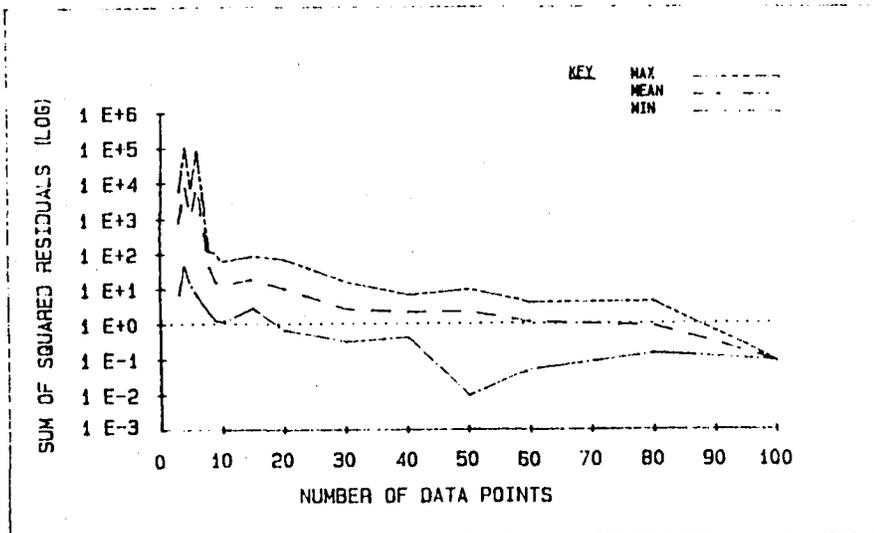


FIGURE 4.9c MULTIQUADRATIC (MQUAD) APPLIED TO TEST1A
SUM OF SQUARED RESIDUALS

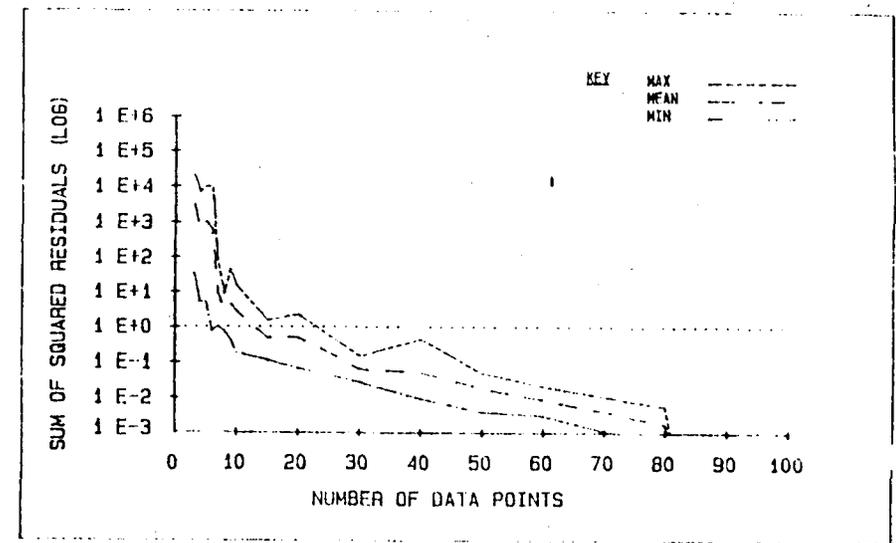


FIGURE 4.9d DISTANCE WEIGHTED (DIST) APPLIED TO TEST1A
SUM OF SQUARED RESIDUALS

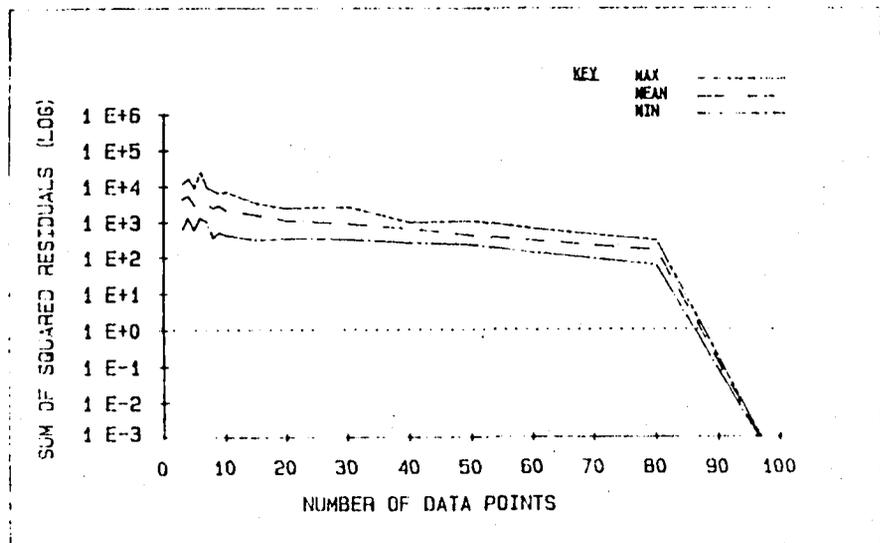


FIGURE 4.10a INVERSE SQUARED DISTANCE (ISD) APPLIED TO TEST2A
SUM OF SQUARED RESIDUALS

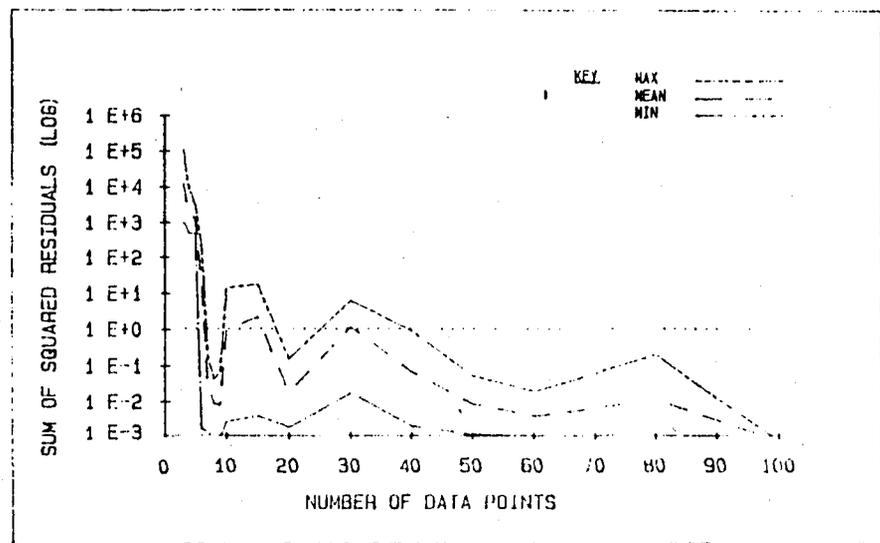


FIGURE 4.10b POLYNOMIAL (POL) APPLIED TO TEST2A
SUM OF SQUARED RESIDUALS

62

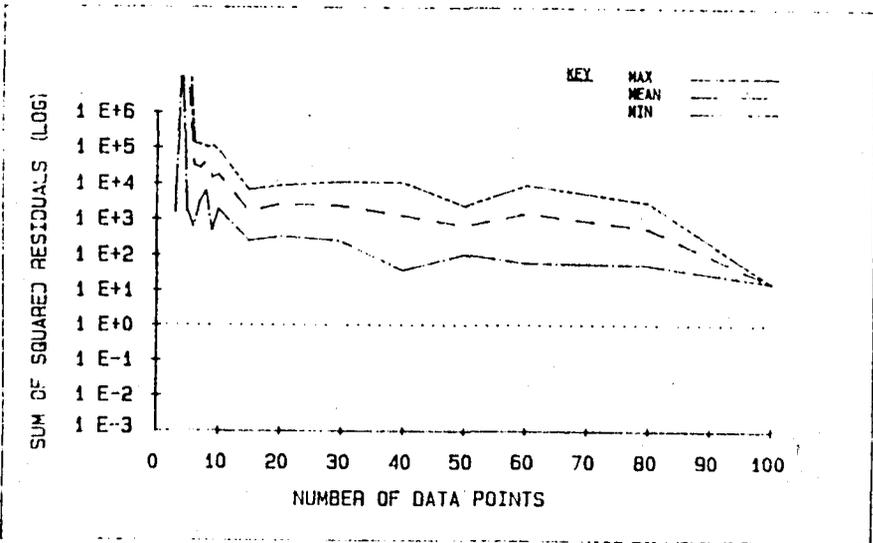


FIGURE 4.10c MULTIQUADRATIC (MQUAD) APPLIED TO TEST2A
SUM OF SQUARED RESIDUALS

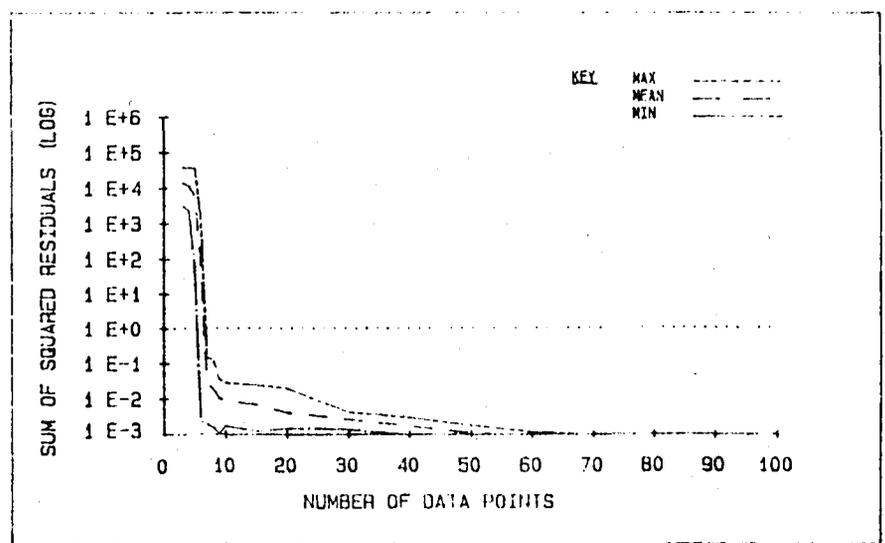


FIGURE 4.10d DISTANCE WEIGHTED (DIST) APPLIED TO TEST2A
SUM OF SQUARED RESIDUALS

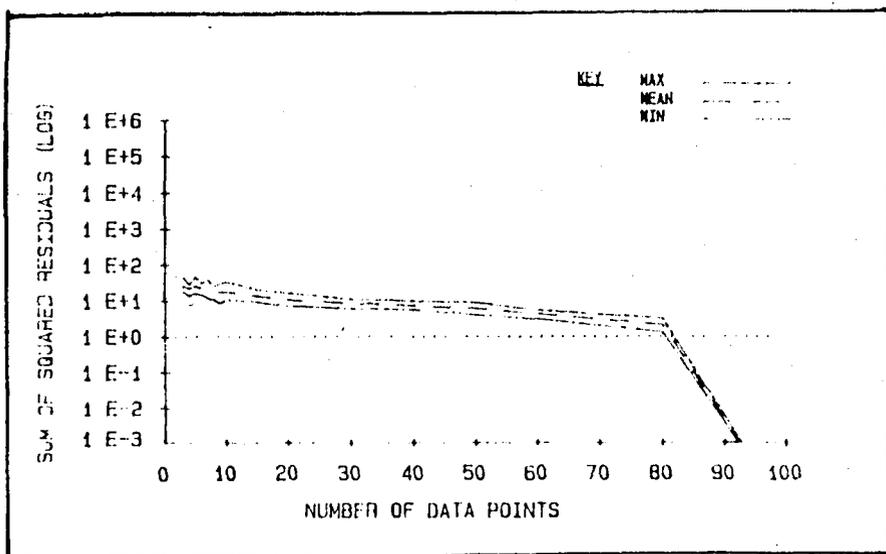


FIGURE 4.11a INVERSE SQUARED DISTANCE (ISD) APPLIED TO TEST3A
SUM OF SQUARED RESIDUALS

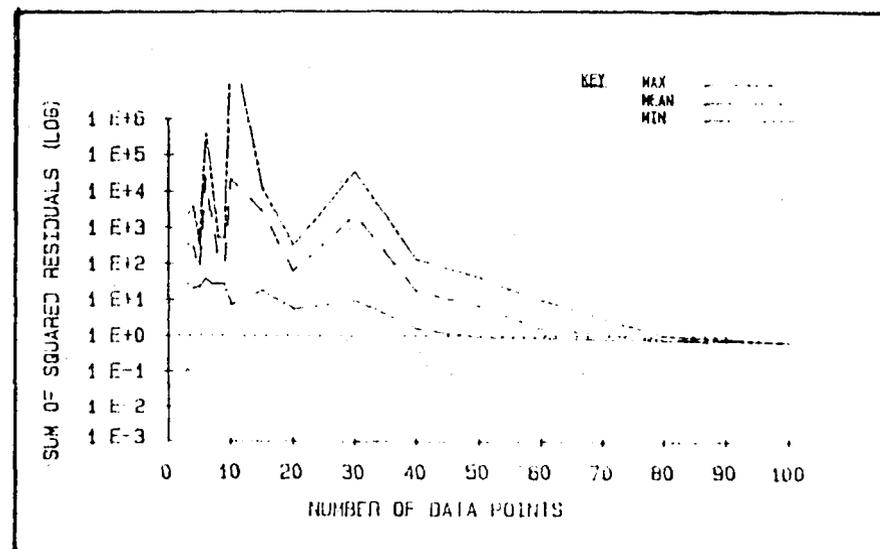


FIGURE 4.11b POLYNOMIAL (POL) APPLIED TO TEST3A
SUM OF SQUARED RESIDUALS

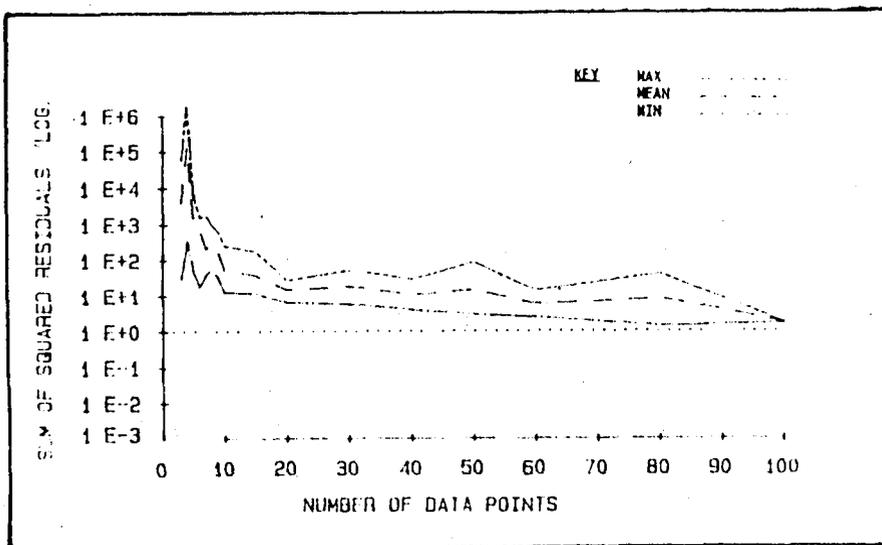


FIGURE 4.11c MULTIQUADRATIC (MQUAD) APPLIED TO TEST3A
SUM OF SQUARED RESIDUALS

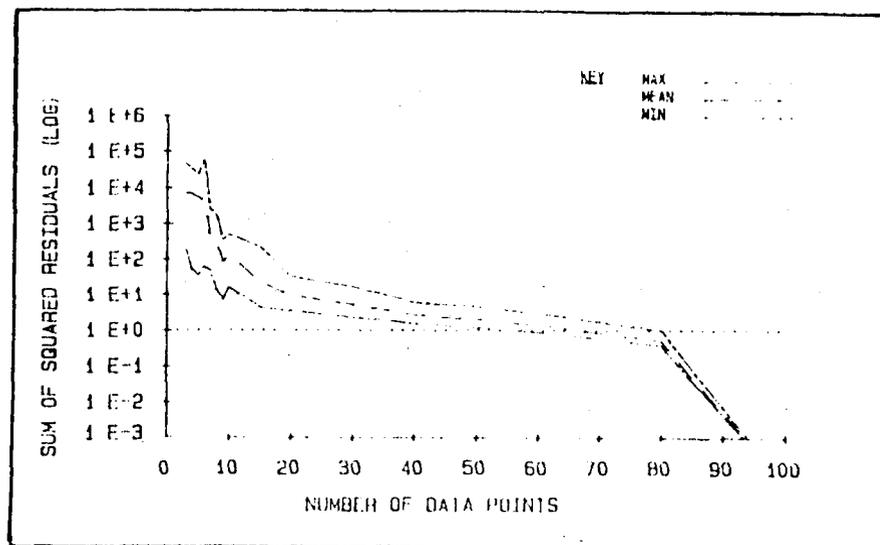


FIGURE 4.11d DISTANCE WEIGHTED (DIST) APPLIED TO TEST3A
SUM OF SQUARED RESIDUALS

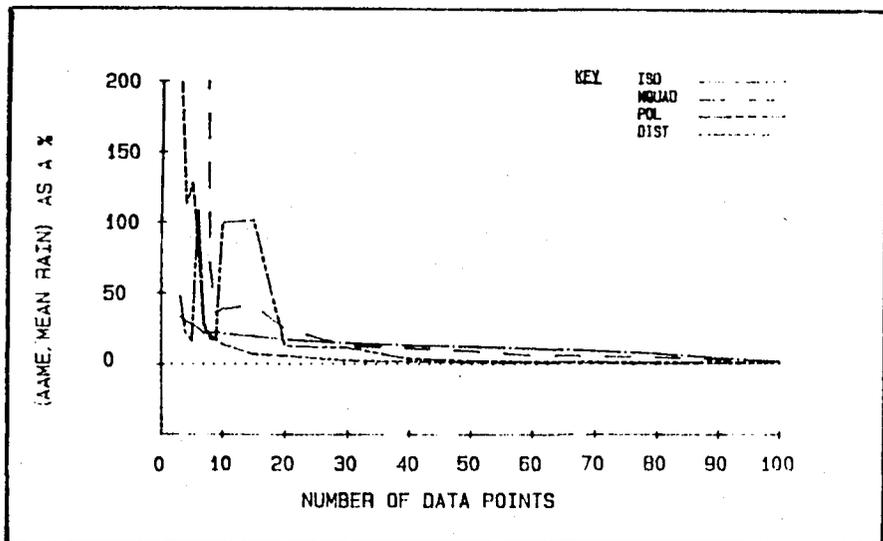


FIGURE 4.12a COMPARISON OF METHODS- TEST1A
PERCENTAGE VALUE OF (AAME/MEAN RAIN)

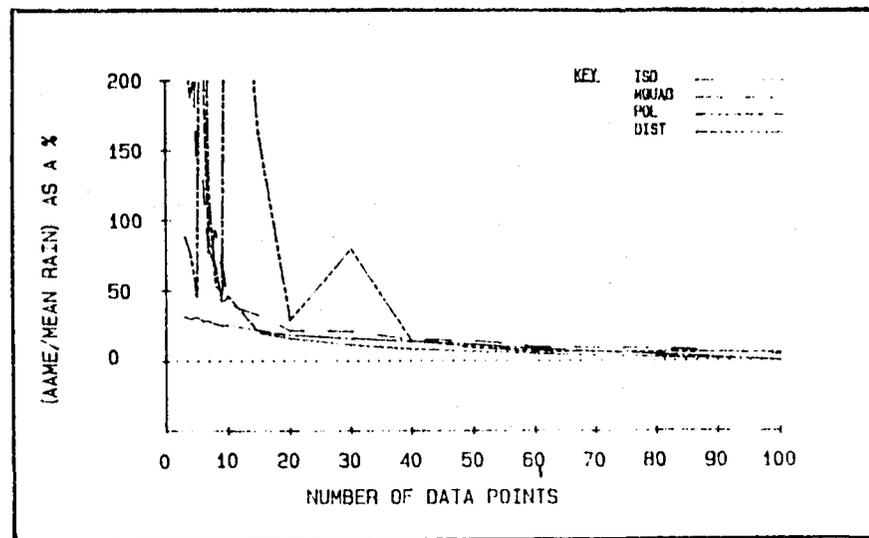


FIGURE 4.12b COMPARISON OF METHODS- TEST3A
PERCENTAGE VALUE OF (AAME/MEAN RAIN)

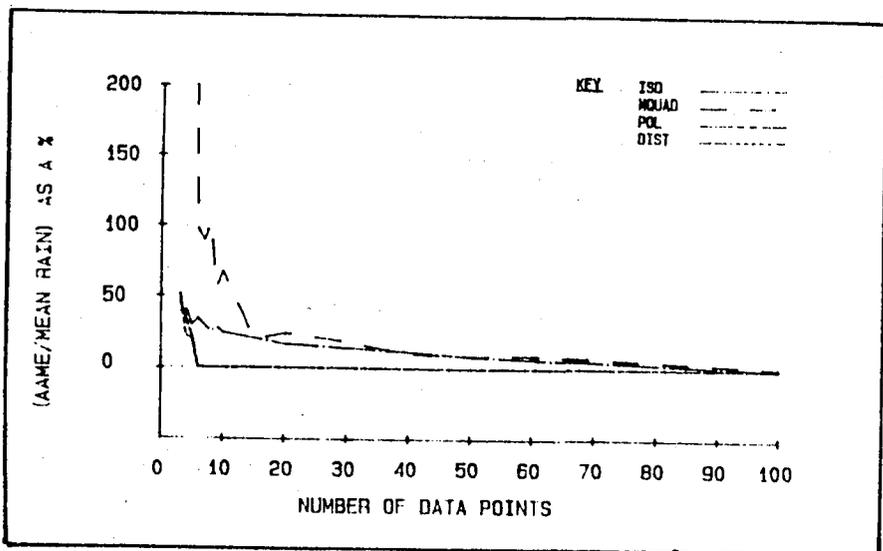


FIGURE 4.12c COMPARISON OF METHODS- TEST2A
PERCENTAGE VALUE OF (AAME/MEAN RAIN)

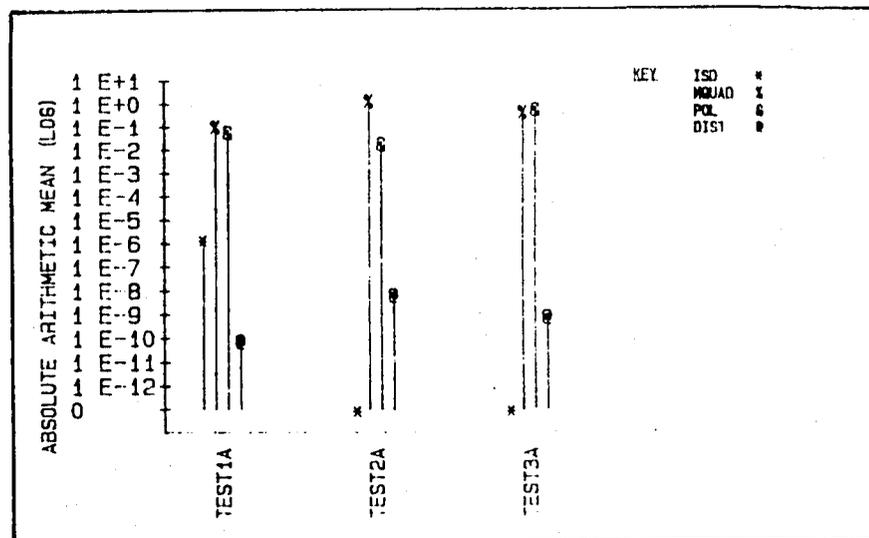


FIGURE 4.13 COMPARISON OF METHODS
AAME (100 DATA POINTS)

contours. For the purposes of this project, this implies a reliability in the ISD method that the other methods do not exhibit - especially in the areas of low numbers of data points (rain-gauges).

In test2a, both the POL and DIST methods behave very well. This is easily explained if one remembers that test2a was derived from a quadratic equation in the first place, and as these methods use polynomials to generate their interpolation values, it is not surprising that they behave well. It must be noted that for the other test data sets, the methods behave very badly in the area of low numbers of data points.

The MQUAD method also behaves consistently in all three test data sets. Unfortunately, the percentage error for less than ten data points is mostly greater than 200%.

The results from figures 4.12 a to c can be summarized as follows:

- . Methods POL and DIST behave both badly and unpredictably for low numbers of data points, except when the original polynomial surface equation used to set up the interpolation is known to follow the trend of the data closely (as in the case of test2a). At higher numbers of data points they behave well.
- . The MQUAD method behaves consistently for all test cases, but has very large errors in areas of low numbers of data points. For high numbers of data points, the method behaves well.
- . The ISD method behaves consistently for all test cases, and although the error is large in low numbers of data points, the method is more consistently accurate than the other three.
- . None of the methods behave exceptionally well in areas of low numbers of data points. The author suggests that this is not a fault of the methods themselves, but rather that they are being expected to perform far beyond the limits of their accuracy for the quantity of data being supplied as an input. This is an intrinsic problem of the process of gridding - availability of enough data points to enable a reliable fit.

Figures 4.14 a to d are included as a qualitative comparison of the methods. The figures show how all four methods interpolate to data set test1a. The four contour maps are very similar to the original data set (figure 4.6b).

It was found that if there existed large relative differences in the data, then the ISD method tended to allow the influence of the high values to "spill over" and introduce false depth-values near these peaks. This characteristic could probably be reduced by re-examining the distance weighting function to allow less influence from data points further away or to introduce a scaling factor for these cases that would reduce the extremes before gridding, and then re-scale the surface afterwards.

The methods behaved very well for the other data sets tested. The one exception to this is the POL method when applied to test3a. This is shown in figure 4.15b, with the original data set shown in figure 4.15a.

The process of selection did in fact become one of rejecting the least suitable methods until there was only one left. The results of figures 4.15a and b, and the unstable nature of the POL method as shown by figures 4.9b, 4.11b, 4.12a and c suggested the rejection of this method at this stage.

The fifth goal mentioned at the beginning of the chapter is that of speed. The whole process of gridding, contouring and plotting of maps is likely to be repeated many times for a storm event. It is therefore desirable to have a fast method. Figure 4.16 gives times taken for the gridding process for the four methods tested. It can be seen that the time taken by the DIST method is in the order of three minutes for about five rain-gauges, which is a typical number of gauges for a small catchment. This time excludes the time taken for contouring and plotting/storing the map information. The DIST method was therefore rejected because of the long times taken for gridding.

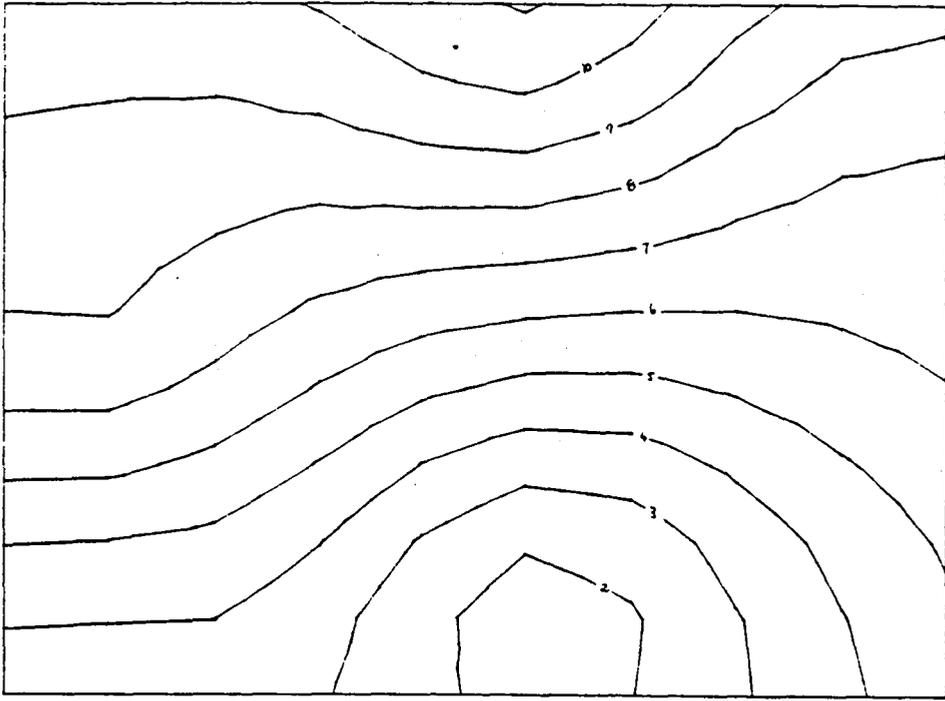


FIGURE 4.14a PLOT OF TEST1A USING ALL 100 GAUGES
METHOD USED IS: ISD

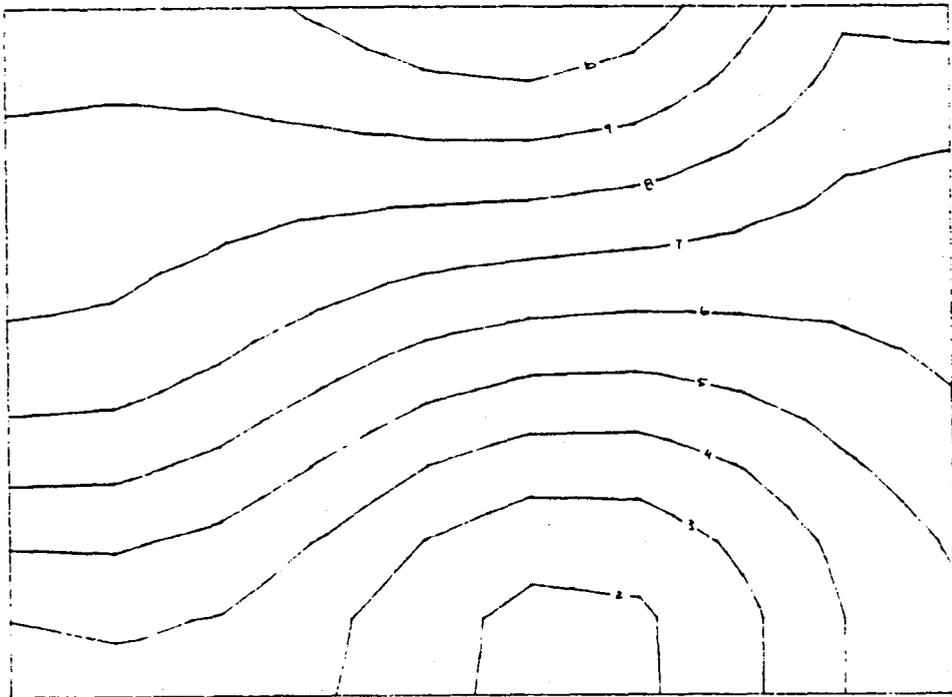


FIGURE 4.14b PLOT OF TEST1A USING ALL 100 GAUGES
METHOD USED IS: POL

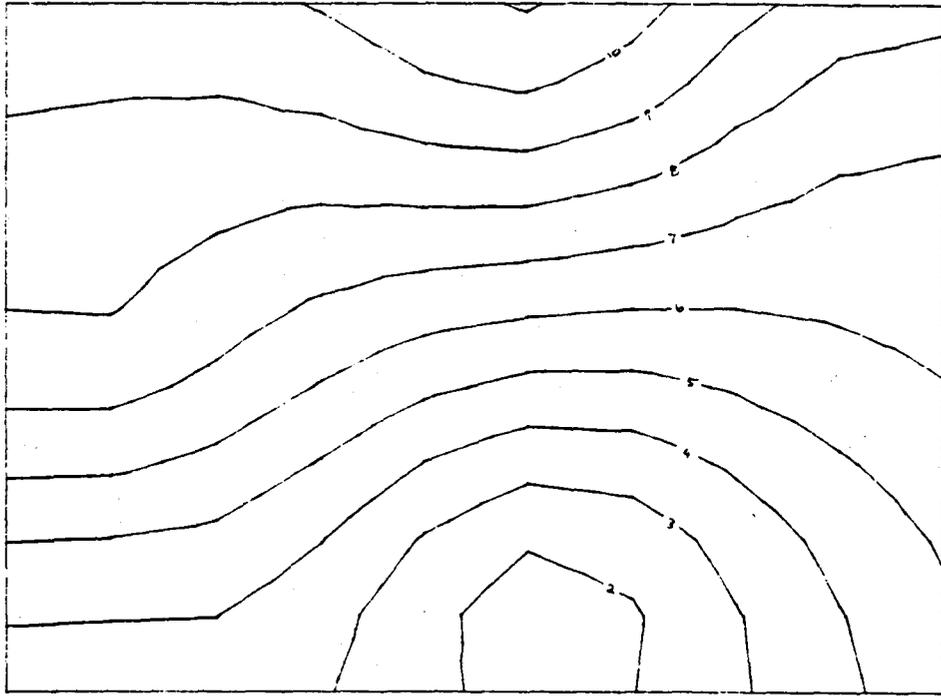


FIGURE 4.14c PLOT OF TEST1A USING ALL 100 GAUGES
METHOD USED IS: MQUAD

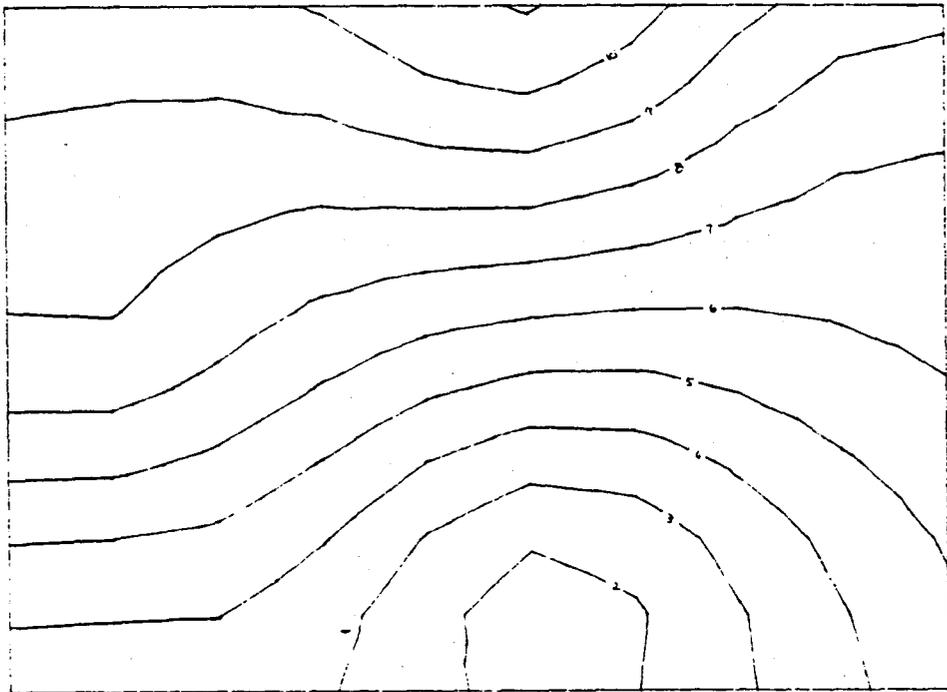


FIGURE 4.14d PLOT OF TEST1A USING ALL 100 GAUGES
METHOD USED IS: DIST

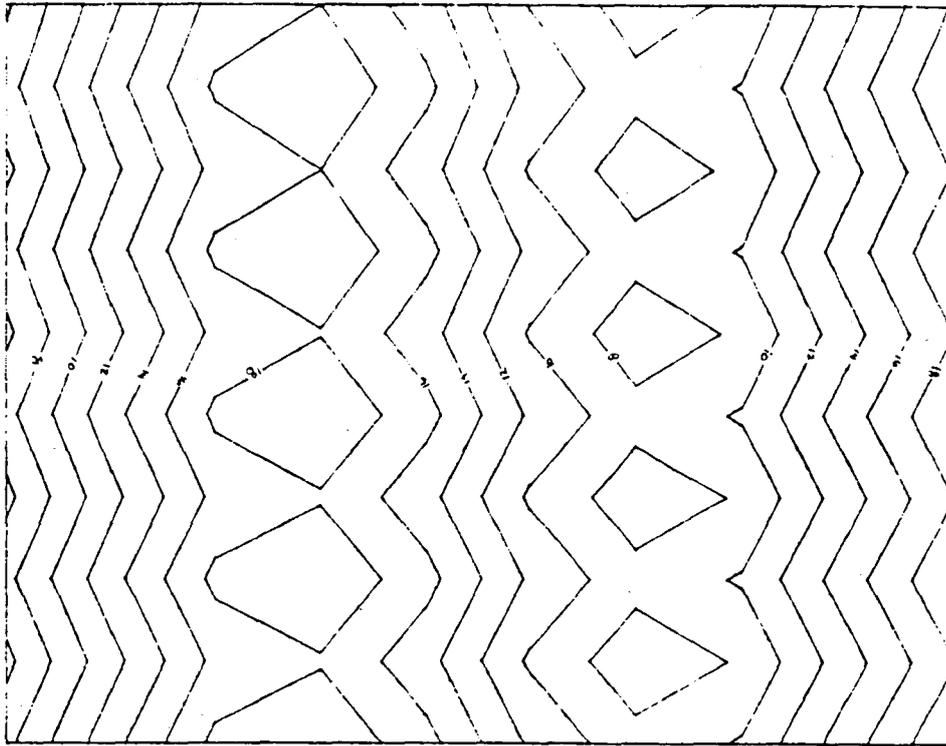


FIGURE 4.15a ACTUAL CONTOURS FOR SET: - TEST3A

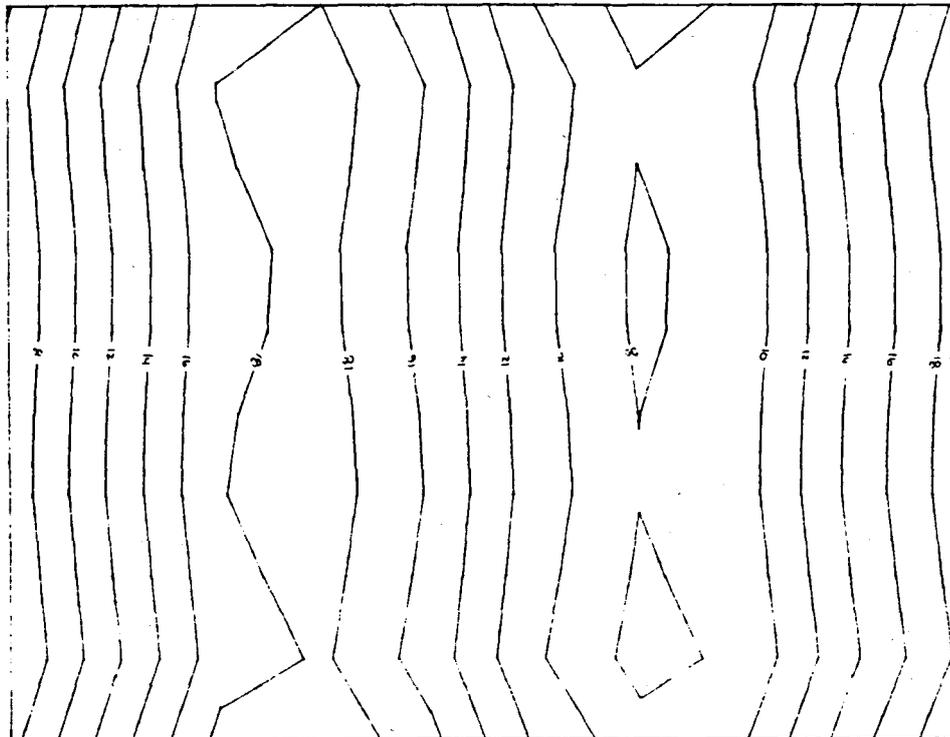


FIGURE 4.15b POLYNOMIAL METHOD APPLIED TO SET: - TEST3A
ALL ONE HUNDRED DATA POINTS USED IN INTERPOLATION

This then left two methods to choose from, the MQUAD method and the ISD method. In figure 4.16 the MQUAD method shows an exponential increase in time taken for gridding. This is explained by the fact that the method has to solve a square matrix of size equal to the number of rain-gauges, and the time taken to do this increase exponentially with the size of the matrix. This would be a drawback if more than about thirty rain-gauges are used in the study area. The study catchment focused on in the latter part of this study had five rain-gauges, so this was not a significant criterion for this project.

However, examination of figure 4.13 shows that the MQUAD method behaved badly when all one hundred data points were available for interpolation. Figures 4.9c, 4.10c and 4.11c show the method to be more unstable and consistently less accurate than the ISD method, as seen in figures 4.9a, 4.10a and 4.11a.

On this basis it was decided to use the ISD method for the remainder of this project; that is the examination of storm movement and storm patterns.

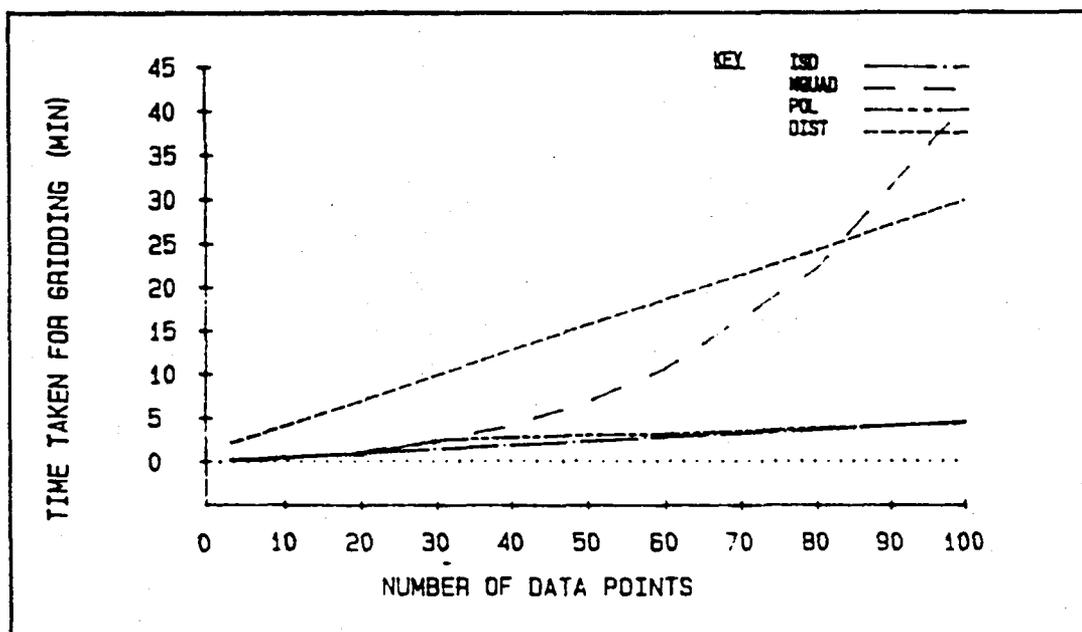


FIGURE 4.16 **COMPARISON OF METHODS**
TIME TAKEN FOR GRIDDING PROCESS IN MINUTES

5 CONTOURING PROCESS

A fast, accurate method of producing contours from a regular grid was needed for this project. A further requirement was the ability to produce a hardcopy of the contour maps for later study. A method using a simple roving three-dimensional surface was developed.

The format of interpolation points for the gridding process was a ten by ten grid of one hundred interpolated depth values. Hence, the output data file from the gridding program made it possible to determine the coordinates of each of the 100 interpolated points, and also the depth value at these interpolation points for the simple 10 by 10 grid system.

Spatially, this resulted in a series of 81 rectangles on a 9 by 9 layout of rectangles, the corners of each rectangle defined by a set of four interpolated depth values. The distance between each interpolation point was the same in the x-direction, and also the same in the y-direction. This enabled considerable computational time-saving, as will be discussed below. A roving three-dimensional surface was fitted to each rectangle in turn, the contours evaluated, and the results printed on the screen directly. The relative position of each rectangle to its neighbours had to be taken into account to print the contours in the correct place on the screen. This was superimposed over the outline of the catchment. The final product was dumped directly to the printer.

The value at the intersection of the diagonals of each rectangle was evaluated by taking the average of the four corner values (this is an application of the ISD method, but because the distances used are the same, it reduces to an arithmetic average calculation). This defined four equilateral triangles of equal area, with differing but constant orientation, and differing inclinations determined by the heights of the vertices.

In addition, each rectangle was converted to a square with dimensions of one hundred units, by dividing the distances between interpolation points by a delta-x and delta-y value. This simplified the coding by

making interpolation calculations completely general for any data set to be interpolated. The coordinates of the contours, once calculated, were post-multiplied by the same delta-x and delta-y values to regain the original coordinate system.

Thus, the roving three-dimensional surface used was simplified to four general plains, as opposed to a higher order polynomial expression. The advantage of this is that defining the four plains in the way described above produced a series of equations that were implicit and could be solved directly. Higher order expressions would have required iterative solutions, and considerably more computational time. An algorithm was developed for evaluating the point of entry and exit any contour would have, if it fell within the limits of the three vertices of the triangle. This was repeated as four sub-portions of the program for the four orientation possible in one square. This process was repeated for each of the 81 squares on the interpolation grid. Because each triangle was joined to it's neighbour by the same slope edge, contours always matched up at the edges of each triangle and square.

Hence two coordinate systems were used: A local coordinate system related to the bottom left hand corner of each square, which remained constant for each square, and a global coordinate system related to the bottom left hand corner of the real system. Each time a square had been completely evaluated, the results were converted to the global coordinate system and plotted on the screen. This method of splitting the contouring into two distinct portions made the determination of contour positions, which required the most calculation, very fast. The conversion to the global coordinate systems was used when each rectangle was completely contoured and the results were superimposed over the catchment map on the screen.

The accuracy of the method was found to be more than adequate for the task, and was more a function of the number of interpolation points than the use of four triangular plane surfaces. The method is simple in concept, but the programming proved to be more complex than anticipated. The complete listing of this program is given in appendix E.

6 STUDY CATCHMENT

A peri-urban catchment with marked changes in elevation and of a suitable area was desired for this study. The catchment used, Montgomery Park, is in the north western area of Johannesburg municipality and contains ten smaller suburbs. At present the land use is peri-urban, with a continuing expansion of residential areas. The area of the catchment is about 10.36km². The change in elevation from lowest to highest point in the catchment is about 200m, which is high for this geographical location.

The catchment is drained by the Montgomery Spruit, a tributary of the Braamfontein Spruit. The Montgomery Spruit flows in a north easterly direction. The northern boundary of the catchment is bordered by a north facing escarpment, with a south facing dip slope of about 10°. Thus the catchment has high areas to the north, west and south, with shallow areas in the centre and east, where the Montgomery Spruit leaves the catchment (see figure 6.1).

The land use is varied, with natural vegetation predominating. Soil is also natural with a small amount of landfill the southern portion of the catchment. There are some light industry stands, but most are residential. A breakdown of land use is seen in table 6.1.

Land Use Percentages - Montgomery Park.	
Land Use Type	Percentage
Rural - Low level vegetation	77,0
Rural - High level vegetation	2,5
Urbanised areas	18,0
Landfill	2,5
Water surfaces	<1,0

TABLE 6.1 LAND USE PERCENTAGES
(After Lambourne and Stephenson, 1986)

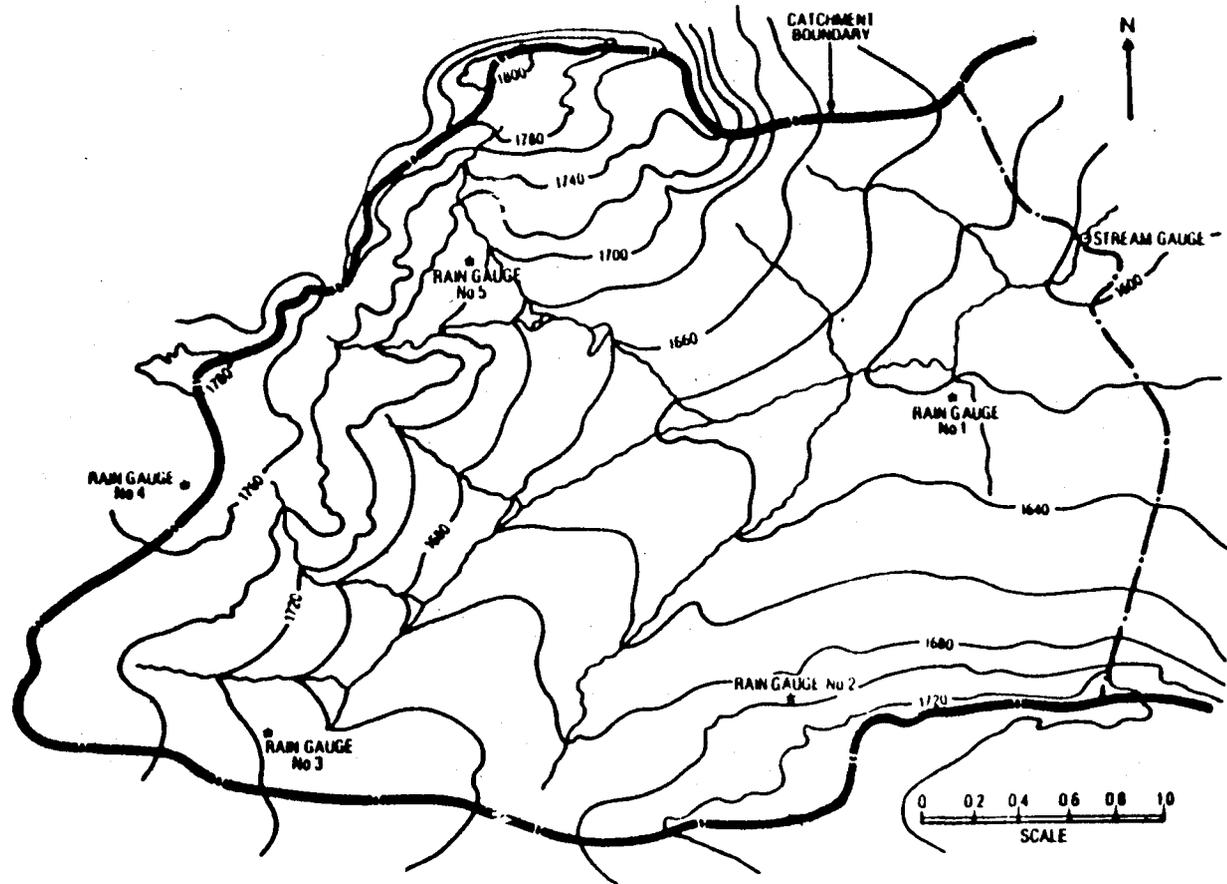


FIGURE 6.1 CONTOUR MAP OF STUDY CATCHMENT, MONTGOMERY PARK
(After Green and Stephenson, 1986)

There are five rain-gauges in the catchment. Four of them are within the boundary of the catchment. Two of the gauges are of the syphon type, while the other three are tipping bucket types, with resolutions of .2mm per tip. All five gauges are logged by means of clockwork recorders. Charts are changed weekly, except for one of the gauges, which is changed monthly.

Various attempts were made to determine the relative influence of the five gauges, such as building a scale relief model and calculating the Thiessen weights (see figure 6.2).

6.1 STORM PATTERN CHARACTERISTICS

As described in chapter four, four gridding methods were tested to determine the most reliable way of producing storm patterns from rain-gauge data. The method chosen, Inverse Squared Distances (ISD), was used to convert the five gauge readings into a grid of interpolated depths over the whole catchment. These depths were then converted to contours of common intensities, using a program written by the author. This entailed three stages:

- . Converting five charts into intensities of rainfall at five minute time intervals and synchronizing the data by real-time into one file. Five minutes was considered a reasonable balance between accuracy of resolution and saving of storage space on disc.
- . Reading this file and interpolating the depths for each five minute time interval over the whole catchment by the ISD method.
- . Taking each five minute interpolation and producing a contour map of rainfall intensities for the whole catchment, and printing this map.

MONTGOMERY PARK

NO. OF POINTS USED= 2441

Y STEPS= 50

TOTAL AREA = 10.36km²

GAUGE	POINTS	WEIGHT (%)	AREA (km ²)
1	736	30.15	3.124
2	529	21.67	2.245
3	373	15.28	1.583
4	220	9.013	0.934
5	583	23.88	2.474

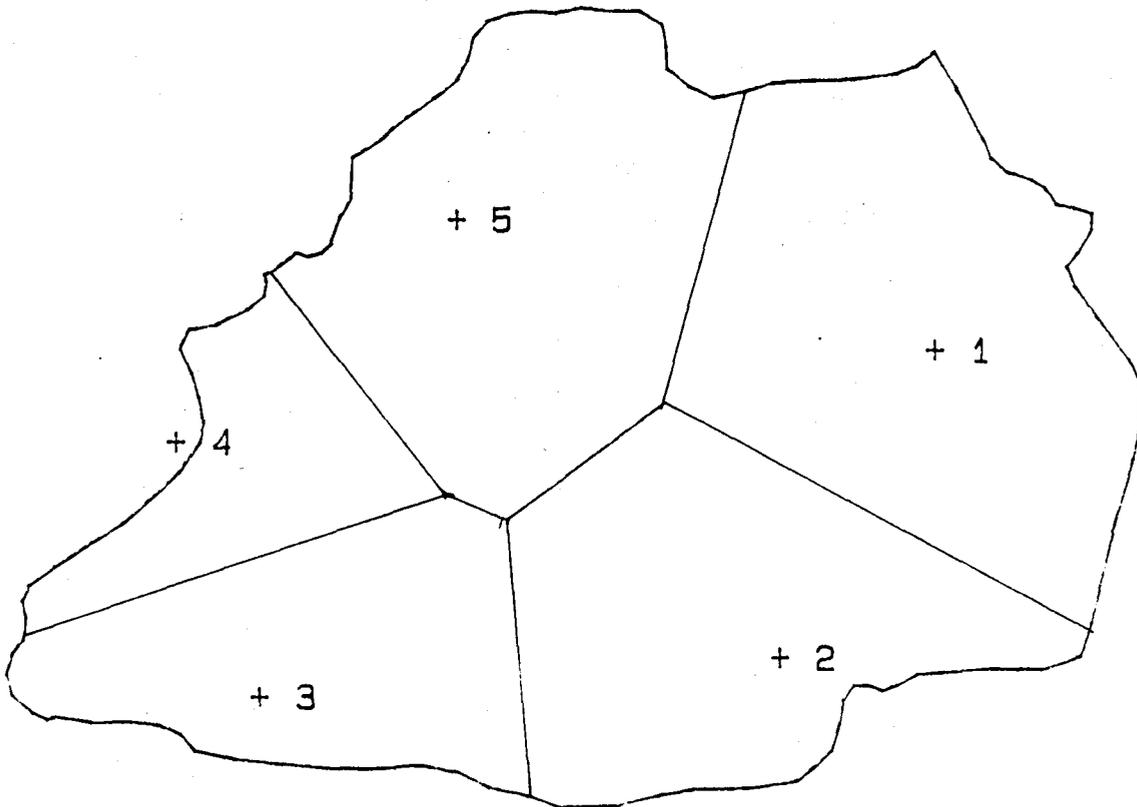


FIGURE 6.2 THIESSEN WEIGHTS AND AREAS FOR MONTGOMERY PARK

The total real-time duration of this file was then from the first onset of rain till the last measurement of rain over all five gauges. The number of maps produced per storm event is the real-time duration in minutes divided by five. Storms with rainfall at four or less gauges were not considered.

Twenty-one storm events were extracted from records starting at the beginning of 1987 and ending in the middle of 1988, and studied in this way. A large variety of events was studied, with mean total depths ranging from about 5mm to about 55mm, and durations varying from about 100 minutes to about 1000 minutes.

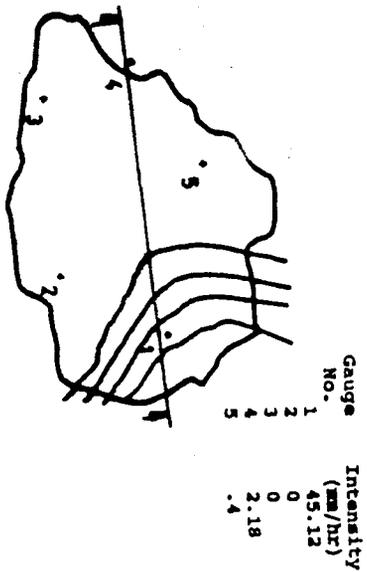
6.1.1 FINDINGS FROM CONTOUR MAPS

With twenty-one storms studied, hundreds of five-minute maps were produced. It is not feasible to include all of them here, and hence some definitive characteristics are highlighted in this section by the author.

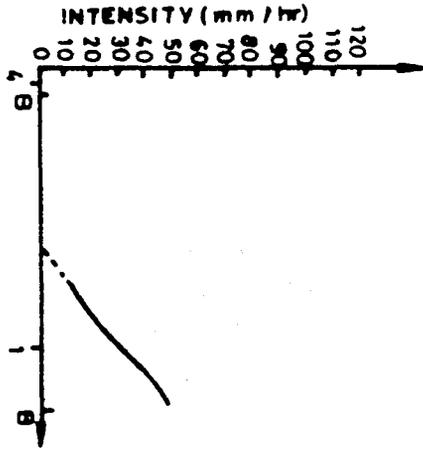
Figure 6.3, from a storm on the 3/02/87, shows a series of eighteen maps in sequence with cross-sections. In the beginning a cell of high intensity forms over gauge one. As this starts to decay, another cell forms at gauge four, and the two of them produce rain at a medium intensity for some twenty-five minutes. These two disappear, and a third cell of high intensity forms over gauge five and remains in effect for some twenty-five minutes. This is fairly typical of the storms studied, with several cells starting and decaying in different parts of the catchment at different times during the duration of one storm event.

The cross-sections give an added dimension to visualize the differences in intensities across the catchment.

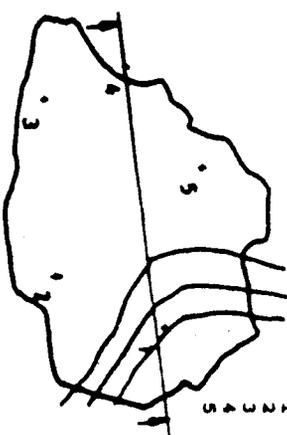
storm start time: 5.15
 contour interval: 10mm/hr
 min passed: 90
 map time step: 5 min



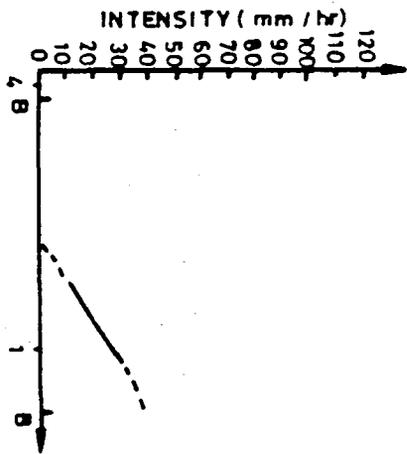
Gauge No.	Intensity (mm/hr)
1	45.12
2	0
3	0
4	2.18
4	.4
5	.4



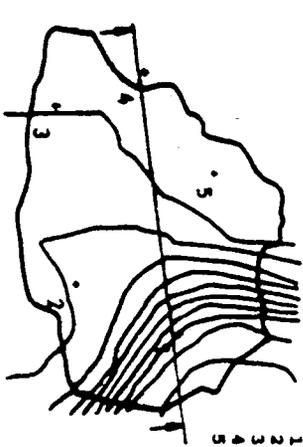
storm start time: 5.15
 contour interval: 10mm/hr
 min passed: 95
 map time step: 5 min



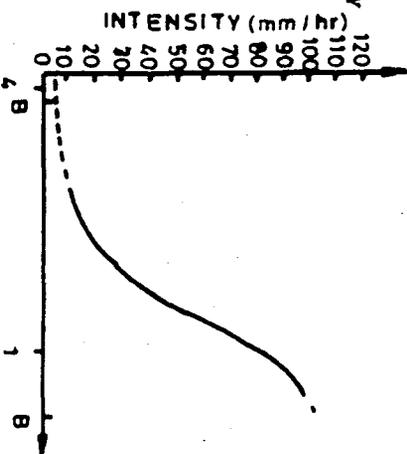
Gauge No.	Intensity (mm/hr)
1	37.98
2	0
3	0
3	5.45
4	0
4	5.45
5	2



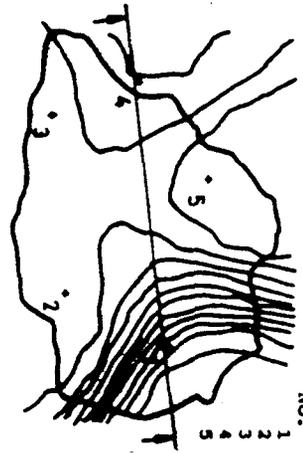
storm start time: 5.15
 contour interval: 10mm/hr
 min passed: 100
 map time step: 5 min



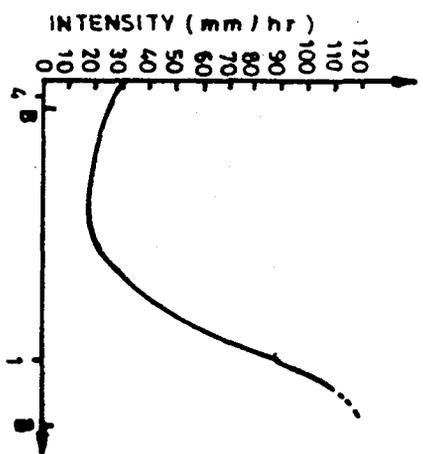
Gauge No.	Intensity (mm/hr)
1	102.9
2	19.2
3	9.9
3	5.76
4	1.2
4	1.2
5	1.2



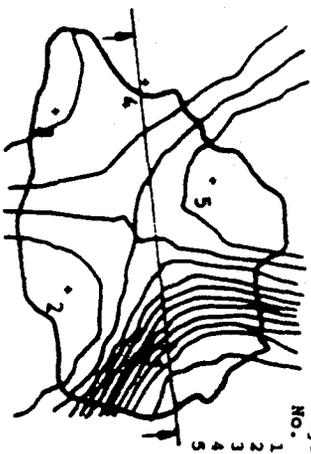
storm start time: 5.15
 contour interval: 10mm/hr
 min passed: 105
 map time step: 5 min



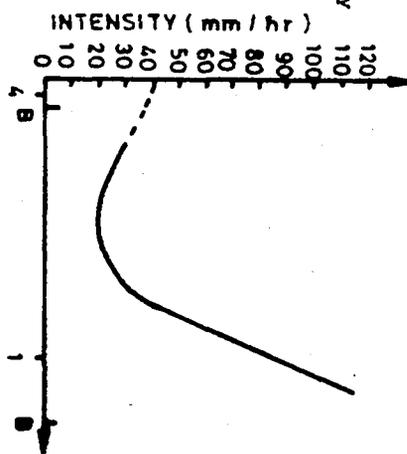
Gauge No.	Intensity (mm/hr)
1	120
2	12.8
3	16.5
3	10.9
4	0
4	0
5	0



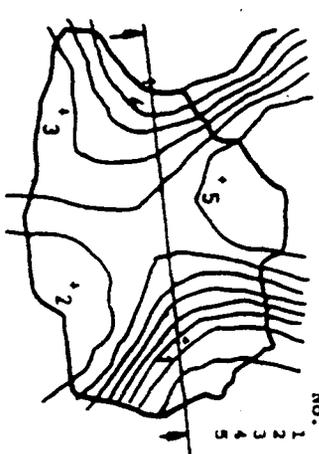
storm start time: 5.15
 contour interval: 10mm/hr
 min passed: 110
 map time step: 5 min



Gauge No.	Intensity (mm/hr)
1	117.60
2	2.8
3	42
3	38.75
4	0
4	0
5	0



storm start time: 5.15
 contour interval: 10mm/hr
 min passed: 115
 map time step: 5 min



Gauge No.	Intensity (mm/hr)
1	84
2	3.9
3	25.38
3	68.95
4	0
4	0
5	0

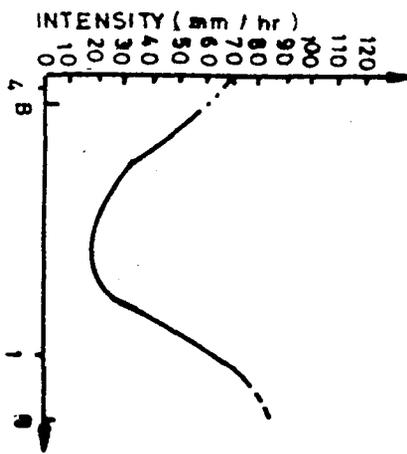
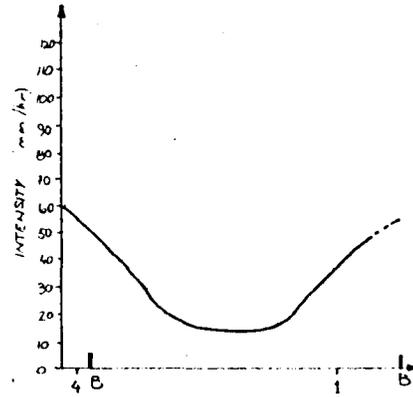
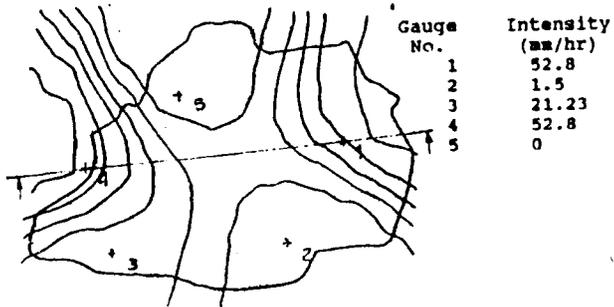
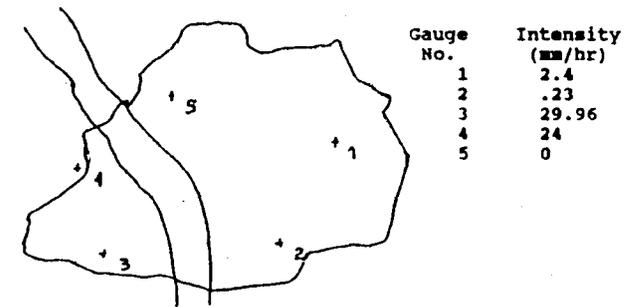


FIGURE 6.3 CONTOURS FROM A SECTIONMENT OF A STORM ON THE 3/02/1987

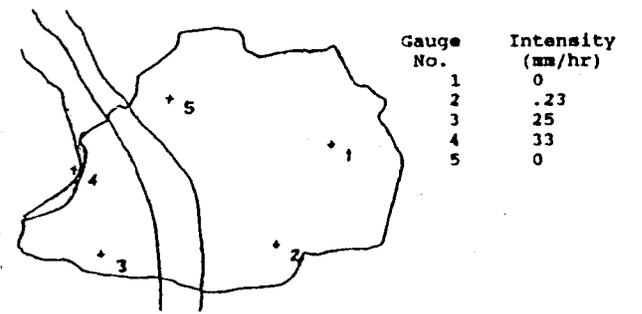
storm start time: 5.15 min passed: 120
 contour interval: 10mm/hr map time step: 5 min



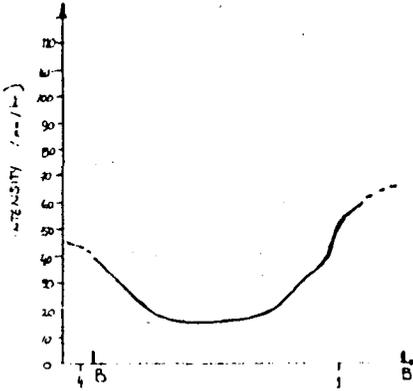
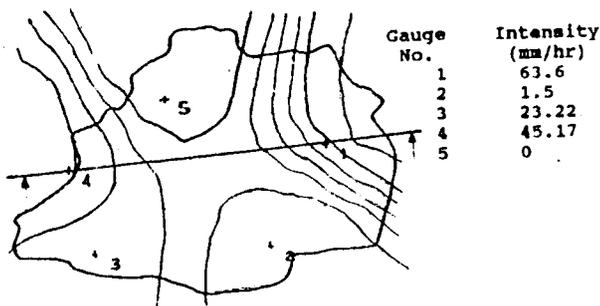
storm start time: 5.15 min passed: 135
 contour interval: 10mm/hr map time step: 5 min



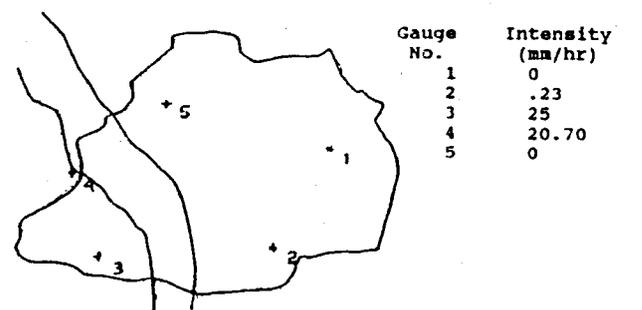
storm start time: 5.15 min passed: 140
 contour interval: 10mm/hr map time step: 5 min



storm start time: 5.15 min passed: 125
 contour interval: 10mm/hr map time step: 5 min



storm start time: 5.15 min passed: 145
 contour interval: 10mm/hr map time step: 5 min



storm start time: 5.15 min passed: 130
 contour interval: 10mm/hr map time step: 5 min

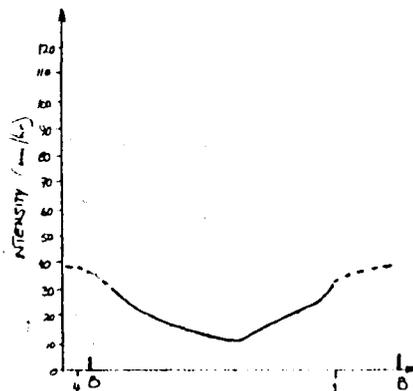
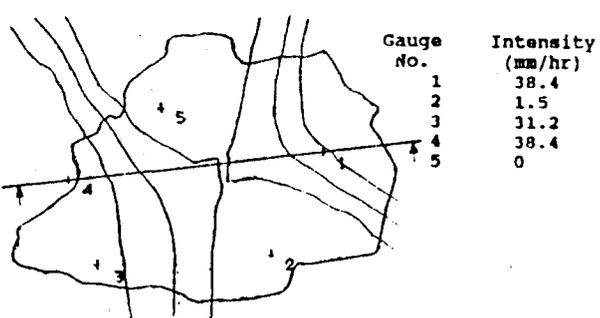


FIGURE 6.3 CONTINUED

storm start time: 5.15
contour interval: 10mm/hr

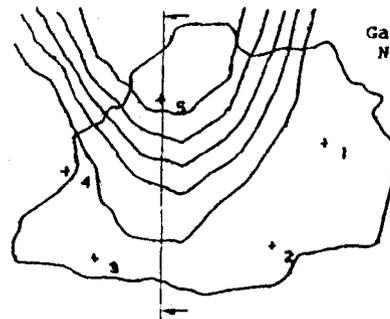
min passed: 150
map time step: 5min



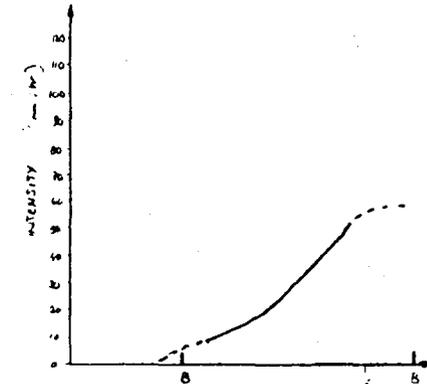
Gauge No.	Intensity (mm/hr)
1	0
2	.23
3	20.71
4	10.2
5	17.7

storm start time: 5.15
contour interval: 10mm/hr

min passed: 165
map time step: 5min

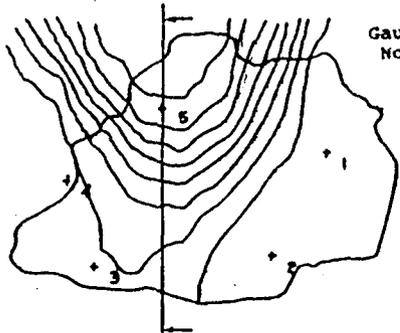


Gauge No.	Intensity (mm/hr)
1	0
2	.23
3	8.25
4	3.27
5	59.14

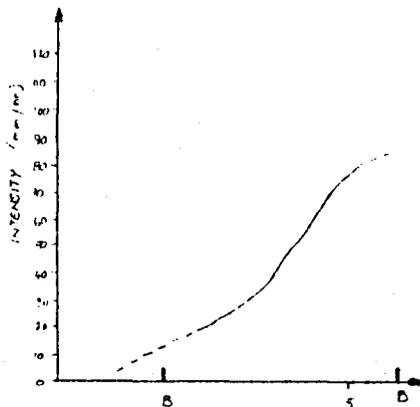


storm start time: 5.15
contour interval: 10mm/hr

min passed: 155
map time step: 5min

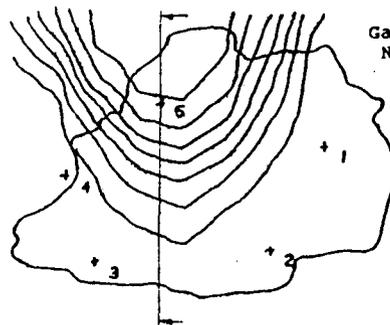


Gauge No.	Intensity (mm/hr)
1	0
2	.23
3	19.63
4	14.04
5	84.7

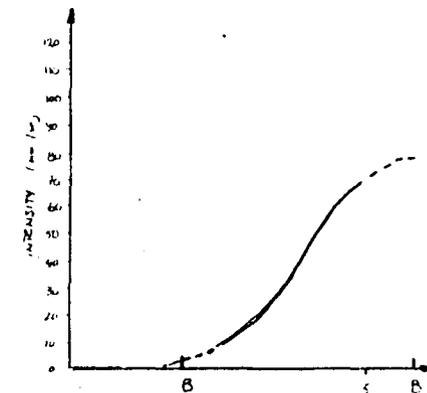


storm start time: 5.15
contour interval: 10mm/hr

min passed: 175
map time step: 5min

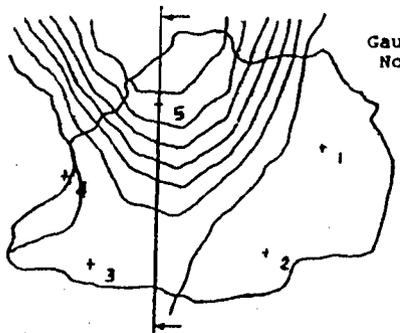


Gauge No.	Intensity (mm/hr)
1	0
2	.23
3	2.82
4	2.57
5	76.62

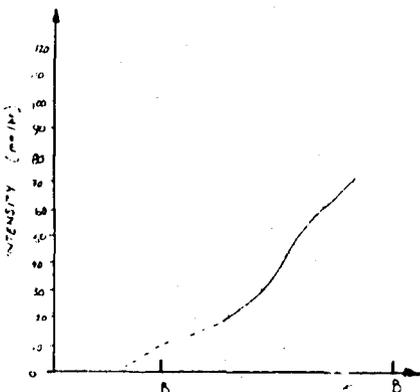


storm start time: 5.15
contour interval: 10mm/hr

min passed: 160
map time step: 5min

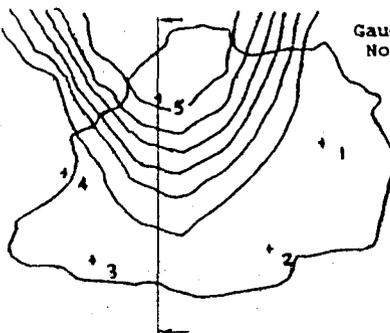


Gauge No.	Intensity (mm/hr)
1	0
2	0.23
3	12.8
4	3.75
5	75.02



storm start time: 5.15
contour interval: 10mm/hr

min passed: 175
map time step: 5min



Gauge No.	Intensity (mm/hr)
1	0
2	.23
3	2.82
4	.25
5	69.16

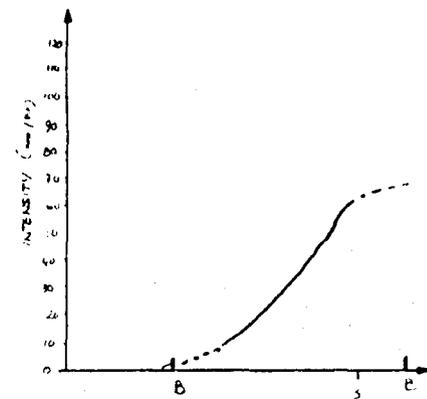


FIGURE 6.3 CONTINUED

An unknown quantity is whether two cells can overlap, or combine partially. It is difficult to measure this from the ground, as local wind effects may produce what appears to be two combined cells. Figure 6.4, a map from a storm on the 19/11/87, could show either one large cell, or two cells overlapping.

STORM START TIME: 3:55 MIN PASSED: 88
 CONTOUR INTERVAL: 18mm/hr MAP TIME STEP: 5min

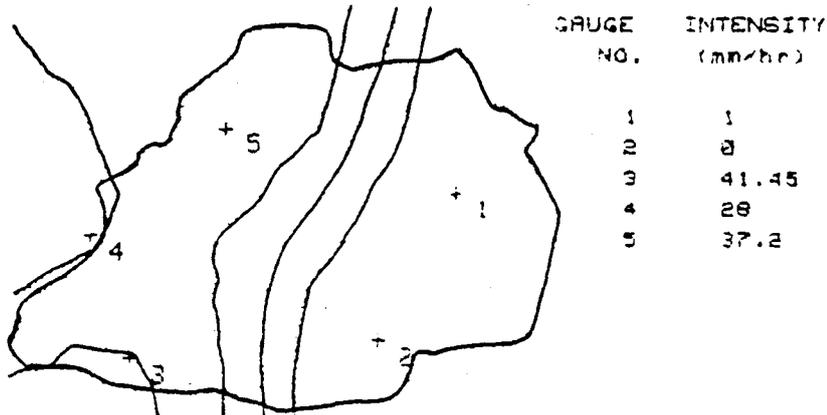


FIGURE 6.4 ABSENCE OF CENTRAL RAINGAUGE MAKES ANALYSIS DIFFICULT IN THIS CASE

There was also evidence of spotty rain in the study catchment. Figure 6.5, from a storm on the 27/11/87, shows a large amount of rainfall appearing at one gauge, and very little elsewhere. Conditions before and after this time interval did not exhibit this disproportionality in intensities.

STORM START TIME: 17:35 MIN PASSED: 185
 CONTOUR INTERVAL: 18mm/hr MAP TIME STEP: 5min

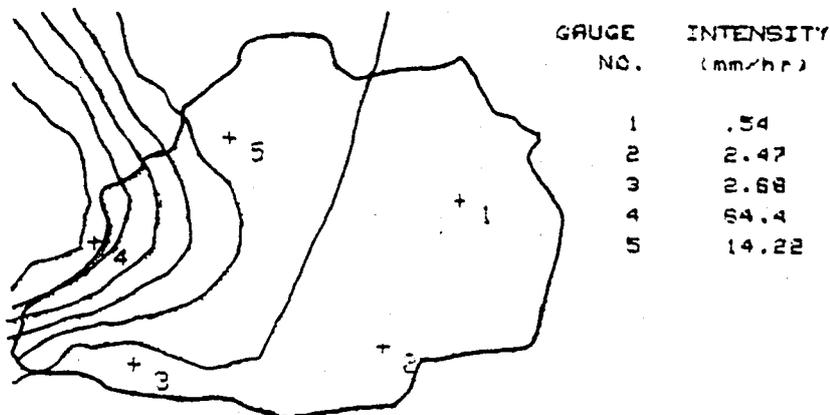
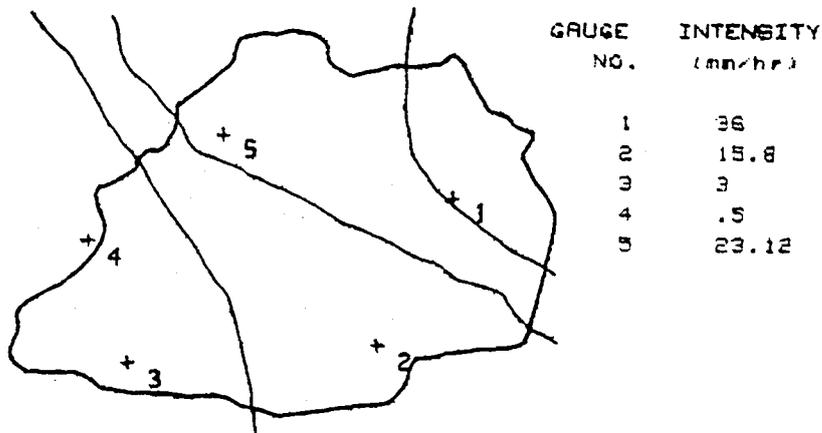


FIGURE 6.5 SPOTTY RAINFALL FOR THE STORM ON 27/11/1987

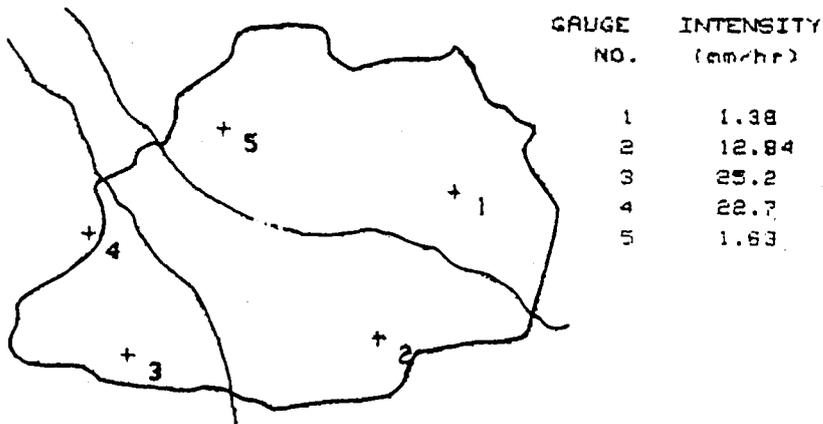
There were some events which did not have a marked cellular structure. Figure 6.6, from a storm on the 21/01/88, shows a more planar distribution of rainfall, as opposed to isolated cellular events.

STORM START TIME: 15:5 MIN PASSED: 40
 CONTOUR INTERVAL: 10mm/hr MAP TIME STEP: 5min



MAP AFTER 45 HAS ALL GAUGE-VALUES LESS THAN 10mm/hr

STORM START TIME: 15:5 MIN PASSED: 50
 CONTOUR INTERVAL: 10mm/hr MAP TIME STEP: 5min



STORM START TIME: 15:5 MIN PASSED: 55
 CONTOUR INTERVAL: 10mm/hr MAP TIME STEP: 5min

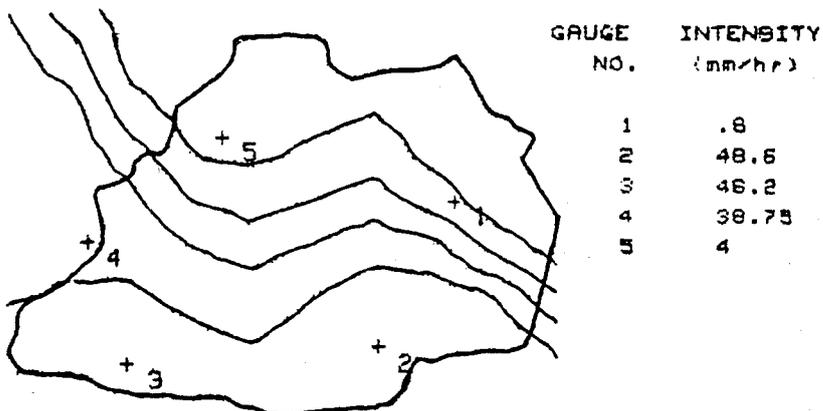


FIGURE 6.6 ABSENCE OF CELLULAR STRUCTURE FOR THE STORM ON 21/01/1988

At the other extreme, there were cases of very convoluted shapes appearing. Figure 6.7, from a storm on the 24/01/88, shows a series of six maps where the shape of one map appears to have little bearing on the shape of the previous or subsequent map.

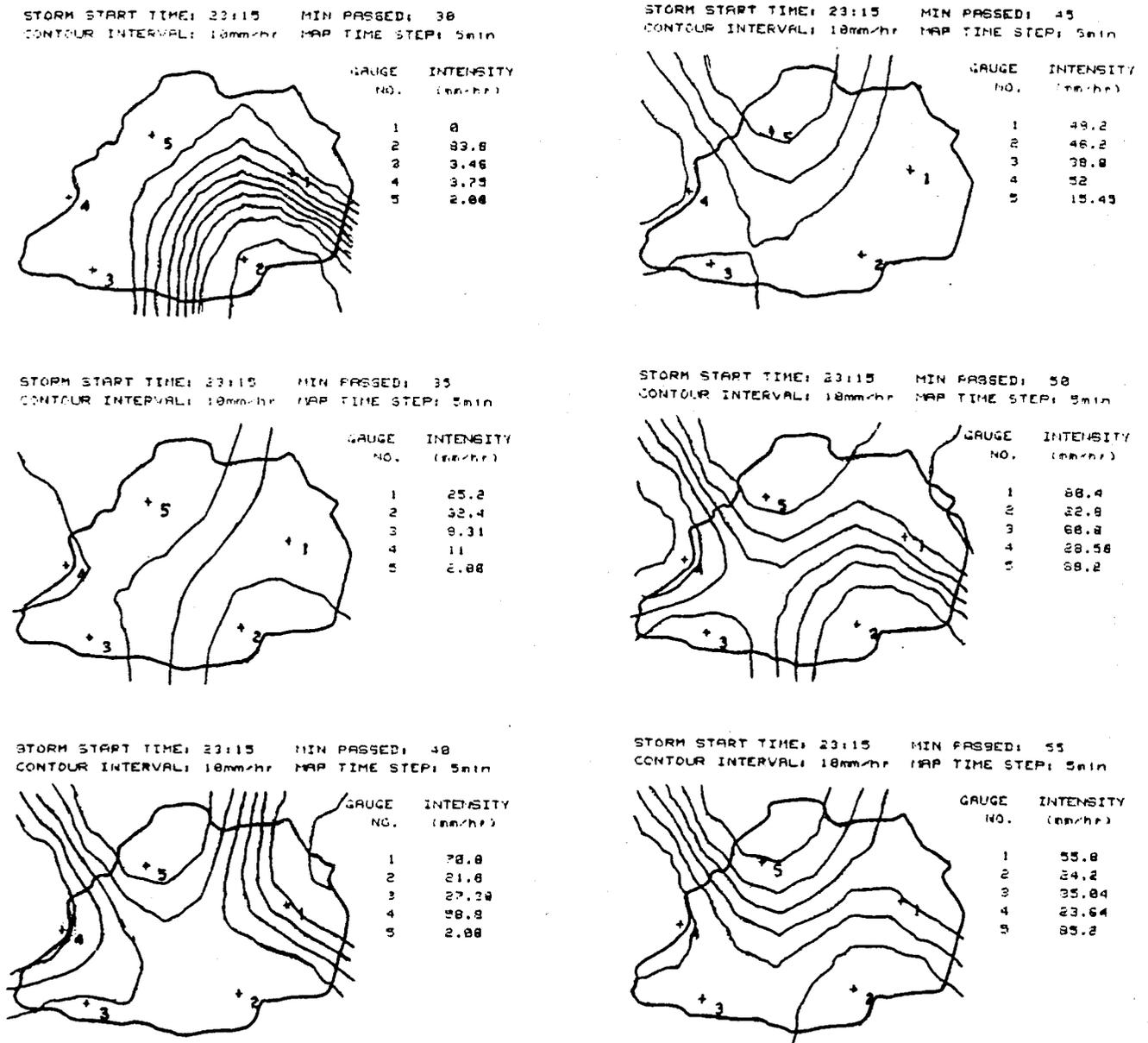
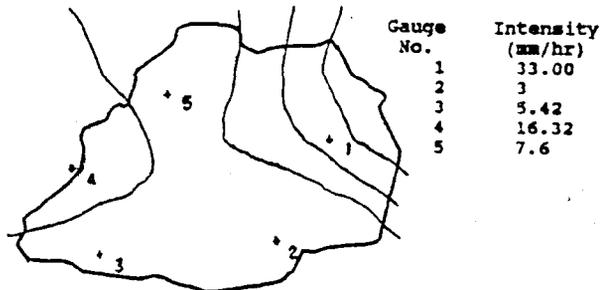


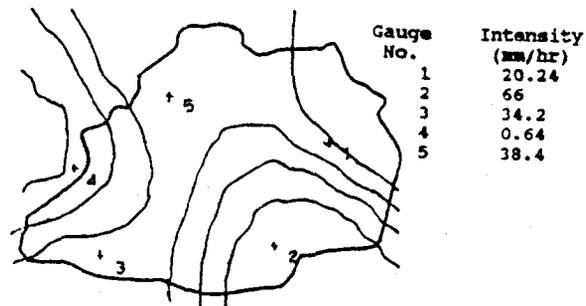
FIGURE 6.7 ERRATIC BEHAVIOUR OF A PORTION OF THE STORM ON 24/01/1988

In some cases it was possible to discern movement of individual cells. Figure 6.8, from a storm on the 2/03/88, shows a cell that originates over gauge number one and drifts south over gauge number two and decays. This sequence lasts about twenty-five minutes, and on the last map one can see the emergence of a new cell over gauge number two.

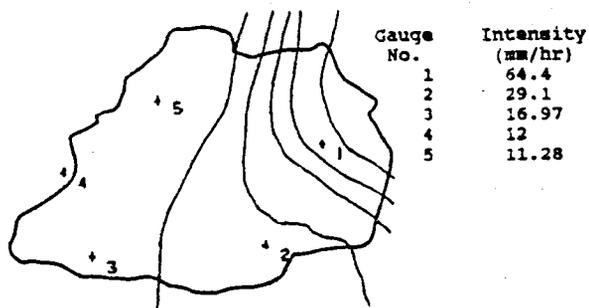
storm start time: 17.45 min passed: 35
 contour interval: 10mm/hr map time step: 5 min



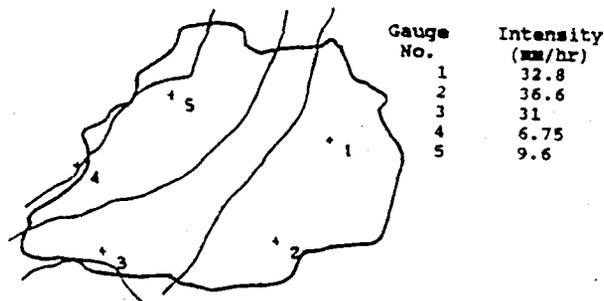
storm start time: 17.45 min passed: 50
 contour interval: 10mm/hr map time step: 5 min



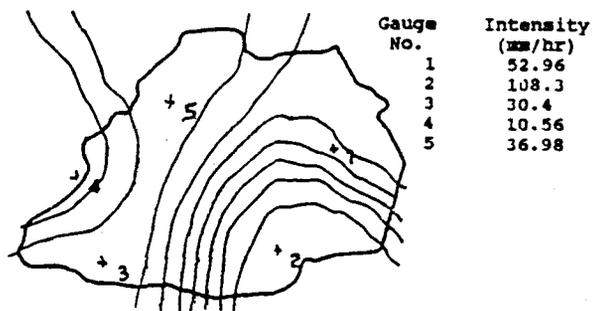
storm start time: 17.45 min passed: 40
 contour interval: 10mm/hr map time step: 5 min



storm start time: 17.45 min passed: 55
 contour interval: 10mm/hr map time step: 5 min



storm start time: 17.45 min passed: 45
 contour interval: 10mm/hr map time step: 5 min



storm start time: 17.45 min passed: 60
 contour interval: 10mm/hr map time step: 5 min

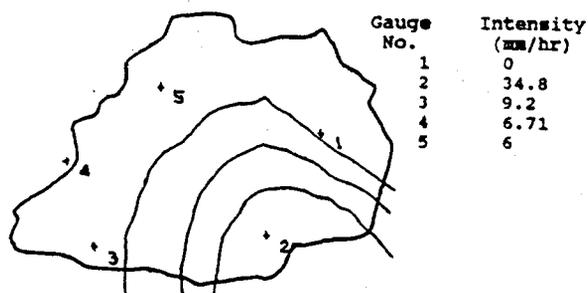


FIGURE 6.8 MOVEMENT OF A CELL OVER THE STUDY CATCHMENT

There were cases when the local storm did not cover the study catchment over a significant area. Figure 6.9, from a storm on the 11/03/88, shows a period of thirty minutes, during which time only gauge number four received any significant rainfall.

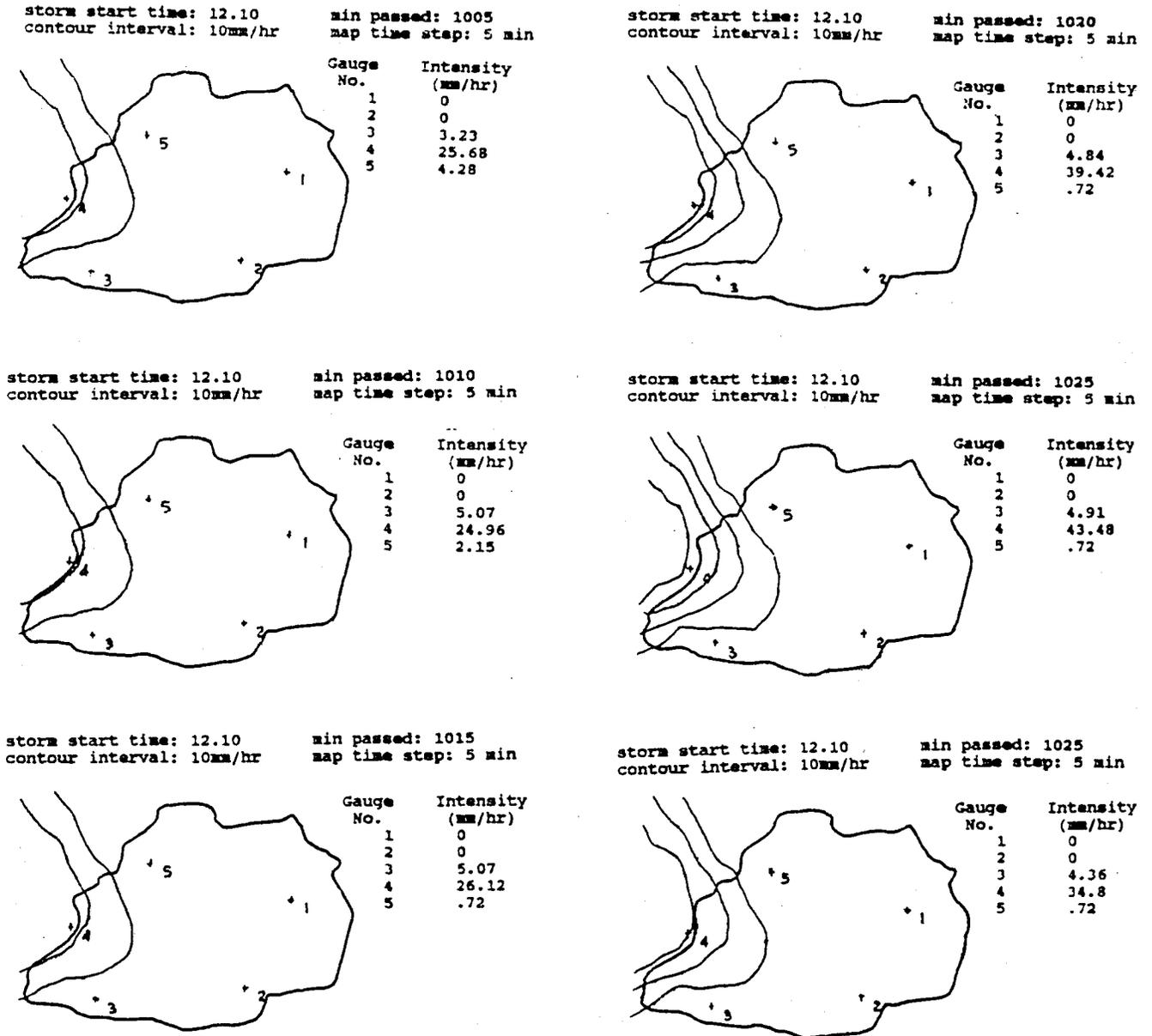
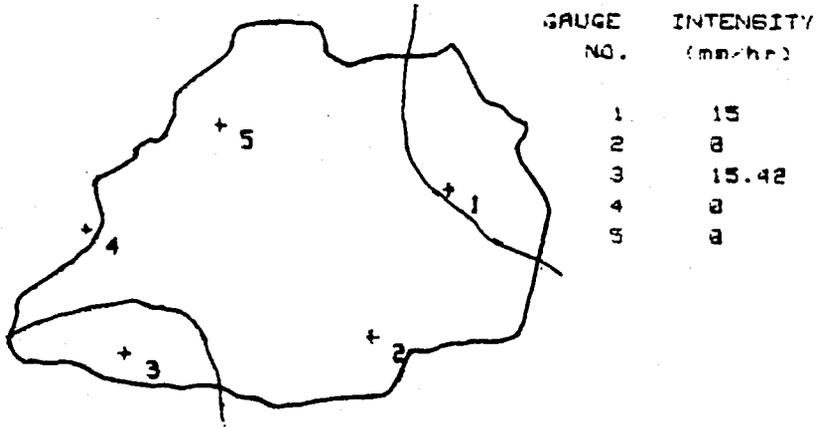


FIGURE 6.9 PARTIAL COVERAGE OF CATCHMENT BY STORM EVENT/CELL

Because the areal coverage of storm cells is in the same order of magnitudes as the distances between the gauges in the study catchment, it was felt that the absence of a gauge in the centre of the catchment may have affected the study. Figure 6.10, from a storm on the 6/11/87

STORM START TIME: 18:30 MIN PASSED: 5
 CONTOUR INTERVAL: 10mm/hr MAP TIME STEP: 5min



STORM START TIME: 18:30 MIN PASSED: 10
 CONTOUR INTERVAL: 10mm/hr MAP TIME STEP: 5min

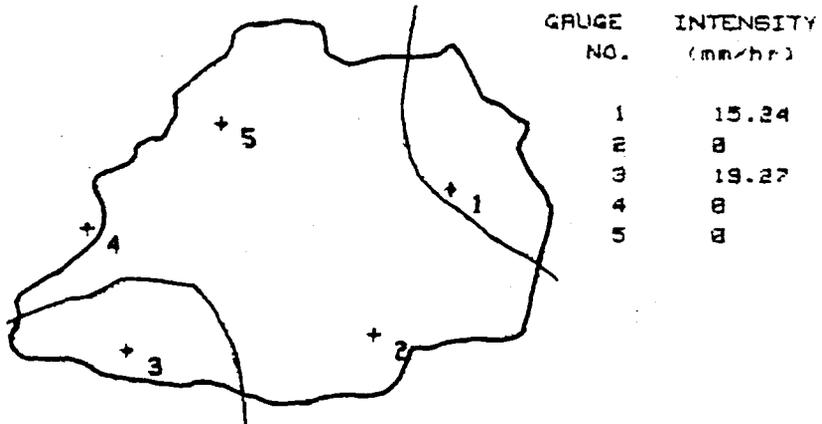


FIGURE 6.10 A CENTRAL RAINGAUGE WOULD CLARIFY THESE TWO CONTOUR MAPS

shows two maps where a long, narrow belt of rain running in a North-East to South-West direction could just as easily have produced the effect shown as two individual cells over gauges one and three. The absence of further definite data makes any attempts at more detailed interpolation spurious.

This is not to be confused with the patterns shown in figure 6.3. In this case it is unlikely that a single cell could produce the high intensities over gauges one and four, and the low intensities over gauges two and five, because of the physical size such intense cells are purported to have (about 3km diameter, from chapter 3).

6.2 TOPOGRAPHICAL EFFECTS

A study was also done on the effects elevation might have on rainfall amounts at the gauges in the catchment. The total depths of rainfall for the twenty-one storms were added up for each gauge and a linear correlation done with height above mean sea level (see figure 6.11). The equation of the line of best fit was found to be:

$$\text{Depth} = 0.9645 \times \text{Height} - 1127$$

where Depth is in mm and Height is in m above sea level.

The correlation coefficient for the fit was 0.63.

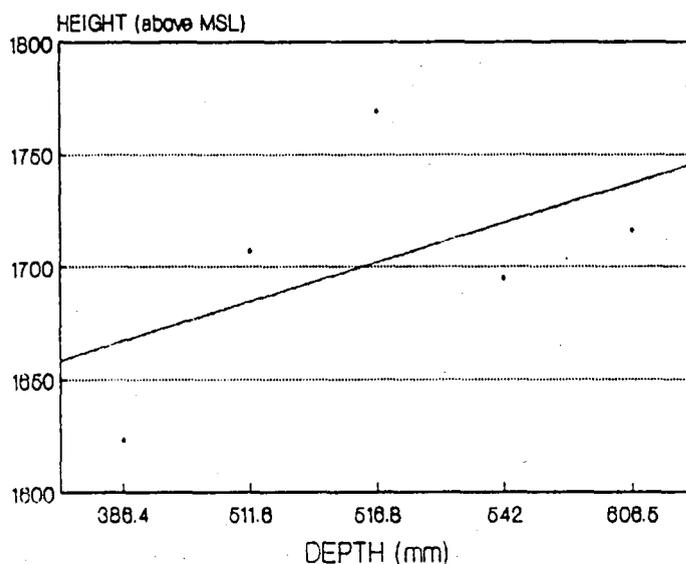


FIGURE 6.11 ELEVATION EFFECTS ON RAINFALL

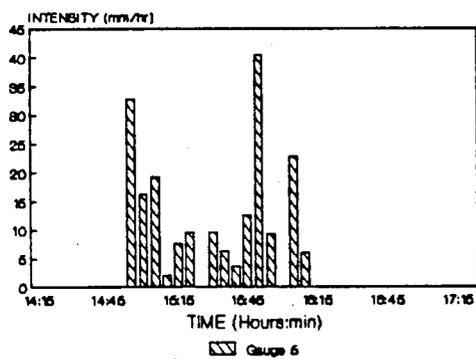
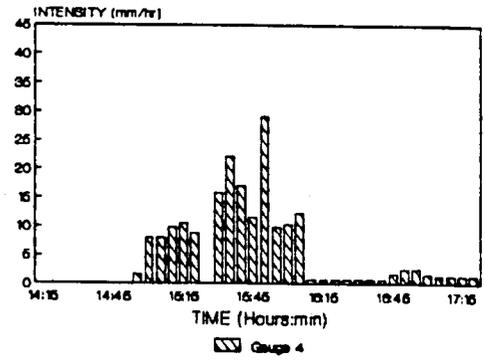
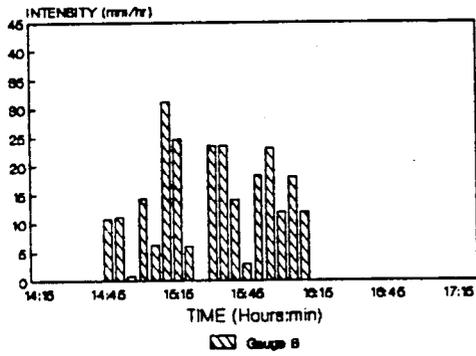
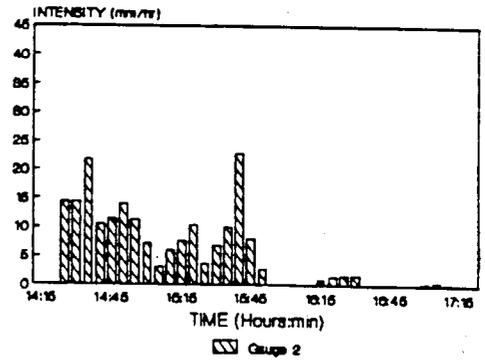
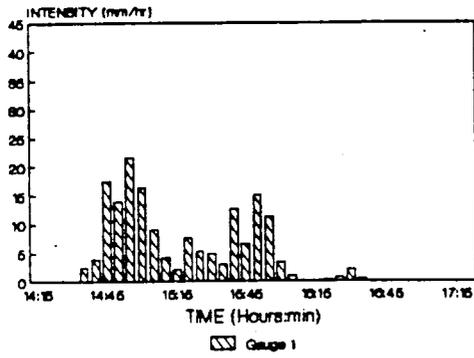
6.3 CELL SIZES AND DURATIONS

Consider figure 6.12, which shows bar graphs of rainfall intensity for each of the five gauges in the study catchment for the same storm event (10/11/87). There are several factors that can be highlighted:

- . The total depths for each of the gauges are of similar magnitudes, and close to the mean depth for the whole catchment.
- . The shapes of the bar graphs for the five gauges are completely different - there is no apparent general storm shape.
- . The start times and end times for the storm event measured at the five gauges are all different.
- . The duration of the event measured at each gauge is different.
- . It is often possible to detect cellular structures, where the storm event is broken into groups of rainfall by periods of no or very little rain (seen clearly at gauges three and five).

This storm event was chosen because it was typical of the storms studied and is of short duration, which made it easier to draw attention to the above points. This particular storm also shows that having the total depths at the measuring gauges being of similar magnitudes does not guarantee similarity of shape characteristics. Spottiness of rainfall would be expected when the depths measured are very different between gauges, but this shows that consistency of total depth is a bad indicator of lack of spottiness.

The fact that the storm was first detected at different times at the five gauges might indicate a trend for the catchment. It was found that some events were initially detected in the same five minute time interval at more than one gauge, and other events at only one gauge. For the twenty-one storms studied, the number of times that rainfall was first detected at the five gauges was eight times for gauge one,



GAUGE DEPTH(mm)

1	14.2
2	15.4
3	22.4
4	16.7
5	18.1

Average depth = 17.1mm (by Innessen weights)

FIGURE 6.12 HYETOGRAPHS FROM THE FIVE GAUGES FOR THE SAME STORM EVENT

eight times for gauge two, six times for gauge three, seven times for gauge four and four time for gauge five.

The number of times rainfall was first detected at one gauge only was also determined and this was four times at gauge one, five times at gauge two, twice at gauge three, three times at gauge four and once at gauge five. The only common factor to emerge from this was that gauge five (which is under the shadow of the highest part of the catchment, but is not the highest gauge) had rainfall starting there significantly less often than the other four gauges.

The duration of the whole event can be taken as the time from the first onset of rain at any gauge until no more rain is measured at any gauge. This is a far more complex situation than design storms of constant intensity and set duration, and even weighted averages of gauge depths for the whole event. This definition of storm duration is however not very useful, as it includes times at the beginning and end of the storm event when there is comparatively little rain. For practical use, though, some means of comparing storm events and relating them to return periods and degrees of severity must be determined.

The contours of storm shapes are useful for a qualitative examination of storm events and cells, but not for a quantitative analysis. For the purposes of individual cell durations, a different technique was used:

For the storms used to determine cell duration, each gauge was studied individually. The duration of cells for several intensities was determined. For example, at gauge number one for the storm on 15/01/87 the durations of cells of intensities greater than or equal to 2.5mm/hr were measured and counted, and four cells of 40, 5, 50 and 15 minutes were found. This was done for the other four gauges as well.

Cells of intensities greater than or equal to 2.5mm/hr, 5mm/hr, 10mm/hr, 15mm/hr, 20mm/hr, 30mm/hr, 40mm/hr,... etc. were measured and counted. This information was then pooled for the storm events and gauges to determine a mean cell duration.

The author found that there was a steadily decreasing number of cells with an increasing cell intensity. When the mean duration was calculated, it was found that there were two definite plateaus; one at a mean duration of about 22 minutes for intensities between 2.5mm/hr and 20mm/hr, and the second at a duration of about 11 minutes for intensities between 20mm/hr and above. Standard deviations were in the same order of magnitude as the mean values.

The author does not suggest that cells of intensities between 2.5mm/hr and 20 mm/hr will have durations of 22 minutes, or that those of intensities greater than 20mm/hr will have durations of about 11 minutes, as this cannot be concluded from the manner in which these figures were derived. What this could indicate, though, is that there may be typical cell characteristics that can be extracted for a catchment and later built into a design model.

In keeping with other studies, examination of contour maps shows cell dimensions of about 3km diameter. This was noticed both at the peak and also for a large portion of the cell's duration, suggesting an upper limit to the areal size of a single cell.

6.4 EFFECTS OF HYETOGRAPHS ON RUNOFF

The above work is enlightening, but needs to have practical implications to the hydrologist for it to be relevant. A means of determining the quantitative effect different storm shapes would have on a catchment was necessary. The most obvious is the characteristics of the resultant runoff hydrograph when a storm event is routed over a catchment.

To this end, a reliable runoff model was needed. The model developed by Green (Green and Stephenson, 1985) called WITWAT was used, for several reasons:

- . The model is used extensively in Southern Africa and has therefore been tested often.

- . The original model was developed on a Hewlett Packard, which was compatible with the Hewlett Packard used by the author. The relevant software was made available to the author by Dr. Green. The original calibrated data sets used for Montgomery Park were made available to the author.
- . It would be necessary to modify WITWAT to accept spatially varying intensities as storm inputs. This had already been done (Lambourne and Stephenson, 1986) and this software was made available to the author by Dr. Lambourne. It was then necessary to convert the measured data from raingauges into a suitable format for this modified version of WITWAT.

Of the twenty-one storms, eight were selected for use with the overland model. This was mostly due to practical considerations of memory space: The real storms had durations of up to 1000 minutes, and the spatial version could only accept up to 108 ordinates (540 minutes). This meant that storms which could not be realistically cut to below 540 minutes duration had to be left out.

Seven runs were made for each storm event. Five of these were routing the measured hyetograph at the five gauges as a spatial event over the whole catchment. In many practical applications, there is a gauge density of less than one gauge per 10km². Data from this type of arrangement must, per force, be applied to large areas surrounding the gauge. These first five tests were done to investigate the possible errors in this.

The sixth run used a weighted event, with Thiessen weights applied to the corresponding intensity at each five minute interval. The whole storm was converted in this way and applied over the whole catchment.

The seventh run used a special data file that was set up using the ISD method. Some way of distributing the rainfall spatially was needed. Each five minute time interval had rainfall intensities interpolated and applied to the center of gravity of each discrete area defined by Green (Green and Stephenson, 1985) for the catchment.

Many studies are done using design storms of constant intensity. To give some means of comparison with other catchments, some constant intensity design storms were also routed over the study catchment. The mean annual precipitation for this area is 720mm, and this gives rise to the Intensity-Duration-Frequency (IDF) curves shown in figure 6.13. This figure was derived using the inland equation developed by op ten Noort and Stephenson (1982).

Intensity - Duration - Frequency Montgomery Park (MAP = 720 mm)

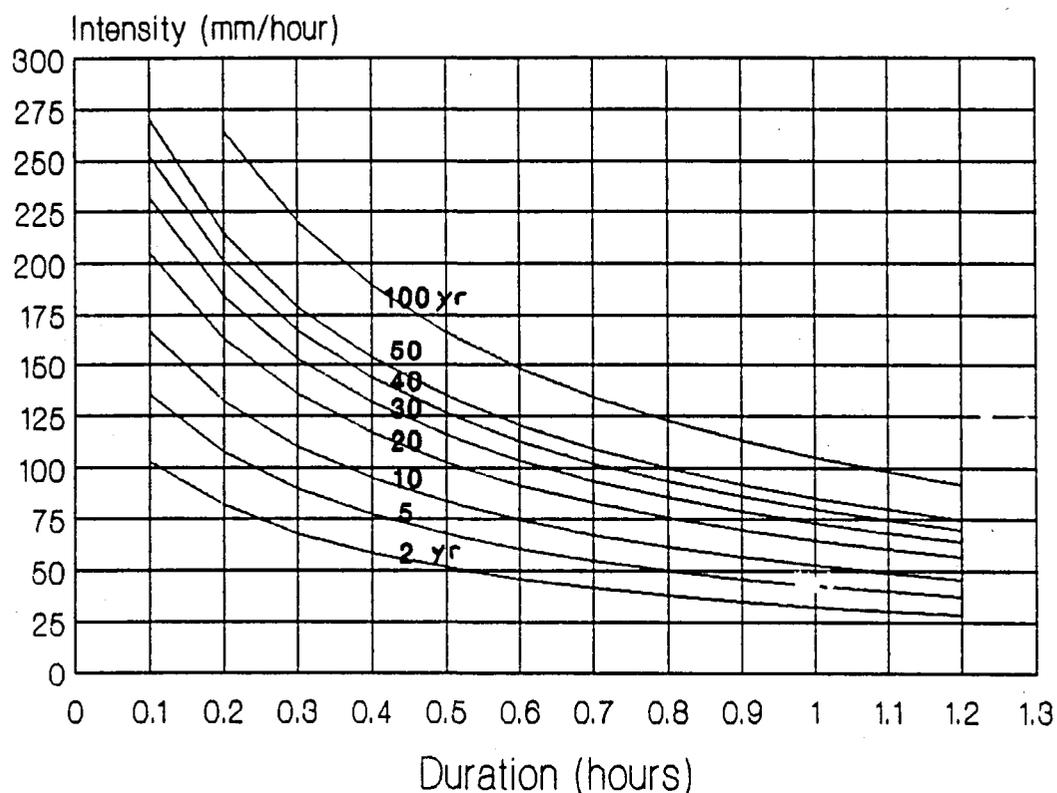


FIGURE 6.13 INTENSITY - DURATION - FREQUENCY CURVES FOR MONTGOMERY PARK (MAP = 720 mm)

Three return periods were examined; two year events, five year events and ten year events. In each return period, the worst combination of intensity and duration was found for the catchment. The resultant hydrographs are shown in figure 6.14.

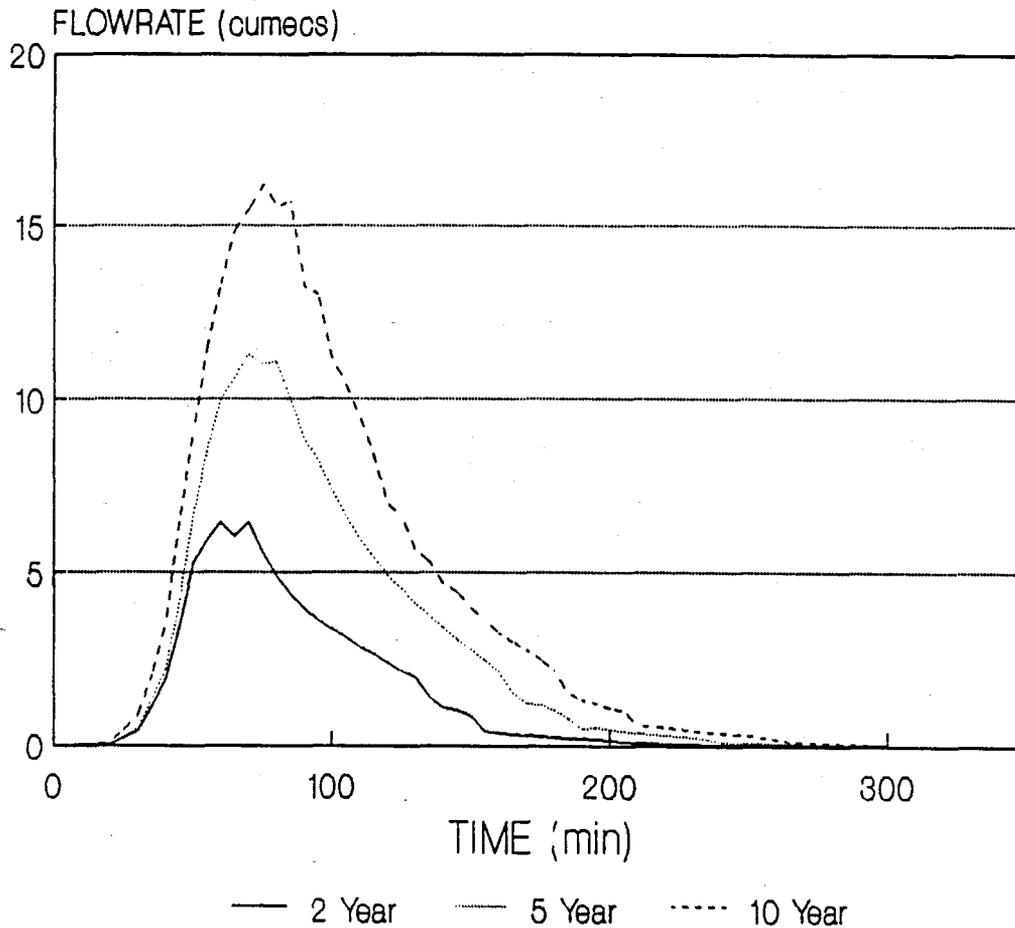


FIGURE 6.14 DESIGN STORM HYDROGRAPHS

The intensities and durations for the three design storms are 42.6mm/hr over 35 minutes, 44.3mm/hr over 50 minutes and 51mm/hr over 55 minutes for the 2, 5 and 10 year events respectively. This calculates out at 24.8mm, 36.9mm and 46.8mm total rainfall depth for the three design storms. The peak flowrates are 6.5 cumecs, 11.3 cumecs and 16.2 cumecs reached at times 60 minutes, 70 minutes and 75 minutes after the start of rainfall. These design storms are included to give some frame of reference to the actual storms studied, and since no real event had a depth exceeding 46.8mm rainfall, no further design storms were used.

Findings from real storms showed a large variation in runoff. Figure 6.15 shows all seven runoff hydrographs from the storm on 15/01/87. It can be seen from the graphs that the lowest peak flowrate is about 0.7 cumecs, while the highest peak flowrate is about 1.5 cumecs - a factor of two difference. This cannot be explained by varying depths of rainfall, as for this event the depths were 16.2mm, 19.3mm, 19.2mm, 15.4mm, 16.7mm and 17.4mm for gauges 1, 2, 3, 4, 5 and the Thiessen weighted option.

The spatial storm, when routed over the catchment, produces a hydrograph somewhere between the highest and lowest hydrographs for individual gauges. If the spatial storm were to exceed the bounds of the highest or lowest hydrographs, then this would indicate a nett loss or gain in rainfall and hence an error (since there cannot be more rain than the maximum value, or less than the minimum value).

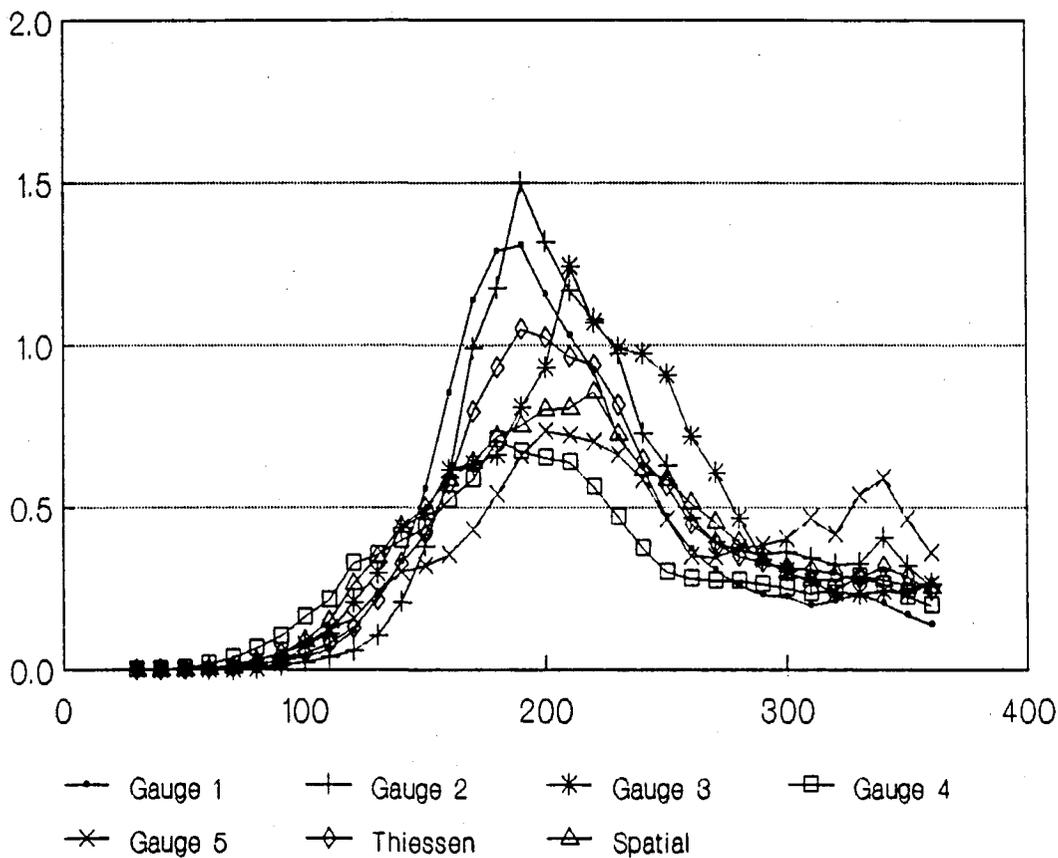


FIGURE 6.15 HYDROGRAPHS FOR THE STORM ON 15/10/1987

As mentioned above, the hyetographs from all five gauges were routed over the catchment as individual storms to simulate what would happen if there were only one gauge in the study catchment. To give an idea of a worst case, the hydrographs from a storm on the 6/11/87 is included here (see figure 6.16). This event has minimum depth recorded as 6mm, and maximum depth recorded as 19.6mm at gauges one and three respectively. The Thiessen weighted mean depth of rainfall is 9.8mm.

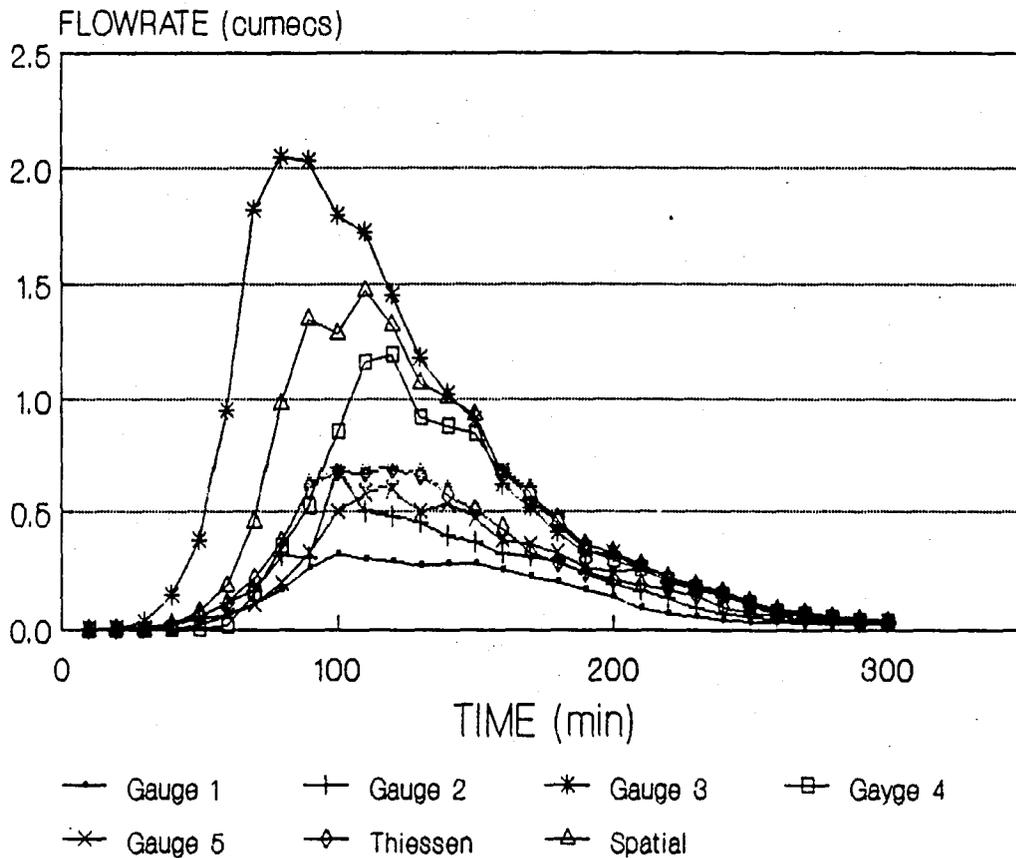


FIGURE 6.16 HYDROGRAPHS FOR THE STORM ON 6/11/1987

As can be seen, the difference in hydrograph characteristics is very large. Using the values from either gauge one or gauge three would result in considerable errors in peak flowrate or runoff volume estimation.

Another event with total depths of similar magnitudes is the storm on the 10/11/87 (see figure 6.17). The five hyetographs for the five raingauges for this event can be found as figure 6.12 above. The difference between the minimum (gauge one) and maximum (gauge five) depths and the Thiessen weighted depth is 17% (less) and 28% (greater) respectively. This can be compared to the resultant peak flowrates comparisons of 22% (less) and 53% (greater), and the resultant volumes comparisons of 19% (less) and 33% (greater) for these same gauges. These trends are similar for all the storm events studied.

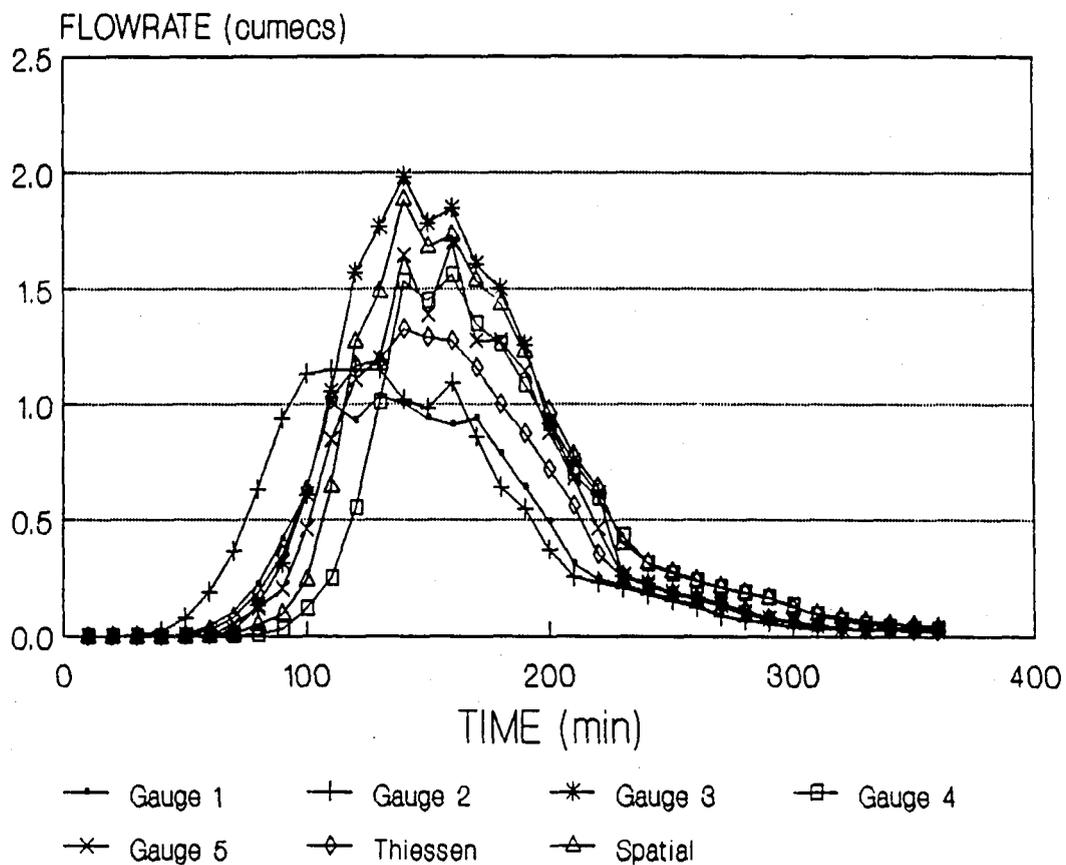


FIGURE 6.17 HYDROGRAPHS FOR THE STORM ON 10/11/1987

If one examines the results of the spatial hydrographs from figures 6.15, 6.16 and 6.17, there appears no easy correlation as with the hydrographs produced by the Thiessen weighted events. Figure 6.15 shows a hydrograph close to the lower bound, figure 6.16 shows a hydrograph close to the middle and figure 6.17 shows a hydrograph close to the upper bound.

6.5 DISCUSSION

Clearly, a single rainfall event is composed of several cells that are the major producer of rainfall during that event. If one is to model rainfall properly then one must introduce this aspect into the model. What could be investigated further is whether there is a typical limit to the areal size of individual cells. From this study, it appears as if there is, and that this size is reached in the early portion of a cell's life. Further growth is apparent in higher intensities within the same area, as opposed to an expanding of areal size.

It has been shown that rainfall varies considerably across a catchment. Weighting discrete time steps of a rainfall event by some means such as Thiessen weights is an improvement on constant intensity storms. However, in many practical situations, such as in the study catchment, the distribution of raingauges is not even. Because of this, some gauges are constantly weighted higher than others, and this lumped rainfall is applied to the whole catchment.

If for some reason (e.g. relief differences or physical obstructions), one gauge had significantly more or less rainfall than others in a catchment, and this gauge also had a significantly high or low weighting, then the rainfall input would have an intrinsic error. The magnitude of this error may be small, but with an increasing pressure on models and planners for accuracy and confidence in results, this should be investigated.

The picture presented by this report of storm movements is a complex one as far as modelling goes. Spotty rain, convoluted cellular distributions of high intensities and planar distributions of intensities have all been found to exist over the study catchment. Sometimes all three existed at some stage in the same storm event.

The trend with computer runoff models appears to be to include higher levels of discretisation of the catchment for descriptions of the physical aspects of the catchment: Ground coverage is divided up by the type of land use (e.g. impervious and characteristics of perviousness), vegetation types and their areal coverage are important (for protection of ground from erosion and the amount of potential evaporation) and the soil characteristics are also evaluated (to determine the degree of ponding, the ability to absorb moisture and transmit subsurface water).

The author feels that the value of this work is largely negated if the rainfall used as an input for the models is not refined to the same degree. As shown above, rainfall is a very variable quantity over short distances and small times. This also introduces a subtle point for calibration of models: If the correct distribution of rainfall is not used when models are being calibrated, then some physical parameters must be artificially adjusted to produce the required performance. This is not desirable.

It has been shown above that considerable errors can occur in runoff calculations if the best estimate of rainfall (i.e. closest to the real event) is not used. The author suggests that this is also true for calibration of models, and that a more refined method of rainfall input be used for calibration.

Sieker (1978) cast doubts on the assumption that a recurrence interval of storm events matched the corresponding recurrence interval for runoff events. For design purposes some form of correlation is required though. A re-evaluation of this position is called for, with a better informed notion of storm characteristics.

7 DISCUSSION

Several points have had attention drawn to them in this project. The most significant of these are discussed in chapters 4.10 and 6.5, but a summary is reproduced here before the final discussion section.

- . A single rainfall event is a composition of several component cells whose individual behaviour is erratic. The description of these cells in parameters useful to hydrologists is very important for the accurate study and understanding of rainfall events.

- . The distribution of raingauges in a catchment are not optimal arrangements for computer interpolation, and uneven representation of gauges can occur. There are seldom enough raingauges for a high level of accuracy when interpolating storm events, which have extreme variations in intensities over small distances - in most cases significantly less than the distances between gauges.

- . The pattern of rainfall intensities from real rainfall events over their duration varies radically, both with time and space. Hyetographs from the same catchment for the same storm event can appear totally unrelated to each other, even though the gauges are all within about 2km of each other, as was found in this study. This makes accurate calibration of models difficult

- . Process models that model the rainfall-runoff process are being produced in package format on micro-computers with increasing levels of discretisation and accuracy with regard to the modeling of soil, slope and vegetation parameters. To benefit from this, the description of distribution of rainfall over the catchment must be to the same accuracy. Rainfall forms the major input for a cause and effect relationship and if this is not taken into account the benefits of discretisation will be lost when the areas are lumped together under averaged rainfall inputs.

. The current use of design storms is based on the assumption that the probability of a rainfall event of a certain magnitude happening is the same as the probability of a runoff event of a certain magnitude happening. This is not necessarily true for real storm events, as for a given average depth of rainfall totally different sections of a catchment can be affected by different storm events, and hence produce totally different hydrograph characteristics.

The above discussion serves to highlight further issues. Addressing these in part form the basis for the discussion in this final chapter:

The number of raingauges in a catchment have a significant effect on the accuracy of rainfall measurement. This is related to both the accuracy of the numerical technique employed, and the number and distance between the gauges available for the study. An important aspect is the requirements of the study; whether it is for small urban studies, drainage regions or water resource management, composed of several drainage regions. An intrinsic assumption is that the raingauges themselves are accurate, and that their time synchronisation is to acceptable levels - the author suggest to within five minutes of each other for meaningful results.

The accuracy benefits of applying the numerical method used in this project to catchments of large aerial sizes is unknown. However, understanding that storm events are erratic in behaviour, and can be physically smaller in area than the study catchment indicates the numerical method's advantage in describing storm events more precisely than current popular techniques. It is therefore reasonable to speculate and expect an increase in accuracy, if such a numerical technique is used for large area studies.

To consider catchments in their correct relation to storm event size, it would be beneficial to classify catchment studies by the same classification applied to storm events, that is microscale for less than about 6km diameter, mesoscale for less than about 50km diameter and synoptic scale for much larger catchment studies.

This would aid in focusing attention on what aspects of the storm event are most influential to that scale of study (i.e. topographical, relief, aerial extent, wide variations in rainfall intensity etc.). Related to this is whether it is possible to link cell characteristics with overall climatic and topographical characteristics.

Two distinct types of storm event are possible: long duration low intensity and short duration high intensity storms. The effects of either are very different. For event models, the soil moisture conditions have to be estimated accurately to produce a calibrated catchment. If combinations of these two types of events have occurred on a real catchment, the resultant runoff can be very different to single events spaced far apart. While the immediate effect of a low intensity storm event is small, it may saturate the soil to a degree that an otherwise moderate storm then event becomes extreme. It was noticed that long duration events studied often exhibited qualities of both types of storm event.

In most runoff modelling studies the start and end time of the storm event is rationalised by some means, giving the same start and end time to each discrete sub-area within the study catchment. By comparison, the real events exhibited large differences in start and end times for the five rain-gauges in the study catchment. The discrepancy between modelled and actual storm behaviour highlights the inability of current techniques to describe storm events to an acceptable level of accuracy. This also corroborates the observed erratic nature of real storm events.

Other questions occurred to the author as a consequence of this project. These are listed below as suggestions for further study.

SUGGESTIONS FOR FURTHER STUDY

- . How much are the shapes of cells (and therefore runoff) influenced by the method used to study them?

- . A method of relating the recurrence of real storm events to the recurrence of runoff events must be determined. This is a complex relationship and the design-storm approach of constant intensities is not to be trusted as representing real events. This relationship should be investigated for planning and design purposes.

- . A new generation of runoff model incorporating the spatial structure of real or reality-based storms is necessary.

- . Initially, a method such as ISD used in this report could be incorporated, but the task expected of these methods exceeds their bounds of accuracy at common levels of raingauge density. The author expects that models based on the physical nature of storms (such as the NWP types) can be developed to model small catchments to sufficient resolution and accuracy. Investigation of this and available alternatives, and the possibility of linking to a computer-based runoff model could be investigated.